

Dirbtinio intelekto metodų tyrimas astrofizikinių objektų klasifikavimui ir/arba svarbių savybių nustatymui

Doktorantas: Tomas Mūžas

Darbo vadovas: prof. dr. Tadas Meškauskas

Darbo konsultantas: dr. Andrius Vytautas Misiukas Misiūnas

Doktorantūros studijų laikotarpis: 2022 – 2026 m.

Studijų planas ir jo vykdymo suvestinė

Metai	Egzaminai		Dalyvavimas konferencijose				Publikacijos		
			Tarptautinės		Nacionalinės		Su citav. rodikl.		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būsena
I (2022/2023)	2	2	-	1	-	-	-	-	-
II (2023/2024)	2	2	-	-	-	-	-	-	-
III (2024/2025)	-	-	1	-	-	-	1	-	-
IV (2025/2026)	-	-	1	1 (10-18)	-	-	1	-	2 straipsniai recenzuojami
Iš viso	4	4	2	2	-	-	2	0	-

2025/2026 m.m. planas ir jo vykdymas

Egzaminai

Planas	Įvykdyta	Būklė
Vaizdų ir duomenų analizė (FF) (2023-06-14)	2023-06-14	Išlaikyta
Informatikos ir informatikos inžinerijos tyrimo metodai ir metodika (2023-06-22)	2023-06-22	Išlaikyta
Fundamentalieji informatikos ir informatikos inžinerijos metodai (2024-01-24)	2024-01-24	Išlaikyta
Gilieji neuroniniai tinklai (2024-06-11)	2024-06-11	Išlaikyta

2025/2026 m.m. planas ir jo vykdymas

Straipsniai

Šiuo metu perengti ir recenzuojami 2 straipsniai:

“Large-Scale Performance Evaluation of Galaxy Image Classification Models Trained on Small Datasets” – *New Astronomy*

“Evaluation of Domain-Specific Convolutional Neural Network Layers for Galaxy Classification Tasks” – *Nonlinear Analysis: Modelling and Control*

Tyrimo objektas

- Astrofizikinių objektų klasifikavimas ir/ar savybių nustatymas
- profesionalų parengti arba savanorių balsavimu grįsti duomenų rinkiniai
- antžeminių ir kosmoso teleskopų nuotraukos

Tyrimo tikslas

Pasiūlyti naują arba patobulintą astrofizikinių objektų klasifikavimo modelį, paremtą dirbtinio intelekto metodais

Tyrimo uždaviniai

1. Sukurti astrofizikinių objektų duomenų analizės metodiką, apibrėžti kriterijus objektų klasifikavimo ir/ar savybių patikimumui įvertinti. Taip pat parengti metodiką sintetinių duomenų generavimui.
2. Atrinkti patikimų duomenų rinkinius, skirtus modelių apmokymui ir validavimui.
3. Apmokyti skirtingus dirbtinio intelekto modelius naudojant parengtą apmokymo duomenų rinkinį. Skaičiavimus atlikti naudojant paskirstyto skaičiavimo resursus bei grafinius procesorius.
4. Išskirti efektyviausius modelius bei jų parametrus, optimizuoti jų taikymą.

Nagrinėjami galaktikų tipai

Spiral (barred)



Spiral (no bar)



Elliptical



Edge-on



Duomenų rinkiniai

4 klasės

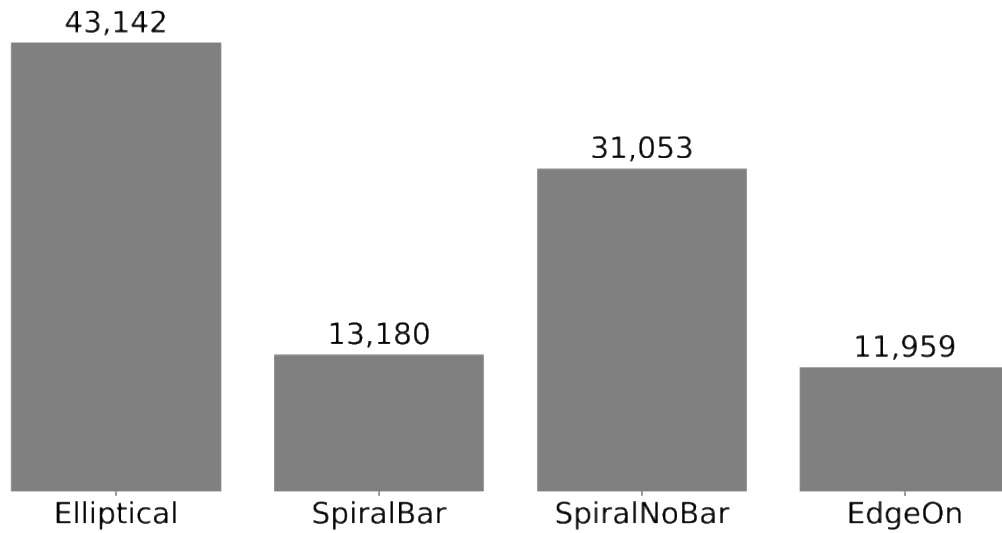
- Buvo siekiama surinkti kuo didesnį patikimai klasifikuotų galaktikų rinkinį
- Buvo naudojami 5 savanorių balsavimu grįsti Galaxy Zoo (GZ) duomenų rinkiniai – GZ1, GZ2, GZ CANDELS, GZ Hubble, GZ DECaLS.
- Paliekamos tik tos unikalios galaktikos, dėl kurių klasės sutaria ne mažiau nei 80% savanorių, ir klasė sutampa visuose duomenų rinkiniuose, kur yra ta galaktika.
- Galutinis galaktikų kiekis – 99,335

11 klasių

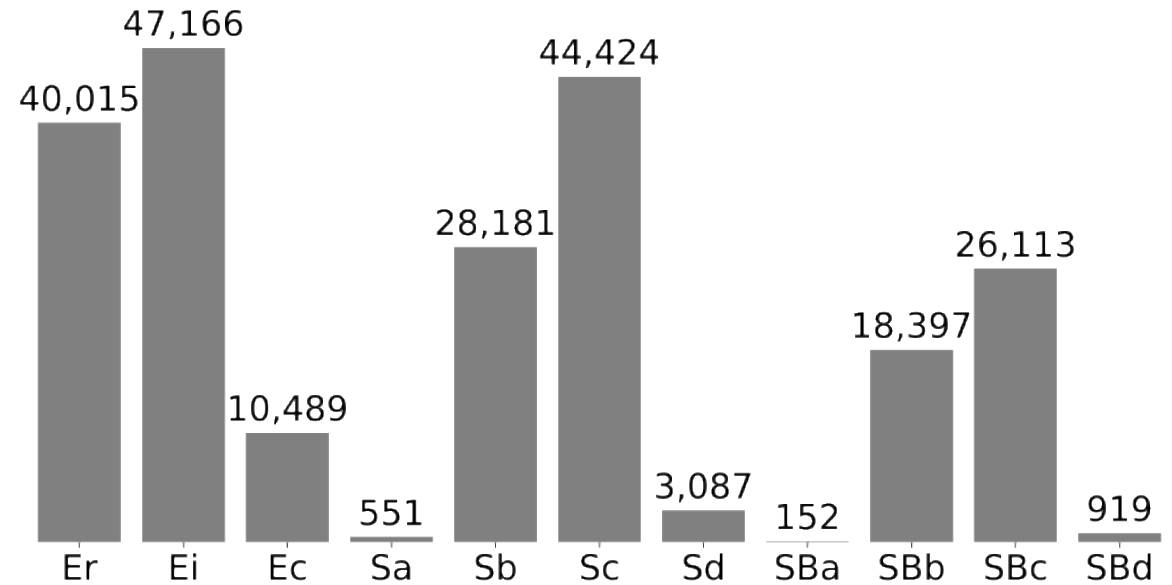
- Naudotas GZ2 duomenų rinkinys, kad rezultatus galima būtų palyginti su Barchi et al.
- Griežtas 4 klasių balsavimo patikimumo filtravimas nebuvo taikomas
- Galutinis galaktikų kiekis - 219,494

Duomenų rinkiniai

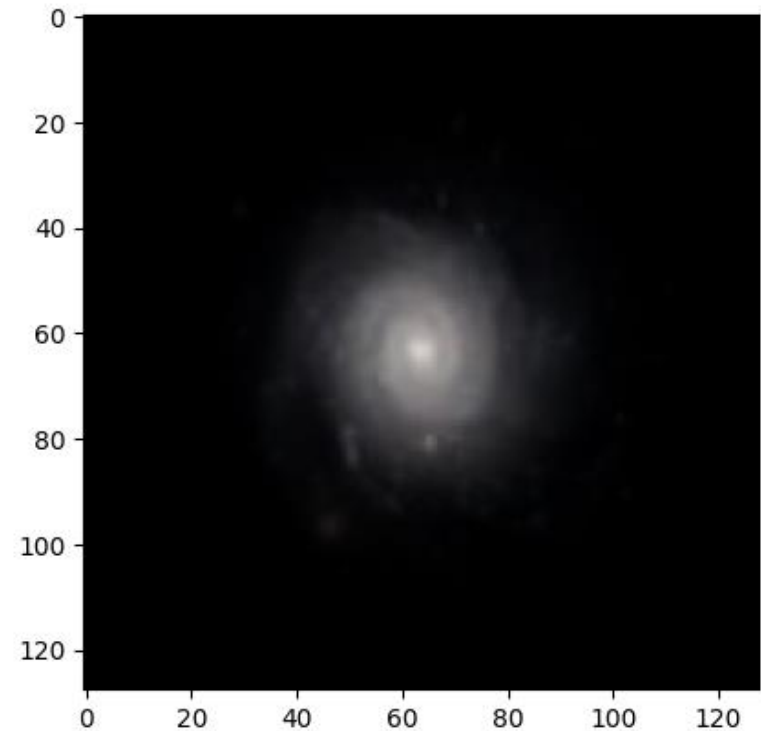
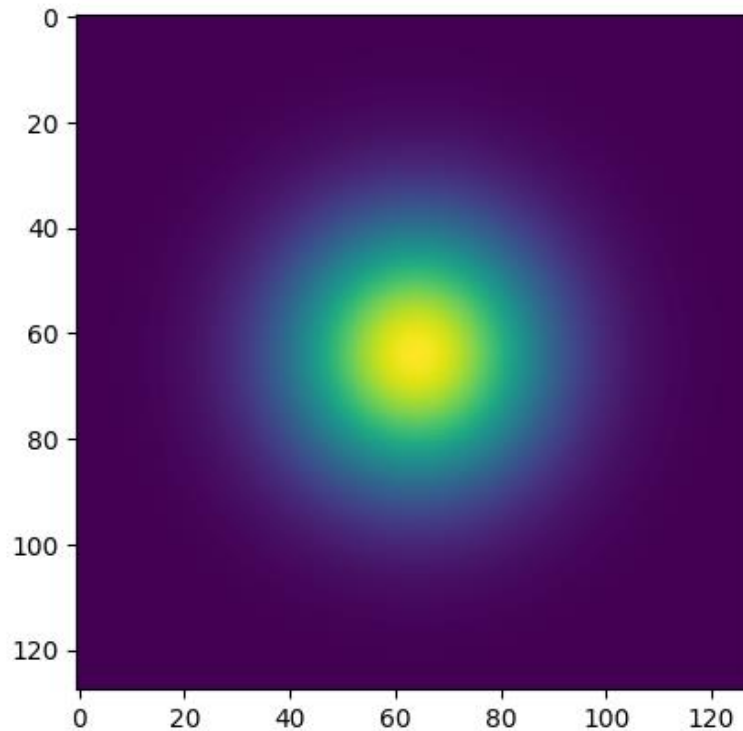
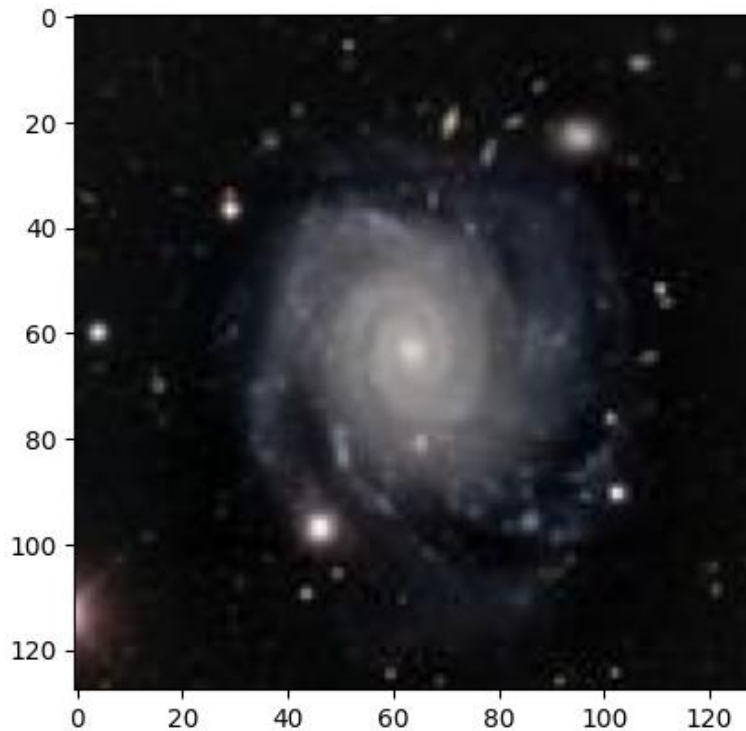
4 klasės



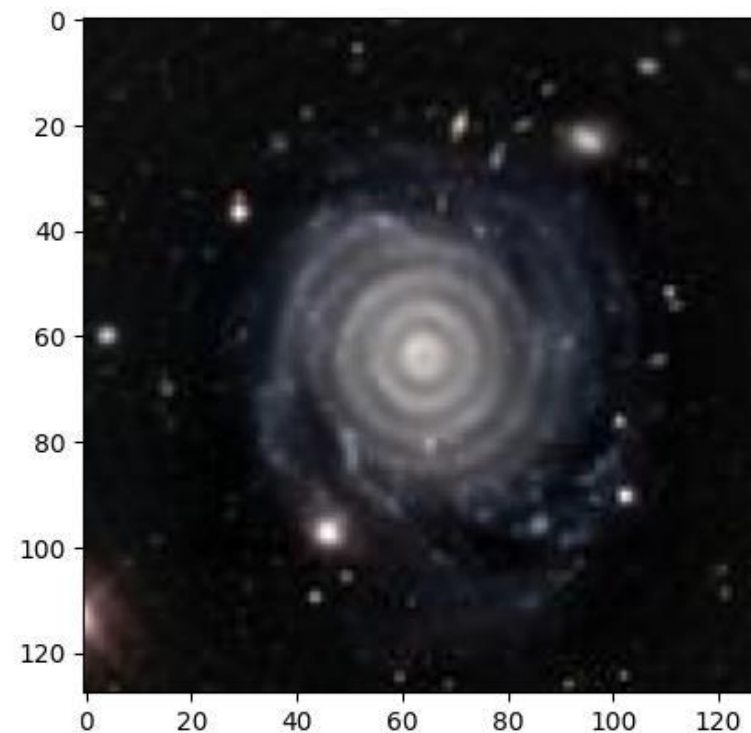
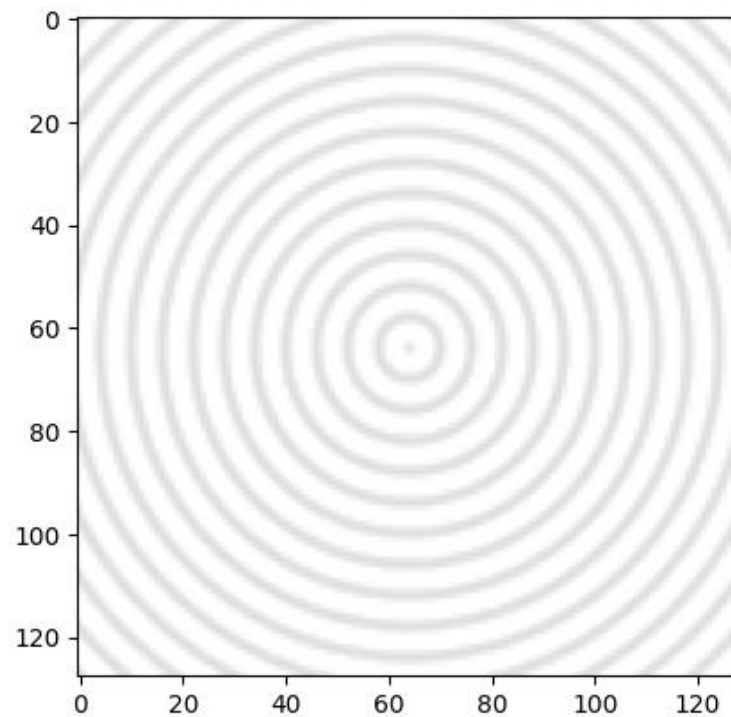
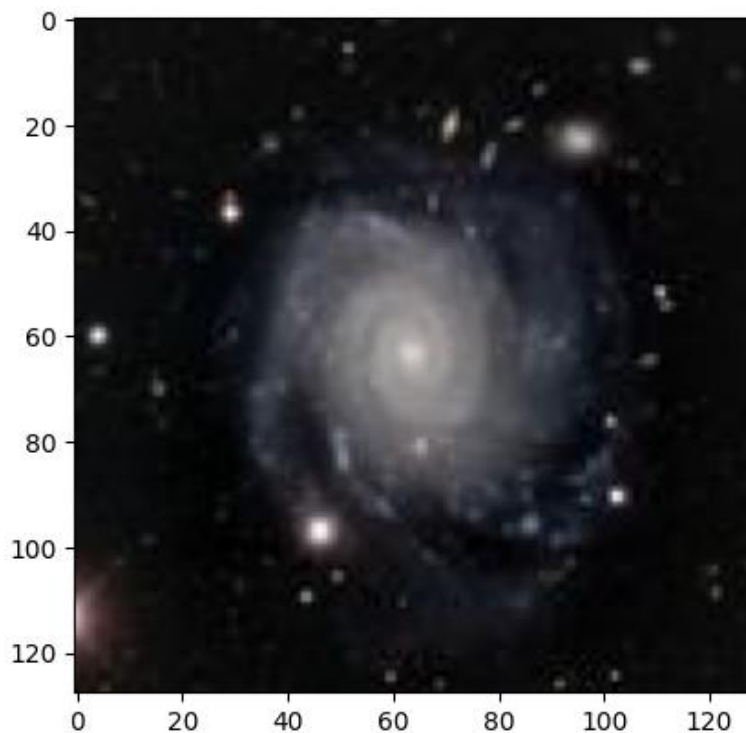
11 klasių



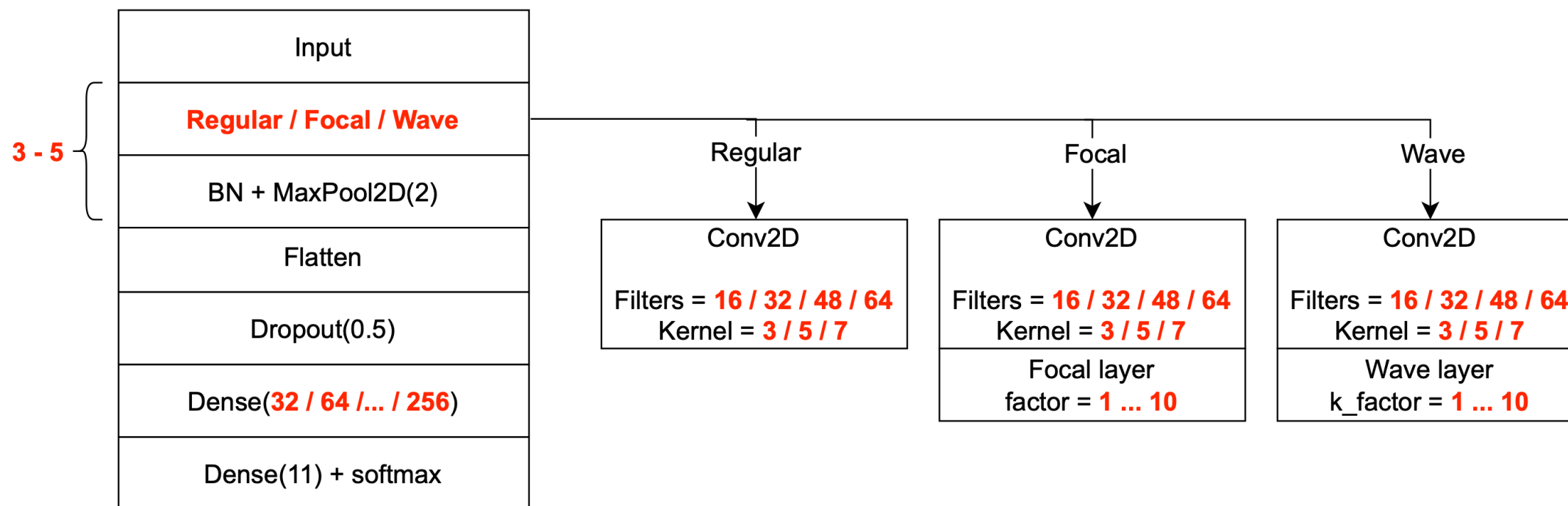
Nauji sluoksniai – Focal conv.



Nauji sluoksniai – Wave conv.



Naujų architektūrų paieška



Geriausi modeliai – 4 klasės

Model	Convolutions	Dense
Best Loss	Focal (64/7/8), Regular (64/7), Regular (48/3), Wave (64/5/5)	224 / 224 / 160
Top 10 Loss	Focal (64/7/8), Focal (64/5/5), Regular (32/7), Wave (3/5/3)	32 / 160 / 160
Best F1	Focal (16/7/7), Focal (48/7), Focal (32/7/6)	224 / 128 / 64
Top 10 F1	Focal (16/7/5), Focal (48/7/3), Focal (32/7/6)	192

Geriausi modeliai – 11 klasių

Model	Convolutions	Dense
Best Loss	Focal (48/3/10), Regular (32/7), Regular (48/7), Regular (32/3)	224 / 192
Top 10 Loss	Focal (48/7/10), Focal (32/7/7), Focal (32/5/8), Focal (32/5/6)	128 / 96 / 64
Best F1	Focal (16/5/6), Focal (64/3/9), Wave (48/7/2)	192 / 160
Top 10 F1	Focal (16/7/6), Focal (32/5/6), Focal (48/7/5)	128 / 160

Rezultatai – 4 klasės

Model	F1 (σ)	Accuracy (σ)	Precision (σ)	Recall (σ)
Cheng	82.89 (0.44)	83.85 (0.27)	83.71 (0.66)	82.37 (0.62)
Cavanagh	85.86 (0.63)	86.09 (0.55)	86.98 (1.14)	85.00 (0.71)
Best Loss	87.36 (0.57)	87.37 (0.52)	87.88 (1.03)	86.99 (0.51)
Top 10 Loss	87.46 (0.50)	87.49 (0.42)	88.08 (0.74)	86.99 (0.57)
Best F1	87.09 (0.47)	87.19 (0.43)	88.10 (0.58)	86.27 (0.69)
Top 10 F1	87.46 (0.50)	87.43 (0.53)	88.12 (0.78)	86.92 (0.46)

Be augmentacijų

Model	F1 (σ)	Accuracy (σ)	Precision (σ)	Recall (σ)
Cheng	88.16 (0.40)	88.20 (0.38)	88.93 (0.32)	87.54 (0.53)
Cavanagh	89.25 (0.20)	89.02 (0.21)	89.92 (0.35)	88.69 (0.31)
Best Loss	89.04 (0.48)	88.82 (0.46)	89.66 (0.50)	88.52 (0.47)
Top 10 Loss	88.55 (0.35)	88.42 (0.34)	89.23 (0.59)	88.01 (0.37)
Best F1	88.92 (0.41)	88.74 (0.35)	89.63 (0.34)	88.33 (0.51)
Top 10 F1	88.92 (0.50)	88.74 (0.50)	89.64 (0.57)	88.32 (0.54)

Su augmentacijomis

Rezultatai – 11 klasių

Model	F1 (σ)	Accuracy (σ)	Precision (σ)	Recall (σ)
Cheng	40.81 (1.05)	60.71 (0.50)	45.74 (3.17)	40.02 (1.14)
Cavanagh	42.68 (0.88)	61.80 (0.42)	46.59 (2.31)	42.09 (1.08)
Best Loss	45.52 (1.06)	63.79 (0.78)	48.65 (3.46)	44.76 (1.25)
Top 10 Loss	44.63 (0.80)	63.43 (0.40)	47.88 (1.45)	43.63 (0.96)
Best F1	45.08 (0.80)	63.48 (0.30)	49.07 (1.17)	43.91 (1.06)
Top 10 F1	4.85 (0.84)	63.35 (0.33)	49.70 (1.30)	43.71 (0.90)

Be augmentacijų

Model	F1 (σ)	Accuracy (σ)	Precision (σ)	Recall (σ)
Cheng	46.06 (1.05)	64.78 (0.18)	50.40 (1.56)	44.71 (1.11)
Cavanagh	46.96 (0.67)	65.29 (0.33)	51.48 (2.07)	45.81 (0.86)
Best Loss	46.04 (1.02)	64.65 (0.64)	49.21 (1.09)	44.83 (1.13)
Top 10 Loss	44.70 (1.28)	63.99 (0.77)	49.93 (2.61)	43.30 (1.23)
Best F1	46.04 (0.32)	64.83 (0.23)	51.20 (2.54)	44.65 (0.35)
Top 10 F1	44.88 (1.18)	64.25 (0.55)	50.21 (1.91)	43.59 (1.11)

Su augmentacijomis

Rezultatai – Ansambliai

Augmentations	F1 (σ)	Accuracy (σ)	Precision (σ)	Recall (σ)
No augmentations	88.75 (0.30)	88.65 (0.32)	89.51 (0.58)	88.11 (0.29)
<i>Best single</i>	<i>87.46 (0.50)</i>	<i>87.49 (0.42)</i>	<i>88.12 (0.78)</i>	<i>86.99 (0.51)</i>
Augmented	89.68 (0.38)	89.44 (0.38)	90.43 (0.41)	89.04 (0.36)
<i>Best single</i>	<i>89.25 (0.20)</i>	<i>89.02 (0.21)</i>	<i>89.92 (0.35)</i>	<i>88.69 (0.31)</i>

4 klasės

Augmentations	F1 (σ)	Accuracy (σ)	Precision (σ)	Recall (σ)
No augmentations	46.59 (0.53)	65.07 (0.31)	50.87 (2.70)	45.29 (0.56)
<i>Best single</i>	<i>45.52 (1.06)</i>	<i>63.79 (0.78)</i>	<i>49.70 (1.30)</i>	<i>44.76 (1.25)</i>
Augmented	46.60 (0.53)	65.46 (0.31)	51.40 (2.38)	45.12 (0.55)
<i>Best single</i>	<i>46.96 (0.67)</i>	<i>65.29 (0.33)</i>	<i>51.48 (2.07)</i>	<i>45.81 (0.86)</i>

11 klasių

Išvados

- Modeliai, apmokyti be augmentacijų ir naudojantys mūsų pasiūlytus sluoksnius rodo statistiškai reikšmingą metrikų pagerėjimą
- Augmentacijų įtaka yra mažesnė modeliams, kurie naudoja mūsų pasiūlytus sluoksnius, tačiau modeliai pasiekia panašius rezultatus lyginant su kitais. Tai reiškia, kad nauji sluoksniai gali sumažinti augmentacijų poreikį modelių apmokymo metu.
- Ansamblis iš modelių be augmentacijų pasiekia panašius rezultatus kaip modeliai su augmentacijomis, o ansamblis iš modelių su augmentacijomis pasiekia tokius pačius ar net geresnius rezultatus nei geriausias modelis.
- Naujai pasiūlytas Focal konvoliucijos sluoksnis yra dažniau pasirenkamas nei standartinės arba Wave konvoliucijos sluoksniai.

Numatomi darbai

- Tęsti tyrimus su naujais konvoliuciniais sluoksniais
- Adaptuoti sluoksnius transformerių architektūroms
- Rengti disertaciją

Ačiū už dėmesį!