



VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

FUNDAMENTINIŲ MOKSLŲ FAKULTETAS
INFORMACINIŲ TECHNOLOGIJŲ KATEDRA

Matas Malickas

**VEIDO SAVYBIŲ NUSTATYMAS VEIDŲ ŠABLONUOSE TAIKANT DIRBTINIO
INTELEKTO METODUS**

**IDENTIFICATION OF FACIAL FEATURES IN FACE TEMPLATES USING
ARTIFICIAL INTELLIGENCE METHODS**

Baigiamasis magistro darbas

Informacinių technologijų studijų programa, valstybinis kodas 6211BX016

Informatikos inžinerijos studijų kryptis

Duomenų gavybos technologijų specializacija

Vilnius, 2025

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS
FUNDAMENTINIŲ MOKSLŲ FAKULTETAS
INFORMACINIŲ TECHNOLOGIJŲ KATEDRA

TVIRTINU
Katedros vedėjas

(Parašas)

(Vardas, pavardė)

(Data)

Matas Malickas

**VEIDO SAVYBIŲ NUSTATYMAS VEIDŲ ŠABLONUOSE TAIKANT
DIRBTINIO INTELEKTO METODUS**

**IDENTIFICATION OF FACIAL FEATURES IN FACE TEMPLATES USING
ARTIFICIAL INTELLIGENCE METHODS**

Baigiamasis magistro darbas

Informacinių technologijų studijų programa, valstybinis kodas 6211BX016

Informatikos inžinerijos studijų kryptis

Duomenų gavybos technologijų specializacija

Vadovas: doc. dr. Pavel Stefanovič _____
(Pedag. vardas, vardas, pavardė) (Parašas) (Data)

Vilnius, 2025

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS
FUNDAMENTINIŲ MOKSLŲ FAKULTETAS
INFORMACINIŲ TECHNOLOGIJŲ KATEDRA

Studijų kryptis: Informatikos inžinerija

TVIRTINU

Studijų programa: Informacinės technologijos , valstybinis kodas 6211BX016 Katedros vedėjas

Specializacija: Duomenų gavybos technologijos

Dmitrij Šešok

2025-05-16

MAGISTRO BAIGIAMOJO DARBO UŽDUOTIS

Nr. DGTfm-23–10641

Vilnius

Studentas (-ė): Matas Malickas

Baigiamojo darbo tema: Veido savybių nustatymas veidų šablonuose taikant dirbtinio intelekto metodus Baigiamojo darbo užbaigimo terminas pagal numatytą studijų kalendorinį grafiką.

BAIGIAMOJO DARBO UŽDUOTIS:

Tikslas: Ištirti mašininio mokymo algoritmų efektyvumą atpažįstant veido savybes naudojant išgautus savybių vektorius.

Uždaviniai:

1. Išanalizuoti veido šablonų išgavimo metodus taikomus panašiems uždaviniams spręsti.
2. Iš anksto apdoroti ir paruošti duomenų aibę eksperimentiniam tyrimui.
3. Eksperimentiškai ištirti skirtingų algoritmų efektyvumą atpažįstant vaizdo savybes iš veidų savybių vektoriaus.

Planuojami rezultatai: Paruošta nauja duomenų aibė. Eksperimentiškai ištirti skirtingi algoritmai sprendžiama uždaviniui ir įvertintas jų efektyvumas tokio tipo uždaviniams.

Vadovas docentas Pavel Stefanovič

<table border="1"> <tr> <td>Vilniaus Gedimino technikos universitetas</td> <td rowspan="3"></td> <td colspan="2"> <table border="1"> <tr> <td>ISBN</td> <td>ISSN</td> </tr> <tr> <td>Egz. sk.</td> <td rowspan="2"></td> </tr> <tr> <td>Data-.....-.....</td> </tr> </table> </td> </tr> <tr> <td>Fundamentinių mokslų fakultetas</td> <td></td> </tr> <tr> <td>Informacinių technologijų katedra</td> <td></td> </tr> </table>	Vilniaus Gedimino technikos universitetas		<table border="1"> <tr> <td>ISBN</td> <td>ISSN</td> </tr> <tr> <td>Egz. sk.</td> <td rowspan="2"></td> </tr> <tr> <td>Data-.....-.....</td> </tr> </table>		ISBN	ISSN	Egz. sk.		Data-.....-.....	Fundamentinių mokslų fakultetas		Informacinių technologijų katedra		Antrosios pakopos studijų Informacinių technologijų programos magistro baigiamasis darbas 4		
Vilniaus Gedimino technikos universitetas			<table border="1"> <tr> <td>ISBN</td> <td>ISSN</td> </tr> <tr> <td>Egz. sk.</td> <td rowspan="2"></td> </tr> <tr> <td>Data-.....-.....</td> </tr> </table>		ISBN	ISSN	Egz. sk.			Data-.....-.....						
ISBN			ISSN													
Egz. sk.																
Data-.....-.....																
Fundamentinių mokslų fakultetas																
Informacinių technologijų katedra																
Pavadinimas	Veido savybių nustatymas veidų šablonuose taikant dirbtinio intelekto metodus															
Autorius	Matas Malickas															
Vadovas	Pavel Stefanovič															
<table border="1"> <tr> <td></td> <td>Kalba: lietuvių</td> </tr> </table>					Kalba: lietuvių											
	Kalba: lietuvių															
Anotacija																
Šiame magistro darbe nagrinėjamas veido savybių nustatymas iš veidų šablonų taikant dirbtinio intelekto metodus. Pirmoje darbo dalyje atlikta veidų atpažinimo technologijų bei veidų šablonų generavimo metodų apžvalga, ypatingą dėmesį skiriant veido savybių vektorių formavimui. Antroje darbo dalyje pasiūlyti keli metodai, leidžiantys klasifikuoti pradinio vaizdo savybes remiantis tik veidų šablonais. Eksperimentams atlikti buvo panaudoti tiek gilieji neuroniniai tinklai, tiek klasikiniai prižiūrimojo mokymo algoritmai. Eksperimentų metu nustatyta, kad lyties atributas yra prognozuojamas tiksliausiai ir geriausias modelis pasiekė 93 % tikslumą. Etninės grupės ir amžiaus klasifikavimas iš veidų šablonų yra sudėtingesnis uždavinys, tikėtina dėl to, kad veidų šablonų generavimo modelių architektūros yra kuriamos taip, kad būtų pašalinama visa informacija, nesusijusi su asmens tapatybe. Palyginus skirtingus mašininio mokymo metodus, nustatyta, kad gilieji neuroniniai tinklai yra labiau tinkami šiam uždaviniui spręsti, lyginant juos su klasikiais klasifikavimo algoritmais.																
Prasminiai žodžiai: veidų savybių vektoriai, veidų šablonai, veidų atpažinimas, veido savybių klasifikavimas, neuroniniai tinklai, prižiūrimasis mokymas																

<table border="1"> <tr> <td>Vilnius Gediminas Technical University</td> <td rowspan="3"></td> <td colspan="2"> <table border="1"> <tr> <td>ISBN</td> <td>ISSN</td> </tr> <tr> <td colspan="2">Copies</td> </tr> <tr> <td colspan="2">No.</td> </tr> <tr> <td colspan="2">Date</td> </tr> <tr> <td colspan="2">.....-.....-.....</td> </tr> </table> </td> </tr> <tr> <td>Faculty of Fundamental Sciences</td> <td></td> </tr> <tr> <td>Department of Information Technologies</td> <td></td> </tr> </table>		Vilnius Gediminas Technical University		<table border="1"> <tr> <td>ISBN</td> <td>ISSN</td> </tr> <tr> <td colspan="2">Copies</td> </tr> <tr> <td colspan="2">No.</td> </tr> <tr> <td colspan="2">Date</td> </tr> <tr> <td colspan="2">.....-.....-.....</td> </tr> </table>		ISBN	ISSN	Copies		No.		Date		-.....-.....		Faculty of Fundamental Sciences		Department of Information Technologies	
Vilnius Gediminas Technical University		<table border="1"> <tr> <td>ISBN</td> <td>ISSN</td> </tr> <tr> <td colspan="2">Copies</td> </tr> <tr> <td colspan="2">No.</td> </tr> <tr> <td colspan="2">Date</td> </tr> <tr> <td colspan="2">.....-.....-.....</td> </tr> </table>		ISBN	ISSN	Copies		No.		Date		-.....-.....							
ISBN		ISSN																	
Copies																			
No.																			
Date																			
.....-.....-.....																			
Faculty of Fundamental Sciences																			
Department of Information Technologies																			
Master Degree Studies Information Technologies study programme Master Graduation Thesis																			
Title	Identification of Facial Features in Face Templates using Artificial Intelligence Methods																		
Author	Matas Malickas																		
Academic supervisor	Pavel Stefanovič																		
<table border="1"> <tr> <td></td> <td>Thesis language: Lithuanian</td> </tr> </table>					Thesis language: Lithuanian														
	Thesis language: Lithuanian																		
Annotation																			
This master's thesis explores the identification of facial features from face templates using artificial intelligence methods. The first part of the thesis provides an overview of facial recognition technologies and face template generation methods, with particular emphasis on the construction of facial feature vectors. In the second part, several methods are proposed for classifying original image attributes based solely on face templates. Both deep neural networks and classical supervised learning algorithms were used for experimentation. The experiments revealed that gender is the most accurately predictable attribute, with the best model achieving 93 % accuracy. Classification of ethnicity and age from face templates provide to be more challenging, likely due to the design of face template generation models, which aim to eliminate information unrelated to identity. A comparison of difference machine learning methods showed that deep neural networks are more suitable for this task than classical classification algorithms.																			
Keywords: face feature vectors, face templates, face recognition, face feature classification, neural networks, supervised learning																			

Turinys

Įvadas	10
1. Veidų atpažinimas.....	13
1.1. Veidų atpažinimo modeliai.....	13
1.1.1. „DeepFace“ veidų atpažinimo modelis.....	13
1.1.2. „FaceNet“ veidų atpažinimo modelis	14
1.1.3. „ArcFace“ veidų atpažinimo modelis	15
1.1.4. „CosFace“ veidų atpažinimo modelis.....	16
1.1.5. „SphereFace“ veidų atpažinimo modelis	17
1.2. Veidų savybių nustatymas	18
1.2.1. Veido atributų nustatymas naudojant giliuosius neuroninius tinklus	18
1.2.2. „SwinFace“ vaizdų transformatorius veidų atpažinime.....	19
1.3. Veidų šablonuose saugoma informacija ir jos rekonstrukcijos galimybės	20
1.3.1. Informacija, saugoma veido tapatybės vektoriuose	20
1.3.2. Veido rekonstrukcija naudojant dekoderius.....	21
1.3.3. Veido atributų rekonstrukcija.....	22
1.4. Skyriaus apibendrinimas.....	23
2. Eksperimentinio tyrimo metodologija	24
2.1. Duomenų aibė aprašymas	24
2.2. Duomenų paruošimas	27
2.3. Tyrimui naudojami algoritmai ir vertinimas.....	27
2.4. Eksperimentinio tyrimo aplinka ir technologijos	29
3. Eksperimentinio tyrimo rezultatai.....	31
3.1. Tyrimo rezultatai naudojant DL metodus.....	31
3.2. Tyrimo rezultatai taikant klasikinius ML metodus.....	48
3.3. Rezultatų palyginimas	51
Išvados	54
Literatūros sąrašas.....	55

Paveikslų sąrašas

1 pav. Lyties atributo pasiskirstymas	24
2 pav. Etninės grupės klasių įrašų skaičius	25
3 pav. Amžiaus klasių pasiskirstymas.....	25
4 pav. Lyties ir etninės grupės pasiskirstymas	26
5 pav. Vidutinis amžius pagal lytį ir etninę grupę	26
6 pav. „TemplateSexFC1Net“ modelio architektūra	33
7 pav. „TemplateSexFC2Net“ tinklo architektūra	33
8 pav. „TemplateSexFC3Net“ tinklo architektūra	34
9 pav. Lyties spėjimo sumaišties matrica.....	36
10 pav. „TemplateEthnicityFC1Net“ tinklo architektūra	37
11 pav. „TemplateEthnicityFC2Net“ tinklo architektūra	37
12 pav. „TemplateEthnicityFC3Net“ tinklo architektūra	37
13 pav. Baltaodžių etninės grupės spėjimo sumaišties matrica	39
14 pav. Juodaodžių etninės grupės spėjimo sumaišties matrica	40
15 pav. Azijiečių etninės grupės spėjimo sumaišties matrica.....	40
16 pav. Ispanų kilmės etninės grupės spėjimo sumaišties matrica	41
17 pav. Indų etninės grupės spėjimo sumaišties matrica	41
18 pav. „TemplateAgeFC1Net“ tinklo architektūra	42
19 pav. „TemplateAgeFC2Net“ tinklo architektūra	42
20 pav. „TemplateAgeFC3Net“ tinklo architektūra	43
21 pav. Pirmos amžiaus grupės spėjimo sumaišties matrica	45
22 pav. Antros amžiaus grupės spėjimo sumaišties matrica.....	45
23 pav. Trečios amžiaus grupės spėjimo sumaišties matrica.....	46
24 pav. Ketvirtos amžiaus grupės spėjimo sumaišties matrica.....	47
25 pav. Penktos amžiaus grupės spėjimo sumaišties matrica	47
26 pav. Šeštosios amžiaus grupės spėjimo sumaišties matrica.....	48

Lentelių sąrašas

1 lentelė. Lyties atpažinimo neuroninių tinklų modelių rezultatai.....	35
2 lentelė. Etninės grupės atpažinimo neuroninių tinklų modelių rezultatai	38
3 lentelė. Amžiaus atpažinimo neuroninių tinklų modelių rezultatai	44
4 lentelė. Lyties atpažinimo atsitiktinių miškų modelio hiperparametrų paieška.....	49
5 lentelė. Lyties atpažinimo sprendimų medžio modelio hiperparametrų paieška	50
6 lentelė. Etninės grupės atsitiktinių miškų modelio hiperparametrų paieška	50
7 lentelė. Etninės grupės sprendimų medžio modelio hiperparametrų paieška	50
8 lentelė. Amžiaus atsitiktinių miškų modelio hiperparametrų paieška	51
9 lentelė. Amžiaus sprendimų medžio modelio hiperparametrų paieška.....	51
10 lentelė. Lyties klasifikavimo modelių rezultatų palyginimas	52
11 lentelė. Etninės grupės klasifikavimo modelių rezultatų palyginimas.....	52
12 lentelė. Amžiaus grupių klasifikavimo modelių rezultatų palyginimas.....	53

Santrumpos

BCE (angl. *binary cross entropy*) – dvejetainė krosentropija, dažniausiai naudojamas terminas kaip nuostolio funkcijos apibrėžimas.

BN (angl. *batch normalization*) – imties normalizacija.

BS (angl. *batch size*) – imties dydis.

CM (angl. *confusion matrix*) – sumaišties matrica.

CNN (angl. *convolutional neural network*) – konvoliucinis neuroninis tinklas.

DL (angl. *deep learning*) – gilusis mokymasis.

DNN (angl. *deep neural network*) – gilusis neuroninis tinklas.

FC (angl. *fully connected*) – pilnai sujungtas, dažniausiai naudojamas apibūdinant giliuosius neuroninius tinklus.

L (angl. *loss*) – praradimo reikšmė, naudojama giliųjų neuroninių tinklų mokyme, nurodo modelio mokymosi efektyvumą.

LR (angl. *learning rate*) – mokymosi greitis.

ML (angl. *machine learning*) – mašininis mokymas.

MLC (angl. *multi-label classification*) – daugiaklasė klasifikacija.

MLCA (angl. *multi-level channel attention*) – daugiapakopis kanalų dėmesio mechanizmas.

MLP (angl. *multi-layer perceptron*) – daugiasluoksnis perceptronas.

IVADAS

Pastaraisiais dešimtmečiais sparčiai augant dirbtinio intelekto ir duomenų apdorojimo technologijoms, kompiuterinė rega tapo viena iš svarbiausių tyrimų ir taikymo sričių informacinių technologijų srityje. Vaizdų atpažinimo sistemos yra plačiai naudojamos tiek akademinėje, tiek pramoninėje aplinkoje – nuo medicininės diagnostikos ir autonominių transporto priemonių iki saugumo sistemų, pramoninės gamybos, žemės ūkio, socialinės medijos analizės ir net menų (Steels & De Mantaras, 2018). Šių sistemų tikslas – automatiškai interpretuoti ir klasifikuoti vaizdų turinį, dažniausiai remiantis giliojo mokymosi (angl. *deep learning*, toliau – DL) metodais.

Gilusis mokymas, ypač konvoliuciniai neuroniniai tinklai (angl. *Convolutional Neural Networks*, toliau – CNN), sukėlė revoliuciją vaizdų analizės srityje (Lecun ir kt., 2015). CNN leidžia kompiuteriams išmokti atpažinti vizualius bruožus neapdorotuose duomenyse. Tokių modelių mokymo metu sistema išmoksta atpažinti, kuri informacija yra reikšminga konkrečiai užduočiai spręsti. Tai pašalina poreikį rankiniu būdu kurti aprašomuosius požymius. Toks metodų universalumas paskatino jų taikymą sprendžiant įvairias problemas kompiuterinės regos srityse, tokiose kaip objektų aptikimas, klasifikavimas, segmentavimas bei veidų analizė.

Vienas reikšmingiausių kompiuterinės regos taikymo atvejų – veido atpažinimas ir su juo susijusių veido savybių analizė. Veido atpažinimas yra procesas, kurio metu sistema identifikuoja ar autentifikuoja asmenį pagal jo veido bruožus. Tokios sistemos plačiai taikomos saugumo, prieigos kontrolės, biometrinės autentifikacijos ir teisėsaugos srityse. Veido atributų atpažinimas yra papildomų semantinių savybių, tokių kaip lytis, amžius, emocija, etninė grupė, šypsena, nustatymas (Ranjan ir kt., 2016). Tokie požymiai gali būti naudingi geresniam tam tikros situacijos konteksto supratimui, vartotojo analizės tikslais arba įvairių paslaugų personalizavimui.

Dažniausiai veido atributai yra nustatomi tiesiogiai iš vaizdų naudojant konvoliucinius neuroninius tinklus, kurie mokomi klasifikuoti tam tikrą savybę remiantis pilnu veido vaizdu. Tačiau biometrinėse sistemose yra įprasta veidą paversti į žemesnės dimensijos veido šabloną (angl. *face template*), kuris yra fiksuoto ilgio vektorinė reprezentacija – taip vadinamas veido savybių vektorius (angl. *embedding*, *feature vector*), kuris yra sugeneruotas neuroninio tinklo. Šie šablonai įprastai yra naudojami tapatybei nustatyti. Tapatybės nustatymas vyksta lyginant veidų savybių vektorius, todėl yra siekiama, kad jie būtų nešališki ir neturintys perteklinės informacijos apie asmenį, išskyrus jo tapatybę.

Naujausi tyrimai rodo, kad veidų savybių vektoriai gali talpinti papildomą informaciją, tokią kaip lytis ar rasė – tai reiškia, kad veido savybės gali būti išskiriamos iš veidų šablonų, o tai gali kelti saugumo ir

konfidencialumo klausimus šioje srityje (Kotwal & Marcel, 2025). Šiame darbe analizuojama, kiek veido savybių informacijos išlieka šiuose šablonuose, ir ar galima šią informaciją išgauti pasitelkus specialiai apmokytus giluminio mokymosi modelius.

Šiame darbe siekiama ištirti, ar įmanoma klasifikuoti žmogaus veido atributus, tokius kaip lytį ir etninę grupę, tik iš veido šablono, neturint originalaus veido vaizdo. Tokie tyrimai ne tik atskleidžia šablonų savybes, bet ir padeda įvertinti kokią informaciją galima atkurti iš tokių šablonų neturint pradinių duomenų.

Šiam tikslui pasiekti bus sukurti keli giluminio mokymo modeliai, priimančys šabloną kaip įvestį ir grąžinantys veido atributus – šiuo atveju, lytį arba etninę grupę. Toliau bus taikomi keli neuroninių tinklų architektūrų variantai, atliekama jų parametrų optimizacija bei lyginami jų rezultatai su kitais žinomais metodais.

Tyrimo objektas – veido savybių atpažinimas.

Darbo tikslas – ištirti mašininio mokymo algoritmų efektyvumą atpažįstant veido savybes naudojant išgautus savybių vektorius.

Darbo uždaviniai:

1. Išanalizuoti veido šablonų išgavimo metodus taikomus panašioms uždaviniams spręsti.
2. Iš anksto apdoroti ir paruošti duomenų aibę eksperimentiniam tyrimui.
3. Eksperimentiškai ištirti skirtingų algoritmų efektyvumą atpažįstant vaizdo savybes iš veidų savybių vektorių.

Temos naujumas – veido savybių nustatymas iš veidų savybių vektorių, skirtu veidų atpažinimui, dar nėra plačiai tyrinėta tema. Dauguma egzistuojančių metodų remiasi tiesioginiu vaizdų apdorojimu, todėl šio darbo požiūris – naudoti tik šablonus – siūlo naują tyrimų kryptį, galinčią padėti mažinti duomenų saugojimo reikalavimus.

Temos aktualumas – šiuolaikinėse biometrinėse sistemose plačiai naudojami veido šablonai, kurie leidžia efektyviai ir patikimai atpažinti asmenis. Tokie šablonai yra kompaktiški, lengvai apdorojami bei dažnai naudojami realiu laiku veikiančiose sistemose. Šiame darbe svarbu ištirti, kiek papildomos informacijos šie šablonai išsaugo apie asmenį, tam, kad nustatyti jų platesnes jų panaudojimo galimybes. Jei savybinė informacija gali būti patikimai nustatyta remiantis tik šablonais, tai gali leisti sumažinti priklausomybę nuo pilnų veidų vaizdų, sumažinti reikalavimus duomenų saugojimui bei padidinti duomenų apdorojimo greiti ir pritaikomumą įvairiose paslaugose.

Darbo santrauka – baigiamojo magistrinio darbo pirmame skyriuje nagrinėjami fundamentiniai veidų atpažinimo metodai, ypatingą dėmesį skiriant veidų šablonų formavimui. Taip pat nagrinėjamos šių šablonų rekonstrukcijos ir veido savybių išgavimo galimybės. Antrajame skyriuje pristatoma duomenų aibė,

atliekama duomenų analizė bei paruošimams eksperimentams. Trečiajame skyriuje pateikiama eksperimentų eiga – sukurtos giliųjų neuroninių tinklų architektūros, taikyti klasikiniai prižiūrimojo mokymo metodai, aprašomi hiperparametrai, su kuriais buvo atlikti eksperimentai. Šiame skyriuje taip pat pateikiami visų tiriamų savybių – lyties, etninės grupės bei amžiaus – klasifikavimo eksperimentų rezultatai.

1. VEIDŲ ATPAŽINIMAS

Veidų atpažinimas yra viena plačiausiai taikomų kompiuterinės regos sričių. Ši technologija leidžia identifikuoti arba autentifikuoti asmenis iš skaitmeninių vaizdų, remiantis unikaliomis jų veido savybėmis. Veidų atpažinimo sistemos yra sudarytos iš kelių pagrindinių žingsnių: veido aptikimas, veido kanoninės pozicijos atstatymas (Meher ir kt., 2023), veido savybių vektoriaus išgavimas bei tapatybės nustatymas. Šiuolaikiniai sprendimai dažniausiai naudoja giliuosius neuroninius tinklus, gebančius efektyviai išmokti išgauti abstrakčius veido bruožus. Toliau darbe bus apžvelgiami keli veidų atpažinimo modeliai, jų pagrindiniai ypatumai, veidų savybių vektorių svarba bei veidų šablonų generavimo metodai.

1.1. Veidų atpažinimo modeliai

1.1.1. „DeepFace“ veidų atpažinimo modelis

Vienas iš didžiausių proveržių veidų atpažinimo srityje buvo pasiektas su „DeepFace“ modeliu (Taigman ir kt., 2014). Šis modelis dažnai įvardijamas kaip vienas iš pirmųjų giliojo mokymosi metodų, parodžiusių artimą žmogaus tikslumui veidų atpažinimo lygį. Jis yra laikomas pamatiniu darbu, kuris padėjo pagrindus šiuolaikinių veidų atpažinimo sistemų architektūrai ir metodologijai. Kadangi šio modelio principai taikomi ir dabartinėse veidų atpažinimo sistemose (M. Li, 2025), juo galima remtis ir tiriant įvairias sistemos dalis ir metodikas, tokias kaip veidų šablonų reprezentacijas.

„DeepFace“ pasiūlė sistema, kuri integruoja kelis svarbiausius komponentus: veido geometrijos normalizavimą (angl. *face alignment*), septynių sluoksnių neuroninį tinklą, bei dvimačių vektorių erdvę, kurioje atliekamas vaizdų palyginimas. Sistemos išskirtinumas buvo tai, kad prieš perduodant vaizdą neuroniniam tinklui, veidas buvo išlyginamas pagal 3 dimensijų modelį, taip užtikrinant, kad tinklas matytų veidus kuo panašesnėje „kanonijėje“ padėtyje. Tai ženkliai sumažino variacijas tarp nuotraukų, susijusias su galvos pasukimu, apšvietimu ir kitais trikdančiais veiksniais, o tai savo ruožtu pagerino klasifikavimo ir kartu veidų atpažinimo tikslumą.

Šis modelis buvo apmokymas su labai dideliu duomenų rinkiniu – daugiau nei 4 milijonais veidų vaizdų, priklausančių daugiau nei 4 000 unikalių žmonių. Tokia duomenų aibė leido tinklui suformuoti stabilias ir generalizuojamas veido reprezentacijas. Modelio klasifikavimo tikslumas „Labeled Faces in the Wild“ duomenų rinkinyje pasiekė 97,35 % tikslumą, kas tuo metu buvo tik 1 % žemiau nei žmogaus atpažinimo gebėjimas. Šis rezultatas buvo laikomas esminiu lūžio tašku – nuo šio laiko giliaji neuroniniai tinklai tapo pagrindiniu metodu veidų atpažinimo srityje.

Svarbus aspektas, susijęs su šiuo darbu, yra būdas, kaip „DeepFace“ saugoja veidus – tai daroma naudojant „embedding“ vektorius, kuriuose yra užkoduojama latentinė informacija apie asmenį. Nors šie

vektoriai buvo sukurti veidų atpažinimo tikslais, jų gebėjimas reprezentuoti unikalius bruožus leidžia svarstyti, ar jie taip pat saugoja papildomas veido savybes, tokias kaip lytis, amžius ar etninė grupė. „DeepFace“ tyrime tiesiogiai nebuvo tiriamas savybių informacijos užsilikimas savybių vektoriuose, tačiau žinant, kad šie vektoriai yra išgauti iš žmonių veidų, galima atlikti tyrimą ir išsiaiškinti ar tokiuose vektoriuose išlieka pašalinės savybės, neaktualios tapatybės nustatyme.

Apibendrinant, „DeepFace“ yra vienas iš pirmųjų modelių, sėkmingai apjungusių kanoninės veido pozicijos atstatymą, giliuosius neuroninius tinklus bei vektorines veido reprezentacijas veido identifikacijai. Nors jo pagrindinis tikslas buvo asmens tapatybės nustatymas, jo tarpinis rezultatas, veidų savybių vektoriai, leidžia šį darbą vertinti kaip svarbų pagrindą tolesniems tyrimams, nagrinėjantiems atributų išgavimą iš veidų šablonų.

1.1.2. „FaceNet“ veidų atpažinimo modelis

Kitas reikšmingas darbas veidų atpažinimo srityje, žymintis perėjimą nuo klasifikacinių sistemų prie tiesioginio veido savybių vektoriaus išgavimo yra „FaceNet“ (Schroff ir kt., 2015). Šio metodo esmė – vietoj klasifikavimo į konkrečias tapatybių klases, tinklas yra apmokomas iš karto projektuoti veidus į euklidinę erdvę taip, kad atstumai tarp jų atspindėtų veidų panašumą. Ši vektorinė reprezentacija leidžia lengvai atlikti tiek veido identifikaciją, tiek autentifikaciją, remiantis atstumo skaičiavimų tarp vektorių. Dėl šių priežasčių, šis algoritmas yra taikomas ir šiuolaikinėse veidų atpažinimo sistemose (S & S, 2021).

Skirtingai nei „DeepFace“, kuris rėmėsi klasifikaciniu požiūriu, „FaceNet“ naudoja trinarę praradimo funkciją (angl. *triplet loss*), kuri priverčia tinklą išmokti projekcijas, kuriose tos pačios tapatybės vaizdai yra arti vienas kito, o skirtingų tapatybių – tolimi. Mokymas vyksta remiantis trimis pavyzdžiais: inkaru (angl. *anchor*), teigiamu ir neigiamu. Tinklas yra mokomas taip, kad atstumas tarp inkarinio ir teigiamo pavyzdžio būtų mažesnis nei tarp inkarinio ir neigiamo pavyzdžio. Tai leidžia sistemai būti lankstesnei ir lengviau pritaikomai realiuose scenarijuose, kur vartotojų skaičius nėra fiksuotas.

Šis modelis iš pradžių naudojo „GoogLeNet“ tinklą (Szegedy ir kt., 2014), tačiau vėlesni metodai naudojo supaprastintas ar efektyvesnes architektūras. Vienas iš svarbiausių aspektų yra tas, kad modelio išėjimo sluoksnis yra tiesiogiai naudojamas kaip veido šablonas – 128 arba 512 matmenų vektorius, kurį galima naudoti lyginimui su kitais veidų vektoriais. Tokiu būdu tinklas tampa ne tik klasifikatoriumi, bet ir veidų „žemėlapiu“, kurioje veidai egzistuoja vieningoje erdvėje.

„FaceNet“ buvo apmokytas su labai dideliu duomenų kiekiu – 8 milijonais vaizdų, apimančių 2,6 milijono unikalių žmonių. Tokia duomenų įvairovė ir kiekis leido tinklui išmokti labai generalizuotą ir efektyvią reprezentaciją. Modelis pasiekė 99,63 % tikslumą „Labeled Faces in the Wild“ duomenų

rinkinyje, ženkliai viršydamas tuo metu buvusių rezultatų. Be to, modelis buvo testuotas ir „YouTube Faces“ duomenų rinkinyje, kuriame taip pat pasiekė įspūdingus rezultatus – 95,12 % tikslumą.

Šiame darbe „FaceNet“ yra ypač aktualus dėl jo naudojamo veido reprezentacijos principo – galutinis modelio sluoksnio vektorius yra tarsi šablonas, tinkamas ne tik tapatybės patikrai, bet ir, teoriškai, veido atributų nustatymui. Nors pats darbas neanalizavo, ar šiuose šablonuose galima identifikuoti požymius apie žmogų, vėlesni tyrimai parodė, kad tokios reprezentacijos dažnai užkoduoja latentinę informaciją apie lytį, amžių ar net emocijas, net jei tai nėra tiesioginis modelio tikslas.

Didelis privalumas atvaizduojant veidą, kaip vektorių, euklidinėje erdvėje yra tai, jog tapatybės lyginimas yra atliekamas skaičiuojant atstumus tarp vektorių. Tai yra ganėtinai paprasta matematine operacija, kas leidžia deklaratyviai aprašyti veido lyginimo procesą biometrinėse identifikavimo sistemose. Kadangi lyginimui nėra naudojamas joks komplikotas neuroninis tinklas ar kitas metodas, reikalaujantis didelių skaičiavimo pajėgumų, veidų identifikacija bei verifikacija naudojant tokius vektorius tampa labai greita.

Šis modelis ne tik ženkliai pagerino veidų atpažinimo tikslumą, bet ir įtvirtino naują paradigmą – asmens veido bruožus užkoduoti į tam tikrą šabloną. Jį galima naudoti kaip duomenis apie žmogaus tapatybę, kuriame potencialiai užkoduoti ir papildomi atributai.

1.1.3. „ArcFace“ veidų atpažinimo modelis

Sekantis modelis, pagerinęs veidų vektorines reprezentacijas yra „ArcFace“. Tai yra vienas pažangiausių metodų, skirtų veidų atpažinimo uždaviniui spręsti, kuris reikšmingai prisidėjo prie veidų vektorių kokybės šuolio (Deng ir kt., 2019). Skirtingai nuo ankstesnių metodų, kurie daugiausiai orientavosi į vektorių sugeneravimą per euklidinę erdvę (pvz., „FaceNet“), „ArcFace“ įveda naują principą – adityvinį kampinį nuostolį (angl. *additive angular margin loss*). Tai reiškia, kad modelis yra mokomas klasifikuoti veidus taip, jog klasifikaciniai sprendimai būtų atskirti ne tik pagal atstumą, bet ir pagal kampinį atstumą tarp jų vektorių, taip pagerinant ryškumą tarp klasių ribų.

Tradiciniai „softmax“ pagrįstos nuostolio funkcijos turi tam tikrų trūkumų – jos ne visada užtikrina pakankamai didelį atskyrimą tarp klasių vektorinėse reprezentacijose, ypač kai veidai yra vizualiai panašūs. „ArcFace“ sprendžia šią problemą transformuodamas tradicinę „softmax“ formulę į kampinę erdvę, kurioje įterpiamas papildomas tarpas (angl. *margin*) tarp skirtingų klasių. Tokiu būdu tinklas yra priverčiamas „stumti“ skirtingų žmonių vektorius toliau vienas nuo kito kampinėje erdvėje. Šis principas ypač naudingas veidų šablonų atskyrimui – skirtingų asmenų tapatybių vektorinės reprezentacijos tampa geriau atskiriamos.

Šis modelis buvo realizuotas naudojant „ResNet-100“ architektūrą, kuri pasižymi didele išraiškos galia, bet taip pat gali pritaikyti supaprastintas versijas mažesnėse ar realaus laiko sistemose (He ir kt.,

2015). Modelio išėjimo sluoksnius – paprastai 512 matmenų vektorius – yra naudojamas kaip šablonas asmens tapatybės reprezentacijai. Ši reprezentacija dažnai naudojama lyginimui su kitais vektoriais, pagrįstu kosinuso atstumu ar euklidiniu atstumu.

„ArcFace“ buvo testuotas su įvairiais žinomais duomenų rinkiniais: „Labeled Faces in the Wild“, „MegaFace“, „YouTube Faces“ ir kt. Su „Labeled Faces in the Wild“ duomenų rinkiniu buvo pasiektas 99,83 % tikslumas – tai vienas geriausių iki tol užfiksuotų rezultatų. Svarbiausia, kad šis modelis parodė aukštesnį tikslumą ir stabilumą didelėse apimties duomenyse, kai reikia identifikuoti veidus tarp milijonų kitų – tai ypač aktualu praktikoje, pavyzdžiui, saugumo ar biometrinėse sistemose.

Šis būdas yra svarbus ne tik kaip metodas, kuriuo galima generuoti kokybiškus šablonus veidų lyginimui, bet ir kaip pavyzdys, kad net ir optimizuoto, tik tapatybės atpažinimui, modelio sugeneruoti vektoriai gali talpinti daug daugiau informacijos – įskaitant lytį, etninę grupę.

„ArcFace“ yra vienas moderniausių ir sėkmingiausių veido šablonų generavimo metodų (Xin ir kt., 2024). Dėl naudojamo kampinio praradimo metodo jo generuojami vektoriai pasižymi itin dideliu skiriamuoju gebėjimu, bet kartu gali turėti latentinės informacijos apie veido atributus.

1.1.4. „CosFace“ veidų atpažinimo modelis

Kitas kitų svarbių veido šablonų generavimo pokyčių per pastarąjį dešimtmetį buvo alternatyvių nuostolių funkcijų (angl. *loss functions*) atsiradimas. Tai leido ženkliai pagerinti neuroninių tinklų gebėjimą atskirti skirtingų asmenų latentines reprezentacijas. Vienas iš reikšmingiausių darbų šioje srityje – „CosFace“ metodas (H. Wang ir kt., 2018). Šiame metode pristatomas didelio nuokrypio kosinusinio nuostolio (angl. *large margin cosine loss*) principas, skirtas pagerinti klasifikacinių ribų aiškumą veido atpažinimo užduotyse. Šis metodas laikomas natūraliu „ArcFace“ pirmtaku ir yra plačiai taikomas šiuolaikinėse veidų atpažinimo sistemose.

Šio metodo idėja paremta hipoteze, kad tradicinės „softmax“ funkcijos neužtikrina pakankamo tarpklasinio atskyrimo „embedding“ erdvėje. Vietoje to, autoriai pasiūlė modifikuotą nuostolio formulę, kurioje įvedamas papildomas kampinis nuokrypis (angl. *cosine margin*) tarp teisingos klasės reprezentacijos ir kitų klasių. Šiame metode kiekvienas vektorius yra normalizuojamas iki vieneto ilgio, todėl jų tarpusavio panašumas vertinamas grynai kampiniu principu, o ne euklidiniu atstumu. Tai užtikrina, kad identiški asmenys „embedding“ erdvėje būtų žymiai arčiau vienas kito nei skirtingi, net ir esant labai mažiems duomenų skirtumams ar įvairiems trikdžiams, tokiems kaip apšvietimas, poza, vaizdo rezoliucija ir panašiai.

Mokymų metu buvo pasitelkti keli didelio masto veidų duomenų rinkiniai, tokie kaip „MS-Celeb-1M“ ir „MegaFace“. Architektūros pagrindui buvo naudota „ResNet-50“ tinklas. Rezultatai parodė reikšmingą pažangą – „CosFace“ pasiekė 97,73 % tikslumą su „Labeled Faces in the Wild“ duomenų

rinkiniu ir viršijo 98 % „top 1“ identifikacijos tikslumą „MegaFace“ iššūkyje, naudodamas tik verifikacijos arba 1:1 palyginimo režimą. Tokie rezultatai ženkliai lenkė ankstesnius metodus, tuo pat metu išlaikant sąlyginai paprastą architektūrą.

Svarbu pabrėžti, kad „CosFace“ ne tik pagerino identifikacinį tikslumą, bet ir turėjo netiesioginės įtakos vėlesnių veido šablonų generavimo metodų raidai. Kampinis nuokrypis, kaip papildomas komponentas „softmax“ funkcijoje, tapo standartu, kuris buvo toliau plėtojamas kituose darbuose (M. Wang ir kt., 2024). Šis metodas taip pat prisideda prie supratimo apie tai, kaip formuojami veido šablonai arba veidų savybių vektoriai, kurie bus nagrinėjami šiame darbe.

1.1.5. „SphereFace“ veidų atpažinimo modelis

Vienas iš ankstyvųjų, tačiau reikšmingų žingsnių veido atpažinimo reprezentacijų tobulinime buvo „SphereFace“ metodas pasiūlytas 2017 metais (W. Liu ir kt., 2017). Šis metodas iš esmės pakeitė veido savybių šablonų mokymo paradigmą, įvedant geometrinį aspektą į vektorialinę erdvę, kurioje formuojamas šios reprezentacijos. „SphereFace“ pristato kampinę „softmax“ (angl. *angular softmax*) funkciją, kuri yra modifikacija tradicinei „softmax“ funkcijai. Ši funkcija padeda mokymo procese didinti tarpklasines vektorių atskirtis ne tik amplitudės, bet ir kampo atžvilgiu. Tai leidžia projekcijas veidų reprezentacijų erdvėje organizuoti ant hipersferos paviršiaus, o tai, kaip rodo autoriai, padeda geriau atskirti veidus reprezentacinėje erdvėje.

Šis metodas pasižymi tuo, kad apriboja reprezentacijas vienos sferos paviršiuje, tokiu būdu pašalindamas priklausomybę nuo vektoriaus ilgio. Vietoj to, tapatybė koduojama tik per kampinį atstumą tarp vektorių, kas leidžia sistemai būti atsparesnei šviesos, pozos bei kitiems pašaliniais vaizdo veiksniams. Autoriai empiriškai parodė, kad toks metodas žymiai pagerina modelio gebėjimą generalizuoti bei atpažinti skirtingus veidus, kurie nebuvo matyti mokymo metu, o tai yra itin svarbu realioms biometrinėms identifikavimo sistemoms.

Šis modelis buvo mokomas naudojant „CASIA-WebFace“ duomenų rinkinį, o vertinimas atliktas pasitelkiant „Labeled Faces in the Wild“, „YouTube Faces“ ir „MegaFace“ duomenų rinkinius. Bandymai atskleidė, jog šis modelis gerokai lenkia klasikinį „softmax“ bei tuometinius pažangius modelius pagal tikslumą. Pavyzdžiui, „Labeled Faces in the Wild“ duomenų rinkinyje pasiektas 99,42 % tikslumas, o tai buvo reikšmingas patobulėjimas nuo kitų tuometinių modelių.

„SphereFace“ modelio įtaka yra reikšminga ne tik dėl tiesioginio pritaikymo rezultatų, bet ir dėl to, kad jis paskatino vėlesnių modelių – tokių kaip „CosFace“ ir „ArcFace“ – atsiradimą, kurie remiasi kampiniu pagrindu veikiančiomis nuostolio funkcijomis, bet įveda skirtingas normalizavimo bei kitas

strategijas. „SphereFace“ galima laikyti pirmuoju metodu, kuris matematiškai griežtai pagrindė hipersferinių „embedding“ erdvių naudą veidų atpažinimo srityje.

Šis metodas yra aktualus todėl, kad rodo, kaip galima modeliuoti veido reprezentacijas taip, kad jos taptų atskiriamos „embedding“ erdvėje net esant ribotam duomenų kiekiui. Taip pat šis metodas vis dar yra naudojamas šiuolaikinėse sistemose (Wu ir kt., 2022). Kadangi šio modelio reprezentacijos yra suformuotos remiantis kampine topologija, galima daryti prielaidą, kad jose išlieka svarbių veidų požymių, tokių kaip lytis, etninė grupė ar net veido struktūros bruožai.

Kaip ir kitų apžvelgtų modelių, „SphereFace“ paskirtis nebuvo suformuoti veidų savybių vektorius tam, kad iš jų būtų galima išskirti veido savybes. Tačiau dėl informacijos išlaikymo „embedding“ vektoriuose galima manyti, jog šie metodai iš dalies užkoduoja ne tik tapatybę, bet ir kitą semantiškai svarbią informaciją.

1.2. Veidų savybių nustatymas

Veidų savybių nustatymas yra kita svarbi kompiuterinės regos sritis, kuri leidžia automatizuoti veido atributų, tokių kaip lytis, amžius, emocija, etninė grupė, nustatymą įvairiose biometrinėse sistemose. Šie požymiai gali būti taikomi įvairiose srityse – nuo vartotojo patirties personalizavimo, iki veido nuotraukos kokybės ar veido išraiškos nustatymo, sprendžiant duomenų kokybės klausimus identifikavimo bei autentifikavimo sistemose. Tradiciniai metodai, tokie kaip geometrijos pagrindu veikiantys algoritmai, anksčiau buvo plačiai taikomi šiems bruožams nustatyti, tačiau pastaraisiais metais šioje srityje imta naudoti giliuosius neuroninius tinklus bei kitus pažangesnius metodus. Šie modeliai leidžia pasiekti žymiai geresnius rezultatus, nes sugeba atpažinti subtilius vizualius skirtumus tarp skirtingų atributų klasių.

1.2.1. Veido atributų nustatymas naudojant giliuosius neuroninius tinklus

Reikšmingas darbas, nagrinėjantis veido atributų atpažinimą iš nuotraukų naudojant giliuosius neuroninius tinklus, yra „Deep Learning Face Attributes in the Wild“ (Z. Liu ir kt., 2014). Šiame darbe pristatoma nauja giliojo mokymosi sistema, skirta veido atributų prognozavimui realiomis sąlygomis, kuriuose veidai gali būti pasukti įvairiais kampais, gali būti skirtingi šviesos lygiai ar uždengti.

Pasiūlyta sistema susideda iš dviejų konvoliucinių neuroninių tinklų: „LNet“ ir „ANet“. Pirmasis tinklas yra atsakingas už veido lokalizavimą nuotraukoje, o antrasis – už veido atributų prognozavimą. Svarbu paminėti, kad abu tinklai buvo iš anksto apmokyti skirtingais duomenimis. „LNet“ tinklas buvo apmokytas su bendrais objektais, siekiant išmokyti lokalizuoti veidus, o „ANet“ mokytas su daugybe skirtingų veidų, siekiant tinklą kuo įmanoma geriau išmokyti generalizuoti. Atributų išgavimas iš nuotraukų buvo pasiektas apjungus šiuos tinklus į vieną.

Vienas iš šio darbo inovatyvių aspektų yra tai, kad „LNet“ gali būti apmokytas lokalizuoti veidus tik pagal vaizdų klases, nereikalaudamas tikslių veido koordinacių. Tai leidžia sumažinti priklausomybę nuo rankiniu būdu sužymėtų duomenų, kurie yra brangūs ir reikalaujantys daug laiko žymėjimui.

„ANet“ tinklas yra atsakingas už veido atributų prognozavimą ir parodė, kad aukšto lygio gilieji neuroniniai tinklai automatiškai atranka semantinius atributus po išankstinio apmokymo (angl. *pre-training*) su daugybe skirtingų veidų. Modelio tikslumas gali būti dar labiau pagerintas po tolimesnio mokymo su sužymėtais atributų duomenimis.

Atliktų eksperimentų rezultatai parodė, kad siūloma sistema žymiai pralenkia ankstesnius metodus veido atributų prognozavimo tikslumu. Tai rodo, kad gilieji neuroniniai tinklai, apmokyti tinkamai, gali efektyviai išmokyti atpažinti įvairius veido atributus net ir sudėtingomis sąlygomis.

1.2.2. „SwinFace“ vaizdų transformatorius veidų atpažinime

Pastaruoju metu vaizdų transformatoriai yra vis plačiau taikomi įvairiose vaizdų atpažinimo srityse. Vienas iš panaudos atvejų yra veidų bei jų atributų atpažinimas, kaip rašoma straipsnyje „SwinFace: A Multi-task Transformer for Face Recognition, Expression Recognition, Age Estimation and Attribute Estimation“ (Qin ir kt., 2023). Šiame darbe pristatoma nauja daugiafunkcinė transformatorių architektūra, kuri leidžia atlikti kelias veido analizės užduotis vienu metu. Šis metodas yra itin aktualus šiuolaikinėse biometrinių sistemų taikymo srityse, nes jis ne tik pagerina atskirų užduočių tikslumą, bet ir padeda sukurti universalų modelį, galintį atlikti įvairias užduotis, naudojant vieną sistemą.

„SwinFace“ modelis yra pagrįstas „Swin Transformer“ architektūra, kuri yra viena iš naujausių ir pažangiausių transformatorių architektūrų, naudojama vaizdų atpažinimo užduotims (Z. Liu ir kt., 2021). Pagrindinė šios architektūros savybė yra gebėjimas efektyviai apdoroti vaizdus įvairiais masteliais ir atsižvelgti į vaizdų struktūrą. Dėl šios savybės ji yra ypač tinkama sudėtingoms veido atpažinimo ir atributų nustatymo užduotims. „SwinFace“ modelis iš esmės praplečia „Swin Transformer“ modelį, įtraukdamas kelis užduočių potinklius (angl. *subnets*), kurie skirti atlikti įvairias užduotis: veido atpažinimą, emocijų atpažinimą, amžiaus spėjimą ir atributų nustatymą.

Šiame modelyje yra naudojamas daugiapakopis kanalų dėmesio mechanizmas (angl. *multi-level channel attention*, toliau – MLCA), kuris leidžia pritaikyti modelį prie įvairių užduočių. Tai yra ypač svarbu, nes skirtingoms užduotims reikia skirtingų požymių ir apdorojimo strategijų. Pavyzdžiui, veido atpažinimas reikalauja didelio tikslumo aptinkant veidą identifikuojančius bruožus, o emocijų atpažinimui svarbūs tam tikri subtilūs veido išraiškos pokyčiai. Daugiapakopis dėmesio mechanizmas leidžia modelyje vienu metu optimizuoti visus šiuos aspektus, todėl jis geba išgauti reikšmingus požymius tiek vienai užduočiai, tiek visoms užduotims kartu.

Straipsnyje pateikti eksperimentiniai rezultatai, kurie rodo, kad šis modelis yra pranašesnis už kitus naudojamus metodus tokioms užduotims spręsti. Modelis pasiekė pažangiausius rezultatus įvairiose užduotyse. Pavyzdžiui, emocijų atpažinimo užduotyje modelis pasiekė 90,97 % tikslumą „RAF-DB“ duomenų rinkinyje, o amžiaus spėjimo užduotyje jo paklaida buvo 0,22 „CLAP2015“ duomenų rinkinyje. Šie rezultatai rodo, jog šis metodas yra itin efektyvus modelis, gebantis atlikti kelias užduotis vienu metu, ir pasiekia aukštą tikslumą, kuris yra reikšmingas realaus pasaulio sistemose.

Nepaisant pažangių rezultatų, šis modelis turi tam tikrų apribojimų. Nors MLCA leidžia efektyviai apdoroti ir nustatyti įvairius požymius, tokio tipo modeliai gali būti brangūs skaičiavimo požiūriu, ypač apdorojant didelius duomenų kiekius ar esant realaus laiko (angl. *real-time*) sistemoms.

1.3. Veidų šablonuose saugoma informacija ir jos rekonstrukcijos galimybės

Kadangi šiuolaikinės veidų atpažinimo sistemos naudoja kompaktiškus veido savybių vektorius tapatybės nustatymui, kyla klausimų, kokia informacija yra užkoduojama tuose vektoriuose ir ar ją galima atkurti. Nors dauguma veidų šablonų generavimo metodų yra kuriami taip, kad tapatybės nustatymui nereikšminga informacija būtų išmetama, tačiau moksliniai tyrimai nustatė, kad tam tikra informacija gali būti atkurama iš šių vektorinių veidų reprezentacijų.

1.3.1. Informacija, saugoma veido tapatybės vektoriuose

Augant biometrinių tapatybės nustatymų sistemų naudojimui kyla susirūpinimas asmens tapatybės saugumu bei privatumu. Siekiant nustatyti rizikas ir saugomos informacijos detales buvo atlikti tyrimai, siekiantys nustatyti kas yra saugojama asmens biometrinėse reprezentacijose. Vienas tokių tyrimų buvo „Beyond Identity: What Information Is Stored in Biometric Face Templates“ (Terhöst ir kt., 2020).

Šiame straipsnyje nagrinėjama, kokia papildoma informacija, be asmens tapatybės, gali būti išgauta iš veido šablonų (angl. *face templates*), kurie yra gauti naudojant pažangius veido atpažinimo metodus, tokius kaip „FaceNet“ ar „ArcFace“. Autoriai iškelia prielaidą, kad veido šablonai gali užkoduoti ne tik tapatybės bruožus, bet ir vadinamuosius „minkštuosius biometrinius požymius“ (angl. *soft-biometrics*), tokius kaip amžius, lytis, šukuosena, makiažas ar net nuotaika.

Šiame tyrime buvo pasitelktas didelės apimties daugiaatributis klasifikatorius (angl. *massive attribute classifier*), kad būtų nustatyta galimybė prognozuoti veido atributus tiesiogiai iš veidų savybių vektorių (angl. *embedding*). Šis klasifikatorius buvo mokomas daugiafunkciniu būdu (angl. *multi-task learning*), kad galėtų panaudoti bendrą požymių erdvę nustatyti skirtingiems atributams. Tyrimas buvo atliktas remiantis dviem žinomais duomenų rinkiniais – „CelebA“ ir „Labeled Faces in the Wild“, o

„embedding“ buvo sugeneruoti naudojant dvi populiarias veido atpažinimo architektūras – „FaceNet“ ir „ArcFace“.

Rezultatai atskleidė, kad net 74 iš 113 atributų buvo bent jau vidutiniškai nuspėjami, o 39 – lengvai prognozuojami nepastovūs atributai, tokie kaip šukuosenos, barzdos, plaukų spalvos, makiažas ar akiniai. Šis atradimas yra itin reikšmingas, nes veidų atpažinimo sistemos paprastai mokomis būti kuo įmanoma atsparesnės tokiems kintamiems požymiams, kurie nėra susiję su žmogaus tapatybe. Tai reiškia, kad net jei sistemos siekia ignoruoti šiuos aspektus, veidų savybių vektoriuose jie vis tiek išlieka ir gali būti nustatomi.

Vienas svarbiausių šio tyrimo akcentų – gebėjimas vertinti klasifikatoriaus prognozės patikimumą, o tai leidžia dar smulkiau diferencijuoti, kurie atributai iš veidų savybių vektorių gali būti patikimai išgaunami. Tokia analizė leidžia spręsti ne tik apie atributų buvimą šablonuose, bet ir apie tai, kokių tikslumu jie gali būti rekonstruojami, o tai turi tiesiogines pasekmes tiek privatumo apsaugai, tiek šališkumo mažinimui technologijose (Terhörst ir kt., 2021).

Darbo autoriai pabrėžia, jog jų klasifikatoriaus architektūra gali būti toliau taikoma norint sukurti privatumo išsaugojimo metodus, nes gebėjimas nustatyti, kokia informacija yra veidų savybių vektoriuose, leidžia kurti būdus, kaip ją šalinti arba transformuoti taip, kad būtų mažinama atitinkamų bruožų rekonstrukcijos rizika.

Šis tyrimas padeda atskleisti realią informaciją ir jos vertę, kuri yra užkoduota veidų šablonuose, ir pabrėžia poreikį kritiškai vertinti jų naudojimą biometrinėse sistemose.

1.3.2. Veido rekonstrukcija naudojant dekoderius

Kitas reikšmingas darbas, analizuojantis informacijos rekonstrukcija iš veido atpažinimo sistemose naudojamų veidų savybių vektorių, yra „Deep Face Decoder“ (Križaj ir kt., 2024). Šio darbo tikslas – ištirti, kokio pobūdžio semantinė informacija yra užkoduota veidų savybių vektoriuose, naudojamuose giliojo mokymosi pagrindi veikiančiose sistemose. Tyrimas siūlo metodą, leidžiantį rekonstruoti veido atvaizdus iš „embedding“ vektorių.

Šiame darbe buvo sukurtas specializuotas giliojo mokymo modelis, kuris buvo apmokytas atkurti pradinį veido atvaizdą ir veido savybių „embedding“, kuris yra generuojamas veidų atpažinimo modelių. Modelio architektūra yra paremta dekoderių, naudojamų „autoencoder“ sistemose (P. Li ir kt., 2023), principu, tačiau ji yra pritaikyta veido rekonstrukcijai.

Šis tinklas priima veidų savybių vektorių, sugeneruota tokiu tinklu, kaip „FaceNet“ ir jį apdoroja pasitelkiant pilnai sujungtus (angl. *fully connected*, toliau – FC) neuroninius tinklus, kurių išvestis yra transformuojama į dvimačio formato tenzorius, atitinkančius veido atvaizdus. Šie tenzoriai toliau

apdorojami per keletą konvoliucinių sluoksnių naudojant „unsampling“ mechanizmus, siekiant atkurti kuo realesnį veido atvaizdą.

Tyrimo metu buvo išbandyti įvairūs žinomo veido atpažinimo modeliai, tokie kaip „FaceNet“ ar „ArcFace“, kurių veidų savybių vektoriai buvo naudojami kaip įvestis dekodieriams. Eksperimentai buvo vykdomi su populiariais duomenų rinkiniais, tokias kaip „Labeled Faces in the Wild“ bei „VGGFace2“, taip užtikrinant pakankama duomenų įvairovę ir skaičių.

Vertinant modelio veikimą buvo analizuojamas ne tik vizualus atitikimas tarp veido rekonstrukcijos ir originalo, bet ir tai, kiek gerai atstatytame atvaizde išlieka semantinės veido savybės. Buvo atlikta atkurto atvaizdo automatinė atributų analizė, kuri atskleidė, jog rekonstrukcija turėjo aukštą semantinį tikslumą – tai reiškia, kad veidų savybių vektoriai užkoduoja šias savybes, net jei jų išsaugojimas nėra tiesioginis šių modelių tikslas.

1.3.3. Veido atributų rekonstrukcija

Veidų savybių vektoriai įprastai yra vertinami tik per tapatybės klasifikacijos arba tapatybės suliginimo (angl. *matching*) tikslumą. Tačiau „embedding“ semantinė struktūra, tai yra, kokia informacija juose užkoduojama be tapatybės, ilgą laiką nebuvo pakankamai ištirta. Straipsnyje „Understanding Deep Face Representation via Attribute Recovery“ remiamasi hipoteze, kad net jei modelis mokytas tik atpažinti tapatybę, jis nevalingai užkoduoja ir kitus semantinius požymius – lytį, amžių, emocijas ar etninę kilmę (Ren ir kt., 2024). Tai veido savybių vektorių paverčia labai jautria informacija, kadangi jis nebėra visiškai „anonimizuotas“.

Šiame darbe yra siūlomas metodas, susidedantis iš dviejų etapų – atributų dekodavimo ir kontrafaktinės analizės. Pirmame etape yra suformuojamas naujas modelis – dekodieris, kuris iš jau sugeneruotų veidų „embedding“ bando rekonstruoti įvairius veido atributus. Veidų „embedding“ yra gaunami iš populiarių veidų atpažinimo modelių, tokių kaip „ArcFace“. Dekoderiai yra mokomi naudojant su sužymėtais atributais iš tokių duomenų rinkinių, kaip „CelebA“ bei „FairFace“, siekiant nustatyti, kiek informacijos apie kiekvieną atributą išlieka veido savybių vektoriuje.

Antru šio modelio etapu, kontrafaktine analize, siekiama nustatyti, kaip kiekviena atributo buvimas veido savybių vektoriuje veikia tapatybės atpažinimo rezultatą. Šiam tikslui pasiekti naudojama „counterfactual inference“ metodika. Joje modeliu simuliuojami veidų savybių vektoriai, kuriuose tam tikras požymis yra pakeičiamas, išlaikant kitą informaciją nepakitusia. Tokie sintetiniai veidų savybių vektoriai naudojami siekiant įvertinti, ar mažos semantinės modifikacijos daro įtaką atpažinimo modelio tikslumui ar identifikacijos rezultatams.

Gauti rezultatai rodo, kad iš veidų savybių vektorių galima gana tiksliai rekonstruoti įvairius veido atributus. Pavyzdžiui, lyties atpažinimo tikslumas siekė virš 95 %, o amžiaus regresijos paklaida buvo nedidelė (4-5 metai). Tai rodo, kad šie atributai yra išsaugomi veidų savybių vektoriuose.

Kontrafaktinės analizės rezultatai rodo, kad tam tikri požymiai, pavyzdžiui amžius ar emocija, darė reikšmingą įtaką veido tapatybės nustatymui. Tai reiškia, kad tie patys asmenys gali būti skirtingai atpažinti priklausomai nuo laikinojo požymio, tokio kaip šypsenos buvimas. Ši įžvalga yra labai svarbi kuriant patikimas biometrines sistemas, kadangi tai gali neigiamai paveikti asmens atpažinimo rezultatus.

1.4. Skyriaus apibendrinimas

Atlikta literatūros analizė parodė, jog veidų atpažinimo srityje didžiausią proveržį lėmė giliojo mokymo metodai, generuojantys kompaktiškus ir informatyvius veido šablonus (veidų savybių vektorius, „embedding“), kurie sėkmingai taikomi tapatybės nustatymo užduotyse. Tuo pačiu metu intensyviai vystomi ir veido savybių atpažinimo metodai, kurie dauguma yra paremti vaizdinių duomenų analize.

Vis dėlto, galima pastebėti, kad mažai darbų nagrinėja veido savybių klasifikacija tiesiogiai iš veidų savybių vektorių. Ši tyrimų kryptis tampa vis aktualesnė, nes leidžia ne tik efektyviau naudoti turimus duomenis, bet ir kelia klausimus dėl veidų šablonų informatyvumo bei privatumo.

Toliau darbe bus tirama, kiek informacijos apie žmogaus savybes galima išgauti iš veido savybių vektoriaus, pateikiant alternatyvų būdą veido savybių nustatymui iš veido šablono.

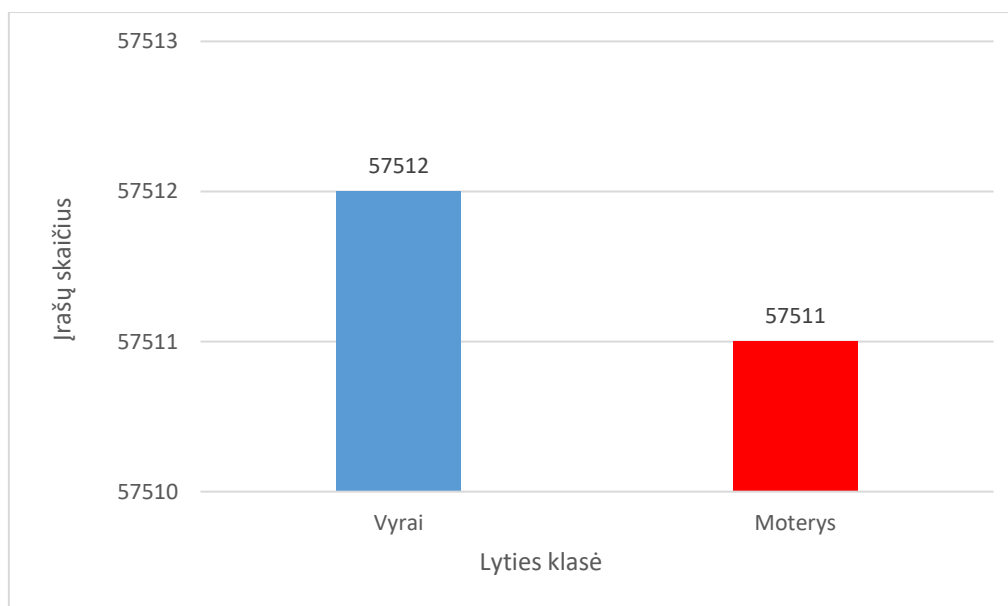
2. EKSPERIMENTINIO TYRIMO METODOLOGIJA

2.1. Duomenų aibė aprašymas

Darbe naudotas duomenų rinkinys buvo pateiktas kompanijos UAB „Neurotechnology“, jis buvo sudarytas iš veidų nuotraukų su metaduomenimis bei sužymėtais veidų atributais, tokiais kaip amžius, lytis ir rasė. Siekiant išgauti kuo įmanoma geresnius rezultatus, duomenų rinkinys buvo apdorotas ir buvo išrinkti tinkami duomenys modelių mokymui. Šiame duomenų rinkinyje buvo išsaugoti veidų savybių vektoriai, išgauti naudojant „MegaMatcher SDK“. Kiekvienas šių vektorių yra sudaryti iš 512 „float32“ tipo skaičių. Buvo atliktas duomenų rinkinio balansavimas, siekiant sudaryti kuo įmanoma pilnesnį rinkinį, turintį didelę atributų įvairovę. Duomenys buvo tolygiai paskirstyti pagal kategorijas ir klases, kurios bus mėginamos nuspėti iš veidų savybių vektorių.

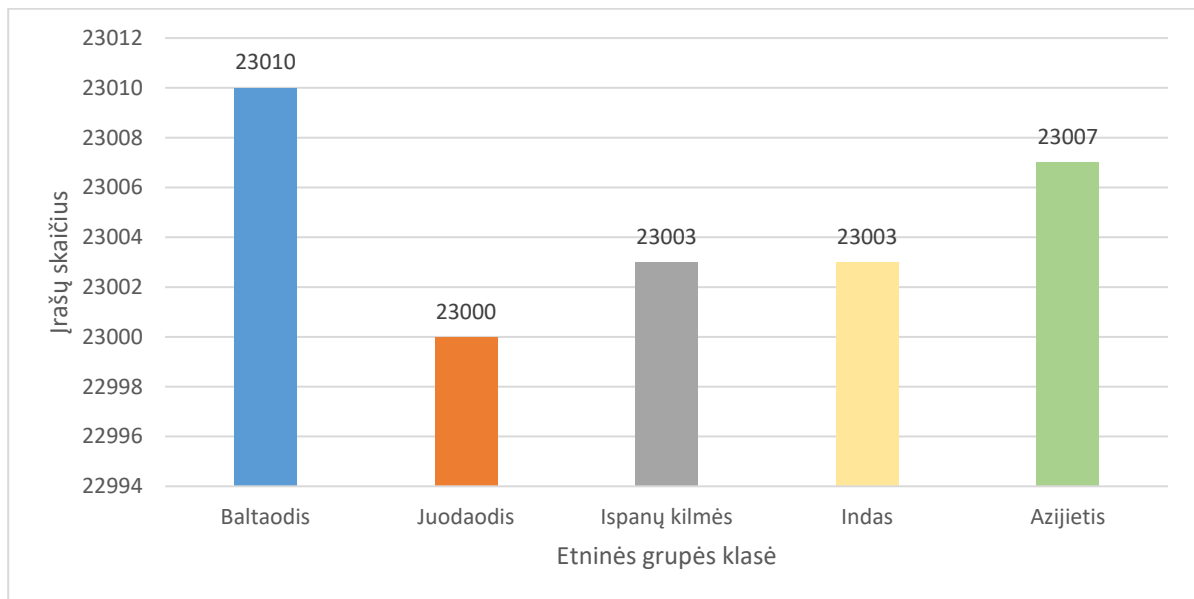
Gautas duomenų rinkinys buvo sudarytas iš 115023 įrašų, kur visi įrašai yra pilnai teisingi, neturintys nevalidžių ar tuščių duomenų. Dauguma įrašų yra anonimizuoti įrašai apie unikalius žmones, tai yra, didžioji dalis duomenų yra unikalios tapatybės.

Lytis duomenų rinkinyje buvo aprašyta kaip dvejetainis atributas. Vyrų duomenų rinkinyje buvo 57512 įrašų, o moterų 57511. Šio atributo pasiskirstymas pavaizduotas 1 pav.



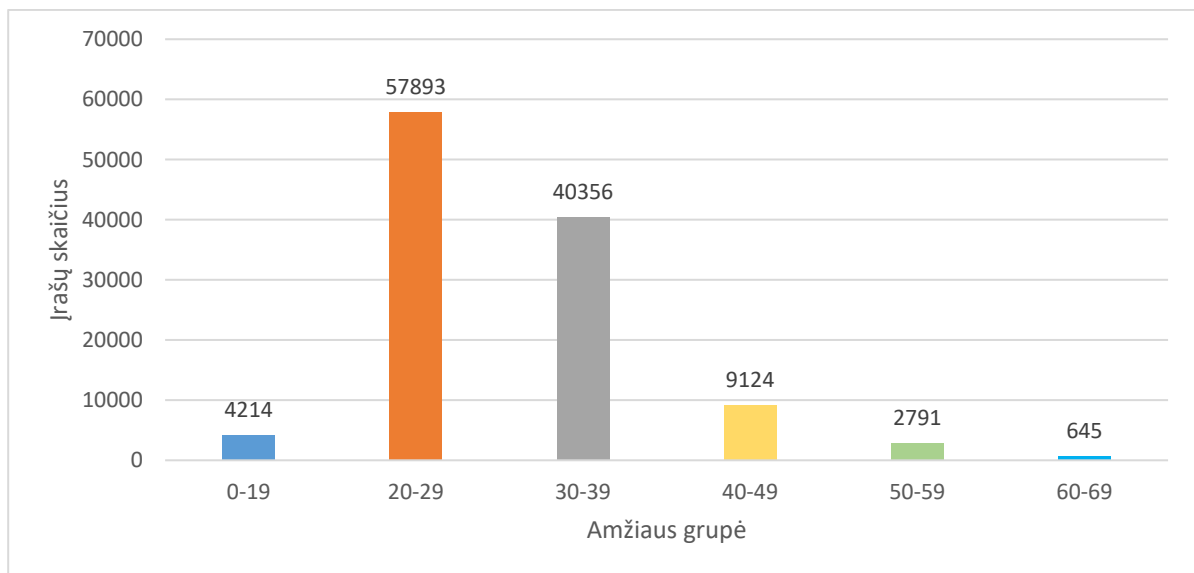
1 pav. Lyties atributo pasiskirstymas

Etninės grupės atributo galimos klasės iš viso buvo 5. Kiekvienos etninės grupės klasės įrašų kiekis yra pavaizduotas 2 pav.



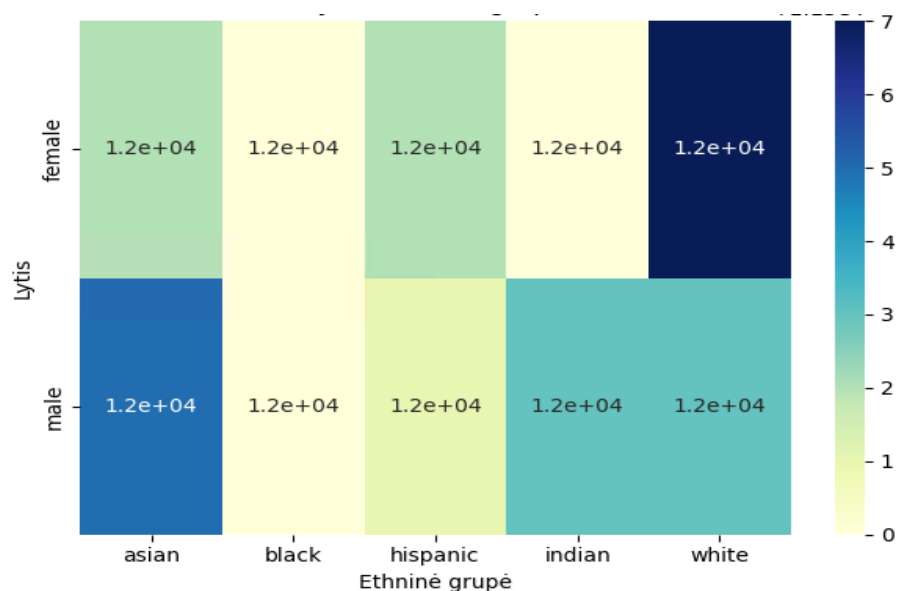
2 pav. Etninės grupės klasių įrašų skaičius

Amžiaus duomenys buvo pateikti kaip realieji skaičiai, priklausantys intervalui [18; 69]. Atlikus keletą eksperimentų, naudojant regresinius modelius, buvo pastebėta, jog regresinių modelių tikslumas yra visiškai nepatenkinamas. Dėl šios priežasties buvo nuspręsta duomenų rinkinį suskirstyti į keletą klasių, kur kiekviena klasė atspindi tam tikrą amžiaus intervalą. Šis duomenų pasiskirstymas yra pavaizduotas 3 pav.



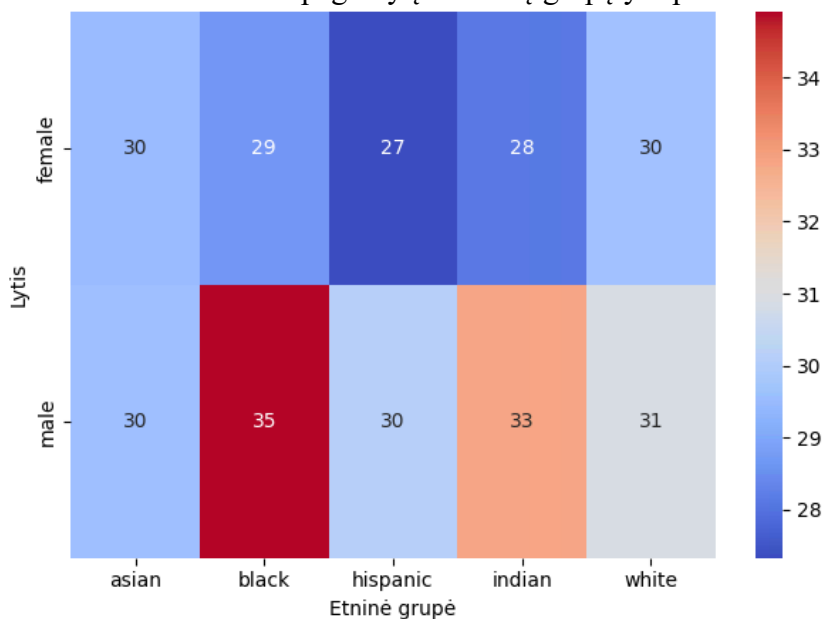
3 pav. Amžiaus klasių pasiskirstymas

Siekiant užtikrinti, kad duomenų rinkinys būtų kuo įmanoma labiau subalansuotas, duomenys buvo subalansuoti pagal klasių kombinacijas. Duomenų rinkinio lyties ir etninės grupės atributų pasiskirstymas yra pavaizduotas 4 pav.



4 pav. Lties ir etninės grupės pasiskirstymas

Duomenų rinkinio vidutinis amžius pagal lytį ir etninę grupę yra pavaizduotas 5 pav.



5 pav. Vidutinis amžius pagal lytį ir etninę grupę

Kaip galima pastebėti iš duomenų pasiskirstymo, surinktas duomenų rinkinys yra ypač gerai subalansuotas lyties bei etninės grupės atributų atžvilgiu. Amžiaus atributo atžvilgiu duomenų rinkinys yra stipriai užpildytas įrašais iš antrosios ir trečiosios klasės, t.y., dominuoja amžiaus grupės nuo 20 iki 39 metų.

Žinant šiuos duomenis, galima prognozuoti, kad nuspėti amžiaus atributo intervalą bus sudėtingesnis uždavinys dėl nesubalansuotų klasių įrašų kiekio.

2.2. Duomenų paruošimas

Pradiniai duomenys buvo laikomi „SQLite“ duomenų bazėje. Norint panaudoti tokią duomenų bazę neuroninių tinklų mokyme ar prižiūrimajame mokyme, reikia sukurti duomenų skaitymo mechanizmą.

Neuroninių tinklų mokymui duomenys buvo užkraunami ir apdorojami naudojant „PyTorch“ „Dataset“ abstrakciją. Duomenų nuskaitymo ir paruošimo procesas buvo sudarytas iš duomenų skaitymo iš releacinės duomenų bazės, grupuojant juos pagal unikalius identifikacijos numerius, kadangi veidų šablonai ir klasifikaciniai atributai buvo išdėstyti per skirtingas lenteles. Nuskaityti veidų šablonai buvo verčiami į „PyTorch“ tenzorius, kurie sudaryti iš „int8“ tipo baitų masyvų. Klasifikaciniai atributai buvo konvertuojami į vienmačius tenzorius, sudarytus iš tiek elementų, kiek yra klasių.

Prižiūrimojo mokymo algoritams duomenys buvo pateikiami kaip „Pandas“ duomenų masyvas. Duomenys buvo nuskaityti iš releacinės duomenų bazės ir jungiami pagal unikalų identifikacijos numerį. Veidų šablonai buvo konvertuojami į „numpy“ vienmačius masyvus ir dedami į bendrą „Pandas“ duomenų masyvą, kuris prieš patenkant į modelį, buvo konvertuojamas į „numpy“ dvimatį masyvą, kuriame kiekviena eilutė yra unikalus įrašas apie veidų savybių vektorių ir jo klases.

2.3. Tyrimui naudojami algoritmai ir vertinimas

Toliau buvo siekiama įvertinti, kiek informacijos apie veido savybes, tokias kaip lytį, amžių bei etninę grupę, yra užkoduojama veido atpažinimo sistemų sugeneruotuose veidų šablonuose. Tyrimo tikslui pasiekti buvo sukurti ir išanalizuoti keli klasifikavimo modeliai, kurių paskirtis – nustatyti tam tikras atributų klases iš veidų savybių vektorių. Modeliai buvo kuriami taikant tiek giliojo mokymo tinklus, tiek klasikinius prižiūrimojo mokymo algoritmus.

Neuroniniai tinklai buvo sukurti naudojant „PyTorch“ programavimo karkasą, leidžiantį lanksčiai modeliuoti įvairias architektūras. Kiekvienas tinklas priima fiksuoto ilgio (512 elementų) vektorių – veido šabloną, kuris yra išgautas naudojantis „MegaMatcher“ veidų atpažinimo algoritmu. Visi tinklai yra pilnai sujungti neuroniniai tinklai, kuriuose yra naudojami įvairūs sluoksniai, tokie kaip grupinės normalizacijos (angl. *batch normalization*), kas padeda pagreitinti modelių konvergavimą ir leidžia tinklui išmokti geriau generalizuoti. Paskutiniame sluoksnyje, priklausomai nuo klasifikavimo modelio tipo, yra naudojamos sigmoidinė arba *softmax* aktyvacijos funkcijos.

Pavyzdžiui, lyties klasifikavimui buvo sukurtas modelis „TemplateSexFC2Net“, sudarytas iš dviejų paslėptųjų sluoksnių. Šis modelis kaip išvestį grąžina vieną reikšmę, interpretuojamą kaip priklausomybę vienai ar kitai lyčiai. Etninės grupės klasifikavimui naudotas tinklas „EthnicityFC1Net“ turi penkis paslėptus sluoksnius bei išvesties sluoksnį su 5 neuronais, atitinkančiais etninių grupių klases. Amžiaus grupės nustatymo modelis buvo kuriamas analogiškai ir išvesties sluoksnyje turėjo 6 neuronus, atitinkančius amžiaus grupes.

Be giliojo mokymo metodų, papildomai buvo atlikti eksperimentai taikant klasikinius prižiūravimo mokymo algoritmus. Tokie metodai kaip „Random Forest“ bei „Decision Tree“ buvo pritaikyti tam pačiam duomenų rinkiniui. Šių metodų įtraukimas į eksperimentus buvo motyvuotas siekiant palyginti giluminio mokymo algoritmų efektyvumą su mažesnio sudėtingumo modeliais, kurie dažnai reikalauja mažiau treniravimo duomenų ir resursų.

Modelių našumo įvertinimui buvo taikytos klasikinės klasifikacijos metrikos, tokios kaip tikslumas (angl. *accuracy*), preciziškumas (angl. *precision*) bei jautrumas (angl. *recall*) bei F1 balas (angl. *F1 score*). Siekiant pagerinti klasikinių modelių rezultatus, buvo atlikta hiperparametrų paieška naudojant „GridSearchCV“ metodą su kryžmine validacija (angl. *cross-validation*) lygia 5, taikant atskiras parametrų aibes kiekvienam algoritmui.

Tikslumas – tai viena paprasčiausių ir dažniausiai naudojamų metrikų, ypač tais atvejais, kai klasės yra pasiskirsčiusios tolygiai. Ši metrika apibrėžia teisingai klasifikuotų pavyzdžių dalį visų pavyzdžių atžvilgiu:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Šioje formulėje:

- TP (angl. *true positives*) – teisingai klasifikuoti teigiami atvejai;
- TN (angl. *true negatives*) – teisingai klasifikuoti neigiami atvejai;
- FP (angl. *false positives*) – klaidingai nuspėti teigiami atvejai;
- FN (angl. *false negatives*) – klaidingai nuspėti neigiami atvejai.

Ši formulė yra taikoma tiek dvejetainiams, tiek daugiaklasiams klasifikavimo atvejams, kai skaičiavimai yra nustatomi apibendrintai visoms klasėms.

Preciziškumas rodo, kokia dalis modelio pateiktų teigiamų spėjimų buvo teisingi. Tai itin svarbi metrika, kai klaidingi teigiami rezultatai turi didelį neigiamą poveikį.

$$Precision = \frac{TP}{TP + FP}$$

Didelis preciziškumas rodo, kad modelis retai „perdeda“ ir žymi tik tuos atvejus, kuriuose yra pakankamai užtikrintas.

Jautrumo (angl. *recall*) metrika rodo, kokią dalį visų realiai teigiamų atvejų modelis sugebėjo identifikuoti:

$$Recall = \frac{TP}{TP + FN}$$

Ši metrika yra svarbi, kai norima aptikti kuo daugiau tikrų teigiamų atvejų.

F1 balas (angl. *F1 score*) yra preciziškumo ir jautrumo vidurkis. Jis subalansuoja abi metrikas ir yra naudingas tuomet, kai duomenų klasės yra nesubalansuotos.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

F1 balas svyruoja tarp 0 ir 1, kur 1 reiškia idealų modelį. Tai viena patikimiausių metrikų, kai neigiamas ir teigiamas klaidingas klasifikavimas turi didelę reikšmę.

Daugiaklasės klasifikacijos (angl. *multi-label classification*, toliau – MLC) atveju (pvz., amžiaus ar etninės grupės klasifikacijoje), metrikos, tokios kaip preciziškumas ir jautrumas yra skaičiuojamos kiekvienai klasei atskirai.

2.4. Eksperimentinio tyrimo aplinka ir technologijos

Eksperimentams atlikti buvo naudojamas stacionarus kompiuteris su „Intel Core i7-14700k“ procesoriumi ir 64 GB DDR5 operatyviosios atminties. Modelių kūrimas ir testavimas buvo vykdomas „Windows 11“ operacinės sistemos aplinkoje naudojant „Python“ 3.9.18 versiją.

Modelių implementavimui ir duomenų apdorojimui buvo naudojami keli „Python“ programavimo kalba paremti karkasai ir bibliotekos. Neuroninių tinklų kūrimui, mokymui ir testavimui buvo naudotas giliojo mokymo karkasas „PyTorch“, duomenų apdorojimui ir paruošimui buvo naudojamos „Pandas“ ir „Numpy“ bibliotekos, prižiūrimumų giliojo mokymo algoritmų kūrimui buvo naudojama „scikit-learn“ biblioteka.

Programavimo karkasas „PyTorch“ buvo naudotas giliųjų neuroninių tinklų kūrimui. Šis karkasas yra plačiai paplitęs neuroninių tinklų kūrimo srityje ir pradžioje buvo sukurtas Meta AI kompanijos. Šis karkasas yra atvirojo kodo projektas, kas leidžia bendruomenei prisidėti prie jo vystymo ir tobulinimo. Atvirojo kodo projektai yra naudingi tuo, kad jie nėra pilnai priklausomi nuo kokios nors kompanijos ir jau egzistuojančios kodo versijos negali būti apmokestintos ar kitaip priverstinai neprieinamos. Šis karkasas leidžia intuityviai dėti neuroninių tinklų architektūras, negaištant laiko prie smulkių techninių detalių. Jis taip pat leidžia žmonėms, kurie nėra programavimo srities ekspertai, lengviau kurti dirbtinio intelekto

sprendimus (Paszke ir kt., 2019). Šiame darbe „PyTorch“ buvo naudojamas neuroninių tinklų, pritaikytų veidų atributų nustatymui iš savybių vektorių, kūrimui bei mokymui.

Duomenų rinkinio apdorojimui buvo pasirinkta naudoti „Pandas“ biblioteka. „Pandas“ yra duomenų analizės ir apdorojimo biblioteka skirta darbui su lentelės formos (angl. *tabular*) duomenimis (Mckinney, 2011). Duomenų rinkinys yra sudarytas iš lentelės formos duomenų, t.y., savybių vektorių galima išdėstyti kaip stulpelius kartu su klasifikacijos klase, todėl tokia apdorojimo biblioteka yra patogi duomenų paruošimui. „Pandas“ pateikia aukštos abstrakcijos duomenų struktūras, tokias kaip „DataFrame“ patogesniai duomenų filtravimui, valymui bei transformavimui. Šios bibliotekos didelis pranašumas yra tai, kad ji yra pritaikyta gerai sąveikauti su kitomis populiariomis bibliotekomis, tokiomis kaip „scikit-learn“ bei „Numpy“.

Viena esminių „Python“ bibliotekų duomenų moksle yra „Numpy“. Ši biblioteka teikia daugybę patogių įrankių darbui su įvairiais masyvais, matricomis ir skaičiais (van der Walt ir kt., 2011). Ši biblioteka siūlo ne tik duomenų struktūras masyvų formos duomenims, bet ir teikia metodus, skirtus matematinių ir kitų operacijų vykdymui su matricomis ir kitomis struktūromis. Ši biblioteka labai gerai sąveikauja su „PyTorch“ programavimo karkasu, kadangi neuroniniai tinklai reikalauja daugybę matricų daugybos operacijų.

„Scikit-learn“ programavimo biblioteka buvo naudojama prižiūrimojo mokymo modelių mokymui bei hiperparameterių paieškai. Ši biblioteka teikia daug pagalbinių metodų darbui su giliojo mokymo algoritmais ir leidžia atlikti įvairius bandymus pasitelkiant aukšto lygio abstrakcijas (Pedregosa ir kt., 2012). Naudojantis šia biblioteka galima lengviau išvesti įvairias modelių metrikas ir rezultatus, kadangi joje yra pateikti metodai ir funkcijos, kaip neimplementuojant matematinių skaičiavimų programuotojui galima gauti tokias metrikas, kaip „F1“ įvertis (angl. *F1-score*), modelio tikslumas, „ROC“ kreivės ir panašiai.

3. EKSPERIMENTINIO TYRIMO REZULTATAI

Šiame skyriuje pateikti eksperimentinio tyrimo rezultatai, kurių tikslas yra įvertinti, ar galima tiksliai nustatyti žmogaus veido savybes, tokias kaip lytį, etninę grupę ir amžių, naudojantis veidų šablonais, gautais iš veidų atpažinimo algoritmo neuroninio tinklo. Tyrime taikyti tiek giliojo mokymo, tiek tradiciniai mašininio mokymo (angl. *machine learning*, toliau – ML) algoritmai, tokie kaip sprendimų medžiai (angl. *decision tree*) bei atsitiktiniai miškai (angl. *Random Forest*).

Visiems metodams taikytas vienodas duomenų padalijimo principas: 80 % duomenų buvo paskirti modelio mokymui ir hiperparametrų optimizavimui, o likę 20 % – validacijai. Siekiant užtikrinti patikimus vertinimo rezultatus prižiūrimojo mokymo metoduose, mokymo duomenų rinkinyje taikytas penkių kartų kryžminės validacijos (angl. *5-fold cross-validation*) metodas.

Norint pasiekti kuo įmanoma geresnius rezultatus, buvo atlikta hiperparametrų optimizacija bei išbandyti įvairūs modeliai bei tinklų architektūros. Prižiūrimojo mokymo metodams buvo pasitelkta tinklelio paieška (angl. *grid search*), o neuroniniams tinklams buvo keičiami parametrai, tokie kaip mokymosi greitis, mokymosi epochų skaičius, imties dydis bei buvo testuojamos įvairios modelių architektūros.

3.1. Tyrimo rezultatai naudojant DL metodus

Vienas iš pagrindinių metodų, taikytų šiame darbe, veido savybių klasifikavimui iš veidų savybių vektorių, buvo giliaji neuroniniai tinklai (angl. *deep neural networks*, toliau – DNN). Eksperimentuose naudoti kelių sluoksnių perceptronai (angl. *multi-layer perceptron*, toliau – MLP), sudaryti iš įėjimo sluoksnio (atitinkančio savybių vektoriaus dimensijas), kelių paslėptųjų sluoksnių ir išėjimo sluoksnio, atitinkančio spėjimų klasių skaičių.

Geriausią hiperparametrų paieškai buvo pasirinkta naudoti *Adam* optimizatorių. Eksperimentų metu buvo išmėginti mokymosi greičiai (angl. *learning rate*, toliau – LR) 0,01, 0,001, 0,0001 bei 0,00001. Taip pat buvo ieškomas optimaliausias imties dydis (angl. *batch size*, toliau – BS), eksperimentuojant su 32, 64, 256 bei 512 imties įrašų kiekiu.

Modeliai buvo mokomi iki 100 epochų, tačiau buvo naudojamas ankstyvo stabdymo (angl. *early stopping*) mechanizmas, kuris sustabdydavo mokymą, jei per 10 epochų nebuvo pasiektas reikšmingas praradimo (angl. *loss*, toliau – L) pagerėjimas validacijos duomenų aibėje. Ankstyvajam stabdymui buvo naudotas pakantumo (angl. *patience*) parametras su reikšme 10, o minimali L reikšmė buvo 0,0001. Šie parametrai reiškia, kad jei mokymosi praradimo funkcijos rezultatas nesumažėja per 0,0001 reikšmę per 10 mokymo epochų, tai modelių mokymas yra stabdomas ir išsaugomas geriausią rezultatą pasiekęs modelis.

Šis metodas leidžia išvengti modelių persimokymo ir paspartina eksperimentų eigą, kadangi nereikia laukti, kol mokymas tęsis visas 100 epochų.

Kiekvienai veido savybei nustatyti buvo kuriami atskiri neuroniniai tinklai su skirtingomis architektūromis ir skirtingais taikytais metodais, siekiant rasti geriausius modelius. Buvo taikytos skirtingos reguliarizacijos technikos, tokios kaip išmetimo sluoksniai (angl. *dropout*), imties normalizacijos sluoksniai (angl. *batch normalization*, toliau – BN).

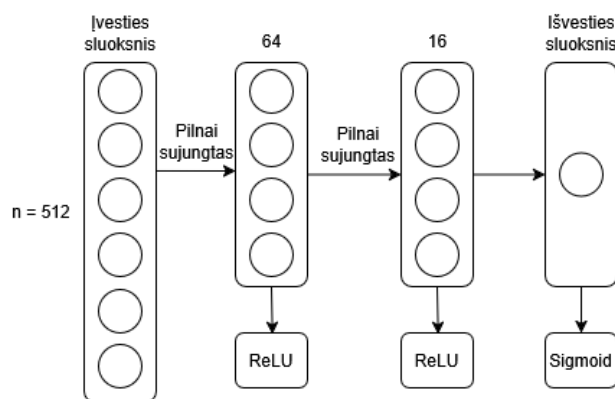
Lyties nustatymui iš viso buvo sukurti ir išmėginti 7 DNN, etninės grupės nustatymui – 6 DNN, o amžiaus grupės nustatymui – 5. Įvertinus gautus modelių rezultatus kiekvienos savybės nustatymo užduotyje, buvo atrinkti 9 neuroninių tinklų modeliai, parodę geriausią rezultatą. Toliau darbe yra pateikiama šių atrinktų modelių ir jų rezultatų analizė.

Iš viso buvo atlikti 144 neuroninių tinklų mokymai atliekant hiperparametrų paiešką, kur kiekvienas mokymas galėjo trukti nuo 10 iki 100 epochų, priklausomai nuo to, kada įsijungia ankstyvo stabdymo mechanizmas ir modelis pasiekia geriausią rezultatą.

Vienas iš šiame darbe naudotų neuroninių tinklų modelių – „TemplateSexFC1Net“ – pilnai sujungtas dirbtinis neuroninis tinklas, pritaikytas bendrajai lyties klasifikacijos užduočiai iš veido šablono. Modelis yra sudarytas iš trijų pilnai sujungtų sluoksnių:

1. Įėjimo sluoksnis – atitinka veido šablono vektorių elementų kiekį – 512.
2. Pirmasis tankus sluoksnis sumažina dimensiją iki 64 neuronų ir pritaiko *ReLU* aktyvacijos funkciją.
3. Antrasis tankus sluoksnis sumažina matmenis iki 16 neuronų ir pritaiko aktyvacijos funkciją.
4. Išvesties sluoksnis turi vieną neuroną ir yra taikoma sigmoidinė aktyvacijos funkcija. Šis sluoksnis leidžia interpretuoti išvestį kaip tikimybę, kokiai lyčiai priklauso žmogaus šablono vektorius.

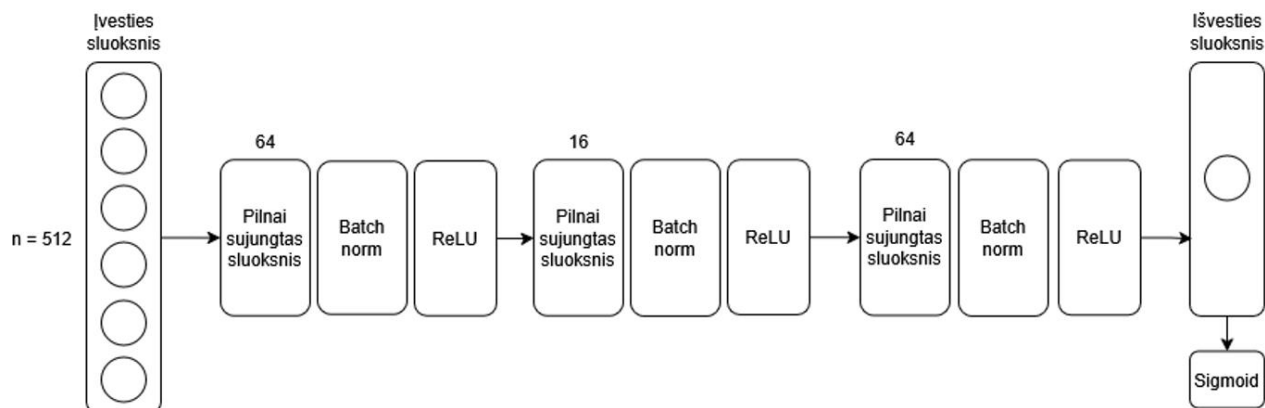
Modelis buvo mokomas naudojant dvejetainės krosentropijos nuostolio (angl. *binary cross entropy*, toliau – BCE) funkciją, kadangi klasifikacijos užduotis yra dvejetainė – vyras arba moteris. Modelio architektūra pavaizduota 6 pav.



6 pav. „TemplateSexFC1Net“ modelio architektūra

Kita bandyta neuroninio tinklo architektūra lyties atpažinimui buvo „TemplateSexFC2Net“. Šis tinklas yra didesnis nei pirmasis tirtas tinklas ir jame yra įtraukiamas naujo tipo metodas – imties normalizacijos sluoksniai. Tai yra vienas esminių architektūros patobulinimų. Šis normalizacijos metodas buvo mėgintas siekiant stabilizuoti ir pagerinti neuroninių tinklų mokymą. Jis sumažina vidinio neuronų aktyvacijų pasislinkimo reiškinį (angl. *internal covariate shift*) ir padeda išlaikyti tinkamesnį gradientų nusileidimą.

Ši technika leidžia naudoti didesnius mokymosi žingsnius, sumažina gradientų sprogimo arba išnykimo riziką ir dažnai pagerina bendrą modelio generalizavimo gebėjimą. Šio modelio architektūra pavaizduota 7 pav.



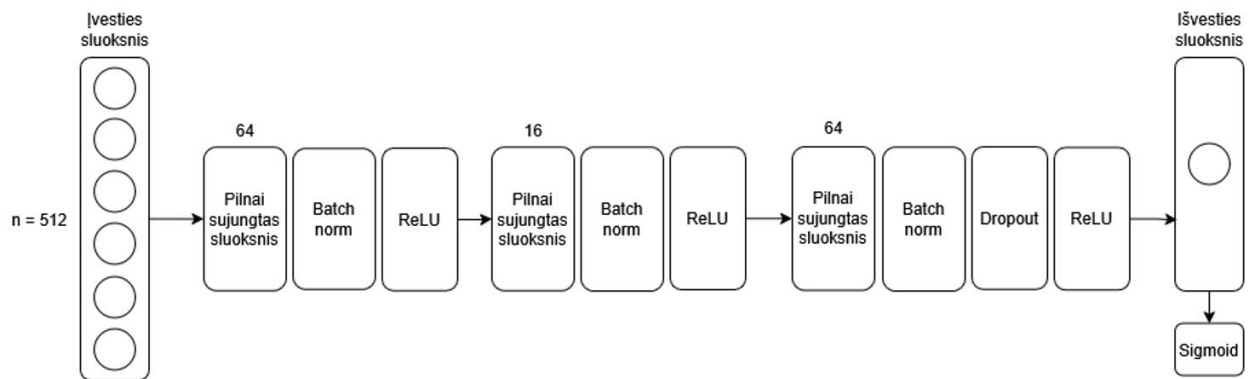
7 pav. „TemplateSexFC2Net“ tinklo architektūra

Trečioji bandyta neuroninio tinklo architektūra, pavadinta „TemplateSexFC3Net“, pasižymi didesniu sluoksnių skaičiumi ir papildomais regularizacijos mechanizmais, skirtais pagerinti modelio generalizavimo gebėjimus. Šiame modelyje išlaikoma ta pati bendroji struktūra, kaip ir ankstesnėse versijose, tačiau esminis skirtumas yra išmetimo (angl. *dropout*) sluoksnio įtraukimas, kuris taikomas po

trečiojo paslėpto sluoksnio prieš aktyvacijos funkciją. Ši technika yra ypač svarbi neuroniniuose tinkluose, kadangi ji padeda sumažinti persimokymo (angl. *overfitting*) riziką.

Vienas iš svarbiausių šios architektūros elementų yra išmetimo metodas. Tai – atsitiktinis tam tikros dalies neuronų išjungimas mokymosi metu, kuris padeda išvengti situacijos, kai modelis tampa pernelyg priklausomas nuo tam tikrų neuronų kombinacijų. Mokymosi metu kiekvienas neuronas turi tam tikrą tikimybę būti išjungtas, t.y., jo reikšmė yra laikoma lygia nuliui. Tai priverčia tinklą išmokti atsparias, mažiau koreliuojančias sąsajas.

Šioje modelio architektūroje išmetimo tikimybė buvo lygi 0,5. Tai reiškia, kad kiekvienas neuronas, esantis tame sluoksnyje, turi 50 % tikimybę būti išjungtas kiekvienos imties metu. Šis sprendimas buvo grindžiamas praktika, kai vidutinės apimties tinkluose 0,5 reikšmė dažnai pasirodo esanti veiksminga kompromiso tarp reguliacijos ir mokymosi efektyvumo prasme. Šio tinklo architektūra pavaizduota 8 pav.



8 pav. „TemplateSexFC3Net“ tinklo architektūra

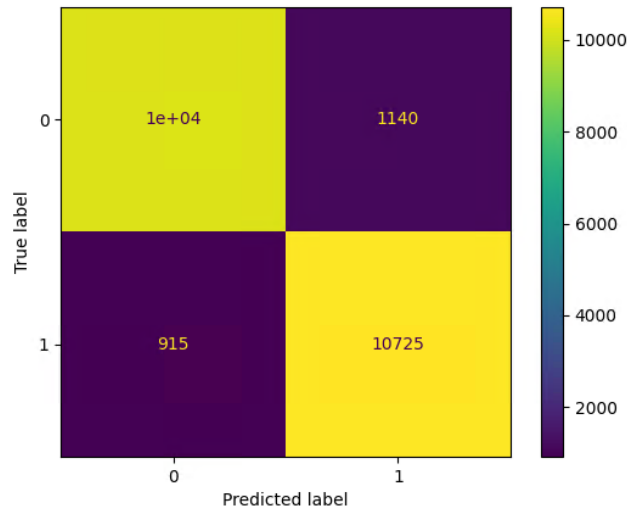
Toliau buvo atliktas visų aprašytų neuroninių tinklų, skirtų lyties atpažinimui iš veidų savybių vektorių, mokymas. Pirmasis modelis pasižymėjo itin paprasta architektūra – vos dviem paslėptaisiais sluoksniais su *ReLU* aktyvacijos funkcija. Šis modelis buvo naudojamas kaip atskaitos taškas, leidžiantis įvertinti sudėtingesnių architektūrų pridėtą naudą. Antrojo modelio struktūra buvo išplėsta įtraukiant imties normalizacijos sluoksnius tarp kiekvieno pilnai sujungto sluoksnio ir aktyvacijos funkcijų. Trečiasis modelis dar labiau išplėtė antrosios architektūros struktūrą ir įtraukė papildomą reguliarizacijos sluoksnį.

Modelių palyginamoji analizė ir hiperparametrų paieška parodė, kad sudėtingesnis modelis nebūtinai reiškia geresnį rezultatą, kaip matoma 1 lentelėje. Antrasis modelis, nors ir turėjo papildomus normalizacijos sluoksnius, parodė prastesnį rezultatą, lyginant su paprasčiausiu, pirmuoju modeliu. Tačiau trečiojo modelio validacijos tikslumas ir F1 balas buvo didesnis nei ankstesnių modelių.

1 lentelė. Lyties atpažinimo neuroninių tinklų modelių rezultatai

Modelis	Ypatybės	Epocha	LR	BS	Tikslumas (%)	F1 balas
TemplateSexFC1Net	FC	4	0,01	32	90,71	0,9046
TemplateSexFC1Net	FC	10	0,01	64	91,09	0,9143
TemplateSexFC1Net	FC	6	0,01	256	90,53	0,9036
TemplateSexFC1Net	FC	5	0,01	512	91,39	0,9156
TemplateSexFC1Net	FC	2	0,001	32	91,65	0,918
TemplateSexFC1Net	FC	3	0,001	64	91,75	0,9187
TemplateSexFC1Net	FC	2	0,001	256	91,75	0,9186
TemplateSexFC1Net	FC	3	0,001	512	91,7	0,9182
TemplateSexFC1Net	FC	5	0,0001	32	91,45	0,9156
TemplateSexFC1Net	FC	7	0,0001	64	91,45	0,9152
TemplateSexFC1Net	FC	11	0,0001	256	91,06	0,9122
TemplateSexFC1Net	FC	13	0,0001	512	92,85	0,9302
TemplateSexFC1Net	FC	24	0,00001	32	91,21	0,9133
TemplateSexFC1Net	FC	36	0,00001	64	91,17	0,913
TemplateSexFC1Net	FC	24	0,00001	256	90,31	0,9037
TemplateSexFC1Net	FC	18	0,00001	512	87,13	0,8738
TemplateSexFC2Net	FC, BN	8	0,01	32	92,49	0,9251
TemplateSexFC2Net	FC, BN	6	0,01	64	92,33	0,9239
TemplateSexFC2Net	FC, BN	2	0,01	256	92,06	0,9216
TemplateSexFC2Net	FC, BN	4	0,01	512	92,12	0,9216
TemplateSexFC2Net	FC, BN	7	0,001	32	92,18	0,9228
TemplateSexFC2Net	FC, BN	4	0,001	64	92,11	0,9218
TemplateSexFC2Net	FC, BN	2	0,001	256	92,12	0,9222
TemplateSexFC2Net	FC, BN	3	0,001	512	91,87	0,9202
TemplateSexFC2Net	FC, BN	8	0,0001	32	92,2	0,9225
TemplateSexFC2Net	FC, BN	11	0,0001	64	91,98	0,921
TemplateSexFC2Net	FC, BN	8	0,0001	256	91,29	0,9145
TemplateSexFC2Net	FC, BN	9	0,0001	512	92,66	0,9282
TemplateSexFC2Net	FC, BN	35	0,00001	32	91,77	0,9179
TemplateSexFC2Net	FC, BN	32	0,00001	64	91,54	0,9161
TemplateSexFC2Net	FC, BN	32	0,00001	256	90,73	0,9077
TemplateSexFC2Net	FC, BN	26	0,00001	512	88,91	0,89
TemplateSexFC3Net	FC, BN, išmetimas	9	0,01	32	92,35	0,9242
TemplateSexFC3Net	FC, BN, išmetimas	5	0,01	64	92,24	0,9237
TemplateSexFC3Net	FC, BN, išmetimas	6	0,01	256	92,12	0,922
TemplateSexFC3Net	FC, BN, išmetimas	4	0,01	512	92,27	0,9239
TemplateSexFC3Net	FC, BN, išmetimas	4	0,001	32	92,15	0,9233
TemplateSexFC3Net	FC, BN, išmetimas	6	0,001	64	91,95	0,9207
TemplateSexFC3Net	FC, BN, išmetimas	3	0,001	256	91,79	0,9191
TemplateSexFC3Net	FC, BN, išmetimas	3	0,001	512	91,71	0,9184
TemplateSexFC3Net	FC, BN, išmetimas	7	0,0001	32	91,72	0,9202
TemplateSexFC3Net	FC, BN, išmetimas	5	0,0001	64	91,66	0,9154
TemplateSexFC3Net	FC, BN, išmetimas	10	0,0001	256	91,39	0,9149
TemplateSexFC3Net	FC, BN, išmetimas	11	0,0001	512	93,04	0,9312
TemplateSexFC3Net	FC, BN, išmetimas	43	0,00001	32	91,89	0,9204
TemplateSexFC3Net	FC, BN, išmetimas	43	0,00001	64	91,75	0,9183
TemplateSexFC3Net	FC, BN, išmetimas	31	0,00001	256	89,85	0,9021
TemplateSexFC3Net	FC, BN, išmetimas	28	0,00001	512	88,2	0,8782

Siekiant detaliau įvertinti klasifikavimo modelių kokybę, buvo sudarytos geriausią rezultatą pasiekusio modelio sumaišties matricos (angl. *confusion matrix*, toliau – CM). Šios matricos leidžia objektyviai įvertinti, kaip dažnai modelis teisingai atpažįsta kiekvieną klasę, ir kokios klasės yra dažniausiai painiojamos tarpusavyje. Tai yra svarbu klasifikavimo užduotyse su keliomis klasėmis, kur bendras tikslumo ar F1 balo įvertinimas gali nublėpti nekorektišką modelio veikimą, pvz., kai viena klasė dominuoja arba kai tam tikros klasės yra nuolat painiojamos.



9 pav. Lyties spėjimo sumaišties matrica

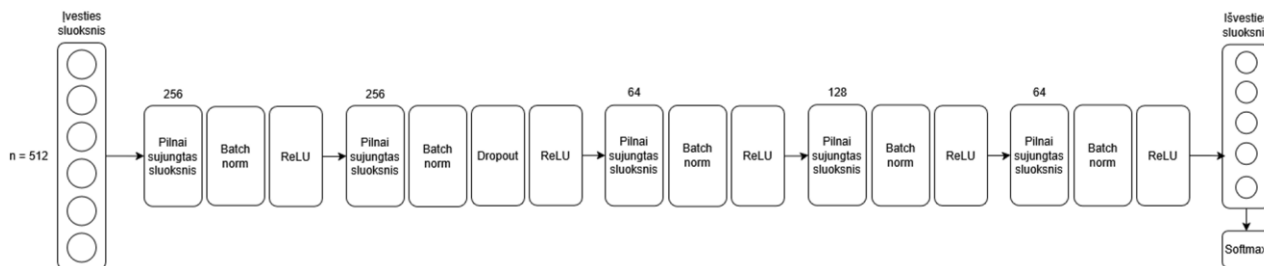
Geriausią tikslumo rezultatą pasiekusio modelio „TemplateSexFC3Net“ lyties klasifikavimo sumaišties matrica rodo, kad modelis gerai atpažįsta abi klases – tiek vyrų, tiek moterų. Paveikslėlyje vyrų klasė yra žymima „0“ reikšme, o moterų klasė yra išreikšta „1“ verte. Klasifikavimo tikslumas yra panašus abejoms klasėms, o klaidų pasiskirstymas tarp klasių skiriasi nedaug (žr. 9 pav.).

Šie duomenys leidžia teigti, kad modelis nėra šališkas (angl. *biased*) nei vienai klasei. Tokie rezultatai rodo, kad mokymo duomenys buvo pakankamai subalansuoti, o modelio architektūra tinkama spęsti šią dvejetainę klasifikavimo užduotį.

Etninės grupės nustatymo užduočiai taip pat buvo sukurti keli pilnai sujungtų neuroninių tinklų modeliai, kurių struktūra buvo panaši į jau naudotus lyties klasifikacijai modelius. Šie tinklai buvo pritaikyti daugiaklasei klasifikacijai. Vienas iš tirtų modelių buvo „EthnicityFC1Net“, kurio esminis skirtumas yra išvesties sluoksnis, naudojantis „softmax“ aktyvacijos funkciją, kad būtų galima išvesti skirtingų klasių spėjimo tikimybes.

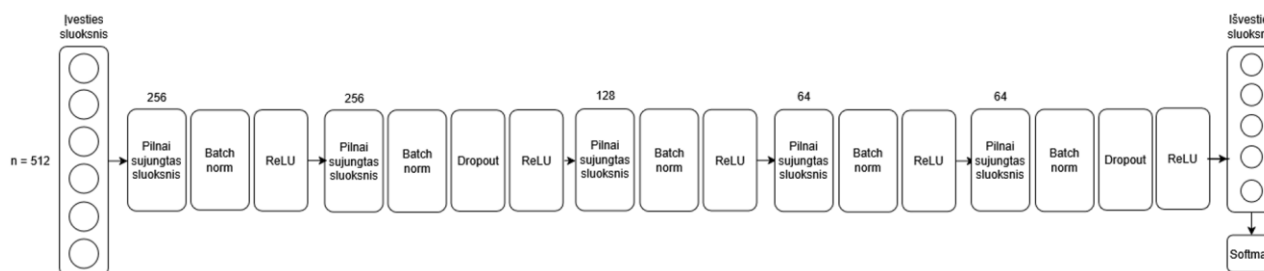
„TemplateEthnicityFC1Net“ tinklas susideda iš penkių paslėptų sluoksnių, kur kiekvieno sluoksnio išvestis yra normalizuojama naudojantis imties normalizacija. Taip pat remiantis ankstesniais duomenimis,

kad lyties spėjimo modelis pasirodė geriausiai naudojant normalizacijos sluoksnius bei išmetimo sluoksnius, buvo įtraukti šie metodai. Šio modelio architektūra yra pavaizduota 10 pav.



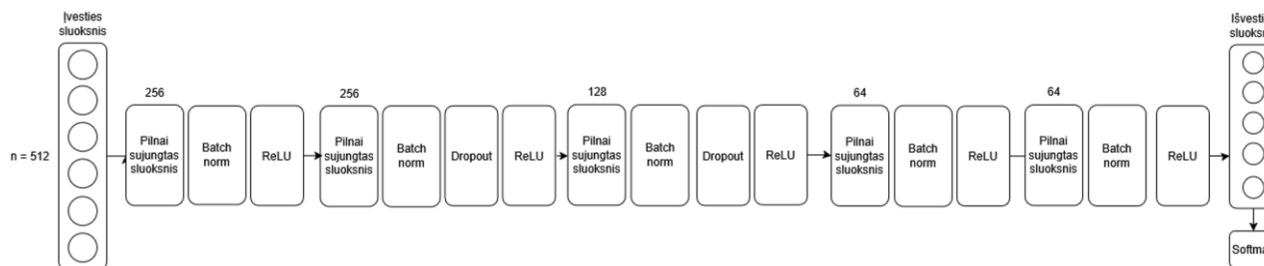
10 pav. „TemplateEthnicityFC1Net“ tinklo architektūra

Antrasis sukurtas neuroninis tinklas, skirtas etninės grupės klasifikavimui, buvo „TemplateEthnicityFC2Net“. Šio modelio architektūra yra labai panaši į pirmojo tinklo, tačiau jis pasižymi keliais esminiais skirtumais. Vienas pagrindinių pokyčių yra padidintas tinklo gylis ir papildomai įtrauktas išmetimo sluoksnis viename iš paskutiniųjų paslėptųjų sluoksnių (žr. 11 pav.).



11 pav. „TemplateEthnicityFC2Net“ tinklo architektūra

Trečiojoje tinklo architektūroje buvo išmėginta išmetimo technika tinklo viduryje. Jo architektūra yra analogiška antrajam tinklui, tačiau buvo pakeista „dropout“ sluoksnio vieta ir pašalintas vienas „dropout“ sluoksnis (žr. 12 pav.).



12 pav. „TemplateEthnicityFC3Net“ tinklo architektūra

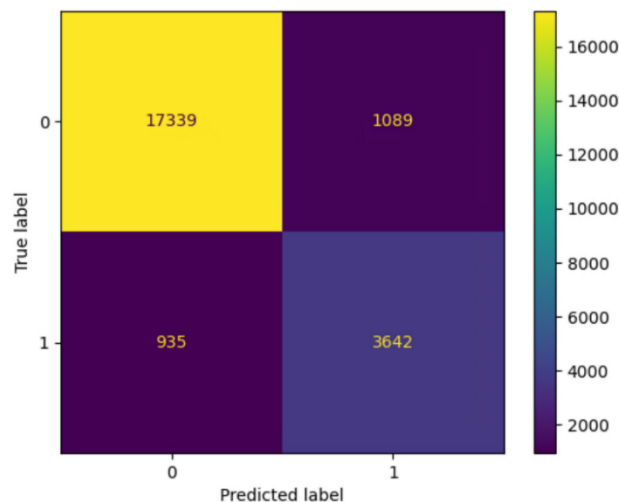
2 lentelė. Etninės grupės atpažinimo neuroninių tinklų modelių rezultatai

Modelis	Ypatybės	Epocha	LR	BS	Tikslumas (%)	F1 balas
TemplateEthnicityFC1Net	FC, BN, išmetimas	34	0,01	32	71,15	0,7125
TemplateEthnicityFC1Net	FC, BN, išmetimas	31	0,01	64	71,69	0,7167
TemplateEthnicityFC1Net	FC, BN, išmetimas	17	0,01	256	70,36	0,7034
TemplateEthnicityFC1Net	FC, BN, išmetimas	13	0,01	512	70,46	0,7047
TemplateEthnicityFC1Net	FC, BN, išmetimas	31	0,001	32	71,81	0,7187
TemplateEthnicityFC1Net	FC, BN, išmetimas	27	0,001	64	71,47	0,7129
TemplateEthnicityFC1Net	FC, BN, išmetimas	12	0,001	256	70,7	0,7075
TemplateEthnicityFC1Net	FC, BN, išmetimas	9	0,001	512	70,92	0,7093
TemplateEthnicityFC1Net	FC, BN, išmetimas	29	0,0001	32	72,47	0,725
TemplateEthnicityFC1Net	FC, BN, išmetimas	21	0,0001	64	71,76	0,7168
TemplateEthnicityFC1Net	FC, BN, išmetimas	13	0,0001	256	69,81	0,6977
TemplateEthnicityFC1Net	FC, BN, išmetimas	20	0,0001	512	70,84	0,7075
TemplateEthnicityFC1Net	FC, BN, išmetimas	58	0,00001	32	71,26	0,7132
TemplateEthnicityFC1Net	FC, BN, išmetimas	49	0,00001	64	70,04	0,6997
TemplateEthnicityFC1Net	FC, BN, išmetimas	32	0,00001	256	66,06	0,6603
TemplateEthnicityFC1Net	FC, BN, išmetimas	31	0,00001	512	64,4	0,6436
TemplateEthnicityFC2Net	FC, BN, išmetimas	24	0,01	32	68,08	0,6824
TemplateEthnicityFC2Net	FC, BN, išmetimas	25	0,01	64	69,61	0,6973
TemplateEthnicityFC2Net	FC, BN, išmetimas	20	0,01	256	70,27	0,7019
TemplateEthnicityFC2Net	FC, BN, išmetimas	12	0,01	512	70,11	0,7003
TemplateEthnicityFC2Net	FC, BN, išmetimas	25	0,001	32	71,27	0,7138
TemplateEthnicityFC2Net	FC, BN, išmetimas	24	0,001	64	71,11	0,7112
TemplateEthnicityFC2Net	FC, BN, išmetimas	17	0,001	256	70,95	0,7091
TemplateEthnicityFC2Net	FC, BN, išmetimas	12	0,001	512	70,9	0,7091
TemplateEthnicityFC2Net	FC, BN, išmetimas	34	0,0001	32	72,38	0,7229
TemplateEthnicityFC2Net	FC, BN, išmetimas	21	0,0001	64	71,81	0,7193
TemplateEthnicityFC2Net	FC, BN, išmetimas	18	0,0001	256	71,22	0,7129
TemplateEthnicityFC2Net	FC, BN, išmetimas	14	0,0001	512	69,81	0,6986
TemplateEthnicityFC2Net	FC, BN, išmetimas	59	0,00001	32	71,42	0,7142
TemplateEthnicityFC2Net	FC, BN, išmetimas	47	0,00001	64	69,64	0,6977
TemplateEthnicityFC2Net	FC, BN, išmetimas	34	0,00001	256	66,5	0,665
TemplateEthnicityFC2Net	FC, BN, išmetimas	30	0,00001	512	63,87	0,6369
TemplateEthnicityFC3Net	FC, BN, išmetimas	44	0,01	32	71,33	0,7135
TemplateEthnicityFC3Net	FC, BN, išmetimas	33	0,01	64	71,48	0,7141
TemplateEthnicityFC3Net	FC, BN, išmetimas	18	0,01	256	70,62	0,7071
TemplateEthnicityFC3Net	FC, BN, išmetimas	11	0,01	512	69,71	0,6982
TemplateEthnicityFC3Net	FC, BN, išmetimas	37	0,001	32	71,68	0,7167
TemplateEthnicityFC3Net	FC, BN, išmetimas	21	0,001	64	71,07	0,7097
TemplateEthnicityFC3Net	FC, BN, išmetimas	13	0,001	256	70,57	0,7063
TemplateEthnicityFC3Net	FC, BN, išmetimas	21	0,001	512	71,07	0,7097
TemplateEthnicityFC3Net	FC, BN, išmetimas	27	0,0001	32	71,95	0,7186
TemplateEthnicityFC3Net	FC, BN, išmetimas	25	0,0001	64	71,25	0,7124
TemplateEthnicityFC3Net	FC, BN, išmetimas	14	0,0001	256	69,55	0,6955
TemplateEthnicityFC3Net	FC, BN, išmetimas	14	0,0001	512	69,33	0,6928
TemplateEthnicityFC3Net	FC, BN, išmetimas	55	0,00001	32	69,92	0,6993
TemplateEthnicityFC3Net	FC, BN, išmetimas	43	0,00001	64	68,86	0,6894
TemplateEthnicityFC3Net	FC, BN, išmetimas	29	0,00001	256	66,18	0,6618
TemplateEthnicityFC3Net	FC, BN, išmetimas	28	0,00001	512	64,5	0,6439

Remiantis atliktais eksperimentais, buvo apmokyti ir įvertinti trys skirtingi neuroninių tinklų modeliai, skirti etninės grupės klasifikavimui pagal veido šabloną. Geriausio modelio radimui buvo atlikta hiperparametrų paieška su 48 unikaliomis kombinacijomis. Visi modeliai yra pagrįsti pilnai sujungtomis architektūromis, kuriose naudojami imties normalizacijos bei išmetimo sluoksniai. 2 lentelėje apibendrinami pagrindiniai šių modelių rezultatai.

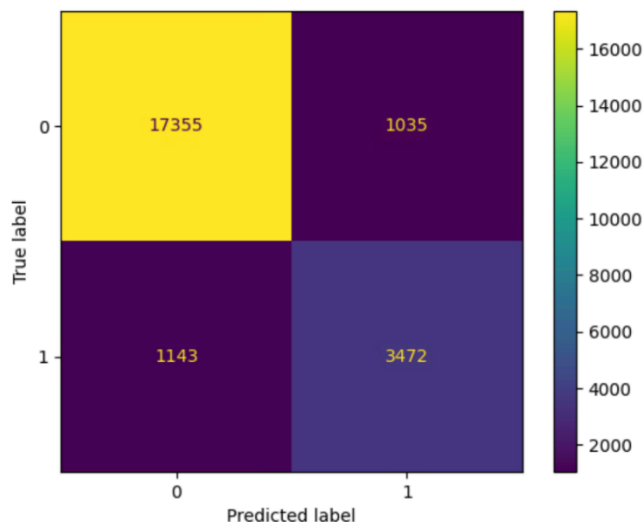
Pirmasis modelis, „TemplateEthnicityFC1Net“, pasirodė geriausiai pagal bendrą klasifikacijos tikslumą bei F1 balą. Jo architektūra pasižymi subalansuotu sudėtingumo lygiu, kuris gerai atitiko turimų duomenų struktūrą. Tuo tarpu antrasis modelis, buvo didesnio gylio – turėjo papildomą pilnai sujungtą bei išmetimo sluoksnį, tačiau tai nesuteikė tikslumo pagerėjimo. Trečiasis modelis pasirodė prasčiausiai tiek pagal validacijos tikslumo, tiek pagal F1 metriką. Tai leidžia daryti prielaidą, kad nors tinklo architektūra buvo pakankamai gili, ji nėra optimali etninės grupės spėjimo užduočiai.

Siekiant išsiaiškinti ar geriausias etninės grupės nustatymo modelis, „TemplateEthnicityFC1Net“ nėra šališkas kuriai nors klasei, buvo apžvelgti kiekvienos klasės sumaišties matricų rezultatai. Kadangi šis modelis atlieka daugiaklasę klasifikaciją, tai CM yra pateiktos kiekvienai klasei kaip atskiri paveikslėliai. Tai leidžia vizualiai apžvelgus matricą nuspręsti ar nėra stipraus šališkumo kuriai nors klasei.



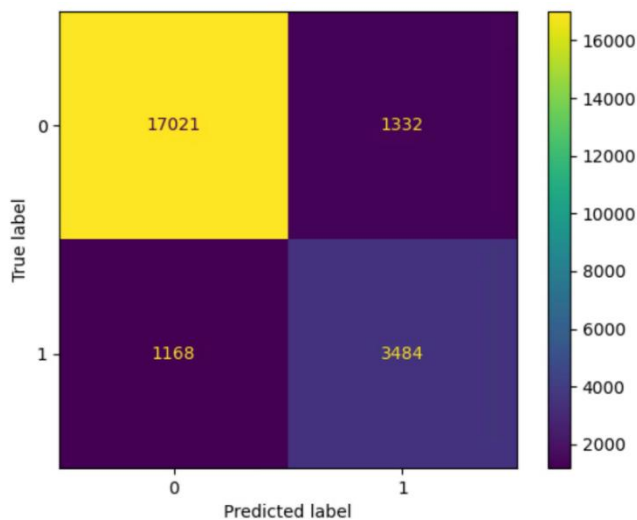
13 pav. Baltaodžių etninės grupės spėjimo sumaišties matrica

Pirmoji pateikta CM atspindi modelio spėjimus, nustatant baltaodžių etninės grupės klasę. Sumaišties matricoje „0“ vertė atspindi neigiamą sprendimą, t.y., modelis spėja, kad tai nėra baltaodžio klasės duomenys, o „1“ vertė reiškia, kad modelis spėja, jog duomenys priklauso baltaodžių etninės grupės klasei (žr. 13 pav.). Pagal šią sumaišties matricą galima spręsti, jog modelis gebėjo atskirti šią klasę pakankamai gerai. Apskaičiavus tikslumą (angl. *precision*) šiai klasei, gaunamas rezultatas yra 79,572 %, o jautrumo (angl. *recall*) reikšmė yra 76,982 %.



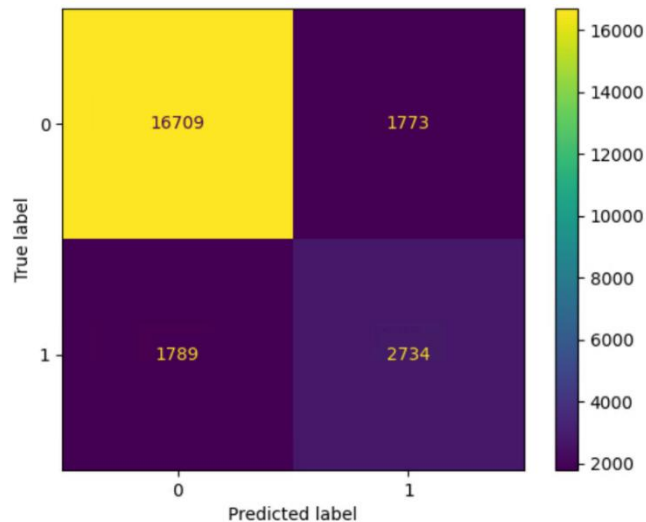
14 pav. Juodaodžių etninės grupės spėjimo sumaištis matrica

Kita modelio nustatoma klasė buvo ar žmogus priklauso juodaodžių kilmės etninei grupei. Geriausio modelio CM yra pateikta 14 pav. Kaip ir pirmajai klasei, taip ir šiai, galima apskaičiuoti pagrindines metrikas. Modelio tikslumas šiai klasei buvo lygus 75,233 %, o jautrumas buvo 77,036 %. Šie rezultatai yra artimi pirmosios klasės spėjimo rezultatams bei bendram pasvertajam modelio F1 balui, kas reiškia, jog modelis geba atpažinti šią klasę ganėtinai patikimai.



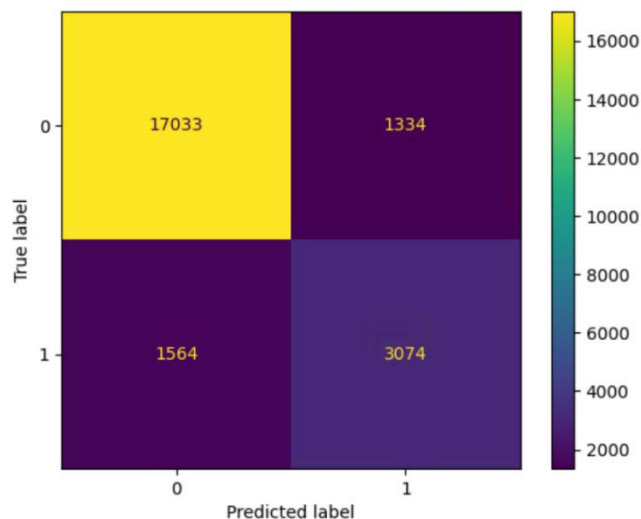
15 pav. Azijiečių etninės grupės spėjimo sumaištis matrica

Kita etninės grupės atpažįstama klasė buvo azijiečių grupė. Šios klasės modelio sumaištis matrica yra pavaizduota 15 pav. Modelio tikslumas šiai klasei buvo 74,893 %, o jautrumo reikšmė – 72,342 %. Šie rezultatai per daug nenukrypsta nuo bendro modelio tikslumo bei anksčiau analizuotų klasių. Šioje CM nėra matoma jokių modelio tikslumo anomalijų.



16 pav. Ispanų kilmės etninės grupės spėjimo sumaišties matrica

Ispanų kilmės etninės grupės atpažinimo modelio sumaišties matrica yra pavaizduota 16 pav. Apskaičiavus geriausio „TemplateEthnicityFC1Net“ modelio tikslumo ir jautrumo metrikas ispanų kilmės etninės grupės klasei, buvo gauti tokie rezultatai: tikslumas – 60,447 %, o jautrumas – 60,661 %. Tai yra prasčiausiais rezultatai iš visų ankščiau apžvelgtų klasių. Tai reiškia, kad šis modelis sunkiau nuspėja veidų, priklausančių šiai etninei grupei, klasę. Viena to priežasčių gali būti tai, kad ispanų kilmės žmonių veidai iš principo yra panašūs į įprastus baltaodžius. Gali būti, jog veido šablone nėra užkoduojama pakankamai informacijos, kuria remiantis būtų galima užtikrintai atskirti šios klasės veidus.



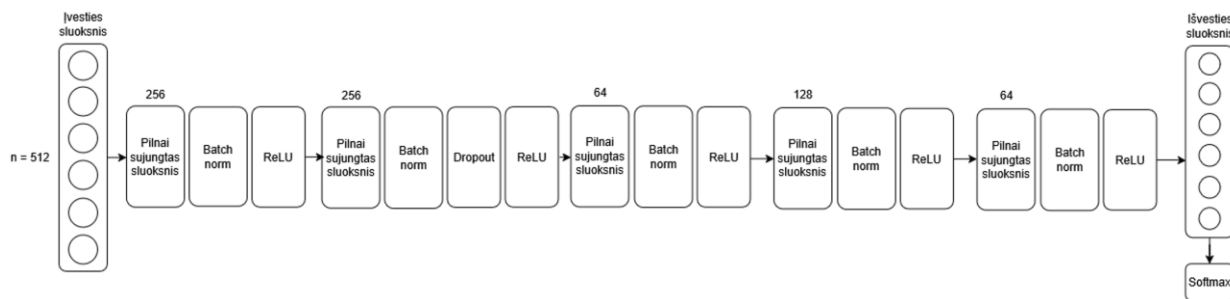
17 pav. Indų etninės grupės spėjimo sumaišties matrica

Paskutinė tiriama etninės grupės nustatymo modelio klasė buvo indų. Modelio validacijos spėjimo rezultatai sumaišties matricos formatu yra pavaizduoti 17 pav. Remiantis pavaizduota sumaišties matrica,

buvo gauti tokie modelio rezultatai šiai klasei: tikslumas – 66,279 %, jautrumas – 69,737 %. Šie rezultatai rodo, kad modelis gebėjo atskirti šios klasės egzempliorius geriau nei ispanų kilmės veidus, tačiau tikslumas bei jautrumas atsiliko nuo baltaodžių, juodaodžių bei azijiečių klasių. Atsižvelgiant į šiuos rezultatus, galima daryti prielaidą, kad veido šablone informacijos, kuria remiantis galima nustatyti šią klasę, yra nepakankamai daug, lyginant su kitomis klasėmis.

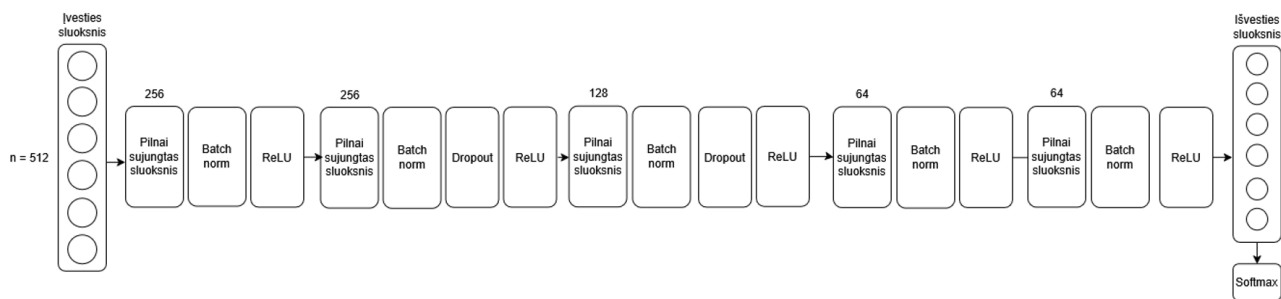
Amžiaus atributo nustatymui buvo sukurti keli neuroniniai tinklai su architektūra, kuri yra panaši etninės grupės nustatymo modeliams. Esminis modelių, skirtų šio atributo nustatymui, skirtumas – išvesties sluoksnio neuronų skaičius. Šiose architektūrose jis buvo lygus amžiaus grupių skaičiui (žr. 3 pav.).

Pirmasis sukurtas neuroninis tinklas buvo „TemplateAgeFC1Net“, kuris buvo sudarytas iš šešių pilnai sujungtų sluoksnių. Šiame tinkle, kiekviename sluoksnyje yra atliekama imties normalizacija, o tinklo viduryje yra įtrauktas išmetimo sluoksnis. Ši architektūra yra pavaizduota 18 pav.



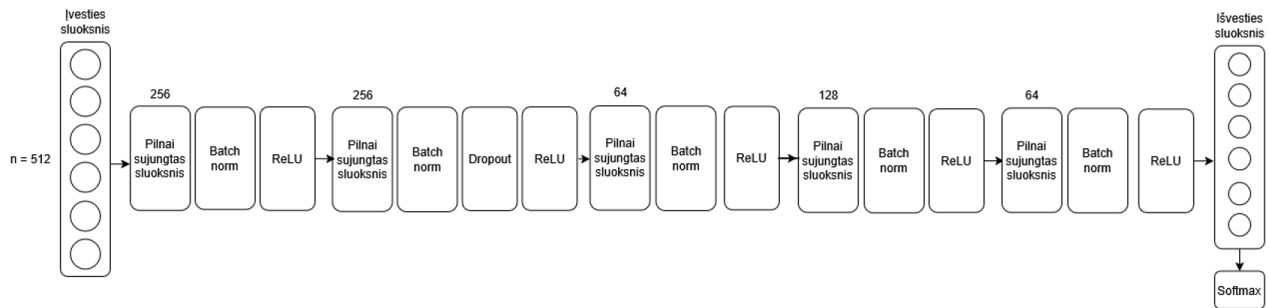
18 pav. „TemplateAgeFC1Net“ tinklo architektūra

Antroji bandyta architektūra, skirta amžiaus atributo nustatymui, buvo „TemplateAgeFC2Net“. Šioje architektūroje neuronų skaičius yra laipsniškai mažėjantis. Jame yra naudojami ankstesniuose tinkluose naudoti normalizacijos ir išmetimo sluoksniai, tačiau skiriasi išvesties sluoksnis. Jis buvo sudarytas iš 6 išvesties neuronų, kurių aktyvacijai buvo naudojama „softmax“ funkcija (žr. 19 pav.).



19 pav. „TemplateAgeFC2Net“ tinklo architektūra

Trečiajai amžiaus spėjimo architektūrai buvo pasirinktas geriausią rezultatą parodęs tinklas etninės grupės atpažinimo uždavinyje. Siekiant pritaikyti jį amžiaus grupės nustatymui, buvo modifikuotas išvesties sluoksnis, kad modelio grąžinami rezultatai atitiktų amžiaus kategorijų skaičių (žr. 20 pav.).



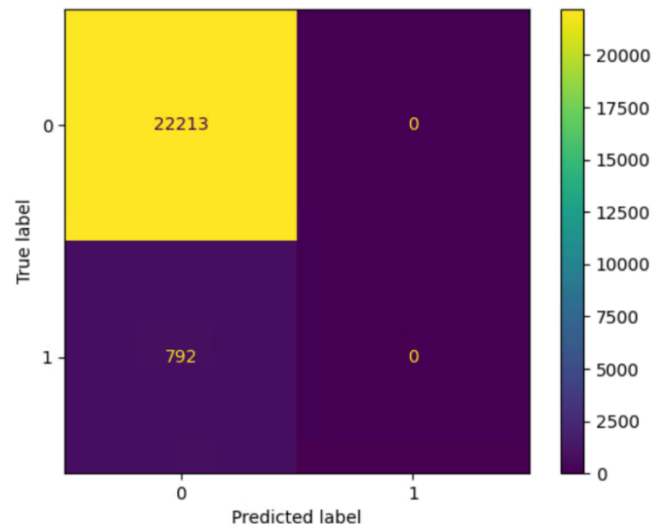
20 pav. „TemplateAgeFC3Net“ tinklo architektūra

Apmokius DNN, skirtus amžiaus atributo atpažinimui, bei atlikus geriausių hiperparametrų paiešką, buvo gauti prasčiausi rezultatai iš visų anksčiau tirtų modelių. Nei vieno modelio tikslumas nepasiekė 70 %, taip pat visų modelių F1 balai buvo žemesni nei modelių validacijos tikslumo įverčiai (žr. 3 lentelė).

Iš apžvelgtų modelių, geriausią rezultatą parodė „TemplateAgeFC1Net“ modelis, kuris pasiekė 65,52 % tikslumą ir jo F1 balas buvo 0,6327. Šis rezultatas yra prastesnis už rezultatus gautus su etninės grupės ir lyties atpažinimo modeliais. Kiti amžiaus nustatymo modelių rezultatai nedaug atsiliko nuo geriausio modelio ir jų tikslumas bei F1 metrika buvo panašūs. Siekiant įvertinti amžiaus nustatymui skirtų modelių efektyvumą, buvo atlikta sumaišties matricių analizė kiekvienai amžiaus grupei. Tai padeda geriau atskleisti ir paaiškinti modelio rezultatus, kadangi modelio bendras tikslumas ar F1 balas gali neparodyti išskirtinai prastų rezultatų nustatant tam tikras klases. Taip pat amžiaus atributo atžvilgiu, duomenų rinkinys nebuvo taip gerai subalansuotas, kaip etninės grupės ar lyties atributai.

3 lentelė. Amžiaus atpažinimo neuroninių tinklų modelių rezultatai

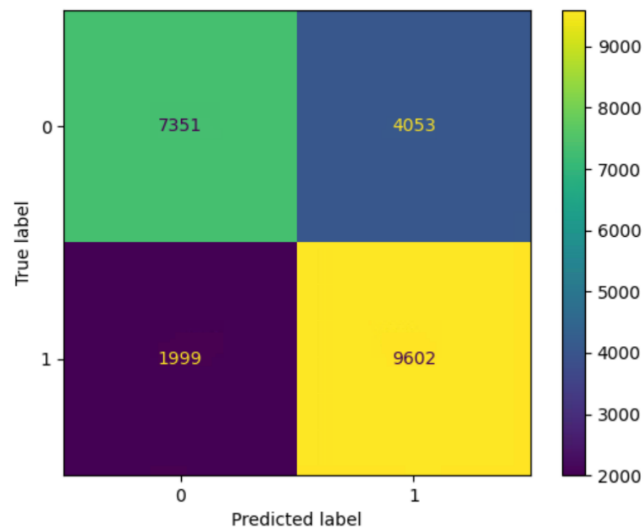
Modelis	Ypatybės	Epocha	LR	BS	Tikslumas (%)	F1 balas
TemplateAgeFC1Net	FC, BN, išmetimas	17	0,01	32	63,57	0,5829
TemplateAgeFC1Net	FC, BN, išmetimas	10	0,01	64	63,28	0,5834
TemplateAgeFC1Net	FC, BN, išmetimas	9	0,01	256	63,43	0,5842
TemplateAgeFC1Net	FC, BN, išmetimas	9	0,01	512	63,45	0,5847
TemplateAgeFC1Net	FC, BN, išmetimas	16	0,001	32	64,46	0,5819
TemplateAgeFC1Net	FC, BN, išmetimas	12	0,001	64	63,95	0,5891
TemplateAgeFC1Net	FC, BN, išmetimas	13	0,001	256	65,52	0,6327
TemplateAgeFC1Net	FC, BN, išmetimas	7	0,001	512	64,96	0,619
TemplateAgeFC1Net	FC, BN, išmetimas	16	0,0001	32	64,96	0,6746
TemplateAgeFC1Net	FC, BN, išmetimas	10	0,0001	64	63,36	0,5826
TemplateAgeFC1Net	FC, BN, išmetimas	11	0,0001	256	64,29	0,6139
TemplateAgeFC1Net	FC, BN, išmetimas	13	0,0001	512	64,67	0,6233
TemplateAgeFC1Net	FC, BN, išmetimas	23	0,00001	32	63,44	0,583
TemplateAgeFC1Net	FC, BN, išmetimas	20	0,00001	64	62,97	0,5806
TemplateAgeFC1Net	FC, BN, išmetimas	24	0,00001	256	62,38	0,5775
TemplateAgeFC1Net	FC, BN, išmetimas	30	0,00001	512	62,5	0,5824
TemplateAgeFC2Net	FC, BN, išmetimas	11	0,01	32	63,54	0,5824
TemplateAgeFC2Net	FC, BN, išmetimas	13	0,01	64	63,74	0,5875
TemplateAgeFC2Net	FC, BN, išmetimas	9	0,01	256	63,2	0,5816
TemplateAgeFC2Net	FC, BN, išmetimas	6	0,01	512	63,27	0,5837
TemplateAgeFC2Net	FC, BN, išmetimas	12	0,001	32	63,48	0,5822
TemplateAgeFC2Net	FC, BN, išmetimas	10	0,001	64	63,19	0,5773
TemplateAgeFC2Net	FC, BN, išmetimas	15	0,001	256	64,9	0,621
TemplateAgeFC2Net	FC, BN, išmetimas	12	0,001	512	63,48	0,5822
TemplateAgeFC2Net	FC, BN, išmetimas	17	0,0001	32	60,62	0,5752
TemplateAgeFC2Net	FC, BN, išmetimas	20	0,0001	64	64,06	0,6176
TemplateAgeFC2Net	FC, BN, išmetimas	7	0,0001	256	63,01	0,6009
TemplateAgeFC2Net	FC, BN, išmetimas	10	0,0001	512	63,88	0,6131
TemplateAgeFC2Net	FC, BN, išmetimas	21	0,00001	32	62,97	0,5803
TemplateAgeFC2Net	FC, BN, išmetimas	20	0,00001	64	62,56	0,5766
TemplateAgeFC2Net	FC, BN, išmetimas	26	0,00001	256	62,74	0,5792
TemplateAgeFC2Net	FC, BN, išmetimas	29	0,00001	512	62,2	0,5993
TemplateAgeFC3Net	FC, BN, išmetimas	16	0,01	32	63,83	0,5885
TemplateAgeFC3Net	FC, BN, išmetimas	13	0,01	64	63,62	0,5844
TemplateAgeFC3Net	FC, BN, išmetimas	9	0,01	256	63,24	0,5831
TemplateAgeFC3Net	FC, BN, išmetimas	8	0,01	512	63,33	0,5842
TemplateAgeFC3Net	FC, BN, išmetimas	18	0,001	32	64,44	0,5837
TemplateAgeFC3Net	FC, BN, išmetimas	10	0,001	64	63,24	0,5827
TemplateAgeFC3Net	FC, BN, išmetimas	6	0,001	256	65,05	0,625
TemplateAgeFC3Net	FC, BN, išmetimas	9	0,001	512	64,86	0,6255
TemplateAgeFC3Net	FC, BN, išmetimas	22	0,0001	32	63,19	0,5814
TemplateAgeFC3Net	FC, BN, išmetimas	18	0,0001	64	64,74	0,6233
TemplateAgeFC3Net	FC, BN, išmetimas	11	0,0001	256	64,62	0,6228
TemplateAgeFC3Net	FC, BN, išmetimas	11	0,0001	512	64,16	0,6161
TemplateAgeFC3Net	FC, BN, išmetimas	21	0,00001	32	63,22	0,5844
TemplateAgeFC3Net	FC, BN, išmetimas	21	0,00001	64	62,86	0,5785
TemplateAgeFC3Net	FC, BN, išmetimas	26	0,00001	256	63,23	0,5982
TemplateAgeFC3Net	FC, BN, išmetimas	26	0,00001	512	60,86	0,5682



21 pav. Pirmos amžiaus grupės spėjimo sumaišties matrica

Atliekant „TemplateAgeFC1Net“ modelio, pasiekusio didžiausią tikslumą, validaciją buvo pastebėta, jog pirmos amžiaus grupės (0-19 metų) spėjimų modelis nepadarė. Vadinasi modelis negebėjo nustatyti nei vieno atvejo, priklausančio šiai klasei. Remiantis šios klasės sumaišties matrica (žr. 21 pav.), modelis apskritai nepadarė nei vieno teigiamo spėjimo, kas reiškia, jog jis visiškai nėra tinkamas šios amžiaus grupės nustatymui.

Iš sumaišties matricos galima matoma, kad modelio tikslumas buvo lygus 0, o jautrumo nepavyko apskaičiuoti, kadangi modelis nepadarė nei vieno teigiamo spėjimo. Šie rezultatai rodo, kad šios amžiaus grupės nustatymui modelis visiškai netinka.

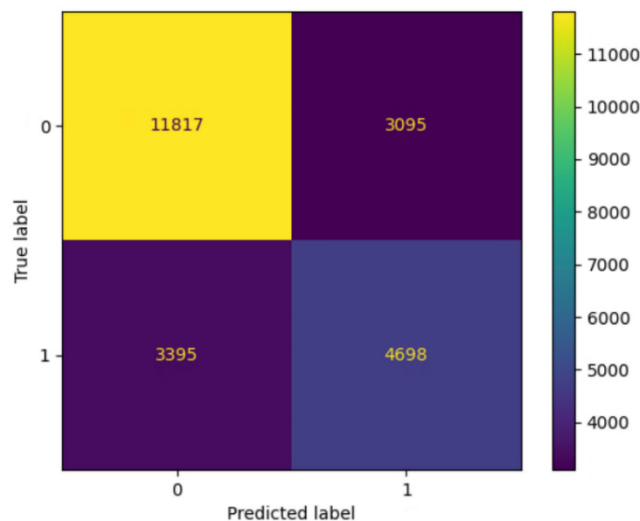


22 pav. Antros amžiaus grupės spėjimo sumaišties matrica

Kita modelio spėjama amžiaus grupė buvo 20-29. Ši grupė turėjo daugiausiai treniravimo ir validacijos duomenų. Tai gali lemti reiškinį, kad modelis bus linkęs „aklai“ teigiamai spėti šią klasę, kadangi atsitiktiniu spėjimo būdu tikimybė atspėti teisingai yra sąlyginai didelė.

Apskaičiavus tikslumą bei jautrumą šiai klasei pagal CM rezultatus (žr. 22 pav.), buvo gauti tokie rezultatai: tikslumas – 82,769 %, jautrumas – 70,319 %. Tai yra sąlyginai aukšti rodikliai, kas įprastai rodytų, jog modelis geba atpažinti veidų šablonus, priklausančius šiai klasei. Tačiau tokie rezultatai gali būti gauti dėl to, kad duomenų rinkinyje dominavo šios amžiaus grupės įrašai. Tikėtina, kad dėl šios priežasties modelis išmoko spėti šią klasę, nes statistiškai tai buvo geriausias pasirinkimas.

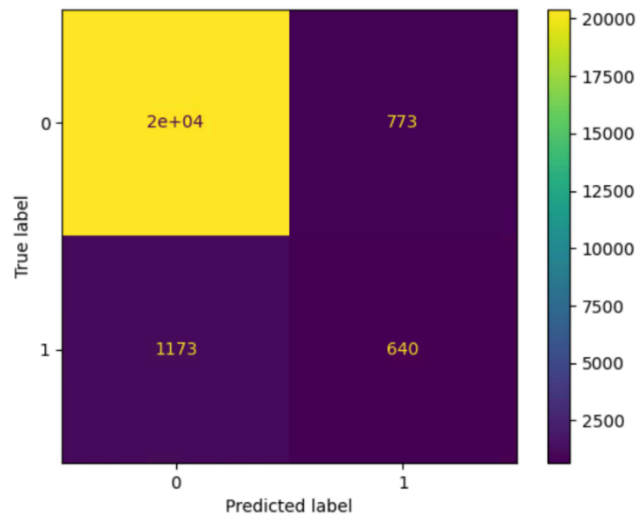
Trečioji modelio spėjama amžiaus grupė buvo 30-39 metų intervale. Šios grupės duomenų buvo šiek tiek mažiau nei 20-29, bet ženkliai daugiau nei likusių klasių. Modelio spėjimų sumaišties matrica šiai klasei pavaizduota 23 pav.



23 pav. Trečios amžiaus grupės spėjimo sumaišties matrica

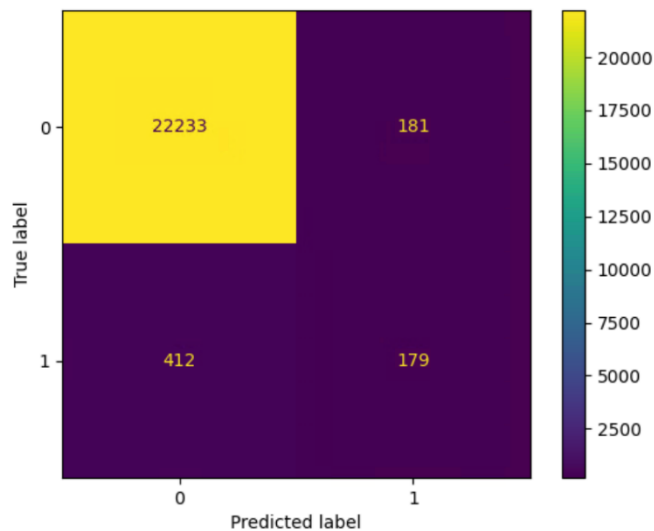
Modelio tikslumas šiai klasei buvo 60,285 %, o jautrumas – 58,05%. Šie rodikliai rodo, kad kaip ir su antrąja amžiaus grupe, modelis buvo linkęs daryti spėjimus šiai klasei. Tikslumas bei jautrumas buvo šiek tiek žemesni nei su antrąja grupe, tačiau atsiliko nedaug. Lyginant šiuos rezultatus galima įžvelgti tam tikrą priklausomybę tarp duomenų kiekio klasėje ir modelio spėjimo rezultatų, kas rodytų, jog modelis tiesiog yra linkęs rinktis didžiausią tikimybę turintį variantą.

Ketvirtoji amžiaus grupė duomenų rinkinyje buvo trečia pagal dydį. Ši grupė, kartu su antrąja ir trečiąja, dominavo duomenų rinkinyje. Kaip šis faktas atsiliepia modelio spėjimams šioje klasėje galima pamatyti 24 pav.



24 pav. Ketvirtos amžiaus grupės spėjimo sumaišties matrica

Modelio tikslumas šios klasės duomenims buvo 35,301 %, o jautrumas – 45,294 %. Ši klasė yra trečia pagal tikslumą, o šios klasės duomenų aibės dydis yra trečias pagal dydį. Kaip ir su ankščiau analizuotomis amžiaus grupėmis, galima įžvelgti sąryšį tarp duomenų aibės dydžio ir modelio tikslumo nustatant konkrečią klasę.

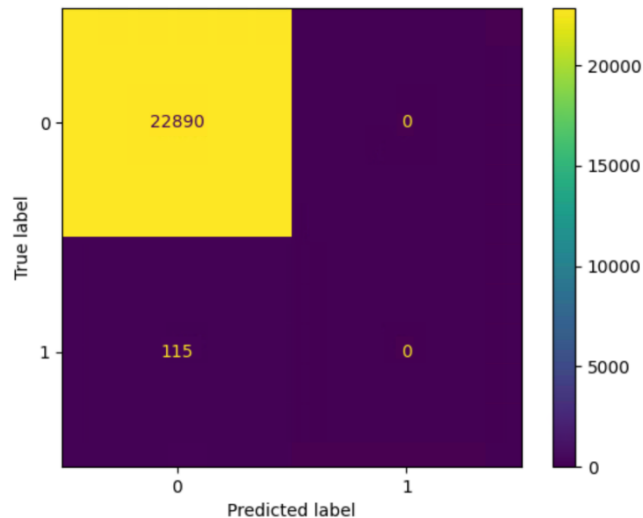


25 pav. Penktos amžiaus grupės spėjimo sumaišties matrica

Sekančiai, penktajai, amžiaus grupei priklausė veidų duomenys, kurie patenka į [50, 59] metų intervalą. Šios amžiaus grupės modelio spėjimų sumaišties matrica pavaizduota 25 pav.

Remiantis šia CM, modelio tikslumas šiai klasei buvo 30,288 %, jautrumas – 49,722 %. Tai yra prastas rezultatas, kas rodo, jog modelis negeba patikimai atpažinti veidų šablonų, priklausančių šiai klasei.

Paskutinė, šeštoji, amžiaus grupė buvo sudaryta iš veidų savybių vektorių, kurie priklausė [60, 69] metų intervalui. Šiai klasei priklausė mažiausias kiekis mokymo duomenų, kas potencialiai turėjo įtakos CM (žr. 26 pav.) matomiems rezultatams.



26 pav. Šeštosios amžiaus grupės spėjimo sumaišties matrica

Šios klasės sumaišties matrica yra labai panaši į pirmosios amžiaus grupės CM. Tikslumas šiai klasei buvo lygus 0, kas rodo, jog modelis nepadarė nei vieno teigiamo spėjimo. Remiantis šiais rezultatais galima teigti, kad šis modelis yra visiškai netinkamas šios amžiaus grupės nustatymui.

Detaliau išanalizavus tiksliausio amžiaus grupės nustatymo modelio rezultatus buvo pastebėta, jog neuroniniai tinklai amžiaus klasifikavimui iš veidų šablonų nėra tinkamas sprendimas. Remiantis sumaišties matricomis galima pastebėti, kad modelio spėjimai ir tikslumas kiekvienai klasei koreliuoja su tam tikrų amžiaus grupių egzempliorių skaičiumi duomenų rinkinyje. Tai reiškia, kad modelis išmoksta atlikti spėjimus grynai remiantis statistika, o ne kažkuriomis duomenų savybėmis.

Viena iš galimų priežasčių, kodėl amžiaus atpažinimo modeliai veikė taip prastai, gali būti veido šablono gavimo principas ir tikslai. Veidų atpažinimo algoritmų, generuojančių tokius šablonus, siekiamybė yra užkoduoti tik tas savybes, kurios yra aktualios žmogaus tapatybės nustatymui. Kadangi žmogaus amžius nėra tapatybės bruožas, tikėtina, kad algoritmai, formuojantys veidų šablonus, šią informaciją ignoruoja ir neišsaugo jos į veido atributų vektorių.

3.2. Tyrimo rezultatai taikant klasikinius ML metodus

Prižiūrimojo mokymo metodai sudaro vieną iš pagrindinių mašininio mokymo šakų, kai modeliai mokomi remiantis sužymėtais duomenimis – kiekvienas įėjimo vektorius turi atitinkamą išvesties žymą arba klasę. Tokiu būdu mokymo metu algoritmas siekia atrasti funkciją, kuri apibrėžia ryšį tarp duomenų

požymiu ir spėjimo reikšmių. Prižiūrimojo mokymo tikslas – išmokyti bendrąsias duomenų pasiskirstymo taisykles, kurios leistų tinkamai prognozuoti naujus, nematytus duomenis. Šiame darbe prižiūrimojo mokymo metodai buvo pasitelkti kaip alternatyva neuroniniams tinklams, siekiant įvertinti, kaip klasikiniai algoritmai (pvz., sprendimų medžiai, atsitiktiniai miškai) veikia su veidų šablonais. Toks tyrimas leidžia ne tik palyginti skirtingų metodų efektyvumą, bet ir įvertinti, ar sudėtingi gilieji neuroniniai tinklai yra būtini šių užduočių sprendimui.

Vienas iš tirtų prižiūrimojo mokymo metodų buvo atsitiktinių miškų klasifikatorius. Šio metodo pagrindinė idėja – apjungti daugelio silpnų klasifikatorių (atskirų sprendimų medžių) sprendimus į vieną stipresnį, naudojant balsavimo principą.

Efektyviausio modelio paieškai buvo atlikta hiperparametrų optimizacija taikant tinklelio paiešką. Šios paieškos metu buvo keičiamas naudojamų medžių skaičius, maksimalus medžių gylis, minimalus pavyzdžių skaičius, reikalingas mazgo padalijimui, minimalus pavyzdžių skaičius lapo mazge bei „bootstrap“ parametras, nustatantis ar reikia naudoti atsitiktinį mėginių pasirinkimą (žr. 4 lentelė). Tyrimo metu iš viso buvo išbandyta 1000 hiperparametrų kombinacijų.

4 lentelė. Lyties atpažinimo atsitiktinių miškų modelio hiperparametrų paieška

Parametras	Bandyti parametrai	Geriausia reikšmė
Medžių skaičius	50, 100, 200, 500	500
Maksimalus medžių gylis	Neribotas, 10, 20, 30, 50	Neribotas
Minimalus pavyzdžių skaičius	2, 5, 10, 20, 50	5
Minimalus pavyzdžių skaičius lapo mazge	1, 2, 4, 8, 16	1
„Bootstrap“	Taip, Ne	Ne

Modelis buvo apmokytas naudojant 80 % visų duomenų, o vertinimas atliktas su likusiais 20 %. Geriausias atsitiktinių miškų modelis pasiekė 89,25 % tikslumą lyties klasifikavime, o bendra F1 metrika siekė 0,89. Stebint kiekvienos klasės rezultatus, pastebėta gana simetriška klasifikavimo kokybė – abiejų klasių tikslumas, jautrumas ir F1 balas išliko beveik vienodi. Tai rodo, kad modelis gerai apdoroja duomenų disbalansą ir nėra šališkas nei vienai klasei. Šie rezultatai rodo, kad atsitiktinių miškų klasifikatorius gali būti veiksmingas alternatyvus metodas savybių klasifikacijai naudojant veidų šablonus.

Lyties klasifikavimo užduočiai taip pat buvo pritaikytas sprendimų medžio klasifikatorius. Šis algoritmas klasifikuoja duomenis priimdamas sprendimus pagal atributų reikšmes, sudarant medžio struktūrą, kur kiekvienas vidinis mazgas atitinka tam tikrą sprendimą (pvz., ar reikšmė yra didesnė už tam tikrą slenkstinę vertę), o lapai žymi prognozuojamas klases.

Šiam modeliui taip pat buvo atlikta geriausių hiperparametrų paieška, naudojant tinklelio paiešką, siekiant užtikrinti optimalų modelio veikimą. Buvo tiriamas maksimalus medžio gylis, minimalus

pavyzdžių skaičius, reikalingas mazgų skaidymui bei maksimalus bruožų skaičius, naudojamas medžio skaidymui (žr. 5 lentelė).

5 lentelė. Lyties atpažinimo sprendimų medžio modelio hiperparametrų paieška

Parametras	Bandyti parametrai	Geriausia reikšmė
Maksimalus medžio gylis	3, 5, 10, Neribotas	5
Minimalus pavyzdžių skaičius, reikalingas mazgų skaidymui	2, 5, 10, 20, 50, 100	10
Maksimalus bruožų skaičius	5, 20, 50, 100, 200, 500, Neribotas	500

Geriausias modelis pasiekė 69,66 % tikslumą, kas rodo, kad šis algoritmas nėra tinkamas lyties klasifikavimui, kadangi jis stipriai atsilieka nuo neuroninių tinklų ar sudėtingesnių prižiūrimento mokymo algoritmų.

Etninės grupės klasifikavimo užduočiai taip pat buvo pritaikytas atsitiktinio miško klasifikatorius. Šio eksperimento metu duomenų rinkinys su visais modeliais buvo padalintas į 80 % duomenų mokymui bei 20 % duomenų testavimui. Hiperparametrų paieška atlikta taikant tinklelio paieškos metodą (žr. 6 lentelė).

6 lentelė. Etninės grupės atsitiktinių miškų modelio hiperparametrų paieška

Parametras	Bandyti parametrai	Geriausia reikšmė
Medžių skaičius	50, 100, 200, 500, 1000	1000
Maksimalus medžių gylis	Neribotas, 10, 20, 30, 50	30
Minimalus pavyzdžių skaičius	2, 5, 10	2
Minimalus pavyzdžių skaičius lapo mazge	1, 2, 4	1
„Bootstrap“	Taip, Ne	Ne

Šis modelis pasiekė bendrą klasifikavimo tikslumą – 69,85 %. Tai rodo, kad šis metodas nedaug atsilieka nuo etninės grupės spėjimui apmokytų neuroninių tinklų.

Kitas etninės grupės klasifikavimui išmėgintas algoritmas buvo sprendimų medžiai. Šis metodas yra patrauktas todėl, nes jo apmokymo laikas yra ženkliai trumpesnis, tačiau įprastai jis yra linkęs persimokyti (angl. *overfitting*), kai medis tampa per gilus. Geriausiam šio modelio radimui buvo atlikta hiperparametrų paieška naudojant tinklelio paieškos metodą (žr. 7 lentelė).

7 lentelė. Etninės grupės sprendimų medžio modelio hiperparametrų paieška

Parametras	Bandyti parametrai	Geriausia reikšmė
Maksimalus medžio gylis	3, 5, 10, Neribotas	5
Minimalus pavyzdžių skaičius, reikalingas mazgų skaidymui	2, 5, 10	2
Maksimalus bruožų skaičius	5, 20	5

Nepaisant atliktos hiperparametrų paieškos, modelio tikslumas siekė vos 38,46 %. Šie rezultatai rodo, kad modelis nesugebėjo išmokyti gerai generalizuoti veidų savybių vektorių bruožų ir susieti jų su tam

etninės grupės klasėmis. Tai reiškia, kad sprendimų medžiai nėra tinkami etninės grupės klasifikacijai iš veidų šablonų, ir kad sudėtingesni metodai, tokie kaip atsitiktiniai miškai ar neuroniniai tinklai pasižymi žymiai geresniu veikimu šioje klasifikacijos užduotyje.

Toliau darbe buvo išbandytas atsitiktinių miškų klasifikavimo metodas amžiaus atributo nustatymui. Atlikus hiperparametrų paiešką, geriausi rezultatai buvo pasiekti naudojant parametrus nurodytus 8 lentelėje. Su šiais parametrais geriausias modelio tikslumas siekė 58,95 % tikslumą, o F1 balas buvo 0,52.

8 lentelė. Amžiaus atsitiktinių miškų modelio hiperparametrų paieška

Parametras	Bandyti parametrai	Geriausia reikšmė
Medžių skaičius	50, 100, 200, 500	200
Maksimalus medžių gylis	Neribotas, 10, 30, 50	30
Minimalus pavyzdžių skaičius	2, 5, 10	2
Minimalus pavyzdžių skaičius lapo mazge	1, 2, 4	1
„Bootstrap“	Taip, Ne	Ne

Gauti geriausio modelio rezultatai rodo, kad šis klasifikatorius nėra tinkamas amžiaus atributo nustatymui, kadangi F1 balas yra arti 0,5, kas rodo, jog modelis negeba teisingai nuspėti amžiaus klasių.

Kitas klasikinis prižiūravimo mokymo algoritmas, naudotas amžiaus grupių klasifikavimui, buvo sprendimų medžio klasifikatorius. Kaip ir kituose eksperimentuose su prižiūravimo mokymo algoritmais, šiame buvo taip pat atliekama hiperparametrų paieška (žr. 9 lentelė).

9 lentelė. Amžiaus sprendimų medžio modelio hiperparametrų paieška

Parametras	Bandyti parametrai	Geriausia reikšmė
Maksimalus medžio gylis	3, 5, 10, 20, 50, 100, Neribotas	5
Minimalus pavyzdžių skaičius, reikalingas mazgų skaidymui	2, 5, 10, 20, 50, 100, Neribotas	2
Maksimalus bruožų skaičius	5, 20, 30, 50, 100, Neribotas	Neribotas

Atlikus mokymą bei geriausių hiperparametrų paiešką, bendras klasifikatoriaus tikslumas siekė vos 50,62 % tikslumą, o F1 balas buvo 0,42. Kaip ir atsitiktinių miškų algoritmas, šis modelis parodė prastą rezultatą ir negebėjimą nustatyti amžiaus atributo iš veidų savybių vektorių. Kaip buvo minėta anksčiau, šių rezultatų priežastis gali būti veido šablono formavimo principas, kadangi veidų aptikimo algoritmų tikslas yra išsaugoti tik su tapatybe susijusius duomenis veido savybių vektoriuje.

3.3. Rezultatų palyginimas

Atliekant skirtingų modelių lyginamąją analizę, siekta įvertinti, kurie metodai geriausiai pritaikomi veido šablonų pagrindu atliekamoms klasifikavimo užduotims. Buvo atlikti eksperimentai tiek su giliojo mokymo neuroniniais tinklais, tiek su tradiciniais prižiūravimo mokymo algoritmais. Kiekvieno metodo veikimas priklausė nuo užduoties pobūdžio – ar tai dvejetainė klasifikacija ar daugiaklasė klasifikacija.

Pastebėta, kad gilieji neuroniniai tinklai pasižymi didesniu lankstumu ir gebėjimu išmokyti sudėtingesnes duomenų struktūras, tačiau reikalauja daugiau bandymų kuriant tinklų architektūrą. Tuo tarpu tradiciniai prižiūrimojo mokymo metodai pasižymi didesniu greičiu ir paprastesniu taikymu, bet gali būti nepakankamai tikslūs klasifikuojant sudėtingesnius atributus.

Lyties atributo klasifikavimo užduotyje eksperimentų metu paaiškėjo modelių veikimo skirtumai. Geriausius rezultatus parodė gilieji neuroniniai tinklai, ypač tie, kurių architektūroje buvo pritaikyta imties normalizacija ir išmetimo sluoksniai. Geriausias neuroninis tinklas sugebėjo pasiekti 93,04 % tikslumą, o prižiūrimojo mokymo algoritmų rezultatai buvo kiek prastesni – geriausias atsitiktinių miškų modelis pasiekė 89,25 % tikslumą (žr. 10 lentelė).

10 lentelė. Lyties klasifikavimo modelių rezultatų palyginimas

Modelis	Tipas	Tikslumas (%)	F1 balas
TemplateSexFC1Net	DNN	92,85	0,9302
TemplateSexFC2Net	DNN	92,66	0,9282
TemplateSexFC3Net	DNN	93,04	0,9312
Sprendimų medis	Klasikinis ML metodas	69,66	0,7
Atsitiktinių miškų modelis	Klasikinis ML metodas	89,25	0,89

Etninės grupės klasifikavimo užduotyje bendras modelių tikslumas buvo prastesnis nei lyties klasifikavime. Geriausią rezultatą pasiekė gilusis neuroninis tinklas „EthnicityFC1Net“, kurio architektūroje buvo panaudoti tiek normalizacijos, tiek išmetimo sluoksniai – tai leido jam pasiekti 72,47 % tikslumą. Kiti giliojo mokymo algoritmai pasiekė labai artimus rezultatus, tačiau jų tikslumas šiek tiek nusileido geriausiam modeliui. Prižiūrimojo mokymo algoritmų rezultatai buvo prastesni – atsitiktinių miškų algoritmas pasiekė 69,85 % tikslumą, o sprendimų medžio algoritmas veikė labai prastai – jo tikslumas nesiekė net 40 % (žr. 11 lentelė). Etninės grupės klasifikavimo modelių rezultatai rodo, kad ši užduotis yra ženkliai sudėtingesnė už lyties klasifikavimą ir reikalauja didesnių neuroninių tinklų ir didesnės parametrų optimizacijos.

11 lentelė. Etninės grupės klasifikavimo modelių rezultatų palyginimas

Modelis	Tipas	Tikslumas (%)	F1 balas
EthnicityFC1Net	DNN	72,47	0,725
EthnicityFC2Net	DNN	72,38	0,7229
EthnicityFC3Net	DNN	71,95	0,7186
Sprendimų medis	Klasikinis ML metodas	38,46	0,26
Atsitiktinių miškų modelis	Klasikinis ML metodas	69,85	0,7

Amžiaus klasifikavimo užduotyje buvo išbandyti keli skirtingi modeliai, kurie parodė nepatenkinamus rezultatus. Geriausią rezultatą parodė DNN – „TemplateAgeFC1Net“ modelis, kurio tikslumas siekė 65,52 %, o F1 balas buvo 0,6327. Eksperimentuose su amžiaus atributo spėjimu išliko tokia

pati tendencija kaip ir su ankščiau tirtomis veido savybėmis. Prižiūrimojo mokymo metodai parodė prastesnius rezultatus nei gilieji neuroniniai tinklai. Geriausias pasiektas rezultatas buvo naudojant atsitiktinių miškų algoritmą – geriausio modelio tikslumas buvo 58,95 %, o F1 balas buvo 0,52 (žr. 12 lentelė). Apibendrinant, nei vienas bandytas modelis neparodė gero rezultato atliekant amžiaus nustatymą iš veidų šablonų. Tai gali būti paaiškinama tuo, kad veidų šablonų generavimo algoritmai yra specialiai kuriami taip, kad sumažintų informacijos, nesusijusios su asmens tapatybę, kiekį veido savybių vektoriuje.

12 lentelė. Amžiaus grupių klasifikavimo modelių rezultatų palyginimas

Modelis	Tipas	Tikslumas (%)	F1 balas
TemplateAgeFC1Net	DNN	65,52	0,6327
TemplateAgeFC2Net	DNN	64,9	0,621
TemplateAgeFC3Net	DNN	65,05	0,625
Sprendimų medis	Klasikinis ML metodas	50,62	0,42
Atsitiktinių miškų modelis	Klasikinis ML metodas	58,95	0,52

Atlikus eksperimentus su visais tirtais metodais ir veidų savybėmis galima pastebėti, kad kai kurių veido savybių klasifikacija iš veidų šablonų yra įgyvendinama, tačiau rezultatai stipriai priklauso nuo prognozuojamo atributo bei pasirinkto spėjimo metodo. Geriausi rezultatai buvo pasiekti nustatant lytį, kur tiek gilieji neuroniniai tinklai, tiek prižiūrimojo mokymo metodai pasiekė aukštą tikslumą. Tai leidžia daryti prielaidą, kad lyties informacija išlieka veidų šablonuose, kadangi ji yra susijusi su gana pastoviomis veido ypatybėmis.

Etninės grupės klasifikavimas buvo sudėtingesnis uždavinys. Geriausias neuroninių tinklų modelis pasiekė 72,47 % tikslumą, kas rodo, jog veido šablonai nebūtinai išsaugo visus bruožus, iš kurių galima spręsti apie žmogaus etninę grupę.

Amžiaus klasifikavimas buvo sudėtingiausias uždavinys – nei gilieji neuroniniai tinklai, nei klasikiniai mašininio mokymo metodai nepasiekė aukštų rezultatų. Geriausias DNN modelis pasiekė 65,52 % tikslumą, tačiau apžvelgus F1 metriką bei sumaišties matricas kiekvienai amžiaus grupei buvo pastebėta, jog modelis negeba patikimai nuspėti šio atributo. Šie rezultatai rodo, kad veido šablono kūrimo metu amžiaus informacija yra sąmoningai išmetama, siekiant padidinti modelių atsparumą žmogaus amžiaus pokyčiui, kas yra aktualu tapatybės identifikacijos uždaviniuose.

IŠVADOS

1. Išanalizavus veidų atpažinimo metodus nustatyta, kad šiuolaikiniai veidų šablonų išgavimo modeliai, tokie kaip „ArcFace“ bei „FaceNet“, naudoja giliuosius neuroninius tinklus, kurie yra apmokyti iš veido vaizdo sugeneruoti kompaktišką vektorinę reprezentaciją. Šie metodai yra plačiai taikomi biometrijoje, tapatybės nustatyme ir kitose praktinėse srityse. Atlikus veidų savybių nustatymo metodų apžvalgą paaiškėjo, jog praktikoje veido savybės dažniausiai yra nustatomos iš pradinio veido vaizdo, o ne veido šablono. Literatūroje egzistuoja mažai darbų, tiriančių veido savybių nustatymą iš veidų savybių vektorių, kadangi veido šablono paskirtis nėra konkrečių veido savybių ar atributų išsaugojimas.
2. Eksperimentinei analizei atlikti buvo paruošta reali veidų šablonų duomenų aibė, susieta su tam tikromis veido savybėmis – lytimi, amžiaus grupe ir etnine grupe. Ruošiant šiuos duomenis buvo atlikti duomenų valymo, balansavimo bei apjungimo veiksmai, kurie užtikrino, kad veido savybių klasių pasiskirstymas būtų kuo įmanoma panašesnis ir duomenys būtų tinkami mokymui ir testavimui. Paruošti duomenys leido atlikti eksperimentus tiek su klasikinais prižiūrimojo mokymo algoritmais, tiek su giliaisiais neuroniniais tinklais.
3. Eksperimentiškai buvo sukurtos 9 giliojo mokymo neuroninių tinklų architektūros bei įvertinti du klasikiniai prižiūrimojo mokymo metodai – sprendimų medis ir atsitiktinių miškų klasifikatorius. Geriausių modelių nustatymui buvo atlikta hiperparametrų paieška. Lengviausiai atpažįstamas atributas buvo lytis – geriausią rezultatą parodė gilusis neuroninis tinklas gebėdamas nustatyti lytį iš veido šablono 93,04 % tikslumu. Eksperimentai su etninės grupės klasifikavimu parodė, jog ši veido savybė yra nustatoma sunkiau – geriausią rezultatą pasiekė gilusis neuroninis tinklas, kurio tikslumas buvo 72,47 %. Amžiaus grupės klasifikavimo rezultatai buvo prasčiausi ir parodė, kad veidų šablone išliekančių duomenų kiekis apie veido amžių yra labai mažas. Geriausio amžiaus grupės spėjimo modelio tikslumas buvo 65,52 %, tačiau atlikus detalesnę rezultatų analizę buvo pastebėta, jog modelis prisitaikė daryti spėjimus pagal statistinį duomenų išsidėstymą duomenų aibėje. Šie rezultatai rodo, kad veido šablonuose išlieka tam tikra informacija apie pradinio vaizdo savybes, tačiau ji turi būti reikšminga veidų atpažinime, kadangi veidų atpažinimo algoritmų siekiamybė yra į veido šabloną sutalpinti tik duomenis, susijusius su tapatybe.

LITERATŪROS SĄRAŠAS

- Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). *ArcFace: Additive Angular Margin Loss for Deep Face Recognition*. <https://github.com/>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. <http://arxiv.org/abs/1512.03385>
- Kotwal, K., & Marcel, S. (2025). *Review of Demographic Bias in Face Recognition*. <http://arxiv.org/abs/2502.02309>
- Križaj, J., Plesh, R. O., Banavar, M., Schuckers, S., & Štruc, V. (2024). Deep Face Decoder: Towards understanding the embedding space of convolutional networks through visual reconstruction of deep face templates. *Engineering Applications of Artificial Intelligence*, 132. <https://doi.org/10.1016/j.engappai.2024.107941>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature* (T. 521, Numeris 7553, p. 436–444). Nature Publishing Group. <https://doi.org/10.1038/nature14539>
- Li, M. (2025). Research and Analysis of Facial Recognition Based on FaceNet, DeepFace, and OpenFace. *ITM Web of Conferences*, 70, 03009. <https://doi.org/10.1051/itmconf/20257003009>
- Li, P., Pei, Y., & Li, J. (2023). A comprehensive survey on design and application of autoencoder in deep learning. *Applied Soft Computing* (T. 138). Elsevier Ltd. <https://doi.org/10.1016/j.asoc.2023.110176>
- Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017). *SphereFace: Deep Hypersphere Embedding for Face Recognition*. <http://arxiv.org/abs/1704.08063>
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). *Swin Transformer: Hierarchical Vision Transformer using Shifted Windows*. <http://arxiv.org/abs/2103.14030>
- Liu, Z., Luo, P., Wang, X., & Tang, X. (2014). *Deep Learning Face Attributes in the Wild*. <http://arxiv.org/abs/1411.7766>
- Mckinney, W. (2011). pandas: a Foundational Python Library for Data Analysis and Statistics. *Python High Performance Science Computer*.
- Meher, J., Allende-Cid, H., & Nordling, T. E. M. (2023). *A survey and classification of face alignment methods based on face models*. <http://arxiv.org/abs/2311.03082>

- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., ... Chintala, S. (2019). *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. <http://arxiv.org/abs/1912.01703>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., & Louppe, G. (2012). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12.
- Qin, L., Wang, M., Deng, C., Wang, K., Chen, X., Hu, J., & Deng, W. (2023). *SwinFace: A Multi-task Transformer for Face Recognition, Expression Recognition, Age Estimation and Attribute Estimation*. <https://doi.org/10.1109/TCSVT.2023.3304724>
- Ranjan, R., Sankaranarayanan, S., Castillo, C. D., & Chellappa, R. (2016). *An All-In-One Convolutional Neural Network for Face Analysis*. <http://arxiv.org/abs/1611.00851>
- Ren, M., Zhu, Y., Wang, Y., Huang, Y., & Sun, Z. (2024). Understanding Deep Face Representation via Attribute Recovery. *IEEE Transactions on Information Forensics and Security*, 19, 6949–6961. <https://doi.org/10.1109/TIFS.2024.3424291>
- S, P., & S, S. (2021). *Security Monitoring System Using FaceNet For Wireless Sensor Network*. <http://arxiv.org/abs/2112.01305>
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). *FaceNet: A Unified Embedding for Face Recognition and Clustering*. <https://doi.org/10.1109/CVPR.2015.7298682>
- Steels, L., & De Mantaras, R. L. (2018). The Barcelona declaration for the proper development and usage of artificial intelligence in Europe. *AI Communications*, 31(6), 485–494. <https://doi.org/10.3233/AIC-180607>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2014). *Going Deeper with Convolutions*. <http://arxiv.org/abs/1409.4842>
- Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 1701–1708. <https://doi.org/10.1109/CVPR.2014.220>

- Terhörst, P., Fährmann, D., Damer, N., Kirchbuchner, F., & Kuijper, A. (2020). *Beyond Identity: What Information Is Stored in Biometric Face Templates?* <http://arxiv.org/abs/2009.09918>
- Terhörst, P., Kolf, J. N., Huber, M., Kirchbuchner, F., Damer, N., Morales, A., Fierrez, J., & Kuijper, A. (2021). *A Comprehensive Study on Face Recognition Biases Beyond Demographics.* <http://arxiv.org/abs/2103.01592>
- van der Walt, S., Colbert, S., & Varoquaux, G. (2011). The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering*, 13, 22–30. <https://doi.org/10.1109/MCSE.2011.37>
- Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., & Liu, W. (2018). *CosFace: Large Margin Cosine Loss for Deep Face Recognition.* <http://arxiv.org/abs/1801.09414>
- Wang, M., Deng, W., Hu, J., & Su, S. (2024). *Marginal Debiased Network for Fair Visual Recognition.* <http://arxiv.org/abs/2401.02150>
- Wu, S., Lin, X., & Yao, T. (2022). Research on Face Recognition Technology of Subway Automatic Ticketing System based on Neural Network and Deep Learning. *IJACSA) International Journal of Advanced Computer Science and Applications* (T. 13, Numeris 12). www.ijacsa.thesai.org
- Xin, Y., Zhou, Y., & Jiang, J. (2024). *Self-Attentioned ArcFace towards Improved Efficiency for Deep Face Recognition.* <https://ssrn.com/abstract=4898101>