

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.

Visų studijų planas ir jo vykdymo suvestinė

Metai	Egzaminai		Dalyvavimas konferencijose		Publikacijos	
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta
I (2022/2023)	2	0 (studijuojama - 2)	-	0 (recenzavimui pateiktas konferencijos straipsnis)	-	-
II (2023/2024)	2	-	-	-	-	-
III (2024/2025)	-	-	1	-	1	-
IV (2025/2026)	-	-	1	-	1	-

Šiuo metu studijuojami dalykai - Informatikos ir informatikos inžinerijos tyrimo metodai ir metodika, Vaizdų ir duomenų analizė (*Fizikos fakultetas*)

Konferencija - ICCSA tarptautinė konferencija Graikijoje, 2023 m. liepos 3-6d.
Konferencijos darbai spausdinami „Lecture Notes of Computer Science“, Springer.

2022/2023 m.m. rudens semestro darbo planas ir jo vykdymas

Egzaminai		Dalyvavimas konferencijose		Publikacijos	
Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta
0	0	0	0	0	0

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.

Pagrindiniai galaktikų tipai

Spiralinė





Eliptinė



Duomenų rinkiniai

- Naudojami GalaxyZoo (GZ) duomenų rinkiniai – GZ1, GZ2, GZ Hubble, GZ DECaLS
- Duomenų rinkiniai grįsti savanorių klasifikacija
- Galaktikos klasė priskiriama pagal balsų daugumą, prieš tai įvertinus raudonojo poslinkio įtaką nuotraukai
- Numatomas patikimų duomenų dydis – apie 400,000 galaktikų



  You should sign in!

TASK

TUTORIAL

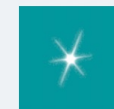
Is the galaxy simply smooth and rounded, with no sign of a disk?



Smooth



Features or Disk



Star or Artifact

Atlikti eksperimentai (1)

Pradinis eksperimentas, jokios augmentacijos netaikytos

Model	Acc. (σ)	Precision (σ)	Recall (σ)	TNR (σ)	F1 (σ)
Simplified	94.72 (0.24)	90.15 (0.51)	91.72 (0.63)	96.62 (0.25)	90.93 (0.41)
C2	94.71 (0.12)	89.80 (1.18)	92.19 (1.47)	96.80 (0.56)	90.96 (0.23)
Dieleman	94.70 (0.12)	89.68 (0.79)	92.28 (0.93)	96.83 (0.36)	90.95 (0.21)
ResNet50	93.53 (0.48)	89,54 (2.27)	88.01 (3.93)	95.21 (1.44)	88.66 (1.14)

Tik standartinės augmentacijos (pasukimas, apvertimas, priartinimas)

Model	Acc. (σ)	Precision (σ)	Recall (σ)	TNR (σ)	F1 (σ)
Simplified	94.56 (0.17)	89.84 (0.61)	91.52 (0.76)	96.54 (0.29)	90.67 (0.30)
C2	95.36 (0.14)	90.73 (0.68)	93.49 (0.95)	97.33 (0.37)	92.09 (0.26)
Dieleman	95.42 (0.17)	91.18 (0.81)	93.17 (0.72)	97.20 (0.28)	92.16 (0.28)
ResNet50	95.21 (0.23)	90.50 (1.11)	93.21 (1.02)	97.22 (0.39)	91.82 (0.36)

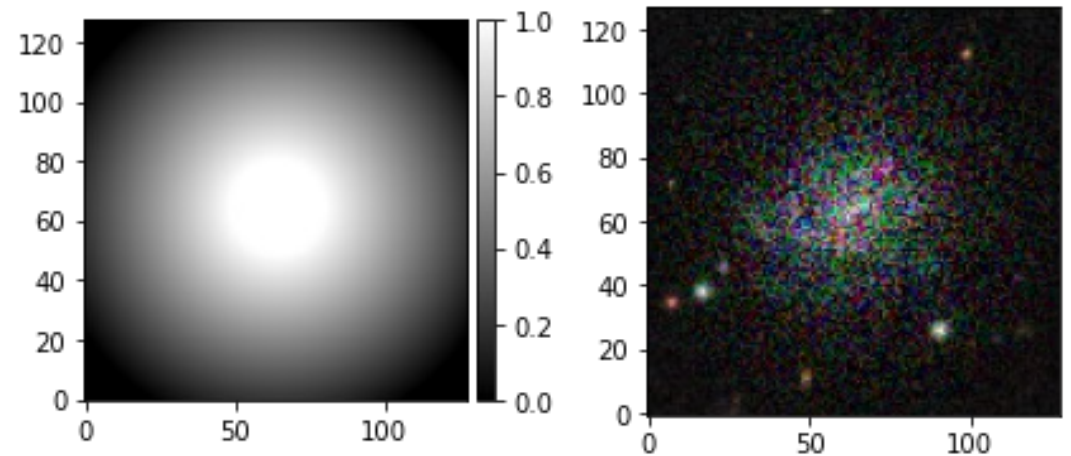
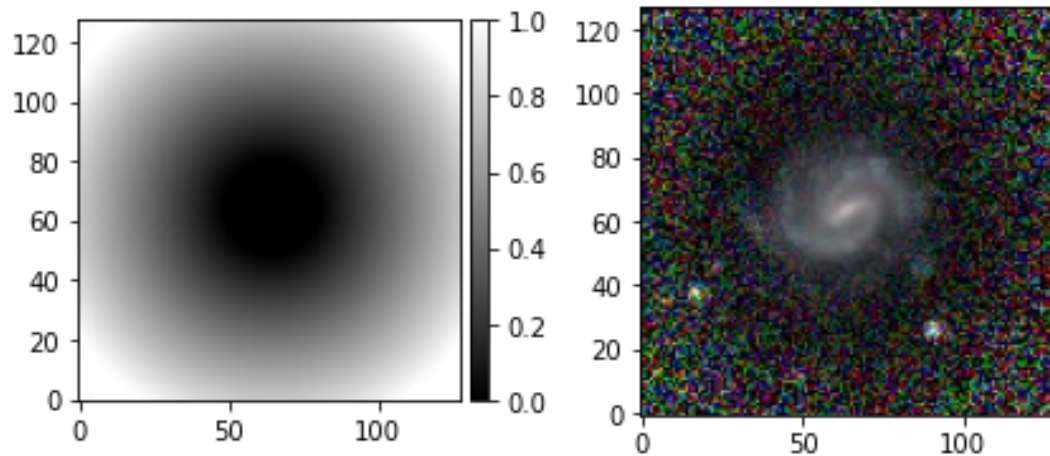
Atlikti eksperimentai (2)

Standartinės augmentacijos + atsitiktinis triukšmas

Model	Acc. (σ)	Precision (σ)	Recall (σ)	TNR (σ)	F1 (σ)
Simplified	94.47 (0.16)	89.80 (0.52)	91.21 (0.73)	96.41 (0.28)	90.49 (0.30)
C2	95.34 (0.10)	90.73 (0.55)	93.42 (0.54)	97.30 (0.21)	92.05 (0.16)
Dieleman	95.31 (0.08)	90.74 (0.46)	93.28 (0.53)	97.24 (0.20)	91.99 (0.14)
ResNet50	95.38 (0.09)	91.32 (0.57)	92.80 (0.53)	97.06 (0.20)	92.05 (0.14)

Rezultatai

- Atlikta galaktikų klasifikavimo į spiralines ir eliptines literatūros apžvalga
- Sudarytas patikimai klasifikuotų GZ1 ir GZ2 galaktikų sąrašas ir sugneruotos jų nuotraukos
- Įvertinta standartinių augmentacijų (pasukimas, apvertimas ir priartinimas) įtaka modelių metrikoms
- Išbandyti du nauji, atsitiktiniu triukšmu paremti augmentacijos būdai:



Išvados

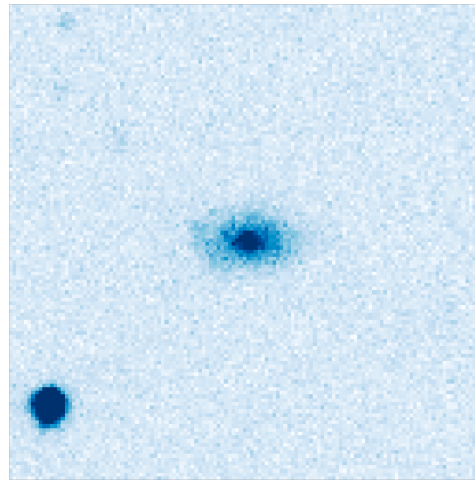
- Standartinės augmentacijos pagerina tikslumą ir F1 įvertį
- Naujos, triukšmu paremtos metrikos nesuteikia jokio statistiškai reikšmingo pagerėjimo visų metrikų atžvilgiu
- Atsitiktinis triukšmas taip pat nepagerina jokių metrikų, tačiau sumažina standartinį nuokrypį (net iki 9 kartų ResNet50 atveju)

Numatomi darbai (1)

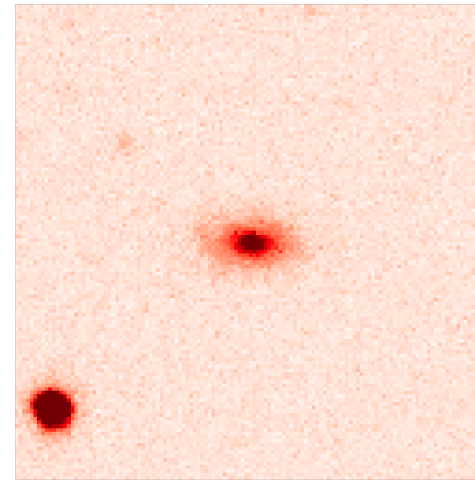
- Atrinkti patikimai klasifikuotas galaktikas visuose GZ duomenų rinkiniuose
- Teleskopų stebimų šviesos bangų ilgius suskirstyti į 3 režius
- Sugeneruoti galaktikos nuotrauką kiekviename iš režijų:



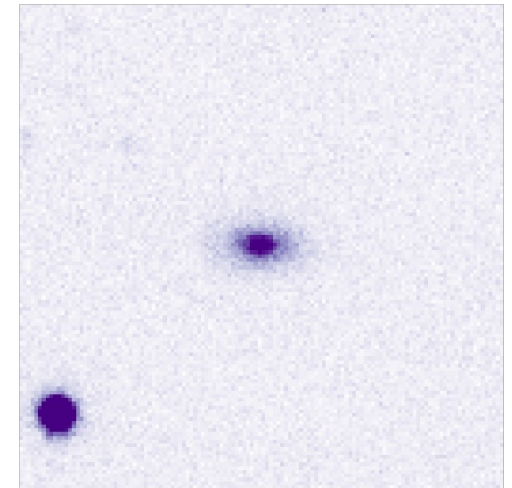
RGB nuotrauka



“Mėlynasis” režis
(*g*)



“Raudonasis” režis
(*r*)



“Infraraudonasis”
režis (*i*)

Numatomi darbai (2)

- Pasirinkti architektūras ir apmokyti modelius su kiekvienu iš režijų
- Patikrinama, ar kuris nors iš režijų neša daugiau informacijos (geresnis tikslumas ir kitos metrikos) nei kiti

Jei nėra vieno režio, nešančio žymiai daugiau informacijos:

- Kiekvienam režiiui parenkamas geriausias modelis
- Modeliai balsuoja dėl galutinės klasės
- Galimai atsižvelgiama į raudonojo poslinkio reikšmę, kiekvienam modeliui priskiriant atitinkamą balso svorį

Ačiū už dėmesį!