



**Kauno technologijos universitetas**

Informatikos fakultetas

**Žmogaus veido ir balso pakeitimo kitu asmeniu sistema ir  
tyrimas**  
**Research on substitution of a human face and voice with that  
of a different person and tool creation**

Baigiamasis magistro studijų projektas

---

**Indrė Dimšė**

Projekto autorė

**Prof. Rytis Maskeliūnas**

Vadovas

---

**Kaunas, 2023**



**Kauno technologijos universitetas**

Informatikos fakultetas

**Žmogaus veido ir balso pakeitimo kitu asmeniu sistema ir  
tyrimas**  
**Research on substitution of a human face and voice with that  
of a different person and tool creation**

Baigiamasis magistro studijų projektas  
Programų sistemų inžinerija (6211BX011)

---

**Indrė Dimšė**

Projekto autorė

**Prof. Rytis Maskeliūnas**

Vadovas

**Prof. Robertas Damaševičius**

Recenzentas

---

**Kaunas, 2023**



**Kauno technologijos universitetas**

Informatikos fakultetas

Indrė Dimšė

**Žmogaus veido ir balso pakeitimo kitu asmeniu sistema ir tyrimas**  
**Research on substitution of a human face and voice with that of a different person and tool creation**

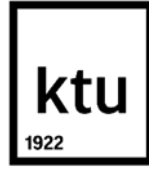
Akademinio sąžiningumo deklaracija

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.

Indrē Dimšē

*Patvirtinta elektroniniu būdu*



**Kauno technologijos universitetas**

Informatikos fakultetas

## **Baigiamojo magistro projekto užduotis**

Projekto tema	<u>Žmogaus veido ir balsų pakeitimo kitu asmeniu sistema ir tyrimas</u>
Reikalavimai ir sąlygos (tikslinti pavadinimą pagal poreikį)	<div style="border: 1px solid black; padding: 10px;"><p>Vertėmis grįstas pasiūlymas: sistema su mašininio mokymo veikimu pagrįsta veido/balsų keitimo analizė</p><ol style="list-style-type: none"><li>1. Skirta žmonėms ir įmonėms, kurie nori pakeisti asmens balsą ir veidą, esančius nuotraukose ir garso įrašuose, kitų asmeniu.</li><li>2. Mūsų siūlomas sprendimas: Veido ir balsų keitimo sistema.</li><li>3. Siūloma sistema galės priimti mašininio mokymo modelius, bei juos naudojant iš originalo failo perkelti balsą/veidą į kitą norimą failą. Sistema bus paruošta naudojimui iki 2023-12-30.</li><li>4. Skirtingai nei rinkoje esančios konkurencinės sistemos, ši sistema galės kokybiškai atlikti veido keitimą nuotraukoje bei balsų keitimą garso įrašuose, bei taip pat naudoti bei dalinti savo sukurtais modeliais.</li></ol></div>
Vadovas / Vadovė	<u>Prof. Rytis Maskeliūnas</u> (vadovo pareigos, vardas, pavardė, parašas) (data)

Dimšė, Indrė. Žmogaus veido ir balsų pakeitimo kitu asmeniu sistema ir tyrimas. Magistro baigiamasis projektas / vadovas prof. Rytis Maskeliūnas; Kauno technologijos universitetas, Informatikos fakultetas

Studijų kryptis ir sritis (studijų krypties grupė): Informatika, programinės įrangos inžinerija.

Reikšminiai žodžiai: neuroninis veido sukeitimas, neuroninis balsų konvertavimas, balsų pakeitimas, veido pakeitimas.

Kaunas, 2023. 150 p.

### **Santrauka**

Šiuo metu žmogaus balsų ir vaizdų generavimo ar modifikavimo naudojant dirbtinį intelektą sritis yra plėtojama ir dar nėra visiškai ištirta ar prieinama plačiajai visuomenei. Šios sistemos taip pat gali būti naudojamos medicinos tikslais, pavyzdžiui, padėti žmonėms, turintiems kalbos sutrikimų, sukuriant sintetinį balsą, kuris skambėtų panašiau į jų natūralų balsą. Jie taip pat gali būti naudojami pramogų pramonėje užsienio filmams įgarsinti arba tikroviškesnei virtualios realybės patirčiai kurti. Jos taip pat gali būti naudojamos interaktyviems ir patraukliems mokymo metodams, pavyzdžiui, atgaivinant garsias istorines asmenybes istorijos pamokose. *Deepfake* technologija jau kuriama ir naudojama asmeniniais ir komerciniais tikslais, tačiau žinių apie tai, kaip atlikti tyrimus šioje srityje, yra minimaliai. Šio projekto tikslas – išplėsti žinias apie veido ir balsų pakeitimą naudojant dirbtinį intelektą ir nustatyti geriausias technologijas kokybiškam veido ir balsų pakeitimui. Šio projekto tikslai – iširti šiuo metu veido ir balsų pakeitimui naudojamas technologijas ir palyginti jų privalumus bei trūkumus. Be to, bus sukurta sistema, kurioje bus naudojami mašininio mokymosi modeliai asmens veidui ir (arba) balsui pakeisti kito asmens veidu ir (arba) balsu.

Dimšė, Indrė. Research on substitution of a human face and voice with that of a different person and tool creation. Master's Final Degree Project / supervisor prof. Rytis Maskeliūnas; Faculty of Informatics, Kaunas University of Technology.

Study field and area: Computing, Software Engineering.

Keywords: face deepfake voice deepfake, neural face swapping, neural voice conversion, voice substitute deep learning, voice substitution, face substitution.

Kaunas, 2023, 129 p.

### **Summary**

The field of generating or modifying human voices and images using artificial intelligence is currently under development and not yet fully explored or available to the general public. These systems can also be used for medical purposes, such as helping people with speech impairments by creating a synthetic voice that sounds more like their natural voice. They can also be used in the entertainment industry for dubbing foreign films or for creating more realistic virtual reality experiences. They can also be used for interactive and engaging teaching methods, such as bringing famous historical figures to life in history classes. Deepfake technology is already being developed and used for personal and commercial purposes, but there is minimal knowledge about how to conduct research in this field. This project aims to expand knowledge about face and voice swapping using artificial intelligence and to identify the best technologies for quality facial and voice substitution. The goals of this project are to explore the current technologies used for face and voice substitution and to compare their strengths and limitations. Additionally, a system will be created that uses machine learning models for substituting a person's face and/or voice with another person's.

## Turinys

Lentelių sąrašas .....	9
Paveikslų sąrašas .....	11
Santrumpų ir terminų sąrašas .....	14
Įvadas.....	17
2. Projekto paraiška .....	18
3. Literatūros analizė.....	25
4. Projektas.....	46
5. Tyrimas.....	109
6. Eksperimentai .....	111
Išvados .....	145
Literatūros sąrašas .....	146
Informacijos šaltinių sąrašas .....	150
Priedai.....	151



## Lentelių sąrašas

2-1 lentelė. Sistemos vertinimo kriterijai .....	23
2-2 lentelė. Projekto biudžetas .....	24
2-3 lentelė. Konkurentų analizė .....	24
3-1 lentelė. Veido keitimo algoritmai .....	26
3-2 lentelė. Veido keitimo duomenų rinkinys.....	35
3-3 lentelė. Balso keitimo algoritmai .....	36
3-4 lentelė. Balso įrašų duomenų rinkiniai .....	44
4-1 lentelė. Detalus projekto planas .....	46
4-2 lentelė. Veiklos suskaidymas.....	51
4-3 lentelė. „Prisijungti“ specifikacija .....	53
4-4 lentelė. PA „Prisijungti“ specifikacija .....	54
4-5 lentelė. PA „Keisti veidą nuotraukoje“ specifikacija.....	54
4-6 lentelė. PA „Keisti balsą garso įrašė“ specifikacija.....	54
4-7 lentelė. PA „Keisti veidą ir balsą vaido įrašė“ specifikacija.....	55
4-8 lentelė. PA „Peržiūrėti visus failus“ specifikacija .....	55
4-9 lentelė. PA „Peržiūrėti failą“ specifikacija .....	55
4-10 lentelė. PA „Atsisiųsti failą“ specifikacija.....	56
4-11 lentelė. PA „Išsaugoti failą“ specifikacija .....	56
4-12 lentelė. PA „Pašalinti failą“ specifikacija .....	56
4-13 lentelė. PA „Peržiūrėti modelius“ specifikacija.....	57
4-14 lentelė. PA „Pašalinti modelį“ specifikacija .....	57
4-15 lentelė. PA „Įkelti modelį“ specifikacija .....	57
4-16 lentelė. PA „Peržiūrėti vartotojus“ specifikacija .....	58
4-17 lentelė. PA „Redaguoti vartotoją“ specifikacija .....	58
4-18 lentelė. PA „Pašalinti vartotoją“ specifikacija.....	58
4-19 lentelė. Greičio reikalavimai.....	59
4-20 lentelė. Pasiiekiamumo reikalavimas.....	59
4-21 lentelė. Klaidų reikalavimas .....	60
4-22 lentelė. Apkrovos reikalavimas .....	60
4-23 lentelė. Spalvos reikalavimai .....	60
4-24 lentelė. Dizaino reikalavimai .....	61
4-25 lentelė. Kalbos reikalavimai .....	61
4-26 lentelė. Modelio atsiminimo reikalavimas.....	61
4-27 lentelė. Mokymo reikalavimai .....	61
4-28 lentelė. Sąvokų reikalavimai.....	61
4-29 lentelė. Išplečiamumo reikalavimai .....	62
4-30 lentelė. Igaamžiškumo reikalavimai .....	62
4-31 lentelė. Fizinės aplinkos reikalavimai.....	62
4-32 lentelė. Suderinamumo reikalavimai .....	63
4-33 lentelė. Sistemos palaikymo reikalavimas .....	63
4-34 lentelė. Prieigos reikalavimas .....	64
4-35 lentelė. Vientisumo reikalavimas.....	64
4-36 lentelė. Privatumas reikalavimas .....	64
4-37 lentelė. Privatumas reikalavimas .....	64

<b>4-38 lentelė.</b> Sistemos kūrimo reikalavimas .....	65
<b>4-39 lentelė.</b> Tikimybių kategorijos .....	67
<b>4-40 lentelė</b> Rizikos tikimybių matrica .....	68
<b>4-41 lentelė</b> Projekto rizikos .....	68
<b>4-42 lentelė</b> Rizikos poveikio matrica .....	73
<b>4-43 lentelė</b> Rizikos ir jų prioritetai .....	73
<b>4-44 lentelė.</b> Išlaidų lentelė .....	74
<b>4-45 lentelė.</b> Registracija bei prisijungimas .....	93
<b>4-46 lentelė.</b> Modelio valdymas .....	93
<b>4-47 lentelė.</b> Failų valdymas .....	93
<b>4-48 lentelė.</b> Vartotojų valdymas .....	94
<b>4-49 lentelė.</b> Veiklų terminai.....	94
<b>4-50 lentelė.</b> Rankinio testavimo rezultatai.....	95
<b>4-51 lentelė.</b> Pakeitimų sąrašas .....	104
<b>4-52 lentelė.</b> Peržiūrų lentelė.....	105
<b>4-53 lentelė.</b> Komandos narių rolės .....	106
<b>4-54 lentelė.</b> Informacijos rinkimo planas .....	106
<b>4-55 lentelė.</b> Kokybės vertinimo duomenų rinkimo planas .....	106
<b>4-56 lentelė.</b> Kokybės vertinimo kriterijai .....	107
<b>6-1 lentelė.</b> Veido keitimo modelių mokymų aprašymas.....	115
<b>6-2 lentelė.</b> Balso keitimo tikrinimo metrikos.....	115
<b>6-3 lentelė.</b> Veido mokymų rezultatai .....	116
<b>6-4 lentelė.</b> Balso keitimo modelių architektūros .....	123
<b>6-5 lentelė.</b> Balso keitimo vokoderių architektūros .....	123
<b>6-6 lentelė.</b> Veido keitimo modelių mokymai .....	130
<b>6-7 lentelė.</b> Veido keitimo vokoderių mokymas .....	131
<b>6-8 lentelė.</b> Veido keitimo tikrinimo metrikos.....	138
<b>6-9 lentelė.</b> Veido keitimo modelių mokymo rezultatai.....	139
<b>6-10 lentelė.</b> Veido keitimo vokoderių mokymo rezultatai .....	139

## Paveikslų sąrašas

2-1 pav. Panaudojimo atveju diagrama.....	21
2-2 pav. Konteksto diagrama .....	22
2-3 pav. Projekto vykdymo grafikas .....	23
3-1 pav. PRISMA atrankos diagrama .....	26
3-2 pav. VAE struktūra [1] .....	27
3-3 pav. Automatinis kodavimo veikimas apkeičiant veidus [6].....	28
3-4 pav. GAN struktūra [1].....	28
3-5 pav. FSGAN veikimas [6] .....	29
3-6 pav. FCCGAN veikimo schema [8] .....	29
3-7 pav. FCCGAN koduotojo schema [15] .....	30
3-8 pav. FCCGAN dekoduoja schema [8] .....	30
3-9 pav. FCCGAN self-attention modulio schema [8].....	30
3-10 pav. Diskriminatoriaus schema.....	31
3-11 pav. RS-GAN veikimo schema [10] .....	31
3-12 pav. FSNet veikimas .....	32
3-13 pav. Veidas prieš pakeitimą, uždėjus šaltinio nuotrauką ir įliejus veidą į nuotrauką [9] .....	32
3-14 pav. Ulyanov architektūra [11]. .....	33
3-15 pav. Spalvos perkėlimo algoritmas [8] .....	34
3-16 pav. Veido spalvos pakeitimo veikimas [8].....	34
3-17 pav. Veido apdorojimas [16] .....	35
3-18 pav. MDCNN veikimo schema [27].....	37
3-19 pav. MDCNN architektūra .....	37
3-20 pav. SVINGE2E architektūra .....	38
3-21 pav. RD-VS architektūra [29].....	38
3-22 pav. WaveNet schema [33] .....	39
3-23 pav. CycleRNN modelio architektūra su WaveNet procedūra [28].....	40
3-24 pav. Seq2Seq veikimo schema [25] .....	41
3-25 pav. SCENT architektūra [31].....	42
3-26 pav. BLSTM architektūra [32].....	43
3-27 pav. BLSTM schema [31].....	43
3-28 pav. Generatoriaus-vokoderio (a) ir WaveRNN (b) veikimo schemas [33].....	44
4-1 pav. Diegimo aplinkos diagrama .....	49
4-2 pav. Veiklos konteksto diagrama.....	50
4-3 pav. Duomenų modelis .....	52
4-4 pav. PA diagrama.....	53
4-5 pav. Kūrimo planas .....	66
4-6 pav. Paketų diagrama.....	76
4-7 pav. PA „Peržiūrėti visus failus“ veiklos diagrama.....	77
4-8 pav. PA „Peržiūrėti failą“ veiklos diagrama.....	77
4-9 pav. PA „Pašalinti failą“ veiklos diagrama.....	78
4-10 pav. PA „Atsisiųsti failą“ veiklos diagrama .....	78
4-11 pav. PA „Išsaugoti failą“ veiklos diagrama .....	79
4-12 pav. Failų keitimo detalūs panaudojimo atvejai .....	79
4-13 pav. PA „Keisti failą“ veiklos diagrama .....	80

<b>4-14 pav.</b> PA „Keisti failą“ sąveikos diagrama.....	81
<b>4-15 pav.</b> PA „Peržiūrėti modelius“ veiklos diagrama .....	82
<b>4-16 pav.</b> PA „Įkelti modelį“ veiklos diagrama .....	82
<b>4-17 pav.</b> PA „Redaguoti modelį“ veiklos diagrama .....	83
<b>4-18 pav.</b> PA „Pašalinti modelį“ veiklos diagrama.....	83
<b>4-19 pav.</b> PA „Pašalinti vartotoją“ veiklos diagrama .....	84
<b>4-20 pav.</b> PA „Redaguoti vartotoją“ veiklos diagrama.....	84
<b>4-21 pav.</b> PA „Peržiūrėti vartotojus“ veiklos diagrama .....	85
<b>4-22 pav.</b> Išdėstymo diagrama .....	85
<b>4-23 pav.</b> Klasių diagrama .....	86
<b>4-24 pav.</b> VAE architektūra [8].....	87
<b>4-25 pav.</b> GAN architektūra veido keitimui [9].....	87
<b>4-26 pav.</b> Veido keitimo detalesnė schema.....	88
<b>4-27 pav.</b> Balso keitimo procesas.....	89
<b>4-28 pav.</b> Balso keitimo į tekstą schema.....	89
<b>4-29 pav.</b> Teksto į kalbą vertimas [42].....	90
<b>4-30 pav.</b> Vaizdo įrašo keitimo procesas .....	91
<b>4-31 pav.</b> Pagrindinis sistemos langas .....	97
<b>4-32 pav.</b> Failų sąrašo langas .....	98
<b>4-33 pav.</b> Pakeistų failų langas .....	98
<b>4-34 pav.</b> Keitimo proceso rezultatas.....	99
<b>4-35 pav.</b> Modelių sąrašo langas.....	100
<b>4-36 pav.</b> Modelių failų sąrašas .....	100
<b>4-37 pav.</b> Veido keitimo modelio langas .....	101
<b>4-38 pav.</b> Balso keitimo modelio langas .....	102
<b>4-39 pav.</b> Vaizdo įrašo modelio keitimo langas.....	103
<b>4-40 pav.</b> Administratoriaus langas .....	103
<b>6-1 pav.</b> Auto-koduotojo architektūra [36] .....	113
<b>6-2 pav.</b> Konvoliuciniai sluoksniai [37].....	114
<b>6-3 pav.</b> AE III mokymo rezultatai.....	117
<b>6-4 pav.</b> CAE III mokymo rezultatai .....	118
<b>6-5 pav.</b> Veido keitimo mokymų rezultatai .....	119
<b>6-6 pav.</b> Veido keitimo mokymų rezultatai po 100 tūkst. epochų .....	119
<b>6-7 pav.</b> Veido keitimo pavyzdžiai naudojant <i>ReLU</i> .....	120
<b>6-8 pav.</b> Veido keitimo pavyzdžiai naudojant <i>CU</i> .....	121
<b>6-9 pav.</b> Spektogramos pavyzdys.....	124
<b>6-10 pav.</b> Seq2Seq architektūros schema [41] .....	125
<b>6-11 pav.</b> Transformerių architektūros schema [43].....	126
<b>6-12 pav.</b> CVAE architektūros schema [44] .....	127
<b>6-13 pav.</b> GAN architektūros schema [45] .....	128
<b>6-14 pav.</b> RNN vokoderio architektūros schema [47] .....	129
<b>6-15 pav.</b> UnivNET architektūros schema [48] .....	130
<b>6-16 pav.</b> Seq2Seq mokymosi nuostoliai .....	131
<b>6-17 pav.</b> Seq2Seq validavimo nuostoliai .....	132
<b>6-18 pav.</b> Seq2Seq epochos laikas.....	132
<b>6-19 pav.</b> Transformerių mokymosi nuostoliai.....	132

<b>6-20 pav.</b> Transformerių validavimo nuostoliai.....	133
<b>6-21 pav.</b> Transformerių epochos laikas .....	133
<b>6-22 pav.</b> <i>CVAE</i> mokymosi nuostoliai .....	133
<b>6-23 pav.</b> <i>CVAE</i> validavimo nuostoliai .....	134
<b>6-24 pav.</b> <i>CVAE</i> epochos laikas .....	134
6-25 pav. <i>RNN</i> vokoderio mokymosi nuostoliai .....	135
<b>6-26 pav.</b> <i>RNN</i> vokoderio validavimo nuostoliai.....	135
<b>6-27 pav.</b> <i>RNN</i> vokoderio epochų laikai .....	135
<b>6-28 pav.</b> <i>GAN</i> vokoderio mokymosi nuostoliai.....	136
<b>6-29 pav.</b> <i>GAN</i> vokoderio validavimo nuostoliai.....	136
<b>6-30 pav.</b> <i>GAN</i> vokoderio epochų laikai .....	137
<b>6-31 pav.</b> <i>UnivNET</i> mokymosi nuostoliai.....	137
<b>6-32 pav.</b> <i>UnivNET</i> validavimo nuostoliai.....	137
<b>6-33 pav.</b> <i>UnivNET</i> epochų laikai .....	138
<b>6-34 pav.</b> Balso keitimo architektūrų palyginimas pagal modelius ir vokoderius.....	140
<b>6-35 pav.</b> Balso keitimo architektūrų palyginimas pagal akcentus.....	141
<b>6-36 pav.</b> Balso keitimo architektūrų palyginimas pagal lytį .....	142
<b>6-37 pav.</b> Balso keitimo metrikų pasiskirstymas pagal mokymus.....	142
<b>6-38 pav.</b> Balso keitimo metrikų pasiskirstymas pagal modelio architektūrą .....	143
<b>6-39 pav.</b> Balso keitimo metrikų pasiskirstymas pagal vokoderius.....	143

## Santrumpų ir terminų sąrašas

### Santrumpos:

LSTM (angl. *short-term memory*) – ilgalaikė trumpalaikė atmintis yra dirbtinė pasikartojanti neuroninio tinklo architektūra, naudojama giluminio mokymosi srityje. Skirtingai nuo standartinių perduotų neuroninių tinklų, LSTM turi grįžtamąjį ryšį. Jis gali apdoroti ne tik pavienius duomenų taškus, bet ir visas duomenų sekas.

CNN (angl. *convolutional neural network*) – konvoliucinis neuroninis tinklas yra dirbtinio neuroninio tinklo rūšis, dažniausiai taikoma kompiuterinės regos srityje, t. y., vaizdų atpažinimui, objektų radimui ar identifikavimui nuotraukose.

DNN (angl. *deep neural network*) – gilusis neuroninis tinklas yra dirbtinis neuroninis tinklas, turintis kelis sluoksnius tarp įvesties ir išvesties sluoksnių.

RNN (angl. *recurrent neural network*) – pasikartojantis neuroninis tinklas yra dirbtinių neuroninių tinklų klasė, kai ryšiai tarp mazgų sudaro nukreiptą arba neorientuotą grafiką pagal laiko seką.

VAE (angl. *variational autoencoders*) – koduotojas, kurio kodavimo paskirstymas yra sureguliuotas mokymo metu, siekiant užtikrinti, kad jo latentinė erdvė turėtų geras savybes, leidžiančias generuoti kai kuriuos naujus duomenis.

GAN (angl. *generative adversarial networks*) – generatyvusis priešininkų tinklas – mašininio mokymosi sistemų klasė, kurioje du neuroniniai tinklai varžosi tarpusavyje žaidime. Turint treniruočių rinkinį, ši technika išmoksta generuoti naujus duomenis naudojant tą pačią statistiką kaip ir mokymo rinkinys.

MFCC (angl. *Mel-frequency cepstral coefficients*) – trumpalaikio garso galios spektro atvaizdas, pagrįstas logaritminio galios spektro tiesine kosinuso transformacija netiesine mel dažnio skale garso apdorojimo metu.

MSE (angl. *Mean Squared Error*) - vidutinis kvadratinis paklaidos rodiklis, naudojamas vertinant skirtumą tarp prognozuotų ir tikrų reikšmių, dažnai naudojamas regresijos modeliuose.

SSIM (angl. *Structural Similarity Index*) - struktūrinis panašumo indeksas, naudojamas matuoti panašumą tarp dviejų vaizdų arba signalų, atsižvelgiant į šešis skirtingus panašumą apibrėžiančius veiksnius: šviesumo, kontrasto, struktūros, panašumo, interfeiso ir vietinio kontrasto.

AE (angl. *Autoencoder*) - automatinis koduotojas, dirbtinio intelekto algoritmas, kuris naudojamas mokymui pačiam išmokyti koduoti duomenis ir jų atkūrimui, panaudojant koduotą reprezentaciją.

SPEC (angl. *Spectral Clustering*) - spektrinis klasterizavimas, klasterizavimo metodas, kuris remiasi spektrinės grafo teorijos principais ir naudojamas grupuoti duomenis pagal jų spektrines savybes.

MCD (angl. *Mel Cepstral Distortion*) - mel cepstralinis iškraipymas, metrika, kuri matuoja skirtumą tarp dviejų garso signalų mel cepstrinėse koeficientų srityse, dažnai naudojama garso apdorojime, pavyzdžiui, garso perdavime arba garso atpažinime.

CAE (angl. *Convolutional Autoencoder*) - konvoliucinis automatinis koduotojas, tai yra automatinio koduotojo variantas, kuris naudoja konvoliucinius sluoksnius kodavimo ir dekodavimo procesuose, dažnai naudojamas vaizdų apdorojime.

LSTM (angl. *Long Short-Term Memory*) - ilgalaikės ir trumpalaikės atminties tinklas, tai yra specialus rekurentinis neuroninis tinklas, kuris turi gebėjimą išsaugoti informaciją per ilgesnį laikotarpį ir tinkamai atkurti ją prognozuojant sekantis elementas sekos.

FR – funkcinis reikalavimas

NFR – nefunkcinis reikalavimas.

PA – panaudojimo atvejis.

### **Terminai:**

Dirbtinis neuroninis tinklas – tarpusavyje sujungtų dirbtinių neuronų grupė. Ši technologija mėgdžioja žmogaus galvos smegenų darbą – tiksliau neuronų veikimą.

Deepfake – giluminiai vaizdai yra sintetinė laikmena, kurioje esamas vaizdas ar vaizdo įrašas pakeičiamas kažkieno panašumu.

API – aplikacijų programavimo sąsaja

Keitimas – procesas, kurio metu naudojant mašininio mokymosi modelius žmogaus vaizdas ir/ar garsas, esantis faile (garso, vaizdo ar vaizdo įrašo) pakeičiamas kito žmogaus vaizdu ir/ar garsu.

Failas – garso, vaizdo ar vaizdo įrašo failas, kuris yra naudojamas keitimui arba yra keitimo rezultatas.

Modelis – mašininio mokymosi modelis naudojamas keitimo procese.

Garso kadras – informacija, kurioje pateikiama amplitudės (garsumo) informacija tam tikru laiko momentu.

Fonema – tai garso vienetas, kuriuo galima atskirti vieną žodį nuo kito tam tikroje kalboje.

Enkoderis/Koduotojas (angl. *encoder*) – algoritmas ar sistema, kuri konvertuoja tam tikrą informaciją į kitą formatą.

Dekoduotojas (angl. *decoder*) – algoritmas ar sistema, kuri atkoduoja koduotą informaciją ir grąžina ją į pradinį formatą.

Vokoderis (angl. *vocoder*) – algoritmas, kuris naudojamas garso sintezei ir kodavimui. Jis susideda iš enkoderio ir dekoduoja daly. Enkoderio dalis analizuoja ir koduoja garso signalą, o dekoduoja dalis dekoduoja ir sintezuoja garso signalą iš koduotų duomenų.

Diskriminatorius (angl. *discriminator*) – algoritmas ar modelis, naudojamas klasifikacijos ir atskyrimo užduotims.

Konvoliuciniai sluoksniai – giliųjų neuroninių tinklų komponentai, kurie naudoja konvoliucijos operaciją tam, kad išgautų svarbius požymius iš įvesties duomenų.

Auto-koduotojas (angl. autoencoder) - algoritmas, kuris naudojamas koduoti įvesties duomenis mažesniu matmeniu ir tada atkurti juos į pradinį formatą.

Rekonstrukcijos paklaidų minimizavimas (angl. Reconstruction error minimization) - proceso tikslas sumažinti skirtumą tarp originalių ir atkurtų duomenų.

Softmax regresija veido atpažinimui (angl. Softmax regression for face recognition) - klasifikavimo metodas, naudojamas veidų atpažinimui, kuris priskiria tikimybinės reikšmės veidų kategorijoms pagal *softmax* funkciją.

Nuostolių funkcija - matematinė funkcija, kuri apibrėžia skirtumą tarp prognozuotų ir tikrų reikšmių, dažnai naudojama optimizavimo algoritmuose.

Interpoliavimo metodas - metodas, kuris leidžia prognozuoti reikšmes tarp esamų duomenų taškų pagal matematinį modelį arba priartinti duomenų reikšmes.

Validavimas - procesas, kuriame įvertinama modelio tikslumo arba veikimo patikimumo laipsnis naudojant atskirus duomenis, kurie nėra naudojami mokymo procese.

*Adabelief* optimizatorius - optimizavimo algoritmas, kuris derina *Adagrad* ir *Adam* optimizavimo metodų privalumus.

Mel-spektrograma - spektrogramos variantas, kuris atlieka dažnių skalę transformaciją į mel skalę, atitinkančią žmogaus girdimo savybes.



## **Įvadas**

Šiuo metu yra kuriamos bei tobulinamos sistemos, kurios gali sugeneruoti ar pakeisti žmogaus balsą ar atvaizdą kitu, tačiau ši sritis dar ne iki galo išplėtota bei iširta bei pateikta plačiajai visuomenei. Tokia sistema gali būti naudojama tiek pramogoms, pavyzdžiui, atkurti vyresnių kartų giminaičius jaunesniajai kartai, tiek kaip priemonė sutaupyti laiką, pavyzdžiui, tėvai galėtų ne skaityti knygas vaikams prieš miegą, o paleisti vaikams programą, kuri jiems knygą skaitytų panašiu balsu. Taip pat tokia sistema gali būti panaudojama mokymo tikslais, padarant pamokas interaktyvesnėmis, pavyzdžiui, istorijos pamokose atgaivinti garsius žmones.

„Deepfake“ technologija jau yra plėtojama ir naudojama tiek asmeniams tikslams, tiek su pelnu susijusiais tikslais. Ateityje ši technologija gali būti dar plačiau ir giliau naudojama kuriant medijų turinį, reklamas, filmus.

### **Priežastis**

Šiuo metu žinių kiekis, kaip atlikti tyrimą, yra minimalus. Ši apžvalga yra reikalinga praplėsti žinias apie dirbtiniu intelektu paremtą veido ir balso keitimą kitu bei, remiantis šiomis žiniomis, atlikti tyrimą bei išsiaiškinti, kurias technologijas geriausiai naudoti norint kokybiškai pakeisti žmogus veidą ir bals kitu.

### **Tikslas**

Šio darbo tikslas – sukurti sistemą, kuri gali naudoti mašininio mokymosi modelius žmogaus veido ir / ar balso keitimui kitu asmeniu bei išsiaiškinti, kokios mašininio mokymosi architektūros šiuo metu yra naudojamos keičiant vieno žmogaus veidą ir balsą ir jas palyginti

## **2. Projekto paraiška**

### **2.1. Poreikis**

#### **2.1.1. Projekto naudotojai ir klientai**

„Deepfake“ technologija naudojama falsifikuotam turiniui sukurti, pakeičiant ar sintetinant veidus, kalbą ir manipuliuojant emocijomis. Ji naudojama norint skaitmeniniu būdu imituoti asmens veiksmą, kurio jis nepadarė ar kurio jis negali atlikti. Kuriamą produktą ir naudotų žmonės ir įmonės, kuriems reikia pakeisti veidą ir / ar balsą savo asmeniniams ar verslo tikslams.

#### **2.1.2. Naudotojų rolės ir tikslai**

#### **2.1.3. Sistemos administratorius**

Auditorija: žmogus (žmonės), kuris matys bei galės perduoti sistemos vartotojų atsiliepimus ir valdys AI modelius.

Tikslas – prižiūrėti sistemą ir jos AI modelius.

Bendros charakteristikos:

- turi sistemų administravimo patirties.

Dalyvavimas projekte: -.

#### **2.1.4. Registruoti plačiosios visuomenės nariai**

Auditorija: plačioji visuomenė, kuriai patiko kuriama sistema ir kuri nori pakeisti vienu žmonių veidą bei balsą į kitų.

Tikslas – naudojantis sistema pakeisti žmogaus veidą bei balsą savo reikmėms bei saugoti naujai sukurtus bei originalius įrašus.

Bendros charakteristikos:

- domisi "deepfake" technologijomis arba dirbtiniu intelektu;
- sistema naudojasi ne pirmą kartą;
- norintys mokėti už papildomas arba platesnes paslaugas.

Dalyvavimas projekte: sukūrus demonstracinę versiją galės nemokamai pradėti naudotis sistema bei teikti atsiliepimus.

#### **2.1.5. Neregistruoti plačiosios visuomenės nariai**

Auditorija: plačioji visuomenė, kuriai patiko kuriama sistema ir kuri nori pakeisti vienu žmonių veidą bei balsą į kitų.

Tikslas – naudojantis sistema pakeisti žmogaus veidą bei balsą savo reikmėms bei saugoti naujai sukurtus bei originalius įrašus.

Bendros charakteristikos:

- sistema naudojasi pirmą kartą.

Dalyvavimas projekte: sukūrus demonstracinę versiją galės nemokamai pradėti naudotis sistema bei teikti atsiliepimus.

### 2.1.6. Rinkos tyrimas

Šia sistema bus galima naudotis įvairiais tikslais (šiuo metu ir ateityje), tačiau galima išskirti šias sritis, iš kurių galima užsidirbti parduodant sistemos sukurtas paslaugas:

1. Švietimo sistema – „*deepfake*“ technologija suteikia daug galimybių švietimo srityje. Mokyklos ir mokytojai jau kurį laiką klasėje naudoja garso ir vaizdo įrašus. „*Deepfakes*“ gali padėti mokytojui vesti novatoriškas pamokas, kurios yra daug patrauklesnės nei tradiciniai vaizdo ir medijos formatai. AI sukurta medija gali prikelti istorinius asmenis į interaktyvesnę klasę. Vaizdo įrašai apie istorinės asmenybės gali turėti didesnę poveikį ir bus geresnė mokymo(si) priemonė. Vien Europos Sąjungoje yra 17,7 mln. mokinių vidurinėse mokyklose (2018 metais) [1].
2. Menas – šiais laikais sudėtingos CGI ir VFX technologijos naudojamos filmuose siekiant papasakoti patrauklesnę istoriją. Tuo tarpu „*deepfake*“ gali pakeisti brangią VFX technologiją bei leisti sukurti daugiau medijos už mažesnę kainą. Kiekvienais metais vien JAV ir Kanadoje yra sukuriama keli šimtai filmų [2].
3. Įgarsinimui – garso pasakojimas ir knygų pasakojimas yra dar vienas geras sintetinio balso naudojimo atvejis. Autoriaus sintetinis balso šriftas gali būti naudojamas autoriaus knygos garso formatui sukurti. Įmonės gali naudoti sintetinius to paties filmo ir animacinių filmukų veikėjų balsus įvairiomis kalbomis, taip keldami sau pelną. Įgarsinimo kaina svyruoja tarp 50–400 dolerių už valandą [3], o naudojant „*deepfake*“ technologiją balsų keitimui galima turėti daug įrašų už mažesnę kainą. O 2020 metais vien JAV jų buvo įgarsinta daugiau kaip 71 tūkstantis [4].
4. Reklamavimas – „*deepfake*“ gali naudoti vieši asmenys (*angl. influencers*), norėdami išplėsti savo pasiekiamumą ir padidinti auditoriją. Naudodamas „*deepfake*“, prekės ženklas gali pasiekti daug klientų, naudodami labai tikslinius ir suasmenintus pranešimus. „*Deepfakes*“ sukurta medija gali tapti nauja mados ir prekės ženklo rinkodaros tendencija.

### 2.1.7. Informacija apie klientus

Potencialūs programinės įrangos naudotojai bus žmonės ir įmonės, kuriems reikia pakeisti veidą ir / ar balsą savo asmeniniams ar verslo tikslams. Daugiau apie įmonių sritis, kurios potencialiai gali naudoti kuriamą produktą, aprašyta 2.1.6 skyrelyje.

## 2.2. Pasiūlymas

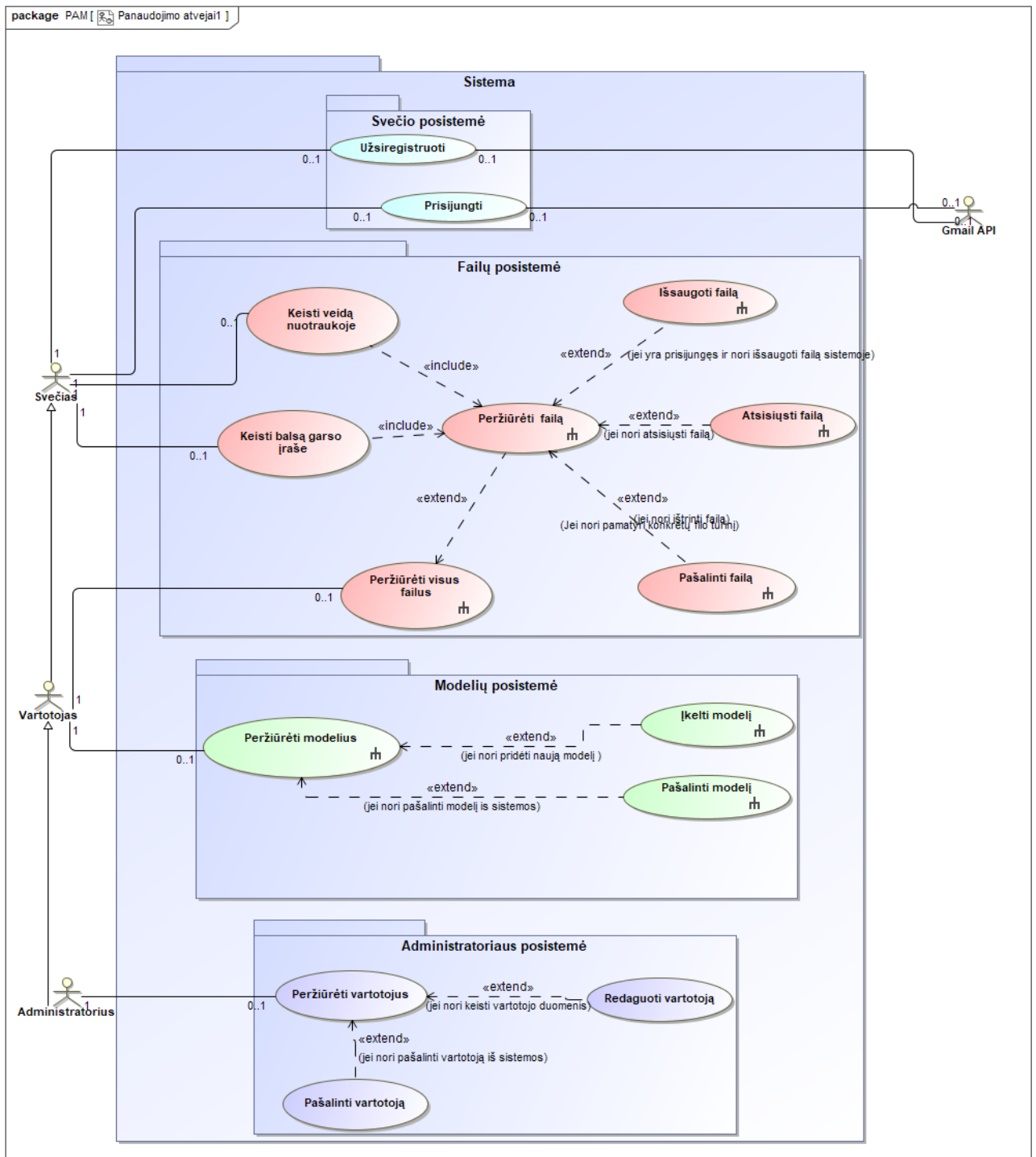
### 2.2.1. Produkto ar paslaugos apibūdinimas

Trumpai apibūdinant kuriamos sistemos veikimą, sistema bus skirta tyrimams su neuroninių tinklų modeliais atlikti bei pakeisti pasirinktą veidą ir / ar balsą kitu. Sistema skirta trijų tipų naudotojams: administratoriams, prisijungusiems bei neprijungusiems vartotojams:

1. Svečias:
  - a. Tikslas – naudojantis sistema pakeisti žmogaus veidą bei balsą savo reikmėms bei atsisiųsti sukurtus failus;
  - b. Funkcijos:
    - i. prisijungti;
    - ii. užsiregistruoti;
    - iii. įkelti paveiksliuko ar garso įrašo failą bei atsisiųsti naujai sugeneruotą failą.
2. Prisijungęs vartotojas:

- a. Tikslas – naudojantis sistema pakeisti žmogaus veidą bei balsą savo reikmėms bei saugoti naujai sukurtus bei originalius įrašus;
  - b. Funkcijos:
    - i. įkelti vaizdo ar garso įrašo ar paveiksluko failą bei atsisiųsti ar išsaugoti naujai sugeneruotą failą pasirinkus norimą modelį;
    - ii. matyti bendrinius bei savo įkeltus modelius, įkelti naują modelį ar ištrinti anksčiau savo įkeltą modelį.
3. Administratorius:
- a. Tikslas – prižiūrėti sistemą ir jos AI modelius ir vartotojus;
  - b. Funkcijos:
    - i. matyti visus modelius, įkelti naują modelį ar ištrinti esamą modelį;
    - ii. matyti visus vartotojus, keisti jų duomenis, ištrinti esamą ar pridėti naują vartotoją.

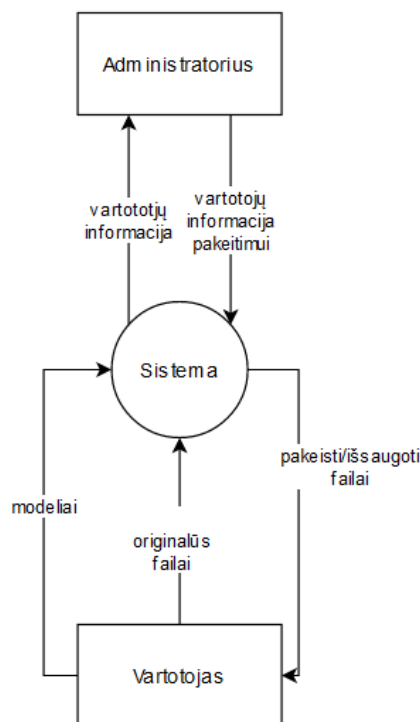
Detalesnės sistemos funkcijos yra aprašytos panaudojimų atvejų diagramoje 2-1 pav.:



2-1 pav. Panaudojimo atveju diagrama

### 2.2.2. Sistemos kontekstas

2-2 pav. pavaizduota sistemos konteksto diagrama, kurioje administratorius iš sistemos gauna vartotojų informaciją, o jai teikia pakeistą informaciją pakeitimui. Tuo tarpu vartotojas sistemai teikia failus bei modelius, o iš sistemos gauna pakeistus failus.



2-2 pav. Konteksto diagrama

### 2.2.3. Techninės galimybės

Atlikti šiai sistemai reikės:

1. Kompiuterio su galinga vaizdo plokšte, tinkama giliajam mokymui (pvz. *RTX 2060 Super* ar *RTX 2080 ti*) arba naudoti mokymo paslaugas debesyse (pvz. *AWS*);
2. Kompiuterio su įrašyta IDE, su kuria galima sukurti informacinę sistemą.

### 2.2.4. Rizika ir apribojimai

Apribojimai:

- Sistema turi sugeneruoti nuotrauką per mažiau nei 10 sekundžių, o garso įrašą – per mažiau nei 150 % originalaus įrašo laiko.
- Sistema turi būti lengva naudotis, t. y. puslapio elementai turi būti aiškiai išdėstyti, suprantamai aprašyti bei atlikti tą funkciją, kurią aprašo.
- Sistema turėtų apdoroti *jpeg* ir *png* formato nuotraukas, *mp3* formato garso įrašus.

Rizikos, problemos ir jų sprendimo būdai:

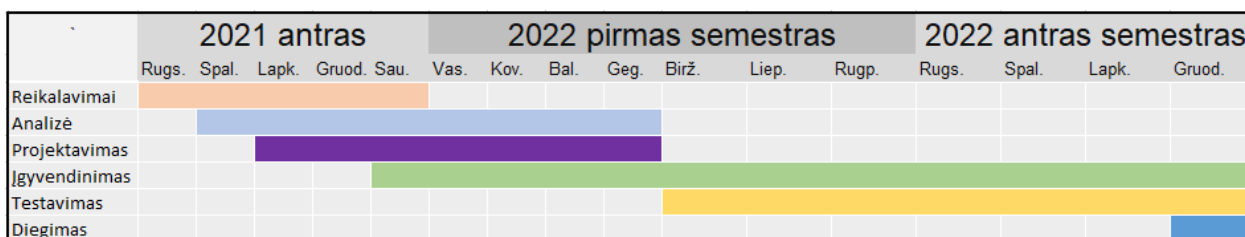
- Konkurentai sukuria tą patį ar net geriau – kad taip neatsitiktų, sistemos kūrimo metu bei sukūrus sistemą stebėti tiek konkurentus, naujienas apie sistemoje naudojamą technologijas bei pagal poreikį tobulinti sistemą. Taip pat projektuojant sistemą numatyti laiko tokiam sistemos tobulinimui.
- Licencijų / įstatymų pokyčiai – kuriant sistemą bei ją įdiegus nuolat atnaujinti licencijas bei projektuojant sistemą numatyti laiko prisitaikyti prie netikėtų licencijų / įstatymų pokyčių.

- Kriptografinės ar fizinės atakos / įsilaužimas – kadangi sistema egzistuos debesyse, fizinių atakų galima nesitikėti. Norint išvengti kriptografinių atakų kuriant sistemą, reikia naudoti patikimas ir saugias technologijas bei diegiant sistemą nepalikti spragų;
- Neteisingas vartotojo elgesys – sistema, patyrusi neteisingą vartotojo elgesį, turi grąžinti aiškią žinutę vartotojui kas blogai ir, jei įmanoma, kaip vartotojui elgtis, norint pasiekti jo norimą rezultatą

### 2.2.5. Projekto įgyvendinimo planai ir kokybės vertinimas

Projekto įgyvendinimo planas pavaizduotas 2-3 pav.:

1. Reikalavimų apibrėžimas – nustatoma pradinė galutinio produkto vizija, pagrindinės funkcijos, kurias sistema turės atlikti (nubraižoma pirminė panaudojimo atvejų diagrama).
2. Analizė – analizuojamam veiklos sritis bei galimos priemonės. Pasirenkamos geriausios priemonės bei detalizuojami panaudojimo atvejai, jų ryšiai tarpusavyje.
3. Projektavimas – suprojektuojama kuriamos sistemos architektūra, nustatoma sistemos struktūra iš posistemių, klasių ir sąsajų, kuriuose paskirstyti panaudojimo atvejai.
4. Realizavimas – realizuojama suprojektuota sistema.
5. Testavimas – sukuriamas testavimo planas bei pagal jį ištestuojama sistema ir, reikalui esant, pataisoma sistema.
6. Diegimas – sistema įdiegiama bei padaroma prieinama visiems.
7. Palaikymas – sistema nuolat prižiūrima, bei, reikalui esant, pataisoma.



2-3 pav. Projekto vykdymo grafikas

Kuriamos sistemos kokybės įvertinimo būdai pateikiami 2-1 lentelė.

2-1 lentelė. Sistemos vertinimo kriterijai

Nr.	Kriterijus	Pagrindimas
1	Išplečiamumas	Numatytos galimybės nesudėtingai pridėti naują funkcionalumą
2	Panaudojamumas	Per kiek laiko IT raštingas naudotojas sugebės išmokyti dirbti su sistema
3	Interaktyvumas	Ar vartotojas gauna grįžtamąjį ryšį iš sistemos
4	Tikslumas	Vartotojas, gavęs pakeistą failą iš sistemos, turi atpažinti sugeneruotą veidą / balsą.

Projektui įgyvendinti reikalingas personalas:

- Vienas programuotojas (Indrė Dimšė), turintis programų sistemų srities bakalauro laipsnį. Su vadovo pagalba aprašo, suprojektuoja, ištestuoja ir realizuoja sistemą. Šiam asmeniui priklausys sukurto produkto autorinės teisės.

- Vadovas (Rytis Maskeliūnas), turintis informacinių technologijų srities daktaro laipsnį .  
Vadovauja kūrimo procesui, padeda studentui su išskylančiomis problemomis ir klausimais.

Projekto biudžeto skaičiavimas pateiktas 2-2 lentelė.

**2-2 lentelė.** Projekto biudžetas

Išlaidos	Vienetas	Vienetų skaičius	Vieneto kaina, Eur	Viso, Eur
Žmonių ištekliai				
Vadovas	Mėnesis	18*0,05 etato	3300	2700
Programuotojas	Mėnesis	18*0,2 etato	1800	6,480
<i>Iš viso žmonių išteklių</i>				9180
2. Įranga ir prekės				
Kompiuteris	Vienetas	1	799,99	799,99
<i>Iš viso įranga ir prekės</i>				799,99
6. Iš viso tiesioginiai projekto kaštai				9,979.99

### 2.3. Nauda

Projekto laukiami rezultatai:

1. tikroviškas veido pakeitimas nuotraukose per 10 sekundžių,
2. tikroviškas balso pakeitimas garso įrašuose per 150 % garso įrašo laiko;

Šių rezultatų laukiama iki studijų trečio semestro pabaigos – 2023-12-23.

### 2.4. Konkurencija ir alternatyvos

Lyginami konkurentai:

- *AutoVC* [1];
- *Deep Nostalgia* [2];
- *ConVoice* [3].

Kaip matome 2-3 lentelė, veido keitimą turi tik *Deep Nostalgia*, o *AutoVC* ir *ConVoice* turi tik balso keitimą, tačiau nei viena neturi abiejų. Taip pat nors *AutoVC* ir *ConVoice* dalinasi kodu, tačiau *Deep Nostalgia* kodas nepasiekiamas. Taip pat nei vienas konkurentas savo internetiniame puslapyje neturi galimybės įkelti atvaizdo, pagal kurį keisti balsą / veidą ir tik *Deep Nostalgia* turi galimybę įkelti originalą, kurį pakeisti.

**2-3 lentelė.** Konkurentų analizė

	<i>AutoVC</i>	<i>Deep Nostalgia</i>	<i>ConVoice</i>	Kuriama sistema
Veido keitimas	Ne	Taip	Ne	Taip
Balso keitimas	Taip	Ne	Taip	Taip
Prieinami modeliai / kodai	Taip	Ne	Taip	Dalinai
Galima įkelti originalą	Ne	Taip	Ne	Taip
Galima įkelti pagal kurį keisti	Ne	Ne	Ne	Taip



### 3. Literatūros analizė

#### 3.1. Literatūros šaltinių paieškos ir apžvalgos metodika

##### 3.1.1. Tinkamumo kriterijai

Straipsnio antraštė ir jos anotacija/aprašymas turi būti apie veido/balso keitimo, „*deepfake*“ technologijas bei straipsniai turi būti iš patikimų šaltinių ir jau cituoti kitų autorių.

##### 3.1.2. Informacijos šaltiniai

Informacijos šaltiniai, kuriuose bus vykdoma paieška:

- <https://www.ieee.org/>;
- <https://www.scopus.com/home.uri>;
- <https://www.acm.org/>;
- <https://www.elsevier.com>;
- <https://scholar.google.com/>;
- <https://www.webofscience.com>;

##### 3.1.3. Paieškos strategija

Kadangi „*deepfake*“ technologija yra gana nauja ir greitai tobulėjanti, bus imami 5 metų senumo šaltiniai. Taip pat bus orientuojamasi į šiuos raktažodžius vykdant šaltinių paiešką:

- *face deepfake*;
- *voice deepfake*;
- *neural face swapping*;
- *neural voice conversion*;
- *voice substitute deep learning*;
- *voice substitution*;
- *face substitution*.

Paieška bus vykdoma anglų kalba, kadangi ji plačiau naudojama ir tikėtina rasti daugiau rezultatų, bei paieška bus vykdoma mokslo publikacijose pagal anksčiau išvardintus terminus.

##### 3.1.4. Duomenų rinkimo procesas

Buvo pasirinktas vieno žmogaus kiekvienos antraštės peržiūrėjimo būdas. Renkantis tinkamus straipsnius buvo kreipiamas dėmesys į antraštę – daug straipsnių nors ir turėjo tinkamus raktažodžius, tačiau jų visos antraštės reikšmė neatitiko tinkamumo kriterijų. Taip pat abejojant straipsnio tinkamumu turi būti perskaityta anotacija/aprašymas.

##### 3.1.4. Duomenys

Renkantis straipsnius buvo išsaugoma straipsnio autorius, metai, leidimo šaltinis bei ISSN numeris. Šie duomenys, priklausomai nuo šaltinio, buvo išsaugomi į *csv* arba *html* formato failus.

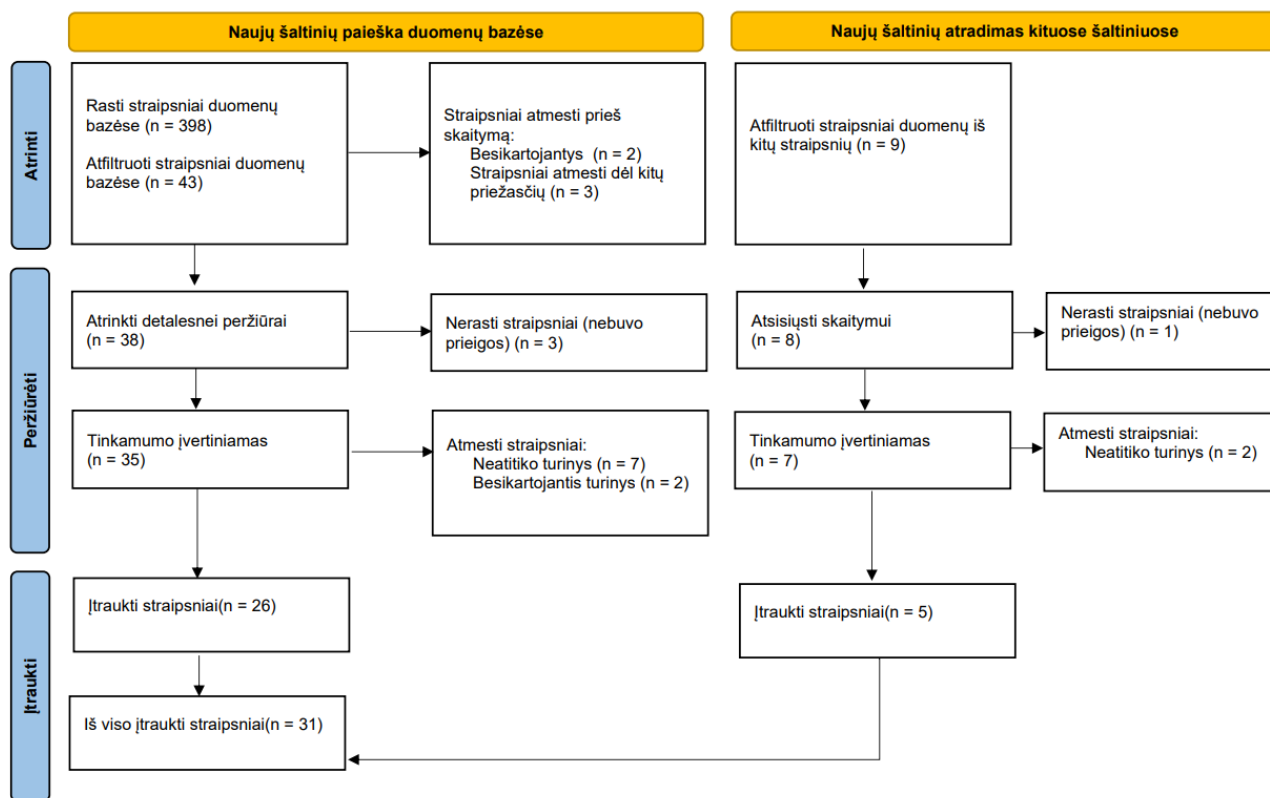
##### 3.1.5. Tikrumo įvertinimas

Straipsnių įrodymų tikrumas vertinamas pagal 5 *GRADE* svarstymus (tyrimo apribojimus, poveikio nuoseklumą, netikslumą, netiesioginumą ir publikavimo šališkumą), kad įvertintume įrodymų

visumos tikrumą, nes jis buvo susijęs su tyrimais. Įrodymų tikrumą įvertinome kaip didelį, vidutinį, žemą arba labai žemą.

### 3.1.6. Šaltinių atranka

3.1 pav. **3-1 pav.** pavaizduota *PRISMA* atrankos diagrama, kurioje pavaizduota kiek straipsnių buvo iš viso peržiūrėta, kiek ir kodėl jų buvo atmesta bei kiek jų įtraukta į literatūros analizę.



3-1 pav. PRISMA atrankos diagrama

## 3.2. Veido keitimo metodai

### 3.2.1. Veido keitimo mašininio mokymosi algoritmai

Veido pakeitimas kitu yra ganėtinai sudėtingas procesas – tam reikia nustatytas veidą nuotraukoje, jį išcentruoti bei segmentuoti tam, kad užtikrinti, kad po veido keitimo veidas yra tinkamoje pozicijoje. Veido keitimo metodų palyginimas yra 3-1 lentelė. Veido keitimo algoritmai. Joje pavaizduoti 3 *GAN* šeimos algoritmai, *FSNet* algoritmas bei metodas, nenaudojantis papildomų bibliotekų – tik *OpenCV* ir *NumPy*. *GAN* šeimos algoritmai gali apdoroti didelės raiškos nuotraukas bei yra gana plačiai naudojama nuotraukų sintezėje [20]. *FSNet* nors irgi naudojama *GAN* architektūra, tačiau labiau remiamasi *CNN* [21].

3-1 lentelė. Veido keitimo algoritmai

	Naudojamas VAE ir GAN	Aukštos kokybės nuotraukos	Komentaras
<i>FS-GAN</i>	+	-	Paprasta VAE ir GAN įgyvendinimas

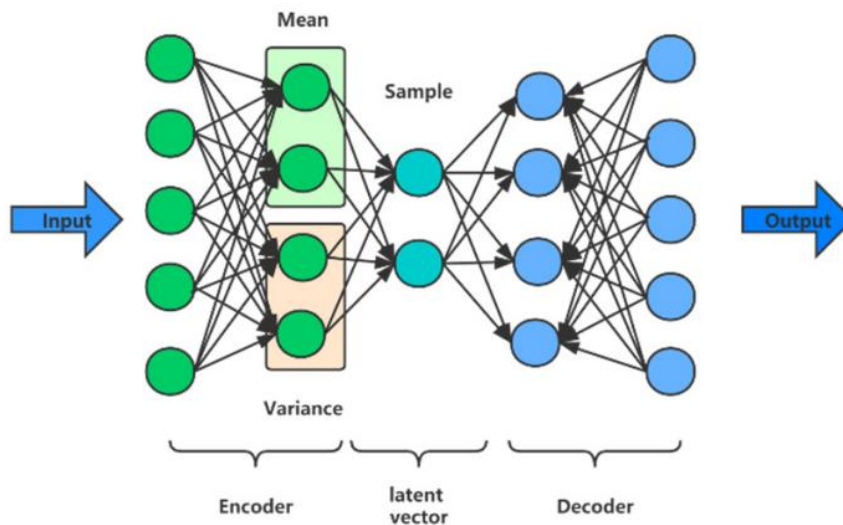
<i>FCC-GAN</i>	+	+	Naudojamas <i>conv2D</i> bei gaunami natūralesni rezultatai su daugiau detalių
<i>RSGAN</i>	+	+	Atskirai atkuria veidą ir plaukus, todėl galima naudoti ir plaukų perkėlimui bei yra naudojami du <i>VAE</i>
<i>FSNet</i>	+	-	Atskiriama veido – beveidė dalis ir jos atskirai keičiamos
Keitimas realiu laiku	-	-	Naudojama tik <i>OpenCV</i> ir <i>NumPy</i> . Keičiama greitai, tačiau nepalaikoma kokybė

### 3.2.2. GAN

Manipuliuojant veidu nuotraukose dažniausiai naudojami *VAE* ir *GAN* modeliai [5].

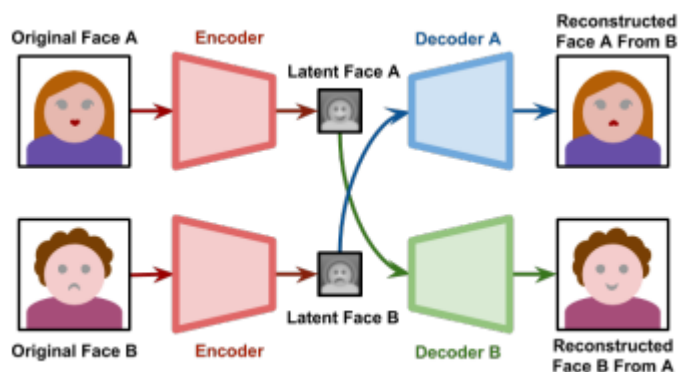
Automatinis kodavimo įrenginys (angl. *Variational auto-encoder*) – tai dirbtinio nervinio tinklo dalis, kurio kodavimo paskirstymas yra sureguliuotas mokymo metu, siekiant užtikrinti, kad jo latentinė erdvė turėtų geras savybes, leidžiančias generuoti naujus duomenis. Kaip pavaizduota 3-2 pav. Automatinio kodavimo įrenginiai priima ir paskutiniam sluoksny turi tiek pat mazgų, tačiau vidury turi labai stiprų susiaurėjimą bei gali būti naudojamas:

- Sutraukti (angl. *compress*) nuotraukas, nors tai ir nėra geriausias šio įrenginio panaudojimo būdas;
- Pašalinti triukšmą iš nuotraukos;
- *VAE* gali generuoti panašias nuotraukas.



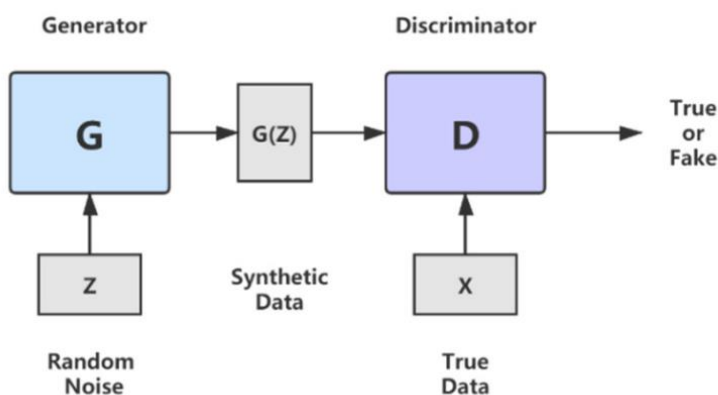
3-2 pav. *VAE* struktūra [1]

Taip pat galiam automatinis kodavimo įrenginį galima naudoti veidų sukeitimams, kaip parodyta pav. 3-3 pav. [6]. Tačiau norint tikroviškesnę vaizdą, t. y. tikslesnę ribą tarp veido ir fono, rekomenduotina naudoti veido segmentavimą nuotrauką praleidžiant per automatinį kodavimą [10].



3-3 pav. Automatinis kodavimo veikimas apkeičiant veidus [6]

Generatyvusis priešininkų tinklas (angl. *Generative adversarial networks*) – tai mašininio mokymo klasė, kurioje 2 neuroniniai tinklai moko vienas kitą.



3-4 pav. GAN struktūra [1]

GAN architektūros [6]:

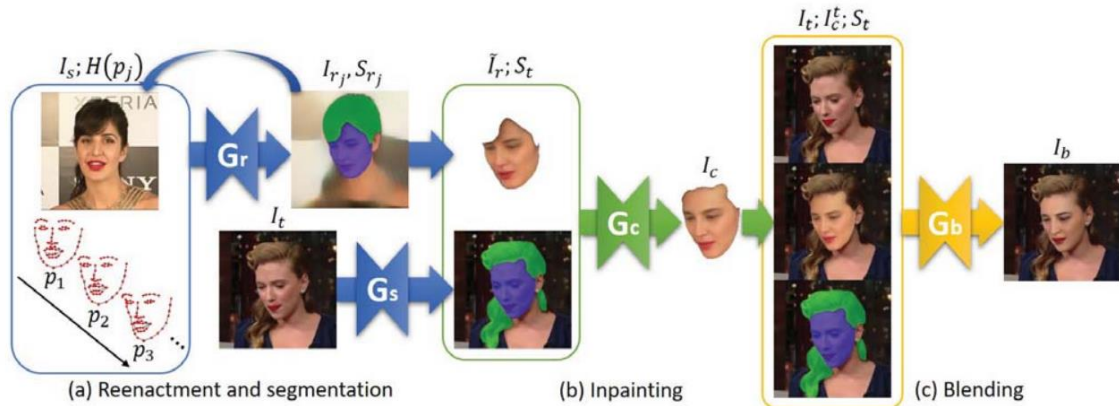
- TARGAN
- ATTGAN
- STYLEGAN
- STYLEGAN2 (Kerras patobulinta STYLEGAN)
- FSGAN
- FCC-GAN

### 3.2.3. FS-GAN

FS-GAN – tai veido keitimo generatyvusis priešininkų tinklas (angl. *Face Swap GAN*), kurio veikimas pavaizduotas 3-5 pav. Sakykime, kad  $I_s$  yra veido nuotraukų šaltinis, o  $I_t$  – tikslas. FS-GAN turi 3 komponentus:

1. Pirmas susideda iš atkūrimo generatoriaus  $G_r$  ir CNN segmentatoriaus  $G_s$  (3-5 pav. a dalis).  $G_r$  įvertina atkurtą veidą  $F_r$  (atkurtas veidas) ir jo segmentaciją  $S_r$ , o  $G_s$  įvertina veido ir plaukų segmentavimo kaukę (3-5 pav. b dalis).
2. Atkurta nuotrauka gali turėti trūkstamų veido dalių, todėl pritaikome veido dažymo tinklą (angl. *Face inpainting network*)  $G_c$  naudodami ir segmentavimą nustatome trūkstamus pikselius.

3. Galiausiai įliejame sukurtą veidą  $F_c$  į tikslo nuotrauką  $I_t$  (3-5 pav. c dalis).

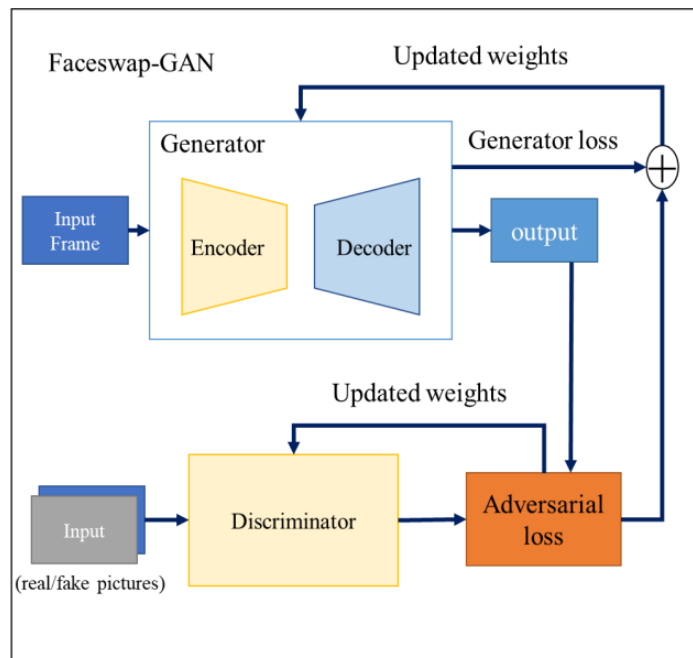


3-5 pav. FSGAN veikimas [6]

Šis metodas apsiriboja apmokymo metu duomenų rinkinio nuotraukų rezoliucija.

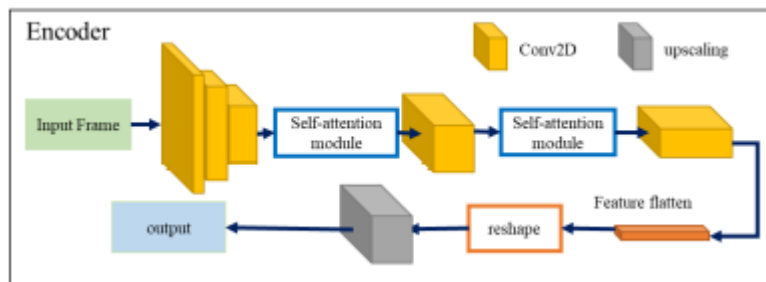
### 3.2.4. FCC-GAN

FCC-GAN – tai pilnai sujungtas ir konvoliucinis GAN. (angl. *A Fully Connected and Convolutional Net Architecture for GAN*). Jis sukurtas remiantis FCGAN, tačiau pasiekia labiau natūralius rezultatus. Originalioje architektūroje self-attention moduliai paprastai konvertuoja veido bruožus iš šaltinio veido į tikslinį veidą su dirbtiniu iškraipymu aplink veidą funkcijos. FCC-GAN naudoja diskriminatorių kaip visiškai sujungtos konvoliucinės struktūros tinklą, todėl rezultatas gali būti sklandesnis ir natūresnis lyginant su rezultatais naudojant originalų veido keitimą. FCC-GAN veikimo schema pavaizduota 3-6 pav. [12]:



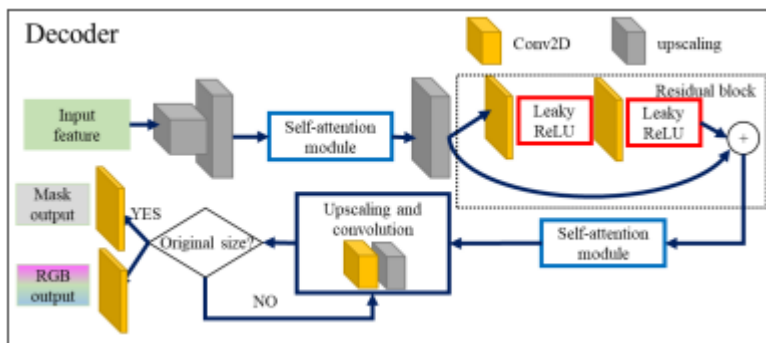
3-6 pav. FCCGAN veikimo schema [8]

3-7 pav. pavaizduotas koduotojas (angl. *encoder*), kuris susideda iš *conv2D* sluoksnių ir *self-attention* modulio derinio, kuris naudojamas tiksliai išryškinti veido bruožus tikslo veidui. Veido bruožų žemėlapis yra išryškinamas naudojant *conv2D* ir *pixelShuffler* metodą.

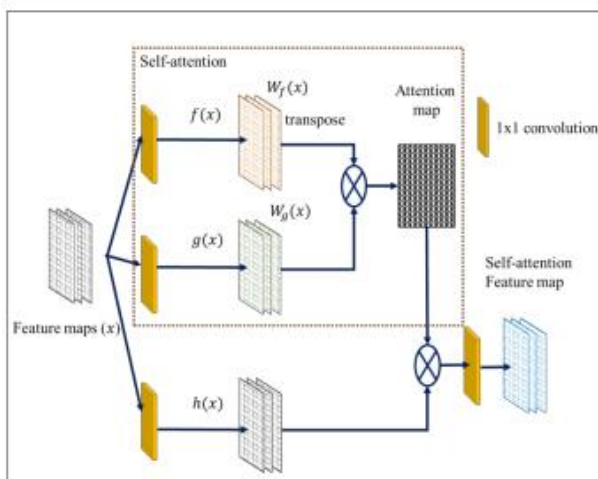


3-7 pav. FCCGAN koduotojo schema [15]

Dekoduotojo veikimo schema pavaizduota 3-8 pav., kurioje naudojama *leaky ReLU* aktyvinimo funkcija, siekiant išvengti išnykusių gradientų problemos. Nuslopinius triukšmą dekoduoja gali geriau atkurti žmogaus veidą.

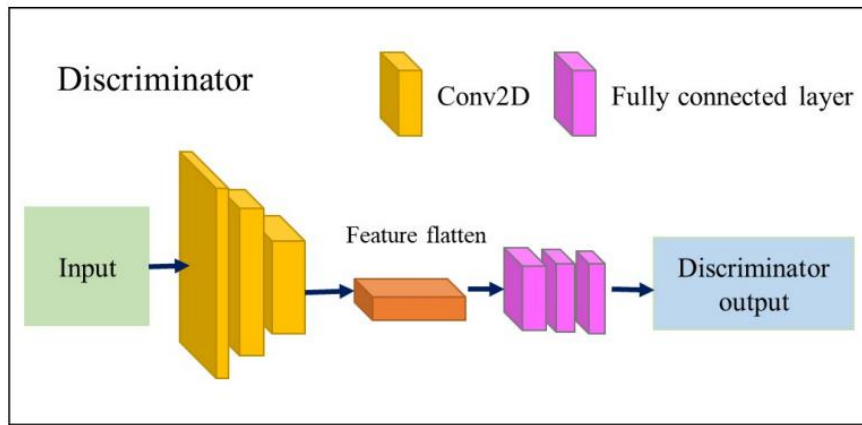


3-8 pav. FCCGAN dekoduoja schema [8]



3-9 pav. FCCGAN *self-attention* modulio schema [8]

3-9 pav. pavaizduota detalesnė diskriminatoriaus schema, kurioje matoma kad ir jame naudojamas konvoliucinis bei pilnai sujungtas tinklai. Dėl ko gali būti sukurtos ir geresnės kokybės nuotraukos, gali būti išvengta staigaus diskriminatoriaus nuostolio mažėjimo bei neturėjimas bendrų svorių gali padėti tinklui mokytis.



3-10 pav. Diskriminatoriaus schema

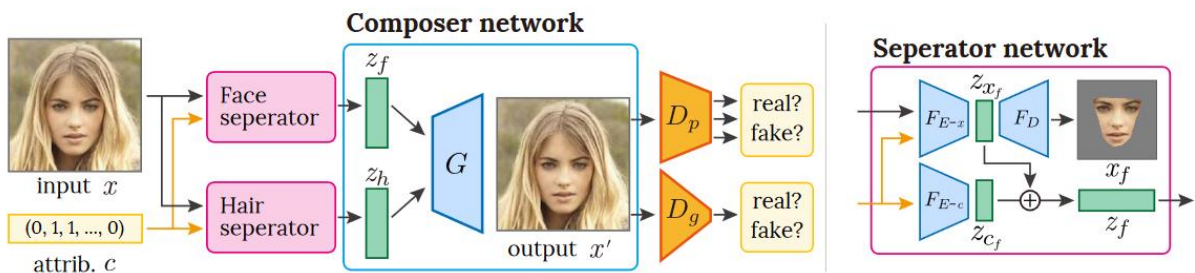
Norint pagerinti *FCCGAN* veikimą, galima išbandyti šiuos būdus [16]:

- Naudoti kelis giliai pilnai sujungtus sluoksnius priešais konvoliucinius sluoksnius generatoriuje;
- Naudoti kelis giliai pilnai sujungtus sluoksnius diskriminatoriuje pakeisti aukštos dimensijos bruožus išgautus konvoliuciniame sluoksnyje į žemesnės dimensijos bruožus;

### 3.2.5. Region-Separate GAN

Šio metodo tikslas yra pasiekti nuotraukos sukūrimą naudojant latentinės erdvės veido ir plaukų vaizdus. Kaip pavaizduota 3-11 pav. *RSGAN* architektūra, kuri sudaryta iš dviejų *VAE* (angl. *Face separator* ir *Hair separator*) bei vieno *GAN* (angl. *Composer network*) [14]:

- Separatorių tinklai išgauna latentinės erdvės mazgus  $z_f$  ir  $z_h$  veidui ir plaukams
- Kompozitoriaus tinklas atkuria įvesties veido vaizdą iš dvi latentinės erdvės mazgų. Rekonstruotas vaizdas  $x'$  ir įvesties vaizdas  $x$  yra vertinami dviejų diskriminatorių tinklų. Visuotinis diskriminatorius  $D_g$  išskiria ar vaizdas tikras, ar netikras, ir atskiria pataisų diskriminatorius  $D_p$  ar vietinės vaizdų segmentai yra tikri, ar netikri.



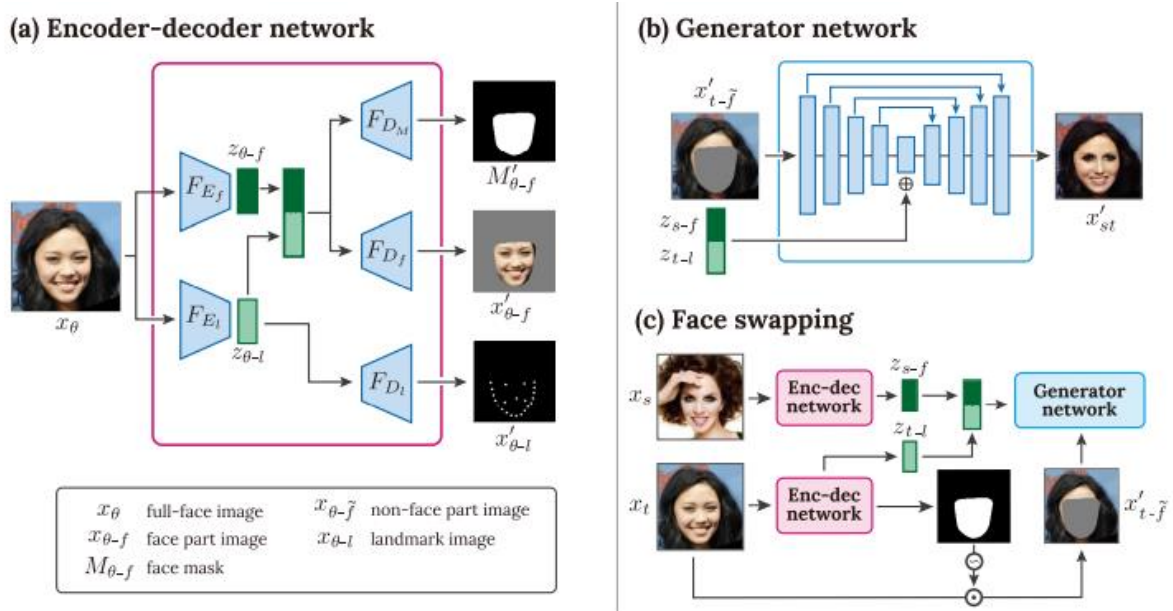
3-11 pav. *RS-GAN* veikimo schema [10]

### 3.2.6. FSNet

*FSNet* architektūra pavaizduota 3-12 pav. kuriame parodytas tinklas yra suskaidytas į dvi dalis, kurios veikia individualiai [13]:

1. Pirmoji dalis paremta anksčiau minėtu *VAE* bei vidurinis (latentinis) sluoksnis yra išsaugomas (3-12 pav. a)

2. Antroje dalyje generatorius susintetina naują veido dalį taip, kad ji atitiktų ne veido dalį tikslo nuotraukoje bei šioje dalyje naudojama *U-Net*. Kitaip nei įprastas *U-Net*, generatorius priima latentiniu kintamuosius ir sujungia juos su bruožų žemėlapiu, duotu viduriniame tinklo sluoksnyje. Taip pat generatorius priima ne įprastą nuotrauką, tačiau jau beveidę (3-12 pav. b ir c)



3-12 pav. FSNet veikimas

### 3.2.7. Keitimas realiu laiku

Norint veido keitimą atlikti kuo greičiau galima naudoti tik *OpenCV* ir *NumPy* bibliotekas su šiais žingsniais [9]:

1. Aptikti veidą nuotraukoje
2. Nuskaityti šaltinio nuotrauką
3. Nustatyti orientacinius taškus šaltinio nuotraukoje
4. Jei reikia, sulygiuoti šaltinio bei tikslo nuotraukų veidus
5. Iškirpti šaltinio veido kontūrus
6. Įdėti iškirptą veidą į tikslo nuotrauką pagal orientacinius taškus.
7. Galiausiai suliejama šaltinio veidas su tikslo nuotrauką

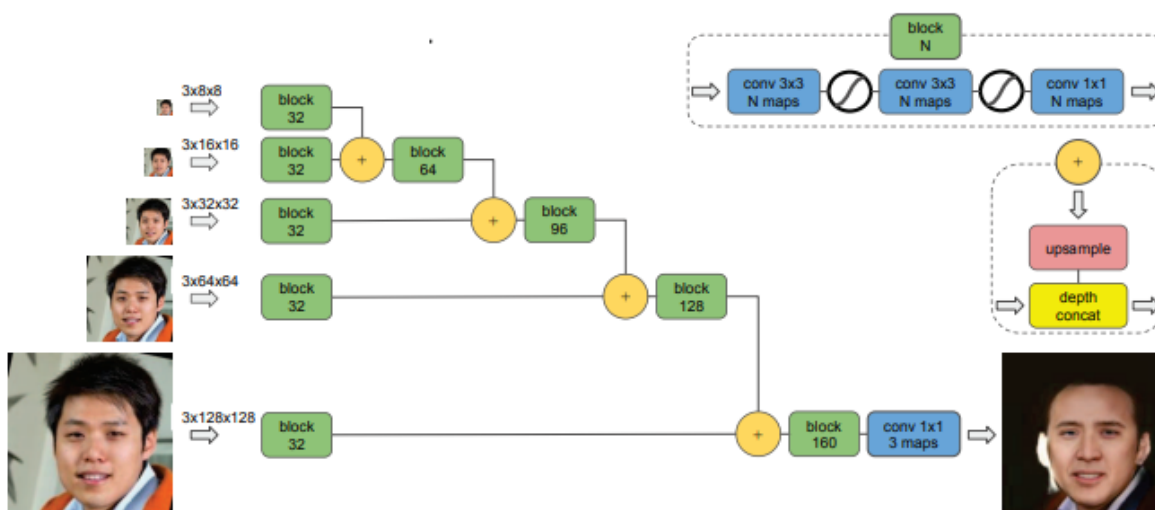


3-13 pav. Veidas prieš pakeitimą, uždėjus šaltinio nuotrauką ir įliejus veidą į nuotrauką [9]



### 3.2.8. Ulyanov

Naudojant *Ulyanov* transformacijos tinklo architektūrą sukuriama daug šakų (priklausomai nuo nuotraukos dydžio). Kiekviena šaka sukuria nulinius konvoliucinius tinklus, po kurių eina linijinis ištaisymas. Vėliau šakos yra sujungiamos naudojant artimiausio kaimyno atranką [11]. Kadangi net pats autorius minėjo, kad šiam metodai dar yra kur tobulėti, jos toliau nenagrinėjamas (3-14 pav.).



3-14 pav. *Ulyanov* architektūra [11].

### 3.2.9. Veido spalvos pritaikymas

Keičiant veidą svarbu ne tik pakeisti veido bruožus, bet taip pat pritaikyti odos spalvą. 3-15 pav. pavaizduotas spalvos perkėlimo algoritmas, kuris iš veido šaltinio  $S$  ištraukia veido spalvą bei perkelia ją į tikslą  $T$  pagal dvejetainę kaukę  $M$ . Šis algoritmas įgyvendinamas remiantis naudingųjų pikselių verčių vidurkiu ir standartinę nuokrypiu  $S$  ir  $T$  [8].

```

1: Convert image S and image T from RGB color space
to CIELAB color space.
2: Initialize four empty arrays:  $L_S, L_T, A_S, A_T, B_S, B_T$ .

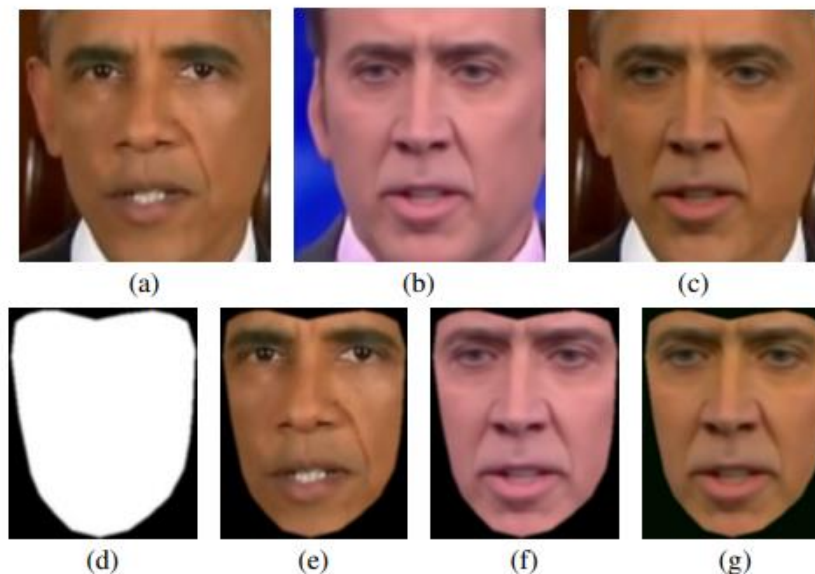
3: for  $i < height$  do
4:   for  $j < width$  do
5:     if  $M(i, j) == 255$  then
6:        $L_S.append(S_L(i, j));$ 
7:        $L_T.append(T_L(i, j));$ 
8:        $A_S.append(S_A(i, j));$ 
9:        $A_T.append(T_A(i, j));$ 
10:       $B_S.append(S_B(i, j));$ 
11:       $B_T.append(T_B(i, j));$ 
12:    end if
13:  end for
14: end for
15: Calculate the mean value of  $L_S, L_T, A_S, A_T, B_S, B_T$ 
and note them as  $L_{Sm}, L_{Tm}, A_{Sm}, A_{Tm}, B_{Sm}, B_{Tm}$ .
16: Calculate the standard deviation of  $L_S, L_T, A_S, A_T,$ 
 $B_S, B_T$  and note them as  $L_{Sstd}, L_{Tstd}, A_{Sstd}, A_{Tstd}, B_{Sstd},$ 
 $B_{Tstd}$ .
17: for  $ch \in [L, A, B]$  do
18:    $ch_T \leftarrow ch_T - ch_{Tm}$ 
19:    $ch_T \leftarrow ch_T \times ch_{Sstd} / ch_{Tstd}$ 
20:    $ch_T \leftarrow ch_T + ch_{Sm}$ 
21:   Clip/Scale the pixel intensities of  $ch_T$  to  $[0, 255]$  if
they fall outside this range.
22: end for
23: Convert image T from CIELAB color space to RGB
color space.

```

3-15 pav. Spalvos perkėlimo algoritmas [8]

3-16 pav. pavaizduotas veido spalvos perkėlimo algoritmo veikimas:

- šaltinio paviršius;
- numatomas taikinio veidas;
- šaltinio paviršius, pakeistas veidu;
- dvejetainė kaukė;
- šaltinio paviršiaus plotas;
- veidas numatomo taikinio veido plotas;
- spalvos perdavimo rezultatas nuo e iki f.



3-16 pav. Veido spalvos pakeitimo veikimas [8]

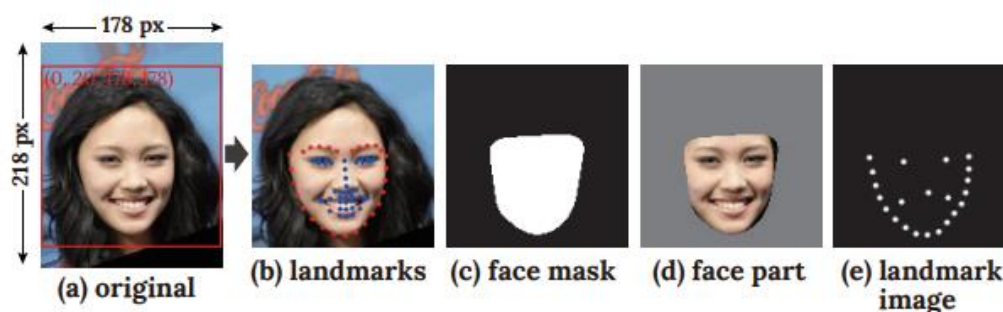
### 3.2.10. Netekties funkcijos

Netekties funkcijos:

- Perceptinė netektis (angl. *perceptual loss*) – išmokomi ir naudojami keli *VGG-19* tinklai veido atpažinimui bei veido savybių klasifikavimui. Plačiai naudojamas darbu su veido sinteze [6];
- Rekonstrukcijos netektis (angl. *reconstruction loss*) – veikia panašiai kaip prieš tai aprašytas, tačiau naudoja  $L_1$  netekties generatorių, dėl ko išvengiama netikslių spalvų [6];
- Priešinė netektis (angl. *adversarial loss*) – naudojamas daugiasektorinis diskriminatorius, kurį sudaro keli diskriminatoriai su skirtingomis vaizdo skiriamosiomis gebomis [6];
- Stiliaus netektis (angl. *style loss*) – naudojamas *VGG* tinklas, tačiau yra fiksuojama informacija apie spalvą/ tekstūrą skirtingais erdviniais masteliais [11];
- Asmenybės netektis (angl. *identity loss*) – metodas, kuris klasifikuoja nuotraukas pagal asmenybę. Tam reiktų, kad būtų naudojamas duomenų rinkinys su asmens etiketėmis (pavyzdžiui *CelebA*). Tačiau net ir tokiu atveju šis būdas nėra labai aptikimas [13].

### 3.2.11. Duomenų apdorojimas ir rinkiniai

Apdorojant veido nuotraukas galima išgauti veido orientacinius taškus naudojant *dlib* biblioteką [11] (3-17 pav. b) Pagal orientacinius taškus galima sukurti veido kaukę (3-17 pav. c) bei išgauti tik veido dalį nuotraukoje (3-17 pav. d). Taip pat vietoj įprastų orientacinių taškų galima naudoti 5 taškus veido bruožams (akims, nosiai, lūpoms) bei  $\pm 17$  taškų veido kontūrams (3-17 pav. e)



3-17 pav. Veido apdorojimas [16]

3-2 lentelė. Veido keitimo duomenų rinkinys

Duomenų rinkinys	Žmonių kiekis	Nuotraukų kiekis	Vid. nuotraukų žmogui
<i>IJB-C</i>	3,531	31,334	6
<i>LFW</i>	5,749	13,233	2
<i>RaFD</i>	67	8,040	120
<i>Multi-PIE</i>	337	+750,000	2225

Ne viename straipsnyje buvo naudotas *IJB-C* duomenų rinkinys ([6], [15], [19]). *RaFD* duomenų rinkinyje kiekvieno žmogaus nuotraukos yra išsaugotos 45, 90 ir 135 laipsnio posūkio kampu bei visos nuotraukos yra kelių formatų bei net yra nuotraukos su veido orientaciniais taškais. Panašus duomenų rinkinys yra *Multi-PIE*, tačiau su daugiau dalyvių ir daugiau nuotraukų vienam žmogui, nes duomenų rinkinys buvo kuriamas net su 15 vaizdo kamerų [16].

### 3.3. Balso keitimo metodai

#### 3.3.1. Balso keitimo mašininio mokymosi algoritmai

Įpratai balso keitimas sudarytas iš trijų žingsnių [24]:

1. Rerezentacinių akustinių savybių išgavimas;
2. Susiejimo (angl. *mapping*) taisyklių tarp originalaus ir tikslo kalbėtojų balsų sudarymas;
3. Tikslo balso sintezė.

Balso keitimo metodų palyginimas yra 3-3 lentelė. Joje matome tris pagrindinius modelius: *RNN*, *DNN* ir *CNN*. Naudojant *RNN* galima sukurti/atkurti garsus su trumpai mokytais modeliais ar neapmokytais mokymų parametrais [21]. Apmokant modelį remiantis *DNN* architektūra galima išgauti gana aukštą kokybę žmogaus balsui [23]. Tuo tarpu *CNN* yra tinkamas naudoti net ir emocionaliam žmogaus kalbėjimui [23].

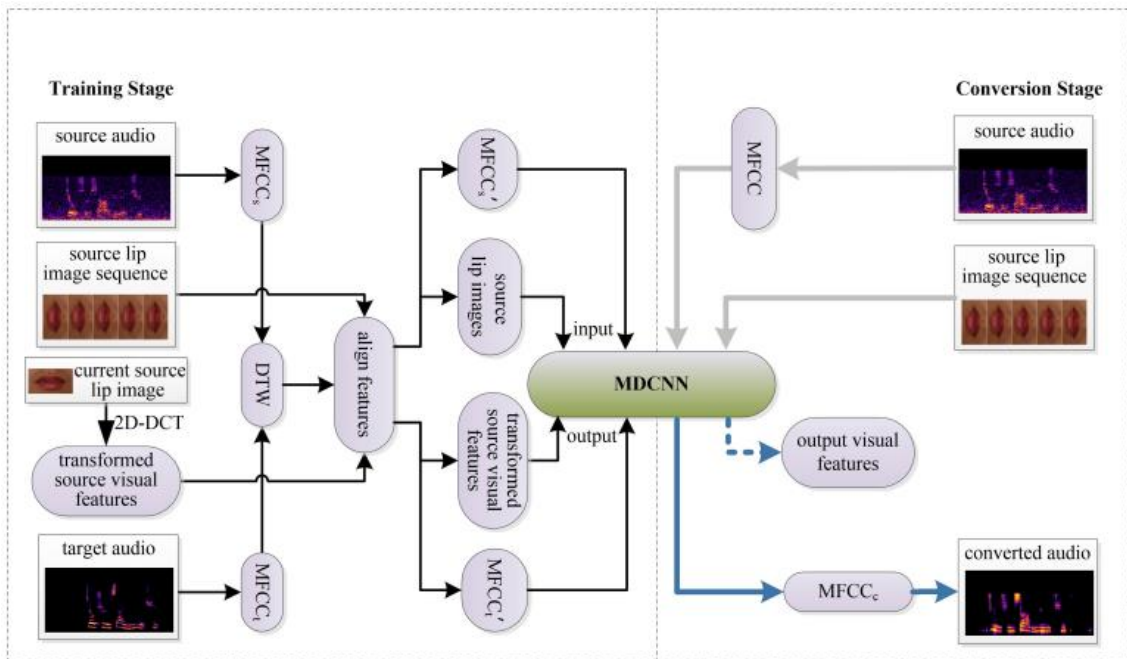
3-3 lentelė. Balso keitimo algoritmai

	Reikalingų įrašų kiekis	Naudojami modeliai	Komentaras
<i>MDCNN</i>	Po vieną	<i>CNN, DNN</i>	Keitimui ir treniravimui naudojamas ne tik garso įrašai, bet ir šaltinio lupų nuotraukos jam kalbant
<i>SVINGE2E</i>	Daug originalo įrašų	<i>CNN</i>	Efektyviai išgauna balso savybes
<i>CycleRNN</i>	Po vieną	<i>VAE, RNN</i>	Naudojamas <i>WaveNet</i> , kuris gerai atkuria balso savybes
<i>Seq2seq</i>	Po vieną	<i>VAE, CNN</i>	Susieja šaltinio bei tikslo savybes bei taip sukuria naują balsą
<i>SCENT</i>	Po vieną	<i>VAE, CNN</i>	Patobulintas <i>Seq2seq</i>
<i>WaveRNN</i>	Po vieną	<i>RNN</i>	Skirtas įrašams skirtingomis kalbomis.

#### 3.3.2. MDCNN

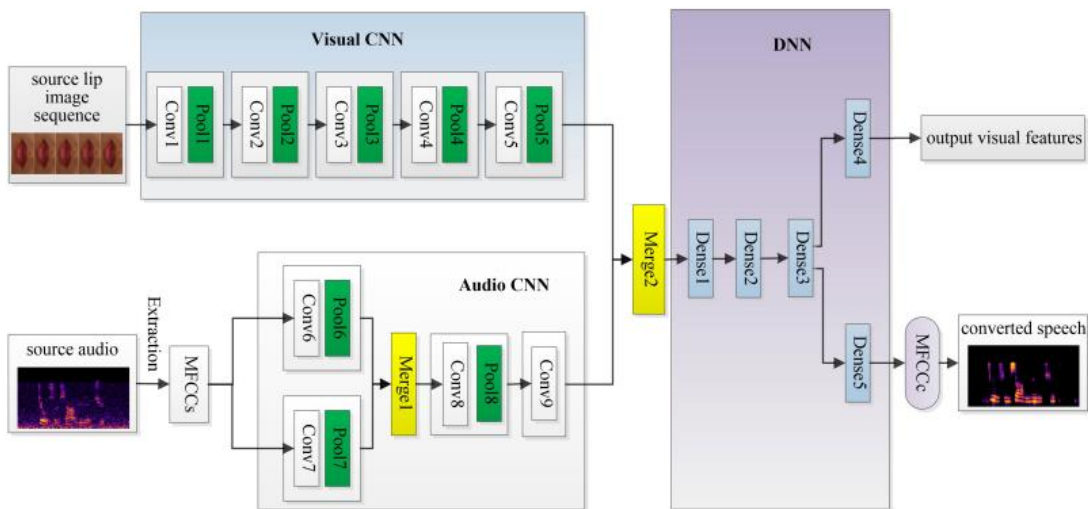
3-19 pav. pavaizduota *MDCNN* (angl. *mel-frequency cepstral coefficients*) veikimo schema. *MFCC* yra pasirinkti kaip garso charakteristikos to sumažinti modelio kompleksiskumą.  $MFCC_s$  ir  $MFCC_t$  yra išgaunami iš šaltinio bei tikslo garso įrašų. Kadangi skirtingų žmonių kalbėjimo greitis yra skirtingas, yra pritaikomas *DTW* (angl. *dynamic time wrapper*) sulyginant abiejų įrašų kalbos charakteristikas. Taip pat kadangi žmogaus kalbėjimas priklauso ir nuo lupų judesių, mokymui naudojamas ne tik garso įrašas, bet ir lupų judesių nuotraukos. Dviejų dimensijų *DCT* (angl. *discrete cosine transform*) naudojamas transformuoti lupų nuotraukas esamame kadre (angl. *frame*), kuris buvo ištiesintas į vienos dimensijos *DCT*.

Mokymui yra pritaikomas lupų nuotraukos,  $MFCC_s$  kaip įvestis bei  $MFCC_t$  ir transformuotos šaltinio vaizdinės charakteristikos kaip išvestis.



3-18 pav. MDCNN veikimo schema [27]

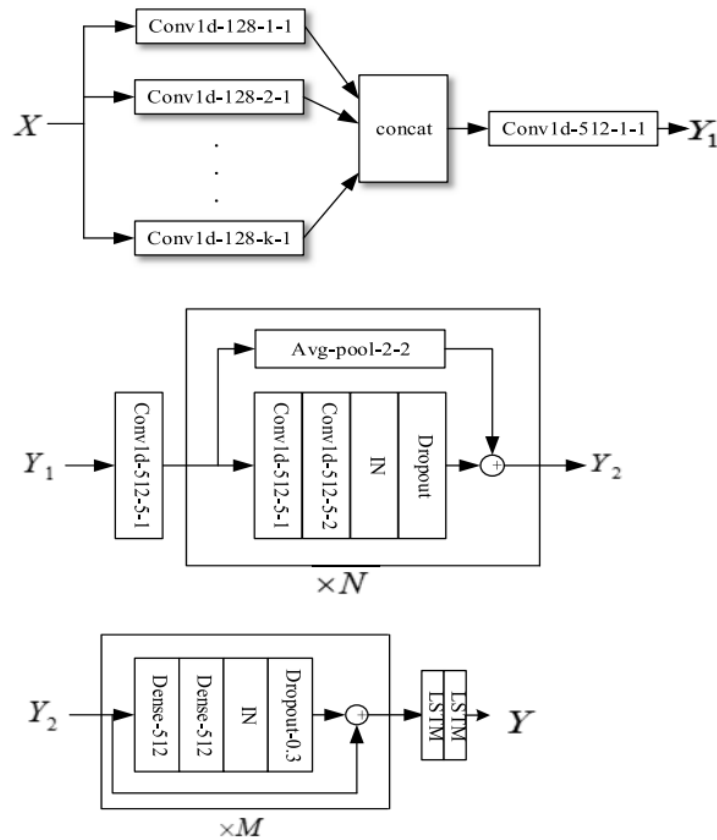
MDCNN gali charakterizuoti ir išmokti susieti garso-vaizdo charakteristikas iš šaltinio įrašo ir garso charakteristikas iš tikslo įrašo. Kaip matoma, MDCNN architektūra sudaryta iš dviejų CNN ir vieno DNN. CNN naudojami ištraukti garso savybes bei vaizdo ypatybes esančias tarp kadro iš šaltinio įrašo. Tuo tarpu DNN naudojamas susieti garso vaizdo iš šaltinio bei garso savybes iš tikslo įrašų.



3-19 pav. MDCNN architektūra

### 3.3.3. SVINGE2E

Optimizuotas SV modelis, kuris turi trijų sluoksnių LSTM (angl. *long short-term memory*) architektūrą kartu su FCNN (angl. *fully connected neural network*) ir normalizavimu (angl. *instance normalization*) viduriniame sluoksnyje. Šitame modelyje taip pat naudojama GE2E (angl. *Generalized End-to-End*) netekties funkcija [27]. SVINGE2E modelis gali efektyviau išgauti balso savybes. 3-20 pav. parodyta SVINGE2E architektūra, kurioje X tam tikra frazė garso įrašė, N – kalbėtojų kiekis bei M – pačių frazių kiekis [26]

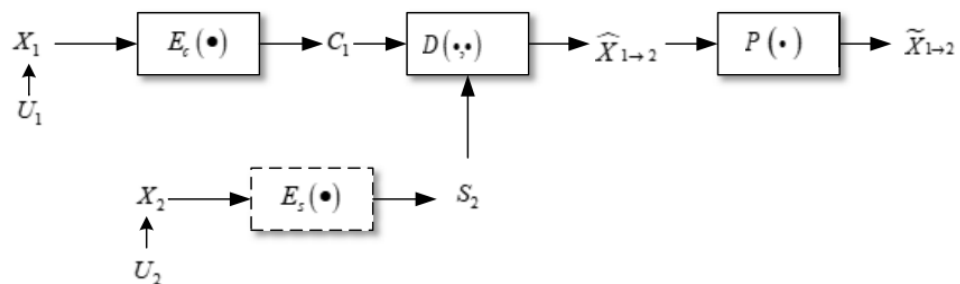


3-20 pav. SVINGE2E architektūra

### 3.3.4. RS-VC

RS (angl. *representation separation*) veido keitimas turi enkoderio-dekoduotojo struktūrą, pavaizduotą 3-21 pav., kuriame yra:

- 2 enkoderiai (originalo bei tikslo garso įrašų), kurie sukuria RS;
- dekoduotojas, kuris naudodamas enkoderių rezultatą sukuria naują garso reprezentaciją;
- po-tinklas (angl. *post-network*) papildo bei patobulina sukurta reprezentaciją.



3-21 pav. RD-VS architektūra [29]

Šiame modelyje naudojamos dvi netekties funkcijos:

- turinio netekties (angl. *content loss*) optimizuoja enkoderį  $E_c$ ;
- rekonstrukcijos praradimo (angl. *reconstruction loss*) padeda dekoderiui susintetinti naujos kalbos spektro dydį;

RS-VC modelio apmokymo metu yra naudojamas tas pats originalo ir tikslo garso įrašas kad modelis išmoktų rekonstruoti ir vėl sukonstruoti balsą, o testavimo metu jau testuojama kaip jam sekėsi tai išmokti.

Taip pat, norint pagerinti šio modelio mokymosi procesą siūloma naudoti progresinį mokymąsi. O pats mokymasis sudarytas iš dviejų etapų:

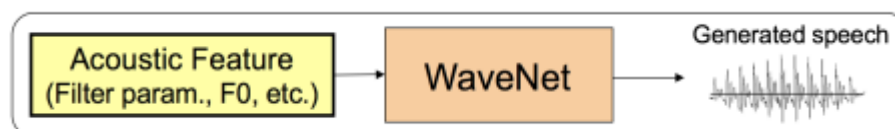
1. spektro rekonstrukcija – leidžia dekoderiui taisyklingai pabaigti Mel spektro rekonstrukciją naudojant originalo bei tikslo įrašų reprezentacijas
2. enkoderio optimizacija – optimizuoja tikslo enkoderį pagal rekonstrukcijos spektrą. Naudojantis tikslo įrašo filtrais bei tikslo turinio netekties funkcija tikslo enkoderis pašalina kalbėtojų informaciją bei pasiima turinio informaciją.

### 3.3.5. WaveNet

*WaveNet* yra autoregresinis generatyvusis tinkas ir yra naudojamas tiesiogiai modeliuoti neapdorotas garsų bangas. Taip pat šį metodą galima naudoti su limituotu mokymo duomenų rinkiniu dviem etapais [29]:

1. Treniruojant *WaveNet* pirminį modelį su kelių kalbėtojų duomenų rinkiniu (akustinių savybių rinkinys naudojamas kaip įvestis)
2. Vėliau treniruojant modelį pakeičiant visus pirminius modelio parametrus naudojant tik mokymo duomenų rinkinio atrinkto kalbėtojų duomenis.

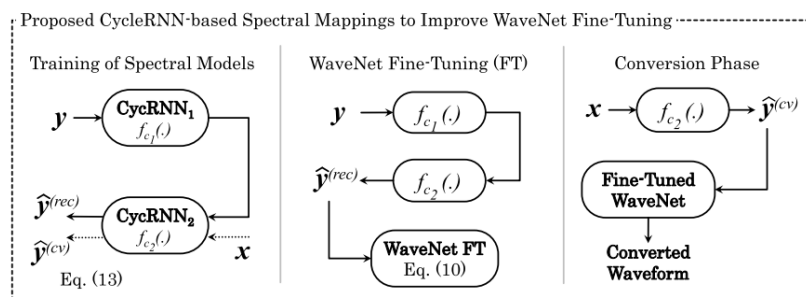
3-22 pav. pavaizduota *WaveNet* veikimo schema. Kadangi akustinės savybės išgaunamos atsižvelgiant į kalbos generavimo procesą, tikimasi, kad atitikimas tarp kalbos bangos formos ir akustinių savybių tinkle bus sukurtas automatiškai, atsižvelgiant į fizinius kalbos generavimo proceso apribojimus. Be to, siūlomas metodas neapima artikuliacinio filtro valdymo pagal sužadinimo signalus, taip pat nereikia jokių matematinių prielaidų, susijusių su duomenimis, pvz., Gauso koeficientu. Todėl galima susintetinti aukštos kokybės kalbą, kuri atkuria išsamią įvairių esamų vokoderių prarastą laiko informaciją [30].



3-22 pav. *WaveNet* schema [33]

### 3.3.6. CycleRNN kartu su WaveNet

3-23 pav. pavaizduotas *CycleRNN* modelis, kuris yra optimizuotas naudojant savaime numatytą tikslinio spektro ciklinę netekties funkciją (angl. *cyclic loss of the self-predicted target spectra*) ir šaltinio-tikslo konversijos praradimo funkciją (angl. *conversion loss of the source-to-target conversion*). *CycleRNN* modelio parametrai yra optimizuoti pagal netekties funkcijas  $\hat{y}^{(conv)}$  ir  $\hat{y}^{(pred)}$ . Tačiau skirtingai nei įprasto *CycleRNN* modelio, schemoje dar naudojama *WaveNet* tikslinimo procedūra, kurioje vyksta savaime numatytas tikslinio spektro savybių optimizavimas. [28]



3-23 pav. CycleRNN modelio architektūra su WaveNet procedūra [28]

### 3.3.7. Bendrinis Seq2Seq

Seq2Seq [25] modelis išmoksta susieti šaltinio bei tikslo savybes. 3-24 pav. pavaizduotas seq2seq modelio architektūra, kuri kaip ir dauguma seq2seq, turi koduotojo-dekoduotojo struktūrą. Koduotojas susieja įvesties garso savybes  $x$  į seką  $H$ :

$$H = \mathbf{h}_{1:n} = \text{Enc}(\mathbf{x}_{1:n}).$$

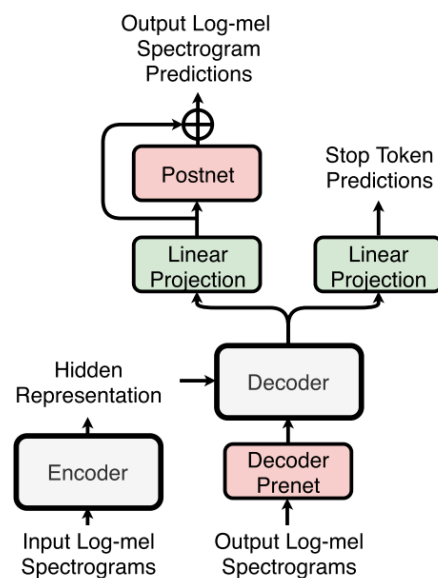
Dekoduotojas yra autoregresinis, t. y. gauti dekodutojo rezultatui  $y_t$ , naudojami praeiti jo rezultatai  $y_{1:t-1}$ :

$$y_t = \text{Dec}(\mathbf{h}_{1:n}, \mathbf{y}_{1:t-1}).$$

Taip pat seq2seq VC modeliui naudojami papildomi elementai, paimti iš TTS modelio:

- *Prenet* – sudarytas iš 2 pilnai sujungtų sluoksnių, kuris naudojamas kaip informacijos susiaurinimas (angl. *bottleneck*) norint naudotis autoregresiniu dekodutoju;
- Linijinė projekcija (angl. *linear projection*) – naudojamas dekoderiui projektuoti išvestis į norimą matmenį;
- *NN postnet* – kompensuoti trūkstamai informacijai.



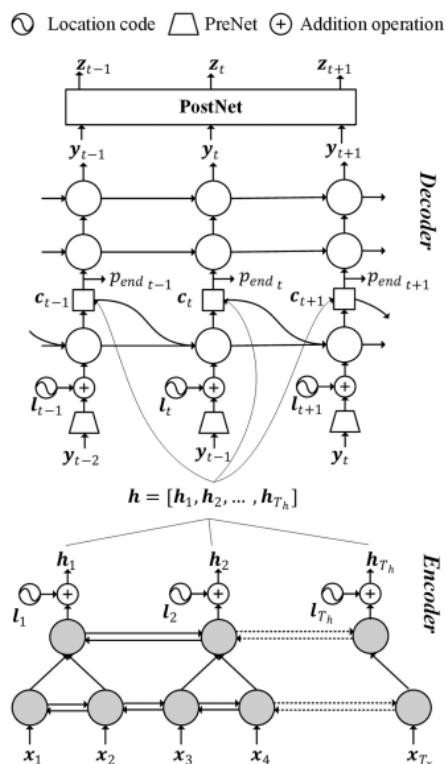


3-24 pav. Seq2Seq veikimo schema [25]

### 3.3.8. SCENT

*SCENT* (angl. *sequence-to-sequence ConvErsion NeTwork*) – modelis, paremtas Seq2Seq veikimu bei skirtas akustiniam balsu konvertavimui. Pasinaudodamas tempimo mechanizmu, *SCENT* modelis nesiremia išankstiniu *DTW* derinimo apdorojimu bei trukmės konvertavimas gali būti pasiektas vienu metu. Taip pat konvertuotos akustinės savybės perduodamos per *WaveNet* vokoderį, kad būtų galima atkurti kalbos bangų formas.

Detalesnė *SCENT* galima veikimo schema pavaizduota 3-25 pav. Schemoje pilki apskritimai koduotoje žymi *LSTM* vienetus su sluoksnio normalizavimu.  $T_x$  ir  $T_h$  yra įvesties sekos ir paslėptų pateikimų kadro skaičiai. Šiame paveiksle pateiktas enkoderis turi sumažinimo dažnį  $M = 2$ , todėl turime  $T_x = 2T_h$ . Natūralios istorijos kontekstai yra generuojamos konversijos metu bei naudojami kaip dekoduojačio įvestys yra treniruočių metu [31].



3-25 pav. SCENT architektūra [31]

Objektyvūs ir subjektyvūs eksperimentų rezultatai parodė šio metodo pranašumą, palyginti su kitais metodais (pvz. *VCC2018*, *DNN*, *JD-GMM* variacijomis), ypač atsižvelgiant į trukmę. Abliacijos (angl. *Ablation*) testai dar labiau įrodė įvesties naudą Mel-spektrogramos.

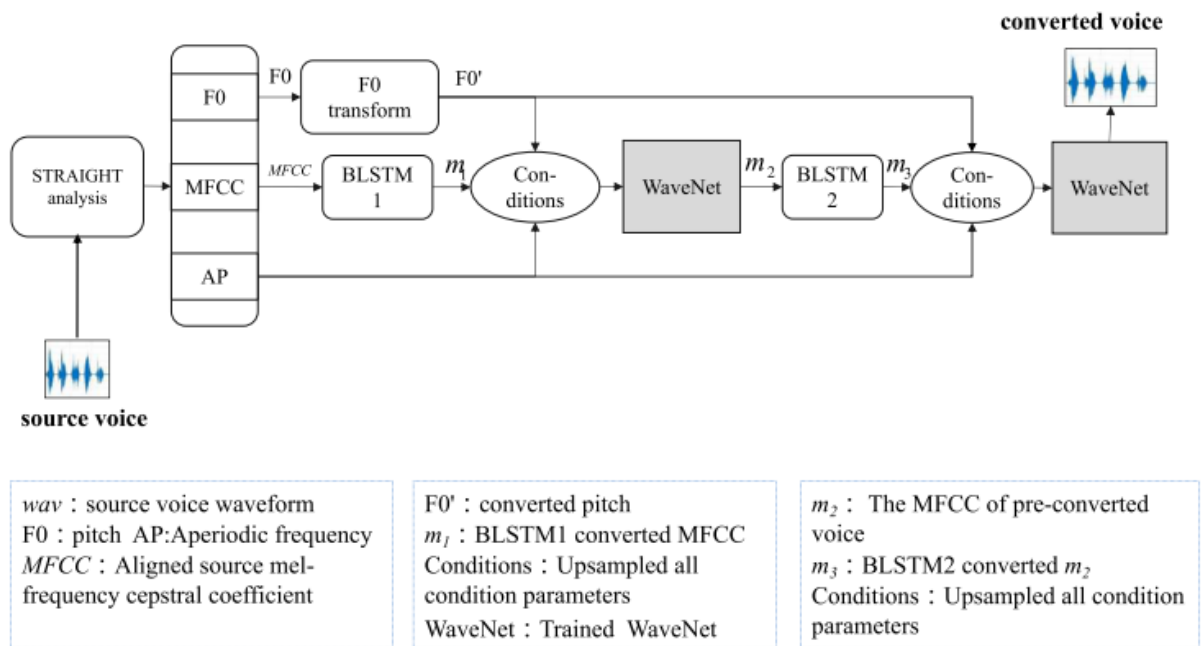
### 3.3.9. BLSTM

3-26 pav. pavaizduota *BLSTM* su *WaveNet* architektūros schema, kurioje balso keitimo procesas sudarytas iš trijų dalių:

1. *WaveNet* bangos formų generavimas – išgaunamas trijų tipų kalbos ypatybės: AP, F0 ir mel dažnio cepstrinius koeficientus (*MFCC*). Tada kiekvieno tipo funkcijos atitinkamai apdorojamos ir konvertuojamos. F0 konvertuojamas logaritmo tiesine transformacija. AP apdorojamas kopijuojant.
2. *BLSTM1* – konvertuoja *MFCC* bei konvertuoti parametrai naudojami kaip *WaveNet*, kad būtų sukurtas iš anksto konvertuotas balsas.
3. *BLSTM2* – ištraukia ir apdoroja iš anksto konvertuotų balsų *MFCC*. Po papildomo apdorojimo naudojami kaip *WaveNet* sąlygos, kad būtų sukurtas galutinis konvertuotas balsas.

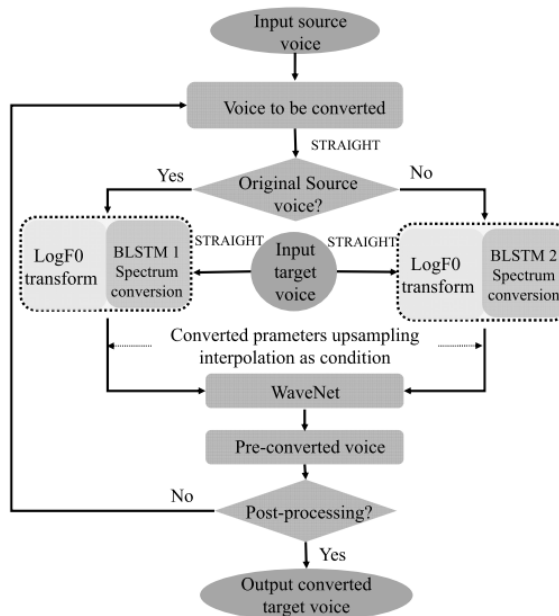
Bangos formos pavyzdžių skaičius laiko srityje paprastai yra didesnis nei kadrų skaičius. Norint suderinti rėmo lygio sąlyginius kintamuosius (*MFCC*, *F0*, *AP*) su laiko srities imties taškais, čia naudojama tiesinė interpoliacija. Tuo pačiu metu minimalios vidutinės kvadratinės paklaidos (*MSE*) kriterijus naudojamas kaip nuostolių funkcija *BLSTM*, o kryžminė entropija naudojama kaip praradimo funkcija *WaveNet* mokymuose. Atsižvelgiant į akustinių savybių konversijos įtaką ir bangos formos sugriuvusią kalbą, kurią sukelia atsitiktinės klaidos bangos formos generavimo

processe, visame tinklo konvertavimo procese naudojami po apdorojimo mokymai ir dviejų etapų balso generavimo procesas. Tokiu būdu tikimasi geresnės konvertuoto balso kokybės.



3-26 pav. BLSTM architektūra [32]

3-27 pav. pavaizduota BLSTM balso konvertavimo schema su sąlyginiu WaveNet mokymo sistema, kur STRAIGHT reiškia savybių analizės ir išgavimo žingsnį, WaveNet yra iš anksto paruošta WaveNet modelis, o punktyrinuose langeliuose yra dvi atskirtos operacijos: logaritminė F0 transformacija ir spektro konvertavimas naudojant BLSTM



3-27 pav BLSTM schema [31]

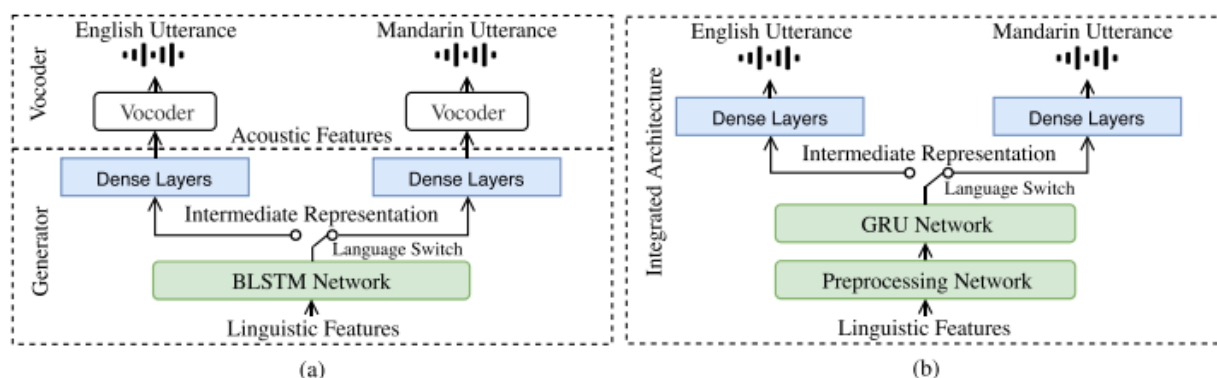
Palyginus su keliais skirtingais balso konvertavimo metodais mandarinų ir anglų kalbos duomenų rinkiniuose, nors šio metodo tobulinimas objektyviose priemonėse yra nereikšmingas, jis žymiai

pagerina subjektyvios klausos poveikį. Taip pat konvertuotas balsas skambėjo natūraliau ir sklandžiau, ypač konvertuojant skirtingų lyčių balsus, kas ne visada išgaunama iš kitų modelių.

### 3.3.10. WaveRNN

*WaveRNN* yra aukštos kokybės kalbos signalų formų generatorius, kuris gali keisti balsus tarp skirtingų kalbų. Jį paprastai sudaro dvi dalys: paslėpta sekcija, kuri susieja įvesties ypatybes su tarpiniais atvaizdais, ir išvesties dalis, susiejanti tarpinius vaizdus su tiksline signalo forma. Paslėpta skyrių sudaro išankstinio apdorojimo tinklas ir *GRU* tinklas. Išankstinio apdorojimo tinklas atrenka įvesties funkcijas į norimą skiriamąją gebą. *GRU* tinklas sudarytas iš vieno tankaus sluoksnio, dviejų ribojamų pasikartojančių vienetų (*GRU*) sluoksnių su praleidžiamomis jungtimis, po kurių seka vienas tankus sluoksnis. Išvesties sekcija susideda iš dviejų tankių sluoksnių. Kryžminis entropijos praradimas tarp numatomos bangos formos ir natūralios tikslinės kalbos priimamas kaip objektyvi funkcija.

3-28 pav. a) dalyje pavaizduotas generatoriaus-vokoderio *VC* schema, kurio generatorius yra daugiafunkcinis modulinis NN. Generatorius naudoja kalbines savybes kaip įvestį ir generuoja akustines savybes kaip išvestį. Vokoderis sintezuoja kalbos bangos formą iš akustinių savybių. b) dalyje pavaizduotas *WaveRNN* veikimo schema, kurioje matoma, generuoja bangos formą tiesiogiai be atskiro vokoderio. Abiejose sistemose naudojama paslėpta sekcija, kurią dalijasi dvi kalbos, ir išvesties sekcija, kurią sudaro dvi atšakos. Kiekviena atšaka mokoma tik kalbančių gimtoji kalba, tačiau naudojama daryti išvadas apie užsienio kalba kalbančių žmonių kalbą [33].



3-28 pav. Generatoriaus-vokoderio (a) ir *WaveRNN* (b) veikimo schemas [33]

### 3.3.11. Duomenų rinkiniai

Žemiau esančioje lentelėje pavaizduoti galimi balso keitimo duomenų rinkiniai. *M-AILABS* duomenų rinkinyje yra balso įrašai įvairiomis kalbomis bei pateiktus tekstus, kas juose sakoma. *VoxCeleb* yra labai didelis duomenų rinkinys, kuriame yra moterų 45% įrašų, tai visai gerai pasiskirstęs duomenų rinkinys [33]. *VCC 2018* duomenų rinkinys nors ir neturi daug įrašų, tačiau įrašai ilgi ir įrašyti studijoje, todėl nėra pašalinių garsų [35].

3-4 lentelė. Balso įrašų duomenų rinkiniai

Duomenų rinkinys	Įrašų kiekis	Žmonių kiekis	Kelios kalbos	Ilgis
<i>M-AILABS</i>	-	-	Taip	iki 20 sekundžių

<i>VCTK</i>	-	110	Įvairūs anglų k. akcentai	apie 400 sakinių
<i>VoxCeleb</i>	153,516	1,251	Ne	-
<i>AVWD</i>	300	3	Ne	2-3 sekundės
<i>VCC 2018</i>	-	20	Ne	13 minutės
<i>CMU Arctic</i>	2834	1132	Ne	-

### **3.4. Atrinkti veido keitimo mašininio mokymosi įrankiai, su kuriais planuojama atlikti eksperimentus**

Eksperimentams atlikti pasirinkti trys architektūros norint išbandyti skirtingus metodus. Pasirinktos konvoliucinių neuroninių tinklų bei auto-koduotojo architektūros kartu su generatyviuoju priešininkų tinklu. Taip pat pasirinktas *FSNet* algoritmas, kuris nors ir nesugeneruoja aukštos kokybės nuotraukų, tačiau nenaudoja sunkių bibliotekų bei teoriškai turėtų veikti greičiau.

Duomenų rinkinys visiems trimis algoritmams bus naudojamas toks pats norint išsiaiškinti geriausią veikimą. Taip pat vienodai bus sukurtos kaukės bei veido kontūrų taškai.

### **3.5. Atrinkti balso keitimo mašininio mokymosi įrankiai, su kuriais planuojama atlikti eksperimentus**

Balso keitimo eksperimentų metu bus išbandomos įvairios architektūros, pavyzdžiui plačiai aprašytos *VAE* – konvoliucinis, pasikartojantis bei gilusis neuroniniai tinklai. Kadangi visi metodai reikalauja tik vieno įrašo, apmokymui bei keitimui bus naudojami tie patys įrašai.

## 4. Projektas

### 4.1. Projekto planas

3 semestrų detalus projekto planas pavaizduotas žemiau esančioje lentelėje.

**4-1 lentelė.** Detalus projekto planas

Nr.	Pavadinimas	Trukmė	Pradžia	Pabaiga
1	Magistrinis projektas	366 d.	10/01/21	2/23/23
2	1 semestras	70 d.	10/01/21	1/05/22
3	Projekto IS	28 d.	10/01/21	11/09/21
4	Projekto IS realizavimas	14 d.	10/01/21	10/20/21
5	Projekto IS pristatymas	1 d.	10/20/21	10/21/21
6	Projekto IS vertinimas	6 d.	10/20/21	10/27/21
7	Projektavimo metodologijos ir technologijų analizė	38 d.	11/15/21	1/05/22
8	Projektavimo metodologijos ir technologijų analizės rašymas	23 d.	10/18/21	11/17/21
9	Projektavimo metodologijos ir technologijų analizės pristatymas	1 d.	11/17/21	11/17/21
10	Projektavimo metodologijos ir technologijų analizės vertinimas	14 d.	11/17/21	12/06/21
11	Projekto paraiška	9 d.	12/01/21	12/13/21
12	Žvilgsnio stebėjimo sistemos projekto paraiškos tarpinis vertinimas ir pataisymai	4 d.	12/01/21	12/06/21
13	Žvilgsnio stebėjimo sistemos projekto paraiškos pristatymas	0 d.	12/06/21	12/06/21
14	Žvilgsnio stebėjimo sistemos projekto paraiškos vertinimas	6 d.	12/06/21	12/13/21
15	Projekto planas	32 d.	12/02/21	1/14/22
16	Projekto plano realizavimas	15 d.	12/02/21	12/22/21
17	1 Semestro planavimas	15 d.	12/02/21	12/22/21
18	2 Semestro planavimas	11 d.	12/05/21	12/20/21
19	3 Semestro planavimas	12 d.	12/03/21	12/20/21
20	Projekto plano pristatymas	3 d.	12/21/21	12/27/21
21	Projekto plano vertinimas	14 d.	12/28/21	1/14/22
22	2 semestras	86 d.	1/26/22	5/25/22
23	Reikalavimų specifikacija	41 d.	1/26/22	3/23/22
24	Funkcinių reikalavimų surinkimas	3 d.	2/01/22	2/03/22
25	Apribojimų surinkimas	3 d.	2/04/22	2/18/22
26	Nefunkcinių reikalavimų surinkimas	3 d.	2/16/22	2/21/22
27	Vartotojo sąsajos specifikavimas	3 d.	2/16/22	2/18/22
28	Sistemos IS administracinės dalies vartotojo sąsajos specifikavimas	2 d.	2/16/22	2/18/22

29	Sistemos kodo redaktoriaus papildinio vartotojo sąsajos specifikavimas	2 d.	2/16/22	2/18/22
30	Duomenų modelio sudarymas	4 d.	2/16/22	2/21/22
31	Esybių ryšių modelio sudarymas	1 d.	2/16/22	2/16/22
32	Duomenų žodyno sudarymas	2 d.	2/18/22	2/21/22
33	Pataisyti ataskaitą bei pabaigti nepabaigtas vietas	7 d.	3/01/22	3/09/22
34	Reikalavimų specifikacijos pristatymas	0 d.	3/10/22	3/10/22
35	Reikalavimų specifikacijos vertinimas	7 d.	3/10/22	3/18/22
36	Sistemos projektavimas	27 d.	3/18/22	4/25/22
37	Panaudos atvejų modelio kūrimas	6 d.	3/18/22	3/25/22
38	Užduoties analizė, ypatingų situacijų analizė	3 d.	3/18/22	3/23/22
39	PA ir dalyvių sąrašai	3 d.	3/18/22	3/23/22
40	PA specifikacijos	3 d.	3/18/22	3/23/22
41	PAM diagrama	3 d.	3/23/22	3/25/22
42	Analitinio modelio kūrimas	3 d.	3/28/22	3/30/22
43	Detalizuotų trasos diagramų braižymas	3 d.	3/28/22	3/30/22
44	Klasių diagramų braižymas	3 d.	3/28/22	3/30/22
45	Sąveikos diagramų braižymas	3 d.	3/28/22	3/30/22
46	Projektinio modelio kūrimas	11 d.	3/31/22	4/14/22
47	Išplėstų trasos diagramų braižymas	3 d.	3/31/22	4/04/22
48	Išplėstų saveikos diagramų braižymas	3 d.	4/05/22	4/07/22
49	Veiksmų eiliškumo diagramų braižymas	3 d.	4/08/22	4/12/22
50	Posistemų diagramos braižymas	2 d.	4/13/22	4/14/22
51	Sistemos architektūros projekto pristatymas	1 d.	4/15/22	4/15/22
52	Sistemos architektūros projekto vertinimas	7 d.	4/15/22	4/25/22
53	Darbų viešo pristatymo medžiagos paruošimas	12 d.	4/18/22	5/03/22
54	Viešo pristatymo medžiagos paruošimas	10 d.	4/18/22	4/29/22
55	Viešojo pristatymo medžiaga	1 d.	5/02/22	5/02/22
56	Viešasis pristatymas	1 d.	5/03/22	5/03/22
57	Programų sistemos prototipas	16 d.	5/04/22	5/25/22
58	Prototipo realizavimas	9 d.	5/04/22	5/16/22
59	Prototipo pristatymas	1 d.?	5/17/22	5/17/22
60	Prototipo vertinimas	7 d.	5/17/22	5/25/22
61	2 semestro pabaiga (Reikalavimų specifikacija, programinės įrangos architektūra, pirmas prototipas)	0 d.	5/25/22	5/25/22
62	3 semestras	104 d.	9/01/22	1/24/23
63	Projekto programų sistemos testavimo planavimas	21 d.	9/01/22	9/29/22
64	Testavimo aplinkos aprašymas	7 d.	9/01/22	9/21/22
65	Testavimo atvejų aprašymas	7 d.	9/21/22	9/29/22
66	Projekto programų sistemos testavimo plano pristatymas	0 d.	9/29/22	9/29/22
67	Projekto programų sistemos testavimo plano vertinimas	7 d.	10/03/22	10/11/22

68	Programų sistemos antrasis prototipas	23 d.	10/03/22	11/02/22
69	Sistemos IS prototipo realizavimas	13 d.	10/17/22	11/02/22
70	Programų sistemos antrojo prototipo pristatymas	0 d.	11/02/22	11/02/22
71	Programų sistemos antrojo prototipo vertinimas	7 d.	10/03/22	10/11/22
72	Testavimo realizavimas	8 d.	11/04/22	11/15/22
73	Komponentinis testavimas	3 d.	11/04/22	11/08/22
74	Integracinis testavimas	3 d.	11/09/22	11/11/22
75	Rankinis žvilgsnio stebėjimo IS testavimas	2 d.	11/14/22	11/15/22
76	Projekto kokybės tyrimas	23 d.	11/16/22	12/16/22
77	Vartotojo vadovo rašymas	17 d.	11/16/22	12/08/22
78	Projekto kokybės tyrimas pristatymas	1 d.	12/09/22	12/09/22
79	Projekto kokybės tyrimo vertinimas	6 d.	12/09/22	12/16/22
80	Programų sistemos įdiegimas	30 d.	11/17/22	12/28/22
81	Programų sistemos įdiegimas	23 d.	11/17/22	12/19/22
82	Programų sistemos įdiegimo pristatymas	1 d.	12/20/22	12/20/22
83	Programų sistemos įdiegimo vertinimas	6 d.	12/21/22	12/28/22
84	Dokumentacijos parengimas	18 d.	12/21/22	1/13/23
85	Projekto techninė dokumentacijos pasirengimas	6 d.	1/02/23	1/09/23
86	Naudotojo vadovo parengimas	5 d.	1/09/23	1/13/23
87	Projekto dokumentacijos pristatymas	0 d.	1/13/23	1/13/23
88	Projekto dokumentacijos vertinimas	6 d.	12/21/22	12/28/22
89	Pasiruošimas programų sistemos gynimui	6 d.	1/16/23	1/23/23
90	Programų sistemos gynimas	1 d.	1/24/23	1/24/23
91	3 semestro pabaiga (Programų sistema, testavimo planas, antras prototipas, pilna dokumentacija ir kt.)	0 d.	1/23/23	1/23/23

## 4.2. Reikalavimų specifikacija

### 4.2.1. Suinteresuoti asmenys

#### Užsakovas(-ai)

Užsakovas yra Indrė Dimšė, kuri yra suinteresuota sukurti šią sistemą, nes sukurta sistema bus pateikta plačiajai visuomenei ir turės nešti pelną.

Atsakomybės:

- stebėti kuriamo produkto eigą bei teikti pastabas ir sukūrus sistemą ją priimti jei ji tenkina reikalavimus;
- nustatyti kuriamos sistemos reikalavimus bei funkcionalumus;
- sukūrus demonstracinę versiją, pateikti ją plačiajai visuomenei bei su programuotojais aptarti gautas pastabas.

#### Potencialūs produkto naudotojai

Potencialūs produkto naudotojai aprašyti 2.1.2 Naudotojų rolės ir tikslai skyrelyje.

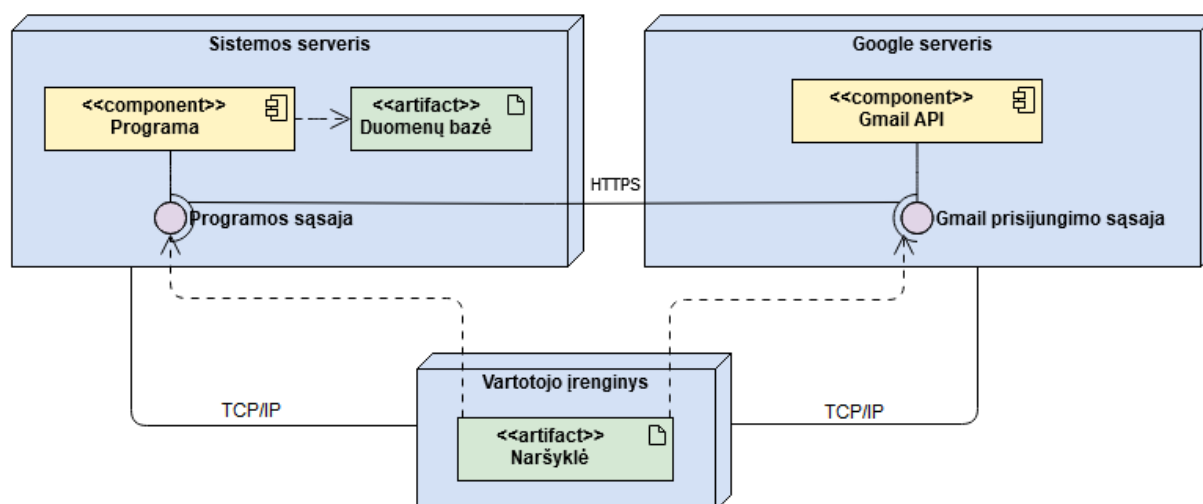


## 4.2.2. Apribojimai

Kadangi sistema kuriama ne tik pelno tikslais, bet ir norint atlikti tyrimą, nėra daug keliamų reikalavimų apart terminų, kada tyrimą reikia pristatyti. Tolesniuose skyriuose bus analizuojami įvairūs sistemos aspektai bei apibrėžti reikalavimai.

### 4.2.2.1. Diegimo aplinka

4-1 pav. pavaizduota diegimo diagrama, kurioje parodyta, kaip vartotojo įrenginys turės bendrauti su sistemos serveriu. Taip pat, jei vartotojas norės, galės kreiptis ir į *Gmail* paskyrą prisijungimui prie kuriamos sistemos.



4-1 pav. Diegimo aplinkos diagrama

### 4.2.2.2. Komunikuojančios sistemos

Naudojant socialinę paskyrą *Gmail* bus galima prisijungti prie sukurtos sistemos. Tokiu būdu bus pagerinta vartotojo patirtis naudojantis sistema, nes užteks paspausti "Prisijungti su *Gmail*" mygtuką vietoj pilnos registracijos.

Bus naudojamas *Gmail* autentifikacijos API:

- vartotojui paspaudus mygtuką, jis yra nusiunčiamas į *Gmail* prisijungimo langą. Unikalus kodas yra internetiniame adrese, su kuriuo yra nukreipiamas vartotojas prisijungti prie *Gmail*;
- vartotojui prisijungus prie savo paskyros, į sistemos serverį ateina vartotojo duomenys, kuriuos naudojant sukuriama vartotojo profilis.

### 4.2.2.3. Numatoma darbo vietos aplinka

Kadangi sistema naudosis plačiosios visuomenės nariai, sistemos naudojimosi aplinka gali būti įvairi. Vienintelis bendras visų vartotojų bruožas – jie turės prieigą prie interneto.

### 4.2.2.4. Sistemos kūrimo terminas(-ai)

Sistemos demonstracinę versiją reikia pateikti iki 2022-05-01, o gautinį produktą iki 2023-01-15.

### 4.2.3. Svarbūs faktai ir prielaidos

#### 4.2.3.1. Faktai

Sistema kuriama kaip programų sistemų inžinerijos baigiamasis magistrinis projektas

#### 4.2.3.2. Prielaidos

- Gali būti išleisti įstatymai, ribojantys/nusakantys žmogaus veido ir/ar balso keitimo kitu asmeniu galimybes.
- Sistemos kūrimo metu gali patobulėti technologijos, todėl gali tekti sistemą pritaikyti prie jų.
- Kitos panašių funkcijų sistemos gali būti kuriamos tuo pačiu metu.
- Planuojama naudoti *Python* programavimo kalbą.
- Įkeliant modelius gali būti naudojami skirtingi formatai.

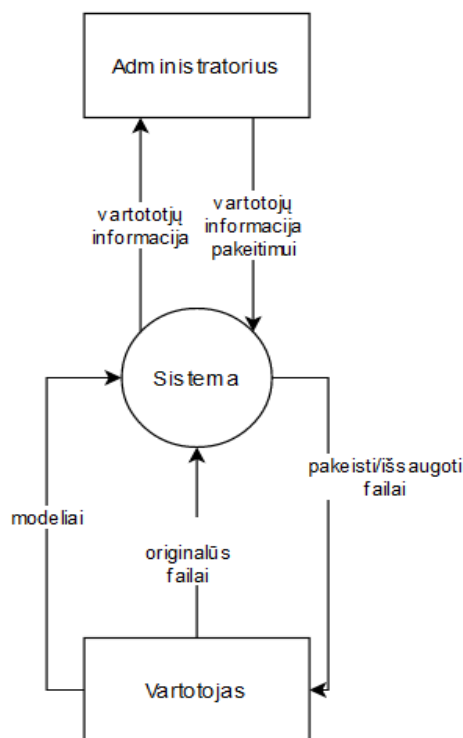
### 4.2.4. Veiklos sudėtis

#### 4.2.4.1. Esama padėtis

Šio projekto metu yra kuriama nauja sistema, todėl šiuo metu veiklos modelis dar neegzistuoja.

#### 4.2.4.2. Veiklos kontekstas

Žemiau pavaizduota projekto konteksto diagrama.



4-2 pav. Veiklos konteksto diagrama

#### 4.2.4.3. Veiklos suskaidymas

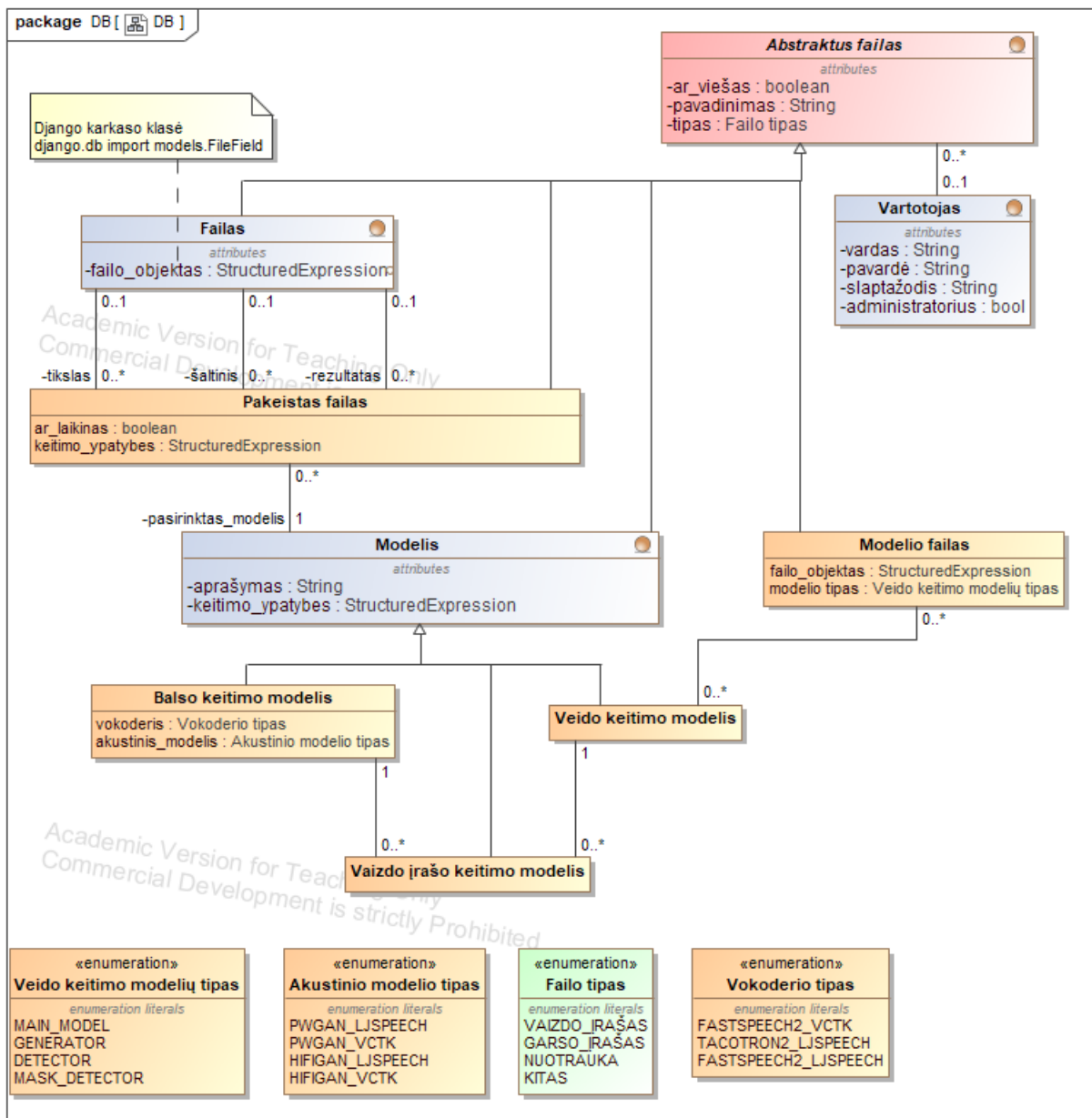
4-2 lentelė. Veiklos suskaidymas

Eil. Nr.	Įvykio pavadinimas	Įeinantys/Išeinantys informacijos srautai
1	modeliai	Vartotojas įkeliamas modelis nuotraukos, garso ar vaizdi įrašų keitimui
2	originalūs failai	Vartotojas įkelia originalius failus, kurie naudojami keitimui
3	pakeisti/išsaugoti failai	Vartotojas gauna tik pakeistus failus ar anksčiau keistus ir išsaugotus
4	vartotojų informacija	Administratorius pasiima vartotojų duomenis
5	pakeista vartotojų informacija	Administratorius pakeičia pasiimtus duomenis ir grąžina į sistemą juos išsaugoti

#### 4.2.5. Duomenų modelis ir jo elementų žodynas

Žemiau esančiame duomenų modelyje pavaizduotos visos sistemoje egzistuojančios klasės:

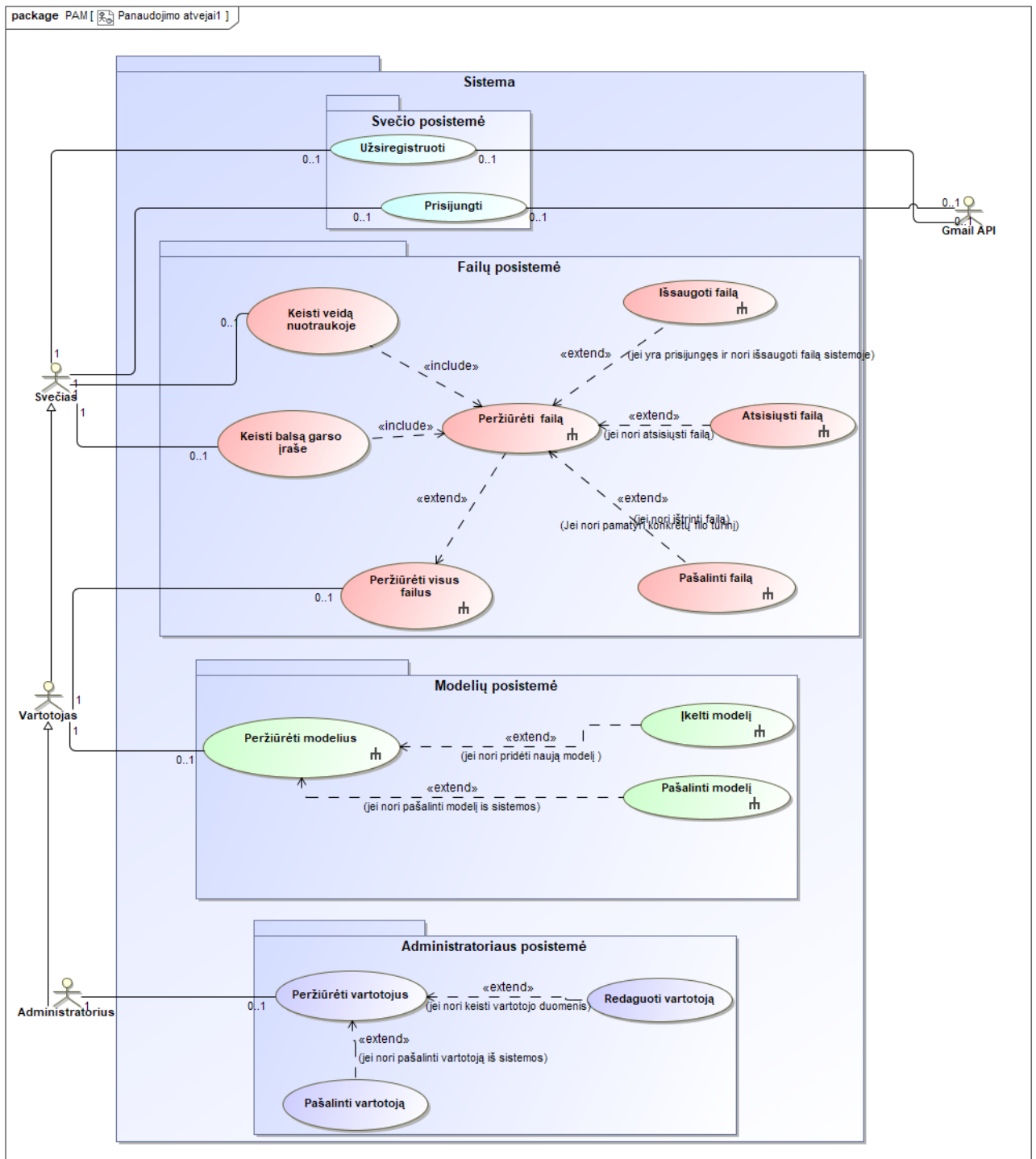
- Vartotojas – registruotas sistemos naudotojas;
- Abstraktus failas – abstrakti klasė, kurios funkcijas naudoja visos likusios klasės;
- Failas – medijos failas (vaizdinio ir/ar garsinio formato);
- Pakeistas failas – objektas, kuris saugo informaciją apie failo keitimo procesą;
- Modelis – klasė, kuri saugo objektus ir informaciją, reikalingą keitimo procesui;
- Modelio failas – veido keitimui naudojamas mašininio mokymosi failas;
- Veido keitimo modelis – klasė, kuri saugo objektus ir informaciją, reikalingą veido keitimo procesui;
- Balso keitimo modelis – klasė, kuri saugo objektus ir informaciją, reikalingą balso keitimo procesui;
- Vaizdo įrašo keitimo modelis – klasė, kuri saugo objektus ir informaciją, reikalingą vaizdo įrašo keitimo procesui.



4-3 pav. Duomenų modelis

#### 4.2.6. Sistemos sudėtis ir panaudojimo atvejai

Žemiau pavaizduota panaudojimo atvejų diagrama, kurioje pavaizduoti visi sistemoje esantys panaudojimo atvejai.



4-4 pav. PA diagrama

4-3 lentelė ir 4-4 lentelė aprašyti prisijungimo bei registracijos panaudojimo atvejai.

4-3 lentelė. PA „Prisijungti“ specifikacija

Panaudojimo atvejis	Nr. 1	Užsiregistruoti
Tikslas		Svečias gali užsiregistruoti sistemoje, kad galėtų valdyti modelius bei išsaugoti failus sistemoje.
Dalyviai		Svečias
Priklausomybės		-

<b>Nefunkciniai reikalavimai</b>	-
<b>Prieš-sąlygos</b>	Ateiti į registracijos langą
<b>Sužadinimo sąlyga</b>	Suvedami duomenys ir paspaudžiamas registracijos mygtukas
<b>Po-sąlyga</b>	Vartotojas gali matyti modelius bei išsaugoti failus sistemoje
<b>Pagrindinis scenarijus</b>	Svečiui suvedus duomenis ir paspaudus registracijos mygtuką, jo duomenys išsaugomi duomenų bazėje bei parodomas pagrindinis vartotojo langas
<b>Alternatyvūs scenarijai</b>	Jei blogi duomenys ar toks vartotojas jau užsiregistravęs, parodomas klaidos pranešimas

4-4 lentelė. PA „Prisijungti“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 2</b>	<b>Prisijungti</b>
<b>Tikslas</b>		Vartotojas gali prisijungti prie sistemos, kad galėtų valdyti modelius bei išsaugoti failus sistemoje.
<b>Dalyviai</b>		Svečias
<b>Priklausomybės</b>		-
<b>Nefunkciniai reikalavimai</b>		-
<b>Prieš-sąlygos</b>		Ateiti į prisijungimo langą
<b>Sužadinimo sąlyga</b>		Suvedami duomenys ir paspaudžiamas prisijungimo mygtukas
<b>Po-sąlyga</b>		Vartotojas gali matyti modelius bei išsaugoti failus sistemoje
<b>Pagrindinis scenarijus</b>		Svečiui suvedus duomenis ir paspaudus prisijungimo mygtuką, jo duomenys išsaugomi duomenų bazėje bei parodomas pagrindinis vartotojo langas
<b>Alternatyvūs scenarijai</b>		Jei blogi prisijungimo duomenys parodomas klaidos pranešimas

4-5 lentelė – 4-12 lentelė aprašyti veiksmai su failu (nuotrauka, garso arba vaizdo įrašu), kurie apima failo valdymą bei veido ir /ar balsų keitimą.

4-5 lentelė. PA „Keisti veidą nuotraukoje“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 3</b>	<b>Keisti veidą nuotraukoje</b>
<b>Tikslas</b>		Vartotojas gali įkelti veido nuotraukos originalą, naują veido nuotrauką skirtą pakeitimui bei atsisiųsti/išsaugoti antros nuotraukos veidą pirmoje nuotraukoje.
<b>Dalyviai</b>		Vartotojas, svečias
<b>Priklausomybės</b>		<i>Peržiūrėti sukurtą failą</i>
<b>Nefunkciniai reikalavimai</b>		Sistema turėtų apdoroti <i>jpeg</i> ir <i>png</i> formato nuotraukas, <i>mp3</i> formato garso įrašus bei <i>mp4</i> formato vaizdo įrašus;  Keitimas turi trukti mažiau nei 10 sekundžių.
<b>Prieš-sąlygos</b>		Ateiti į keitimo langą ir pasirinkti nuotraukos tipą
<b>Sužadinimo sąlyga</b>		Įkelti nuotrauką ir paspausti apdorojimo mygtuką
<b>Po-sąlyga</b>		Parodyti pakeistą nuotrauką, leisti ją išsaugoti sistemoje arba atsisiųsti
<b>Pagrindinis scenarijus</b>		Vartotojas įkelia originalią nuotrauką bei nuotrauką su norimu pakeisti veidu, paspaudžia apdorojimo mygtuką ir sistemai sukūrus naują nuotrauką vartotojas gali ją pamatyti, atsisiųsti ar išsaugoti sistemoje
<b>Alternatyvūs scenarijai</b>		Padavus blogo formato failą, parodomas klaidos pranešimas

4-6 lentelė. PA „Keisti balsą garso įrašė“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 4</b>	<b>Keisti balsą garso įrašė</b>
<b>Tikslas</b>		Turi būti galimybė įkelti garso įrašo originalą, naują garso įrašą skirtą pakeitimui bei atsisiųsti/išsaugoti antro įrašo balsą pirmame įrašė;
<b>Dalyviai</b>		Vartotojas, svečias

<b>Priklausomybės</b>	<i>Peržiūrėti sukurtą failą</i>
<b>Nefunkciniai reikalavimai</b>	Sistema turėtų apdoroti <i>mp3</i> formato garso įrašus ;  Sistema turi garso įrašą – per mažiau nei 150% originalaus įrašo laiko.
<b>Prieš-sąlygos</b>	Ateiti į keitimo langą ir pasirinkti balso tipą
<b>Sužadinimo sąlyga</b>	Įkelti garso įrašus ir paspausti apdorojimo mygtuką
<b>Po-sąlyga</b>	Parodyti pakeistą garso įrašą, leisti jį išsaugoti sistemoje arba atsisiųsti
<b>Pagrindinis scenarijus</b>	Vartotojas įkelia originalų garso įrašą bei garso įrašą su norimu pakeisti balsu, paspaudžia apdorojimo mygtuką ir sistemai sukūrus naują garso įrašą vartotojas gali jį pamatyti, atsisiųsti ar išsaugoti sistemoje
<b>Alternatyvūs scenarijai</b>	Padavus blogo formato failą, parodomas klaidos pranešimas

4-7 lentelė. PA „Keisti veidą ir balsą vaido įrašė“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 5</b>	<b>Keisti veidą ir balsą vaido įrašė</b>
<b>Tikslas</b>		Turi būti galimybė įkelti vaizdo įrašo originalą, naują vaizdo įrašą skirtą pakeitimui bei atsisiųsti/išsaugoti antro vaizdo įrašo žmogaus veidą ir balsą esantį pirmame įrašė;
<b>Dalyviai</b>		Vartotojas
<b>Priklausomybės</b>		<i>Peržiūrėti sukurtą failą</i>
<b>Nefunkciniai reikalavimai</b>		Sistema turėtų apdoroti <i>mp4</i> formato vaizdo įrašus;  Sistema turi vaizdo įrašą – per mažiau nei 200% vaizdo įrašo laiko
<b>Prieš-sąlygos</b>		Ateiti į keitimo langą ir pasirinkti vaizdo įrašo tipą
<b>Sužadinimo sąlyga</b>		Įkelti vaizdo įrašus ir paspausti apdorojimo mygtuką
<b>Po-sąlyga</b>		Parodyti pakeistą vaizdo įrašą, leisti jį išsaugoti sistemoje arba atsisiųsti
<b>Pagrindinis scenarijus</b>		Vartotojas įkelia originalų vaizdo įrašą bei vaizdo įrašą su norimu pakeisti balsu ir veidu, paspaudžia apdorojimo mygtuką ir sistemai sukūrus naują vaizdo įrašą vartotojas gali jį pamatyti, atsisiųsti ar išsaugoti sistemoje
<b>Alternatyvūs scenarijai</b>		Padavus blogo formato failą, parodomas klaidos pranešimas

4-8 lentelė. PA „Peržiūrėti visus failus“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 6</b>	<b>Peržiūrėti visus failus</b>
<b>Tikslas</b>		Peržiūrėti visus galimus išsaugotus failus
<b>Dalyviai</b>		Vartotojas
<b>Priklausomybės</b>		<i>Peržiūrėti failą</i>
<b>Nefunkciniai reikalavimai</b>		-
<b>Prieš-sąlygos</b>		Būti prisijungus prie sistemos
<b>Sužadinimo sąlyga</b>		Nueiti į failų peržiūros langą
<b>Po-sąlyga</b>		Parodomi visi esami failai
<b>Pagrindinis scenarijus</b>		Vartotojas pasirenka failų peržiūros langą ir sistema jam parodo jam galimus matyti failus
<b>Alternatyvūs scenarijai</b>		Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas

4-9 lentelė. PA „Peržiūrėti failą“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 7</b>	<b>Peržiūrėti failą</b>
<b>Tikslas</b>		Peržiūrėti naujai sukurtą failą
<b>Dalyviai</b>		Vartotojas, svečias, svečias, svečias
<b>Priklausomybės</b>		<i>Keisti veidą nuotraukoje, Keisti balsą garso įrašė, Keisti veidą ir balsą vaido įrašė, Atsisiųsti vaizdo įrašą, Išsaugoti failą, Peržiūrėti visus</i>  <i>Failus, Pašalinti failą</i>

<b>Nefunkciniai reikalavimai</b>	-
<b>Prieš-sąlygos</b>	Sėkmingai įsivykdyti <i>Keisti veidą nuotraukoje, Keisti balsą garso įrašė ar Keisti veidą ir balsą vaido įrašė</i> panaudojimo atvejis arba paspausti konkretų failą jų peržiūros lange
<b>Sužadinimo sąlyga</b>	Sistema baigia failo apdorojimą arba vartotojas pasirenka konkretų failą.
<b>Po-sąlyga</b>	Parodomas failas
<b>Pagrindinis scenarijus</b>	Įvyksta sužadinimo sąlyga ir vartotojas mato failą bei gali jį atsisiųsti ar išsaugoti (jei buvo tik sugeneruotas) bei ištrinti
<b>Alternatyvūs scenarijai</b>	Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas

4-10 lentelė. PA „Atsisiųsti failą“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 8</b>	<b>Atsisiųsti failą</b>
<b>Tikslas</b>		Atsisiųsti naujai sukurtą failą
<b>Dalyviai</b>		Vartotojas, svečias
<b>Priklausomybės</b>		<i>Peržiūrėti failą</i>
<b>Nefunkciniai reikalavimai</b>		-
<b>Prieš-sąlygos</b>		Sėkmingai įsivykdyti <i>Keisti veidą nuotraukoje, Keisti balsą garso įrašė ar Keisti veidą ir balsą vaido įrašė</i> panaudojimo atvejis
<b>Sužadinimo sąlyga</b>		Vartotojas paspaudžia atsisiuntimo mygtuką
<b>Po-sąlyga</b>		Failas yra atsiunčiama į vartotojo kompiuterį
<b>Pagrindinis scenarijus</b>		Vartotojas failo peržiūros lange paspaudžia atsisiuntimo mygtuką ir jam failas yra atsiunčiamas
<b>Alternatyvūs scenarijai</b>		Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas

4-11 lentelė. PA „Išsaugoti failą“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 9</b>	<b>Išsaugoti failą</b>
<b>Tikslas</b>		Išsaugoti naujai sukurtą failą sistemoje ir vėliau galėti jį peržiūrėti
<b>Dalyviai</b>		Vartotojas
<b>Priklausomybės</b>		<i>Peržiūrėti failą</i>
<b>Nefunkciniai reikalavimai</b>		-
<b>Prieš-sąlygos</b>		Sėkmingai įsivykdyti <i>Keisti veidą nuotraukoje, Keisti balsą garso įrašė ar Keisti veidą ir balsą vaido įrašė</i> panaudojimo atvejis
<b>Sužadinimo sąlyga</b>		Vartotojas paspaudžia išsaugojimo mygtuką
<b>Po-sąlyga</b>		Failas yra išsaugomas sistemoje
<b>Pagrindinis scenarijus</b>		Vartotojas failo peržiūros lange paspaudžia išsaugojimo mygtuką ir pakeičia pavadinimą jei esamas netinka
<b>Alternatyvūs scenarijai</b>		Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas

4-12 lentelė. PA „Pašalinti failą“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 10</b>	<b>Pašalinti failą</b>
<b>Tikslas</b>		Pašalinti išsaugotą failą iš sistemos
<b>Dalyviai</b>		Vartotojas
<b>Priklausomybės</b>		<i>Peržiūrėti failą</i>
<b>Nefunkciniai reikalavimai</b>		-
<b>Prieš-sąlygos</b>		Sėkmingai įsivykdyti <i>Keisti veidą nuotraukoje, Keisti balsą garso įrašė ar Keisti veidą ir balsą vaido įrašė</i> panaudojimo atvejis
<b>Sužadinimo sąlyga</b>		Vartotojas paspaudžia pašalinimo mygtuką



<b>Po-sąlyga</b>	Failas pašalinimas iš sistemos
<b>Pagrindinis scenarijus</b>	Vartotojas failo peržiūros lange paspaudžia pašalinimo mygtuką ir sistemai pasitikslinus ar tikra vartotojas nori pašalinti failą sutinka
<b>Alternatyvūs scenarijai</b>	Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas

4-13 lentelė – 4-15 lentelė aprašyti veiksmai su keitimo modeliais.

**4-13 lentelė.** PA „Peržiūrėti modelius“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 11</b>	<b>Peržiūrėti modelius</b>
<b>Tikslas</b>		Peržiūrėti visus galimus modelius
<b>Dalyviai</b>		Vartotojas
<b>Priklausomybės</b>		<i>Įkelti modelį, Pašalinti modelį</i>
<b>Nefunkciniai reikalavimai</b>		-
<b>Prieš-sąlygos</b>		Būti prisijungus prie sistemos
<b>Sužadinimo sąlyga</b>		Nueiti į modelių peržiūros langą
<b>Po-sąlyga</b>		Parodomi visi esami modeliai
<b>Pagrindinis scenarijus</b>		Vartotojas pasirenka modelių peržiūros langą ir sistema jam parodo jam galimus matyti modelis
<b>Alternatyvūs scenarijai</b>		Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas

**4-14 lentelė.** PA „Pašalinti modelį“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 12</b>	<b>Pašalinti modelį</b>
<b>Tikslas</b>		Pašalinti anksčiau įkeltą modelį iš sistemos
<b>Dalyviai</b>		Vartotojas
<b>Priklausomybės</b>		<i>Peržiūrėti modelius</i>
<b>Nefunkciniai reikalavimai</b>		-
<b>Prieš-sąlygos</b>		Būti modelių peržiūros lange
<b>Sužadinimo sąlyga</b>		Vartotojas paspaudžia pašalinimo mygtuką
<b>Po-sąlyga</b>		Modelis pašalinimas iš sistemos
<b>Pagrindinis scenarijus</b>		Vartotojas modelių peržiūros lange paspaudžia pašalinimo mygtuką ir sistemai pasitikslinus ar tikra vartotojas nori pašalinti modelį, sutinka. Modelis pašalinamas iš sistemos
<b>Alternatyvūs scenarijai</b>		Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas

**4-15 lentelė.** PA „Įkelti modelį“ specifikacija

<b>Panaudojimo atvejis</b>	<b>Nr. 13</b>	<b>Įkelti modelį</b>
<b>Tikslas</b>		Įkelti naują modelį į sistemą
<b>Dalyviai</b>		Vartotojas, Administratorius
<b>Priklausomybės</b>		<i>Peržiūrėti modelius</i>
<b>Nefunkciniai reikalavimai</b>		-
<b>Prieš-sąlygos</b>		Būti modelių peržiūros lange
<b>Sužadinimo sąlyga</b>		Vartotojas paspaudžia įkėlimo mygtuką
<b>Po-sąlyga</b>		Modelis išsaugomas sistemoje sistemos
<b>Pagrindinis scenarijus</b>		Administratorius paspaudžia įkėlimo mygtuką, pasirenka norimą failą iš kompiuterio ir paspaudžia išsaugojimo mygtuką.
<b>Alternatyvūs scenarijai</b>		Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas

Galiausiai 4-16 lentelė – 4-18 lentelė aprašyti administratoriaus veiksmai su vartotojais.

**4-16 lentelė.** PA „Peržiūrėti vartotojus“ specifikacija

Panaudojimo atvejis	Nr. 14	Peržiūrėti vartotojus
Tikslas	Peržiūrėti visus vartotojus, užsiregistravusius sistemoje	
Dalyviai	Administratorius	
Priklausomybės	<i>Redaguoti vartotoją, Pašalinti vartotoją</i>	
Nefunkciniai reikalavimai	-	
Prieš-sąlygos	Būti prisijungus prie sistemos	
Sužadinimo sąlyga	Nueiti į vartotojų peržiūros langą	
Po-sąlyga	Parodomi visi esami vartotojai	
Pagrindinis scenarijus	Administratorius pasirenka vartotojų peržiūros langą ir sistema jam parodo vartotojų sąrašą	
Alternatyvūs scenarijai	Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas	

**4-17 lentelė.** PA „Redaguoti vartotoją“ specifikacija

Panaudojimo atvejis	Nr. 15	Redaguoti vartotoją
Tikslas	Keisti vartotojo informaciją	
Dalyviai	Administratorius	
Priklausomybės	<i>Peržiūrėti vartotojus</i>	
Nefunkciniai reikalavimai	-	
Prieš-sąlygos	Būti vartotojų peržiūros lange	
Sužadinimo sąlyga	Administratorius paspaudžia redagavimo mygtuką	
Po-sąlyga	Išsaugoma pakeista vartotojo informacija	
Pagrindinis scenarijus	Administratorius vartotojų peržiūros lange paspaudžia redagavimo mygtuką ir pakeičia norimus duomenis ir paspaudžia išsaugojimo mygtuką	
Alternatyvūs scenarijai	Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas	

**4-18 lentelė.** PA „Pašalinti vartotoją“ specifikacija

Panaudojimo atvejis	Nr. 16	Pašalinti vartotoją
Tikslas	Pašalinti vartotoją iš sistemos	
Dalyviai	Administratorius	
Priklausomybės	<i>Peržiūrėti vartotojus</i>	
Nefunkciniai reikalavimai	-	
Prieš-sąlygos	Būti vartotojų peržiūros lange	
Sužadinimo sąlyga	Administratorius paspaudžia pašalinimo mygtuką	
Po-sąlyga	Vartotojas pašalinamas iš sistemos	
Pagrindinis scenarijus	Administratorius vartotojų peržiūros lange paspaudžia pašalinimo mygtuką ir sistemai pasitikslinus ar tikra administratorius nori pašalinti vartotoją, sutinka. Vartotojas pašalinamas iš sistemos	
Alternatyvūs scenarijai	Įvykus nenumatytai aplinkybei parodomas klaidos pranešimas	

## 4.2.7. Funkciniai reikalavimai

### 4.2.7.1. Reikalavimai užduočių vykdymo greičiui

Sistema turi sugeneruoti nuotrauką per mažiau nei 10 sekundžių, o garso įrašą – per mažiau nei 150% originalaus įrašo laiko bei vaizdo įrašą – per mažiau nei 200% vaizdo įrašo laiko.

4-19 lentelė. Greičio reikalavimai

<b>Reikalavimas</b>	FR-1	<b>Reikalavimo tipas</b>	V9
<b>Panaudojimo atvejai</b>	3		
<b>Aprašymas</b>	Sistema turi sugeneruoti nuotrauką per mažiau nei 10 sekundžių		
<b>Pagrindimas</b>	Suteikti vartotojui greitą rezultatą		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Tikinimo kriterijus</b>	Nuotraukos veidas turi būti pakeičiamas per 10 sekundžių		
<b>Užsakovo tenkinimas</b>	32	<b>Užsakovo netenkinimas</b>	2
<b>Priklausomybės</b>		<b>Konfliktai</b>	Nėra
<b>Reikalavimas</b>	FR-2	<b>Reikalavimo tipas</b>	V9
<b>Panaudojimo atvejai</b>	45		
<b>Aprašymas</b>	Sistema turi sugeneruoti garso įrašą – per mažiau nei 150% originalaus įrašo laiko		
<b>Pagrindimas</b>	Suteikti vartotojui greitą rezultatą		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Tikinimo kriterijus</b>	Garso įrašas turi būti pakeičiamas per 150% originalaus įrašo laiko		
<b>Užsakovo tenkinimas</b>	3	<b>Užsakovo netenkinimas</b>	2
<b>Priklausomybės</b>		<b>Konfliktai</b>	Nėra
<b>Reikalavimas</b>	FR-3	<b>Reikalavimo tipas</b>	V9
<b>Panaudojimo atvejai</b>	5		
<b>Aprašymas</b>	Sistema turi sugeneruoti vaizdo įrašą per mažiau nei 200% vaizdo įrašo laiko.		
<b>Pagrindimas</b>	Suteikti vartotojui greitą rezultatą		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Tikinimo kriterijus</b>	Vaizdo įrašas turi būti pakeičiamas per 200% originalaus įrašo laiko		
<b>Užsakovo tenkinimas</b>	32	<b>Užsakovo netenkinimas</b>	
<b>Priklausomybės</b>		<b>Konfliktai</b>	Nėra
<b>Papildoma medžiaga</b>	Administratorius paspaudžia pašalinimo mygtuką		
<b>Istorija</b>	Vartotojas pašalinimas iš sistemos		

### 4.2.7.2. Patikimumas ir pasiekiamumas

Žemiau aprašyti pasiekiamumo ir atsako į klaidas reikalavimai.

4-20 lentelė. Pasiekiamumo reikalavimas

<b>Reikalavimas</b>	FR-4	<b>Reikalavimo tipas</b>	V9
<b>Panaudojimo atvejai</b>	Visi		
<b>Aprašymas</b>	Sistema turi būti visada pasiekiamas, jei yra naudojama produkcinėje aplinkoje		
<b>Pagrindimas</b>	Plačiosios visuomenės nariai gali bet kuriuo metu užsimanyti naudotis sistema		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Tikinimo kriterijus</b>	Vartotojai bet kuriuo jiems patogiu metu turi galimybę naudotis sistema		
<b>Užsakovo tenkinimas</b>	32	<b>Užsakovo netenkinimas</b>	
<b>Priklausomybės</b>		<b>Konfliktai</b>	Nėra

#### 4-21 lentelė. Klaidų reikalavimas

<b>Reikalavimas</b>	FR-5	<b>Reikalavimo tipas</b>	V9
<b>Panaudojimo atvejai</b>	Visi		
<b>Aprašymas</b>	Sistema, įvykus klaidai turi pranešti vartotojui apie įvykusią klaidą. Taip pat jei klaida įvyko dėl vartotojo klaidos, turi pranešti kaip išvengti tos klaidos		
<b>Pagrindimas</b>	Užtikrinti kuo geresnę vartotojo patirtį		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Tikinimo kriterijus</b>	Įvykus klaidai vartotojas gauna įspėjimą apie ją bei kaip to gali išvengti.		
<b>Užsakovo tenkinimas</b>	32	<b>Užsakovo netenkinimas</b>	
<b>Priklausomybės</b>		<b>Konfliktai</b>	Nėra

#### 4.2.7.3. Reikalavimai apdorojamų duomenų apimtims

Sistema vienu metu turi galėti patenkinti 150 vartotojų bei 50 keitimo operacijų vienu metu

#### 4-22 lentelė. Apkrovos reikalavimas

<b>Reikalavimas</b>	NFR-9	<b>Reikalavimo tipas</b>	V9
<b>Panaudojimo atvejai</b>	Visi		
<b>Aprašymas</b>	Sistema vienu metu turi galėti patenkinti 150 vartotojų bei 50 keitimo operacijų, jei yra naudojama produkcinėje aplinkoje		
<b>Pagrindimas</b>	-		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Tikinimo kriterijus</b>	-		
<b>Užsakovo tenkinimas</b>	2	<b>Užsakovo netenkinimas</b>	2
<b>Priklausomybės</b>		<b>Konfliktai</b>	Nėra

#### 4.2.8. Reikalavimai sistemos išvaizdai

##### 4.2.8.1. Išvaizdos reikalavimai

Sistemos išvaizdai specialių reikalavimų nėra

##### 4.2.8.2. Stiliaus reikalavimai

Stiliui šioje sistemoje nebuvo aprašyta daug reikalavimų – tik spalvų.

#### 4-23 lentelė. Spalvos reikalavimai

<b>Reikalavimas</b>	NFR-1	<b>Reikalavimo tipas</b>	V10
<b>Aprašymas</b>	Sistemos pagrindinį spalvos turi būti mėlyna ir violetinė		
<b>Pagrindimas</b>	Mėlyna spalva asocijuojasi su patikimumu, o violetinė su kūrybiškumu		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Atitikimo kriterijus</b>	Sistemos pagrindinį spalvos turi būti mėlyna ir violetinė		

#### 4.2.9. Reikalavimai panaudojamumui

Žemiau esančiose lentelėse aprašyti dizaino, kalbos, vartotojo apmokymo bei personalizavimo reikalavimai.

#### 4.2.9.1. Naudojimosi paprastumas

4-24 lentelė. Dizaino reikalavimai

Reikalavimas	NFR-2	Reikalavimo tipas	V12
Aprašymas	Dizainas ir mygtukai turi būti informatyvūs ir turi būti naudojama paprasta anglų kalba pagrindinėms funkcijoms.		
Pagrindimas	Produktu gali naudotis žmonės nepilnai suprantantys anglų kalbą		
Šaltinis	Sistemos užsakovas		
Atitikimo kriterijus	Vartotojas, besinaudojantis sistema, turi lengvai suprasti kaip atlikti užduotis nežinodamas/ mažai žinodamas anglų kalbą tačiau turintis kompiuterinį raštingumą		

#### 4.2.9.2. Personalizavimo ir kalbos konfigūravimo reikalavimai

4-25 lentelė. Kalbos reikalavimai

Reikalavimas	NFR-3	Reikalavimo tipas	V12
Aprašymas	Sistema turi vartotojui pasirinkti norimą kalbą iš kelių, jei yra naudojama produkcinėje aplinkoje		
Pagrindimas	Ne visi žmonės moka anglų kalbą ir ne visiems ji patogi		
Šaltinis	Sistemos užsakovas		
Atitikimo kriterijus	Vartotojas gali pasirinkti jam priimtinausią sistemos kalbą		

4-26 lentelė. Modelio atsiminimo reikalavimas

Reikalavimas	NFR-4	Reikalavimo tipas	V12
Aprašymas	Sistema turi atsiminti paskutinį naudotą modelį bei jį duoti kaip numatytąjį sesijos metu		
Pagrindimas	Kad būtų vartotojas galėtų greičiau ir efektyviau naudotis sistema.		
Šaltinis	Sistemos užsakovas		
Atitikimo kriterijus	Pakeitus failą naudojantis modelių ir vėl nuėjus į keitimo langą, tą patį modelį turi mesti pirmą		

#### 4.2.9.3. Mokymosi reikalavimai

4-27 lentelė. Mokymo reikalavimai

Reikalavimas	NFR-5	Reikalavimo tipas	V12
Aprašymas	Vartotojai turi gebėti naudotis sistema iš karto be apmokymų, todėl sistemos išvaizda turi būti paprasta, o mygtukai aiškiai suprantami.		
Pagrindimas	Sistema bus skirta plačiosios visuomenės nariams		
Šaltinis	Sistemos užsakovas		
Atitikimo kriterijus	Vartotojas be specialaus išsilavinimo turi gebėti pakeisti nuotrauką		

#### 4.2.9.4. Suprantamumo ir mandagumo reikalavimai

4-28 lentelė. Sąvokų reikalavimai

Reikalavimas	NFR-6	Reikalavimo tipas	V12
Aprašymas	Sąvokos bei simboliai, kurie gali būti nežinomi visiems žmonėms, turi būti paaiškinti.		
Pagrindimas	Sistemoje naudojama kalba turi būti suprantama plačiajai visuomenei.		
Šaltinis	Sistemos užsakovas		
Atitikimo kriterijus	Vartotojas be specialaus išsilavinimo turi gebėti pakeisti nuotrauką		

#### 4.2.9.5. Prieinamumo neįgaliesiems reikalavimai

Prieinamumo neįgaliesiems reikalavimų nėra

#### 4.2.9.6. Suteikiami patogumai

Suteikiamų patogumų sistemoje nenumatyta

#### 4.2.10. Efektyvumo ir našumo reikalavimai

##### 4.2.10.1. Reikalavimai išplečiamumui

Sistemos vartotojų kiekio reikalavimas aprašytas žemiau esančioje lentelėje.

4-29 lentelė. Išplečiamumo reikalavimai

Reikalavimas	NFR-10	Reikalavimo tipas	V13
Aprašymas	Per pirmus metus sistema turi pajėgti apdoroti iki 1000 vartotojų. Po dviejų metų šitas skaičius tikimasi išsaugoti iki 2500 vartotojų		
Pagrindimas	-		
Šaltinis	Sistemos užsakovas		
Atitikimo kriterijus	-		

##### 4.2.10.2. Reikalavimai produkto ilgaamžiškumui

Sistemos ilgaamžiškumo reikalavimas aprašytas žemiau esančioje lentelėje.

4-30 lentelė. Igaamžiškumo reikalavimai

Reikalavimas	NFR-11	Reikalavimo tipas	V13
Aprašymas	Tikimasi, kad sistema veiks neviršydama biudžeto 2 metus		
Pagrindimas	-		
Šaltinis	Sistemos užsakovas		
Atitikimo kriterijus	-		

#### 4.2.11. Reikalavimai veikimo sąlygoms

##### 4.2.11.1. Numatoma fizinė aplinka

Sistemos fizinės aplinkos reikalavimas aprašytas žemiau esančioje lentelėje.

4-31 lentelė. Fizinės aplinkos reikalavimai

Reikalavimas	NFR-12	Reikalavimo tipas	V14
Aprašymas	Leidžiant vaizdo įrašus turi būti mygtukas nutildantis garsą.		
Pagrindimas	Sistema gali būti naudojama viešoje vietoje		
Šaltinis	Sistemos užsakovas		
Atitikimo kriterijus	Leidžiant vaizdo įrašą specialus mygtukas išjungia garsą		
Reikalavimas	NFR-13	Reikalavimo tipas	V14
Aprašymas	Vartotojo sąsaja turi būti pritaikyta įvairiems įrenginiams		

<b>Pagrindimas</b>	Vartotojas naudosis sistema savo įrenginyje (telefone, kompiuteryje)
<b>Šaltinis</b>	Sistemos užsakovas
<b>Atitikimo kriterijus</b>	Sistema turi veikt įvairiuose įrenginiuose

#### 4.2.11.2. Aplinkosauginiai reikalavimai

Aplinkosauginių reikalavimų sistemoje nenumatyta.

#### 4.2.11.3. Reikalavimai darbui su gretimomis sistemomis

Sistemos suderinamumo reikalavimas aprašytas žemiau esančioje lentelėje.

4-32 lentelė. Suderinamumo reikalavimai

<b>Reikalavimas</b>	NFR-14	<b>Reikalavimo tipas</b>	V14
<b>Aprašymas</b>	Sistema turi veikti šioje naršyklėse: Google Chrome, Mozilla Firefox ir Safari		
<b>Pagrindimas</b>	Vartotojas gali naudotis įvairiomis naršyklėmis		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Atitikimo kriterijus</b>	Sistema turi veikt įvairiose naršyklėse		

#### 4.2.11.4. Reikalavimai sistemos platinimo/gamybos formatui

Reikalavimų sistemos platinimui nėra

#### 4.2.11.5. Reikalavimai leidybos procesui

Reikalavimų leidybos procesui nėra

#### 4.2.11.6. Reikalavimai atvirkštiniam suderinamumui

Reikalavimų atvirkštiniam suderinamumui nėra

#### 4.2.12. Reikalavimai sistemos priežiūrai

##### 4.2.12.1. Sistemos aptarnavimas

Sistemos aptarnavimui reikalavimų nėra

##### 4.2.12.2. Sistemos palaikymas

Sistemos palaikymo reikalavimas aprašytas žemiau esančioje lentelėje.

4-33 lentelė. Sistemos palaikymo reikalavimas

<b>Reikalavimas</b>	NFR-15	<b>Reikalavimo tipas</b>	V15
<b>Aprašymas</b>	Sistemoje turi būti įkeltas bent vienas modelis kiekvienam keitimo tipui		
<b>Pagrindimas</b>	Be modelių nebus galima keisti failų – atlikti pagrindinės sistemos funkcijos		
<b>Šaltinis</b>	Sistemos užsakovas		

<b>Atitikimo kriterijus</b>	Paleidus sistemą joje turi būti mažiausiai trys modeliai – po vieną kiekvienam keitimo tipui
-----------------------------	--

#### 4.2.12.3. Pritaikymo kitoms platformoms reikalavimai

Reikalavimo pritaikymui kitoms sistemoms nėra, nes sistema bus prieinama per naršyklę

#### 4.2.13. Reikalavimai saugumui

##### 4.2.13.1. Prieigos reikalavimai

Sistemos prieigos reikalavimas aprašytas žemiau esančioje lentelėje.

4-34 lentelė. Prieigos reikalavimas

Reikalavimas	NFR-16	Reikalavimo tipas	V16
<b>Aprašymas</b>	Tik sistemos administratoriai turi prieigą prie vartotojų duomenų		
<b>Pagrindimas</b>	Išlaikyti vartotojų konfidencialumą		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Atitikimo kriterijus</b>	Vartotojai nemato vienas kitų asmeninių duomenų		

##### 4.2.13.2. Vientisumo reikalavimai

Sistemos vientisumo reikalavimas aprašytas žemiau esančioje lentelėje.

4-35 lentelė. Vientisumo reikalavimas

Reikalavimas	NFR-17	Reikalavimo tipas	V16
<b>Aprašymas</b>	Sistema turi validuoti įvedamus duomenis		
<b>Pagrindimas</b>	Kad išvengti galimų klaidų bei nesklandumų vartotojams		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Atitikimo kriterijus</b>	Netaisyklingi duomenys turi būti validuojami		

##### 4.2.13.3. Reikalavimai privatumui

Sistemos privatumo reikalavimas aprašytas žemiau esančioje lentelėje.

4-36 lentelė. Privatumas reikalavimas

Reikalavimas	NFR-18	Reikalavimo tipas	V16
<b>Aprašymas</b>	Sistema turi informuoti vartotojus apie duomenų saugojimo politiką prieš registraciją		
<b>Pagrindimas</b>	Kad išvengti galimų klaidų bei nesklandumų vartotojams bei sistemos savininkams		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Atitikimo kriterijus</b>	Prieš registraciją parodoma duomenų saugojimo nuostatai		

##### 4.2.13.4. Audito reikalavimai

Sistemos privatumo reikalavimas aprašytas žemiau esančioje lentelėje.

4-37 lentelė. Privatumas reikalavimas

Reikalavimas	NFR-19	Reikalavimo tipas	V16
<b>Aprašymas</b>	Sistemos auditas turi būti atliekamas kartą per mėnesį		



<b>Pagrindimas</b>	Kad surasti sistemos spragas bei gerinti vartotojo patirtį
<b>Šaltinis</b>	Sistemos užsakovas
<b>Atitikimo kriterijus</b>	Kas mėnesį atliekamas sistemos auditas

#### 4.2.13.5. Reikalavimai savisaugai nuo išorinių grėsmių

Sistemos kūrimo reikalavimas aprašytas žemiau esančioje lentelėje.

**4-38 lentelė.** Sistemos kūrimo reikalavimas

<b>Reikalavimas</b>	NFR-20	<b>Reikalavimo tipas</b>	V16
<b>Aprašymas</b>	Sistema turi būti kuriama naudojant naujas bei patikimas technologijas		
<b>Pagrindimas</b>	Kad palaikyti sistemos saugumą		
<b>Šaltinis</b>	Sistemos užsakovas		
<b>Atitikimo kriterijus</b>	Sukūrus sistemą bandyti ją nulaužti		

#### 4.2.14. Kultūriniai reikalavimai

Kultūrinių reikalavimų nėra

#### 4.2.15. Atitikties reikalavimai

Reikalavimų atitikties standartams nėra

#### 4.2.16. Egzistuojantys sprendimai

##### 4.2.16.1. Prieinamos sistemos

Šiuo metu yra jau egzistuojančio sistemos, kurios gali pasiūlyti balso ar veido keitimą:

- *AutoVC* (<https://auspicious3000.gib.io/autovc-demo/>);
- *Deep Nostalgia* (<https://www.myheritage.lt/deep-nostalgia>).

##### 4.2.16.2. Prieinami komponentai

Šiuo metu „*deepfake*“ technologija šiuo metu populiarėja ir yra sukuriama vis daugiau sprendimų, komponentų bei bibliotekų veido ir balso keitimui galima rasti lengvai. Atsirinkti tinkamas bibliotekas bus atlikti išsami jų analizė

Taip vartotojo sąsajai galima rasti šablonų, kurie pagreitintų jos sukūrimą

##### 4.2.16.3. Kopijuotini sprendimai

Balso keitimui galima naudoti ConVoice projektą (<https://rebryk.gib.io/convoice-demo>).

#### 4.2.17. Naujos problemos

##### 4.2.17.1. Poveikis diegimo aplinkai

Sistema kuriama naujai, todėl poveikio diegimo aplinkai neturės.

##### 4.2.17.2. Poveikis esamoms sistemoms

Sistema kuriama naujai, todėl poveikio esamoms sistemoms neturės.

### 4.2.17.3. Probleminė naudotojų reakcija

Sistema kuriama naujai, todėl naudotojai nepatirs problemų dėl pasikeitusios aplinkos ar funkcionalumo.

### 4.2.17.4. Apribojimai diegimo aplinkoje

Sistema kuriama naujai, todėl diegimo aplinka bus jai pritaikyta ir problemų nekils.

### 4.2.17.5. Kitos potencialios problemos

Sistema kuriama naujai, naujų potencialių problemų nenusimato.

## 4.2.18. Uždaviniai

### 4.2.18.1. Sistemos kūrimo procesas

Technologijos – kuriant sistemą bus naudojama:

- *Linux* operacinė sistema;
- *Python* ir *JavaScript* programavimo kalbos.

Priemonės:

- „*PyCharm*“ ir „*Visual Studio Code*“ integruotos kūrimo aplinkos;
- „*MagicDraw*“ modeliavimo įrankis;

Žmonės:

- Programuotojas, kuris turi patirties kuriant informacines sistemas bei kuris yra susipažinęs su dirbtiniu intelektu;
- Mašininio mokymosi ekspertas, kuris konsultuos programuotoją DI klausimais;

Kadangi sistema nėra didelė, bei jai atlikti yra skirtas ilgas laiko tarpas, ją sukurs vienas žmogus. Žinoma, karto nuo karto jis konsultuosis su mašininio mokymosi ekspertu. Projekto vykdymo grafikas pavaizduotas 4-5 pav. Kūrimo planas., projektas truks beveik 1.5 metus. Projekto pabaigoje turi būti sukurta visa informacinė sistema su veikiančiais mašininio mokymosi modeliais.

## 4.2.19. Kūrimo planas



4-5 pav. Kūrimo planas

1. Reikalavimų apibrėžimas – nustatoma pradinė galutinio produkto vizija, pagrindinės funkcijos, kurias sistema turės atlikti (nubraižoma pirminė panaudojimo atvejų diagrama);

2. Analizė – analizuojamam veiklos sritis bei galimos priemonės. Pasirenkamos geriausios priemonės bei detalizuojami panaudojimo atvejai, jų ryšiai tarpusavyje;
3. Projektavimas – suprojektuojama kuriamos sistemos architektūra, nustatoma sistemos struktūra iš posistemių, klasių ir karkasų, kuriuose paskirstyti panaudojimo atvejai;
4. Realizavimas – realizuojama suprojektuota sistema;
5. Testavimas – sukuriamas testavimo planas bei pagal jį ištestuojama sistema ir, reikalui esant, pataisoma sistema;
6. Diegimas – sistema įdiegiama bei padaroma prieinama visiems;
7. Palaikymas – sistema nuolat prižiūrima, bei, reikalui esant, pataisoma.

#### **4.2.20. Migravimas į naują produktą**

##### **4.2.20.1. Migravimo į naują produktą reikalavimai**

Kadangi esamo produkto nėra, o yra kuriamas naujas, migravimas į naują produktą nereikalingas

##### **4.2.20.2. Duomenų transformavimo reikalavimai**

Kadangi esamo produkto nėra, o yra kuriamas naujas, duomenų transformavimas nereikalingas

#### **4.2.21. Rizikų įvertinimas**

##### **4.2.21.1. Rizikų tikimybių apibrėžimai**

Žemiau yra parodyta šio projekto tikimybių kategorijos ir kiek tai kategorijai priklausoma tikimybės skaitinė reikšmė, kur reikšmė  $\geq 0$ , bet  $\leq 1$ .

**4-39 lentelė.** Tikimybių kategorijos

<b>Tikimybės kategorija</b>	<b>Tikimybė</b>	<b>Apibūdinimas</b>
Labai aukšta	0.90	Tikėtinas rizikos įvykis
Aukšta	0.70	Rizikos įvykis labiau tikėtinas nei neįvykis
Vidutinė	0.50	Rizikos įvykis gali įvykti arba neįvykti
Žema	0.30	Rizikos įvykis yra mažiau tikėtinas nei neįvykis
Labai žema	0.10	Rizikos įvykio nesitikima

##### **4.2.21.2. Rizikos tikimybės ir poveikio matrica**

Toliau yra pateikta rizikos poveikio matrica, parodanti rizikos poveikio ir tikimybės kombinaciją. Rizikos, kurios patenka į raudoną langelį turi didžiausią prioritetą, o į žalią – mažiausią.

Rizikos yra paskirstytos tokia tvarka:

- tikimybė mažesnė arba lygi 0.05 – mažo prioriteto;
- tikimybė didesnė už 0.05 bet mažesnę ar lygi 0.3 – vidutinio prioriteto;
- tikimybė didesnę už 0.3 – aukšto prioriteto.

#### 4-40 lentelė Rizikos tikimybių matrica

Tikimybė	Poveikis				
0.90	0.05	0.09	0.18	0.36	0.72
0.70	0.04	0.07	0.14	0.28	0.56
0.50	0.03	0.05	0.10	0.20	0.40
0.30	0.02	0.03	0.06	0.12	0.24
0.10	0.01	0.01	0.02	0.04	0.08
	0.05	0.10	0.20	0.40	0.80

#### 4.2.21.3. Rizikų aprašymas

Toliau yra pateiktos projekto rizikos:

#### 4-41 lentelė Projekto rizikos

Nr.	Rizika	Poveikis	Tikimybė	Poveikis
1	Kuriamos sistemos bibliotekos nepalaikomos iOS ar Windows OS	Katastrofinis	Vidutinė	Negalima naudotis nepalaikoma OS, eikvojamas laikas rasti sprendimui
2	Sistema neatitinka tikslinių reikalavimų	Ribinis	Žema	Priklausomai nuo neatitikimo dydžio krenta sistemos reputacija
3	Mokymosi rezultatų praradimas naudojant Google Colab	Kritinis	Žema	Dėl platformos taisyklių sunaudojama daugiau laiko nei planuota, prarandami eksperimentų duomenys
4	Netinkamas eksperimentų etapų dokumentavimas	Ribinis	Vidutinė	Prarandami eksperimentų rezultatai, kainuoja laiko ir lėšų, jei platforma, kurioje vykdomi eksperimentai kainuoja
5	Galimai pasenusios Python ar naudojamų bibliotekų versijos	Ribinis	Labai žema	Neveikiančios ir /ar nepalaikomos bibliotekos/resursai. Naujų bibliotekų/resursų paieška bei pritaikymas kainuoja
6	Surasti geresni sprendimai nei suprojektuota planavimo metu	Ribinis	Vidutinė	Reikalingi papildomas laikas įgyvendinti naują sprendimą arba sukurta sistema/komponentas greičiausiai bus prastesnės kokybės palyginus su egzistuojančia rinka
7	Blogai sukurtos UML diagramos	Nežymus	Aukšta	Kuriant bei palaikant sistemą nėra tinkamos dokumentacijos, išsauga laiko kaštai
8	Audito metu nesurandamos visos problemos	Kritinis	Žema	Nerastos problemos toliau egzistuoja, kurios gali atnešti rimtesnių (finansinių ir baudžiamųjų) padarinių; gadinama reputacija
9	Sistemos bei komponento kūrimo laikas numatytas per daug optimistiškas	Katastrofinis	Aukšta	Per mažai numatytas laiko nespėjama pilnai ir kokybiškai įgyvendinti numatytų reikalavimų
10	Prasta, neaiški arba nepilna vartotojo dokumentacija	Ribinis	Žema	Vartotojui nėra aišku, kaip naudotis sistema – nepatenkinti vartotojai, krenta sistemos/komponento reputacija
11	Dėl nenumatytų aplinkybių laikas reikalingas papildomoms užduotims	Katastrofinis	Vidutinė	Nepasiekiamas pristatymo data, trūksta numatyto funkcionalumo

12	Prastas sistemos dizainas, kuriuo sunku naudotis	Nežymus	Labai žema	Nepatenkinti vartotojai, krenta sistemos/komponento reputacija
13	Blogi komponento įdiegimo failai	Katastrofinis	Labai žema	Vartotojas turi įdėti daugiau pastangų norėdamas naudotis komponentu, dėl ko krenta reputacija
14	Problematiškas komponento įkėlimas į saugyklą	Katastrofinis	Labai žema	Vartotojai neturi prieigos prie komponento bei projektas nėra įgyvendintas
15	Kvalifikacijos trūkumas kuriant mašininio mokymosi modelius	Ribinis	Žema	Mašininio mokymosi modeliai sukuriami prastos kokybės bei juos naudojant nėra pasiekiamas norimo rezultato bei vartotojai yra nepatenkinti
16	Neįgyvendinti failų keitimo laiko reikalavimai	Nežymus	Žema	Failai yra ilgiau keičiami nei vartotojams būtų patogiu ir rinka gali pasiūlyti, dėl to vartotojai gali pasirinkti alternatyvią sistemą/komponentą
17	Per maža pakeistų failų kokybė (vaizdo bei garso kokybė)	Ribinis	Žema	Per maža vaizdo ir garso kokybė gali netenkinti vartotojų ir jie gali ieškoti alternatyvių sistemų
18	Per maža pakeistų failų kokybė (sukurtas failas neatitinka norimo rezultato)	Kritinis	Vidutinė	Failo keitimas yra svarbiausias kuriamas funkcionalumas, todėl pačio keitimo rezultato kokybė turi būti pakankama norint turėti užbaigtą produktą
19	Per maža techninė dokumentacija	Ribinis	Aukšta	Neturint tikslios ir pilnos techninės dokumentacijos sunkėja sistemos/komponento palaikymas bei kūrimo ir palaikymo kaštai.
20	Per maža sistemos kūrėjų kvalifikacija	Kritinis	Labai žema	Per maža kvalifikacija gali lemti nekokybišką sistemą arba net viso projekto žlugimą

#### 4.2.21.4. Rizikų sprendimas (RMMM)

Atsakas į rizikas (rizikų sprendimas) aprašytas naudojantis *RMMM* metodologija – rizikų vengimas, stebėjimas ir valdymas ( angl. *Risk Mitigation, Monitoring and Management*)

1. Kuriamos sistemos bibliotekos nepalaikomos iOS ar Windows OS:
  - i. Vengimas – dėl nepalaikomų bibliotekų neveiks sistema/komponentas, todėl svarbu iš anksto tam pasiruošti:
    - i. Turėti galimybę paleisti atskiras sistemos/komponento dalis (ypač svarbu komponentui);
    - ii. Iš anksto pranešti apie nepalaikomas operacines sistemas bei kompiuterių modelius;
  - ii. Stebėjimas – stebėti sistemos veikimą iš veikimo pranešimų (angl. *logs*) (nepalaikomos sistemos aplinkos pranešimas turi būti kritinio lygio). Taip pat kuo greičiau (per savaitę), reaguoti į komponento problemas bei pranešimus apie jį.
  - iii. Valdymas – atsiradus nebepalaikomumo klaidai, tiek sistemoje, tiek komponente patikrinti tos bibliotekos veikimą bei pakeisti bibliotekos versiją į veikiančią. Jei veikiančios nėra – ieškoti kitokio būdo įgyvendinti norimą funkcionalumą.
2. Sistema neatitinka tikslinių reikalavimų:
  - i. Vengimas – laiku atnaujinti reikalavimų dokumentaciją bei pažymėti, kada ir kodėl tai buvo atnaujinta;

- ii. Stebėjimas – kuriant sistemą bei dirbant prie reikalavimų dokumentacijos stebėti, ar ji yra atnaujinta. Jei nėra atnaujinta – pačiam atnaujinti arba raštu pranešti/pažymėti, kur reikia atnaujinti.
  - iii. Valdymas – suradus neatitikimą tarp reikalavimų bei sistemos įgyvendinimo išsiaiškinti, ar sistemos į
  - iv. gyvendintas variantas yra galimas bei atnaujinti reikalavimų dokumentaciją ir pažymėta kada ir kodėl ji buvo atnaujinta.
3. Mokymosi rezultatų praradimas naudojant *Google Colab*:
- i. Vengimas – tarpinius apmokymo rezultatus saugoti į Google diską;
  - ii. Valdymas – praradus duomenis pasižymėti apmokymo parametrus bei bandyti apmokyti iš naujo;
4. Netinkamas eksperimentų etapų dokumentavimas:
- i. Vengimas – aprašyti eksperimentų rezultatus tik įgyvendinus eksperimentą, o neturint tam laiko – išsaugot eksperimento rezultatus ir parašyti nors kelis komentarus.
  - ii. Stebėjimas – stebėti ar visi rezultatai tinkamai aprašyti.
  - iii. Valdymas – suradus neaprašytą eksperimentą – pabandyti jį aprašyti, jei jo rezultatai geri. Jei rezultatai geri ir eksperimentas nebesuprantamas – jį įgyvendinti iš naujo
5. Galimai pasenusios *Python* ar nebepalaikomos, tačiau veikiančios bibliotekų versijos:
- i. Vengimas – pradėdant naudoti naują biblioteką – naudoti kuo naujesnę jos versiją;
  - ii. Stebėjimas – stebėti sistemos/komponento įspėjimus.
  - iii. Valdymas – Pastebėjus įspėjimą apie senstančią bibliotekos versiją pasiruošti ją atnaujinti:
    - i. Skubiai – jei biblioteka nebepalaikoma;
    - ii. Neskubiai – jei biblioteka dar palaikoma, tačiau gaunami įspėjimai
6. Surasti geresni sprendimai nei suprojektuota planavimo metu:
- i. Vengimas – prieš projektuojant sistemą/komponentą pasidaryti gerą jau egzistuojančių sprendimų analizę;
  - ii. Stebėjimas – reikia domėtis jau egzistuojančiais arba naujai kuriamais sprendimais tam, kad sukurtas produktas būtų kokybiškesnis ir patrauklesnis;
  - iii. Valdymas – radus geresnį sprendimą nei yra suprojektuota jį išanalizuoti bei nuspręsti, ar naujo sprendimo sekimas atneš daugiau naudos ar žalos.
7. Blogai sukurtos UML diagramos:
- i. Vengimas – gerai išanalizuoti reikalavimus bei egzistuojančius sprendimus prieš kuriant UML diagramas. Nusprendus nesekti diagramų pavyzdžiu – jas atnaujinti bei pridėti priedą kada ir kodėl jos buvo atnaujintos
  - ii. Stebėjimas – atidžiai kurti diagramas bei pastebėjus klaidas – jas užfiksuoti raštiškai;
  - iii. Valdymas – UML diagramų klaidas užfiksuoti raštiškai bei pagal individualų atvejį jas tvarkyti bei dokumentuoti tvarkymą.
8. Audito metu nesurandamos visos problemos:
- i. Vengimas – turėti pilną bei tikslų audito planą ir jį kruopščiai vykdyti;
  - ii. Stebėjimas – stebėti ar audito plane, jo įgyvendinimo metu bei sistemoje nėra spragų.
  - iii. Valdymas – suradus audito spragas atnaujinti audito vykdymo planą.
9. Sistemos bei komponento kūrimo laikas numatytas per daug optimistiškas:
- i. Vengimas – kruopščiai išanalizuoti reikalavimus bei suprojektuotą architektūrą ir pagal tai numatyti galimą sistemos ir komponento sukūrimo laiką, jei įmanoma duoti užduotims daugiau laiko nei numanoma kad joms prireiks;

- ii. Stebėjimas – stebėti ar vykdymo planas ir suplanuoti laikai atitinka realybę;
  - iii. Valdymas – atsiliesiant nuo numatyto grafiko iš naujo nustatyti prioritetus bei atnaujinti grafiką arba jei yra galimybė – ieškoti papildomų resursų;
10. Prasta, neaiški arba nepilna vartotojo dokumentacija:
- i. Vengimas – sukūrus sistemą/komponentą pilnai dokumentuoti vartotojui suprantama kalba, bei pridėjus papildomo funkcionalumo – dokumentuoti ir tai;
  - ii. Stebėjimas – stebėti, ar vartotojai yra patenkinti esama dokumentacija bei ar ją supranta. Taip pat galima pašaliniams žmonėms duoti išsibandyti sistemą bei gauti iš jų atsiliepimus;
  - iii. Valdymas – laiku ar neaiškiai dokumentavus funkcionalumą kuo greičiau tai ištaisyti bei pridėjus naują funkcionalumą jį dokumentuoti.
11. Dėl nenumatytų aplinkybių laikas reikalingas papildomoms užduotims:
- i. Vengimas – kuriant užduočių atlikimo grafiką skirti papildomai laiko nenumatytoms užduotims.
  - ii. Stebėjimas – stebėti ar laikomasi projekto pradžioje numatyto užduočių atlikimo grafiko;
  - iii. Valdymas – pastebėjus, kad nepavyksta laikytis grafiko iš naujo nustatyti prioritetus ir sukurti užduočių atlikimo grafiką.
12. Prastas sistemos dizainas, kuriuo sunku naudotis:
- i. Vengimas – atlikti detalią įvairių sistemų analizę bei sukurti šios sistemos/komponento dizainą;
  - ii. Stebėjimas – įgyvendinant dizainą stebėti ar jis ir realybėje yra patogus vartotojui;
  - iii. Valdymas – pastebėjus dizaino klaidų ar galimų patobulinimų juos dokumentuoti ir atsiradus galimybei – įgyvendinti.
13. Blogi komponento įdiegimo failai:
- i. Vengimas – naudoti komponentų saugyklos šabloną bei ištestuoti komponento įdiegimą skirtingose aplinkose (mašinos);
  - ii. Stebėjimas – reguliariai pratestuoti komponento įdiegimą lokaliaje mašinoje;
  - iii. Valdymas – nepavykus įdiegti komponento ieškoti problemų bei jas spręsti
14. Problematiškas komponento įkėlimas į saugyklą:
- i. Vengimas – išsiaiškinti saugyklos reikalavimus bei pilnai paruošti komponentą įkėlimui, naudoti komponentų saugyklos šabloną;
  - ii. Stebėjimas – testuoti, ar įkeltas komponentas taisyklingai gali būti įrašytas į lokalią mašiną;
  - iii. Valdymas – atsiradus problemoms komponento įkėlimui į komponentų saugyklą arba jo įsirašymui į lokalią mašiną ieškoti problemų bei jas ištaisyti.
15. Kvalifikacijos trūkumas kuriant mašininio mokymosi modelius:
- i. Vengimas – įvertinti esamą kvalifikaciją prieš pradėdant eksperimentus bei jos trūkstant dalyvauti mokymuose ir ją pasikelti;
  - ii. Stebėjimas – stebėti, ar vykdant eksperimentus užtenka žinių juo kokybiškai vykdyti, suprasti ir dokumentuoti;
  - iii. Valdymas – trūkstant kvalifikacijos dalyvauti mokymuose, jei turima papildomai laiko, arba ieškoti žmonių su turima kvalifikacija bei iš jų sužinoti reikiamus sprendimus/žinias;
16. Neįgyvendinti failų keitimo laiko reikalavimai:

- i. Vengimas – prieš nustatant reikalavimus išanalizuoti esamus veido ir balso keitimo sprendimus bei išsiaiškinti jų vykdymo laikus ir pagal tai nustatyti reikalavimus. Taip pat numatyti, kad skirtingose aplinkose failų keitimo laikas gali skirtis;
  - ii. Stebėjimas – parašyti automatizuotus testus bei stebėti failų pakeitimo laiką
  - iii. Valdymas – pastebėjus, kad failų keitimo laiko reikalavimai yra neįgyvendinti išsiaiškinti to priežastį ir bandyti tai išspręsti, jei turima resursų. Neturint pakankamai resursų dokumentuoti faktą bei jo priežastis;
17. Per maža pakeistų failų kokybė (vaizdo bei garso kokybė) :
- i. Vengimas – prieš kuriant sistemą bei jos metu ieškoti geriausių sprendimų, padedančių palaikyti kuo geresnę vaizdo ir garso failų kokybę;
  - ii. Stebėjimas – parašyti automatizuotus testus, kurie tikrina ir failų kokybę bei pastebėjus prastą kokybę šį faktą užfiksuoti raštiškai;
  - iii. Valdymas – išsiaiškinti, kodėl failų kokybė yra per maža bei ieškoti sprendimų, kaip ją pakelti.
18. Per maža pakeistų failų kokybė (sukurtas failas neatitinka norimo rezultato) :
- i. Vengimas – skirti pakankamai laiko mašininio mokymosi modelių kūrimui, eksperimentavimui su jais ir jų tobulinimui. Eksperimentavimo metu išsiaiškinti geresnio ar blogesnio jų veikimo priežastis. Skirti papildomo laiko eksperimentavimui;
  - ii. Stebėjimas – stebėti bei išsaugoti skirtingų eksperimentų bei sistemos/komponento klasių veikimą siekiant išsiaiškinti kada išgaunama geriausia pakeistų failų kokybė;
  - iii. Valdymas – esant per mažai pakeistų failų kokybei skirti papildomai laiko eksperimentavimui, tiksliai išsiaiškinti kaip veikia dabartinis sprendimas, ieškoti naujų metodų/sprendimų failo keitimui.
19. Per maža techninė dokumentacija:
- i. Vengimas – kuriant naują funkcionalumą ar keičiant sistemos reikalavimus iškart tai dokumentuoti jei yra laiko, o jei nėra – raštiškai pažymėti, ką reikia dokumentuoti;
  - ii. Stebėjimas – kuriant ir palaikant produktą stebėti, ar techninė dokumentacija atitinka realų produktą;
  - iii. Valdymas – suradus produkto ir dokumentacijos neatitikimus juos ištaisyti. Tuo metu neturint laiko – raštiškai pažymėti, kada buvo rastas neatitikimas ir jį apibūdinti.
20. Per maža sistemos kūrėjų kvalifikacija:
- i. Vengimas – sistemos projektavimo metu išsiaiškinti, kokių reikia kvalifikavimų bei skirti laiko tų kvalifikaciją įgijimui. Taip pat numatyti papildomo laiko trūkstamų kvalifikacijų įgijimui;
  - ii. Stebėjimas – stebėti, ar atliekant užduotis užtenka žinių.
  - iii. Valdymas – turint laiko įgyti reikalingų kvalifikacijų, o neturint – ieškoti žmonių, kurie turi trūkstama kvalifikacijas ir gali jomis pasidalinti.

#### **4.2.21.5. Rizikų poveikio ir tikimybių analizė**

Šiame skyriuje pavaizduota rizikų tikimybių poveikio matrica bei rizikų prioritetai su sprendimo laiko reikalavimais



4-42 lentelė Rizikos poveikio matrica

Tikimybė	Labai aukšta				
	Aukšta	7	19		9
	Vidutinė		4, 6	18	1, 11
	Žema	16	10, 15, 17	8	
	Labai žema	12	2, 5	3, 20	13, 14
		Nežymus	Ribinis	Kritinis	Katastrofinis
		<b>Poveikis</b>			

4-43 lentelė Rizikos ir jų prioritetai

Rizikos prioritetas	Rizikos	Reagavimo laikas
Aukštas	Rizikos 1, 9, 11	1 d.d
Vidutinis	Rizikos 8, 18, 19	Savaitė
Žemas	Rizikos 2, 3, 4, 5, 6, 7, 10, 12, 13, 14, 15, 16, 17, 20	Mėnesis

#### 4.2.21.6. Projekto priežiūra ir auditas

Sistemos priežiūra:

- Kas pusė metų tikrinama *Python* programavimo kalbos bei naudojamų bibliotekų versijos ir sąveika. Esant naujesnėms versijoms – jos atnaujinamos;

Komponento priežiūra

- Kas pusė metų tikrinama *Python* programavimo kalbos bei naudojamų bibliotekų versijos ir sąveika. Esant naujesnėms versijoms – jos atnaujinamos;
- Kartą per mėnesį tikrinama komponento būseną bei iškilusias problemas.

Numatoma vykdyti sistemos auditą atsižvelgiant į šios punktus:

- Saugumo auditas – pasirūpinti, kad tiek sistema, tiek pateiktas komponentas užtikrintų atitinkamą prieigos kontrolę.
- Informacinių sistemų auditas – siekiant įsitikinti, kad pakeisti failai būtų atitinkamai pažymėti metaduomenyse.
- Projekto auditas – sukurtas produktas bei kūrimo procesas atitinka kūrimo reikalavimus bei įdiegta sistema veikia tinkamai.

#### 4.2.21.7. Kaštai

Projektui įgyvendinti reikalingas personalas:

- Vienas programuotojas (Indrė Dimšė), turinti bakalauro laipsnį programų sistemų srityje. Su vadovo pagalba aprašo, suprojektuoja, ištestuoja ir realizuoja sistemą. Šiam asmeniui priklausys sukurto produkto autorinės teisės.
- Vadovas (Rytis Maskeliūnas), turintis daktaro laipsnį informacinių technologijų srityje. Vadovauja kūrimo procesui, padeda studentui su iššylančiomis problemomis ir klausimais.

Išlaidos pavaizduotos žemiau esančioje lentelėje.

#### 4-44 lentelė. Išlaidų lentelė

Išlaidos	Vienetas	Vienetų skaičius	Vieneto kaina, Eur	Viso, Eur
Žmonių ištekliai				
Vadovas	Mėnesis	18*0,05 etato	3300	2700
Programuotojas	Mėnesis	18*0,2 etato	1800	6,480
<i>Iš viso žmonių išteklių</i>				9180
2. Įranga ir prekės				
Kompiuteris	Vienetas	1	799,99	799,99
<i>Iš viso Įranga ir prekės</i>				799,99
6. Viso tiesioginiai projekto kaštai				9,979.99

#### 4.2.22. Naudotojo dokumentacija ir apmokymas

##### 4.2.22.1. Reikalavimai naudotojų dokumentacijai

Vartotojo dokumentacija – dokumentacija, kuri padės vartotojui sklandžiau naudotis sistema bei kurioje bus aprašyta kaip ja naudotis, kokie failų tipai gali būti naudojami bei kaip paruošti modelius įkėlimui.

Vidinė dokumentacija – dokumentacija, kurioje bus aprašytos sistemos architektūra, struktūra bei kiti svarbūs faktai, reikalingi palaikyti bei tobulinti sistemą.

##### 4.2.22.2. Reikalavimai naudotojų apmokymui

Sistema bus skirta plačiajai visuomenei, todėl apmokymų kaip ja naudotis nereikės. Visa informacija, kuri reikalinga vartotojui, bus pateikta vartotojo dokumentacijoje.

#### 4.2.23. Perspektyviniai reikalavimai

Ateityje būtų galima ne tik pakeisti failą, bet ir jį pakeitus redaguoti toje pačioje sistemoje.

#### 4.2.24. Idėjos sprendimams

Keičiant vaizdo įrašą, galima atskirai sukurti vaizdo įrašo be garso keitimą bei naudoti balso keitimą, taip sutaupant modelio treniravimą.

### 4.3. Architektūros specifikacija

Architektūros specifikacijoje pateikiamas architektūrinis kuriamos sistemos vaizdas, kuriame surinkti ir pateikti svarbiausi architektūriniai sprendimai, kurie buvo atlikti projektuojant sistemą. Šio dokumento pagrindu vėliau bus sudaroma detali projekto architektūra, kuria remiantis bus atliekama realizacija. Šis dokumentas tarnauja kaip bendravimo medžiaga tarp programinės įrangos architekto ir kitų komandos narių dėl architektūrinių sistemos kūrimo sprendimų.

Šiame skyriuje aprašyta kuriamos sistemos architektūra naudojantis UML modeliavimo įrankiu *MagicDraw*. Sistemos architektūra pavaizduota panaudojimo atvejų, veiklos, klasių ir paketų diagramomis. Kadangi kuriama sistema nėra sudėtinga bei visas jos funkcionalumas bus paimtas iš

*Django* karkaso administratoriaus sistemos funkcionalumo, architektūros vaizdas nėra labai detaliai specifikuojamas. Taip pat aprašoma kaip bus palaikoma sistemos kokybė.

#### **4.3.1. Architektūros pateikimas**

Dokumente kuriamos sistemos architektūros vaizdas pakeikiamas naudojant UML kalba bei grafinio atvaizdavimo programa *MagicDraw*. Sistemos architektūra pavaizduota šiomis diagramomis:

- Panaudojimo atvejų vaidas;
  - Panaudojimo atvejų diagrama;
- Statinis vaidas;
  - Sistemos išskaidymo į paketus diagrama;
- Dinaminis vaidas;
  - Veiklos diagramos.
- Išdėstymo vaidas;
  - Išdėstymo diagrama;
- Duomenų vaidas;
  - Klasių diagrama;

Taip pat pateikiamas veido ir/ar balso keitimo procesų architektūros.

#### **4.3.2. Architektūros tikslai ir apribojimai**

Sprendimus, susijusius su sistemos architektūros projektavimu, įtakoiantys reikalavimai:

Realizuota sistema turi būti paprasta, gerai dokumentuota ir suprantama kitiems programuotojams, kad jie galėtų ją panaudoti savo reikmėms;

Sistema turi sugebėti priimti 3 tipų failus (garso, vaizdo ir vaizdo įrašo) bei būti lengvai įgyvendinama priimti skirtingų formatų modelius;

Sistema turi turėti ne tik grafinę sąsają vartotojams, bet ir aplikacijų programavimo sąsają (API).

#### **4.3.3. Panaudojimo atvejų vaizdas**

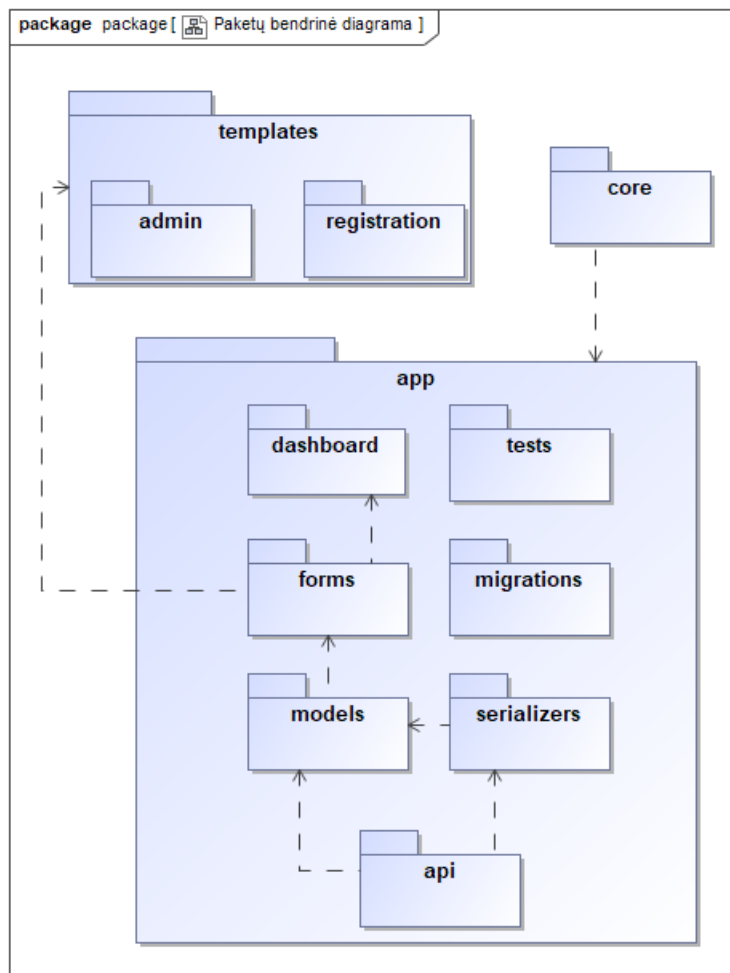
Panaudojimo atvejų diagrama ir visi atvejai aprašyti 4.2.6 skyrelyje.

### **3.6. Sistemos statinis vaizdas**

#### **3.6.1. Apžvalga**

4-6 pav. pavaizduota sistemos paketų diagrama, kurioje galima išvelgti Django karkasą:

- Pagrindinis paketas *core*;
- Aplikacija *app*, kurioje įgyvendintas funkcionalumas;
- *templates* paketas, kuriame yra perrašyti *Django* karkaso vaizdai.



4-6 pav. Paketų diagrama

Pavaizduotoje paketų diagramoje *core* paketas savyje turi tik nustatymus, reikalingus sistemos veikimui. *templates* pakete laikomi perrašomi Django karkaso administravimo lango šablonai, kuriuos tiesiogiai naudoja formos. Jei kažkuris šablonas nėra aprašytas, jis yra paimamas iš Django bibliotekos.

*app* pakete pavaizduotas pagrindinis sistemos veikimas:

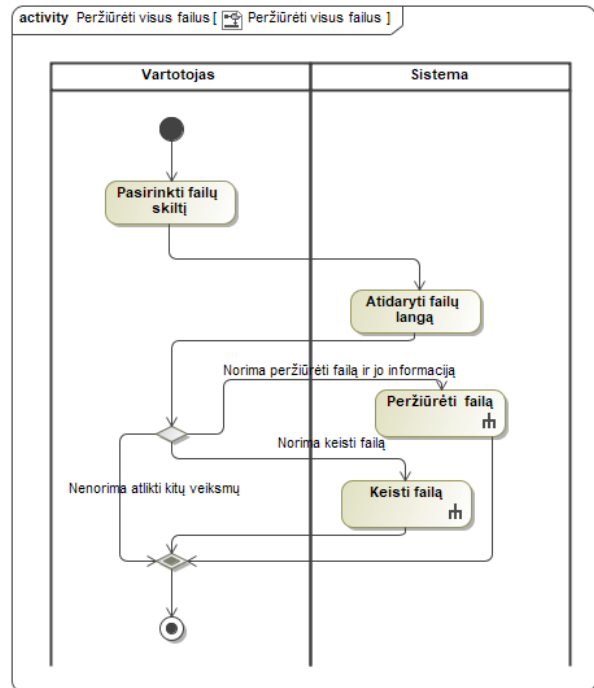
- *Models* – modeliai bei jų ryšiai;
- *Api* – aplikacijų programavimo sąsaja;
- *Serializers* – API naudojami funkcijos;
- *Forms* – formos, reikalingos įgyvendinti *dashboard* netradicinį karkaso funkcionalumą;
- *Dashboard* – funkcijos, kurios įgyvendina netradicinį karkaso funkcionalumą.

#### 4.3.4. Sistemos dinaminis vaizdas

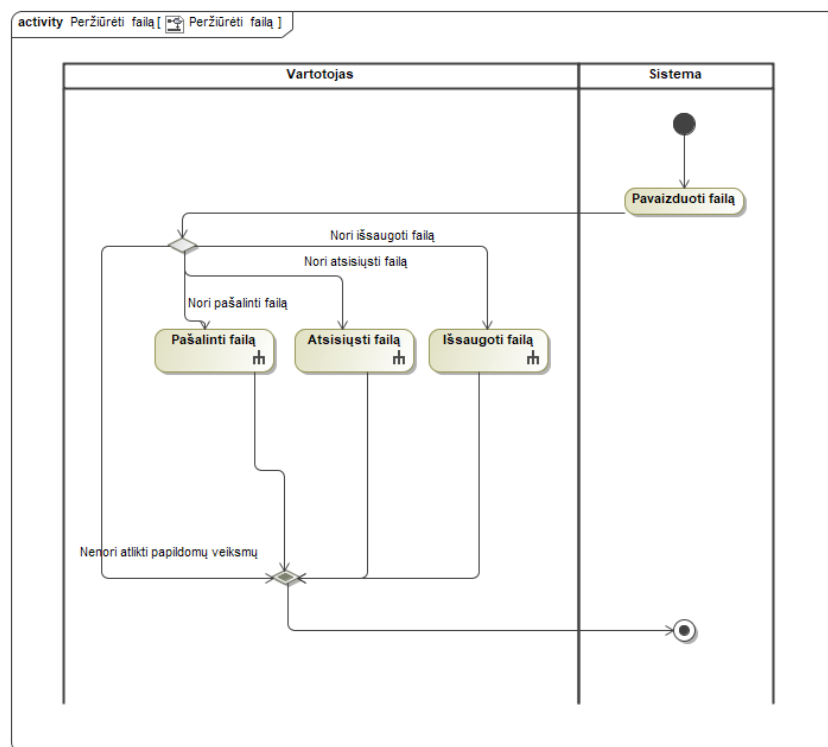
Šiame skyriuje aprašomas dinaminis sistemos vaizdas naudojant tik veiklos diagramas, nes sistema yra paprasta bei jai sukurti naudojamas Django karkasas bei jo administracinė dalis. Taip pat magistrinis darbas yra skirtas išsiaiškinti, kokios technologijos šiuo metu yra naudojamos keičiant vieno žmogaus veidą ir balsą, kitu bei kaip sukurti tikslesnius mašininio mokymosi modelius, kurie apdoroja vaizdą bei garsą, todėl pati sistema kuriama kaip papildomas produktas.

Kuriama sistema naudos Django karkaso funkcijas. Svečio posistemėje naudojamas numatytos funkcijos prisijungimui bei registracijai.

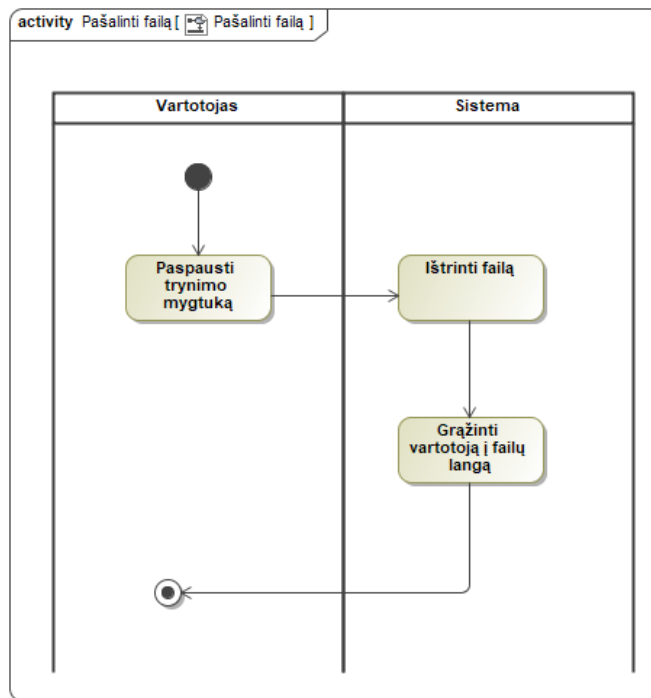
Failų posistemės panaudojimo atvejų diagramos pavaizduotos 4-7 pav. – 4-11 pav. Kadangi karkaso duotų funkcijų neužtenka įgyvendinti projekto reikalavimų, failų keitimo panaudojimo atvejai bus įgyvendinti papildomai prie duoto funkcionalumo.



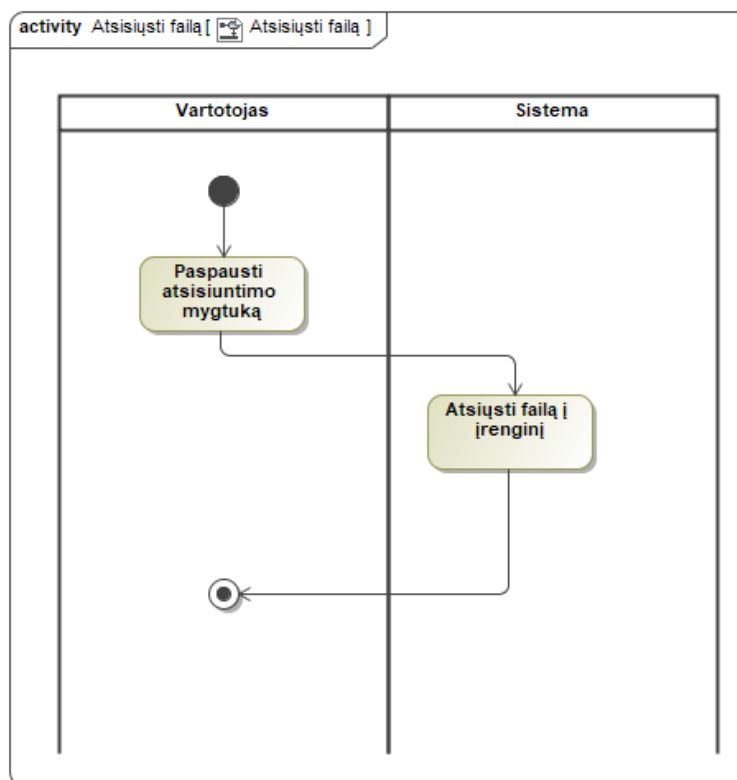
4-7 pav. PA „Peržiūrėti visus failus“ veiklos diagrama



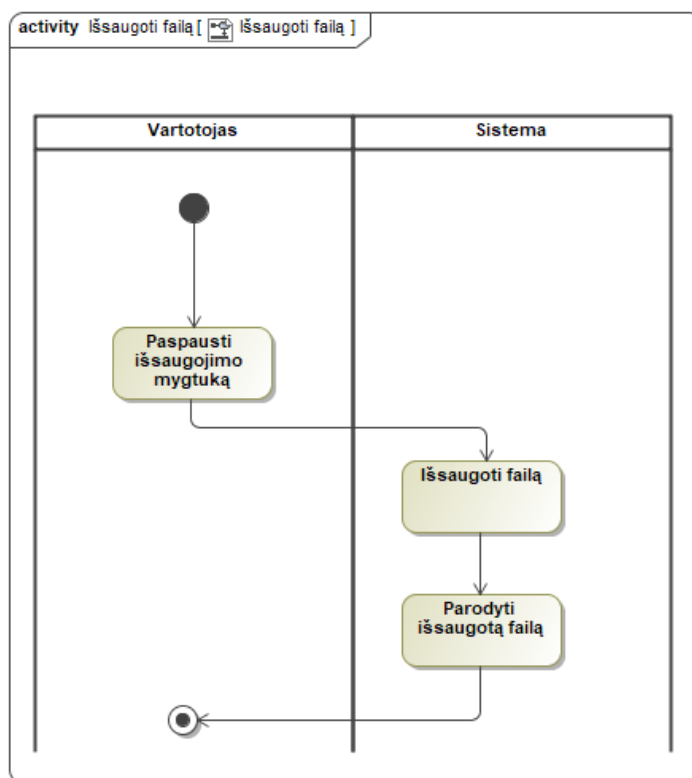
4-8 pav. PA „Peržiūrėti failą“ veiklos diagrama



4-9 pav. PA „Pašalinti failą“ veiklos diagrama

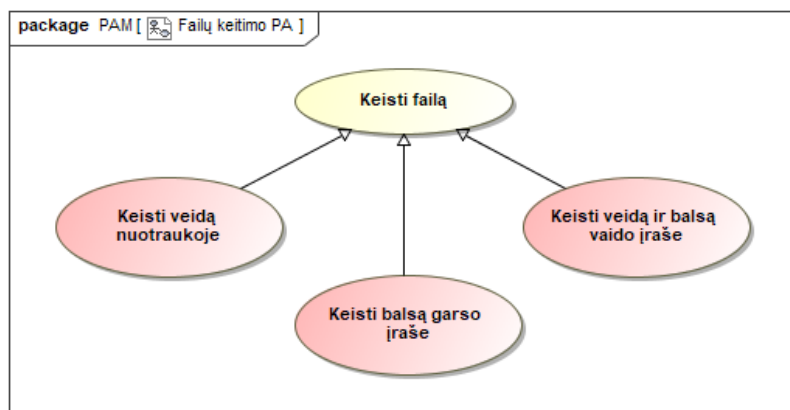


4-10 pav. PA „Atsisiųsti failą“ veiklos diagrama

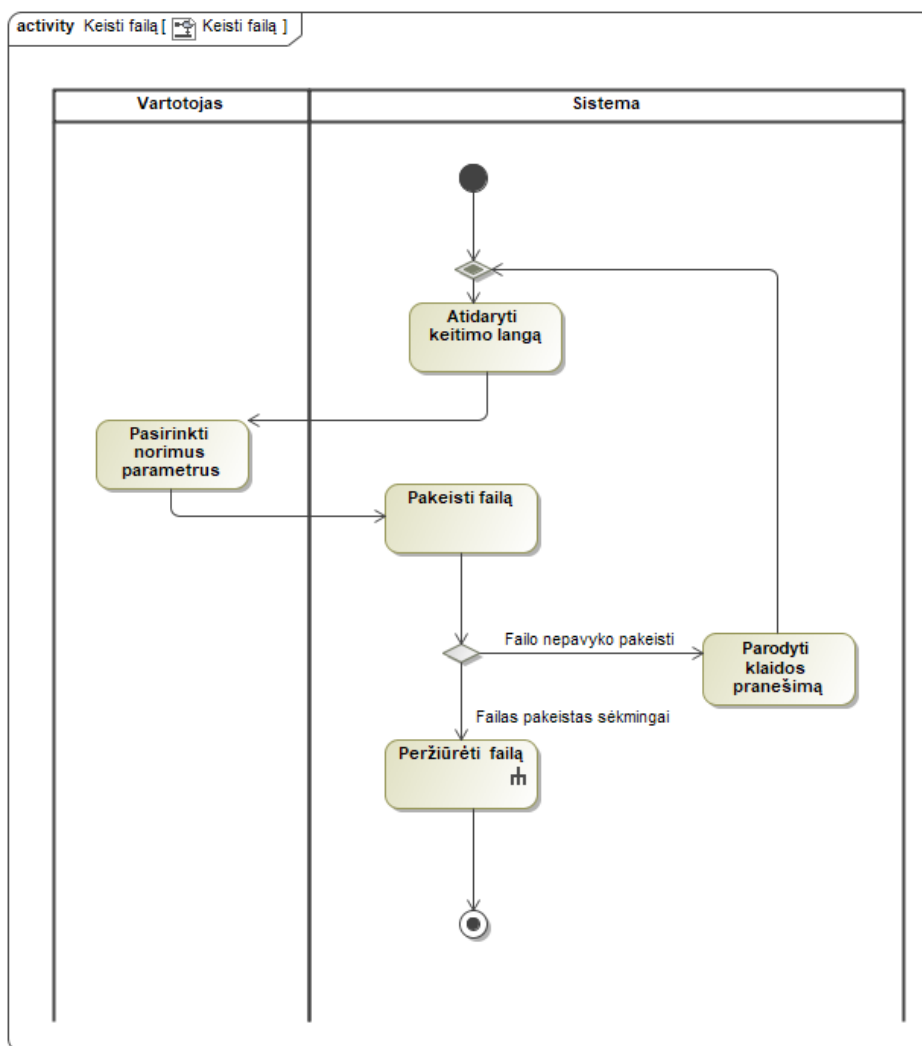


4-11 pav. PA „Išsaugoti failą“ veiklos diagrama

4-12 pav. pavaizduotas failų keitimo panaudojimo atvejo išskaidymas pagal tipus. Tačiau šio proceso veiksmų eiga yra vienoda ir pavaizduota 4-13 pav. veiklos diagramoje bei 4-14 pav. sąveikos diagramoje.

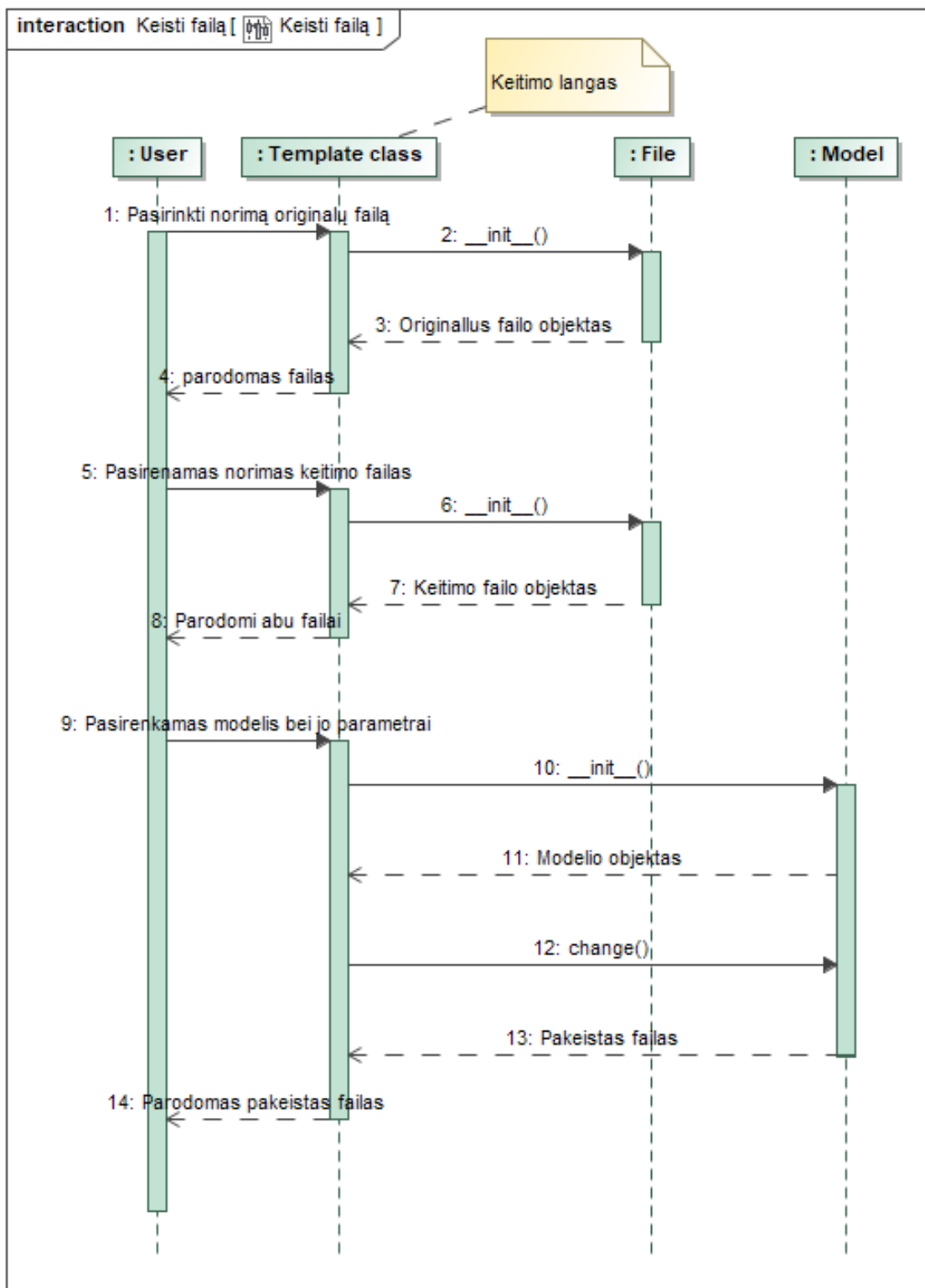


4-12 pav. Failų keitimo detalūs panaudojimo atvejai



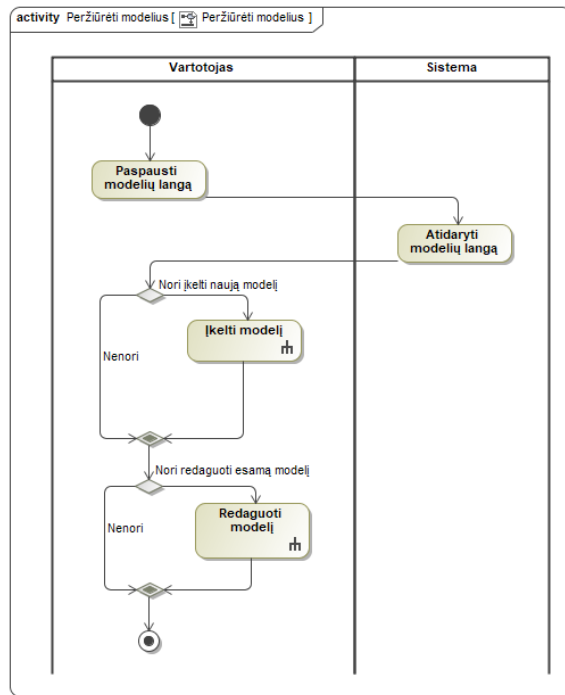
4-13 pav. PA „Keisti failą“ veiklos diagrama



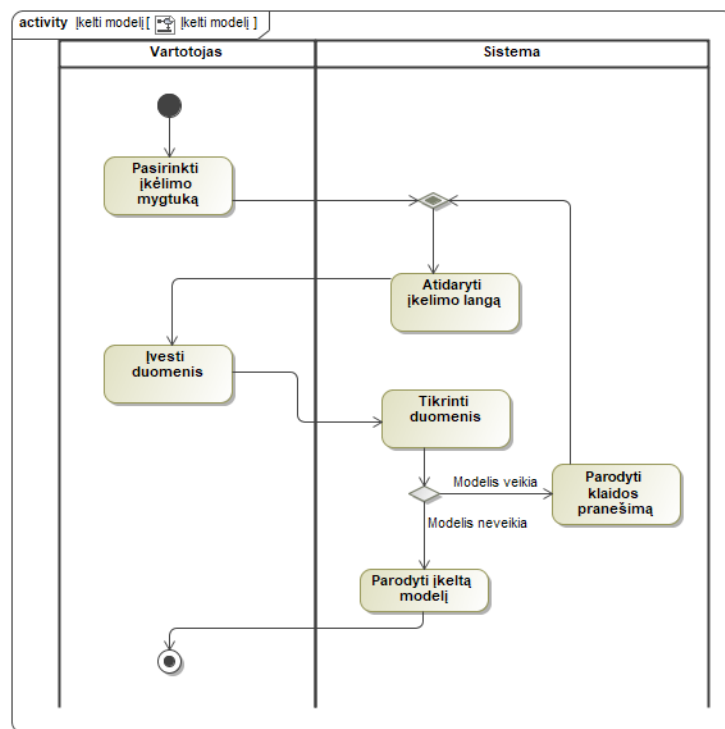


4-14 pav. PA „Keisti failą“ sąveikos diagrama

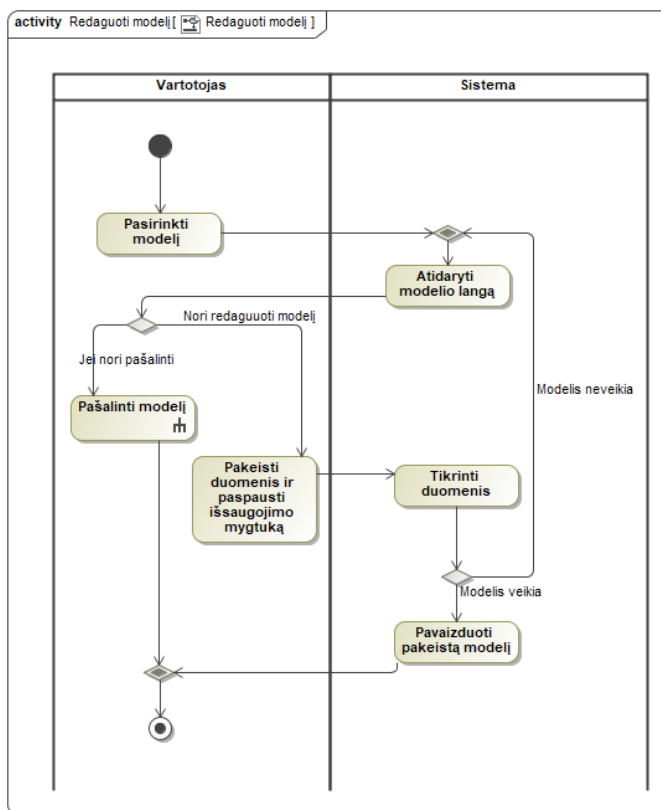
Tačiau modelių ir administratoriaus posistemėje karkaso funkcionalumo užtenka įgyvendinti reikalavimus. Jų veiklos diagramos pavaizduotos 4-15 pav. – 4-21 pav.



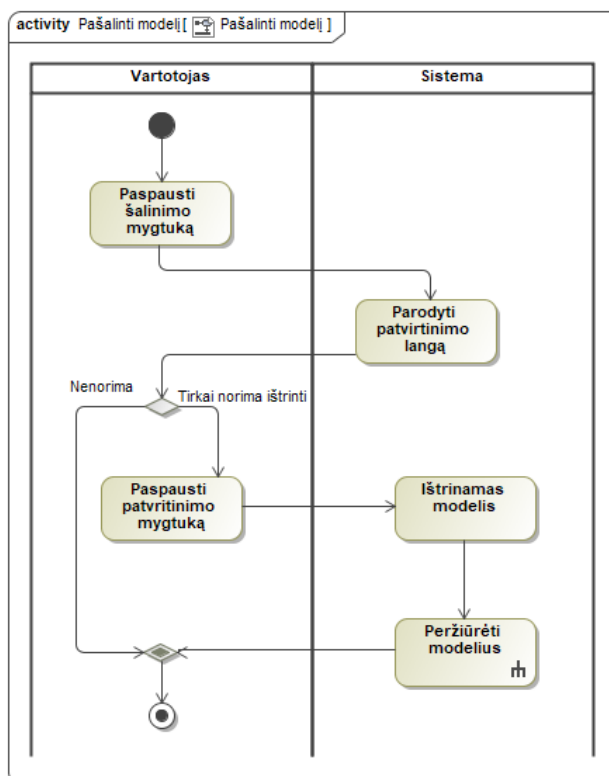
4-15 pav. PA „Peržiūrėti modelius“ veiklos diagrama



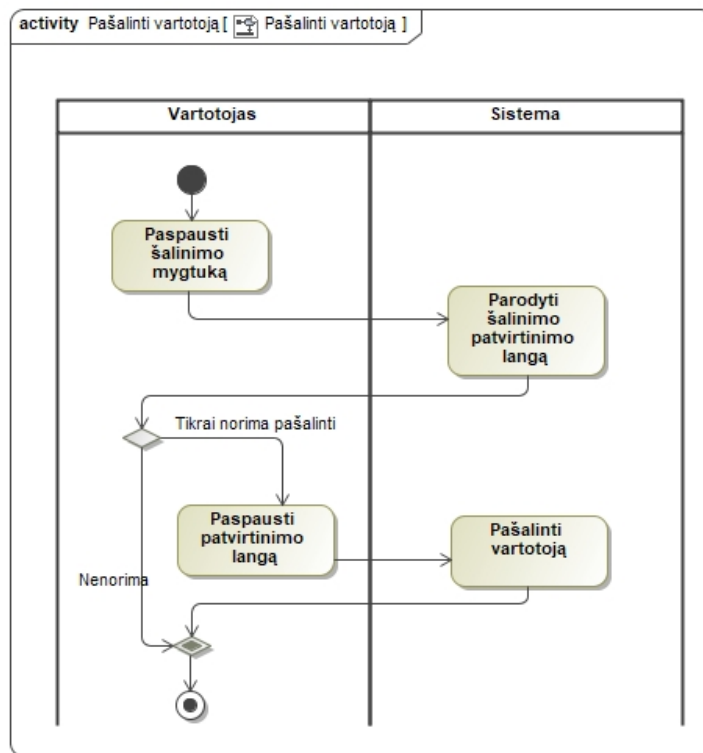
4-16 pav. PA „Įkelti modelį“ veiklos diagrama



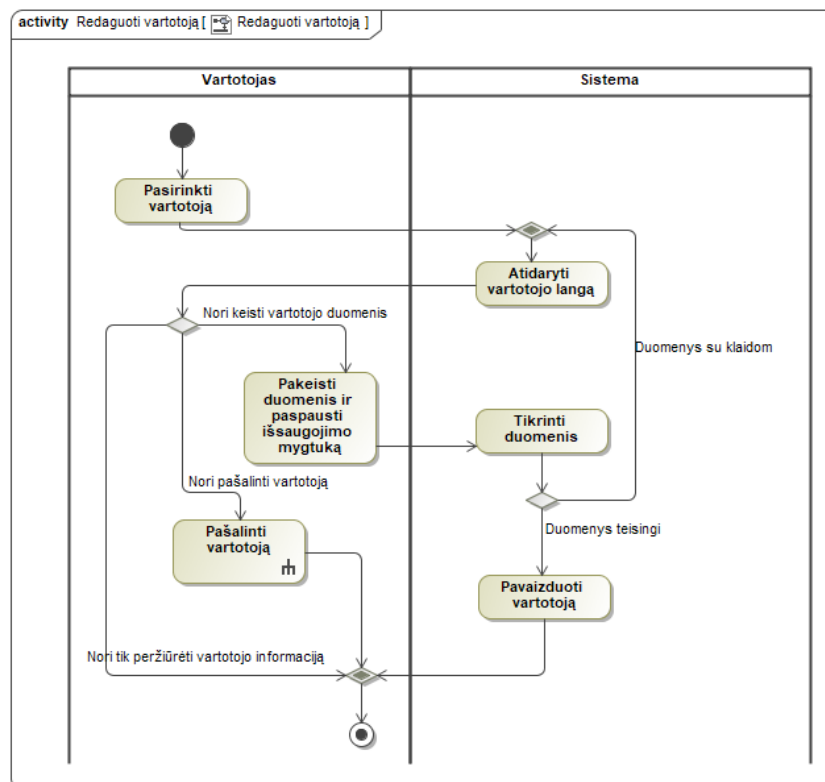
4-17 pav. PA „Redaguoti modelį“ veiklos diagrama



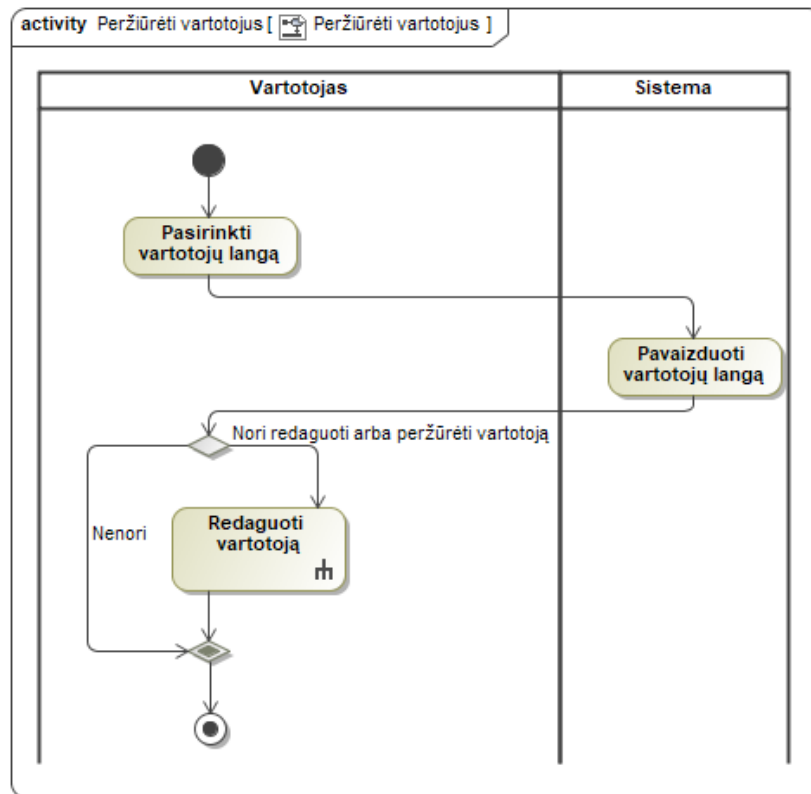
4-18 pav. PA „Pašalinti modelį“ veiklos diagrama



4-19 pav. PA „Pašalinti vartotoją“ veiklos diagrama



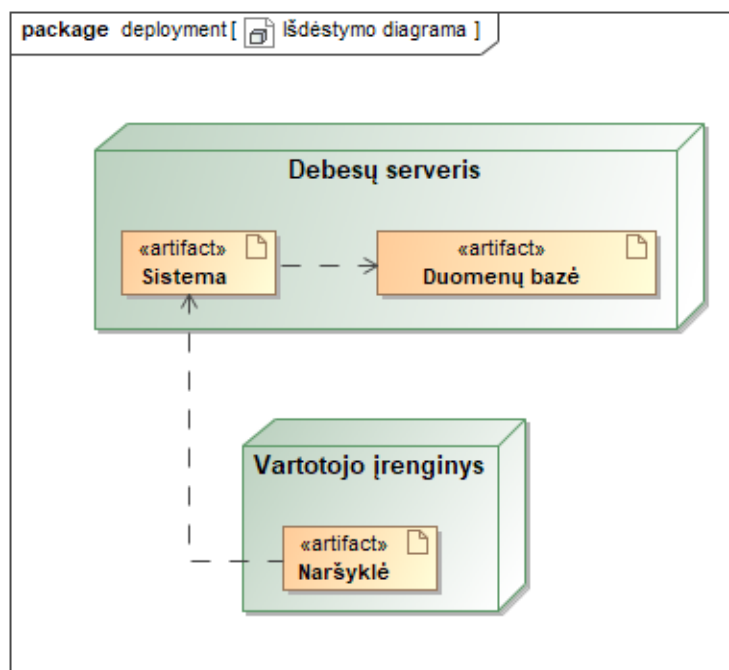
4-20 pav. PA „Redaguoti vartotoją“ veiklos diagrama



4-21 pav. PA „Peržiūrėti vartotojus“ veiklos diagrama

#### 4.3.5. Išdėstymo vaizdas

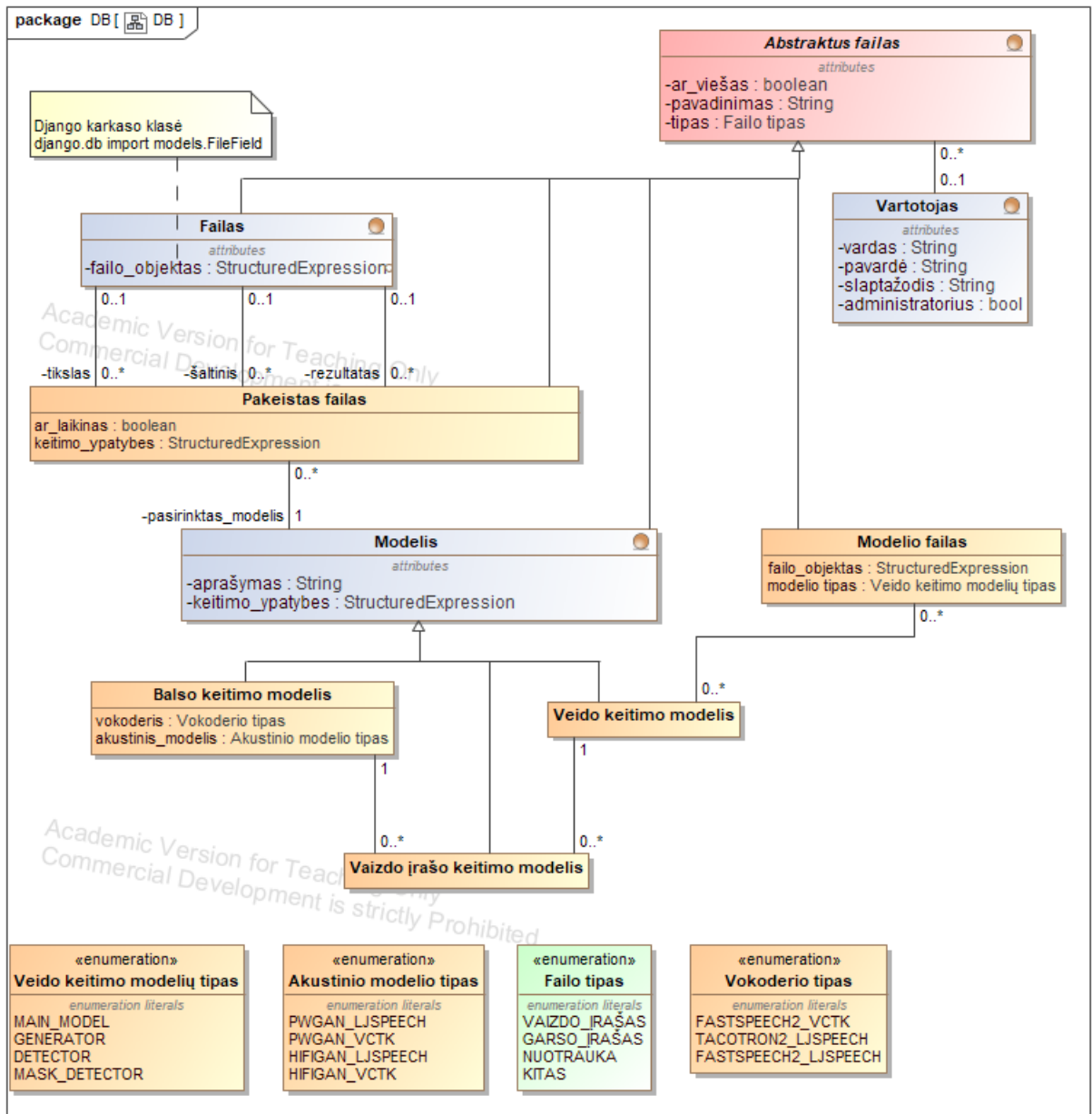
4-22 pav. pavaizduota sistemos išdėstymo diagrama. Kuriama sistema nebus pririšta prie specifinių paslaugų tiekėjų. Jei bus norima sistemą įdiegti į debesų serverį, tiek sukurta sistema, tiek kitų programuotojų iš naujo panaudota sistema minimaliai turės turėti pačios sistemos bei duomenų bazės artefaktus.



4-22 pav. Išdėstymo diagrama

### 4.3.6. Duomenų vaizdas

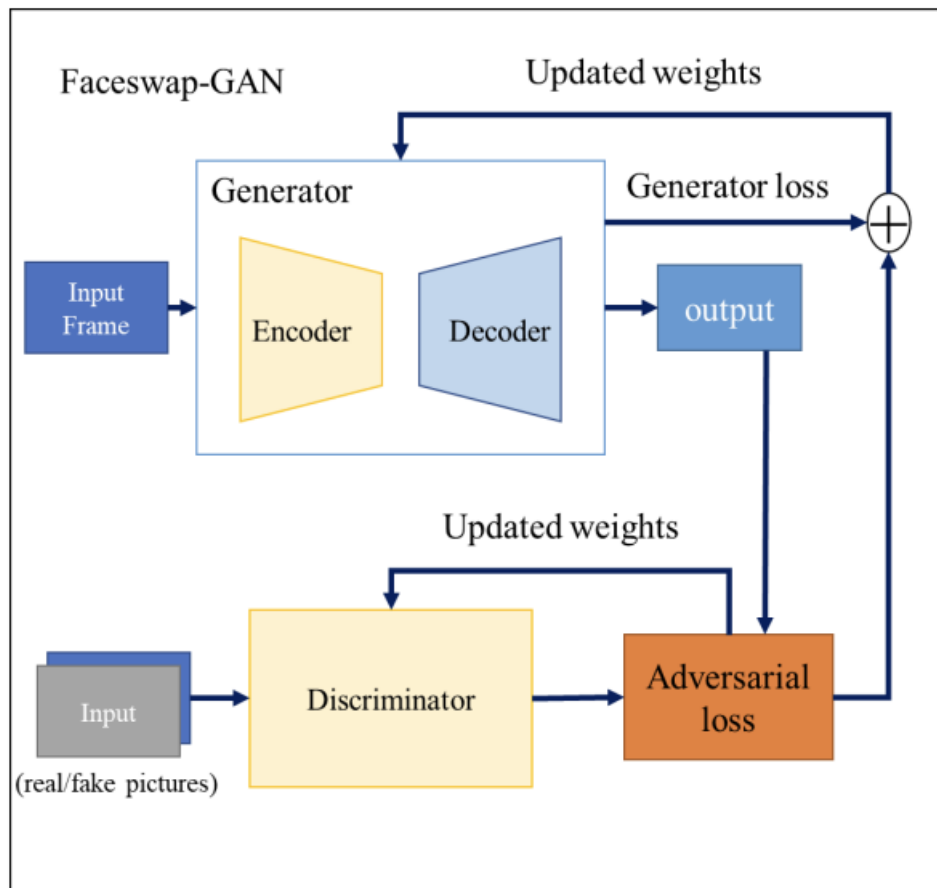
4-23 pav. pavaizduota klasių diagrama, kurioje *User* objektas yra numatytas *Django* karkaso vartotojas, kuris gali manipuluoti failais (*AbstractFile*). Ši klasė yra abstrakti ir ją paveldi konkrečios klasės *Model* ir *File*. Abi šios klasės turi atributą *object*, tačiau kiekvienai klasei jis yra skirtingas.



4-23 pav. Klasių diagrama

### 4.3.7. Modelių architektūros

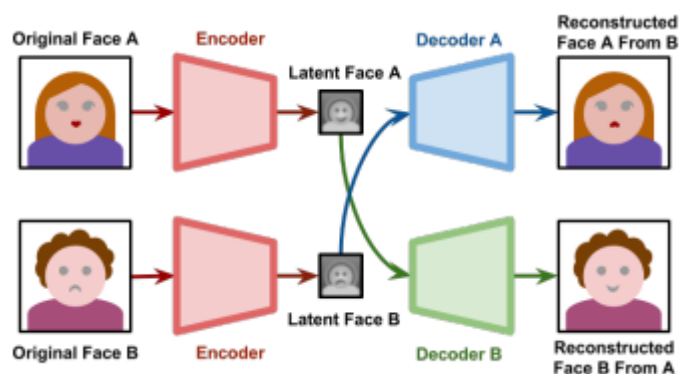
Veido ir balso keitimui naudojama GAN architektūra. GAN sudarytas iš generatoriaus ir diskriminatoriaus. Generatorius sukuria naujus duomenis, šiuo atveju iš gautos nuotraukos sugeneruoja naują. Tuo tarpu diskriminatorius lygina sugeneruotą nuotrauką su realia. Jei diskriminatorius atpažįsta sugeneruotą nuotrauką, ji yra nusiunčiama taisyti generatoriui tol, kol diskriminatorius neatpažįsta sugeneruotos nuotraukos.



4-24 pav. VAE architektūra [8]

Generatoriui naudojamas VAE architektūra, kurioje koduojamas išmoksta kad bruožai (šiuo atveju veido) egzistuoja bei juos atskiria ir sukuria latentinę erdvę, o dekoduoja iš latentinės erdvės sugeneruoja naują nuotrauką, kuri yra siunčiama diskriminatoriui:

- *CNN* (konvoliucinis neuroninis tinklas);
- *RNN* (rekurentinis neuroninis tinklas).



4-25 pav. GAN architektūra veido keitimui [9]

Nuotraukos (veido) keitimas:

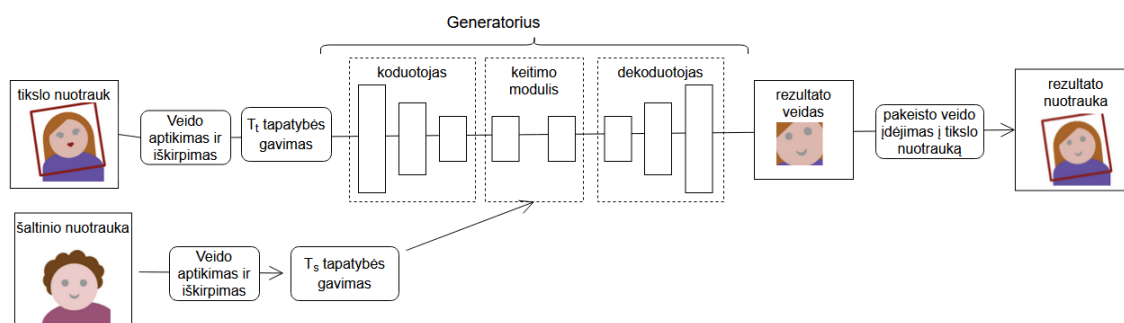
- Naudojama *PyTorch* [11] biblioteka
- Naudojami 4 modeliai (failai) veido keitimui:
  - a. Detektorius – tikrina, ar veidas egzistuoja duotoje nuotraukoje bei jį iškerpa, jei egzistuoja

- b. Atpažinėjas (angl. *recognitor*) – Iš iškirptos nuotraukos ištraukiamos veido savybės, kurios bus naudojamos keitimui
- c. Generatorius – sukuria naują veidą iš šaltinio veido savybių į tikslo veido išraišką
- d. Kaukės detektorius (angl. *mask detector*) – ištraukia veido kontūro kaukę veido keitimo metu šaltinio nuotraukoje (šiai galima ir bibliotekas naudoti, kai CV2), bet modelis tikslesnis kiek pastebėjau

Generatoriaus veikimas – atpažinimo modelis išgauna nuotraukoje esančio veido tapatybę ( $T_s$  – šaltinio arba  $T_t$  – tikslo).

Patį veido keitimą vykdo generatorius, kurį sudaro:

- Koduotojas (angl. *encoder*) – išgauna tikslo veido savybes  $Sav_t$
- Identifikacijos keitimo modulis – perkelia tapatybę, išgautą iš šaltinio nuotraukos  $T_s$ , į  $Sav_t$
- Dekoduotojas (angl. *decoder*) – Atkuria modifikuotą  $Sav_t$  (tikslo nuotraukos savybę) su  $T_s$  (šaltinio tapatybe) vaizdą.



4-26 pav. Veido keitimo detalesnė schema

Tyrinėjant veido keitimo modelių architektūras, bus tyrinėjamos ne tik pačios architektūros, bet ir šių hiperparametrų įtaka joms:

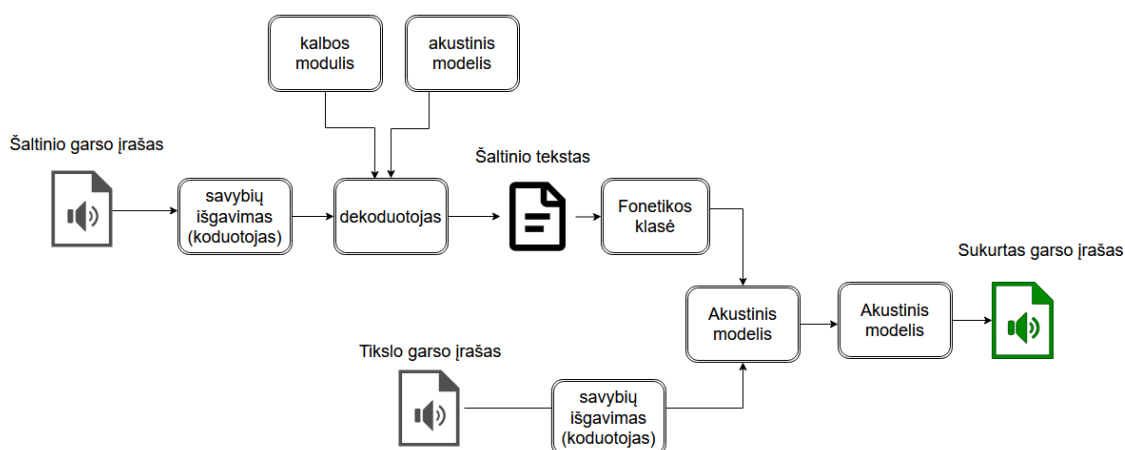
- AE bei GAN dimensijos;
- Atsitiktinis duomenų išmetimas
- Optimizatorius;
- Aktyvacijos ir netekties funkcijaos;
- GAN stiprumas;
- Spalvų ir šviesos intensyvumas;
- Veido bruožų prioritetai.

Garso (balso) keitimas:

- Naudojama *paddlespeech* [12] biblioteka, kuri balso keitimui gali naudoti jau egzistuojančius arba naujai sukurtus modelius:
- Naudojami modeliai balso keitimui projekte:
  - Kalba į tekstą:
    1. Transformuotojas – iš šaltinio garso paimamas tekstas;
  - Tekstas į kalbą
    1. Fonetikos klasė – tekstą konvertuoja į fonetinius duomenis;
    2. Koduotojas – ištraukia tikslo balso savybes;
    3. Akustinis modelis – šaltinio teksto savybes sugretina su tikslo garso savybėmis;



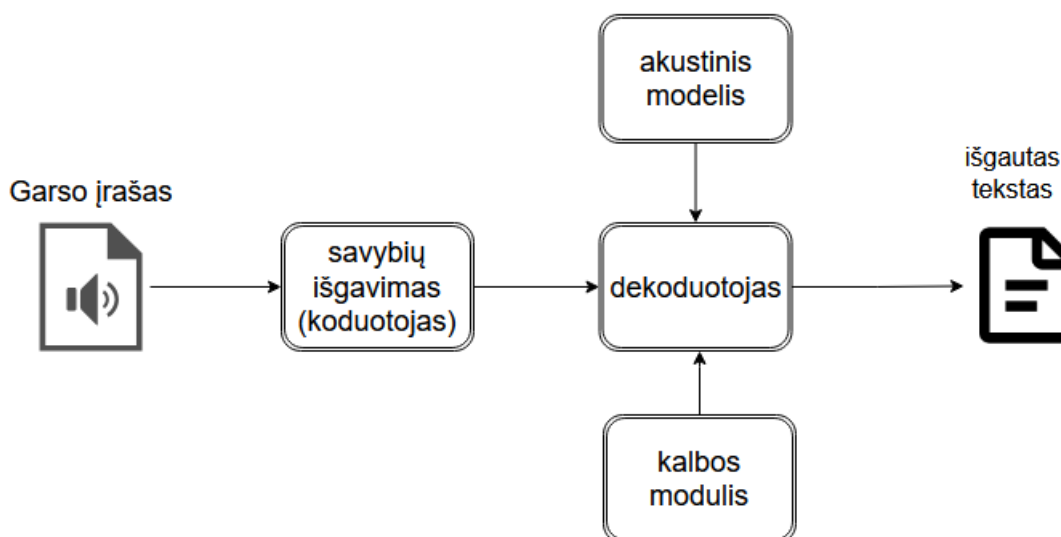
4. Vokoderis – teksto savybes konvertuota į balso įrašą pagal tikslo balso savybes.



4-27 pav. Balso keitimo procesas

Kalbos keitimas į tekstą:

- Koduotoją sudaro du 2D konvoliucijos pogrupio sluoksniai, kurie iš neapdoroto garso požymio išgauna požymio atvaizdavimą ir kartu sumažina garso požymio ilgį, ir keli vienas ant kito sudėti vienos krypties *RNN* sluoksniai, kurie garso požymius paverčia kadrais
- Dekoduotojas iš kodavimo įrenginio išvestą kiekvieną kadrą įveda į projektavimo sluoksnį, kuriame akustinis modelis kadrus paverčia fonemas. Projekcinio sluoksnio išėjimo dimensija yra tokia pati kaip žodyno dydis. Po projekcinio sluoksnio, kadrų lygmens požymiai paverčiami kalbos atpažinimo ženklais.

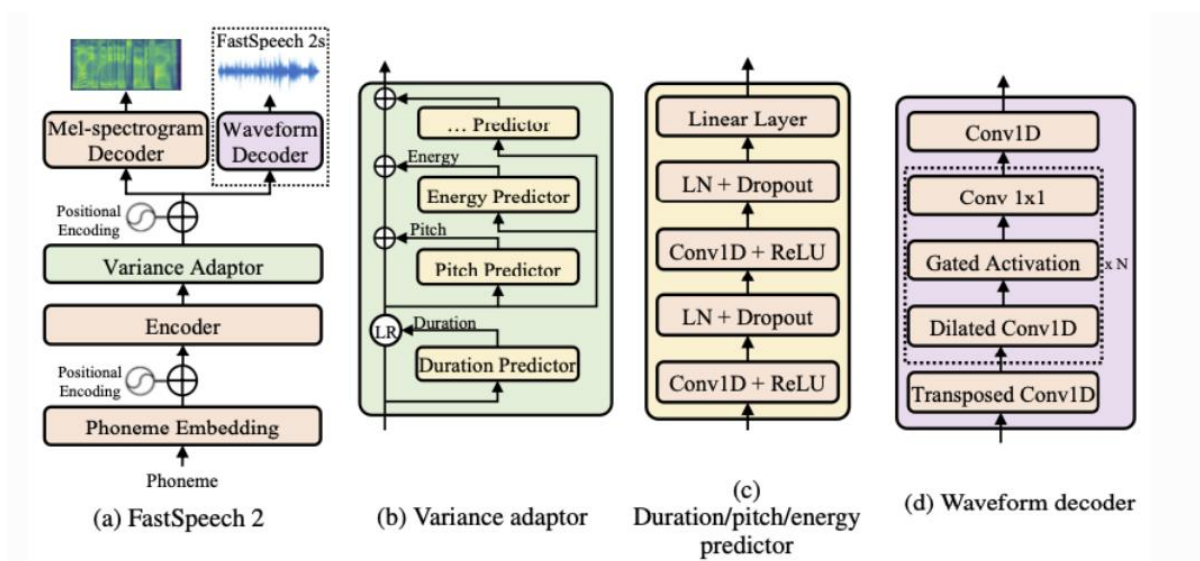


4-28 pav. Balso keitimo į tekstą schema

Teksto keitimas į kalbą:

- Fonemų įterpimo metu yra užfiksuojamos nustatoma garsų akustinė informacija (pvz., tarimo ypatybes) ir ji pateikiama fonemų vektoriais;
- Kodavimo įrenginys konvertuoja fonemų įterpimo seką į fonemų paslėptą seką;

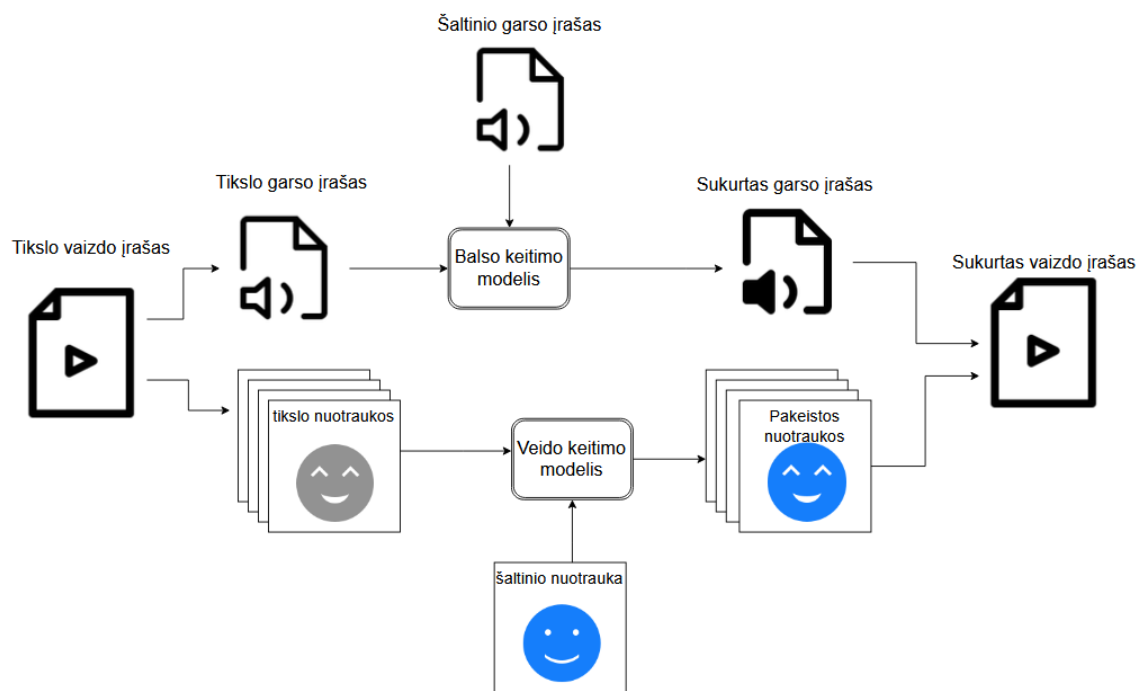
- Dispersijos adapteris į paslėptą seką prideda įvairią dispersijos informaciją, pavyzdžiui, trukmę, aukštį ir energiją;
- mel-spektrogramos dekoduoja lygiagrečiai konvertuoja adaptuotą paslėptą seką į mel-spektrogramos seką, kuri paverčiama garso įrašu.



4-29 pav. Teksto į kalbą vertimas [42]

Vaizdo įrašo (veido ir/ar balso) keitimas:

1. Atskiriamas tikslo vaizdo įrašo garso (jei egzistuoja);
2. Jei šaltinis vaizdo įrašas – paimamas pirmas kadras, kuriame yra žmogus (jei šaltinis nuotrauka – nieko nedaroma);
3. Keičiamas veidas vaizdo įrašė (jei šaltinis yra nuotrauka arba vaizdo įrašas):
  - a. kartojama per kiekvieną kadra;
  - b. kadre (nuotraukoje) pakeičiamas veidas (kaip paprastoje nuotraukoje);
  - c. naujas kadras (tikslo nuotrauka su šaltinio veidu) išsaugoma sąrašė;
  - d. po iteracijos kadrų sąrašas paverčiamas vaizdo įrašu;
4. Keičiamas balsas vaizdo įrašė (jei šaltinis vaizdo įrašo arba garso failas):
  - a. Jei tikslas ar šaltinis neturėjo balso – nieko nedarom
  - b. Jei yra tik tikslo balsas – jis pridamas prie sukurto vaizdo įrašo;
  - c. Jei yra ir šaltinio, ir tikslo balsas – šaltinio balsas pakeičiamas į tikslo (kaip balso keitime).



4-30 pav. Vaizdo įrašo keitimo procesas

Tyrinėjant balso keitimo modelių architektūras, bus tyrinėjamos ne tik pačios architektūros, bet ir šių parametų įtaka joms:

- Netekties funkcijos
- Optimizatorius;
- Aktyvacijos funkcija;
- Skirtingų modelių ir vokoderių deriniai.

#### 4.3.8. Kokybė

Sukurtos sistemos kokybė bus įgyvendinta:

- Sukurta sistema galės būti paprastai praplečiama norint pridėti naujo funkcionalumo dėl sukurtų klasių sąsajų;
- Bus sukurta *pip* biblioteka, kurioje bus aprašytas funkcionalumas bei kitiems programinės įrangos vartotojas leis naudotis sukurta sistema ir ją pritaikyti savo reikmėms;
- Sukurtos sistemos kodas bus analizuojamas statiniu kodo analizės įrankiu siekiant sumažinti kodo sudėtingumą ir galimų klaidų skaičių bei pagerinti skaitomumą.

### 4.4. Testavimas

#### 4.4.1. Testavimo planas

Šis skyrius yra skirtas specifikuoti kuriamo projekto „Žmogaus veido ir balso pakeitimo kitu asmeniu sistemos“ testavimą. Skyrius pateikia testavimo planą, bendrą testavimo specifikaciją bei apibrėžia testavimo procedūras. Tolimesni sistemos testavimo darbai bus atliekami remiantis šiuo skyriumi, kuris padeda nustatyti sistemos reikalavimų išbaigtumą, teisingumą.

#### 4.4.1.1. Testavimo elementai

Sistemą sudarys tik vienas artefaktas, kuriame bus įgyvendintas pagrindinė kuriamos sistemos logika, API bei grafinė vartotojo sąsaja

#### 4.4.1.2. Programinės įrangos rizikos klausimai

Toliau bus pateiktos kritinės bei svarbios vietos, į kurias būtinas atkreipti dėmesį testavimo metu:

- naujos bibliotekų versijos bei senų versijų nepalaikomumas;
- įvairūs failų bei modelių formatai;
- sugedus ar kitaip sutrikus sklandžiam sensoriaus veikimui naudotojas turi būti apie tai informuotas.

#### 4.4.1.3. Testuotinos funkcijos

Toliau bus pateiktas sąrašas funkcijų, kurios bus atliekamos testavimo metu, iš naudotojo perspektyvos kartu su rizikos lygiu:

- prisijungimas, registracija bei atsijungimas (**aukšta**);
- failų įkėlimas bei keitimas grafinėje vartotojo sąsajoje (**aukšta**);
- modelių įkėlimas bei keitimas grafinėje vartotojo sąsajoje (**vidutinė**);
- failų įkėlimas bei keitimas per API (**vidutinė**);
- modelių įkėlimas bei keitimas API(**vidutinė**);
- vartotojų valdymas (**žema**).

### 4.4.2. Strategija

#### 4.4.2.1. Statinė kodo analizė

Sistemos kūrimo metu bus naudojamas *Black* [14] statinis kodo analizavimo įrankis. Programinės įrangos kūrimo metu naudojantis statinio kodo analizavimo įrankiu bus siekiama parašyti kuo labiau skaitomesnį kodą bei išvengti nenaudojamo ar blogai naudojamo kodo fragmentų

#### 4.4.2.2. Automatinis testavimas

Kuriama sistema bus testuojama naudojantis *pytest* [15] testavimo karkasu, kuriuo galima ištestuoti didžiąją dalį sukurtos programinės įrangos.

#### 4.4.2.3. Komponentų testavimas

Vienetų testai bus parašyti sistemos kūrimo metu sistemos kūrėjo toms failų keitimo funkcijoms:

- Nuotraukos keitimas;
- Garso įrašo keitimas;
- Vaizdo įrašo keitimas.

Šio testavimo tikslas yra patikrinti, ar šios funkcijos sugeba:

- Priimti reikalavimuose įvardintų formatų failus;
- Sėkmingai pakeisti gero formato failą;
- Gavus netinkamo formato failą nenuklūžti bei grąžinti tinkamą klaidos pranešimą.

#### 4.4.2.4. Integracinis testavimas

Kadangi kuriant sistemą yra naudojamas *Django* karkasas bei jo administravimo funkcionalumas, kurio tik dalis yra rašoma sistemos kūrėjo, o kita dalis paimta iš *Django* karkaso, todėl integracinio testavimo metu bus ištestuoti API endpointai tuo pačiu ištestuojant ir sistemos kūrėjo parašytą funkcionalumą. Testai bus parašyti failų bei modelių įkėlimo, keitimo, pasiėmimo bei šalinimo funkcijoms. Šių testų metu bus siekiama užtikrinti, kad sąveika tarp skirtingų sistemos komponentų (programėlės, serverio ir sensoriaus) veikia taip, kaip ir tikėtasi.

#### 4.4.2.5. Rankinis testavimas

Rankinio testavimo metu bus ištestuotas sukurtas funkcionalumas grafinėje vartotojo sąsajoje siekiant patikrinti funkcionalumą, kurio nepavyko ištestuoti automatinio testavimo metu bei patikrinti ar vartotojas sugebės patogiai naudotis sistema. Rankinio testavimo atvejai suskirstyti į grupes:

- Registracija bei prisijungimas (4-45 lentelė);
- Modelių valdymas (4-46 lentelė);
- Failų valdymas (4-47 lentelė);
- Vartotojų valdymas (4-48 lentelė).

##### 4-45 lentelė. Registracija bei prisijungimas

Atvejo nr	Aprašymas	Tikėtinas rezultatas
1.1.	Vartotojas bando registruotis su naujais korektiškais duomenimis	Registracija įvykdoma
1.2.	Vartotojas bando registruotis su jau užregistruotais, bet nekorektiškais duomenimis	Atitinkamas klaidos pranešimas
1.3.	Vartotojas bando registruotis su nekorektiškais	Atitinkamas klaidos pranešimas
1.4.	Vartotojas bando prisijungti su korektiškais duomenimis	Prisijungimas įvykdomas
1.5.	Vartotojas bando prisijungti su nekorektiškais duomenimis	Atitinkamas klaidos pranešimas

##### 4-46 lentelė. Modelio valdymas

Atvejo nr.	Aprašymas	Tikėtinas rezultatas
2.1.	Vartotojas bando įkelti korektišką modelį	Veiksmas įvykdomas
2.2.	Vartotojas bando įkelti nekorektišką modelį	Atitinkamas klaidos pranešimas
2.3.	Vartotojas bando įkelti pakeisti modelio failą/informaciją į korektišką	Veiksmas įvykdomas
2.4.	Vartotojas bando įkelti pakeisti modelio failą/informaciją į nekorektišką	Atitinkamas klaidos pranešimas
2.5.	Vartotojas, turintis teises, bando ištrinti modelį	Veiksmas įvykdomas
2.6.	Vartotojas, neturintis teisių, bando ištrinti modelį	Atitinkamas klaidos pranešimas
2.7.	Vartotojas bando pasiimti modelio informaciją	Veiksmas įvykdomas
2.8.	Vartotojas bando pasiimti kelių/visų modelių informaciją	Veiksmas įvykdomas

##### 4-47 lentelė. Failų valdymas

Atvejo nr.	Aprašymas	Tikėtinas rezultatas
3.1.	Vartotojas bando įkelti korektišką failą	Veiksmas įvykdomas
3.2.	Vartotojas bando įkelti nekorektišką failą	Atitinkamas klaidos pranešimas
3.3.	Vartotojas bando įkelti pakeisti failą/informaciją į korektišką	Veiksmas įvykdomas

3.4.	Vartotojas bando įkelti pakeisti failą/informaciją į nekorektišką	Atitinkamas klaidos pranešimas
3.5.	Vartotojas, turintis teises, bando ištrinti failą	Veiksmas įvykdomas
3.6.	Vartotojas, neturintis teisių, bando ištrinti failą	Atitinkamas klaidos pranešimas
3.7.	Vartotojas bando pasiimti failo informaciją	Veiksmas įvykdomas
3.8.	Vartotojas bando pasiimti kelių/visų failų informaciją	Veiksmas įvykdomas
3.9.	Vartotojas bando pakeisti failą su korektiškais duomenimis	Veiksmas įvykdomas
3.10.	Vartotojas bando pakeisti failą su nekorektiškais duomenimis	Atitinkamas klaidos pranešimas

#### 4-48 lentelė. Vartotojų valdymas

Atvejo nr.	Aprašymas	Tikėtinas rezultatas
4.1.	Administratorius bando peržiūrėti kelis/visus vartotojus	Veiksmas įvykdomas
4.2.	Administratorius bando peržiūrėti vartotojo informaciją	Veiksmas įvykdomas
4.3.	Vartotojas bando įkelti pakeisti vartotojo į korektišką	Veiksmas įvykdomas
4.4.	Vartotojas bando įkelti pakeisti vartotojo į nekorektišką	Atitinkamas klaidos pranešimas
4.5.	Vartotojas, turintis teises, bando ištrinti vartotoją	Veiksmas įvykdomas
4.6.	Vartotojas, neturintis teisių, bando ištrinti vartotoją	Atitinkamas klaidos pranešimas

#### 4.4.2.6. Priėmimo testavimas

Priėmimo testavimo metu bus tikrinama, ar visi funkciniai bei nefunkciniai reikalavimai yra įgyvendinti sukurtoje sistemoje. Šio testavimo metu peržiūrima sukurta sistema bei radus neatitiktumus su aprašytais reikalavimais sistema keičiama tol, kol atitinka užsibrėžtus reikalavimus.

#### 4.4.2.7. Elemento patenkinimo / nepatenkinimo kriterijai

Automatinių testų pabaigimo kriterijus bus pasiektas tada, kai programinis kodas pasieks 80 procentų kodo padengimo bei visi testai bus sėkmingi.

Rankinio testavimo pabaigimo kriterijus bus pasiektas tada, kai visi testavimo atvejai veiks kai aprašyti 5.3 skyriuje.

Priėmimo testavimo pabaigimo kriterijus bus pasiektas tada, kai bus patikrinta, kad visi funkciniai bei nefunkciniai reikalavimai yra įgyvendinti sukurtoje sistemoje.

#### 4.4.2.8. Bandymų rezultatai

Testavimo pabaigoje rezultatai turi būti išsaugoti bei pateikti kitame skyrelyje.

#### 4.4.2.9. Tvarkaraštis

Toliau bus pateiktas šio testavimo plano tvarkaraštis:

#### 4-49 lentelė. Veiklų terminai

Veikla	Terminas
Testavimo plano parengimas	2022 m. gegužės 31 d.
Automatinis testavimas	2022 m. spalio 1 d.
Rankinis testavimas	2022 m. lapkričio 1 d.
Priėmimo bei streso testavimas	2022 m. lapkričio 15 d.

#### 4.4.2.10. Rizikos ir nenumatytų atvejų planavimas

Rizikų analizė aprašyta 4.2.21 *Rizikų įvertinimas* skyrelyje.

#### 4.4.3. Testavimo rezultatai ir išvados

##### 4.4.3.1. Automatinio testavimo rezultatai

Automatinio testavimo metu parašyti 43 testai, kurie padengia 85% testuojamo kodo.

##### 3.6.1.1. Rankinio testavimo rezultatai

Atliktas rankinis testavimas pagal sudarytą planą be jo rezultatai aprašyti lentelėje

4-50 lentelė. Rankinio testavimo rezultatai

Atvejo nr	Aprašymas	Tikėtinas rezultatas	Gautas rezultatas
1.1.	Vartotojas bando registruotis su naujais korektiškais duomenimis	Registracija įvykdoma	Rezultatas tenkinamas
1.2.	Vartotojas bando registruotis su jau užregistruotais, bet nekorektiškais duomenimis	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
1.3.	Vartotojas bando registruotis su nekorektiškais	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
1.4.	Vartotojas bando prisijungti su korektiškais duomenimis	Prisijungimas įvykdomas	Rezultatas tenkinamas
1.5.	Vartotojas bando prisijungti su nekorektiškais duomenimis	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
2.1.	Vartotojas bando įkelti korektišką modelį	Veiksmas įvykdomas	Rezultatas tenkinamas
2.2.	Vartotojas bando įkelti nekorektišką modelį	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
2.3.	Vartotojas bando įkelti pakeisti modelio failą/informaciją į korektišką	Veiksmas įvykdomas	Rezultatas tenkinamas
2.4.	Vartotojas bando įkelti pakeisti modelio failą/informaciją į nekorektišką	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
2.5.	Vartotojas, turintis teises, bando ištrinti modelį	Veiksmas įvykdomas	Rezultatas tenkinamas
2.6.	Vartotojas, neturintis teisių, bando ištrinti modelį	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
2.7.	Vartotojas bando pasiimti modelio informaciją	Veiksmas įvykdomas	Rezultatas tenkinamas
2.8.	Vartotojas bando pasiimti kelių/visų modelių informaciją	Veiksmas įvykdomas	Rezultatas tenkinamas
3.1.	Vartotojas bando įkelti korektišką failą	Veiksmas įvykdomas	Rezultatas tenkinamas
3.2.	Vartotojas bando įkelti nekorektišką failą	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
3.3.	Vartotojas bando įkelti pakeisti failą/informaciją į korektišką	Veiksmas įvykdomas	Rezultatas tenkinamas
3.4.	Vartotojas bando įkelti pakeisti failą/informaciją į nekorektišką	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
3.5.	Vartotojas, turintis teises, bando ištrinti failą	Veiksmas įvykdomas	Rezultatas tenkinamas

3.6.	Vartotojas, neturintis teisių, bando ištrinti failą	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
3.7.	Vartotojas bando pasiimti failo informaciją	Veiksmas įvykdomas	Rezultatas tenkinamas
3.8.	Vartotojas bando pasiimti kelių/visų failų informaciją	Veiksmas įvykdomas	Rezultatas tenkinamas
3.9.	Vartotojas bando pakeisti failą su korektiškais duomenimis	Veiksmas įvykdomas	Rezultatas tenkinamas
3.10.	Vartotojas bando pakeisti failą su nekorektiškais duomenimis	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
4.1.	Administratorius bando peržiūrėti kelis/visus vartotojus	Veiksmas įvykdomas	Rezultatas tenkinamas
4.2.	Administratorius bando peržiūrėti vartotojo informaciją	Veiksmas įvykdomas	Rezultatas tenkinamas
4.3.	Vartotojas bando įkelti pakeisti vartotojo į korektišką	Veiksmas įvykdomas	Rezultatas tenkinamas
4.4.	Vartotojas bando įkelti pakeisti vartotojo į nekorektišką	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas
4.5.	Vartotojas, turintis teises, bando ištrinti vartotoją	Veiksmas įvykdomas	Rezultatas tenkinamas
4.6.	Vartotojas, neturintis teisių, bando ištrinti vartotoją	Atitinkamas klaidos pranešimas	Rezultatas tenkinamas

#### 4.4.4. Testavimo išvados

Atliekant testavimą buvo laikomasi nustatytų metodikų. Projektavimo metu parengtas testavimo planas. Dauguma klaidų ir neatitikimų buvo pašalintos programavimo etape. Rankinio testavimo metu visi aprašyti testavimo atvejai įgyvendinti, o automatiniam testavimui parašyti 43 testai, kurie padengia 86% testuojamo kodo.

#### 4.5. Vartotojo dokumentacija

##### 4.5.1. Sistemos funkcinis aprašymas

Skirta žmonėms ir įmonėms, kurie nori pakeisti asmens balsą ir veidą, esančius nuotraukose ir garso įrašuose, kitų asmeniu. Sistema gali priimti mašininio mokymo modelius, bei juos naudojant iš originalo failo perkelti balsą/veidą į kitą norimą failą. Skirtingai nei rinkoje esančios konkurencinės sistemos, ši sistema galės kokybiškai atlikti veido keitimą nuotraukoje bei balso keitimą garso įrašuose, bei taip pat naudoti bei dalinti savo sukurtais modeliais

#### Pagrindinės funkcijos

Trumpai apibūdinant kuriamos sistemos veikimą, sistema bus skirta tyrimams su neuroninių tinklų modeliais atlikti bei pakeisti duotą veidą/balsą kitu. Sistema skirta trijų tipų naudotojams: administratoriams, prisijungusiems bei neprisijungusiems vartotojams:

1. Svečias
  - a. prisijungti
  - b. užsiregistruoti,
  - c. įkelti paveiksluko ar garso įrašo failą bei jį bei atsisiųsti naujai sugeneruotą failą
2. Prisijungęs vartotojas:



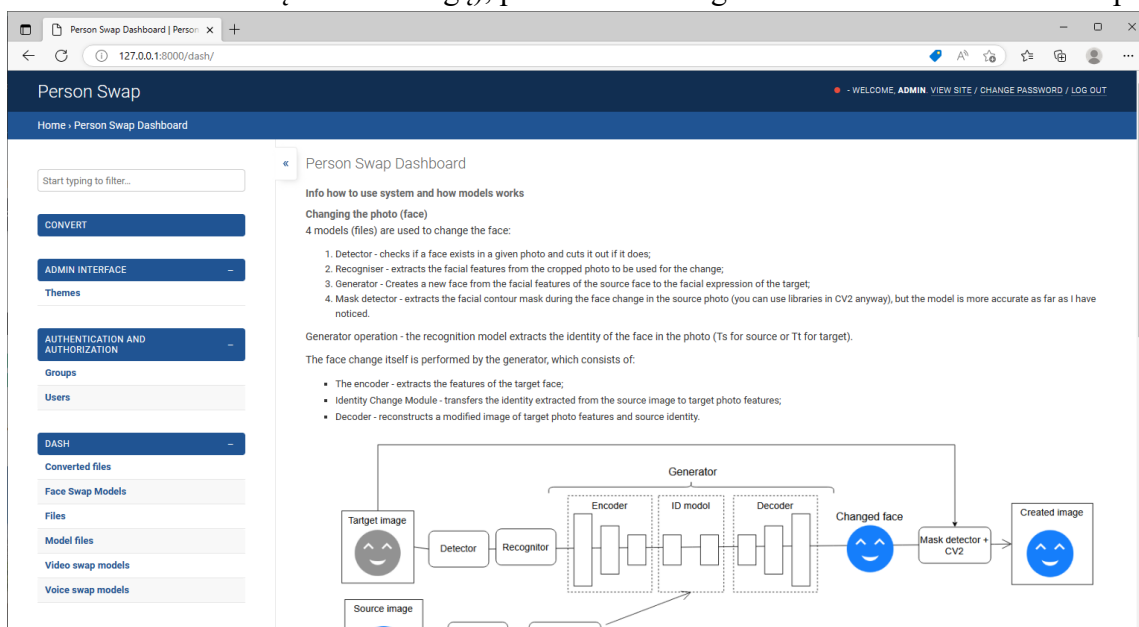
- a. įkelti vaizdo ar garso įrašo ar paveiksluko failą bei jį bei atsisiųsti ar išsaugoti naujai sugeneruotą failą pasirinkus norimą modelį;
  - b. matyti bendrinius bei savo įkeltus modelius, įkelti naują modelį ar ištrinti anksčiau savo įkeltą modelį;
3. Administratorius
    - a. matyti visus modelius, įkelti naują modelį ar ištrinti esamą modelį;
    - b. matyti visus vartotojus, keisti jų duomenis, ištrinti esamą ar pridėti naują vartotoją.

## 4.5.2. Naudotojo dokumentacija

Šiame skyriuje pateikiami pagrindiniai veiksmai, kuriuos gali atlikti vartotojas besinaudojantis sistema.

### 4.5.2.1. Pagrindinis sistemos langas

Pagrindiniame sistemos lange matoma glaustas sistemos aprašymas, meniu (kurio pirmas mygtukas „Convert“ iš karto nukelia į keitimo langą), pradedantis žmogaus veido ir/arba also keitimo procesą.



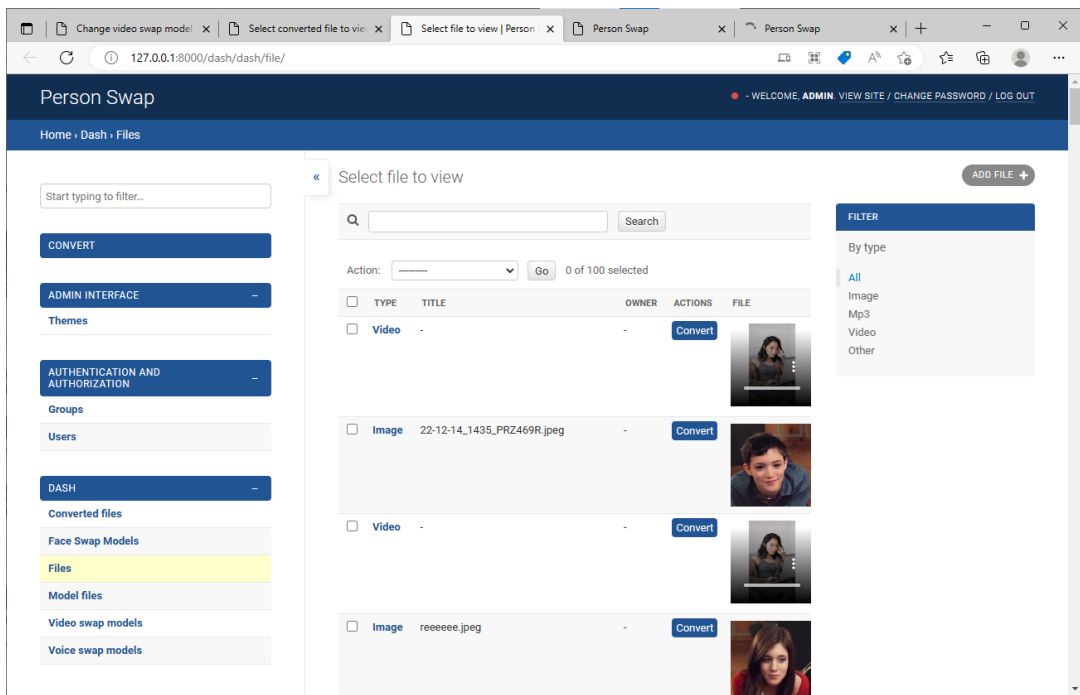
4-31pav. Pagrindinis sistemos langas

### 4.5.2.2. Failų sąrašai

Failų sąrašo lange (4-33 pav.) matome visi egzistuojantys failai. Norint vykdyti keitimo procesą, prie norimo failo reikia paspausti keitimo mygtuką (daugiau informacijos sekančiame skyriuje 4.5.2.3 Failo keitimas).

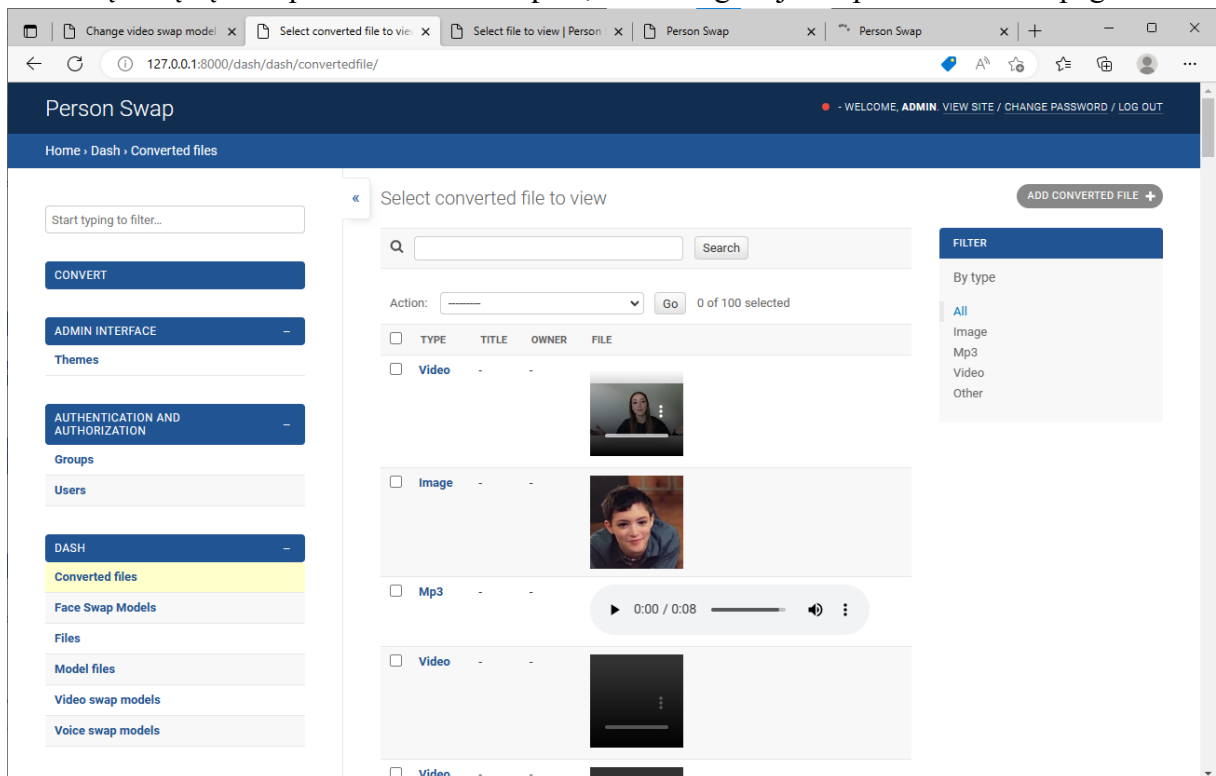
Galimi modelių filtrai:

- Pagal tipą:
  - Nuotraukos;
  - Garso įrašai;
  - Vaizdo įrašai;
  - Kiti.



4-32 pav. Failų sąrašo langas

Pakeistų failų sąrašas pavaizduotas 4-33 pav., kuriame galioja tas pats filtravimas pagal failo tipą



4-33 pav. Pakeistų failų langas

#### 4.5.2.3. Failo keitimas

Žmogaus veidą ir/ar balsą galima pakeisti tiek iš jau sistemoje egzistuojančių failų, tiek įkeliant failus iš įrenginio.

Keitimo procesas su jau pasirinktu failu:

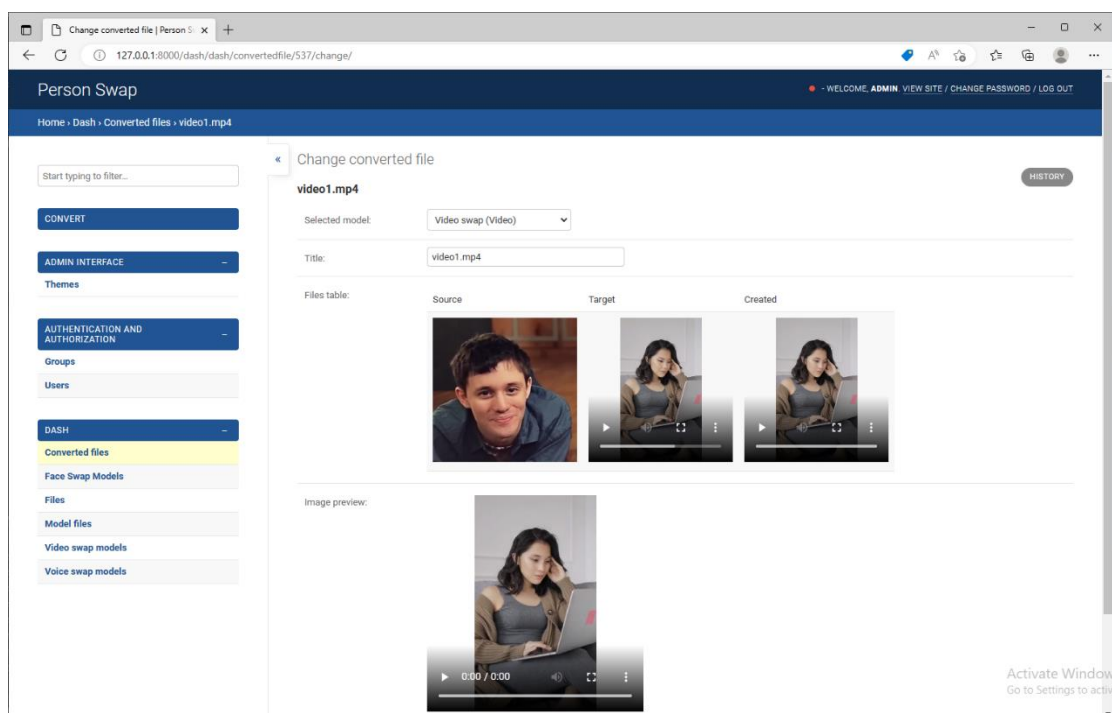
1. Prie norimo failo paspaudžiamas „Convert“ mygtukas

2. Pasirenkamas keitimo modelis
3. Pasirenkamas pasirinkto failo tipas
  - a. „*source*“ – pasirinktas failas naudojamas kaip šaltinis keitimo procese;
  - b. „*target*“ – pasirinktas failas naudojamas kaip tikslas keitimo procese;
4. Pasirenkamas likęs failo tipas iš egzistuojančių sistemoje failų arba įrenginio (šaltinio arba tikslo failas)
5. Paspaužiamas keitimo mygtukas (keitimas gali užtrukti, todėl negalima uždaryti sistemos lango)
6. Nustojus vykdyti keitimo procesą ekrane parodomi šaltinio, tikslo ir sukurtas failai.

Keitimo procesas be pasirinkto failo:

1. Pagrindiniame lange arba meniu skiltyje paspaudžiamas „*Convert*“ mygtukas
2. Pasirenkamas keitimo modelis
3. Pasirenkamas šaltinio ir tikslo failai
4. Paspaužiamas keitimo mygtukas (keitimas gali užtrukti, todėl negalima uždaryti sistemos lango)
5. Nustojus vykdyti keitimo procesą ekrane parodomi šaltinio, tikslo ir sukurtas failai.

Keitimo proceso rezultatas pavaizduotas 4-34 pav.



4-34 pav. Keitimo proceso rezultatas

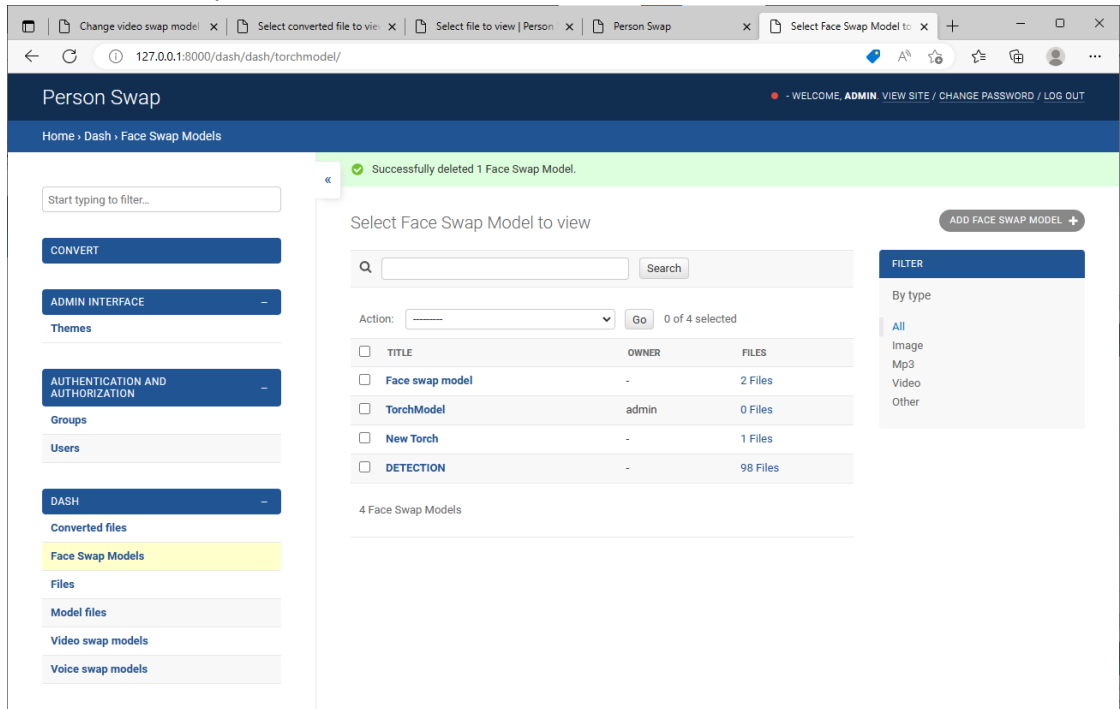
#### 4.5.2.4. Modelių langas

Modelių sąrašo lange (4.35 pav.) matome visi egzistuojantys modeliai bei failų, kurie buvo keičiami naudojantis konkrečiu modeliu, kiekis ir nuoroda į juos.

Galimi modelių filtrai:

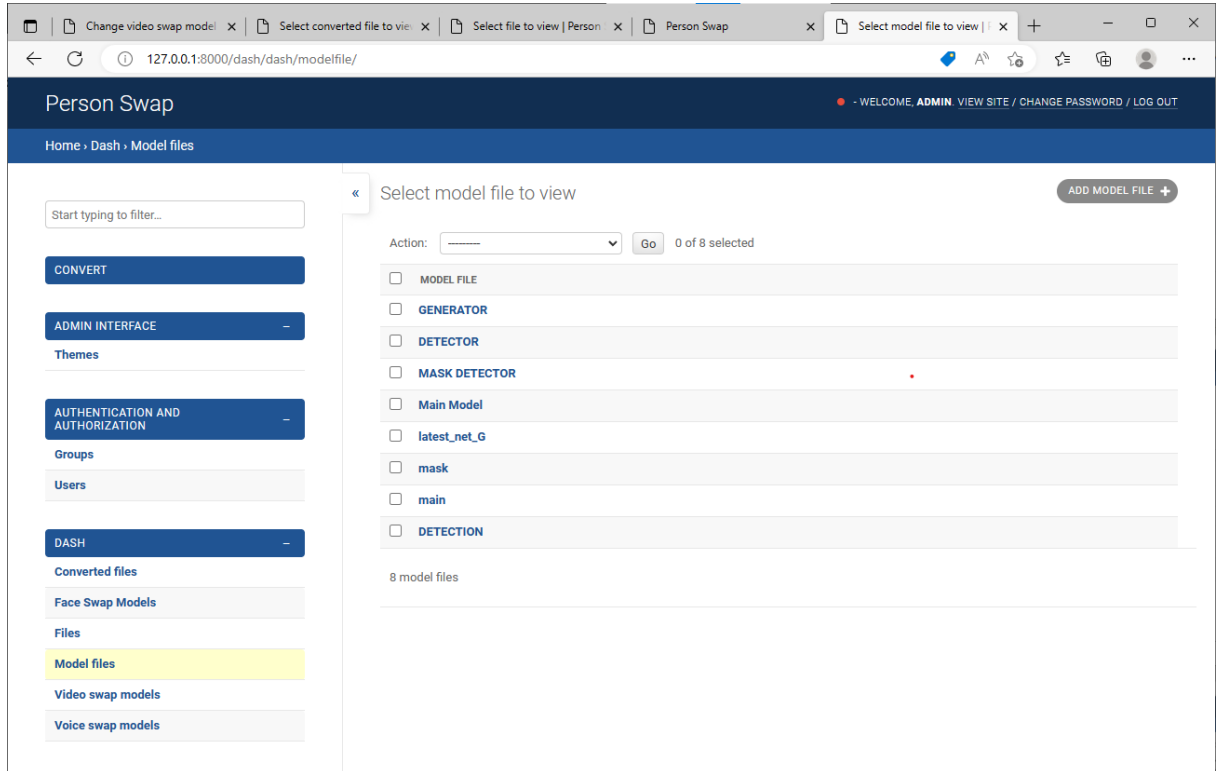
- Pagal tipą:
  - Nuotraukos;

- Garso įrašai;
- Vaizdo įrašai;



4-35 pav. Modelių sąrašo langas

Veido keitimo modeliai naudoja modelių keitimo failus, kurių sąrašas pavaizduotas 4-36 pav.

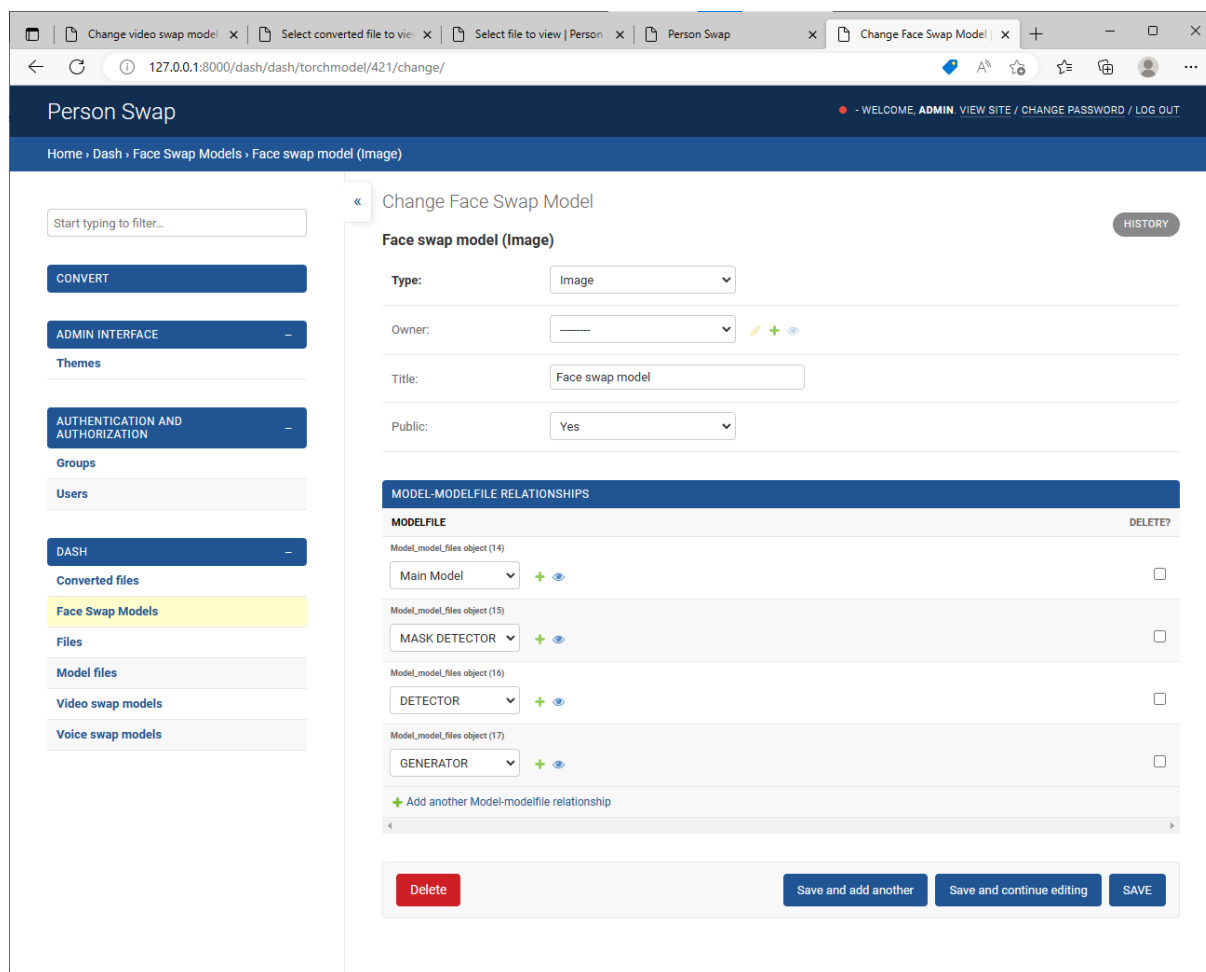


4-36 pav. Modelių failų sąrašas

#### 4.5.2.5. Modelio kūrimas

Modelių kūrimo bei keitimo langai atrodo vienodai. 4-37 pav. pavaizduotas veido modelio kūrimo/keitimo langas. Šis modelis taip pat naudoja modelio failus. Modelio failai gali būti paimami iš:

- [https://drive.google.com/drive/folders/1jV6\\_0FIMPC53FZ2HzZNJZGMe55bbu17R](https://drive.google.com/drive/folders/1jV6_0FIMPC53FZ2HzZNJZGMe55bbu17R) ( autoriai: Renwang Chen, Xuanhong Chen, Bingbing Ni, Yanhao Ge)
- <https://gib.com/deepinsight/insightface/blob/6562e324a0f13b065560a558f123a03f50ab500c/detection/retinaface/README.md> ( autoriai: Deng, Jiankang, Guo, Jia, Liu, Tongliang, Gong, Mingming, Zafeiriou, Stefanos)

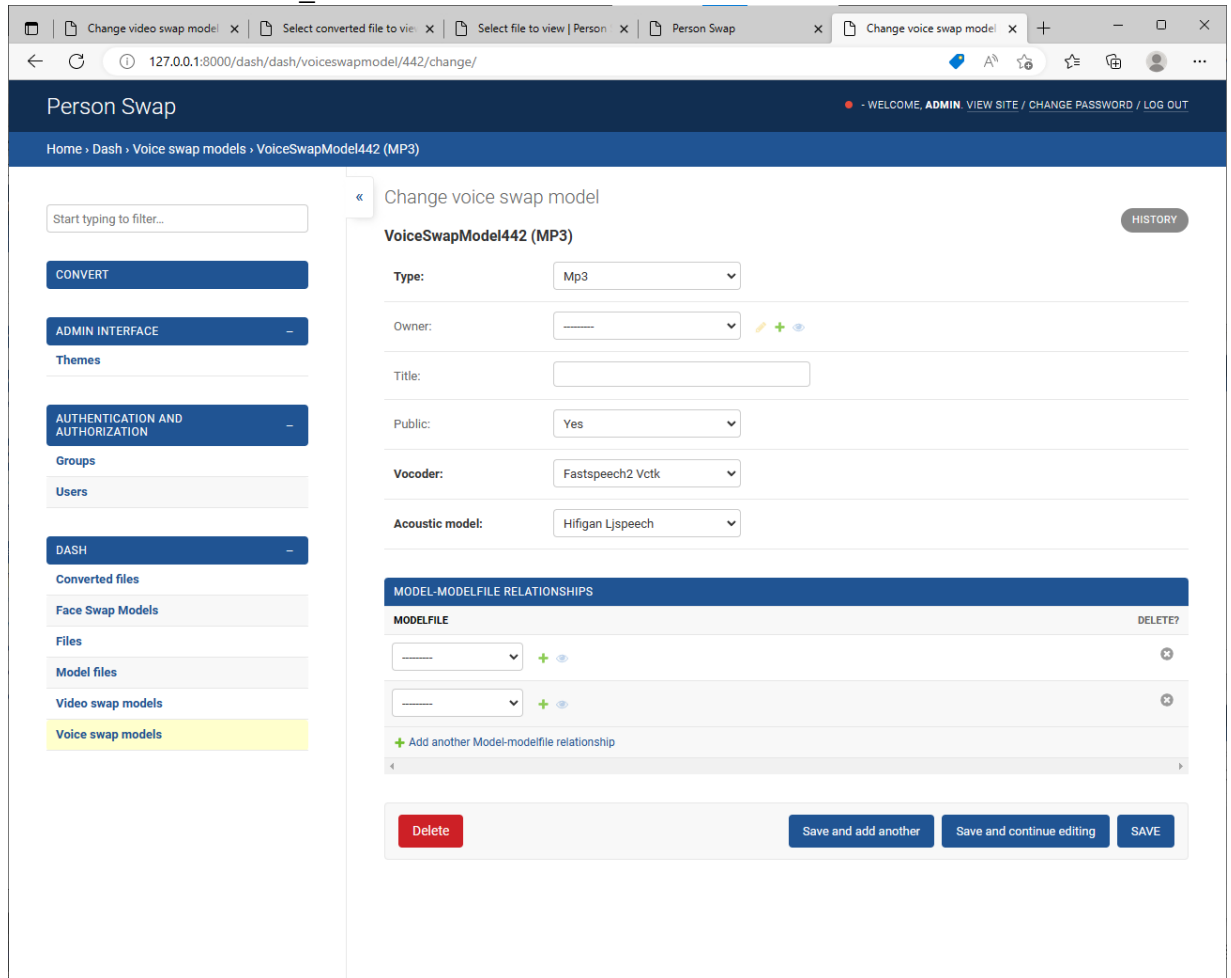


4-37 pav. Veido keitimo modelio langas

4-38 pav. pavaizduotas balsų keitimo modelio kūrimo/keitimo langas. Jis, kitaip nei veido keitimo modelis, prašo pasirinkti vokoderį bei akustinį modelius, kurie paimami iš *PaddleSpeech* [10] bibliotekos ir gali būti:

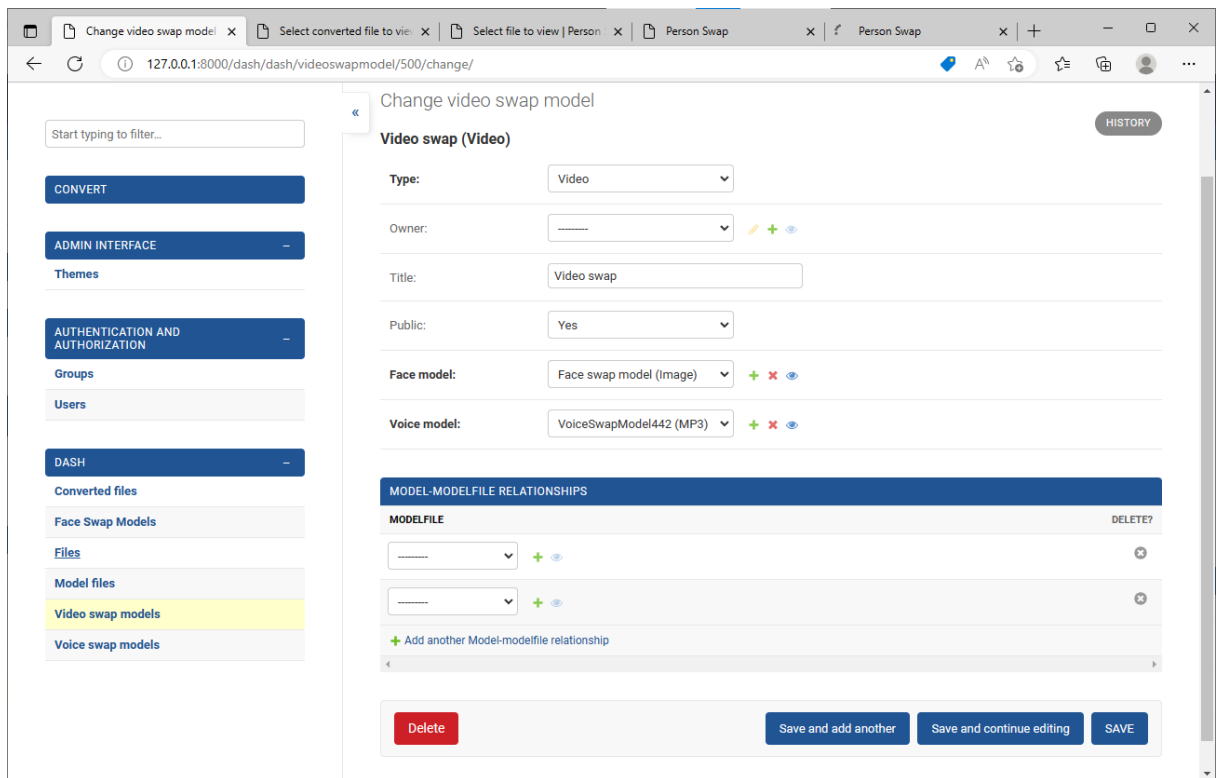
- Vokoderiai
  - FASTSPEECH2\_VCTK
  - TACOTRON2\_LJSPEECH
  - FASTSPEECH2\_LJSPEECH
- Akustiniai modeliai
  - PWGAN\_LJSPEECH
  - PWGAN\_VCTK

- HIFIGAN\_LJSPEECH
- HIFIGAN\_VCTK



4-38 pav. Balso keitimo modelio langas

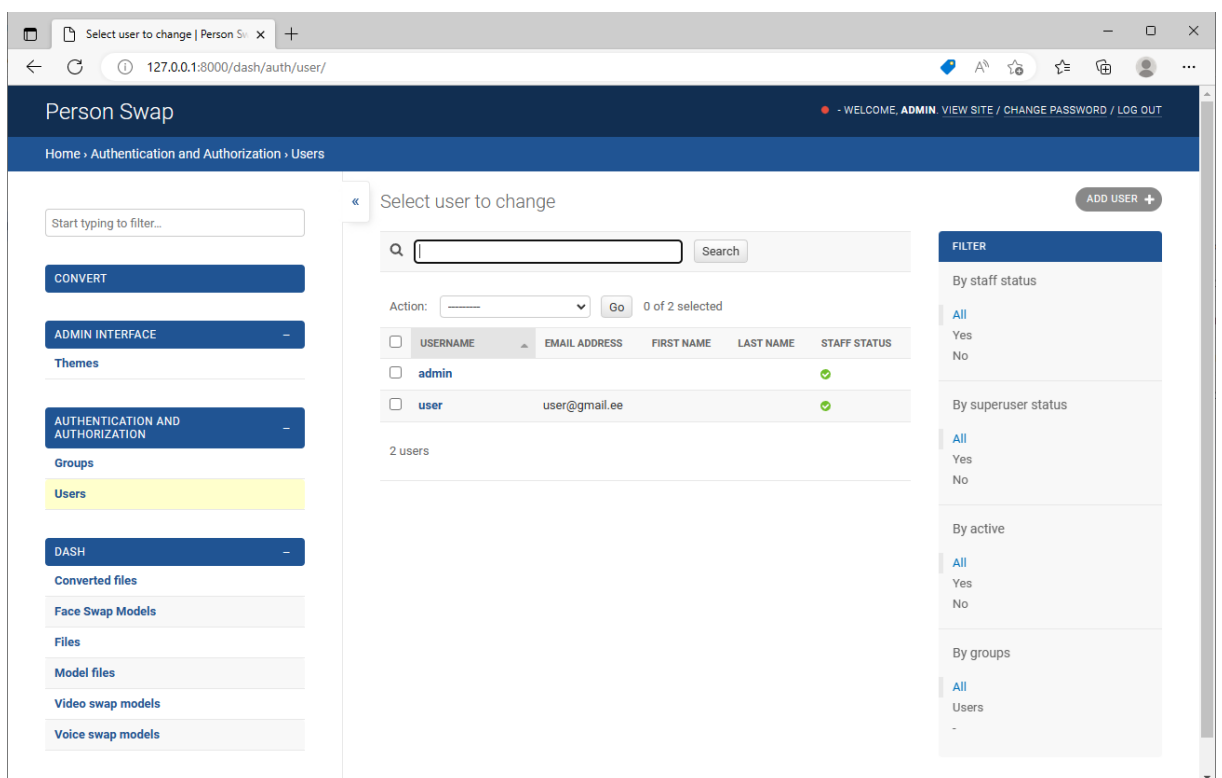
4-39 pav. pavaizduotas vaizdo įrašo kūrimo/keitimo langas, kuriame kitaip nei prieš tai aprašytuose kūrimo/keitimo languose, reikia pasirinkti veido ir balso keitimo modelius.



4-39 pav. Vaizdo įrašo modelio keitimo langas

#### 4.5.2.6. Administratoriaus langas

4-40 pav. pavaizduotas vartotojų valdymo langas.



4-40 pav. Administratoriaus langas

### 4.5.3. Sistemos diegimas ir paleidimas

Norint instaliuoti bei paleisti sistemą reikia atlikti šiuos veiksmus:

1. Sukurti virtualią aplinką, kuri naudotų *Python 3.8* versiją bei ją aktyvuoti;
2. Paleisti instaliavimo komandą *python setup.py install*, kuri instaliuoja visas reikiamas bibliotekas.
3. Paleisti komandą *python collectstatic*, kuri sukurs statinių failų aplanką;
4. (Nebūtina) Jei norima, galima paleisti komandą *python download\_models.py*, kuri atsiunčia mašininio mokymosi modelius, su kuriais galima išsibandyti veido ir/ar balso keitimą;
5. Sukurti duomenų bazės lenteles naudojant komandą: *python manage.py migrate*;
6. Paleisti sistemą naudojant komandą *python manage.py runserver*.

### 4.5.4. Sistemos diegimas iš paketų valdymo (PyPI)

Norint instaliuoti bei paleisti sistemą reikia atlikti šiuos veiksmus:

1. Sukurti virtualią aplinką, kuri naudotų *Python 3.8* versiją bei ją aktyvuoti;
2. Sukurti Django projektą;
3. Paleisti instaliavimo komandą *pip install person-swap*, kuri instaliuoja visas reikiamas bibliotekas.
4. Pridėti komponentą prie sukurto projekto url's nuorodų;
5. Paleisti komandą *python collectstatic*, kuri sukurs statinių failų aplanką;
6. (Nebūtina) Jei norima, galima paleisti komandą *python download\_models.py*, kuri atsiunčia mašininio mokymosi modelius, su kuriais galima išsibandyti veido ir/ar balso keitimą;
7. Sukurti duomenų bazės lenteles naudojant komandą: *python manage.py migrate*;
8. Sukurti duomenų bazės lenteles naudojant komandą: *python manage.py migrate*;
9. Paleisti sistemą naudojant komandą *python manage.py runserver*.

### 4.5.5. Sistemos licencija

Pasirinkta MIT licencija, kurio tekstas pateikiamas prie priedų.

## 4.6. Pakeitimų sąrašas

4-51 lentelė. Pakeitimų sąrašas

Pokytis	Skyrius	Priežastis	Inicijavo
Pridėtos papildomos klasės	4.5 ir 5.7 skyriai	Sistemos realizavimo metu išaiškėjo, kad reikia daugiau klasių	Indrė Dimšė



FR-4, NFR-9, NFR-3 reikalavimai atnaujinti	4 skyrius	Sistema, egzistuojanti skirtingose aplinkose gali veikti skirtingai	Indrė Dimšė
--	-----------	---	-------------

#### 4.7. Kokybės vertinimas

Šiame skyriuje pateikiama projekto kokybės vertinimo proceso ir procedūrų aprašymai, vertinimo rezultatai, išvados. Nustatyta ar programų sistema atitinka standartus bei reikalavimų specifikaciją, ar išpildyti užsakovo lūkesčiai.

##### 4.7.1. Atlikto darbo kokybės analizės tikslai

###### **Aptikti klaidas funkcionavime, logikoje, realizacijoje.**

Sukurtos programinės įrangos klaidos ir veikimo netikslumai buvo aptikti naudojantis procedūromis, aprašytomis testavimo plane.

###### **Patikrinti ar programų sistema atitinka reikalavimų specifikaciją.**

Sukurtos programų sistemos atitikimo tikrinimas reikalavimo specifikacijomis atliktas priėmimo testavimo metu.

###### **Įsitikinti, ar programų sistema sukurta pagal standartus.**

Sukurtos programų sistemos atitikimas standartams aprašytas šiame skyriuje, kuriame yra apžvelgtas atitikimas geriausiomis praktikomis bei standartams.

##### 4.7.2. Kokybės vertinimo procesas

###### 4.7.2.1. Peržiūros

Projekto kūrimo pabaigoje vykdytojų komanda peržiūri visą projektą, surenka informaciją apie projektą bei pateikia projekto sprendimų apžvalgą. Apžvalgoje pateikiami sprendimai, kurie galės būti naudojami ir kituose projektuose. Taip pat aprašomi ir blogi, netinkami sprendimai, kuriuos reikia tobulinti.

Ši apžvalga dažniausiai atliekama projekto pabaigoje, jei projektas yra nedidelės apimties, arba po kiekvienos svarbesnės dalies pabaigimo, jei projektas yra didelis.

Projektavimo komandos narių peržiūrų aprašymas pavaizduotas 4.52 lentelė.

**4-52 lentelė.** Peržiūrų lentelė

Peržiūra	Aprašymas
Naudojamo proceso apibrėžimas	Išrinkti asmenis, kurie turėtų dalyvauti apžvalgoje
Sukurti duomenų surinkimo priemonės	Identifikuoti renkamą informaciją ir nustatyti jos rinkimo įrankius.
Duomenų surinkimas	Surinkti duomenis apžvalgai
Peržiūros santraukos peržiūra	Peržvelgti peržiūros rezultatus su visais dalyviais ir pasiekti vieningų išvadų

Rolės ir atsakomybės pavaizduotos 4-53 lentelė ir 4-54 lentelė.

#### 4-53 lentelė. Komandos narių rolės

Rolės	Asmuo
Moderatorius	Projekto vadovas
Dalyviai	Projekto kūrėjas
Raštininkas	Projekto kūrėjas

#### 4-54 lentelė. Informacijos rinkimo planas

Atsakingas asmuo	Atliekamos atsakomybės
Moderatorius	Nustatyti, kokia informacija yra surinkta: <ul style="list-style-type: none"><li>○ Projekto biudžetas.</li><li>○ Numatytas darbuotojų kiekis ir, kiek iš tikro darbuotojai skyrė laiko projekto įgyvendinimui.</li><li>○ Projekto įgyvendinimo tvarkaraštis.</li><li>○ Reikalavimų specifikacijos pakeitimai.</li></ul> Paruošti apklausos anketą.

#### 4-55 lentelė. Kokybės vertinimo duomenų rinkimo planas

Atsakingas asmuo	Atliekamos atsakomybės
Dalyviai	Užpildo paruoštas anketas.
Moderatorius	<ul style="list-style-type: none"><li>• Patvirtina galutinį anketos variantą.</li><li>• Užtikrina, kad visi dalyviai užpildytų anketas.</li><li>• Suorganizuoja pasitarimą su dalyviais.</li><li>• Pasitarimo metu veda diskusiją apie tai kas pavyko ir kas nepavyko.</li></ul>

#### 4.7.2.2. Formalios techninės peržiūros

Formalios techninės peržiūros buvo vykdomos naudojant *Python* programavimo kalbos *flake8* biblioteka bei *PyCharm* įrankį bei patikrinta:

- Ar metodams paduodami tinkami parametrai (pagal tipus);
- Ar metodai grąžina tinkamas reikšmes;
- Ar nėra pasikartojančio kodo;
- Ar nėra nenaudojamų kintamųjų;
- Ar nėra nenaudojamų funkcijų;
- Ar nėra funkcijų, kurios gali nepriklausyti klasei arba būti statinės;
- Ar visos importuojamos bibliotekos yra naudojamos;
- Ar nėra pasenusių ir artimiausiu metu pašalintų kintamųjų/metodų;
- Ar kodo stilius atitinka PEP8 standartą.

#### 4.7.3. Vertinimo rezultatai

Gauti vertinimo rezultatai pagal aprašytus vertinimo kriterijus. Rezultatus patartina pateikti lentelėmis ir grafikais. Taip pat realiai atlikto darbo kaina, darbo našumas.

Kokybės vertinimas atliktas pagal pagal randamumo, prieinamumo, sąveikumo bei pakartotinio naudojimo (angl. *FAIR*) principus. Vertinimo rezultatais priskiriamos 5 reikšmės:

- Puikus
- Geras
- Patenkinamas
- Nepatenkinamas;
- Numatytas.

**4-56 lentelė.** Kokybės vertinimo kriterijai

Parametras	Aprašymas	Vertinimas
Randamumas	Sistemos programinis kodas bus versijuojamas Gib kodo saugykloje.	Puikus
	Sistemoje saugomi duomenys bus saugomi pagal vartotojo norimą formatą, t.y vartotojas pats pasirinktą norimą duomenų bazės bei failų saugyklos tipą. Į sistemą įkelti failai bus papildomai apsaugoti šiais būdais: <ul style="list-style-type: none"> <li>• Failai bei modeliai sistemoje bus pavadinti vartotojo suteiktu vardu arba, jei vartotojas nesuteiks vardo: <ul style="list-style-type: none"> <li>○ modelio pavadinimas susidės iš tipo (vaizdo ar garso) ir įkėlimo datos;</li> <li>○ failo – iš modelio vardo (jei failas sukurtas sistemoje keitimo metu) bei įkėlimo datos.</li> </ul> </li> <li>• Pakeistiems failams prie meta datos bus pridėta sistemos informacija, modelio pavadinimas bei specifikacijos bei vartotojo, kuris keitė failą, informacija.</li> </ul>	Puikus
	Eksperimentų metu sukurti modelių pavadinimas susidės iš naudojamos architektūros pavadinimo bei modelio sukūrimo datos. Sukurti modeliai bus saugojami privačioje Gib saugykloje	Geras
	Eksperimentų metu naudojamos nuotraukos bei garso įrašai bus suskirstyti į atskirtus aplankus, kurių pavadinimai turi atspindėti aplankalo turinį. Duomenys bus paimami iš interneto, todėl jų saugojimu nereikia rūpintis	Patenkinamas
Prieinamumas	Sistemos programinis kodas bus pateiktas viešoje kodo saugykloje su laisvo naudojimo licencija	Puikus
	Į sistemą įkelti failai pagal autoriaus pageidavimą bus pažymimi kaip viešo/privataus naudojimo	Puikus
	Mašininio mokymosi modeliai bus įkelti į viešą saugyklą su laisvo naudojimo licencija	Numatytas
	Eksperimentų metu naudojamos nuotraukos bei garso įrašai nebus atskirai saugomi, nes jie jau egzistuoja internete ir yra viešai prieinami, tačiau bus aprašyti	Numatytas
Sąveikumas	Programinis kodas bei sukurti modeliai bus saugoti viešoje saugykloje ir visiems prieinami.	Puikus
Pakartotinis naudojimas:	Sukurtai sistemai bei modeliams bus suteikta MIT licencija bei nebus varžomas jų naudojimas bei šie duomenys bus saugomi viešoje saugykloje ir visiems prieinami.	Puikus
	Programinė įranga gerai dokumentuota bei paruošta pakartotiniam naudojimui	Puikus
Duomenų saugojimo kaina	Duomenys bus saugojami sistemos kūrėjo Gib saugykloje, todėl saugojimas nieko nekainuos.	Puikus

### Realio darbo kaina

Projektas vykdytas nekomerciniais tikslais, todėl darbo kaina neaktuali ir neskaičiuota.

**Darbo našumas**

Projektas vykdytas įvykdytas laiku.

## 5. Tyrimas

Tyrimo metu naudojantis *Python* kodo kokybės ir saugumo bibliotekomis buvo nuskanuotas projektas bei rastos galimos projekto problemos.

### 5.1. Kodo kokybė

„*Pylint*“ [4] yra *Python* biblioteka, kuri padeda nustatyti įvairias kodo problemas, įskaitant sintaksės klaidas, nenaudojamus kintamuosius ir kodavimo stiliaus pažeidimus. Ši biblioteka atrado šias problemas:

- (E5110) `person_swap/__init__.py` faile rodo, kad *Django* nebuvo tinkamai sukonfigūruotas.
- (C0305) `person_swap/settings.py` faile rodo, kad failo pabaigoje yra baigiamųjų naujųjų eilučių.
- (C0413) `person_swap/settings.py` faile rodo, kad *mimetypes* importas turėtų būti patalpintas modulio viršuje.
- `Person_swap/urls.py` faile yra kelios problemos, susijusios su importo tvarka ir grupavimu.
  - (E0401) rodo, kad nebuvo rastas brūkšnelio modulis
  - (C0411) rodo, kad trečiosios šalies importas turėtų būti pateiktas prieš vietinį importą
  - (C0412) rodo, kad importai iš *Django* paketo turėtų būti grupuojami atskirai nuo kitų importų.
- (C0304) `person_swap/wsgi.py` faile rodo, kad trūksta paskutinės naujos eilutės

„*Pylint*“ kodą įvertino 7,09/10 balo. Nors įvertinimas nėra tobulas, jis rodo, kad projektas iš yra pakankamai gerai struktūrizuotas ir atitinka gerą kodavimo praktiką.

„*Vulture*“ [5] yra *Python* biblioteka, naudojama nenaudojamam ir negyvam kodui aptikti *Python* programose. Ji statiškai analizuoja *Python* kodą ir nustato nenaudojamas klases, funkcijas, kintamuosius ir modulius, taip pat mirusį kodą, t. y. kodą, kuris niekada nevykdomas dėl nepasiekiamų ar perteklinių teiginių.

Šiame projekte buvo rastos nenaudojamos:

- Klasės – 17
- Funkcijos – 4
- Kintamieji – 12

### 5.2. Saugumas

„*Safety*“ [6] biblioteka – tai paketas, kuris skenuoja *Python* aplinką ir tikrina, ar projekte nėra žinomų pažeidžiamumų. Ją galima naudoti siekiant užtikrinti, kad jūsų programoje nebūtų naudojami paketai su žinomomis saugumo problemomis, o tai gali padėti išvengti atakų prieš jūsų programą.

Paleidus šią biblioteką buvo rastos penki pažeidžiamumai ir keturios rekomendacijos, ištaisyti pažeidžiamumus.

- „*Setuptools*“ versija 49.2.1 – dėl šios spragos nuotoliniai įsilaužėliai gali paleisti savavališką kodą arba vykdyti atsisakymo aptarnauti atakas. Ši problema ištaisyta 65.5.1 versijoje.
- „*Gitpython*“ 3.0.6 versija – dėl šio pažeidžiamumo nuotoliniai įsilaužėliai gali sukelti atsisakymo aptarnauti (*DoS*) ataką per sukurtą reguliariąją išraišką „*git*“ saugykloje. Problema ištaisyta 3.1.27 versijoje.
- „*Gitpython*“ 3.0.6 versija – dėl šios spragos nuotoliniai įsilaužėliai gali paleisti savavališką kodą per sukurtą saugyklos URL. Ši problema ištaisyta 3.1.30 versijoje

- „*Httpx*“ versija 0.22.0 – dėl šios spragos nuotoliniai įsilaužėliai gali paleisti savavališką kodą per sukurtą URL adresą. Problema ištaisyta 0.23.0 versijoje.
- „*Cryptography*“ 35.0.0 versija – dėl šios spragos nuotoliniai įsilaužėliai gali paleisti savavališką kodą per sukurtą *Cipher.update\_into()* funkcijos įvestį. Problema ištaisyta 39.0.1 versijoje.

„*Bandit*“ [7] yra kita *Python* kodo saugumo biblioteka, kuri aptinka įvairių tipų saugumo spragas, įskaitant galimas XSS (kryžminio svetainių skriptavimo) spragas, kietai užkoduotus įgaliojimus ir kt. Ji naudoja skirtingus įskiepius įvairių tipų problemoms aptikti ir priskiria joms rimtumo lygį (žemas, vidutinis, aukštas).

Šiame projekte buvo aptiktos šios problemos:

1. Trys vietos, kuriose naudojami *assert*, kuriuos „*Bandit*“ pažymi kaip galimas problemas, kurias reikėtų peržiūrėti, siekiant užtikrinti, kad jos nesukurtų saugumo spragų. Šios problemos laikomos mažo pavojingumo.
2. Dvi funkcijos *mark\_safe()*, kurios pažymėtos kaip potencialios XSS spragos. *mark\_safe()* yra *Django* funkcija, kuri pažymi eilutę kaip saugią HTML, tačiau ja galima piktnaudžiauti, todėl gali kilti potencialių saugumo problemų. Šios problemos pažymėtos kaip vidutinio pavojingumo.

### 5.3. Išvados

Projektas buvo patikrintas naudojant kelias *Python* bibliotekas:

- „*Pylint*“, kuris aptiko keletą kodo problemų, tačiau apskritai projekte laikomasi geros kodavimo praktikos.
- „*Safety*“ biblioteka aptiko penkis pažeidžiamumus ir pateikė rekomendacijas, kaip juos ištaisyti. Nenaudojamam kodui aptikti naudotas.
- „*Vulture*“ nustatė, kad projekte yra nenaudojama 17 klasių, 4 funkcijos ir 12 kintamųjų.
- „*Bandit*“ aptiko kelių tipų saugumo spragų.

Apibendrinus galima teigti, jog projektą reikia šiek tiek patobulinti, tačiau jis turi gerą pagrindą ir gali būti toliau plėtojamas.

## 6. Eksperimentai

Mašininio mokymosi modeliai pasižymi puikiais rezultatais įvairiose srityse, įskaitant vaizdų ir kalbos apdorojimą. Viena iš sričių, kurioje mašininis mokymasis padarė didelę pažangą, yra veidų ir balsų keitimas, leidžiantis manipuluoti veido ir balso išraiškomis tam tikrame vaizdo ar garso įrašė. Veido ir balso keitimas gali būti taikomas įvairiai – nuo pramogų ir reklamos iki stebėjimo ir saugumo.

Šiame skyriuje siekiame palyginti ir eksperimentuoti su dvejais veidų keitimo mašininio mokymosi modeliais ir trimis balso keitimo mašininio mokymosi modeliais su trimis vokoderiais, kad nustatytume jų santykinį veiksmingumą keičiant veidus ir balsus.

### 6.1. Veido keitimas

Pastaraisiais metais manipuliacijos veidais sulaukė didelio dėmesio dėl plataus pritaikymo pramogų, saugumo ir kriminalistikos srityse. Ypač populiarus tapo veidų keitimo metodas, kai nuotraukoje ar vaizdo įrašė asmens veidas pakeičiamas kitu veidu. Veido keitimo algoritmų sėkmė priklauso nuo gebėjimo tiksliai užfiksuoti pradinio ir tikslinio veido bruožus bei detales ir sklandžiai juos perkelti. Šiai užduočiai spręsti pasiūlytos įvairios giliojo mokymosi architektūros, kurių kiekviena turi savų stipriųjų ir silpnųjų pusių. Šiame darbe lyginame auto-koduotojo ir konvoliucinio auto-koduotojo kartu su architektūras, skirtas veidams sukeisti vietomis. Siekiame įvertinti kiekvienos architektūros našumą pagal generuojamų vaizdų kokybę, apdorojimo greitį ir atsparumą įvairių tipų veido išraiškoms, pozoms ir apšvietimo sąlygoms. Šis palyginimas padės išsiaiškinti kiekvienos architektūros privalumus ir trūkumus bei padės ateityje kurti veiksmingesnius ir efektyvesnius veidų sukeitimo algoritmus.

Eksperimentas buvo pasitelktos šios bibliotekos:

- *Tensorflow* [8];
- *Pytorch* [9];
- *DeepFaceLab* [16];
- *CV2* [17];
- *h5py* [18].

#### 6.1.1. Duomenų rinkinys

Eksperimentams naudoju „*DeepFakeVFX*“ [22] pateiktus įžymybių veidų duomenų rinkinius, kuriuose yra aukštos kokybės įžymybių veidų atvaizdai. Konkrečiai pasirinkau „viso veido“ veidų tipą, į kurį įeina visas veidas, ir pasirinkau 512x512 raiškos vyrų ir moterų vaizdus.

Pasirinkti 2 duomenų rinkiniai – vienas vyro ir vienas moters po 10000 atvaizdų. Duomenis suskirsčiau į mokymo, tikrinimo ir testavimo rinkinius, 80 % duomenų panaudojau mokymui, 10 % – tikrinimui ir 10 % – testavimui.

#### 6.1.2. Duomenų apdorojimas

Prieš mokydamas veidų keitimo modelius atlikau kelis pirminio apdorojimo veiksmus, kad paruoščiau duomenis. Pirmiausia nuotraukų dydį iki 128x128 pikselių, kad pagreitinti mokymąsi. Šis žingsnis taip pat užtikrino, kad visi vaizdai buvo vienodo dydžio, jei per žmogišką klaidą duomenų rinkinyje buvo palikti kitokio dydžio vaizdai.

Taip pat, siekiant išvengti perteklinio pritaikymo ir pagerinti apibendrinimo našumą, taikiau atsitiktinius sukimus, vertimus ir priartinimus vaizdus mokymo metu. Šios transformacijos prideda dirbtinio kintamumo mokymo duomenims, o tai padeda modeliams išmokti patikimų atvaizdų įvesties duomenis.

### 6.1.3. Modelių architektūra

Abi architektūros turi savo stipriųjų ir silpnųjų pusių, susijusių su veidų keitimu arba giliomis klastotėmis. Automatinio kodavimu pagrįstus metodus galima greičiau apmokyti, tačiau jie gali duoti mažiau tikroviškus rezultatus. Konvoliuciniu auto-koduotoju pagrįsti metodai gali duoti tikroviškesnius rezultatus, tačiau jiems reikia daugiau duomenų ir skaičiavimo išteklių. Architektūros pasirinkimas priklausys nuo konkrečių projekto poreikių ir turimų išteklių.

6-1 lentelė. Veido keitimo modelių architektūros

Architektūra	Auto-koduotojas (AE)	Konvoliucinis Auto-koduotojas (CAE)
Latentinis erdvės dydis	Mažesnė dimensija	Didelė dimensija
Skaičiavimo resursų reikalavimai	Mažesnis	Didesnis
Duomenų reikalavimai	Mažesnis	Didesnis
Skaičiavimo sudėtingumas	Santykinai žemas	Santykinai didelis
Taikymas	Vaizdų interpoliacija, vaizdų užbaigimas	Veido atpažinimas, veido patvirtinimas, veido aptikimas
Stiprybės	Greitesnis mokymas, reikia mažiau duomenų	Aukštos kokybės išvestis, daugiau realizmo
Silpnybės	Mažiau realistiški rezultatai	Ilgesnis mokymasis, reikia daugiau duomenų
Vertinimo funkcija	Rekonstrukcijos paklaidų minimizavimas (angl. <i>Reconstruction error minimization</i> )	<i>Softmax</i> regresija veido atpažinimui (angl. <i>Softmax regression for face recognition</i> )

#### 6.1.3.1. Auto-koduotojas (AE)

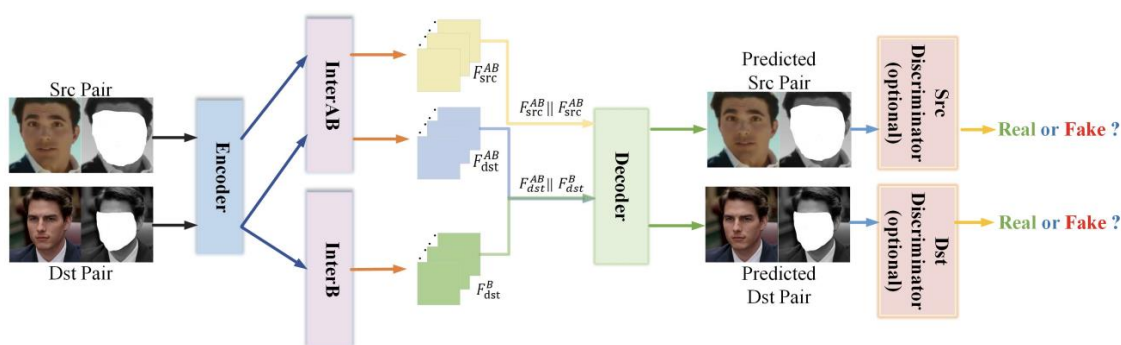
Auto-koduotojo architektūra yra populiarus gilaus mokymosi metodas, naudojamas aukštos kokybės veido atvaizdams generuoti. Šią architektūrą sudaro du pagrindiniai komponentai: kodavimo tinklas, kuris įvesties vaizdą atvaizduoja į mažo matmens požymių atvaizdavimą, ir dekodavimo tinklas, kuris požymių atvaizdavimą atvaizduoja atgal į išvesties vaizdą [36].

Tada modelis mokomas naudojant nuostolių funkciją, pagrįstą vidutine kvadratine paklaida (*MSE*) tarp pradinio įvesties vaizdo ir atkurto išvesties vaizdo. Mokymo procesas apima šios nuostolių funkcijos minimizavimą naudojant grįžtamąjį skleidimą, kuris koreguoja neuroninio tinklo svorius ir šališkumą, kad pagerintų rekonstruotų vaizdų tikslumą.

Šaltinio ir tikslinių vaizdų modeliui apmokyti naudojamos skirtingos interpoliavimo funkcijos ir diskriminatoriai, nes abiejų vaizdų rinkinių savybės ir požymiai skiriasi, o skirtingų metodų naudojimas gali padėti optimizuoti modelį kiekvienam vaizdų rinkiniui atskirai. Be to, kiekvienam vaizdų rinkiniui naudojamas interpoliavimo metodas gali būti optimizuotas konkrečioms įvesties duomenų aspektams, pavyzdžiui, veido orientyrams išsaugoti arba iškraipymams sumažinti. Naudojant skirtingas interpoliavimo funkcijas ir diskriminatorius pradiniam ir tiksliniam vaizdams,



modelį galima veiksmingiau apmokyti, kad būtų sukurti aukštos kokybės išvesties vaizdai, kuriuose tiksliai užfiksuoti pageidaujami veido bruožai ir išraiškos.



6-1 pav. Auto-koduotojo architektūra [36]

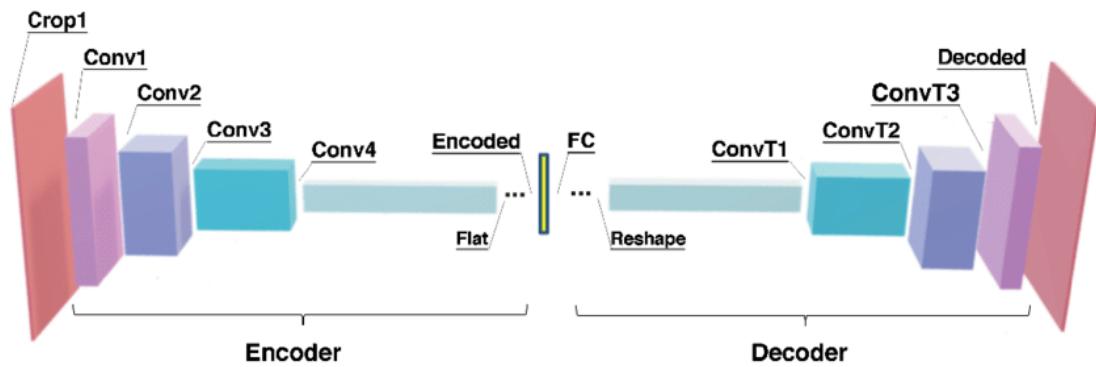
### 6.1.3.2. Konvoliucinis auto-koduotojas

Konvoliucinio automatinio kodavimo (*CAE*) architektūra yra neuroninio tinklo architektūra, kurią sudaro du pagrindiniai komponentai: koduotojas ir dekoduoja. Koduotojas priima įvesties vaizdą ir suspaudžia jį į mažo matmens požymių atvaizdavimą, o dekoduoja priima šį požymių atvaizdavimą ir atkuria originalų vaizdą. *AE* modelį irgi pagrinde sudaro koduotojas ir dekoduoja, tačiau *CAE* modelyje į koduotoją ir dekoduoja įterpami konvoliuciniai sluoksniai, todėl jis geriau tinka vaizdų apdorojimo užduotims. Konvoliuciniai sluoksniai leidžia *CAE* modeliui užfiksuoti įvesties vaizdo erdvinės savybes, todėl galima tiksliau atkurti vaizdą ir geriau atlikti tokias užduotis kaip veidų keitimas. Konvoliuciniai sluoksniai pavaizduoti 6-2 pav. [37].

Koduotoją sudaro keli konvoliuciniai sluoksniai, kurių kiekvienas įvesties vaizdai taiko mokytis leidžiančių filtrų rinkinį, kad iš jo išgautų požymius. Po šių konvoliucinių sluoksnių seka keli visiškai sujungti sluoksniai, kurie konvoliucinių sluoksnių išskirtus didelio matmens požymius atvaizduoja į mažesnio matmens požymių atvaizdavimą. Kodavimo įrenginio išvestis yra suspaustas įvesties vaizdo atvaizdas, kuriame yra tik svarbiausi požymiai [36].

Tada dekoduoja paima šį suspaustą atvaizdą ir iš jo atkuria originalų vaizdą. Jį sudaro keli dekonvoliucijos sluoksniai, kurie taiko atvirkštinę konvoliucijos operaciją suspaustam požymių atvaizdavimui padidinti. Šie sluoksniai palaipsniui didina požymių atvaizdo skiriamąją gebą, kol gaunamas galutinis išvesties vaizdas.

*CAE* architektūra sumažina skirtumą tarp įvesties vaizdo ir atkurto išvesties vaizdo, naudojant nuostolių funkciją. Mokymo proceso metu modelio svoriai ir šališkumas koreguojami naudojant atgalinį sklaidimą, kad būtų sumažinta ši nuostolių funkcija.



6-2 pav. Konvoliuciniai sluoksniai [37]

#### 6.1.4. Modelių mokymas

Visų modelių mokymosi metu buvo nustatyti šie parametrai:

- Naudojant mažo dydžio vaizdus (128x128), galima sugeneruoti pakankamos kokybės vaizdus tyrimui, nes įvesties ir išvesties vaizduose galima matyti daugiau detalių. Taip pat dėl mažesnio dydžio mokymasis vyksta greičiau;
- Modelio parametras *face\_type* nustatytas į "wf" (pilnas veidas, angl. *whole face*), kuris reiškia, kad įvesties vaizdo veido sritis yra sulygiuota su vaizdo centru ir turi fiksuotą plotį bei aukštį.

Kai kuriuose modelių mokymuose naudojamas *Adabelief* optimizatorius, kuris yra Adamo optimizatoriaus variantas. Šis optimizatorius gerai veikia įvairiose gilaus mokymosi užduotyse, įskaitant vaizdų generavimą.

Bus išbandytos dvi aktyvacijos funkcijos [38]:

- Kosinuso vienetas (angl. *Cosine Unit* arba *CU*), fiksuojantis kampinį panašumą ir periodinius modelius, gali būti naudingas vaizdų generavimo užduotims, susijusioms su cikliniais ar tekstūriniais vaizdais.
- *ReLU*, pasižymintis nelineiškumu ir gebėjimu fiksuoti sudėtingus vaizdo požymius, paprastai naudojamas vaizdų generavimo modeliuose, leidžiantis jiems mokytis ir generuoti tikroviškus ir įvairius vaizdus.

Šiems modeliams apmokyti reikia nemažai kompiuterinių išteklių, ypač jei naudojami didelės skiriamosios gebos vaizdai. Paprastai mokymas atliekamas naudojant galingus *GPU*, pavyzdžiui, *NVIDIA „Tesla“* arba „*GeForce*“ modelius. Apmokymas vyko „*Google Colab*“ aplinkoje, kurioje yra „*Tesla4*“ *GPU* modelis su daugiau nei 13GB atmintimi.

*CAE* mokymo procesas:

1. Veidai iš nuotraukų perduodami giliam neuroniniam tinklui, kurį sudaro koduotojas ir dekoduoja. Kodavimo įrenginys suspaudžia įvesties vaizdą į mažesnio matmens požymių atvaizdavimą, o dekoduoja atkuria vaizdą iš šio atvaizdavimo.
2. Mokymo metu naudojama nuostolių funkcija yra rekonstrukcijos nuostolių, priešingų nuostolių ir požymių nuostolių derinys. Rekonstrukcijos nuostoliai matuoja skirtumą tarp įvesties ir išvesties vaizdų, priešingi nuostoliai skatina generatorių kurti tikroviškus vaizdus, o požymių nuostoliai užtikrina, kad sukurtame vaizde būtų išsaugoti tikslingo vaizdo veido bruožai.
3. Mokymo procesas paprastai vykdomas naudojant duomenų rinkinį, kurį sudaro suporuoti šaltinio ir tikslo vaizdai, o generatorius mokomas transformuoti šaltinio vaizdus į tikslo vaizdus.

AE mokymo procesas:

1. AE naudoja linijinio auto-koduotojo architektūrą, kurią sudaro koduotojas ir dekoduojuojas, sujungti linijinės transformacijos matrica. AE apmokytas taip, kad būtų sumažinta įvesties ir išvesties vaizdų atstatymo paklaida.
2. Mokymo metu koduotojas išmoksta retą požymių atvaizdavimą, o dekoduojuojas pagal šį atvaizdavimą generuoja išvesties vaizdą.
3. AE modelyje taip pat naudojamas priešingas nuostolis, skatinantis generatorių kurti tikroviškai atrodančius vaizdus.

Kiekvienos architektūros modeliai bus apmokyti keliais būdais (su skirtingais parametrais) tam, kad nustatyta, kaip geriausia apmokyti veido keitimo modelius.

**6-2 lentelė.** Veido keitimo modelių mokymų aprašymas

Modelis	AE I	CAE I	AE II	CAE II	AE III	CAE III	AE VI	CAE VI	AE V	CAE V
AE dimensija	128	128	64	64	128	128	64	64	64	64
Atsitiktinis duomenų išmetimas	+	-	+	+	+	-	+	+	+	+
Naudojamas adabelief	+	+	+	+	+	+	+	+	-	-
Aktyvacijos funkcija	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	CU	CU	CU	CU
GAN spalvų plečių dydis	16	16	32	32	16	16	32	32	32	32
GAN dimensija	16	16	32	32	16	16	32	32	32	32
GAN stiprumas	0	0	3	3	2	0.1	3	3	3	3
Atsitiktinis intensyvumas	0.05	0.05	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Akių/burnos prioritetas	-	-	-	-	+	+	-	-	-	-

Lentelėje esančių stulpelių paaiškinimai

- GAN spalvų plečių dydis – GAN sluoksnio spalvų kanalų skaičių.
- GAN dimensija – GAN sluoksnio dimensijų skaičius.
- GAN stiprumas – nusako, kaip stipriai neuroninis tinklas mokosi smulkių veido detalių.
- Atsitiktinis intensyvumas – atsitiktinis atspalvis / sodrumas / šviesos intensyvumas.

### 6.1.5. Modelių vertinimas

Visuose modeliuose naudojama mišri nuostolių funkcija, kuri yra dviejų nuostolių funkcijų derinys: vidutinės kvadratinės klaidos (*MSE*) ir struktūrinio panašumo indekso (*SSIM*) nuostolių derinys [].

**6-3 lentelė.** Balso keitimo tikrinimo metrikos

Metrikos pavadinimas	Vidutinio kvadrato paklaidos ( <i>MSE</i> )	Struktūrinio panašumo indeksas ( <i>SSIM</i> )
----------------------	---	--

Procesas	Matuoja vidutinį kvadratinį skirtumą tarp dviejų vaizdų pikselių	Lygina dviejų vaizdų struktūrinį panašumą pagal šviesumą, kontrastą ir struktūrą
Interpretacija	Mažesnės reikšmės rodo didesnį panašumą tarp dviejų vaizdų	Reikšmės, artimesnės 1, rodo didesnį panašumą tarp dviejų vaizdų
Privalumai	Lengva apskaičiuoti, suteikia kiekybinę informaciją apie vaizdo kokybės skirtumus	Atsižvelgia į žmogaus suvokimą apie vaizdo kokybę ir panašumą
Ribojimai	Nepaiso struktūrinės informacijos arba suvokimo skirtumų tarp vaizdų	Gali būti jautrus kontrasto ar šviesumo pokyčiams, gali netikti visiems vaizdų palyginimams
Naudojimo atvejis	Naudinga paprastiems vaizdų palyginimo uždaviniams, pvz., nustatant, ar du vaizdai yra tapatybės atžvilgiu identiški	Naudinga sudėtingesniems vaizdų palyginimo uždaviniams, pvz., vertinant panašumą tarp dviejų vaizdų, kurių šviesumas ar kontrastas skiriasi

*MSE* ir *SSIM* nuostolių derinys padeda rasti kompromisą tarp apibendrinimo ir aiškumo. Šio derinio motyvacija – gauti naudos iš abiejų: *SSIM* greičiau apibendrina žmonių veidus, o *MSE* užtikrina didesnį aiškumą. Šių dviejų nuostolių svorius galima reguliuoti, kad būtų rasta geriausia apibendrinimo ir aiškumo pusiausvyra.

### 6.1.6. Rezultatai

Žemiau esančioje lentelėje pavaizduoti visi veido keitimo modelių mokymai. Plačiau šie rezultatai pavaizduot antrame priede.

6-4 lentelė. Veido mokymų rezultatai

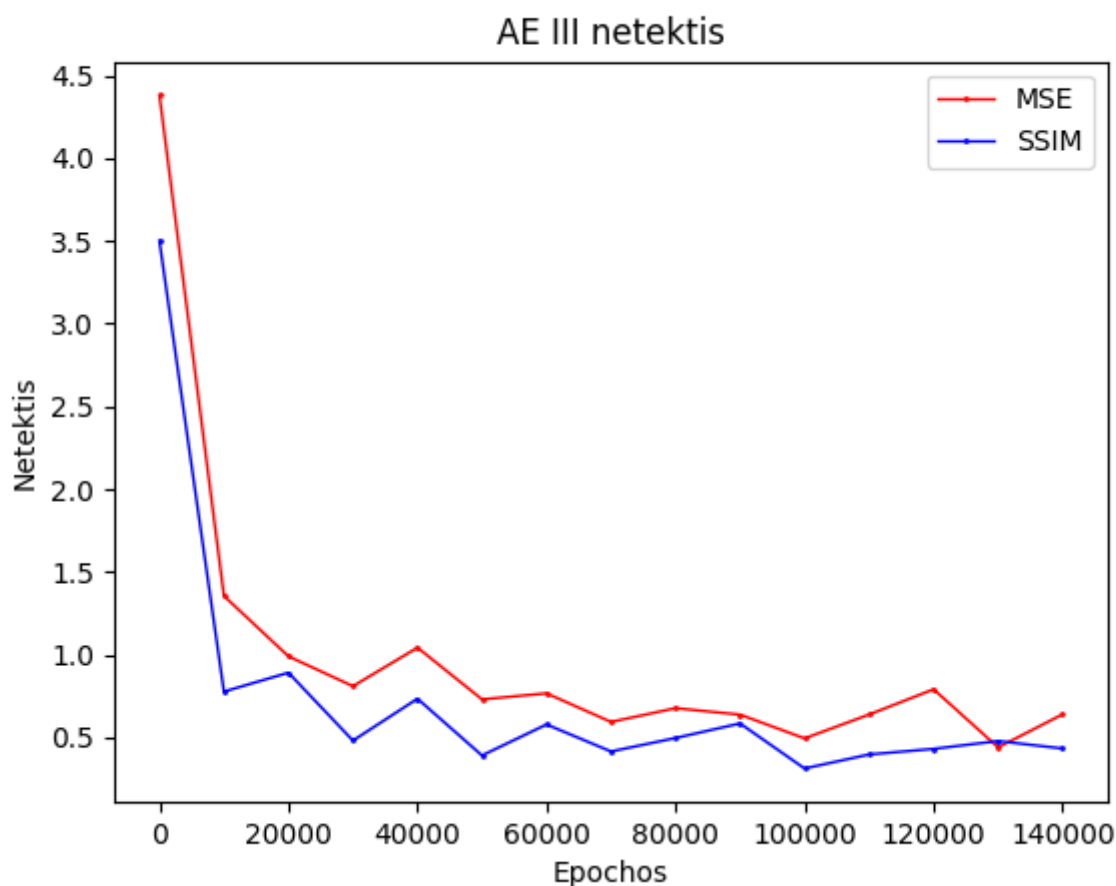
Modelio mokymai	<i>SSIM</i>	<i>MSE</i>	Epochos	<i>MSE</i> (100k)	<i>SSIM</i> (100k)	Modelio dydis (MB)
<i>AE I</i>	0.522	0.489	380000	0.747	0.929	508
<i>AE II</i>	0.724	0.985	300000	1.934	0.913	212
<i>AE III</i>	0.313	0.440	140000	0.495	0.313	513
<i>AE IV</i>	1.006	1.380	100000	1.380	1.006	209
<i>AE V</i>	1.041	1.609	100000	1.609	1.041	896
<i>CAE I</i>	0.483	0.620	680000	1.038	0.758	479
<i>CAE II</i>	0.742	1.028	260000	1.337	1.147	106
<i>CAE III</i>	0.450	0.783	150000	0.896	0.626	483
<i>CAE IV</i>	0.991	1.718	110000	1.718	0.991	208
<i>CAE V</i>	1.320	1.825	100000	1.825	1.320	141

*AE III* ir *CAE III* mokymų rezultatai pilnai aprašyti žemiau esančiuose skyreliuose, nes šie modeliai pasiekė geriausius rezultatus per trumpiausią laiką.

Tolesniame skyrelyje lyginamos pirmos 100000 skirtingų mokymų epochos. Mokymai skiriasi dimensijomis, pasukimais, optimizatoriais, aktyvacijos funkcijomis bei veido bruožų prioritetais.

### 6.1.6.1. AE III

6-3 pav. esančiame grafe pavaizduotas *AE III* mokymas, nes jis iš auto-koduotojų serijos parodė geriausius rezultatus **6-3 pav.**



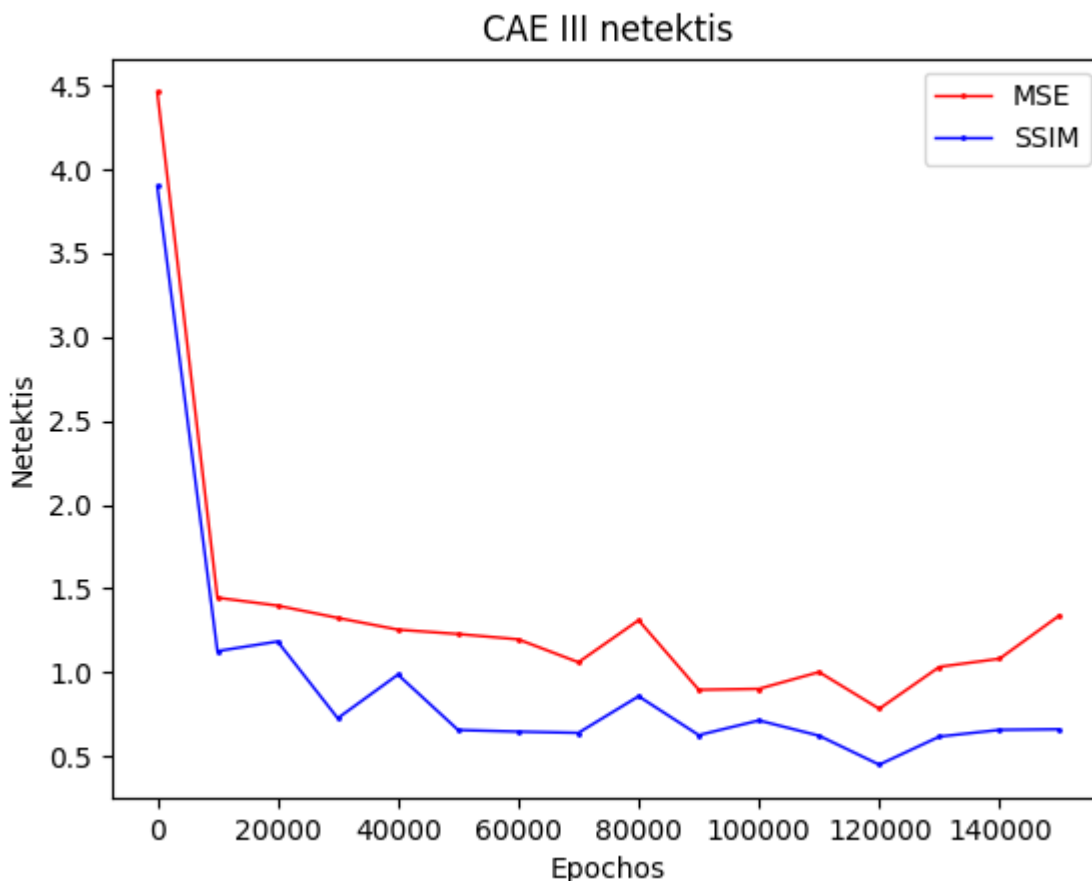
**6-3 pav.** *AE III* mokymo rezultatai

*AE III* mokymo santrauka:

- Mokymo eigoje tiek *MSE*, tiek *SSIM* vertės gerėja. *MSE* mažėja, o tai rodo geresnę rekonstrukcijos tikslumą, o *SSIM* mažėjimas rodo geresnę originalo ir rekonstruotų vaizdų struktūrinį panašumą. Kadangi modelis nebuvo apmokytas, iki 1000 epochos netekties reikšmės stipriai krenta.
- Vykstant mokymui atsiranda svyravimai ir pokyčiai. Tai gali lemti tokie veiksniai kaip įvesties vaizdų sudėtingumas, mokymo duomenų rinkinio įvairovė ir modelio mokymosi dinamika.
- Modelis visą laiką gerai mokosi ir jį būtų galima toliau mokyti ir pagerinti rezultatus. Geriausios modelio pasiektos netekties reikšmės yra *MSE* 0,44 ir *SSIM* 0,31, o tai rodo, kad modelis pasiekia gana gerų rezultatų vaizdo rekonstrukcijos ir struktūrinio panašumo požiūriu
- Galutinė *MSE* vertė 0,64 ir galutinė *SSIM* vertė 0,43.

### 6.1.6.2. CAE III

6-4 pav. esančiame grafe pavaizduotas *CAE III* mokymas, nes jis iš konvoliucinių auto-koduotojų serijos parodė geriausius rezultatus.



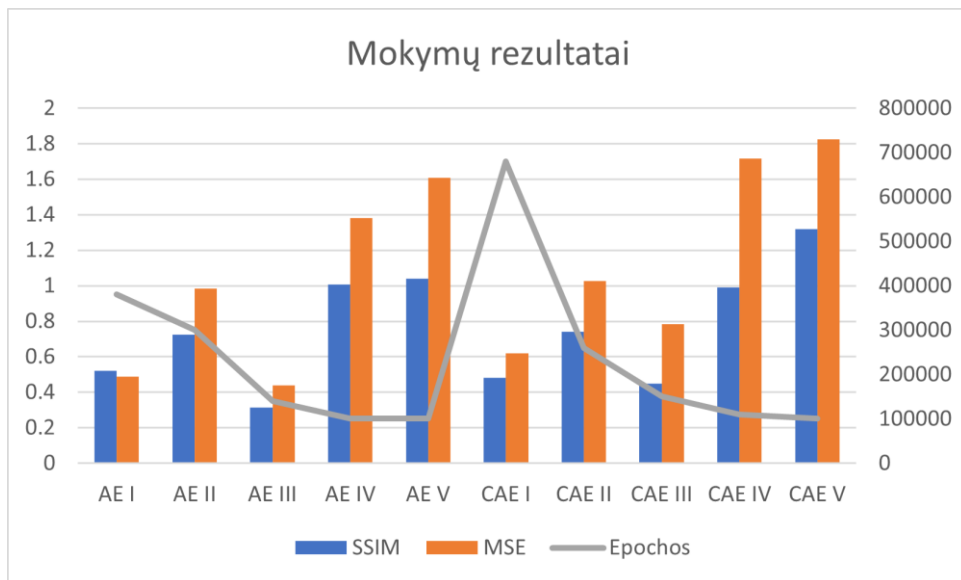
6-4 pav. CAE III mokymo rezultatai

CAE III mokymo santrauka:

- Nuo 0 iki 1000 epochos modelio kuriami rezultatai greitai gerėja – tai rodo, kad modelis neapmokytas.
- Mokymo eigoje *MSE* ir *SSIM* vertės svyruoja. Nors yra svyravimų, galime pastebėti bendrą *MSE* reikšmių gerėjimo tendenciją, rodančią geresnę rekonstrukcijos tikslumą, ir kintančias *SSIM* reikšmių tendencijas, rodančias pradinių ir rekonstruotų vaizdų struktūrinio panašumo pokyčius.
- Atrodo, kad modelio veikimas stabilizuojasi ties 12000 epocha ir toliau modelis pradeda persimokyti. Šiuo metu *MSE* vertė yra 0,45 ir *SSIM* vertė – 0,78.
- Galutinė *MSE* vertė 1,34 ir galutinė *SSIM* vertė 0,66 rodo modelio našumą, pasiektą po nurodyto epochų skaičiaus.

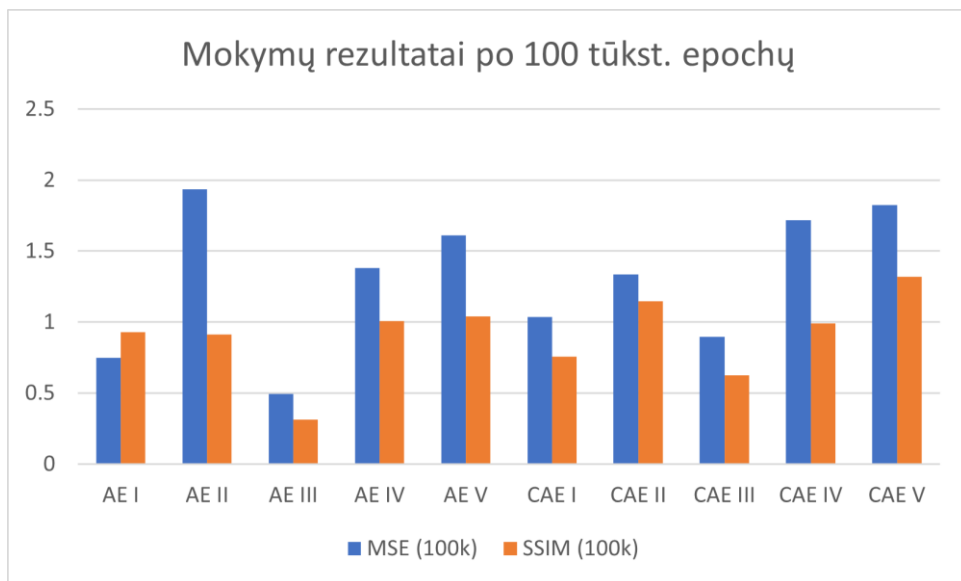
### 6.1.6.3. Parametrų palyginimas

Žemiau esančiame grafe gerai matome modelių kokybės pasiskirstymą. Geriausiai pasirodė AE III modelis, nors jis ir neturėjo daug iteracijų.



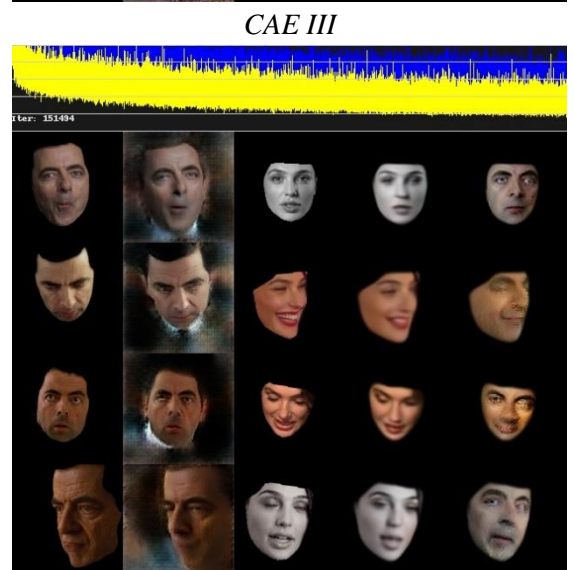
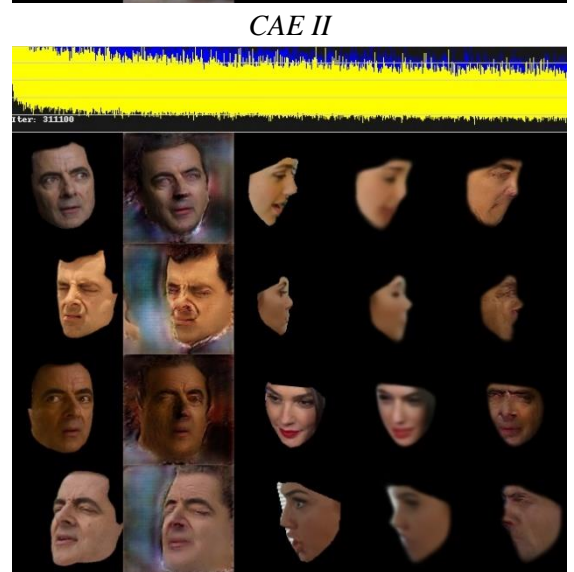
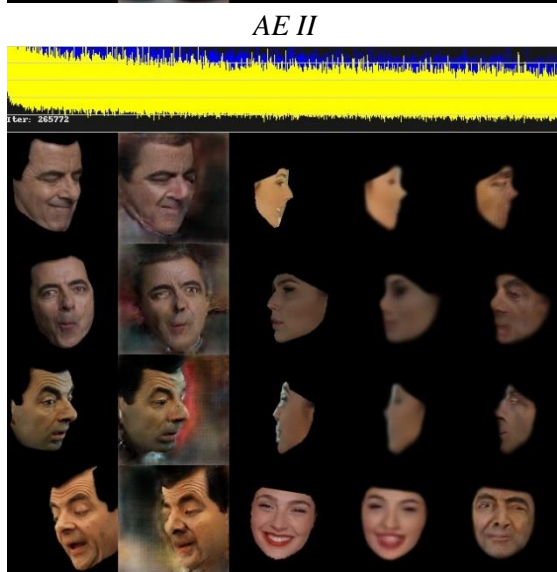
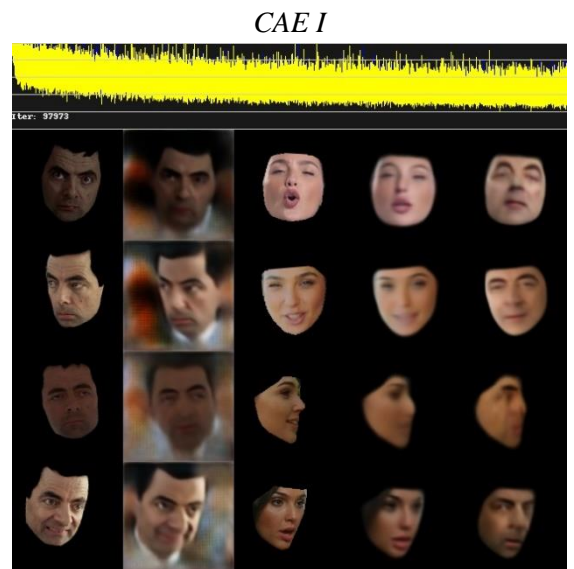
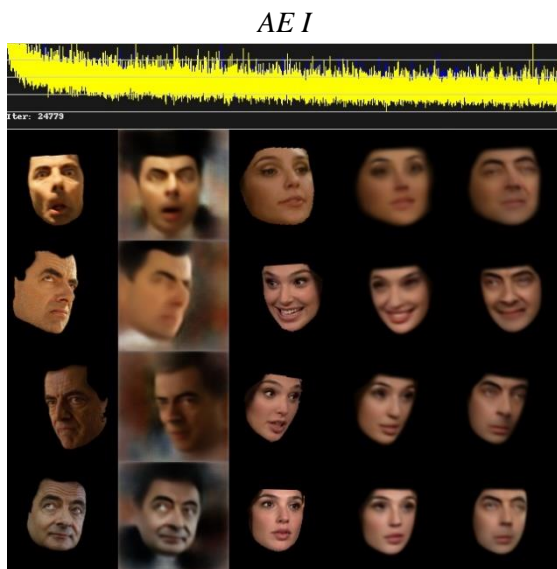
**6-5 pav.** Veido keitimo mokymų rezultatai

Žemiau pavaizduotas grafas ima tik pirmas 100000 epochų, kuriame matomas *AE II* lyderis. *CAE III* irgi greitai rodo neblogus rezultatus, kurie gana panašūs į *CAE I*, kuris turi daug daugiau iteracijų



**6-6 pav.** Veido keitimo mokymų rezultatai po 100 tūkst. epochų

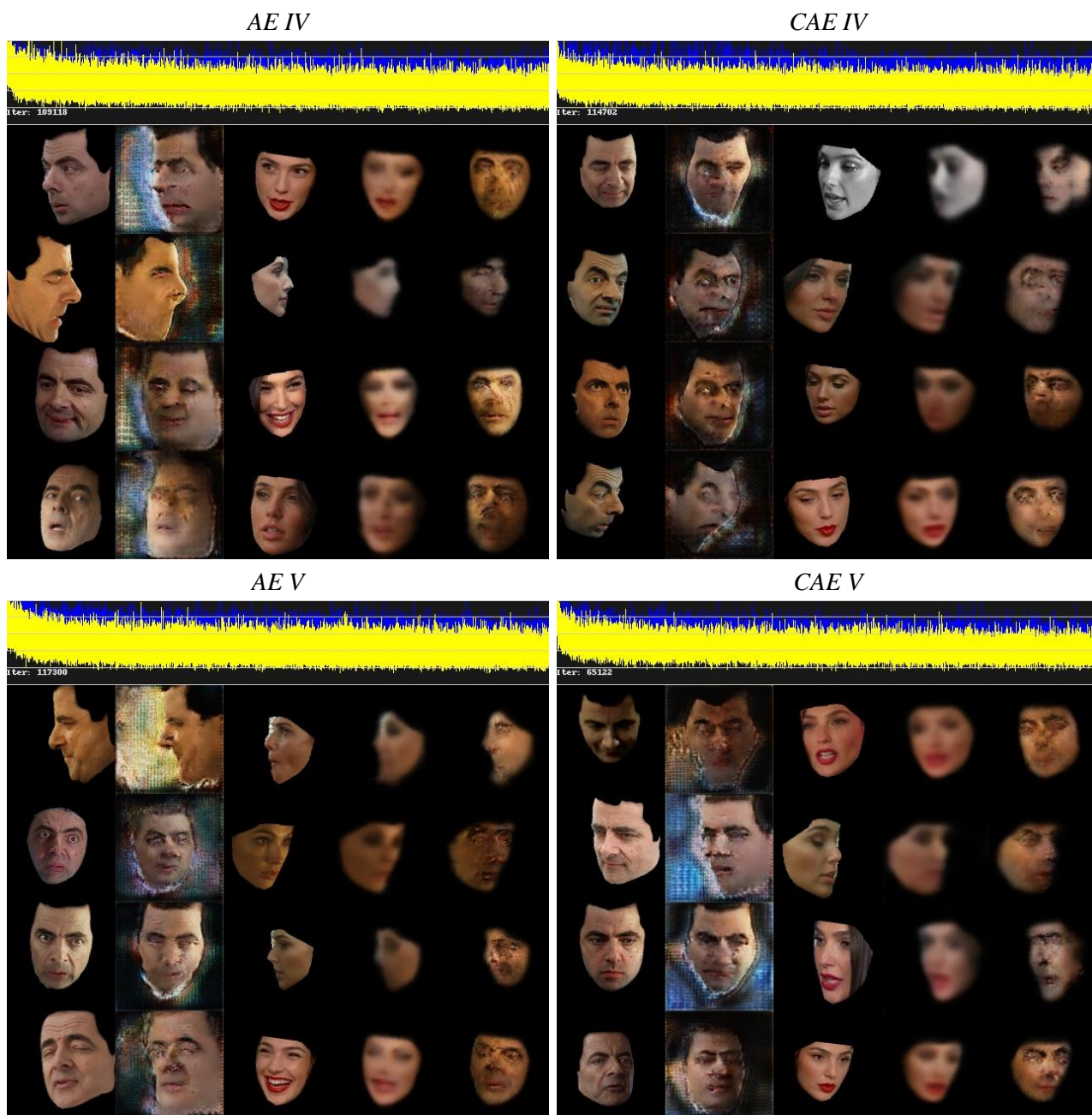
Žemiau pavaizduotos visų mokymų paskutinių epochų nuotraukų pavyzdžiai. 6-7 pav. pavaizduoti modelių sugeneruotos nuotraukos. *AE II* ir *CAE II* naudojamos mažesnės dimensijos bei matome, kad sugeneruotų nuotraukų rezultatai yra prastesni. Tuo tarpu *AE I* ir *CAE I* pasiekia gerus rezultatus su daug epochų, o tuo tarpu *AE III* ir *CAE III* pasiekia gerus rezultatus su mažiau epochų – todėl galime sakyti, jog stipresnis *GAN* naudojimas pagerina veido keitimo rezultatus.



6-7 pav. Veido keitimo pavyzdžiai naudojant *ReLU*



6-7 pav. pavaizduoti modelių sugeneruotos nuotraukos, kurie naudoja *ReLU* aktyvacijos funkcijos, o 6-6 pav – *CU* aktyvacijos funkciją. Iš karto galime matyti, jog *ReLU* aktyvacijos funkcija geriau tinka veido keitimui.



6-8 pav. Veido keitimo pavyzdžiai naudojant *CU*

### 6.1.7. Išvados

Abiejų architektūrų modeliai buvo apmokyti naudojant mažo dydžio vaizdus (128x128), kad būtų gauti tyrimui tinkamos kokybės vaizdai. Toks dydis pasirinktas todėl, kad įvesties ir išvesties vaizduose galima išvelgti daugiau detalių, be to, mokymasis vyksta greičiau. Buvo eksperimentuojama su pilnu veidu, kuris sulygiuoja įvesties vaizdo veido sritį su vaizdo centru ir turi fiksuotą plotį bei aukštį. Skirtingiems mokymams buvo keistos dimensijos, optimizatoriaus, aktyvacijos funkcija, *GAN* stiprumas bei veido bruožų prioritetai.

Veido keitimo eksperimentų metu geriau pasirodė modeliai, turintys didesnę auto-koduotojo, tačiau mažesnę *GAN* dimensiją. Taip pat spalvų virpesių naudojimas padėjo padidinti sukurtų vaizdų

tikroviškumą, o atsitiktinio duomenų iškritimo nebuvimas padėjo išsaugoti daugiau informacijos (požymių) veidų nuotraukose. Be to, šie modeliai siekė sutelkti dėmesį į tikroviškai atrodančių akių ir burnos generuojamuose veido atvaizduose generavimą.

Žinoma *ReLU* aktyvacijos funkcija bei *Adabelief* optimizatorius buvo naudojami geriau pasirodžiusiuose modeliuose, dėl ko galima spręsti, kad šie parametrai turi įtakos vaizdų generavime.

Buvo nustatyta, kad *AE* modelis yra skaičiavimo požiūriu efektyvesnis ir jam reikia mažiau mokymo duomenų, palyginti su *CAE*. Taip pat buvo pastebėta, kad *CAE* modelis linkęs labiau persimokyti nei *AE*.

## 6.2. Balso keitimas

Balso keitimo eksperimentų tikslas – palyginti skirtingas balso mainų architektūras. Eksperimento metu išbandysiu tris mašininio mokymosi modelius: Modeliai bus išbandyti su skirtingais vokoderiais (pagal galimybes): *MelGAN*, *WaveRNN* ir *UnivNET*.

Modelio ir vokoderio derinio našumą vertinsiu naudodamas objektyvius vertinimo rodiklius, tokius kaip *MCD* ir *LSD*.

Eksperimentas buvo pasitelktos šios bibliotekos:

- *Tensorflow* [8];
- *Pytorch* [9];
- *TTS* [19]
- *Torchaudio* [20];
- *Trainer* [21].

### 6.2.1. Duomenų rinkinys

*LJSpeech* [23] – tai viešai prieinamas duomenų rinkinys, kurį sudaro maždaug 13 100 anglų kalbos garso įrašų, kurių kiekvieno vidutinė trukmė yra apie 5 sekundės. Duomenų rinkinys buvo gautas iš *LibriVox* [23] projekto, kuris yra savanorių pastangos kurti nemokamas garso knygas iš viešai prieinamų knygų. Įrašus skaito įvairūs kalbėtojai ir juose aptariamos įvairios temos.

*LJSpeech* duomenų rinkinys buvo atsisiųstas iš oficialios saugyklos ir iš anksto apdorotas, kad jį būtų galima naudoti su mašininio mokymosi modeliais. Duomenų rinkinys buvo suskirstytas į mokymo, tikrinimo ir testavimo rinkinius santykiu 80:10:10. Mokymo rinkinys buvo naudojamas mašininio mokymosi modeliams mokyti, o tikrinimo rinkinys – hiperparametrus derinti ir ankstyvam sustabdymui. Galiausiai testavimo rinkinys buvo naudojamas apmokytų modelių veikimui įvertinti.

Duomenų aibė buvo aukštos kokybės, daugumoje įrašų kalba buvo aiški ir nuosekli. Kai kuriuose įrašuose buvo foninio triukšmo, iškraipymų ar kitų problemų, kurios galėjo turėti įtakos modelių veikimui, todėl duomenys buvo apdoroti, prieš naudojant juos modelių mokymui.

### 6.2.2. Duomenų apdorojimas

Duomenų apdorojimas vykdomas prieš kiekvieną modelio mokymo procesą ir apima šiuos veiksmus:

1. Garso failai buvo patikrinti, ar juose nėra klaidų ir neatitikimų, galinčių turėti įtakos modelių veikimui. Visi garso įrašai, kurie buvo per daug triukšmingi arba kuriuose buvo įrašymo artefaktų, buvo pašalinti, o transkripcijos klaidos atitinkamuose tekstiniuose failuose ištaisytos.
2. Garso duomenys buvo normalizuoti, siekiant užtikrinti, kad įvesties vertės patektų į tam tikrą intervalą. Tai buvo daroma siekiant užtikrinti, kad modeliai galėtų veiksmingai apdoroti duomenis. Normalizavimo procesas apėmė garso failų pergrupavimą iki 22,05 kHz diskretizavimo dažnio ir visų garso failų pradžioje ir pabaigoje esančios tylos apkarpyimą.
3. Duomenų rinkinys buvo padalytas į tris rinkinius: mokymo rinkinį (80 % duomenų), patvirtinimo rinkinį (10 % duomenų) ir testavimo rinkinį (10 % duomenų). Duomenų rinkinys buvo padalytas atsitiktine tvarka, siekiant užtikrinti, kad kiekviename rinkinyje būtų reprezentatyvi duomenų imtis. Mokymo aibės tikslas buvo apmokyti modelius, tvirtinimo aibė buvo naudojama modelio hiperparametrus derinti ir mokymo eigai stebėti, o testavimo aibė buvo naudojama apmokytų modelių veikimui vertinti naudojant nematytus duomenis.

### 6.2.3. Modelių architektūra

Balso keitimo modelių architektūros pavaizduotos žemiau esančioje lentelėje.

**6-5 lentelė.** Balso keitimo modelių architektūros

Savybės	Pasikartojanti <i>Seq2Seq</i>	Pilnai lygiagreti transformerių	<i>CVAE</i>
Pritaikymas	Kalbos sintezė	Balso dydžio/aukščio keitimas	Balso keitimas
Netekties funkcija	<i>MCD</i>	<i>MSE + F0</i> netektis	<i>SPEC + MCD</i>
Sudėtingumas	Vidutinis	Paprastas	Sudėtingas
Architektūra	Rekurentinė neuronų tinklo ( <i>RNN</i> ) briauninė architektūra	Tinklo architektūra, skirta greitam tonų pašalinimui	Konvoliucinis variacinis autoenkoderis
Duomenų reikalavimai	Tekstas ir audio failai	Tekstas ir audio failai	Tekstas ir audio failai

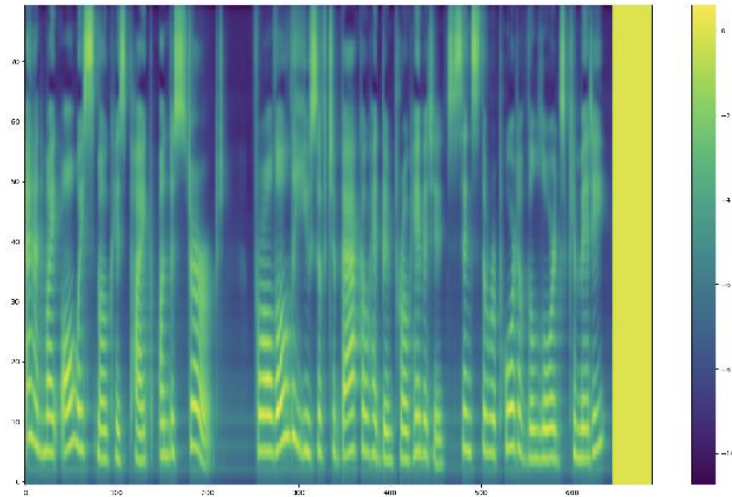
Vokoderių architektūros pavaizduotos žemiau esančioje lentelėje.

**6-6 lentelė.** Balso keitimo vokoderių architektūros

Savybės	<i>WaveRNN</i>	<i>MelGAN</i>	<i>UnivNET</i>
Architektūra	<i>RNN</i> (Rekurentinis neuroninis tinklas) su spektrogramos perdavimu į garsą	<i>GAN</i> (generatyvi konkurencija) su signalo perdavimu į garsą	Giliojo konvoliucinio neuroninio tinklo architektūra
Tikslus vertinimas	<i>MCD, MOS</i>	<i>MCD, MOS</i>	<i>SFTP</i>
Lankstumas	Gali būti naudojamas su kitais TTS modeliais, bet reikalauja spektrogramos kaip įvesties duomenų.	Gali būti naudojamas su kitais TTS modeliais, bet reikalauja signalo kaip įvesties duomenų.	Gali būti naudojamas su kitais TTS modeliais, bet reikalauja spektrogramos kaip įvesties duomenų.

#### 6.2.3.1. Pasikartojanti *Seq2Seq*

Pasikartojanti *Seq2Seq* – tai gilaus mokymosi architektūros modelis, kuris priima teksto seką kaip įvestį ir išveda atitinkamą spektrogramų seką, kuri atspindi kalbos bangos formos dažnio turinį laikui bėgant. Spektrogramas galima įsivaizduoti kaip 2D kalbos atvaizdus, kur x ašis rodo laiką, o y ašis – dažnį **6-9 pav.** Spektrogramos pavyzdys *Seq2Seq* modelis mokomas pagal įvesties teksto ir tikslinių spektrogramų poras, kad išmokytų sukurti tinkamą spektrogramą tam tikram įvesties tekstui [40].



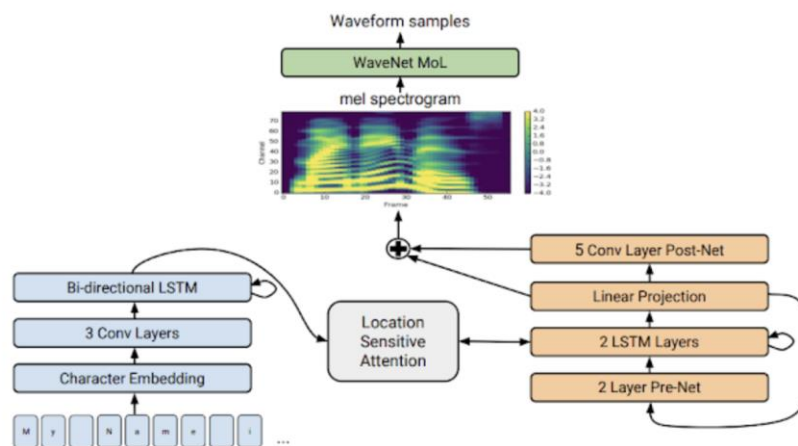
**6-9 pav.** Spektrogramos pavyzdys

Kai *Seq2Seq* sukuria spektrogramų seką, kitas žingsnis – paversti jas į bangos formą, kurią galima atkurti kaip kalbą. Tam pasitelkiamas vokoderio modelis. Vokoderis yra neuroninio tinklo tipas, kuris priima spektrogramą kaip įvestį ir išveda atitinkamą bangos formą.

*Seq2Seq* modelį sudaro kodavimo įrenginys, dėmesio mechanizmas ir dekoduoja, kurie pavaizduoti 6-10 pav.:

- Kodavimo įrenginys priima teksto seką ir, naudodamas konvoliucinių sluoksnių steką, sukuria aukšto lygio įvesties atvaizdavimą;
- Dėmesio mechanizmas naudoja užkoduotą tekstą ir ankstesnį dekoduoja išėjimą, kad apskaičiuotų konteksto vektorius, kuriame užfiksuotos svarbiausios įvesties dalys dabartiniame laiko žingsnyje;
- Dekoduoja naudoja konteksto vektorius ir ankstesnio dekoduoja išvestį kitam mel-spektrogramos kadrai nuspėti.

Modelis mokomas naudojant vidutinės kvadratinės paklaidos (*MSE*) ir dvejetainės kryžminės entropijos (*BCE*) nuostolių derinį tarp prognozuojamos mel-spektrogramos ir pagrindinės tiesos. Mokymo metu naudojamas mokytojo priverstinis vertimas, t. y. kiekvienu laiko žingsniu į modelį įvedama pagrindinė tiesa, o ne ankstesnio dekoduoja išvestis.

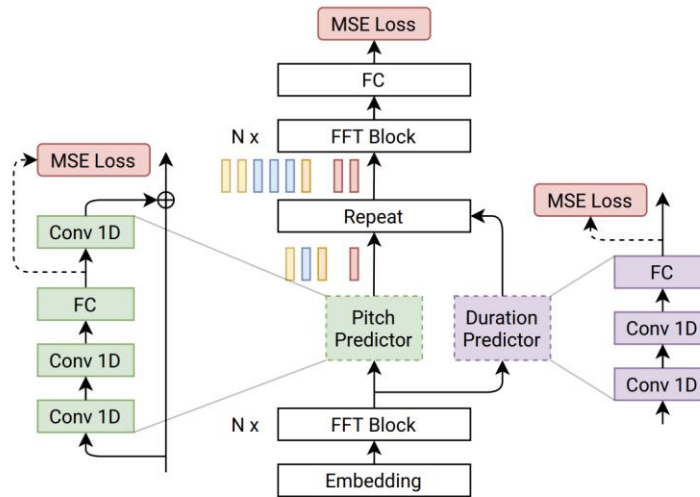


6-10 pav. Seq2Seq architektūros schema [41]

### 6.2.3.2. Pilnai lygiagreti transformerių architektūra

Modelyje naudojama pilnai lygiagreti transformerių architektūra su garso aukščio prognozavimo moduliu, kad būtų sukurta natūraliai skambanti prozodija pasižyminti kalba. Modelio architektūrą, kuri pavaizduota 6-11 pav., galima suskirstyti į kelis pagrindinius komponentus [42]:

- Teksto koduotojas (angl. *embedding*) – įvesties tekstas pirmiausia perduodamas *LSTM* sukurti aukšto lygio paslėptų reprezentacijų seką, kuri atspindi semantinę ir sintaksinę įvesties teksto informaciją.
- Skambesio prognozuotojas (angl. *pitch predictor*) – aukščio prognozavimo modulis priima užkoduotą teksto atvaizdavimą kaip įvestį ir numato generuojamos kalbos aukščio kontūrą. Tai atliekama naudojant *LSTM*, po kurio seka *multi-head* savaiminio dėmesio mechanizmas. Aukščio kontūras yra pagrindinių dažnių seka, valdanti kalbos bangos aukštį, ir jis naudojamas generuojamos kalbos prozodijai moduluoti.
- Trukmės prognozuotojas (angl. *duration predictor*) – numato kiekvienos fonemos trukmę įvesties tekste, kuri naudojama generuojamos kalbos greičiui valdyti. Trukmės prognozavimo modulis kaip įvestį priima tą pačią užkoduotą teksto reprezentaciją ir prognozuoja kiekvienos fonemos trukmę naudodamas vieno sluoksnio dvikryptį *LSTM*, po kurio seka kelių galvučių savasties mechanizmas. Tada prognozuojamos trukmės naudojamos akustinio dekodavimo generuojamoms mel-spektrogramoms suderinti pagal laiką, taip užtikrinant, kad generuojama kalba būtų tinkamo laiko ir ritmo.
- Akustinis dekodavimo modulis – akustinis dekodavimo modulis priima užkoduotą teksto atvaizdavimą ir garso aukščio kontūrą kaip įvestį ir generuoja akustinius požymius, atitinkančius kalbos bangos formą. Dekoduotoją sudaro vienakrypčiai *LSTM* sluoksniai, kurie yra įtakojami aukščio kontūro, naudojant dėmesio mechanizmą. Dekoduotojo išvestis yra mel-spektrogramų seka, kuri atspindi kalbos signalo dažnių turinį.
- Transformeriai, kurių tinklą sudaro savarankiško dėmesio ir tiesioginio kreipimo sluoksniai, leidžia modeliui fiksuoti tolimąsias priklausomybes ir konteksto informaciją akustinėse savybėse. Pirmas veikia įvesties ženklų skiriamąja gėba, kad užkoduotų įvesties tekstą, o antrasis – išvesties kadrų skiriamąja gėba, kad iškoduotų akustines kalbos ypatybes.



6-11 pav. Transformerių architektūros schema [43]

Mokymo procesas pradamas nuo įvesties teksto pirminio apdorojimo ir žymėjimo. Tuomet iš pagrindinio garso įrašo sukuriamos mel-spektrogramos ir garso aukščio kontūrai. Garso aukščio modelis mokomas nuspėti garso aukščio kontūrus iš pagrindinio garso įrašo, o šie nuspėti garso aukščio kontūrai naudojami garso aukščio kontūrams generuoti iš nuspėtų mel-spektrogramų. Trukmės modelis yra apmokytas prognozuoti kiekvienos fonemos trukmę iš teksto įvesties. Tuomet mel-spektrogramų modelis mokomas generuoti mel-spektrogramas iš prognozuojamo garso aukščio ir trukmės. Modelio svoriai atnaujinami pagal apskaičiuotą nuostolį kiekviename mokymo etape, o procesas kartojamas tol, kol apdorojami visi mokymo pavyzdžiai ir nuostolis konverguoja.

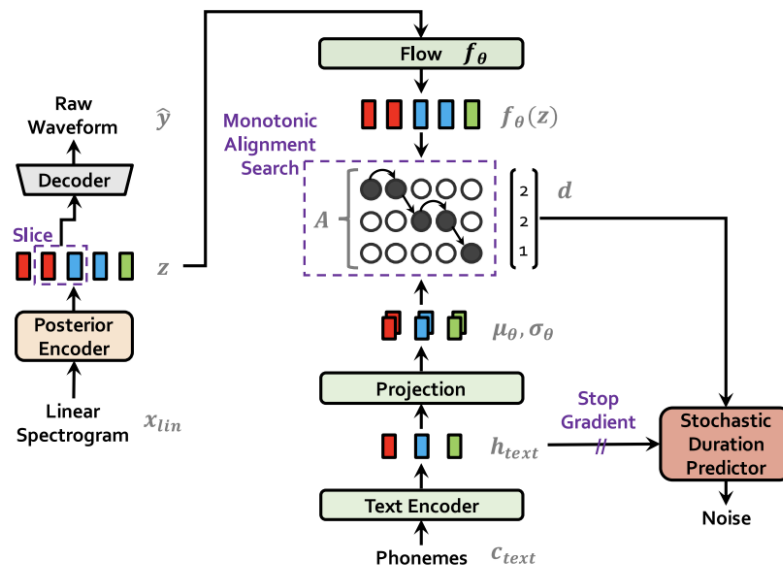
### 6.2.3.3. CVAE

Tai teksto sintezės modelis, kuriame naudojamas sąlyginis variacinis autokoderis (*CVAE*) su priešpriešiniu mokymusi. Šis modelis skirtas aukštos kokybės ir natūraliai skambančiai kalbai iš teksto įvesties generuoti [44].

*CVAE* mokymo procesas apima kelių komponentų, įskaitant teksto koduotoją, kalbėtojo koduotoją, *CVAE* ir diskriminatorių, mokymą naudojant MSE, priešingų nuostolių ir rekonstrukcijos nuostolių derinį.

Šį modelį, kurio architektūra pavaizduota 6-12 pav., sudaro trys pagrindinės sudedamosios dalys:

- Teksto kodavimo modulis, kuris priima tekstą kaip įvestį ir išveda užkoduotą teksto atvaizdavimą, naudodamas transformatoriais pagrįstą architektūrą.
- Kodavimo modulis (angl. *posterior encoder*), kuris priima tikslinės kalbos mel-spektrogramos ir prognozuojamos dabartinės išvesties mel-spektrogramos sandūrą ir sukuria koduotą kalbos požymių tikimybių atvaizdavimą.
- Srautu pagrįstas generatyvinis modelis (angl. *residual coupling blocks*), kuris priima užkoduotas tikimybes ir generuoja prognozuojamą kalbos išvesties mel-spektrogramą.
- Trukmės generavimo modelis (angl. *duration predictor*), numatantis kiekvienos kalbos ypatybės trukmę, atsižvelgiant į užkoduotą teksto atvaizdavimą ir prognozuojamą mel-spektrogramą.
- Generatorius, kuris kaip įvesties duomenis priima prognozuojamą mel-spektrogramą, informaciją apie trukmę ir triukšmą ir išveda atitinkamą kalbos bangos formą.



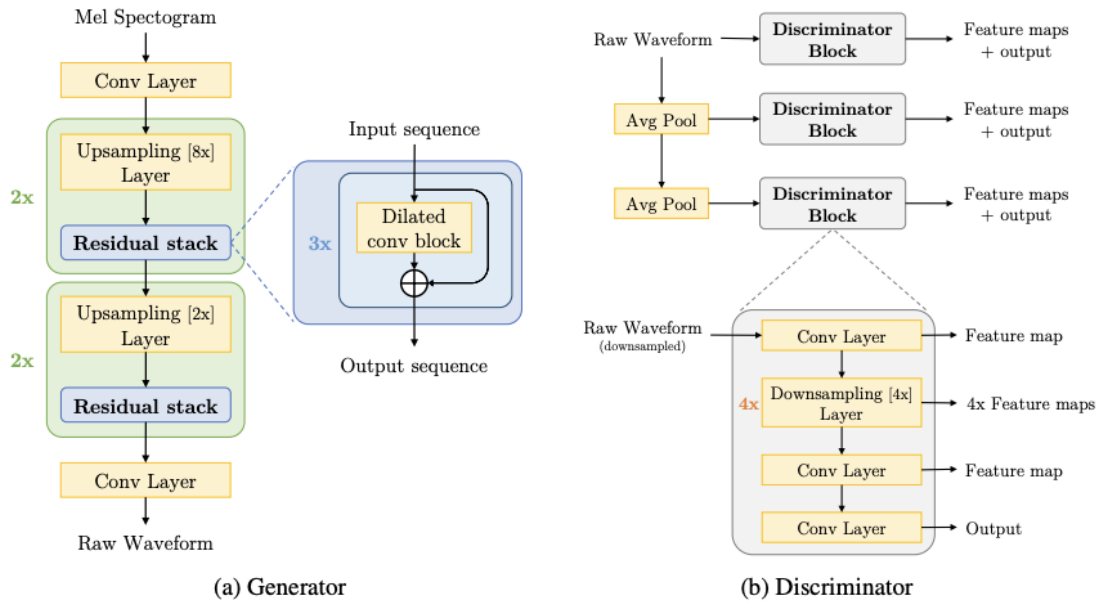
6-12 pav. CVAE architektūros schema [44]

Siekiant pagerinti generuojamos kalbos kokybę, į mokymo procesą įtraukiami priešingi nuostoliai. Diskriminacinis tinklas mokomas atskirti sugeneruotas mel-spektrogramas nuo realių mel-spektrogramų iš mokymo duomenų. Tada generatoriaus tinklas, kurį sudaro teksto koduotojas, kalbėtojo koduotojas ir dekoduoja, mokomas taip, kad diskriminatoriaus nuostoliai būtų kuo mažesni, o generuojamos mel-spektrogramos būtų kuo artimesnės tikrosioms mel-spektrogramoms.

#### 6.2.3.4. GAN vokoderis

*MelGAN* yra generatyvinio priešpriešinio tinklo architektūra, skirta aukštos kokybės garso bangų formoms generuoti, atsižvelgiant į mel-spektrogramas, kurios dažniausiai naudojamos garso signalams atvaizduoti kalbos ir muzikos apdorojimo procesuose [45].

Šią architektūrą sudaro generatoriaus ir diskriminatoriaus tinklai, kurie yra apmokyti priešpriešiniu būdu, kad generuotų aukštos kokybės garso bangų formas, neatskiriamas nuo tikrų garso pavyzdžių. Generatoriaus tinklas priima mel-spektrogramą ir generuoja atitinkamą garso bangos formą, o diskriminatoriaus tinklo tikslas – atskirti tikrus ir generuotus garso pavyzdžius.



6-13 pav. GAN architektūros schema [45]

Generatoriaus tinklas yra paremtas transponuotais konvoliuciniams sluoksniams, kurie naudojami mel-spektrogramai padidinti iki pageidaujamo garso bangos formos ilgio. Po transponuotų konvoliucinių sluoksnių seka eilė likutinių blokų, kurie skirti padėti generatoriui modeliuoti ilgalaikės garso signalo priklausomybes. Paskutiniame generatoriaus tinklo sluoksnyje sugeneruotai garso bangos formai taikoma aktyvavimo funkcija.

Diskriminatoriaus tinklas pagrįstas konvoliuciniams sluoksniams, kurie naudojami garso bangos formai sumažinti iki mažo matmens požymių atvaizdavimo. Tada sumažintos savybės perduodamos per keletą visiškai sujungtų sluoksnių, kad būtų apskaičiuotas skaliarinis išėjimas, rodantis, ar įvesties garso pavyzdys yra tikras, ar sugeneruotas.

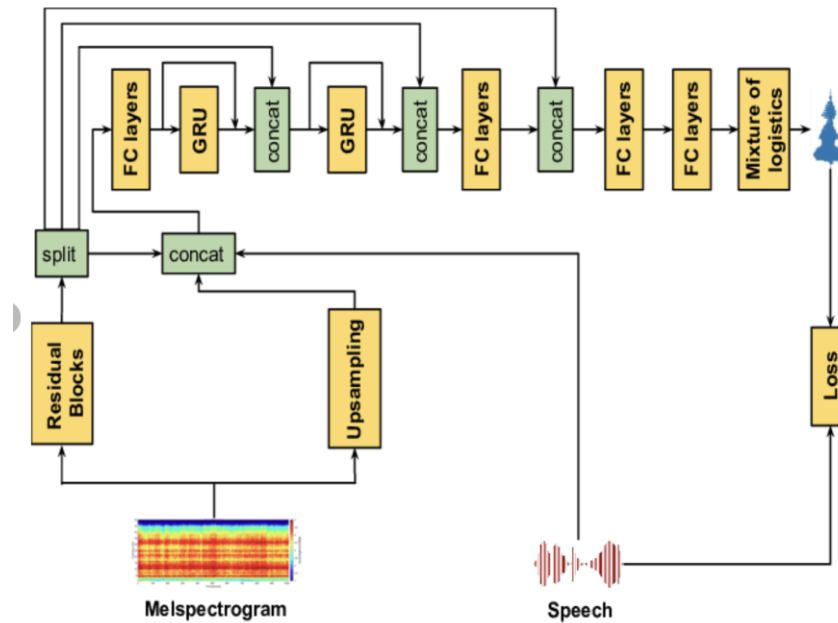
Mokymo metu generatoriaus tinklas mokomas generuoti garso bangų formas, kurios nesiskiria nuo tikrų garso pavyzdžių, o diskriminatoriaus tinklas mokomas atskirti tikrus ir generuotus garso pavyzdžius. Šie du tinklai mokomi priešpriešiniu būdu, kai generatoriaus tinklas bando generuoti garso bangų formas, kurios apgauna diskriminatorių, o diskriminatoriaus tinklas bando teisingai atskirti tikrus ir sugeneruotus garso pavyzdžius.

### 6.2.3.5. RNN vokoderis

*WaveRNN* – tai neuroninis garso sintezės modelis, kuris generuoja aukštos kokybės garso pavyzdžius. Modelis pagrįstas pasikartojančio neuroninio tinklo architektūra, kuri modeliuoja bendrą garso bangos formos ir jos atitinkamo kvantinio atvaizdavimo pasiskirstymą.

*WaveRNN* susideda iš dviejų pagrindinių komponentų: sąlyginio tinklo ir pasikartojančio tinklo. Sąlyginis tinklas priima pageidaujamą garso įrašą apibūdinančių požymių seką, pavyzdžiui, lingvistinius ar muzikinius įvesties duomenis. Pasikartojantis tinklas pateikia paslėptų būsenų rinkinį, kuris naudojamas imties generavimo tinklui nustatyti [46].





6-14 pav. RNN vokoderio architektūros schema [47]

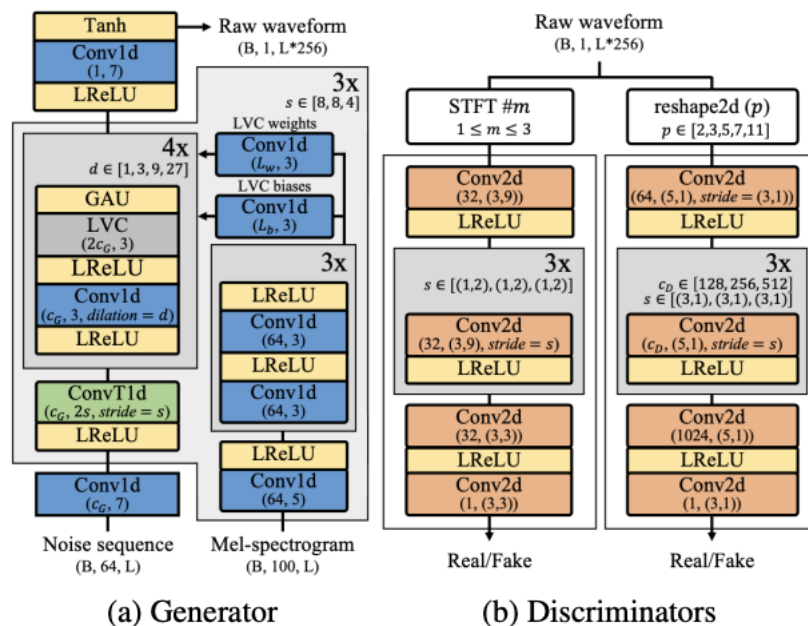
Pasikartojantis neuroninis tinklas generuoja garso bangos formą pavyzdys po pavyzdžio. Kiekvienu laiko žingsniu tinklas kaip įvestį priima ankstesnį garso pavyzdį ir atitinkamą kvantuotą garso signalo atvaizdavimą. Tinklas taip pat priima paslėptas būsenas iš kondicionavimo tinklo kaip papildomą įvestį. Tada tinklas generuoja kitą garso pavyzdį, numatydamas kitos kvantuotos vertės tikimybinį pasiskirstymą.

Kad generuotų aukštos kokybės garso įrašus, *WaveRNN* naudoja metodą, vadinamą autoregresiniu modeliavimu, kai kiekvienas pavyzdys generuojamas atsižvelgiant į anksčiau generuotus pavyzdžius. Modelyje taip pat naudojama hierarchinė kvantavimo schema, kad ištisinis garso signalas būtų diskretizuotas į mažesnę galimų reikšmių rinkinį. Tai padeda sumažinti galimų išvesties reikšmių skaičių, todėl generavimo procesas tampa efektyvesnis.

### 6.2.3.6. UnivNET vokoderis

*UnivNet* yra neuroninis vokoderis, kuris priima mel-spektrogramą, ir generuoja didelio tikslumo garso bangos formą. *UnivNet* sukurta generuoti aukštos kokybės garsą mažomis skaičiavimo sąnaudomis ir pagal garso kokybę gali lenkti anksčiau aprašytus neuroninius vokoderius [48].

*UnivNet* sudaro trys pagrindiniai komponentai: kodavimo tinklas, dekodavimo tinklas ir daugialypės skiriamosios gebos spektrogramų diskriminatoriai. Kodavimo tinklas priima laiko srities garso signalą kaip įvestį ir sukuria mel-spektrogramą. Dekodavimo tinklas priima koduotojo sukurtą mel-spektrogramą ir sukuria didelės išraiškos garso bangos formą. Daugialypės skiriamosios gebos spektrogramų diskriminatoriai mokymo metu teikia grįžtamąjį ryšį kodavimo ir dekodavimo tinklams, padėdami užtikrinti, kad sukurtas garsas būtų aukštos kokybės ir neatskiriamas nuo tikro garso.



6-15 pav. UnivNET architektūros schema [48]

UnivNet kodavimo tinklas naudoja išplėstines konvoliucijas požymiams iš garso signalo išgauti. Dekoduotojo tinkle garso bangos formai generuoti naudojami apverčiamieji neuroniniai tinklai. Ir kodavimo, ir dekodavimo tinklai sukurti taip, kad būtų skaičiavimo požiūriu efektyvūs, todėl *UnivNet* gali generuoti aukštos kokybės garsą realiuoju laiku.

Daugialypės skiriamosios gebos spektrogramos diskriminatoriai *UnivNet* sukurti taip, kad grįžtamasis ryšys būtų teikiamas skirtinguose spektrogramos lygiuose, todėl tinklas gali išmokti generuoti aukštos kokybės garsą plačiame dažnių diapazone. Diskriminatorius sudaro konvoliucinių neuroninių tinklų, veikiančių skirtinguose spektrogramos dažnių diapazonuose, rinkinys. Mokymo metu diskriminatoriai teikia grįžtamąjį ryšį kodavimo ir dekodavimo tinklams, lygindami generuojamą garso bangos formą su pagrindine garso bangos forma.

#### 6.2.4. Modelių mokymas

Žemiau esančioje lentelėse palygintas modelių bei vokoderių mokymas.

6-7 lentelė. Veido keitimo modelių mokymai

Architektūra	Pasikartojanti <i>Seq2Seq</i>	Pilnai lygiagreti transformerių	<i>CVAE</i>
Mokymosi greitis	pradedama nuo 0,001 ir kas 10000 žingsnių sumažinama perpus	pradedama nuo 0,001 ir kas 10000 žingsnių sumažinama perpus	pradedamas nuo 0,0005 ir mažinamas taikant kosinusinio atkaitinimo grafiką.
Partijos dydis	32	32	32
Optimizavimo algoritmas	Adam optimizatorius, kurio $\beta_1 = 0,9$ , $\beta_2 = 0,98$ ir $\epsilon = 10e-9$		
Nuostolių funkcija	naudoja įvairių nuostolių funkcijų derinį, pavyzdžiui, vidutinės kvadratinės klaidos (MSE) nuostolį tiesinėms spektrogramoms, dvejetainės kryžminės entropijos (BCE) nuostolį stop ženklų nuspėjimui ir trukmės nuostolį.		

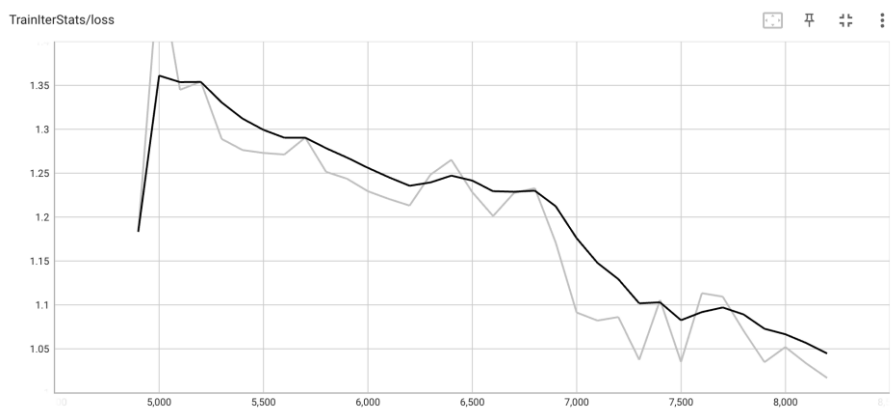
## 6-8 lentelė. Veido keitimo vokoderių mokymas

Savybės	<i>WaveRNN</i>	<i>MelGAN</i>	<i>UnivNET</i>
Architektūra	<i>RNN</i> (Rekurentinis neuroninis tinklas) su spektrogramos perdavimu į garsą	<i>GAN</i> (generatyvi konkurencija) su signalo perdavimu į garsą	Giliojo konvoliucinio neuroninio tinklo architektūra
Optimizavimo algoritmas	Adam optimizatorius	Adam optimizatorius	Adam optimizatorius
Netekties funkcija	Binarinės kryžminės entropijos nuostoliai	<i>MSE</i>	<i>MSE</i>
Mokymosi sparta	1e-4 generatoriui ir 4e-5 diskriminatoriui	1e-4	1e-4
Partijos dydis	32	32	32
Optimizavimo algoritmas	Adam optimizatorius su beta1=0,5, beta2=0,999 ir svorio mažėjimu 1e-6		
Mokymosi kriterijai	Mokymosi kriterijai: vertinamas kas 1000 žingsnių naudojant patvirtinimo rinkinį, o mokymas nutraukiamas, jei patvirtinimo nuostoliai nesumažėja per 5 vertinimus iš eilės.		

Žemiau pavaizduoti ir aprašyti modelių bei vokoderių mokymosi rezultatai. Grafai paimti naudojant „*tensorboard*“ biblioteką. Grafuose x ašis atitinka epochų skaičių, y ašis – skaitinį rezultatą (nuostolius arba laiką).

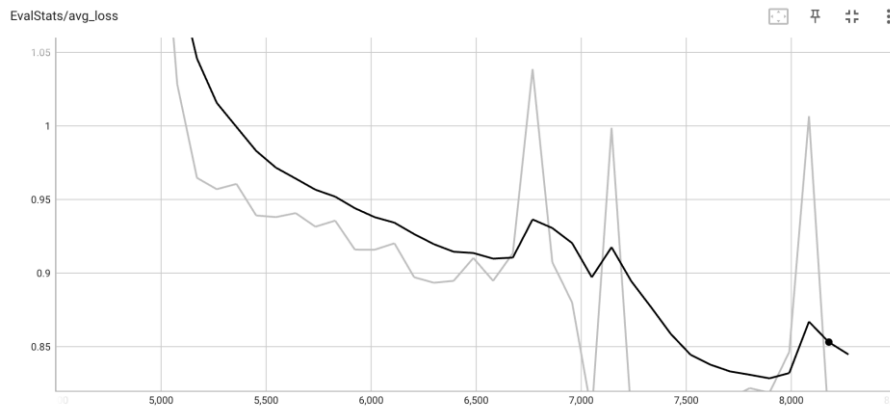
### 6.2.4.1. Pasikartojanti Seq2Seq mokymo rezultatai

Šios architektūros modelis buvo treniruotas per du kartus, pirmą kartą iki 5000 epochos. Tęsiant mokymą, nuostoliai pakilo, tačiau greitai susinormalizavo ir žemėjo. Mažiausias treniravimosi nuostolis 1,05921.



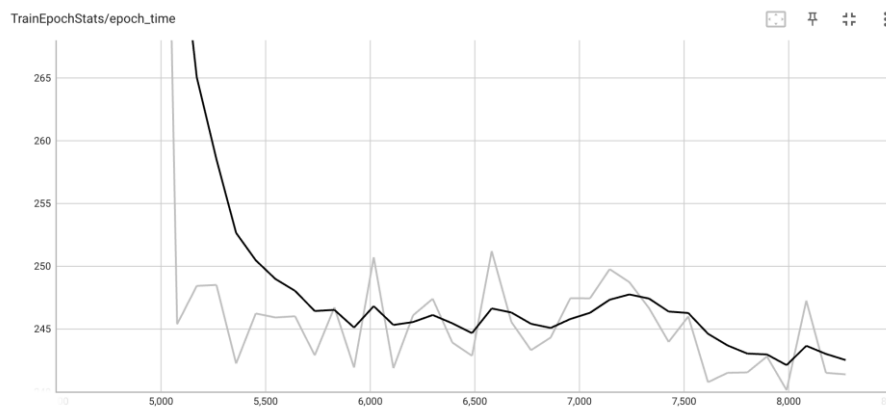
6-16 pav. *Seq2Seq* mokymosi nuostoliai

Validuojant modelį mokymosi metu nuostoliai taip pat mažėjo ir laikėsi žemesni nei treniravimosi.



**6-17pav.** *Seq2Seq* validavimo nuostoliai

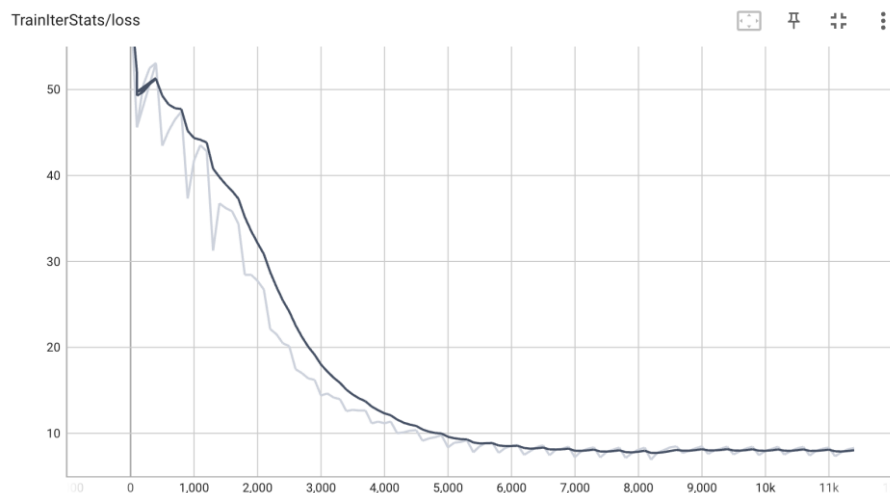
Žemiau pavaizduotas grafas vaizduoja epochos laiką, kuris vidutiniškai yra apie 3 minutes.



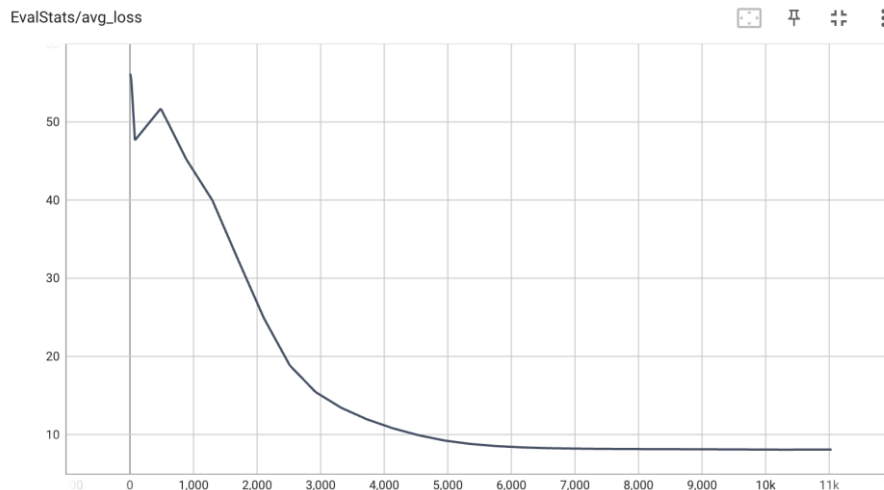
**6-18 pav.** *Seq2Seq* epochos laikas

#### 6.2.4.2. Pilnai lygiagreti transformerių mokymo rezultatai

Iš mokymosi ir validavimo nuostolių diagramų galime matyti, jog modelis sulėtėjo mokytis nuo 5000 epochos, nustojo mokytis nuo 8000 epochos.

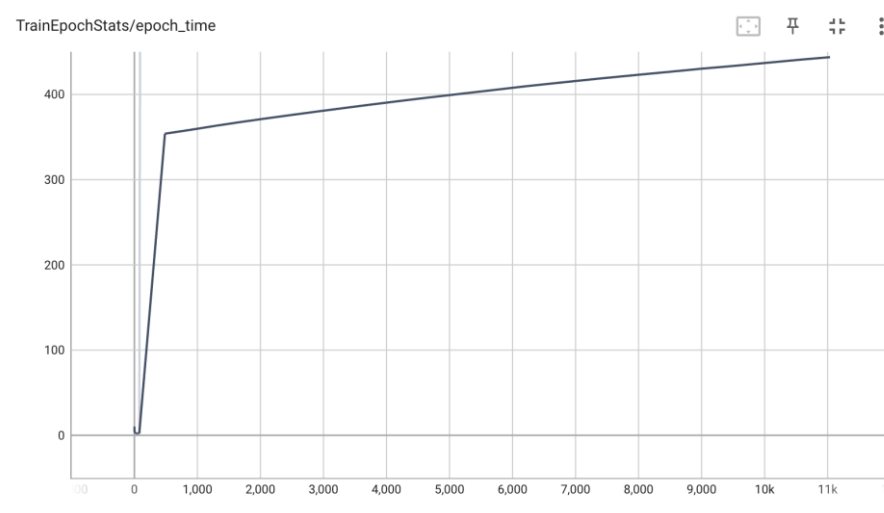


**6-19 pav.** Transformerių mokymosi nuostoliai



**6-20 pav.** Transformerių validavimo nuostoliai

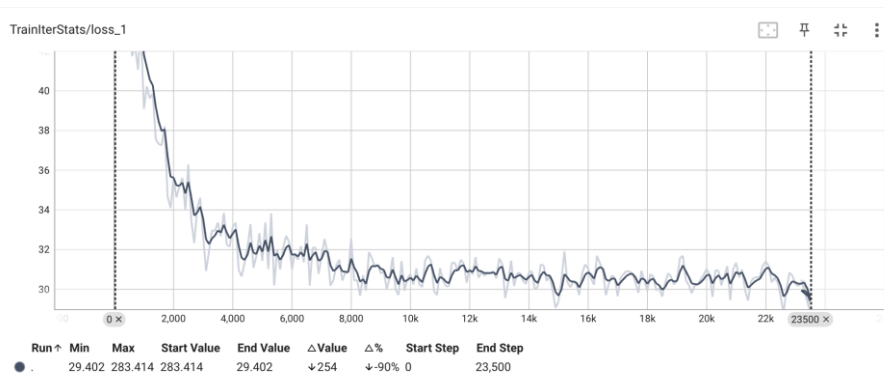
Taip pat viena epocha vidutiniškai trunka nuo 6 minučių ir jos laikas iš lėto auga.



**6-21 pav.** Transformerių epochos laikas

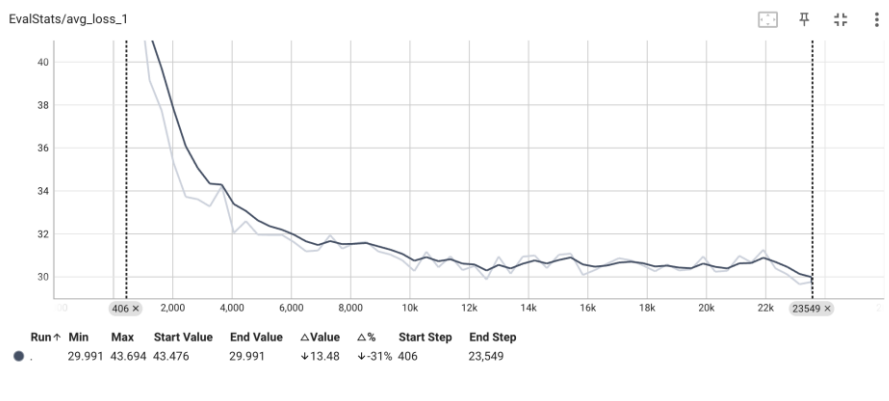
### 6.2.4.3. CVAE mokymo rezultatai

Šios architektūros modelis mokosi iki 8000 epochos (mokymosi nuostolis mažėja).



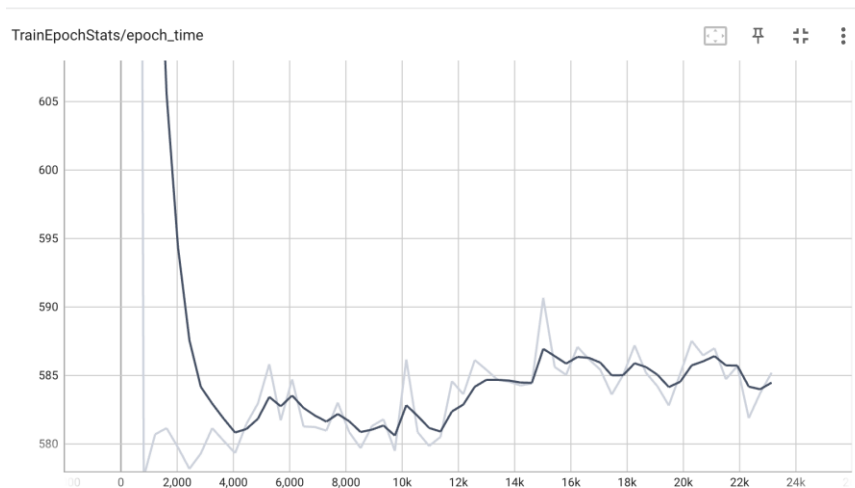
**6-22 pav.** CVAE mokymosi nuostoliai

Tačiau validacijos nuostolis mažėja iki pat mokymosi pabaigos – šiuo atveju iki 12000 epochos. Dėl to galime teigti, jog modelis gerai atpažįsta naujus modelius bei gali gebėti geriau pakeisti balsą nei kiti modeliai.



**6-23 pav.** CVAE validavimo nuostoliai

Nors modelis geriau susidorėjo su balso keitimu, tačiau jo epochos laikas ilgesnis – iki 10 minučių.



**6-24 pav.** CVAE epochos laikas

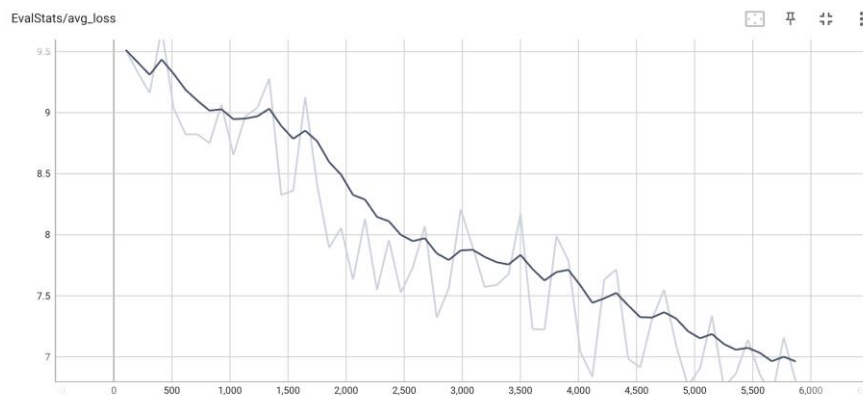
#### 6.2.4.4. RNN vokoderis

RNN vokoderis mokosi lėtai, tačiau stabiliai. Deja per beveik 6000 epochų jis nepasiekė gerų nuostolio rezultatų.



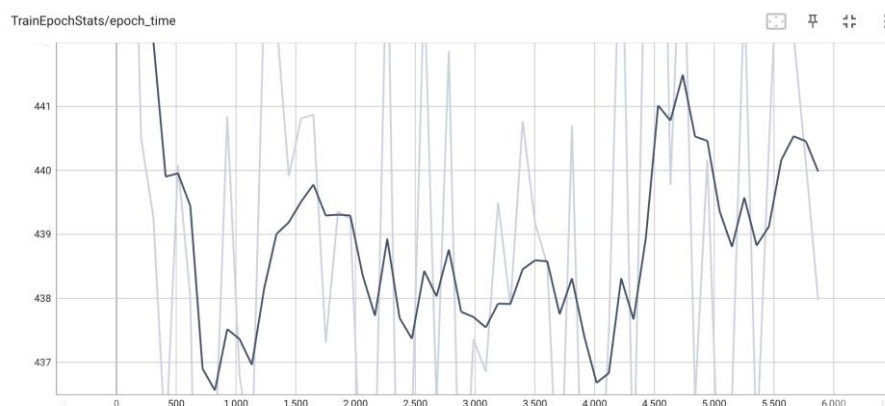
6-25 pav. RNN vokoderio mokymosi nuostoliai

Validavimo nuostoliai labai panašūs į treniravimosi, todėl galima manyti, jog modelis mokosi kokybiškai.



6-26 pav. RNN vokoderio validavimo nuostoliai

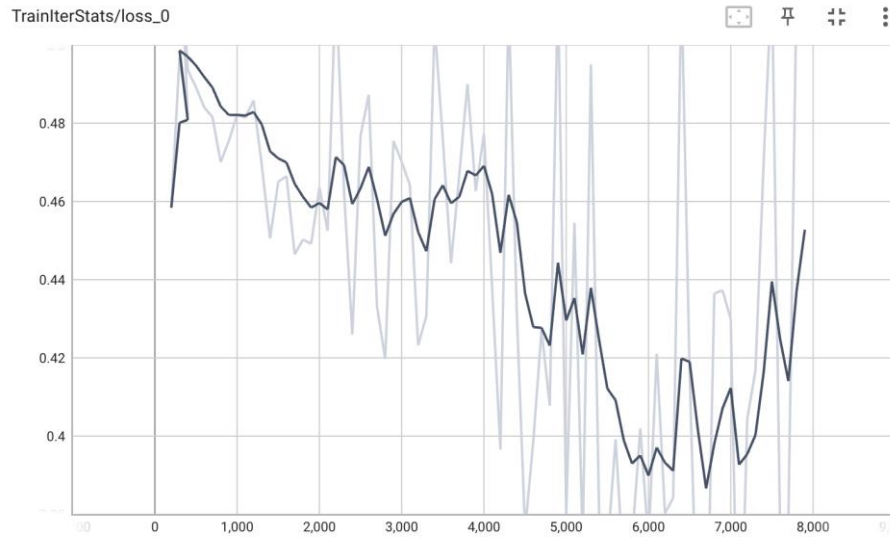
Vienos epochos treniravimosi laikas apie 7-8 minutes.



6-27 pav. RNN vokoderio epochų laikai

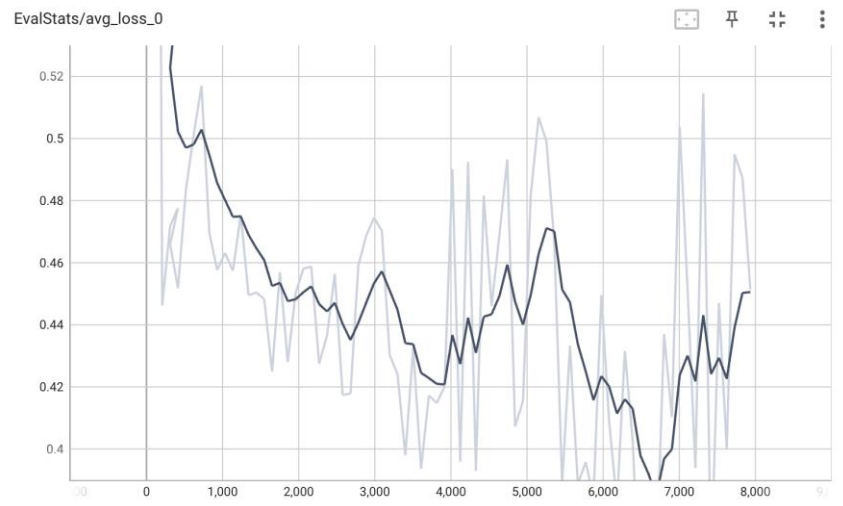
#### 6.2.4.5. GAN vokoderis

Žemiau pavaizduoto mokymasis parodo, jog modelis greitai pasiekia gana gerus rezultatus, tačiau, nuo 6300 epochos pradeda persimokyti.



**6-28 pav.** GAN vokoderio mokymosi nuostoliai

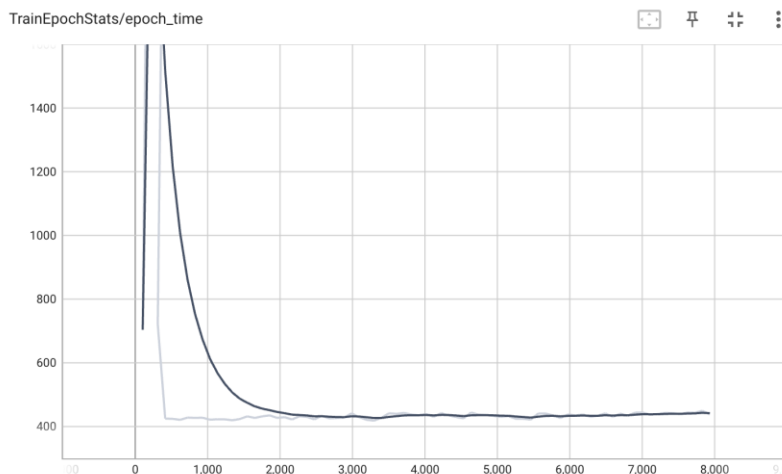
Iš validavimo grafo galima teigti, jog modelis pradeda persimokyti nuo 7000 epochos.



**6-29 pav.** GAN vokoderio validavimo nuostoliai

Viena modelio epocha vidutiniškai trunka apie 7 minutes.

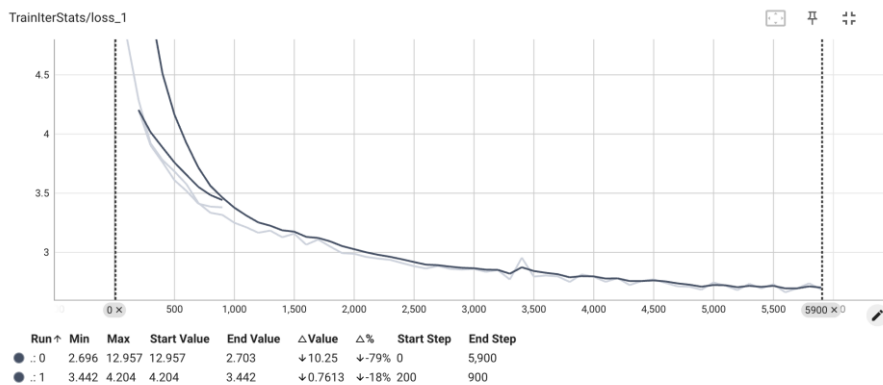




6-30 pav. GAN vokoderio epochų laikai

### 6.2.4.6. UnivNET vokoderis

Modelis sugebėjo mokyti iki 6000 epochos, po kurios sulūžo ir iš naujo pradėjo skaičiuoti epochas, todėl bus vertinamas modelis iki šios epochos.



6-31 pav. UnivNET mokymosi nuostoliai

Validavimo nuostoliai labai panašūs į mokymosi nuostolius, todėl galima teigti, jog iki 6000 epochos modelis mokėsi gerai. Žinoma, jo mokymosi greitis sumažėjo, tačiau jis toliau mokėsi.



6-32 pav. UnivNET validavimo nuostoliai

Modelio mokymosi epocha vidutiniškai trunka apie 9 minutes. Matoma, jog nuo tada, kai modelis nustojo veikti tinkamai, jo vidutinis epochos laikas pailgėjo.



6-33 pav. UnivNET epochų laikai

### 6.2.5. Modelių vertinimas

Modeliai bei vokoderiai yra lyginami pagal tai, kaip gerai sugeba įgarsinti/pakeisti balsą. Norint naudoti modelį balso keitimui, naudojamas stiliaus perdavimas, su kuriuo galima keisti sintezuotos kalbos prozodiją, įtraukiant informaciją iš išorinio garso failo, kuriame yra norimos prozodijos etaloninis balsas.

Balso keitimo procesas naudojant šiuos modelį:

1. Įvesties tekstas pirmiausia konvertuojamas į mel-spektrogramą naudojant apmokytą balso keitimo modelį. Mel-spektrogramoje kalba pateikiama dažnių ir laiko srityje.
2. Balso stiliaus modelis priima spektrogramas ir tikslo (stiliaus) failą kaip įvesties duomenis ir pateikia stiliaus įterpimo vektorius, kuriame koduojama norima prozodija iš tikslo failo.
3. Sukurtas stiliaus įterpimo vektorius sujungiamas su spektrogramomis ir gaunama stiliaus įterptos spektrogramos. Šiose spektrogramose yra ir įvesties teksto turinys, ir prozodijos informacija iš tikslo failo.
4. Tada sugeneruota mel-spektrograma konvertuojama į garso bangos formą naudojant vokoderį. Vokoderis priima mel-spektrogramą kaip įvestį ir generuoja bangos formą, kuri tiksliai atitinka mel-spektrogramos charakteristikas.

Šiame tyrime vertinant balso keitimo modelių ir vokoderių našumą, bus naudojamos *Mel Cepstral Distortion (MCD)* ir *Log-Spectral Distortion (LSD)* metrikos [49]. *MCD* ir *LSD* balai bus apskaičiuojami tarp originalios ir sintezuotos kalbos pavyzdžių. Be to, bus atliekami subjektyvūs klausymo testai, siekiant įvertinti sintezuotos kalbos natūralumą ir panašumą į tikslinį balsą. Šių metrikų palyginimas – lentelėje.

6-9 lentelė. Veido keitimo tikrinimo metrikos

Metrika	<i>MCD</i>	<i>LSD</i>
Procesas	Apskaičiuojamas kaip euklidinio atstumo tarp dviejų signalų	Matuoja vidutinį kvadratinį skirtumą tarp dviejų signalų log-spektrų.
Interpretacija	Mažesnis reikšmė rodo geresnę sintezės kokybę, nes jis reiškia, kad dviejų garso signalų spektrai yra artimesni vienas kitam.	Mažesnė reikšmė rodo, kad tikrasis signalas ir prognozuojamas signalas yra panašesni. Daugiausia naudojamas vertinant garso perdavimo kokybę per skaitmeninę liniją (pavyzdžiui, internetu ar per radiją).

Privalumai	Gerai žinomas ir dažnai naudojamas matas kalbos sintezės srityje.	Gerai tinka vertinti garso perdavimo kokybę, ypač kai garso signalas skaitmeninėje linijoje yra papildomai suskaidytas.
Ribojimai	Vertinimas yra labai priklausomas nuo garso signalo garsiųjų. Tai reiškia, kad ne visada yra patikimas vertinant garso kokybę, kai garso signalo garsiųjų lygis skiriasi.	jautrus skirtingų dažnių riboms, o tai gali lemti nesutapimų ir neteisingų rezultatų atveju, kai signalų spektrai yra iškraipyti.
Naudojimo atvejis	Dažnai naudojamas kalbos sintezės kokybės vertinimui ir perfekcionavimui.	Plačiai naudojamas garso perdavimo kokybės vertinimui per skaitmenines linijas.

### 6.2.6. Rezultatai

Žemiau esančiose lentelėse pavaizduoti balso keitimo modulių bei vokoderių mokymosi rezultatai. Pilni rezultatai pavaizduoti trečiame priede.

**6-10 lentelė.** Veido keitimo modelių mokymo rezultatai

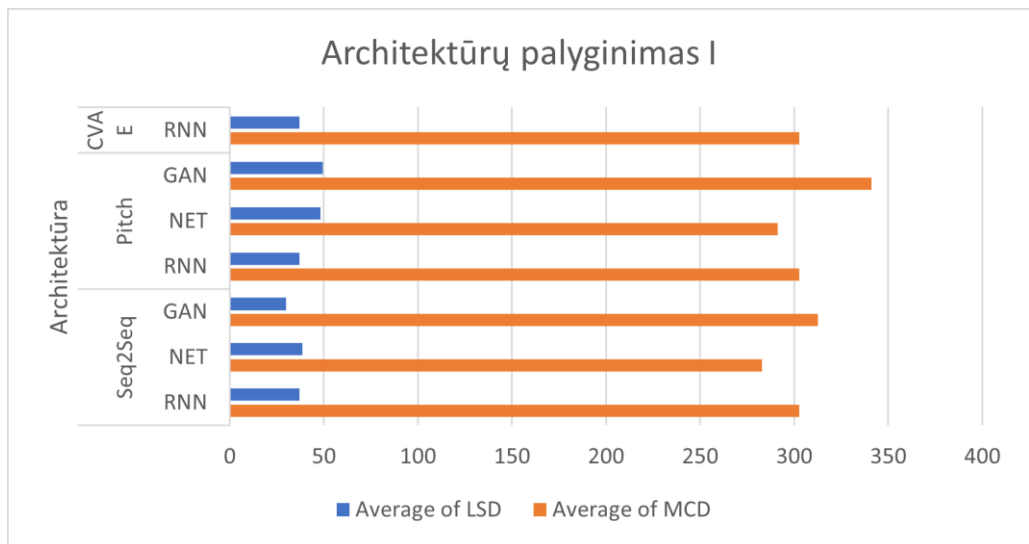
Architektūra	Pasikartojanti <i>Seq2Seq</i>	Pilnai lygiagreti transformerių	<i>CVAE</i>
Epochų kiekis	8300	8	29,4
Mokymosi nuostoliai	1,02	8	29,9
Validavimo nuostoliai	0,87	11000	23500
Nustojo mokytis	-	6000	12000
Vidutinis epochos laikas	3	6	10
Modelio dydis (MB)	347	436	951
Hiperparametrų kiekis	30339130	37517410	83059180

**6-11 lentelė.** Veido keitimo vokoderių mokymo rezultatai

Savybės	<i>WaveRNN</i>	<i>MelGAN</i>	<i>UnivNET</i>
Epochų kiekis	5900	8000	5900
Mokymosi nuostoliai	7.4	0,39	2.7
Validavimo nuostoliai	7	0,39	2.7
Nustojo mokytis	-	7000	-
Vidutinis epochos laikas	7	7	6
Modelio dydis (MB)	46	79	328
Hiperparametrų kiekis	4233673	6894446	56232399

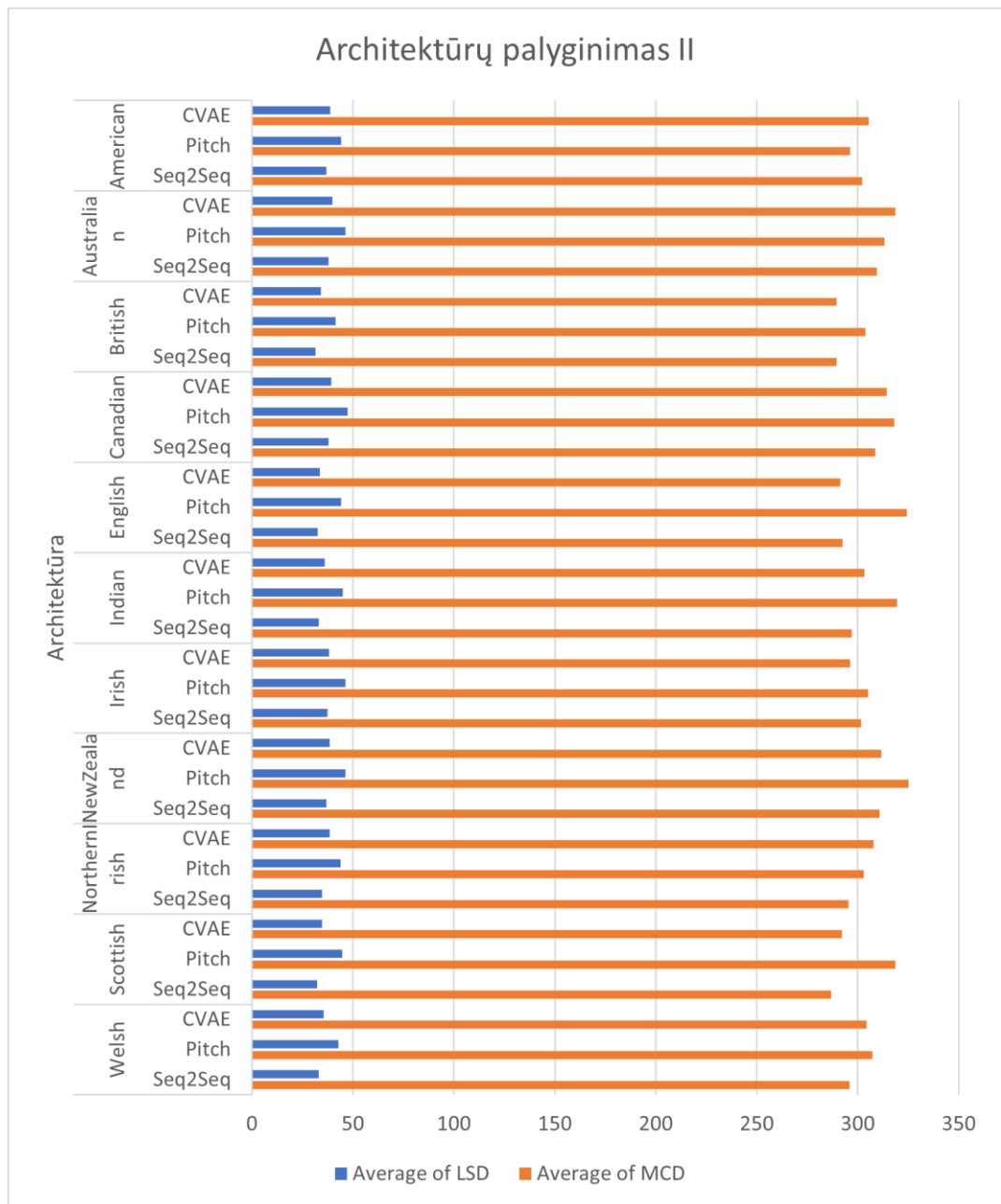
Galiausiai naudojant *LSD* ir *MCD* metrikas buvo palyginti modeliai su galimais vokoderiais naudojant *VCTK* duomenų rinkinį, kuriame yra 12 skirtingų anglų kalbos akcentų. Kiekvienam akcentui buvo paimti po 5 vyro ir moters garso įrašai (jei tiek egzistuoja, jei ne, papildyti kitos lyties įrašais). Galiausiai naudojant modelius keistas balsas pagal tų įrašų tekstą, ir jei modelis leidžia, pagal jų balso stilių.

Žemiau esančiame grafe pavaizduotas bendrinis architektūros vertinimas, iš kuriame matoma, jog *Seq2Seq* su *NET* vokoderiu geriausiai pakeitė balsą, o transformerių (grafuose *Pitch*) su *GAN* – blogiausiai. Lyginant su kitais vokoderiais, *GAN* vokoderis sukūrė blogiausius rezultatus.



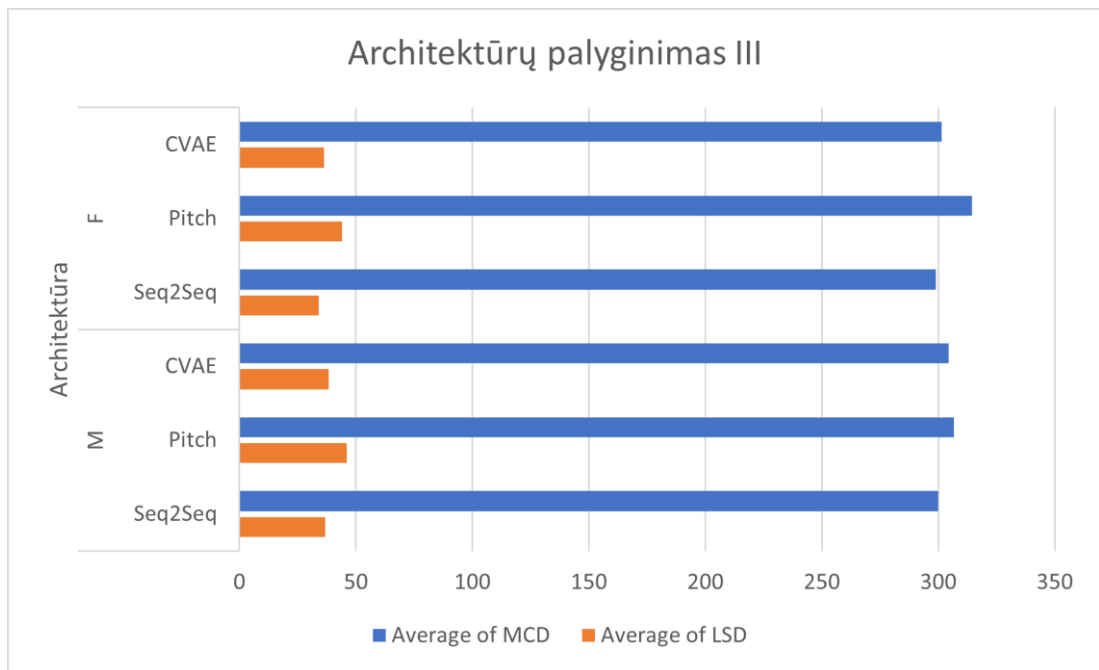
**6-34 pav.** Balso keitimo architektūrų palyginimas pagal modelius ir vokoderius

Žemiau pavaizduotas grafas rodo architektūrų rezultatus pagal šaltinio garso įrašo akcentą. Jame matome, kad skirtingos architektūros skirtingai susidoroja su skirtingais akcentais.



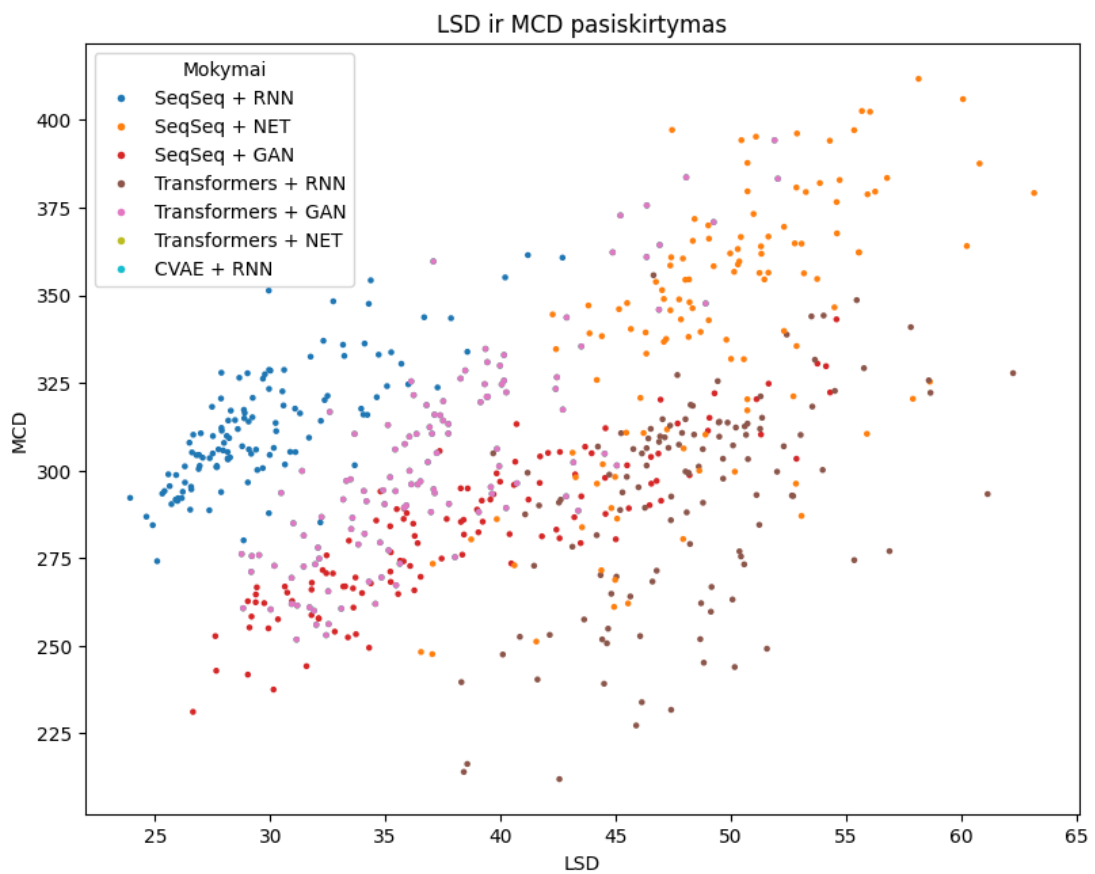
**6-35 pav.** Balso keitimo architektūrų palyginimas pagal akcentus

Galiausiai žemiau esančiame grafe matome, *CVAE* architektūros modeliai geriau susidoroja su moterišku balsu, transformerių – su vyrišku, o *Seq2Seq* lytis didelio skirtumo neturi. Nors skirtumai tarp skirtingų lyčių ir architektūrų egzistuoja, jie yra minimalūs.



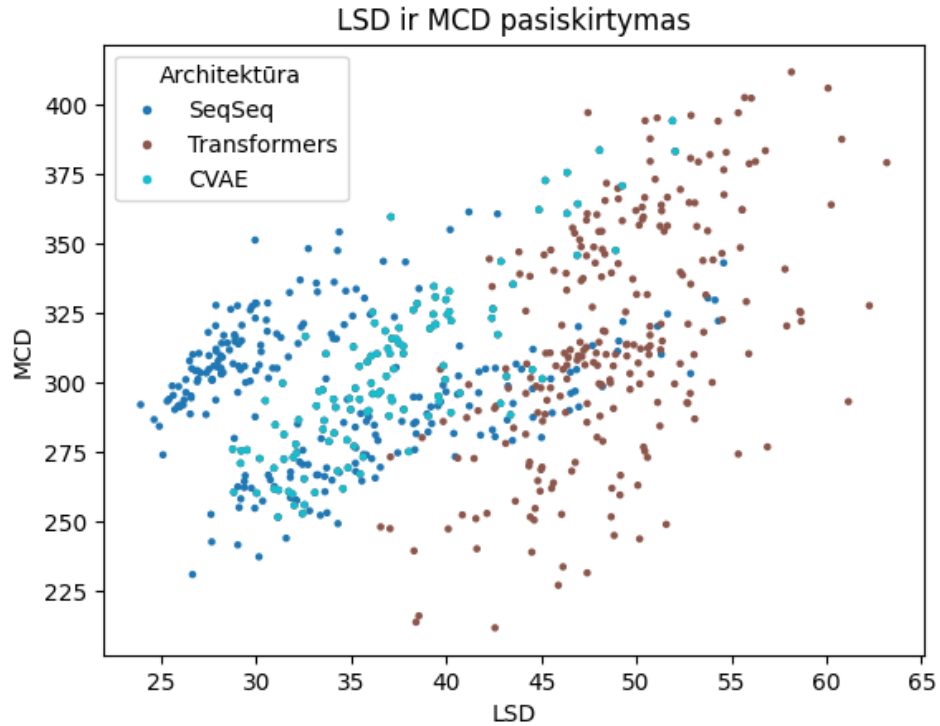
6-36 pav. Balso keitimo architektūrų palyginimas pagal lytį

6-37 pav. pavaizduotas balso keitimo metrikų pasiskirstymas pagal sugeneruotus garso įrašus, t.y. pagal kartu veikiančius veido keitimo modelius ir vokoderius.



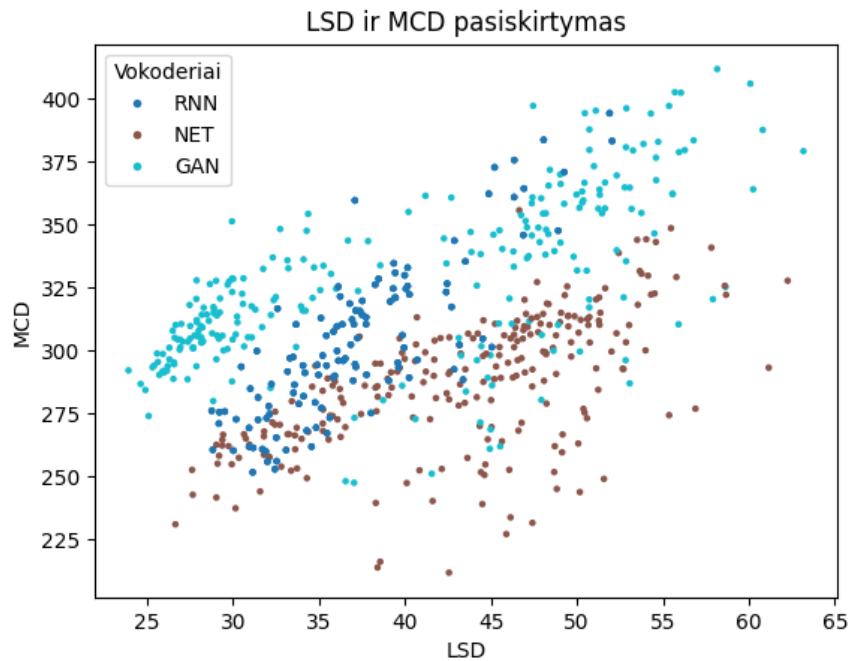
6-37 pav. Balso keitimo metrikų pasiskirstymas pagal mokymus

6-38 pav. pavaizduotas balso keitimo metrikų pasiskirstymas pagal balso keitimo modelius.



6-38 pav. Balso keitimo metrikų pasiskirstymas pagal modelio architektūrą

6-39 pav. pavaizduotas balso keitimo metrikų pasiskirstymas pagal balso keitimo vokoderius.



6-39 pav. Balso keitimo metrikų pasiskirstymas pagal vokoderius

Iš šių taškinių diagramų matome tą patį, jog *Seq2Seq* modelio su *NET* vokoderiu rezultatai buvo geriausi, o transformerių modelio su *GAN* vokoderiu – blogiausi. Be to, rezultatai, gauti su *GAN* vokoderiu, apskritai buvo prasčiausi.

### 6.2.7. Išvados

Atlikus eksperimentus galima teigti, kad visi modeliai skirtingai mokėsi pakeisti vieno žmogaus balsą kitu. *Seq2Seq* architektūra buvo apmokyta du kartus, o mažiausias pasiektas mokymo nuostolis buvo 1,059. Visiškai lygiagretaus transformatoriaus modelis nustojo mokytis po 5000 epochų, o *CVAE* architektūra tobulėjo iki 8000 epochų. *RNN* vokoderis mokėsi lėtai, bet tolygiai, o *GAN* vokoderis pradžioje pasiekė gerų rezultatų, tačiau maždaug po 4000 epochų pradėjo per daug prisitaikyti. *UnivNET* vokoderis galėjo gerai mokytis iki 6000 epochų, kol susidūrė su problemomis. Kiekvieno modelio mokymo trukmė taip pat skyrėsi: pasikartojanti *Seq2Seq* architektūra užtruko apie 3 minutes per epochą, o *UnivNET* vokoderis – iki 9 minučių per epochą.

Rezultatai rodo, kad *Seq2Seq* su *NET* vokoderiu veikė geriausiai, o transformerių su *GAN* vokoderiu – prasčiausiai. Be to, su *GAN* vokoderio rezultatai apskritai buvo prasčiausi. Taip pat buvo atrasta, kad skirtingos architektūros veikia skirtingai su įvairiais akcentams. Taip pat garso įrašo kalbėtojo lytis gali įtakoti architektūrų rezultatus – *CVAE* architektūra geriau veikia su moteriškais balsais, transformerių – su vyriškais balsais, o *Seq2Seq* skirtumai tarp lyčių yra minimalūs.



## Išvados

1. Analizuojant literatūrą apie veido ir balso keitimą naudojant neuroninius tinklus išsiaiškinta, kad šiam tikslui yra sukurta daug neuroninių tinklų architektūrų, kurios skiriasi sudėtingumu, paskirtimi bei rezultatų kokybe.
2. Suprojektuota ir sukurta programinė įranga, skirta veido ir balso keitimui bei sukurtas *pip* komponentas, kurį galima įdiegti į kitą programinę įrangą, parašytą su *Django* karkasu.
3. Atliktas programinės įrangos kokybės tyrimas, kuriuo metu išsiaiškinta, kad programinė įranga yra pakankamai kokybiškai ir saugiai sukurta. Kodo kokybė buvo įvertina 7/10 balo, o saugumo atžvilgiu nebuvo rasta didelių problemų.
4. Atlikti eksperimentai su veido keitimo modeliais, kurių metu išsiaiškinta, kokie parametrai gali įtakoti šių modelių rezultatus. Pavyzdžiui, auto-koduotoju grįsti modeliai yra efektyvus ir kuria kokybiškus rezultatus, tačiau auto-koduotojas su konvoliuciniais sluoksniais reikalauja daugiau resursų apsimokymui bei jo rezultatai nėra tokie geri. Taip pat *ReLU* aktyvacijos funkcija, *Adabelief* optimizatorius, didesnė *AE*, tačiau mažesnė *GAN* dimensija turėjo teigiamos įtakos rezultatams.
5. Atlikti eksperimentai su balso keitimo modeliais, kurių metu išsiaiškinta, kad *RNN* balso keitimo modeliai greičiausiai ir kokybiškiausiai apsimoka, tuo tarpu *CVAE* – prasčiausiai. Taip pat buvo pastebėta, jog *GAN* vokoderis duota prasčiausius rezultatus, todėl nėra tinkamas balso keitimui su nagrinėtais modeliais. Eksperimentai parodė, jog renkantis modelius balso keitimui svarbu atkreipti dėmesį į kalbėtojo lytį bei akcentą, nes jie turi įtakos.

## Literatūros sąrašas

- [1] „Secondary education statistics“ [žiūrėta 2021 12 04]. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Secondary\\_education\\_statistics](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Secondary_education_statistics)
- [2] „Number of movies released in the United States and Canada from 2000 to 2020“ [žiūrėta 2021 12 04]. <https://www.statista.com/statistics/187122/movie-releases-in-north-america-since-2001/>
- [3] „How much does it cost to record an audiobook?“ [žiūrėta 2021 12 04]. <https://krystalwascher.com/blog-audiobooks-for-authors/how-much-does-it-cost-to-create-an-audiobook>
- [4] „Number of audiobook titles published in the United States from 2007 to 2020“ [žiūrėta 2021 12 04]. <https://www.statista.com/statistics/261185/number-of-audiobooks-published-in-the-us/>
- [5] X. Tong; L. Wang; X.Pan; J. Wang. „An Overview of Deepfake: The Sword of Damocles in AI“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/9270435>
- [6] Nirkin, Y. Keller, T. Keller. „FSGAN: Subject Agnostic Face Swapping and Reenactment“ (2019). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1908.05932>
- [7] P. Swathi, S. Saritha. „DeepFake Creation and Detection: A Survey“ (2021). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1909.11573>
- [8] X. Fan, D. Bonatto, G. Lafruit. „Consistent Long Sequences Deep Faces“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/8975999>
- [9] A. Datta, O. Yadav, Y. Singh. „Real-Time Face Swapping System using OpenCV“ (2021). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/9545010>
- [10] T. Zhang, L. Deng, L. Zhang, X. Dang. „Deep Learning in Face Synthesis: A Survey on Deepfakes“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/9213159>
- [11] Y. Korshunova, W. Shi, J. Dambre, L. Theis. „Fast Face-swap Using Convolutional Neural Networks“ (2017). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1611.09577>
- [12] B. Lin, D. Hsu, C. Shen, H. Hsiao. „Using Fully Connected and Convolutional Net for GAN-Based Face Swapping“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9301665>
- [13] R. Natsume , T. Yatagawa, S. Morishima. „FSNet: An Identity-Aware Generative Model for Image-based Face Swapping“ (2018). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1811.12666>
- [14] R. Natsume , T. Yatagawa, S. Morishima. „RSGAN: Face Swapping and Editing using Face and Hair Representation in Latent Spaces“ (2018). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1804.03447>
- [15] B. Maze, J. Adams, J. Duncan, N. Kalka, T. Miller. „Face Dataset and Protocol“ (2018). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/8411217>
- [16] S. Barua , S. Erfani, J. Bailey. „FCC-GAN: A Fully Connected and Convolutional Net Architecture for GANs“ (2019). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1905.02417>

- [17] J. Zhang, X. Zeng, M. Wang, Y. Pan. „FreeNet: Multi-Identity Face Reenactment“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1905.11805>
- [18] D. Chen, Q. Chen, J. Wu, X. Yu, T. Jia. „Face Swapping: Realistic Image Synthesis Based on Facial Landmarks Alignment“ (2019). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: [https://www.researchgate.net/publication/331786996\\_Face\\_Swapping\\_Realistic\\_Image\\_Synthesis\\_Based\\_on\\_Facial\\_Landmarks\\_Alignment](https://www.researchgate.net/publication/331786996_Face_Swapping_Realistic_Image_Synthesis_Based_on_Facial_Landmarks_Alignment)
- [19] Y. Zhang, L. Zheng, V. Thing. „Automated Face Swapping and Its Detection“ (2017). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8124497>
- [20] A. Brock, J. Donahue, K. Simonyam. „Large scale gan training for High fidelity natural image synthesis“ (2019). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/1809.11096.pdf>
- [21] Y. Yang, J. Yu, N. Jojic, J. Huan, T. Huang. „FSNet: Compression of Deep Convolutional Neural Networks by Filter Summary“ (2018). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1902.03264>
- [22] M. Kawanaka, Y. Koizumi, R. Miyazaki, K. Yatabe. „Stable Training of Dnn for Speech Enhancement Based on Perceptually-Motivated Black-Box Cost Function“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą:
- [23] J. Zhao, X. Mao, L. Chen. „Speech emotion recognition using deep 1D & 2D CNN LSTM networks“ (2018). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/abstract/document/9054578>
- [24] J. „Multimodal Voice Conversion Under Adverse Environment Using a Deep Convolutional Neural Network“ (2019). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/8913498>
- [25] W. Huang, T. Hayashi, Y. Wu. „Pretraining Techniques for Sequence-to-Sequence Voice Conversion“ (2021). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9314100>
- [26] C. Deng, Y. Chen, H. Deng. „One-Shot Voice Conversion Algorithm Based on Representations Separation“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/9240913>
- [27] L. Wan, Q. Wang, A. Papir, I. Moreno. „Generalized end-to-end loss for speaker verification“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1710.10467>
- [28] T. Tobing, Y. Wu, T. Hayashi, K. Kobayashi. „Voice Conversion With CycleRNN-Based Spectral Mapping and Finely Tuned WaveNet Vocoder“ (2019). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/8913551>
- [29] L. Liu, Z. Ling, Y. Jiang, L. Dai. „WaveNet Vocoder with Limited Training Data for Voice Conversion“ (2018). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: [https://www.isca-speech.org/archive/pdfs/interspeech\\_2018/liu18\\_interspeech.pdf](https://www.isca-speech.org/archive/pdfs/interspeech_2018/liu18_interspeech.pdf)
- [30] A. Tamamori, T. Hayashi, K. Kobayashi, K. Takeda, T. Toda. „Speaker-dependent WaveNet vocoder“ (2017). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: [https://www.researchgate.net/publication/319184793\\_Speaker-Dependent\\_WaveNet\\_Vocoder](https://www.researchgate.net/publication/319184793_Speaker-Dependent_WaveNet_Vocoder)
- [31] J. Zhang, Z. Ling, L. Liu, Y. Jang. „Sequence-to-Sequence Acoustic Modeling for Voice Conversion“ (2019). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1810.06865>

- [32] X. Miao, X. Zhang, M. Sun, C. Zheng. „A BLSTM and WaveNet-Based Voice Conversion Method With Waveform Collapse Suppression by Post-Processing“ (2019). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/document/8695725?source=au>
- [33] Y. Zhou, X. Tian, H. Li. „Multi-Task WaveRNN With an Integrated Architecture for Cross-Lingual Voice Conversion“ (2020). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: [https://www.researchgate.net/publication/342919288\\_Multi-Task\\_WaveRNN\\_With\\_an\\_Integrated\\_Architecture\\_for\\_Cross-Lingual\\_Voice\\_Conversion](https://www.researchgate.net/publication/342919288_Multi-Task_WaveRNN_With_an_Integrated_Architecture_for_Cross-Lingual_Voice_Conversion)
- [34] Nagrani, J. Chung, A. Zisserman. „VoxCeleb: a large-scale speaker identification dataset“ (2018). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1706.08612>
- [35] J. Lorenzo-Trueba, J. Yamagishi, T. Toda, D. Saito. „Promoting Development of Parallel and Nonparallel Methods“ (2018). [žiūrėta 2021 m. lapkričio 17 d.] Prieiga per internetą: <https://arxiv.org/abs/1804.04262>
- [36] Perov I., Gao D., Chervoniy N, Liu K., Marangonda U., Umé C. „DeepFaceLab: Integrated, flexible and extensible face-swapping framework“ (2020). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/abs/2005.05535>
- [37] Snover D., Johnson C. W., Bianco M. J., Gerstoft P. „Deep Clustering to Identify Sources of Urban Seismic Noise in Long Beach, California“ (2020). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: [https://www.researchgate.net/figure/Convolutional-autoencoder-CAE-architecture-The-encoder-compresses-the-input-images-to\\_fig2\\_347540445](https://www.researchgate.net/figure/Convolutional-autoencoder-CAE-architecture-The-encoder-compresses-the-input-images-to_fig2_347540445)
- [38] Noal M., Trivedi A., Dutta P. „Growing Cosine Unit“ (2023). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/2108.12943.pdf>
- [39] Sara U., Akter M., Uddin M. S., „Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study“ (2019). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: [https://www.scirp.org/pdf/JCC\\_2019030117485323.pdf](https://www.scirp.org/pdf/JCC_2019030117485323.pdf)
- [40] Shen J., Pang R., Weiss R., Jaitly N. „Natural TTS synthesis by conditioning wavenet on mel spectrogram predictions“ (2018). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/1712.05884.pdf>
- [41] Shen J., Pang R. „Tacotron 2: Generating Human-like Speech from Text“ (2017). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://ai.googleblog.com/2017/12/tacotron-2-generating-human-like-speech.html>
- [42] Łancucki A. „Fastpitch: parallel text-to-speech with pitch prediction“ (2021). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/2006.06873.pdf>
- [43] Choi H. Heo H., Lee J. Lee K. „Phase-aware Single-stage Speech Denoising and Dereverberation with U-Net“ (2020). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/2006.0687.pdf>
- [44] Kim J., Kong J., Son, J., „Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech“ (2021). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/2106.06103.pdf>
- [45] Kumar K., Kumar R. m Gestin L, Toeh W. Sotelo J. „MelGAN: Generative Adversarial Networks for Conditional Waveform Synthesis“ (2019). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/1910.06711.pdf>
- [46] Kalchbrenner N., Elsen E., Siminyan K., Noury S. „Efficient Neural Audio Synthesis“ (2018). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/1802.08435v1.pdf>

- [47] Dipjoyoti P. Shifas M., Pantazis Y. „Enhancing Speech Intelligibility in Text-To-Speech Synthesis using Speaking Style Conversion“ (2020). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą:  
[https://www.researchgate.net/publication/343648872\\_Enhancing\\_Speech\\_Intelligibility\\_in\\_Text-To-Speech\\_Synthesis\\_using\\_Speaking\\_Style\\_Conversion](https://www.researchgate.net/publication/343648872_Enhancing_Speech_Intelligibility_in_Text-To-Speech_Synthesis_using_Speaking_Style_Conversion)
- [48] Jang W., Lim D., Yoon J., Kim B., Kim J. „UnivNet: A Neural Vocoder with Multi-Resolution Spectrogram Discriminators for High-Fidelity Waveform Generation“ (2021). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/2106.07889.pdf>
- [49] Ji S., Luo J., Yang X. „A Comprehensive Survey on Deep Music Generation: Multi-level Representations, Algorithms, Evaluations, and Future Directions“ (2020). [žiūrėta 2023 m. gegužės 17 d.] Prieiga per internetą: <https://arxiv.org/pdf/2011.06801.pdf>

## Informacijos šaltinių sąrašas

- [1] AutoVC [žiūrėta 2021 m. gruodžio 4 d.]. Prieiga per internetą: <https://auspicious3000.gib.io/autovc-demo/>
- [2] Deep Nostalgia [žiūrėta 2021 m. gruodžio 4 d.]. Prieiga per internetą: <https://www.myheritage.lt/deep-nostalgia>
- [3] ConVoice [žiūrėta 2021 m. gruodžio 4 d.]. Prieiga per internetą: <https://rebryk.gib.io/convoice-demo>
- [4] Pylint [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://pypi.org/project/pylint>
- [5] Vulture [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://pypi.org/project/vulture>
- [6] Safety [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://pypi.org/project/safety/>
- [7] Bandit [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://pypi.org/project/bandit/>
- [8] Tensorflow [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://www.tensorflow.org/>
- [9] Pytorch [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://pytorch.org/>
- [10] PaddleSpeech dokumentacija [žiūrėta 2023 m. sausio 13 d.] Prieiga per internetą: [https://paddlespeech.readthedocs.io/en/latest/tts/models\\_introduction.html](https://paddlespeech.readthedocs.io/en/latest/tts/models_introduction.html)
- [11] PyTorch biblioteka [žiūrėta 2023 m. sausio 13 d.] Prieiga per internetą: <https://pytorch.org/>
- [12] PaddleSpeech biblioteka [žiūrėta 2023 m. sausio 13 d.] Prieiga per internetą: <https://paddlespeech.readthedocs.io/en/latest/>
- [13] PaddlePaddle įrankis [žiūrėta 2023 m. sausio 13 d.] Prieiga per internetą: <https://raw.githubusercontent.com/PaddlePaddle/PaddleSpeech/develop/docs/images/fastspeech2.png>
- [14] Black biblioteka [žiūrėta 2023 m. sausio 13 d.] Prieiga per internetą: <https://pypi.org/project/black/>
- [15] Pytest biblioteka [žiūrėta 2023 m. sausio 13 d.] Prieiga per internetą: <https://docs.pytest.org/en/7.2.x/>
- [16] DeepFaceLab [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://gib.com/iperov/DeepFaceLab/>
- [17] CV2 [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://pypi.org/project/opencv-python/>
- [18] h5py [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://www.h5py.org/>
- [19] TTS [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://tts.readthedocs.io/en/latest/>
- [20] Torchaudio [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://pytorch.org/audio/stable/index.html>
- [21] Trainer [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://gib.com/coqui-ai/Trainer>
- [22] DeepfakeVFX [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://www.deepfakevfx.com>
- [23] LJ-Speech duomenų rinkinys [žiūrėta 2023 m. gegužės 13 d.]. Prieiga per internetą: <https://keithito.com/LJ-Speech-Dataset>

## **Priedai**

### **1 priedas. Licencija**

MIT License

Copyright (c) 2023 Indrė Dimšė

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

## 2 priedas. Veido keitimo eksperimentų rezultatai

Mokymas	Epocha	MSE	SSIM
AE II	0	5.061679	5.406548
AE II	10000	2.258605	1.483945
AE II	20000	2.664814	1.490764
AE II	30000	2.248652	1.45037
AE II	40000	2.129727	1.424484
AE II	50000	2.258826	1.704803
AE II	60000	2.230191	1.233818
AE II	70000	1.944137	1.849289
AE II	80000	1.933966	0.913171
AE II	90000	2.001261	2.551352
AE II	100000	2.127165	1.458572
AE II	110000	1.460737	1.067376
AE II	120000	2.050212	1.724917
AE II	130000	1.95519	1.354248
AE II	140000	1.872617	1.16412
AE II	150000	2.155118	1.544595
AE II	160000	1.433814	1.699129
AE II	170000	1.452943	1.356114
AE II	180000	1.656147	1.212901
AE II	190000	1.939741	0.72415
AE II	200000	1.580081	1.28912
AE II	210000	1.411413	1.122568
AE II	220000	1.322126	1.040451
AE II	230000	1.487332	1.522761
AE II	240000	1.363842	1.215398
AE II	250000	1.155271	1.279889
AE II	260000	1.708463	1.243073
AE II	270000	1.426653	1.364778
AE II	280000	1.544151	1.465434
AE II	290000	1.887779	1.137332
AE II	300000	0.98529	0.826303
AE IV	0	6.19062	4.604189
AE IV	10000	1.845824	1.890674
AE IV	20000	1.895177	1.566743
AE IV	30000	1.707512	1.840921
AE IV	40000	1.813156	1.006398
AE IV	50000	1.857474	1.538371
AE IV	60000	1.380341	2.151584
AE IV	70000	1.694557	1.62836
AE IV	80000	2.334585	1.823364
AE IV	90000	1.603665	1.274552
AE IV	100000	2.561093	1.689048
AE V	0	6.08281	5.254317
AE V	10000	2.220895	2.405038
AE V	20000	2.448142	2.065784
AE V	30000	2.015881	1.862443
AE V	40000	2.160964	1.632535
AE V	50000	2.256997	1.040854
AE V	60000	1.836541	1.117291
AE V	70000	2.196489	2.079695
AE V	80000	2.324358	1.33793
AE V	90000	1.94712	1.779808
AE V	100000	1.608818	1.309976
AE III	0	4.38054	3.500222
AE III	10000	1.355307	0.777406
AE III	20000	0.99037	0.891724
AE III	30000	0.810367	0.481394
AE III	40000	1.043337	0.734544
AE III	50000	0.73067	0.391552



AE III	60000	0.766981	0.578897
AE III	70000	0.594353	0.415699
AE III	80000	0.677318	0.497021
AE III	90000	0.638338	0.585003
AE III	100000	0.49488	0.312885
AE III	110000	0.639041	0.397613
AE III	120000	0.79146	0.431537
AE III	130000	0.440288	0.478679
AE III	140000	0.640549	0.43361
AE I	0	6.785233	5.194282
AE I	10000	1.513281	1.240303
AE I	20000	1.336686	1.074535
AE I	30000	1.026655	1.24891
AE I	40000	0.747337	0.943818
AE I	50000	0.956244	1.077735
AE I	60000	1.114126	0.929289
AE I	70000	1.047987	1.503653
AE I	80000	0.929668	1.087074
AE I	90000	1.032059	1.247743
AE I	100000	1.266139	1.17826
AE I	110000	1.171615	0.852692
AE I	120000	1.21254	1.261104
AE I	130000	0.714295	0.923333
AE I	140000	0.890119	0.802123
AE I	150000	1.414937	0.909181
AE I	160000	0.74462	0.700306
AE I	170000	0.894363	0.693148
AE I	180000	0.657213	1.040092
AE I	190000	0.968631	0.816631
AE I	200000	0.988918	0.678469
AE I	210000	0.679229	0.820917
AE I	220000	0.729871	1.211038
AE I	230000	0.99843	0.984242
AE I	240000	0.651355	0.890098
AE I	250000	0.539059	0.673095
AE I	260000	0.661177	0.671701
AE I	270000	0.652214	0.719785
AE I	280000	1.166182	0.774031
AE I	290000	0.658655	0.69528
AE I	300000	0.705515	0.576065
AE I	310000	0.833656	0.747564
AE I	320000	0.623049	0.703548
AE I	330000	0.737828	0.548487
AE I	340000	0.580841	0.553377
AE I	350000	0.876662	0.530439
AE I	360000	0.748031	0.644693
AE I	370000	0.620614	0.521911
AE I	380000	0.489073	0.79172
CAE II	0	4.573033	5.266249
CAE II	10000	2.03278	1.8058
CAE II	20000	2.026486	1.610907
CAE II	30000	1.821256	1.147107
CAE II	40000	1.771545	1.288048
CAE II	50000	1.680531	1.245563
CAE II	60000	1.94538	1.315131
CAE II	70000	1.967171	1.436449
CAE II	80000	1.337234	1.228067
CAE II	90000	1.835077	1.22839
CAE II	100000	1.657492	1.522265
CAE II	110000	1.276855	1.828265
CAE II	120000	1.372112	1.619281

CAE II	130000	2.058326	0.741692
CAE II	140000	1.737425	1.327488
CAE II	150000	1.88394	1.153954
CAE II	160000	1.682802	1.158741
CAE II	170000	1.637854	0.847677
CAE II	180000	1.494132	1.456372
CAE II	190000	1.81831	1.10883
CAE II	200000	1.480072	0.888122
CAE II	210000	1.512204	1.015012
CAE II	220000	1.531987	0.837964
CAE II	230000	1.028389	0.907144
CAE II	240000	1.482721	1.20506
CAE II	250000	1.134108	1.162876
CAE II	260000	1.472222	1.082797
CAE IV	0	5.603387	5.81399
CAE IV	10000	2.058641	1.439366
CAE IV	20000	2.945758	1.874956
CAE IV	30000	2.351121	0.991238
CAE IV	40000	2.328864	1.604133
CAE IV	50000	1.717831	1.180865
CAE IV	60000	2.67844	2.186853
CAE IV	70000	2.490469	1.215006
CAE IV	80000	2.332032	1.632203
CAE IV	90000	2.440942	1.878625
CAE IV	100000	2.391428	2.035027
CAE IV	110000	2.396674	1.939695
CAE V	0	7.042601	5.366327
CAE V	10000	2.430672	2.452591
CAE V	20000	2.500626	1.724538
CAE V	30000	2.496747	2.716511
CAE V	40000	1.825448	2.043958
CAE V	50000	2.298932	1.768485
CAE V	60000	2.119287	1.555371
CAE V	70000	2.284882	1.767439
CAE V	80000	1.996962	2.226608
CAE V	90000	2.10315	1.320095
CAE V	100000	2.747225	1.476115
CAE III	0	4.462295	3.901119
CAE III	10000	1.445431	1.126757
CAE III	20000	1.398205	1.18426
CAE III	30000	1.324935	0.726472
CAE III	40000	1.254919	0.986581
CAE III	50000	1.228732	0.656462
CAE III	60000	1.196936	0.646464
CAE III	70000	1.060341	0.639329
CAE III	80000	1.310934	0.857487
CAE III	90000	0.895979	0.625714
CAE III	100000	0.901751	0.713312
CAE III	110000	1.001862	0.62213
CAE III	120000	0.782864	0.449609
CAE III	130000	1.033054	0.617091
CAE III	140000	1.081553	0.657011
CAE III	150000	1.341039	0.660713
CAE I	0	6.290395	5.043916
CAE I	10000	1.818781	1.555788
CAE I	20000	1.6704	2.01189
CAE I	30000	1.382946	1.042863
CAE I	40000	1.464935	1.308721
CAE I	50000	1.293124	1.633367
CAE I	60000	1.466346	0.7578
CAE I	70000	1.037639	1.268766

CAE I	80000	1.102195	1.114832
CAE I	90000	1.370568	1.75366
CAE I	100000	1.22139	1.014223
CAE I	110000	1.289323	1.568776
CAE I	120000	1.00198	0.907653
CAE I	130000	1.391971	1.11368
CAE I	140000	0.908886	1.344801
CAE I	150000	1.1237	1.094055
CAE I	160000	1.042496	0.947667
CAE I	170000	1.074941	0.760997
CAE I	180000	0.985608	0.970393
CAE I	190000	1.136436	0.694124
CAE I	200000	1.086653	0.940793
CAE I	210000	0.796677	1.056858
CAE I	220000	1.100568	0.861201
CAE I	230000	0.901817	0.992548
CAE I	240000	1.006423	1.03337
CAE I	250000	0.905447	0.80486
CAE I	260000	1.015054	0.754838
CAE I	270000	0.973291	0.785704
CAE I	280000	0.90972	0.885338
CAE I	290000	0.752307	0.925233
CAE I	300000	1.073217	0.665679
CAE I	310000	0.825404	1.061404
CAE I	320000	0.70042	0.783306
CAE I	330000	0.983446	0.984025
CAE I	340000	0.952178	0.981481
CAE I	350000	1.019938	0.775712
CAE I	360000	0.929358	0.90512
CAE I	370000	0.853974	0.90971
CAE I	380000	0.970292	0.661889
CAE I	390000	0.677571	1.208951
CAE I	400000	0.754414	0.803972
CAE I	410000	0.698611	0.589287
CAE I	420000	0.632464	0.650454
CAE I	430000	0.687246	1.232633
CAE I	440000	0.619875	0.746548
CAE I	450000	0.732893	0.764947
CAE I	460000	0.835287	0.895823
CAE I	470000	0.762054	0.775864
CAE I	480000	1.263913	0.902937
CAE I	490000	0.698619	0.734245
CAE I	500000	0.785153	0.537001
CAE I	510000	0.745649	0.907114
CAE I	520000	0.693809	0.658236
CAE I	530000	0.734967	0.747787
CAE I	540000	1.087817	0.52112
CAE I	550000	0.709429	0.858817
CAE I	560000	0.802443	0.585882
CAE I	570000	0.726221	0.744643
CAE I	580000	0.692497	0.6846
CAE I	590000	0.86773	0.880367
CAE I	600000	0.634419	0.866709
CAE I	610000	0.844161	0.483013
CAE I	620000	1.127326	0.845756
CAE I	630000	0.785479	0.633427
CAE I	640000	0.65923	0.534882
CAE I	650000	0.665211	0.61759
CAE I	660000	0.804586	1.11087
CAE I	670000	0.841024	0.713261
CAE I	680000	0.873335	0.531499

### 3 priedas. Balso keitimo eksperimentų rezultatai

Architektūra	Vokoderis	Akcentas	Lytis	MCD	LSD
SeqSeq	RNN	English	F	260.3799	30.02154
SeqSeq	RNN	English	F	260.9638	31.7087
SeqSeq	RNN	English	F	310.5578	37.7368
SeqSeq	RNN	English	F	359.6662	37.08592
SeqSeq	RNN	English	F	290.4194	34.96936
SeqSeq	RNN	English	F	278.9194	33.82704
SeqSeq	RNN	English	M	297.5127	33.49255
SeqSeq	RNN	English	M	272.6088	31.45284
SeqSeq	RNN	English	M	269.3785	30.92981
SeqSeq	RNN	English	M	312.9836	35.11102
SeqSeq	RNN	Scottish	M	321.5176	36.23052
SeqSeq	RNN	Scottish	M	261.9694	34.56777
SeqSeq	RNN	Scottish	M	334.711	39.34938
SeqSeq	RNN	Scottish	M	286.7405	32.22967
SeqSeq	RNN	Scottish	M	296.0772	36.13046
SeqSeq	RNN	Scottish	M	291.2472	34.18153
SeqSeq	RNN	Scottish	F	265.5407	32.51915
SeqSeq	RNN	Scottish	F	273.2494	31.95457
SeqSeq	RNN	Scottish	F	298.0945	35.20948
SeqSeq	RNN	Scottish	F	294.1194	35.43943
SeqSeq	RNN	NorthernIrish	F	279.4613	34.77303
SeqSeq	RNN	NorthernIrish	F	268.2895	34.28716
SeqSeq	RNN	NorthernIrish	F	286.5145	33.57223
SeqSeq	RNN	NorthernIrish	F	375.6408	46.35688
SeqSeq	RNN	NorthernIrish	M	306.1705	39.84719
SeqSeq	RNN	NorthernIrish	M	317.3793	42.71001
SeqSeq	RNN	NorthernIrish	M	326.2437	38.26785
SeqSeq	RNN	NorthernIrish	M	360.9236	46.3488
SeqSeq	RNN	NorthernIrish	M	296.9236	36.52649
SeqSeq	RNN	NorthernIrish	M	260.0272	31.90167
SeqSeq	RNN	Indian	F	325.6499	40.15692
SeqSeq	RNN	Indian	F	267.2869	35.46554
SeqSeq	RNN	Indian	F	307.5809	35.84544
SeqSeq	RNN	Indian	F	309.6869	36.04488
SeqSeq	RNN	Indian	F	256.1896	32.53725
SeqSeq	RNN	Indian	F	302.4248	36.84725
SeqSeq	RNN	Indian	F	316.7225	32.58582
SeqSeq	RNN	Indian	F	319.7948	37.51697
SeqSeq	RNN	Indian	F	294.1597	34.06716
SeqSeq	RNN	Indian	F	332.9633	40.15636
SeqSeq	RNN	Welsh	F	364.353	46.91721
SeqSeq	RNN	Welsh	F	372.7736	45.21051
SeqSeq	RNN	Welsh	F	306.0236	35.42592
SeqSeq	RNN	Welsh	F	318.6635	36.78993
SeqSeq	RNN	Welsh	F	284.9596	31.00399
SeqSeq	RNN	Welsh	F	275.8669	29.51745
SeqSeq	RNN	Welsh	F	277.2458	35.14555
SeqSeq	RNN	Welsh	F	299.8904	31.38493
SeqSeq	RNN	Welsh	F	291.7575	33.1723
SeqSeq	RNN	Welsh	F	251.7896	31.14471
SeqSeq	RNN	Irish	F	276.1844	28.75827
SeqSeq	RNN	Irish	F	321.0232	39.40724
SeqSeq	RNN	Irish	F	394.2239	51.9029
SeqSeq	RNN	Irish	F	271.1082	29.17481
SeqSeq	RNN	Irish	F	294.3906	34.90329
SeqSeq	RNN	Irish	F	260.7135	28.82402
SeqSeq	RNN	Irish	F	305.01	37.23314
SeqSeq	RNN	Irish	F	292.6204	42.85366
SeqSeq	RNN	Irish	F	283.3079	33.5188
SeqSeq	RNN	Irish	F	290.0571	35.91616

SeqSeq	RNN	Irish	M	326.622	42.45082
SeqSeq	RNN	Irish	M	273.4116	35.62862
SeqSeq	RNN	Irish	M	304.8015	44.52589
SeqSeq	RNN	Irish	M	301.2356	39.94515
SeqSeq	RNN	Irish	M	288.5488	43.38083
SeqSeq	RNN	Irish	M	293.5985	37.14486
SeqSeq	RNN	Irish	M	269.5083	34.83251
SeqSeq	RNN	Irish	M	302.33	43.16847
SeqSeq	RNN	Irish	M	288.1461	39.0569
SeqSeq	RNN	Irish	M	289.3168	40.24462
SeqSeq	RNN	American	F	255.9949	31.99937
SeqSeq	RNN	American	F	299.9251	36.05113
SeqSeq	RNN	American	F	290.4845	37.76287
SeqSeq	RNN	American	F	293.3677	39.62923
SeqSeq	RNN	American	F	315.9406	37.13428
SeqSeq	RNN	American	F	383.6575	48.06765
SeqSeq	RNN	American	M	296.3655	40.71579
SeqSeq	RNN	American	M	294.6605	37.0984
SeqSeq	RNN	American	M	347.667	48.91279
SeqSeq	RNN	American	M	275.607	29.20129
SeqSeq	RNN	Canadian	M	322.3232	40.25586
SeqSeq	RNN	Canadian	M	272.8255	30.18183
SeqSeq	RNN	Canadian	M	335.4238	43.51154
SeqSeq	RNN	Canadian	M	310.4514	36.35465
SeqSeq	RNN	Canadian	F	324.5533	40.05473
SeqSeq	RNN	Canadian	F	281.9556	34.09995
SeqSeq	RNN	Canadian	F	329.9191	39.98169
SeqSeq	RNN	Canadian	F	275.2954	38.02515
SeqSeq	RNN	Canadian	F	343.6708	42.87375
SeqSeq	RNN	Canadian	F	345.8899	46.8854
SeqSeq	RNN	NewZealand	F	295.2897	39.58217
SeqSeq	RNN	NewZealand	F	289.4121	35.822
SeqSeq	RNN	NewZealand	F	277.9962	32.07537
SeqSeq	RNN	NewZealand	F	297.0721	33.30819
SeqSeq	RNN	NewZealand	F	328.5334	38.47115
SeqSeq	RNN	NewZealand	F	296.0346	36.40699
SeqSeq	RNN	NewZealand	F	288.163	36.98545
SeqSeq	RNN	NewZealand	F	370.8704	49.26068
SeqSeq	RNN	NewZealand	F	362.2488	44.87347
SeqSeq	RNN	NewZealand	F	310.5138	37.03005
SeqSeq	RNN	Australian	M	383.2427	52.04856
SeqSeq	RNN	Australian	M	330.9115	39.43051
SeqSeq	RNN	Australian	M	314.2385	37.50971
SeqSeq	RNN	Australian	M	261.4906	31.17269
SeqSeq	RNN	Australian	M	301.4445	45.05277
SeqSeq	RNN	Australian	M	323.2881	42.39913
SeqSeq	RNN	Australian	M	319.5202	39.13591
SeqSeq	RNN	Australian	M	312.3754	36.92452
SeqSeq	RNN	Australian	M	325.4268	36.1255
SeqSeq	RNN	Australian	M	315.7601	37.38139
SeqSeq	RNN	British	F	310.449	33.66387
SeqSeq	RNN	British	F	313.2741	37.74027
SeqSeq	RNN	British	F	321.0585	39.44936
SeqSeq	RNN	British	F	324.6578	39.31625
SeqSeq	RNN	British	F	260.6014	33.08647
SeqSeq	RNN	British	F	253.0311	32.43106
SeqSeq	RNN	British	F	281.4615	31.45585
SeqSeq	RNN	British	F	261.9505	30.93158
SeqSeq	RNN	British	F	274.864	32.12665
SeqSeq	RNN	British	F	293.6362	30.47682
SeqSeq	NET	English	F	262.7026	29.02986
SeqSeq	NET	English	F	265.1836	30.74109
SeqSeq	NET	English	F	296.8212	39.98575

SeqSeq	NET	English	F	305.5728	37.35934
SeqSeq	NET	English	F	315.038	49.07781
SeqSeq	NET	English	F	254.0308	32.80429
SeqSeq	NET	English	M	258.76	31.79382
SeqSeq	NET	English	M	270.7026	32.43652
SeqSeq	NET	English	M	269.6802	36.53253
SeqSeq	NET	English	M	264.7702	35.5572
SeqSeq	NET	Scottish	M	244.1811	31.57593
SeqSeq	NET	Scottish	M	273.4912	40.4712
SeqSeq	NET	Scottish	M	266.9151	33.17333
SeqSeq	NET	Scottish	M	252.4143	33.37668
SeqSeq	NET	Scottish	M	249.4378	34.29687
SeqSeq	NET	Scottish	M	237.5109	30.14495
SeqSeq	NET	Scottish	F	258.3348	29.18938
SeqSeq	NET	Scottish	F	242.8797	27.65727
SeqSeq	NET	Scottish	F	279.2886	36.40418
SeqSeq	NET	Scottish	F	271.6024	32.28028
SeqSeq	NET	NorthernIrish	F	274.111	35.79452
SeqSeq	NET	NorthernIrish	F	257.5617	30.33881
SeqSeq	NET	NorthernIrish	F	275.7947	32.45017
SeqSeq	NET	NorthernIrish	F	281.8564	40.39518
SeqSeq	NET	NorthernIrish	M	280.6835	42.56716
SeqSeq	NET	NorthernIrish	M	287.6292	44.56249
SeqSeq	NET	NorthernIrish	M	272.7746	36.07741
SeqSeq	NET	NorthernIrish	M	297.9166	44.54337
SeqSeq	NET	NorthernIrish	M	260.9006	33.61501
SeqSeq	NET	NorthernIrish	M	253.3024	33.72533
SeqSeq	NET	Indian	F	291.4261	39.19068
SeqSeq	NET	Indian	F	267.9787	31.80855
SeqSeq	NET	Indian	F	266.0175	31.795
SeqSeq	NET	Indian	F	287.8557	35.92946
SeqSeq	NET	Indian	F	286.1922	35.79643
SeqSeq	NET	Indian	F	284.1017	35.21156
SeqSeq	NET	Indian	F	284.8444	36.21293
SeqSeq	NET	Indian	F	241.7754	29.03016
SeqSeq	NET	Indian	F	294.034	34.79093
SeqSeq	NET	Indian	F	293.2369	39.69914
SeqSeq	NET	Welsh	F	291.3897	46.97675
SeqSeq	NET	Welsh	F	295.8888	40.58809
SeqSeq	NET	Welsh	F	270.6501	32.72453
SeqSeq	NET	Welsh	F	271.1586	35.21895
SeqSeq	NET	Welsh	F	266.8954	30.6365
SeqSeq	NET	Welsh	F	291.7181	39.56246
SeqSeq	NET	Welsh	F	266.9488	33.25467
SeqSeq	NET	Welsh	F	279.9824	33.41879
SeqSeq	NET	Welsh	F	257.8645	32.1008
SeqSeq	NET	Welsh	F	255.2061	29.10101
SeqSeq	NET	Irish	F	262.4153	29.36986
SeqSeq	NET	Irish	F	291.7081	41.21797
SeqSeq	NET	Irish	F	343.1341	54.59403
SeqSeq	NET	Irish	F	266.3934	33.59802
SeqSeq	NET	Irish	F	294.915	38.27441
SeqSeq	NET	Irish	F	231.1443	26.6469
SeqSeq	NET	Irish	F	281.3317	36.27623
SeqSeq	NET	Irish	F	282.3797	39.04345
SeqSeq	NET	Irish	F	254.9684	29.924
SeqSeq	NET	Irish	F	286.1785	37.66517
SeqSeq	NET	Irish	M	275.9878	38.34338
SeqSeq	NET	Irish	M	292.6351	43.51208
SeqSeq	NET	Irish	M	310.2764	51.31379
SeqSeq	NET	Irish	M	279.3049	43.48266
SeqSeq	NET	Irish	M	322.2792	54.3163
SeqSeq	NET	Irish	M	286.6775	43.20951

SeqSeq	NET	Irish	M	290.1139	46.47666
SeqSeq	NET	Irish	M	289.2237	45.57505
SeqSeq	NET	Irish	M	303.3574	52.85516
SeqSeq	NET	Irish	M	298.6551	48.21159
SeqSeq	NET	American	F	294.8842	38.5765
SeqSeq	NET	American	F	288.986	35.49733
SeqSeq	NET	American	F	298.9049	43.24544
SeqSeq	NET	American	F	281.2224	41.8022
SeqSeq	NET	American	F	285.3947	39.23
SeqSeq	NET	American	F	329.765	54.13987
SeqSeq	NET	American	M	330.5093	53.76545
SeqSeq	NET	American	M	281.7371	38.41539
SeqSeq	NET	American	M	280.3888	45.00938
SeqSeq	NET	American	M	257.7523	32.10928
SeqSeq	NET	Canadian	M	296.259	46.55272
SeqSeq	NET	Canadian	M	289.2124	35.80828
SeqSeq	NET	Canadian	M	297.0369	46.7823
SeqSeq	NET	Canadian	M	267.8234	34.37
SeqSeq	NET	Canadian	F	302.509	40.64354
SeqSeq	NET	Canadian	F	269.4769	33.70403
SeqSeq	NET	Canadian	F	312.0801	44.56096
SeqSeq	NET	Canadian	F	303.8799	46.53759
SeqSeq	NET	Canadian	F	322.0232	49.29744
SeqSeq	NET	Canadian	F	320.3541	51.12321
SeqSeq	NET	NewZealand	F	313.4196	47.69502
SeqSeq	NET	NewZealand	F	305.0283	42.06796
SeqSeq	NET	NewZealand	F	264.6346	29.37758
SeqSeq	NET	NewZealand	F	296.4614	41.70413
SeqSeq	NET	NewZealand	F	313.2549	40.69824
SeqSeq	NET	NewZealand	F	285.741	34.6053
SeqSeq	NET	NewZealand	F	285.9511	38.40072
SeqSeq	NET	NewZealand	F	324.7872	51.64635
SeqSeq	NET	NewZealand	F	320.2127	46.95407
SeqSeq	NET	NewZealand	F	306.7765	43.6775
SeqSeq	NET	Australian	M	311.3495	49.0127
SeqSeq	NET	Australian	M	304.9315	43.96125
SeqSeq	NET	Australian	M	304.0204	41.72225
SeqSeq	NET	Australian	M	274.8851	37.44436
SeqSeq	NET	Australian	M	304.8166	46.85854
SeqSeq	NET	Australian	M	283.1317	42.42903
SeqSeq	NET	Australian	M	285.3431	38.28059
SeqSeq	NET	Australian	M	301.4841	45.49956
SeqSeq	NET	Australian	M	305.3	42.58304
SeqSeq	NET	Australian	M	299.1963	39.83477
SeqSeq	NET	British	F	276.6756	35.23958
SeqSeq	NET	British	F	268.2131	35.23306
SeqSeq	NET	British	F	265.8257	36.26024
SeqSeq	NET	British	F	288.853	38.98507
SeqSeq	NET	British	F	252.7559	27.6181
SeqSeq	NET	British	F	274.2755	35.64341
SeqSeq	NET	British	F	262.7889	30.95112
SeqSeq	NET	British	F	266.6743	29.41927
SeqSeq	NET	British	F	262.1723	29.74135
SeqSeq	NET	British	F	265.0407	33.98967
SeqSeq	GAN	English	F	291.6912	25.89151
SeqSeq	GAN	English	F	305.168	26.60452
SeqSeq	GAN	English	F	304.4505	29.9787
SeqSeq	GAN	English	F	351.3426	29.93834
SeqSeq	GAN	English	F	337.0165	32.31144
SeqSeq	GAN	English	F	306.9178	28.78436
SeqSeq	GAN	English	M	294.1522	25.41304
SeqSeq	GAN	English	M	295.5525	25.63321
SeqSeq	GAN	English	M	313.9763	29.06339

SeqSeq	GAN	English	M	300.9882	27.66651
SeqSeq	GAN	Scottish	M	317.3253	28.85994
SeqSeq	GAN	Scottish	M	305.9625	29.38556
SeqSeq	GAN	Scottish	M	317.6365	33.95168
SeqSeq	GAN	Scottish	M	306.0369	29.07549
SeqSeq	GAN	Scottish	M	318.5563	30.57809
SeqSeq	GAN	Scottish	M	309.8687	28.11355
SeqSeq	GAN	Scottish	F	293.9316	26.20241
SeqSeq	GAN	Scottish	F	307.8827	27.99458
SeqSeq	GAN	Scottish	F	315.201	29.22869
SeqSeq	GAN	Scottish	F	309.2548	28.21812
SeqSeq	GAN	NorthernIrish	F	314.3803	28.47226
SeqSeq	GAN	NorthernIrish	F	303.5448	27.37375
SeqSeq	GAN	NorthernIrish	F	292.0984	25.99697
SeqSeq	GAN	NorthernIrish	F	327.4698	29.77125
SeqSeq	GAN	NorthernIrish	M	280.0777	28.84556
SeqSeq	GAN	NorthernIrish	M	306.4555	30.15587
SeqSeq	GAN	NorthernIrish	M	305.274	31.0955
SeqSeq	GAN	NorthernIrish	M	332.488	31.75438
SeqSeq	GAN	NorthernIrish	M	288.8314	26.52757
SeqSeq	GAN	NorthernIrish	M	290.4249	25.71222
SeqSeq	GAN	Indian	F	305.2875	28.15216
SeqSeq	GAN	Indian	F	311.839	28.59485
SeqSeq	GAN	Indian	F	320.7329	29.2411
SeqSeq	GAN	Indian	F	296.5852	26.29904
SeqSeq	GAN	Indian	F	316.4513	28.8929
SeqSeq	GAN	Indian	F	301.3346	26.93721
SeqSeq	GAN	Indian	F	317.0805	28.28954
SeqSeq	GAN	Indian	F	303.1624	28.1638
SeqSeq	GAN	Indian	F	303.6636	27.0682
SeqSeq	GAN	Indian	F	303.4418	28.15248
SeqSeq	GAN	Welsh	F	314.1764	32.20087
SeqSeq	GAN	Welsh	F	328.7272	29.95117
SeqSeq	GAN	Welsh	F	306.0339	27.75637
SeqSeq	GAN	Welsh	F	300.1515	29.44031
SeqSeq	GAN	Welsh	F	304.3906	26.80037
SeqSeq	GAN	Welsh	F	318.1864	27.47434
SeqSeq	GAN	Welsh	F	313.5694	30.23612
SeqSeq	GAN	Welsh	F	293.4209	25.31435
SeqSeq	GAN	Welsh	F	305.7396	27.95657
SeqSeq	GAN	Welsh	F	300.4445	26.89344
SeqSeq	GAN	Irish	F	304.4501	26.91291
SeqSeq	GAN	Irish	F	309.3688	31.68702
SeqSeq	GAN	Irish	F	361.4601	41.19121
SeqSeq	GAN	Irish	F	292.1903	23.918
SeqSeq	GAN	Irish	F	326.2015	29.68519
SeqSeq	GAN	Irish	F	292.1073	26.16001
SeqSeq	GAN	Irish	F	304.8637	27.50363
SeqSeq	GAN	Irish	F	348.2742	32.74065
SeqSeq	GAN	Irish	F	309.8265	27.51222
SeqSeq	GAN	Irish	F	323.2844	29.94989
SeqSeq	GAN	Irish	M	335.8857	33.16321
SeqSeq	GAN	Irish	M	336.228	34.10086
SeqSeq	GAN	Irish	M	333.8779	38.56304
SeqSeq	GAN	Irish	M	320.9201	34.62846
SeqSeq	GAN	Irish	M	343.7277	36.68806
SeqSeq	GAN	Irish	M	317.6556	31.07979
SeqSeq	GAN	Irish	M	323.6939	37.26764
SeqSeq	GAN	Irish	M	330.4619	35.69722
SeqSeq	GAN	Irish	M	333.7437	35.25594
SeqSeq	GAN	Irish	M	285.2059	32.17991
SeqSeq	GAN	American	F	301.0442	27.69008
SeqSeq	GAN	American	F	274.1445	25.09159



SeqSeq	GAN	American	F	287.8203	29.94228
SeqSeq	GAN	American	F	303.7487	28.25767
SeqSeq	GAN	American	F	311.357	27.89634
SeqSeq	GAN	American	F	347.5645	34.28188
SeqSeq	GAN	American	M	333.0493	34.71097
SeqSeq	GAN	American	M	296.6092	29.02945
SeqSeq	GAN	American	M	324.079	35.07314
SeqSeq	GAN	American	M	301.1932	26.29405
SeqSeq	GAN	Canadian	M	315.8852	34.21764
SeqSeq	GAN	Canadian	M	286.8353	24.62382
SeqSeq	GAN	Canadian	M	301.7094	30.64146
SeqSeq	GAN	Canadian	M	301.3166	27.67942
SeqSeq	GAN	Canadian	F	305.3256	30.86887
SeqSeq	GAN	Canadian	F	294.6456	26.57501
SeqSeq	GAN	Canadian	F	332.699	33.21985
SeqSeq	GAN	Canadian	F	324.5964	35.99786
SeqSeq	GAN	Canadian	F	316.3296	31.29616
SeqSeq	GAN	Canadian	F	360.7462	42.69593
SeqSeq	GAN	NewZealand	F	320.0988	32.36536
SeqSeq	GAN	NewZealand	F	310.6287	26.98558
SeqSeq	GAN	NewZealand	F	298.7117	25.91699
SeqSeq	GAN	NewZealand	F	328.7103	30.60478
SeqSeq	GAN	NewZealand	F	291.3342	25.97199
SeqSeq	GAN	NewZealand	F	328.5172	30.01006
SeqSeq	GAN	NewZealand	F	314.2676	28.17836
SeqSeq	GAN	NewZealand	F	355.0931	40.20447
SeqSeq	GAN	NewZealand	F	354.2725	34.36875
SeqSeq	GAN	NewZealand	F	293.8482	27.86919
SeqSeq	GAN	Australian	M	343.4646	37.85342
SeqSeq	GAN	Australian	M	300.6783	29.67748
SeqSeq	GAN	Australian	M	321.3024	32.49103
SeqSeq	GAN	Australian	M	288.6564	27.35665
SeqSeq	GAN	Australian	M	315.9595	34.04763
SeqSeq	GAN	Australian	M	301.4981	33.67131
SeqSeq	GAN	Australian	M	304.7884	29.17539
SeqSeq	GAN	Australian	M	322.5875	30.53114
SeqSeq	GAN	Australian	M	312.1641	27.8823
SeqSeq	GAN	Australian	M	314.9025	28.87418
SeqSeq	GAN	British	F	310.2096	26.65843
SeqSeq	GAN	British	F	311.2386	30.26038
SeqSeq	GAN	British	F	327.7542	29.01452
SeqSeq	GAN	British	F	327.9098	27.88051
SeqSeq	GAN	British	F	320.5529	27.86428
SeqSeq	GAN	British	F	295.377	26.56327
SeqSeq	GAN	British	F	307.9421	26.49203
SeqSeq	GAN	British	F	298.9153	25.57832
SeqSeq	GAN	British	F	326.4713	28.65913
SeqSeq	GAN	British	F	284.4128	24.89951
Transformers	RNN	English	F	260.3799	30.02154
Transformers	RNN	English	F	260.9638	31.7087
Transformers	RNN	English	F	310.5578	37.7368
Transformers	RNN	English	F	359.6662	37.08592
Transformers	RNN	English	F	290.4194	34.96936
Transformers	RNN	English	F	278.9194	33.82704
Transformers	RNN	English	M	297.5127	33.49255
Transformers	RNN	English	M	272.6088	31.45284
Transformers	RNN	English	M	269.3785	30.92981
Transformers	RNN	English	M	312.9836	35.11102
Transformers	RNN	Scottish	M	321.5176	36.23052
Transformers	RNN	Scottish	M	261.9694	34.56777
Transformers	RNN	Scottish	M	334.711	39.34938
Transformers	RNN	Scottish	M	286.7405	32.22967
Transformers	RNN	Scottish	M	296.0772	36.13046

Transformers	RNN	Scottish	M	291.2472	34.18153
Transformers	RNN	Scottish	F	265.5407	32.51915
Transformers	RNN	Scottish	F	273.2494	31.95457
Transformers	RNN	Scottish	F	298.0945	35.20948
Transformers	RNN	Scottish	F	294.1194	35.43943
Transformers	RNN	NorthernIrish	F	279.4613	34.77303
Transformers	RNN	NorthernIrish	F	268.2895	34.28716
Transformers	RNN	NorthernIrish	F	286.5145	33.57223
Transformers	RNN	NorthernIrish	F	375.6408	46.35688
Transformers	RNN	NorthernIrish	M	306.1705	39.84719
Transformers	RNN	NorthernIrish	M	317.3793	42.71001
Transformers	RNN	NorthernIrish	M	326.2437	38.26785
Transformers	RNN	NorthernIrish	M	360.9236	46.3488
Transformers	RNN	NorthernIrish	M	296.9236	36.52649
Transformers	RNN	NorthernIrish	M	260.0272	31.90167
Transformers	RNN	Indian	F	325.6499	40.15692
Transformers	RNN	Indian	F	267.2869	35.46554
Transformers	RNN	Indian	F	307.5809	35.84544
Transformers	RNN	Indian	F	309.6869	36.04488
Transformers	RNN	Indian	F	256.1896	32.53725
Transformers	RNN	Indian	F	302.4248	36.84725
Transformers	RNN	Indian	F	316.7225	32.58582
Transformers	RNN	Indian	F	319.7948	37.51697
Transformers	RNN	Indian	F	294.1597	34.06716
Transformers	RNN	Indian	F	332.9633	40.15636
Transformers	RNN	Welsh	F	364.353	46.91721
Transformers	RNN	Welsh	F	372.7736	45.21051
Transformers	RNN	Welsh	F	306.0236	35.42592
Transformers	RNN	Welsh	F	318.6635	36.78993
Transformers	RNN	Welsh	F	284.9596	31.00399
Transformers	RNN	Welsh	F	275.8669	29.51745
Transformers	RNN	Welsh	F	277.2458	35.14555
Transformers	RNN	Welsh	F	299.8904	31.38493
Transformers	RNN	Welsh	F	291.7575	33.1723
Transformers	RNN	Welsh	F	251.7896	31.14471
Transformers	RNN	Irish	F	276.1844	28.75827
Transformers	RNN	Irish	F	321.0232	39.40724
Transformers	RNN	Irish	F	394.2239	51.9029
Transformers	RNN	Irish	F	271.1082	29.17481
Transformers	RNN	Irish	F	294.3906	34.90329
Transformers	RNN	Irish	F	260.7135	28.82402
Transformers	RNN	Irish	F	305.01	37.23314
Transformers	RNN	Irish	F	292.6204	42.85366
Transformers	RNN	Irish	F	283.3079	33.5188
Transformers	RNN	Irish	F	290.0571	35.91616
Transformers	RNN	Irish	M	326.622	42.45082
Transformers	RNN	Irish	M	273.4116	35.62862
Transformers	RNN	Irish	M	304.8015	44.52589
Transformers	RNN	Irish	M	301.2356	39.94515
Transformers	RNN	Irish	M	288.5488	43.38083
Transformers	RNN	Irish	M	293.5985	37.14486
Transformers	RNN	Irish	M	269.5083	34.83251
Transformers	RNN	Irish	M	302.33	43.16847
Transformers	RNN	Irish	M	288.1461	39.0569
Transformers	RNN	Irish	M	289.3168	40.24462
Transformers	RNN	American	F	255.9949	31.99937
Transformers	RNN	American	F	299.9251	36.05113
Transformers	RNN	American	F	290.4845	37.76287
Transformers	RNN	American	F	293.3677	39.62923
Transformers	RNN	American	F	315.9406	37.13428
Transformers	RNN	American	F	383.6575	48.06765
Transformers	RNN	American	M	296.3655	40.71579
Transformers	RNN	American	M	294.6605	37.0984

Transformers	RNN	American	M	347.667	48.91279
Transformers	RNN	American	M	275.607	29.20129
Transformers	RNN	Canadian	M	322.3232	40.25586
Transformers	RNN	Canadian	M	272.8255	30.18183
Transformers	RNN	Canadian	M	335.4238	43.51154
Transformers	RNN	Canadian	M	310.4514	36.35465
Transformers	RNN	Canadian	F	324.5533	40.05473
Transformers	RNN	Canadian	F	281.9556	34.09995
Transformers	RNN	Canadian	F	329.9191	39.98169
Transformers	RNN	Canadian	F	275.2954	38.02515
Transformers	RNN	Canadian	F	343.6708	42.87375
Transformers	RNN	Canadian	F	345.8899	46.8854
Transformers	RNN	NewZealand	F	295.2897	39.58217
Transformers	RNN	NewZealand	F	289.4121	35.822
Transformers	RNN	NewZealand	F	277.9962	32.07537
Transformers	RNN	NewZealand	F	297.0721	33.30819
Transformers	RNN	NewZealand	F	328.5334	38.47115
Transformers	RNN	NewZealand	F	296.0346	36.40699
Transformers	RNN	NewZealand	F	288.163	36.98545
Transformers	RNN	NewZealand	F	370.8704	49.26068
Transformers	RNN	NewZealand	F	362.2488	44.87347
Transformers	RNN	NewZealand	F	310.5138	37.03005
Transformers	RNN	Australian	M	383.2427	52.04856
Transformers	RNN	Australian	M	330.9115	39.43051
Transformers	RNN	Australian	M	314.2385	37.50971
Transformers	RNN	Australian	M	261.4906	31.17269
Transformers	RNN	Australian	M	301.4445	45.05277
Transformers	RNN	Australian	M	323.2881	42.39913
Transformers	RNN	Australian	M	319.5202	39.13591
Transformers	RNN	Australian	M	312.3754	36.92452
Transformers	RNN	Australian	M	325.4268	36.1255
Transformers	RNN	Australian	M	315.7601	37.38139
Transformers	RNN	British	F	310.449	33.66387
Transformers	RNN	British	F	313.2741	37.74027
Transformers	RNN	British	F	321.0585	39.44936
Transformers	RNN	British	F	324.6578	39.31625
Transformers	RNN	British	F	260.6014	33.08647
Transformers	RNN	British	F	253.0311	32.43106
Transformers	RNN	British	F	281.4615	31.45585
Transformers	RNN	British	F	261.9505	30.93158
Transformers	RNN	British	F	274.864	32.12665
Transformers	RNN	British	F	293.6362	30.47682
Transformers	GAN	English	F	354.5152	48.1902
Transformers	NET	English	F	306.389	46.3481
Transformers	GAN	English	F	363.187	50.30756
Transformers	NET	English	F	318.7414	48.2593
Transformers	GAN	English	F	338.341	44.40368
Transformers	NET	English	F	306.8248	44.23299
Transformers	GAN	English	F	397.1735	47.4544
Transformers	NET	English	F	355.7189	46.64744
Transformers	GAN	English	F	383.4723	56.79027
Transformers	NET	English	F	329.2024	55.78763
Transformers	GAN	English	F	361.852	51.33228
Transformers	NET	English	F	299.7132	48.03812
Transformers	GAN	English	M	359.6645	50.37292
Transformers	NET	English	M	306.4768	47.29799
Transformers	GAN	English	M	348.9449	47.09786
Transformers	NET	English	M	314.6386	47.04111
Transformers	GAN	English	M	364.725	53.07667
Transformers	NET	English	M	313.9341	49.50714
Transformers	GAN	English	M	379.4569	53.26476
Transformers	NET	English	M	315.058	51.36372
Transformers	GAN	Scottish	M	361.9591	49.97701

Transformers	NET	Scottish	M	307.3942	47.75153
Transformers	GAN	Scottish	M	379.5715	56.27201
Transformers	NET	Scottish	M	310.1558	53.03727
Transformers	GAN	Scottish	M	356.3777	51.24973
Transformers	NET	Scottish	M	312.2721	50.22478
Transformers	GAN	Scottish	M	369.5298	52.31679
Transformers	NET	Scottish	M	288.0675	48.08943
Transformers	GAN	Scottish	M	376.5962	54.60039
Transformers	NET	Scottish	M	311.9183	51.27698
Transformers	GAN	Scottish	M	354.6551	53.75478
Transformers	NET	Scottish	M	297.6893	49.67167
Transformers	GAN	Scottish	F	353.8378	46.75234
Transformers	NET	Scottish	F	306.1213	45.76981
Transformers	GAN	Scottish	F	356.6932	50.14997
Transformers	NET	Scottish	F	306.0949	48.78853
Transformers	GAN	Scottish	F	348.8316	47.76088
Transformers	NET	Scottish	F	298.2073	45.48729
Transformers	GAN	Scottish	F	346.0404	45.14817
Transformers	NET	Scottish	F	291.6783	42.61665
Transformers	GAN	NorthernIrish	F	354.4242	48.02983
Transformers	NET	NorthernIrish	F	308.1759	46.89008
Transformers	GAN	NorthernIrish	F	358.2883	49.2604
Transformers	NET	NorthernIrish	F	310.742	47.88774
Transformers	GAN	NorthernIrish	F	379.6515	50.73026
Transformers	NET	NorthernIrish	F	309.7521	46.93618
Transformers	GAN	NorthernIrish	F	301.7794	44.44403
Transformers	NET	NorthernIrish	F	269.7289	45.03781
Transformers	GAN	NorthernIrish	M	286.2734	45.06699
Transformers	NET	NorthernIrish	M	239.1754	44.50059
Transformers	GAN	NorthernIrish	M	268.8125	44.98263
Transformers	NET	NorthernIrish	M	233.8987	46.13985
Transformers	GAN	NorthernIrish	M	320.7198	46.0782
Transformers	NET	NorthernIrish	M	270.1666	44.34994
Transformers	GAN	NorthernIrish	M	289.2966	44.81602
Transformers	NET	NorthernIrish	M	252.7656	46.06728
Transformers	GAN	NorthernIrish	M	333.3536	46.33494
Transformers	NET	NorthernIrish	M	293.7633	45.29899
Transformers	GAN	NorthernIrish	M	342.8737	49.04577
Transformers	NET	NorthernIrish	M	292.5775	47.41735
Transformers	GAN	Indian	F	296.2456	44.19154
Transformers	NET	Indian	F	251.8112	44.42071
Transformers	GAN	Indian	F	366.1498	49.05686
Transformers	NET	Indian	F	314.6425	48.03532
Transformers	GAN	Indian	F	338.1255	48.16986
Transformers	NET	Indian	F	299.3215	48.17055
Transformers	GAN	Indian	F	356.462	51.63532
Transformers	NET	Indian	F	318.2463	53.54467
Transformers	GAN	Indian	F	402.5732	55.69974
Transformers	NET	Indian	F	338.7794	52.43288
Transformers	GAN	Indian	F	394.0965	54.30338
Transformers	NET	Indian	F	303.3549	49.66515
Transformers	GAN	Indian	F	396.1581	52.87913
Transformers	NET	Indian	F	310.2435	48.40955
Transformers	GAN	Indian	F	305.0916	43.13242
Transformers	NET	Indian	F	253.1066	42.14113
Transformers	GAN	Indian	F	366.6336	50.44089
Transformers	NET	Indian	F	321.1492	51.2974
Transformers	GAN	Indian	F	331.7667	50.57145
Transformers	NET	Indian	F	284.5043	51.24218
Transformers	GAN	Welsh	F	247.6204	37.0408
Transformers	NET	Welsh	F	216.2803	38.55959
Transformers	GAN	Welsh	F	273.3991	37.05993
Transformers	NET	Welsh	F	239.6279	38.30344

Transformers	GAN	Welsh	F	337.3189	49.81173
Transformers	NET	Welsh	F	290.8068	48.79731
Transformers	GAN	Welsh	F	320.3359	50.72352
Transformers	NET	Welsh	F	273.229	50.58702
Transformers	GAN	Welsh	F	394.2524	50.46373
Transformers	NET	Welsh	F	327.2459	47.68346
Transformers	GAN	Welsh	F	331.8397	50.01515
Transformers	NET	Welsh	F	279.0405	48.23288
Transformers	GAN	Welsh	F	364.8076	52.78195
Transformers	NET	Welsh	F	310.1812	50.66618
Transformers	GAN	Welsh	F	387.7647	50.72793
Transformers	NET	Welsh	F	318.4765	48.35331
Transformers	GAN	Welsh	F	340.3747	45.66407
Transformers	NET	Welsh	F	272.8363	41.4612
Transformers	GAN	Welsh	F	346.3019	48.34206
Transformers	NET	Welsh	F	301.424	46.34938
Transformers	GAN	Irish	F	351.4785	47.02007
Transformers	NET	Irish	F	309.4596	47.1365
Transformers	GAN	Irish	F	356.269	53.18171
Transformers	NET	Irish	F	299.6665	52.01442
Transformers	GAN	Irish	F	298.0941	43.26601
Transformers	NET	Irish	F	264.837	44.81537
Transformers	GAN	Irish	F	344.5404	42.26706
Transformers	NET	Irish	F	304.9005	39.6868
Transformers	GAN	Irish	F	397.1021	55.36223
Transformers	NET	Irish	F	343.976	53.50725
Transformers	GAN	Irish	F	339.1539	43.86724
Transformers	NET	Irish	F	299.365	41.17888
Transformers	GAN	Irish	F	373.2037	50.99187
Transformers	NET	Irish	F	309.6748	49.50341
Transformers	GAN	Irish	F	280.3541	38.72766
Transformers	NET	Irish	F	247.5077	40.10365
Transformers	GAN	Irish	F	365.5244	48.35408
Transformers	NET	Irish	F	309.1473	46.42284
Transformers	GAN	Irish	F	402.3516	56.05197
Transformers	NET	Irish	F	344.2194	54.03632
Transformers	GAN	Irish	M	321.168	52.72239
Transformers	NET	Irish	M	275.4824	50.44667
Transformers	GAN	Irish	M	310.2577	48.90403
Transformers	NET	Irish	M	254.8747	44.6752
Transformers	GAN	Irish	M	310.4721	55.91557
Transformers	NET	Irish	M	249.1647	51.57415
Transformers	GAN	Irish	M	320.4627	57.91275
Transformers	NET	Irish	M	293.287	61.15867
Transformers	GAN	Irish	M	325.3476	58.64937
Transformers	NET	Irish	M	277.0019	56.89716
Transformers	GAN	Irish	M	306.304	47.95522
Transformers	NET	Irish	M	263.206	50.07873
Transformers	GAN	Irish	M	271.5359	44.38236
Transformers	NET	Irish	M	251.8816	48.67786
Transformers	GAN	Irish	M	379.179	63.17799
Transformers	NET	Irish	M	327.7796	62.26316
Transformers	GAN	Irish	M	287.0625	53.07357
Transformers	NET	Irish	M	243.9346	50.1738
Transformers	GAN	Irish	M	296.2452	52.83318
Transformers	NET	Irish	M	274.4317	55.36588
Transformers	GAN	American	F	336.7234	47.09667
Transformers	NET	American	F	299.1698	48.17296
Transformers	GAN	American	F	348.019	48.21147
Transformers	NET	American	F	295.2267	45.99643
Transformers	GAN	American	F	364.0433	60.25457
Transformers	NET	American	F	322.1764	58.67645
Transformers	GAN	American	F	248.2334	36.54812

Transformers	NET	American	F	214.011	38.40878
Transformers	GAN	American	F	310.6999	46.21557
Transformers	NET	American	F	257.4993	43.63504
Transformers	GAN	American	F	298.1875	44.97812
Transformers	NET	American	F	264.0753	45.64035
Transformers	GAN	American	M	261.1035	44.937
Transformers	NET	American	M	227.2653	45.89506
Transformers	GAN	American	M	354.5253	51.4626
Transformers	NET	American	M	313.7055	49.47982
Transformers	GAN	American	M	262.0896	45.54238
Transformers	NET	American	M	231.7293	47.4137
Transformers	GAN	American	M	337.5388	47.19696
Transformers	NET	American	M	289.3765	46.27655
Transformers	GAN	Canadian	M	362.231	55.56689
Transformers	NET	Canadian	M	292.7354	52.71186
Transformers	GAN	Canadian	M	347.09	43.82057
Transformers	NET	Canadian	M	290.0231	41.67742
Transformers	GAN	Canadian	M	362.231	55.56689
Transformers	NET	Canadian	M	292.8677	52.65807
Transformers	GAN	Canadian	M	358.7974	50.31469
Transformers	NET	Canadian	M	301.1387	48.6127
Transformers	GAN	Canadian	F	367.6551	54.61327
Transformers	NET	Canadian	F	322.7284	54.52325
Transformers	GAN	Canadian	F	360.481	47.92453
Transformers	NET	Canadian	F	298.8356	44.72714
Transformers	GAN	Canadian	F	346.5544	54.50652
Transformers	NET	Canadian	F	300.1661	53.99463
Transformers	GAN	Canadian	F	387.5821	60.80925
Transformers	NET	Canadian	F	325.7348	58.60194
Transformers	GAN	Canadian	F	299.6736	50.17915
Transformers	NET	Canadian	F	262.1451	48.71521
Transformers	GAN	Canadian	F	280.4502	47.93588
Transformers	NET	Canadian	F	245.1763	48.8317
Transformers	GAN	NewZealand	F	405.9568	60.08915
Transformers	NET	NewZealand	F	340.9038	57.82262
Transformers	GAN	NewZealand	F	339.3836	46.30469
Transformers	NET	NewZealand	F	310.6653	45.2395
Transformers	GAN	NewZealand	F	334.642	42.41229
Transformers	NET	NewZealand	F	287.4918	41.07287
Transformers	GAN	NewZealand	F	382.0209	53.87278
Transformers	NET	NewZealand	F	312.5834	50.60905
Transformers	GAN	NewZealand	F	343.1194	47.83747
Transformers	NET	NewZealand	F	307.1567	50.1019
Transformers	GAN	NewZealand	F	371.782	48.43074
Transformers	NET	NewZealand	F	311.9005	46.52716
Transformers	GAN	NewZealand	F	411.7567	58.16092
Transformers	NET	NewZealand	F	348.6248	55.47392
Transformers	GAN	NewZealand	F	299.9756	48.63644
Transformers	NET	NewZealand	F	259.7138	49.13895
Transformers	GAN	NewZealand	F	311.7297	47.23522
Transformers	NET	NewZealand	F	271.4232	46.7844
Transformers	GAN	NewZealand	F	378.7999	55.94205
Transformers	NET	NewZealand	F	306.9125	52.3048
Transformers	GAN	Australian	M	283.8211	43.54391
Transformers	NET	Australian	M	250.7034	44.62418
Transformers	GAN	Australian	M	339.8149	52.31504
Transformers	NET	Australian	M	293.0182	51.12087
Transformers	GAN	Australian	M	363.9527	51.31463
Transformers	NET	Australian	M	313.3819	50.72991
Transformers	GAN	Australian	M	339.5471	48.6918
Transformers	NET	Australian	M	285.8385	47.40236
Transformers	GAN	Australian	M	251.2152	41.55796
Transformers	NET	Australian	M	211.9587	42.56046

Transformers	GAN	Australian	M	317.1669	50.72237
Transformers	NET	Australian	M	266.7863	49.16807
Transformers	GAN	Australian	M	366.7343	51.63689
Transformers	NET	Australian	M	313.6683	49.86481
Transformers	GAN	Australian	M	335.5149	52.86042
Transformers	NET	Australian	M	276.9423	50.38032
Transformers	GAN	Australian	M	380.742	52.8532
Transformers	NET	Australian	M	303.2577	50.75069
Transformers	GAN	Australian	M	382.8767	54.73135
Transformers	NET	Australian	M	331.6159	53.65729
Transformers	GAN	British	F	325.8405	44.18391
Transformers	NET	British	F	281.5559	44.18032
Transformers	GAN	British	F	310.8056	45.46365
Transformers	NET	British	F	268.3245	46.60167
Transformers	GAN	British	F	272.961	40.60774
Transformers	NET	British	F	240.3848	41.60835
Transformers	GAN	British	F	286.1368	39.8401
Transformers	NET	British	F	252.5643	40.8439
Transformers	GAN	British	F	347.7742	45.50037
Transformers	NET	British	F	290.8633	42.54111
Transformers	GAN	British	F	345.72	47.38045
Transformers	NET	British	F	288.7368	45.19894
Transformers	GAN	British	F	395.2488	51.0973
Transformers	NET	British	F	325.5037	49.43621
Transformers	GAN	British	F	358.5045	47.37684
Transformers	NET	British	F	304.8731	46.14535
Transformers	GAN	British	F	369.931	49.02641
Transformers	NET	British	F	312.8346	47.36878
Transformers	GAN	British	F	360.8411	47.41605
Transformers	NET	British	F	278.2704	43.12532
CVAE	RNN	English	F	260.3799	30.02154
CVAE	RNN	English	F	260.9638	31.7087
CVAE	RNN	English	F	310.5578	37.7368
CVAE	RNN	English	F	359.6662	37.08592
CVAE	RNN	English	F	290.4194	34.96936
CVAE	RNN	English	F	278.9194	33.82704
CVAE	RNN	English	M	297.5127	33.49255
CVAE	RNN	English	M	272.6088	31.45284
CVAE	RNN	English	M	269.3785	30.92981
CVAE	RNN	English	M	312.9836	35.11102
CVAE	RNN	Scottish	M	321.5176	36.23052
CVAE	RNN	Scottish	M	261.9694	34.56777
CVAE	RNN	Scottish	M	334.711	39.34938
CVAE	RNN	Scottish	M	286.7405	32.22968
CVAE	RNN	Scottish	M	296.0772	36.13046
CVAE	RNN	Scottish	M	291.2472	34.18154
CVAE	RNN	Scottish	F	265.5407	32.51915
CVAE	RNN	Scottish	F	273.2495	31.95457
CVAE	RNN	Scottish	F	298.0945	35.20948
CVAE	RNN	Scottish	F	294.1194	35.43943
CVAE	RNN	NorthernIrish	F	279.4613	34.77304
CVAE	RNN	NorthernIrish	F	268.2895	34.28716
CVAE	RNN	NorthernIrish	F	286.5145	33.57223
CVAE	RNN	NorthernIrish	F	375.6408	46.35688
CVAE	RNN	NorthernIrish	M	306.1704	39.84718
CVAE	RNN	NorthernIrish	M	317.3793	42.71001
CVAE	RNN	NorthernIrish	M	326.2438	38.26785
CVAE	RNN	NorthernIrish	M	360.9236	46.3488
CVAE	RNN	NorthernIrish	M	296.9236	36.52649
CVAE	RNN	NorthernIrish	M	260.0272	31.90167
CVAE	RNN	Indian	F	325.6499	40.15692
CVAE	RNN	Indian	F	267.2869	35.46554
CVAE	RNN	Indian	F	307.5809	35.84544

CVAE	RNN	Indian	F	309.6869	36.04488
CVAE	RNN	Indian	F	256.1896	32.53724
CVAE	RNN	Indian	F	302.4248	36.84725
CVAE	RNN	Indian	F	316.7225	32.58583
CVAE	RNN	Indian	F	319.7948	37.51697
CVAE	RNN	Indian	F	294.1597	34.06716
CVAE	RNN	Indian	F	332.9633	40.15636
CVAE	RNN	Welsh	F	364.353	46.91721
CVAE	RNN	Welsh	F	372.7736	45.21051
CVAE	RNN	Welsh	F	306.0236	35.42592
CVAE	RNN	Welsh	F	318.6635	36.78992
CVAE	RNN	Welsh	F	284.9596	31.00399
CVAE	RNN	Welsh	F	275.8669	29.51745
CVAE	RNN	Welsh	F	277.2458	35.14555
CVAE	RNN	Welsh	F	299.8904	31.38492
CVAE	RNN	Welsh	F	291.7575	33.1723
CVAE	RNN	Welsh	F	251.7896	31.14471
CVAE	RNN	Irish	F	276.1844	28.75827
CVAE	RNN	Irish	F	321.0232	39.40724
CVAE	RNN	Irish	F	394.2239	51.9029
CVAE	RNN	Irish	F	271.1082	29.17481
CVAE	RNN	Irish	F	294.3906	34.90329
CVAE	RNN	Irish	F	260.7135	28.82402
CVAE	RNN	Irish	F	305.01	37.23314
CVAE	RNN	Irish	F	292.6204	42.85366
CVAE	RNN	Irish	F	283.3079	33.5188
CVAE	RNN	Irish	F	290.0571	35.91616
CVAE	RNN	Irish	M	326.622	42.45082
CVAE	RNN	Irish	M	273.4116	35.62862
CVAE	RNN	Irish	M	304.8015	44.52589
CVAE	RNN	Irish	M	301.2356	39.94515
CVAE	RNN	Irish	M	288.5488	43.38083
CVAE	RNN	Irish	M	293.5985	37.14485
CVAE	RNN	Irish	M	269.5083	34.83251
CVAE	RNN	Irish	M	302.33	43.16847
CVAE	RNN	Irish	M	288.1461	39.05691
CVAE	RNN	Irish	M	289.3168	40.24462
CVAE	RNN	American	F	255.9949	31.99937
CVAE	RNN	American	F	299.9251	36.05113
CVAE	RNN	American	F	290.4845	37.76286
CVAE	RNN	American	F	293.3677	39.62923
CVAE	RNN	American	F	315.9406	37.13428
CVAE	RNN	American	F	383.6575	48.06765
CVAE	RNN	American	M	296.3655	40.7158
CVAE	RNN	American	M	294.6606	37.0984
CVAE	RNN	American	M	347.667	48.91279
CVAE	RNN	American	M	275.607	29.20129
CVAE	RNN	Canadian	M	322.3232	40.25586
CVAE	RNN	Canadian	M	272.8255	30.18183
CVAE	RNN	Canadian	M	335.4238	43.51153
CVAE	RNN	Canadian	M	310.4514	36.35465
CVAE	RNN	Canadian	F	324.5533	40.05473
CVAE	RNN	Canadian	F	281.9556	34.09995
CVAE	RNN	Canadian	F	329.9191	39.98168
CVAE	RNN	Canadian	F	275.2954	38.02515
CVAE	RNN	Canadian	F	343.6708	42.87374
CVAE	RNN	Canadian	F	345.8899	46.8854
CVAE	RNN	NewZealand	F	295.2897	39.58217
CVAE	RNN	NewZealand	F	289.4121	35.822
CVAE	RNN	NewZealand	F	277.9962	32.07537
CVAE	RNN	NewZealand	F	297.0721	33.30819
CVAE	RNN	NewZealand	F	328.5334	38.47115
CVAE	RNN	NewZealand	F	296.0347	36.40699



CVAE	RNN	NewZealand	F	288.163	36.98545
CVAE	RNN	NewZealand	F	370.8704	49.26068
CVAE	RNN	NewZealand	F	362.2488	44.87347
CVAE	RNN	NewZealand	F	310.5138	37.03005
CVAE	RNN	Australian	M	383.2427	52.04856
CVAE	RNN	Australian	M	330.9115	39.43052
CVAE	RNN	Australian	M	314.2385	37.50971
CVAE	RNN	Australian	M	261.4906	31.17269
CVAE	RNN	Australian	M	301.4445	45.05277
CVAE	RNN	Australian	M	323.2881	42.39913
CVAE	RNN	Australian	M	319.5202	39.13591
CVAE	RNN	Australian	M	312.3754	36.92453
CVAE	RNN	Australian	M	325.4268	36.1255
CVAE	RNN	Australian	M	315.7601	37.38139
CVAE	RNN	British	F	310.449	33.66387
CVAE	RNN	British	F	313.2741	37.74027
CVAE	RNN	British	F	321.0585	39.44936
CVAE	RNN	British	F	324.6578	39.31625
CVAE	RNN	British	F	260.6014	33.08647
CVAE	RNN	British	F	253.0311	32.43106
CVAE	RNN	British	F	281.4615	31.45585
CVAE	RNN	British	F	261.9505	30.93158
CVAE	RNN	British	F	274.864	32.12665
CVAE	RNN	British	F	293.6362	30.47682