

Mašininio mokymosi metodai patikimam galaktikų nuotraukų klasifikavimui

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	1	Publikuota
IV (2025/2026)	-	-	1	1	-	-	1	1 (+ 1)	Publikuota, priimta
Iš viso	4	4	2	2	-	-	2	2	-

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

2026 m.m. planas ir jo vykdymas

Publikuoti straipsniai:

T. Mūžas, A. V. Misiukas Misiūnas, T. Meškauskas

Evaluation of Domain-Specific Convolutional Neural Network Layers for Galaxy Classification Tasks

DOI: 10.15388/namc.2026.31.45923

Nonlinear Analysis: Modelling and Control, Vol. 31, p. 1-18

T. Mūžas, A. V. Misiukas Misiūnas, T. Meškauskas

Large-scale performance evaluation of galaxy image classification models trained on small datasets

DOI: 10.15388/namc.2026.31.47651

Nonlinear Analysis: Modelling and Control, Vol. 31, p. 1–22.

2026 m.m. planas ir jo vykdymas

Priimtas straipsnis:

T. Mūžas, A. V. Misiukas Misiūnas, T. Meškauskas

Ablation Study of Domain-Specific Focal Layer for Galaxy Image Classification with Convolutional Neural Networks.

Baltic Journal of Modern Computing.

Tyrimo objektas

Savanorių klasifikuotos skaitmeninės galaktikų nuotraukos, gautos naudojant antžeminius bei kosminius teleskopus.

Tyrimo tikslas

Ištirti ir patobulinti galaktikų morfologinį klasifikavimą naudojant konvoliucinius neuroninius tinklus ir didelės apimties duomenų rinkinius

Tyrimo uždaviniai

1. Sukurti galaktikų duomenų analizės metodologiją siekiant nustatyti patikimai suklasifikuotas galaktikas Galaxy Zoo duomenų rinkiniuose.
2. Sukurti didelės apimties duomenų rinkinius, sudarytus tik iš patikimai suklasifikuotų galaktikų.
3. Sukurti duomenų augmentacijos metodologiją siekiant dar labiau praplėsti prieinamų duomenų kiekį.
4. Ištirti šiuo metu galaktikų klasifikavimui naudojamas CNN architektūras.
5. Pagerinti egzistuojančias CNN architektūras siekiant tikslesnio ir patikimo galaktikų klasifikavimo.
6. Įvertinti, kaip tiksliai CNN modeliai geba suklasifikuoti didelės apimties duomenų rinkinius.
7. Atlikti reikiamus eksperimentus rezultatų verifikavimui.

Nagrinėjami galaktikų tipai

Spiralinė



Eiptinė

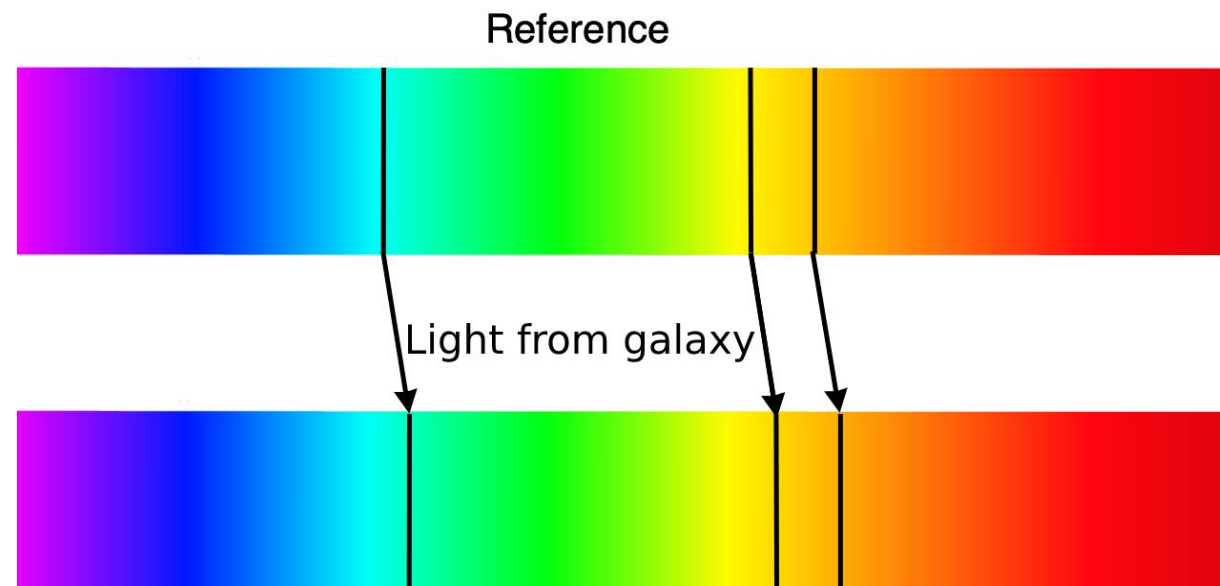


Raudonasis poslinkis (*redshift*, z)

Tai fizikinis dydis, apibūdinantis, kiek šviesa iš galaktikos pasislenka į raudoną šviesos spektro pusę, jai keliaujant kosmosu.

Kuo didesnis z , tuo toliau yra galaktika.

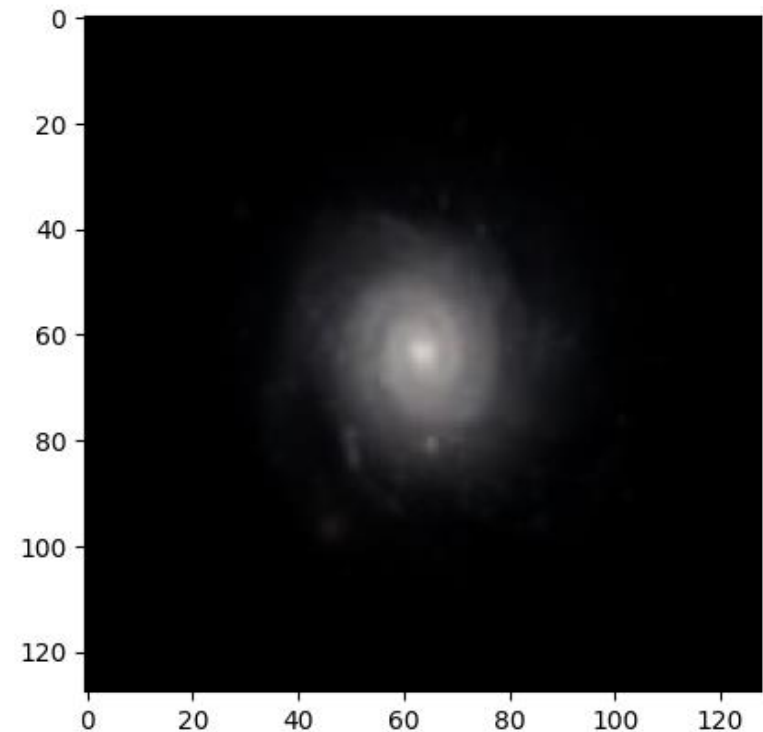
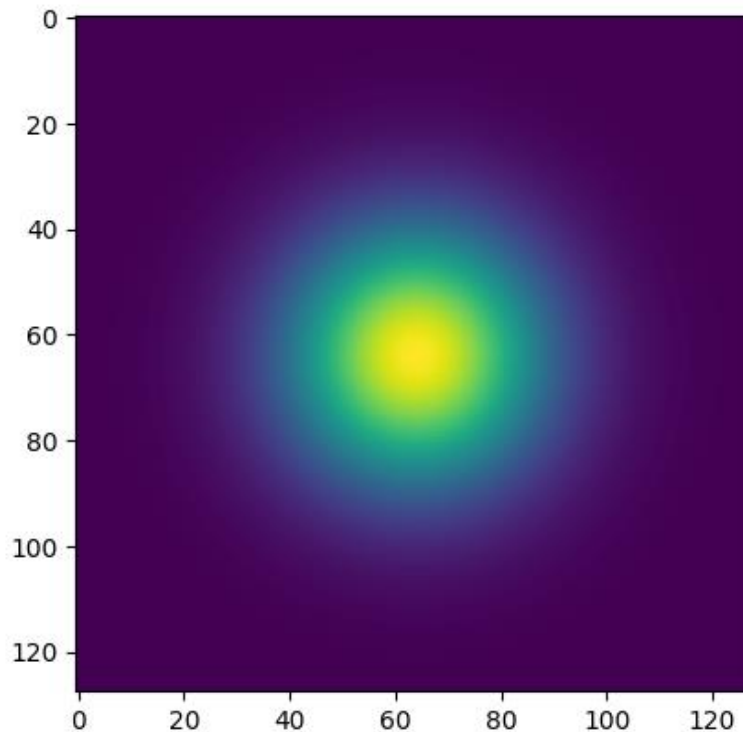
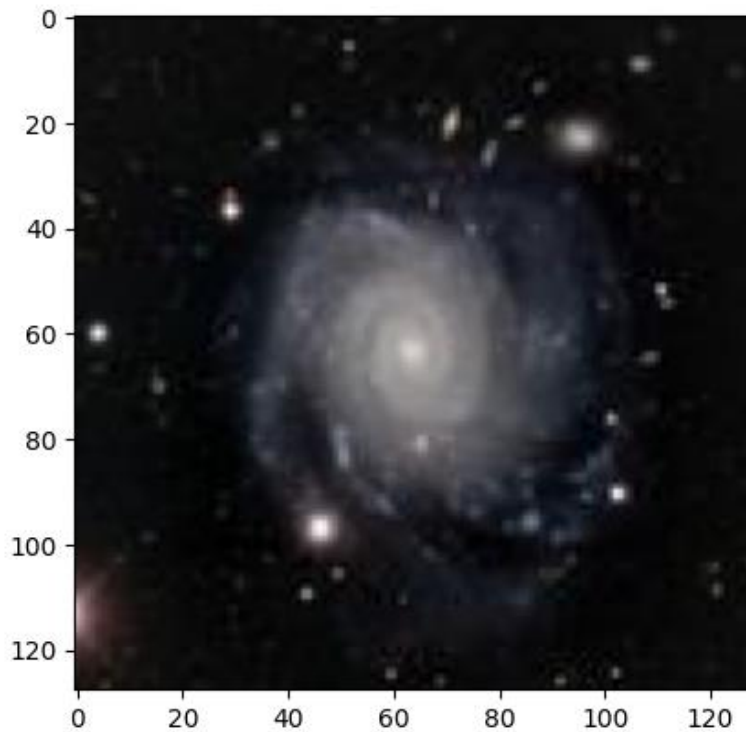
$$z = \frac{\lambda_0 - \lambda}{\lambda}$$



Patikimai klasifikuotų galaktikų duomenų rinkinys

- Naudojami 5 savanorių balsavimu grįsti Galaxy Zoo (GZ) duomenų rinkiniai – GZ1, GZ2, GZ CANDELS, GZ Hubble, GZ DECaLS.
- 418,757 unikalių, patikimai suklasifikuotų galaktikų.
- Padalintas į 5 rinkinius pagal teleskopą ir z reikšmę:
 - SDSS ($z \leq 0.25$)
 - DES ($z \leq 0.15$)
 - Hubble 1 ($0 < z < 0.3$)
 - *Hubble 2* ($0.3 \leq z < 1$)
 - *Hubble 3* ($1 \leq z < 4$)

Focal conv sluoksnis

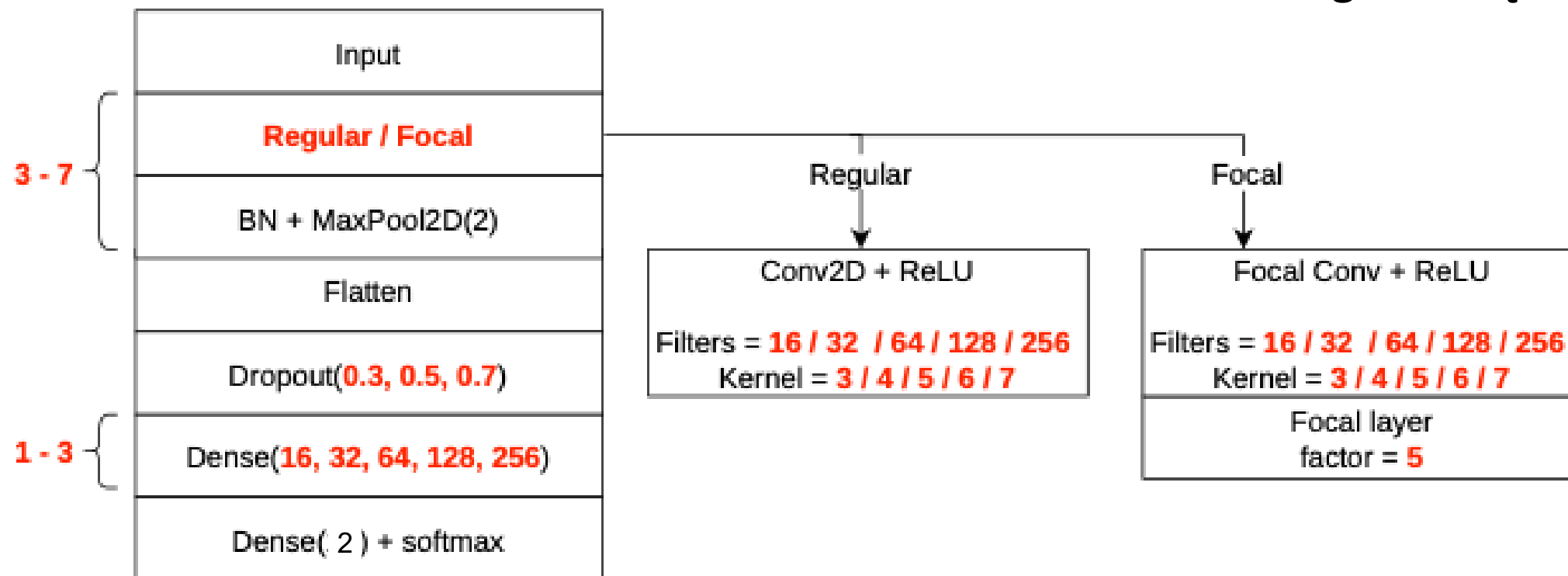


I etapas

Kombinacijų: 11,250

Išbandyta: 1,000 atsitiktinių

Naudota galaktikų: 10,000

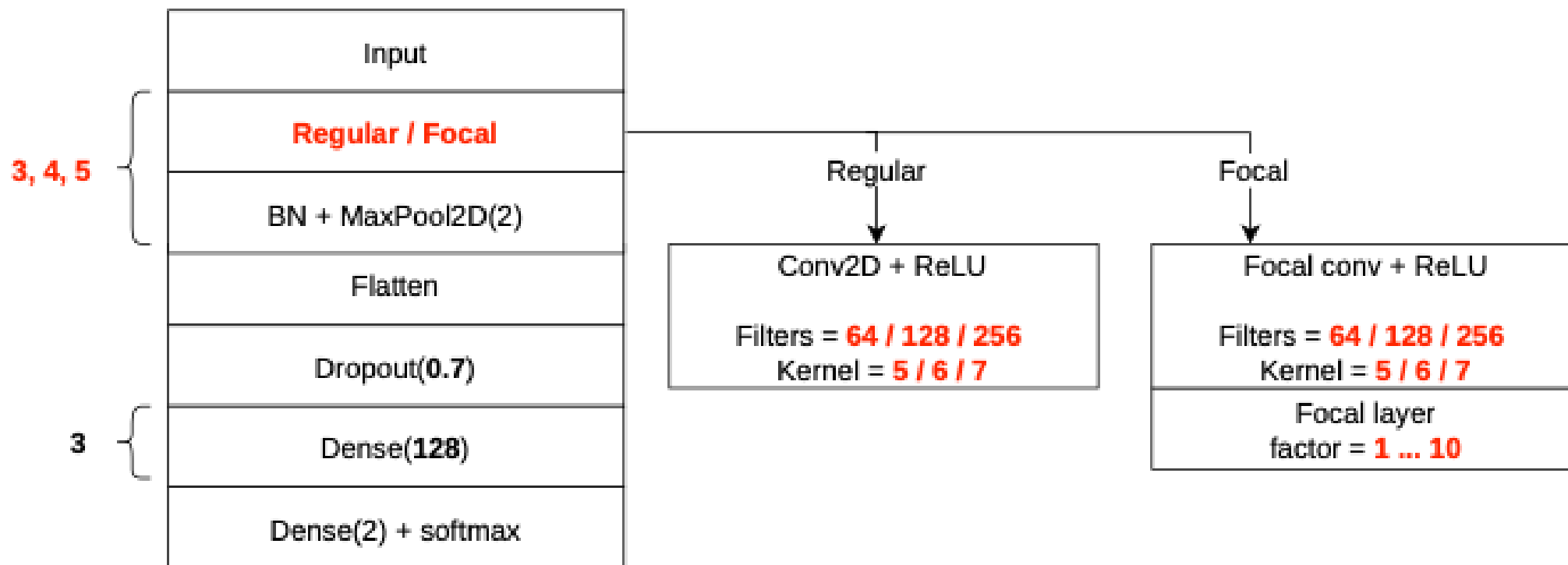


II etapas

Kombinacijų: 297

Išbandyta: 297

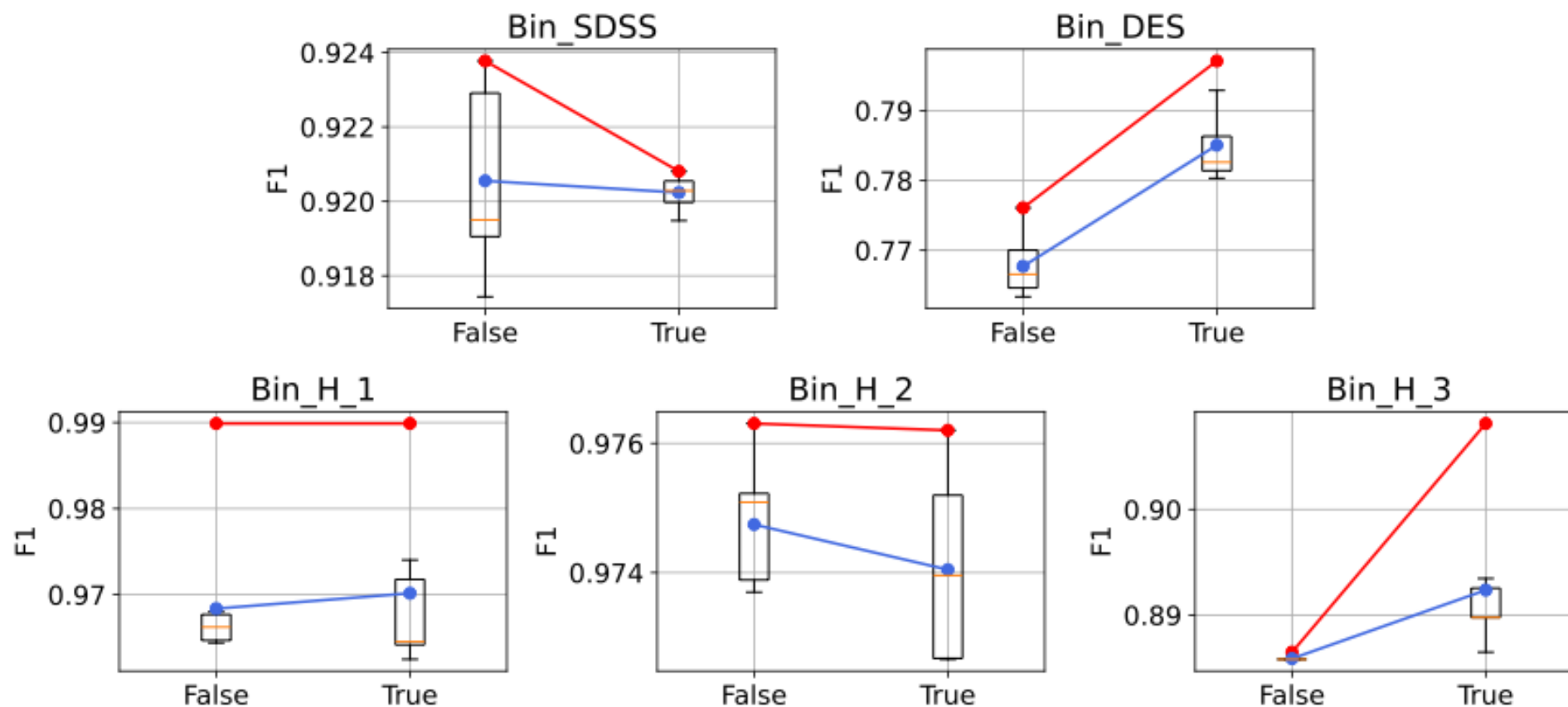
Naudota galaktikų: 10,000



Tarpiniai rezultatai – I etapas

● Max F1 value ● Top 10 F1 values, avg

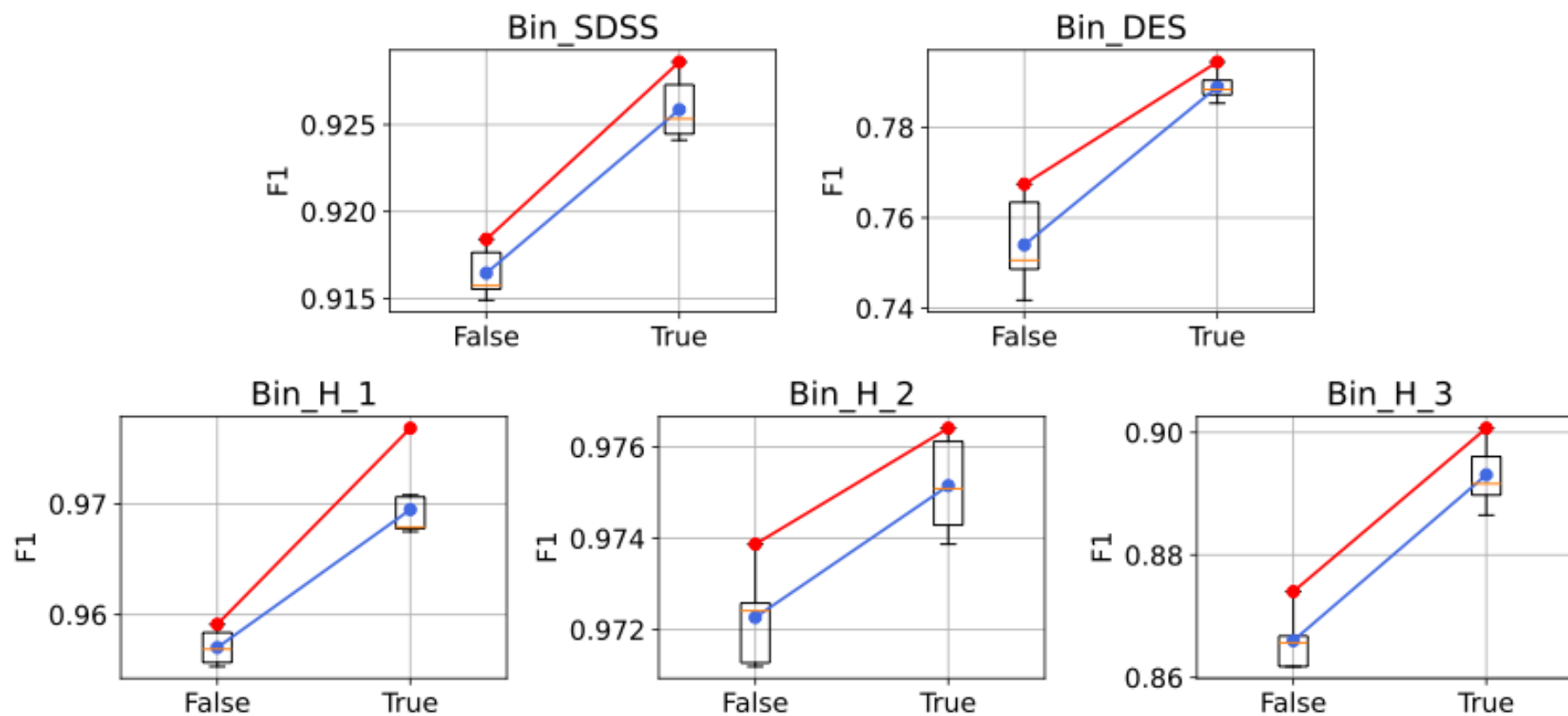
Focal



Tarpiniai rezultatai – II etapas

—●— Max F1 value —●— Top 10 F1 values, avg

Focal



Focal sluoksnio abliacijos tyrimas

Aspektai:

1. įvertinti, ar įprasta konvoliucija galėtų būti pakeista Focal sluoksniu kitų autorių architektūrose
2. įvertinti Focal sluoksnio įtaką atskirai, naudojant mūsų ankstesnes architektūras.

Naudojamas tas pats duomenų rinkinys, tik šį kartą, visi duomenys.

Focal sluoksnio abliacijos tyrimas

Pirmam aspektui patikrinti buvo naudojamos Cavanagh et al. bei Cheng et al. architektūros.

- Bazinė variacija – originali architektūra
- Focal variacija – visos konvoliucijos pakeistos Focal

Antram aspektui buvo naudojamos dvi architektūros iš mūsų ankstesnio tyrimo, kurios turėjo tiek Focal, tiek įprastas konvoliucijas.

- Bazinė variacija – visos konvoliucijos įprastos
- Originali variacija – tiek įprastos, tiek Focal, kaip aprašyta tyrime
- Focal variacija – visos konvoliucijos pakeistos Focal

Rezultatai

1. Kitų autorių architektūrų atveju, naudojant antžeminių teleskopų nuotraukas, rezultatai pagerėjo statistiškai reikšmingai
 - SDSS – Cavanagh: 80.34 → 81.56, Cheng: 78.99 → 81.60 (F1)
 - DES – Cavanagh: 93.56 → 94.07, Cheng: 93.44 → 94.14 (F1)
 - Visų rezultatų Wilcoxon testo $p = 0.002$
2. Mūsų architektūrų atveju, naudojant antžeminių teleskopų nuotraukas, geriausi rezultatai buvo pasiekti su originaliomis variacijomis
 1. SDSS – rezultatai panašūs
 2. DES – rezultatai geresni nei bazinė variacija su $p = 0.006$ bei $p = 0.01$
3. Naudojant kosmoso duomenis, visi rezultatai buvo panašūs arba net prastesni.

Išvados

1. Focal sluoksnis yra naudingas klasifikuojant antžeminių teleskopų nuotraukas, tačiau jis turi būti naudojamas konservatyviai.
2. Remiantis mūsų architektūrų sandara, Focal sluoksnis turėtų būti naudojamas modelio pradinuose sluoksniuose.
3. Focal sluoksnio įtaka yra minimali kuomet naudojamos kosminio teleskopo nuotraukos. Šio reiškinių priežastis gali būti tai, jog kosminių teleskopų nuotraukose yra mažiau foninių objektų bei atmosferos triukšmo.

Ačiū už dėmesį!