



VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
INFORMATIKOS INSTITUTAS
KOMPIUTERINIO IR DUOMENŲ MODELIAVIMO KATEDRA

Magistro baigiamasis darbas

Morfologinis galaktikų klasifikavimas naudojant konvoliucinius neuroninius tinklus

Morphological Galaxy Classification Using Convolutional Neural Networks

Atliko:

Rūta Marija Trečekauskaitė

parašas

Vadovas:

Prof. dr. Tadas Meškauskas

Vilnius
2026

Turinys

Sutartinis terminų žodynas	4
Dirbtinio intelekto naudojimo deklaracija	5
Santrauka	6
Summary	7
Įvadas	8
1. Galaktikų klasifikavimo uždavinys	10
1.1. Galaktikų klasifikavimas	10
1.2. <i>Galaxy Zoo</i> projektas	11
2. Literatūros apžvalga	16
2.1. Galaktikų klasifikavimas pasitelkiant <i>ResNet</i> konvoliucinių tinklų architektūrą	16
2.2. Tikimybinis <i>Bayesian</i> morfologinis klasifikavimas naudojant konvoliucinius neuroninius tinklus	16
2.3. Konvoliucinių neuroninių tinklų galaktikų klasifikavimo palyginimas keičiant klasių kieki	17
2.4. <i>Deep Galaxy</i> gilusis konvoliucinis neuroninis tinklas	18
2.5. <i>Xception</i> neuroninio tinklo panaudojimas galaktikų klasifikavimui	18
2.6. ViT panaudojimas morfologiniam galaktikų klasifikavimui	18
2.7. Galaktikų klasifikavimas pasitelkiant CvT	19
2.8. Konvoliucinis neuroninis tinklas galaktikų klasifikavimui	19
2.9. Galaktikų klasifikavimas naudojant jungtinį <i>GalaxyZoo</i> duomenų rinkinį	20
3. Mašininio mokymosi teorijos apžvalga	21
3.1. Giliojo mokymo principai	21
3.2. Dirbtiniai neuroniniai tinklai	23
3.3. Konvoliuciniai neuroniniai tinklai	24
3.4. Likutiniai neuroniniai tinklai	25
3.5. Transformatoriai	26
3.6. Regos transformatoriai	26
3.7. Konvoliuciniai regos transformatoriai	27
3.8. Hiperparametrų derinimas	28
3.9. Vertinimo metodai ir rodikliai	29
4. Darbo metodika	32
4.1. Darbo įrankiai ir duomenų rinkiniai	32
4.2. Modelių architektūros	36
4.2.1. <i>GalaxyCNN</i> architektūra	36
4.2.2. <i>GalaxyCvT</i> architektūra	38
4.3. Modelio vertinimo metodika	39
Rezultatai	41

Bendras klasifikacijos vertinimas	41
Mokymo vertinimas ir nuostolių funkcijos konvergavimas	41
Klasifikavimo kokybės analizė	42
Jautrumo duomenų kiekiui analizė	46
Modelių sprendimų vizualinio stabilumo analizė	47
Skaičiavimo efektyvumo analizė	50
Išvados ir rekomendacijos	52
Ateities tyrimų planas	53
Literatūros šaltiniai	54

Sutartinis terminų žodynas

- ANN (angl. *Artificial Neural Network*) – dirbtiniai neuroniniai tinklai, mašininio mokymosi poaibis paremtas pagrindinio smegenų ir nervų sistemos vieneto – neurono struktūra ir veikimo principu.
- BALD (angl. *Bayesian Active Learning by Disagreement*) – Bajeso aktyvaus mokymosi metodas paremtas nesutarimo principu.
- CNN (angl. *Convolutional Neural Networks*) – konvoliuciniai neuroniniai tinklai.
- CTE (angl. *Convolutional Token Embedding*) – konvoliucinis žetonų įterpimas.
- CvT (angl. *Convolutional Vision Transformers*) – konvoliuciniai regos transformatoriai.
- FFN (angl. *feed-forward network*) – tiesioginio perdavimo tinklas.
- Grad-CAM (angl. *Gradient-weighted Class Activation Map*) – populiari metodika vizualizuoti modelio "matomas" vietas priimant sprendimus.
- MHA (angl. *multi-head self-attention mechanism*) – daugiasluoksnis dėmesio mechanizmas.
- MLP (angl. *multi-layer perceptron*) - daugiasluoksnis perceptronas.
- ResNet (angl. *Residual Networks*) – dirbtiniai neuroniniai tinklai paremti liekamaisiais blokais.
- SDSS DR (angl. *Sloan Digital Sky Survey Data Release*) – Sloano skaitmeninis dangaus tyrimas yra didelio masto daugiaspektris vaizdavimo ir spektroskopinių raudonųjų poslinkių tyrimas, atliekamas naudojant specialų 2,5 m plačiakampį optinį teleskopą *Apache Point* observatorijoje Naujojoje Meksikoje, Jungtinėse Amerikos Valstijose. Septintajame pagrindiniame duomenų leidinyje pateikiami vaizdai, vaizdų katalogai, spektrai ir raudonieji poslinkiai.
- SGD (angl. *Stochastic Gradient Descent*) – stochastinis gradientų nusileidimas, modelių optimizacijos metodas.
- SSIM (angl. *Structural Similarity Index Measurement*) – vaizdo vertinimo rodiklis (naudojamas palyginti dviejų vaizdų panašumą).
- SVM (angl. *Support Vector Machine*) – palaikomųjų vektorių mašinos - tai prižiūrimo mokymosi metodų rinkinys, naudojamas klasifikavimo, regresijos ir kitų uždavinių sprendimui.
- ViT (angl. *Vision Transformers*) – regos transformatoriai.
- Centrinis telkinys (angl. *central bulge*) - glaudžiai susitelkusi žvaigždžių grupė matoma didesniojo žvaigždžių darinio centre.
- Skersėtosios spiralinės galaktikos - galaktikos pasižyminčios skerse (angl. *bar*) esančia galaktikos centre (angl. *Barred spiral galaxies*).

Dirbtinio intelekto naudojimo deklaracija

Šiame darbe buvo naudoti dirbtinio intelekto (DI) įrankiai:

- *Google NotebookLM* (Google LLC) <https://notebooklm.google/>,
- *ChatGPT* versija: GPT-5.2 (OpenAI Inc.) <https://chatgpt.com/>,
- *Copilot* versija: GPT-5 mini (Microsoft Corporation) <https://copilot.microsoft.com/>,
- *DeepL Translate* versija: Q4 2025 (DeepL SE) <https://www.deepl.com/en/translator>,
- *Gemini* versija: 3 (Google LLC) <https://www.google.com/>.

DI buvo pasitelktas šiais tikslais:

- akademinio stiliaus teksto ir kalbos koregavimui,
- formuluočių aiškumo gerinimui, sakinių struktūros peržiūrai ir naudojamų sąvokų vertimui,
- mokslinių straipsnių palyginimui, informacijos apibendrinimui bei paieškai,
- kodo taisymui, klaidų sprendimui, kodo rašymo gairėms.

Santrauka

Galaktikų klasifikavimas ilgą laiką buvo atliekamas rankiniu būdu, surinktų duomenų peržiūras vykdant specialistų ar savanorių grupėms. Dėl pažangos elektronikos srityje nulemtu eksponentinio informacijos augimo astronomijos srityje, klasifikavimo uždavinio sprendimui imta naudoti automatizuotus sprendimus, tačiau grupavimo uždavinys žymiai susiaurintas klasių atžvilgiu. Atsižvelgiant į tai, šiame darbe nagrinėjamas konvoliucinių neuroninių tinklų taikymas automatiniam galaktikų morfologijos klasifikavimo uždaviniui spręsti. Darbe siekiama sukurti, įgyvendinti ir įvertinti CNN pagrįstą modelį, skirtą elipsinių, spiralinių (su skersine struktūra ir be jos) bei šonu stebimų galaktikų klasifikavimui, naudojant *Galaxy Zoo* ir kitų kosmologinių tyrimų duomenis. Darbe pateikiamas galaktikų klasifikavimo uždavinio sprendimas pasitelkiant konvoliucinius neuroninius tinklus – sukurtas *GalaxyCNN* modelis. Atlikta išsami eksperimentinė analizė, atsižvelgiant į ryškų klasių disbalansą, skirtingus duomenų rinkinius ir ribotą klasifikuotų duomenų kiekį. Siūlomas *GalaxyCNN* modelis pasiekė aukštą bendrą klasifikavimo tikslumą ($F1_{macro} = 0.897$) ir nežymiai pranoko palyginamąjį konvoliucinio regos transformatoriaus architektūra paremtą modelį. Rezultatai parodė, kad CNN sprendimai pasižymi erdviu stabilumu ir atsparumu foniniam triukšmui, o didžioji dalis klaidų kyla dėl pačios galaktikų morfologijos neapibrėžtumo. Darbe patvirtinta, jog CNN architektūros yra efektyvus ir patikimas sprendimas galaktikų morfologijos analizės uždaviniams, ypač esant ribotiems ir heterogeniškiems duomenims.

Summary

Morphological Galaxy Classification Using Convolutional Neural Networks

Galaxy morphology classification has traditionally been performed manually by experts or groups of volunteers. However, the rapid growth of astronomical data driven by advances in observational technology has made manual classification increasingly impractical. Although automated approaches have been introduced, they are often limited in classification scope due to data complexity, class imbalance, and variability across astronomical surveys. In this context, this research focuses on the application of convolutional neural networks (CNNs) as an effective and robust solution for automated galaxy morphology classification.

The main objective of this work is to design, implement, and evaluate a CNN-based model for classifying galaxies according to their morphological features. The study utilizes data from the Galaxy Zoo project and multiple cosmological surveys, focusing on elliptical (smooth-round, cigar-shaped), spiral (barred and unbarred), and edge-on galaxies derived from the corresponding Galaxy Zoo decision trees. In addition, a convolutional vision transformer model (GalaxyCvT) is implemented solely for comparative purposes, representing a newer architectural approach.

A comprehensive experimental evaluation is conducted under conditions of significant class imbalance, using the macro-averaged F1 score as the primary metric. Model performance is assessed across multiple datasets with differing image acquisition methods, spectral ranges, and observational equipment, and additional experiments analyze sensitivity to the amount of available training data. Model interpretability is examined using Grad-CAM visualizations to evaluate spatial stability and feature focus under image transformations and background noise, with results compared against attention-based visualizations from the CvT model. Furthermore, computational efficiency is evaluated in terms of training and inference performance.

The proposed GalaxyCNN model achieves a high overall classification performance $F1_{macro} = 0.897$, outperforming the comparative GalaxyCvT model ($F1_{macro} = 0.884$), with particularly strong results for challenging classes (i.e. barred spiral galaxies) and improved robustness when trained on smaller datasets.

Most misclassifications occur between morphologically similar classes, indicating that misclassifications are primarily caused by intrinsic ambiguity in galaxy morphology rather than architectural limitations. Visualization results confirm that the CNN model bases its decisions on stable, local morphological features that are robust to background noise, while the CvT model exhibits greater sensitivity to input variations.

In conclusion, this thesis confirms that CNNs remain a highly effective and reliable approach for automated galaxy morphology classification, particularly in scenarios with limited and heterogeneous data. While transformer-based models represent a promising direction, their advantages are not fully realized under current data constraints; therefore, future work should focus on enhancing CNN architectures, expanding labeled datasets, integrating additional spectral information, and improving robustness and interpretability for large-scale astronomical surveys.

Įvadas

Žvaigždžių, jų sistemų, tarpžvaigždinių medžiagų (dulkių, dujų, kosminių spindulių) ir tam-siosios medžiagos sistemos, kurias sieja tarpusavio traukos jėgos, sudaro galaktikas. Stebimojoje visatoje matomos galaktikos išsiskiria savo išvaizda, kurią lemia pirminės formavimosi sąlygos ir įvairios sąveikos su kitais atitinkamo mastelio dangaus kūnais. Dėl tokios šių dangaus kūnų prigimties vis dar nėra neprieštaringos galaktikų grupavimo schemos, o šių dienų klasifikavimo pamatu vis dar galima laikyti 1926 m. E.P. Hubble pasiūlytą schemą [1], kurioje grupavimą nulemia galaktikų formos. Klasifikavimas padeda daryti pirmines prielaidas apie galaktikos gimimą, evoliuciją ir visatos prigimtį.

Galaktikų klasifikavimas ilgą laiką buvo atliekamas rankiniu būdu, surinktų duomenų peržiūras vykdant specialistų grupėms, tačiau augant duomenų kiekiui toks sprendimas tapo neveiksmingas. 2007 m. pradėtame projekte *Galaxy Zoo* [2] [3] grupavimo uždavinio sprendimui buvo pasitelktas savanoriškas visuomenės prisidėjimas, kuriame iki 2019 m. buvo surinkta virš 400 000 savanorių [4]. Nepaisant to, dėl pažangos elektronikos srityje, kuri nulėmė eksponentinį informacijos astronomijos srityje augimą [5], grupavimo atlikimas rankiniu būdu tapo per brangus laiko ir žmogiškųjų išteklių atžvilgiu. Dėl šios priežasties pastaraisiais metais vis daugiau dėmesio skiriama automatizuotiems sprendimams, pasitelkiant dirbtinius neuroninius tinklus, Bayesian metodus ir kitus mašininio mokymosi algoritmus [6] [7] [8] [9] [10].

Atsižvelgiant į šias problemas, šio darbo motyvacija – įvertinti konvoliucinių neuroninių tinklų tinkamumą automatiniam galaktikų morfologijos klasifikavimui, ypač esant ribotam ir nevienalyčiam duomenų kiekiui. **Darbo tikslas** – sukurti, įgyvendinti ir eksperimentiškai įvertinti giliojo mokymosi metodą, paremtą konvoliuciniais neuroniniais tinklais, skirtą galaktikų klasifikavimui pagal jų morfologinius požymius (eliptiškumą, spiralines struktūras ir stebėjimo kampą). Darbo tikslui pasiekti išskiriami šie esminiai uždaviniai:

1. Vaizdinių ir tekstinių duomenų analizė ir apdorojimas – atlikti vaizdinių ir savanorių balsavimo duomenų analizę bei apdorojimą, atrenkant tinkamus duomenų poaibius ir pritaikant problematikai adekvačius išankstinio apdorojimo metodus;
2. Tinklų architektūros modeliavimas – suprojektuoti ir įgyvendinti konvoliucinio neuroninio tinklo architektūrą, pritaikytą galaktikų morfologinių požymių išgavimui;
3. Sukurto modelio įvertinimas – pagal nagrinėjamos problemos ypatybes tinkamai parinkti vertinimo metodus ir rodiklius, išanalizuoti ir įvertinti modelio efektyvumą, stabilumą ir našumą.

Šis darbas tęsia autorės ankstesniame magistro studijų semestre vykdyto mokslo tiriamojo darbo projekto metu pradėtus tyrimus. Ankstesniame darbe buvo atlikta pirminė galaktikų vaizdinių duomenų analizė ir išbandytas konvoliuciniais neuroniniais tinklais paremtas daugiaužduotinio modelio kūrimas spiralinių galaktikų poaibiui klasifikuoti. Dėl riboto duomenų kiekio sudėtingesnėms poskyrio klasėms pasiekti rezultatai buvo riboti, o eksperimentinė analizė – siauresnės apimties. Šiame darbe ankstesnio tyrimo pagrindu buvo panaudota galaktikų klasifikavimo problemos formuluotė (1 skyrius), literatūros apžvalga (2 skyrius), bei mašininio mokymosi teoriniai pagrindai (3 skyrius), kurie buvo išplėtoti ir pritaikyti platesniam galaktikų morfologijos klasifikavimo uždaviniui.

Šis darbas, atsižvelgiant į tikslą ir pagrindinius uždavinius, suskirstytas į penkias pagrindines dalis: Galaktikų klasifikavimo problemų ir *Galaxy Zoo* projekto apžvalgą (1 skyrius), literatūros

šaltinių analizę (2 skyrius), darbui aktualios teorijos apžvalgą (3 skyrius), darbo metodologijos aprašymą (4 skyrius) bei rezultatus ir išvadas. Pirmajame skyriuje pateikiamas teorinis įvadas apie galaktikų klasifikavimą, pristatoma *Galaxy Zoo* projekto metodologija. Antrajame skyriuje aptariami anksčiau vykdyti tyrimai galaktikų grupavimo uždaviniui spręsti. Trečiasis skyrius pristato tyrime naudotų metodų, vertinimo rodiklių bei mokslo sričių teorinį pagrindą. Metodologijos skyriuje pateikiama informacija apie duomenų apdorojimo ir analizės procesus, tinklo modeliavimo metodus, bei vertinimo metodiką ir rodiklius. Tolimesniuose skyriuose pateikiami tyrimo rezultatai, išvados bei ateities tyrimų planas. Darbe cituojamų literatūros šaltinių sąrašas pateikiamas darbo pabaigoje.

Šiame darbe buvo sukurta ir įgyvendinta konvoliuciniais neuroniniais tinklais paremta galaktikų morfologijos klasifikavimo sistema, skirta darbui su dideliais, klasių disbalansą turinčiais astronominiais duomenų rinkiniais. Tyrimo metu atlikta išsami eksperimentinė analizė, apimanti skirtingų duomenų rinkinių palyginimą, modelio jautrumo mokymo duomenų kiekiui vertinimą, sprendimų interpretavimo analizę bei skaičiavimo našumo įvertinimą. Taip pat buvo išspręstas galaktikų morfologinės klasifikacijos automatizavimo uždavinys, parodant, kad konvoliuciniai neuroniniai tinklai leidžia patikimai išgauti lokalius morfologinius požymius net esant ribotam duomenų kiekiui. Gauti rezultatai patvirtino siūlomo CNN sprendimo efektyvumą ir stabilumą, o palyginamoji analizė su konvoliucinių regos transformatorių modeliu leido įvertinti architektūrinių sprendimų tinkamumą praktiniams astronominiais taikymams.

1. Galaktikų klasifikavimo uždavinys

Šiame skyriuje aprašoma galaktikų klasifikavimo sistemos ir aptariamas *Galaxy Zoo* projekto savitumas – projektų sprendimų medžiai, projektuose surinktų duomenų pirminis apdorojimas ir pagrindinės kliūtys, susietos su savanorių šališkumu ar vertinimo paklaidomis dėl įvairių priežasčių, bei jų sprendimo būdai.

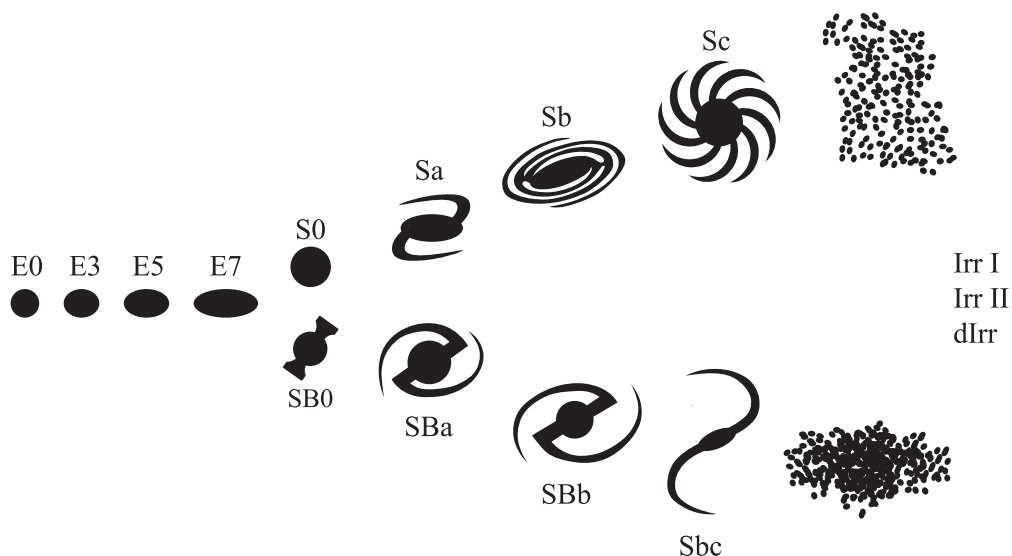
1.1. Galaktikų klasifikavimas

Pamatinis galaktikų klasifikavimas atliekamas pagal Hubble kamertono diagramą [11] (1 paveikslas) kuris yra grindžiamas fotografiniu (morfologiniu) galaktikų vaizdų. Galaktikų klasifikavimas pagal Hubble kamertono diagramą nėra absoliutus – klasifikavimas neatsižvelgia į neįprastus atributus ir iš savo pagrindų turi neteisingų prielaidų [12]. Dėl šių priežasčių galaktikų klasės dažnai papildomos atsižvelgiant į Elmegreen spiralinių galaktikų klasifikavimą [13] arba papildant kamertono diagramą pagal De Vaucouleurs [14].

Skiriami šeši galaktikų tipai: elipsinės, spiralinės, lėšinės, diskinės, skersėtosios spiralinės, netaisyklingos. Elipsinės galaktikos – sukimosi arba triašio elipsoido formos, žymimos En , kur

$$n = 10 \cdot \left(1 - \frac{b}{a}\right), \quad (1.1)$$

elipsiškumas n – sveikas skaičius, a , b – galaktikos dydis išilgai elipsoido didžiosios ir mažosios ašių atitinkamai. Elipsinės galaktikos skurdžios dujomis, neturi jaunų žvaigždžių, žvaigždžių sistemos juda lėtai. Daroma prielaida, kad susiformuoja kaip kelių galaktikų susiliejimo pasekmė.



1 pav. Hubble kamertono diagrama.

Spiralinės galaktikos – plonas diskas (centrinis telkinys) sudarytas iš žvaigždžių ir tarpžvaigždinės medžiagos su nusitęsiančiomis spiralinėmis vijomis. Spiralinės galaktikos (žymimos raide S) skirstomos į spirales, skersėtasias spirales (žymimos SB). Iškiriami trys potipiai – a , b , c , kurie nusako centrinio telkinio dydį ir vijų glaudumą (a – didžiausias centrinis telkinys, labai glaudžios vijos, b – vidutinio dydžio centrinis telkinys, "laisvos" vijos, c – mažiausias centrinis telkinys, labai "laisvos" vijos). Skersėtosios spiralinės galaktikos pasižymi vijomis išeinančiomis iš centre esančios skersės galų, taip pat grupuojamos pagal potipius. Spiralinių galaktikų vijose gausu dujų ir dulkių, dažnos žvaigždėdaros sritys.

Pereinamasis tipas tarp elipsinės ir spiralinės galaktikos – turi centrinę telkinį tačiau neturi vijų – diskinės galaktikos žymimos S0. Lėšinės galaktikos – diskinės galaktikos turinčios skersę, žymimos SB0. S0, SB0 tipo galaktikos, kaip ir elipsinės neturi žvaigždėdaros regionų, dėl to daroma prielaida, jog yra spiralinės galaktikos evoliucijos fazė. Nesimetrinės sandaros galaktikos vadinamos netaisyklingomis (žymima Irr). Skirstomos į tris tipus: Irr I – manoma, jog tai – taisyklingos galaktikos, kurių struktūrą išdaro stochastiškas aktyvių žvaigždėdaros sričių išsidėstymas, Irr II – chaotiška forma nulemta sąveikų su kitomis galaktikomis, dIrr – nykštukinės netaisyklingosios galaktikos – nemetalingos, panašios į ankstyvosios visatos galaktikas.

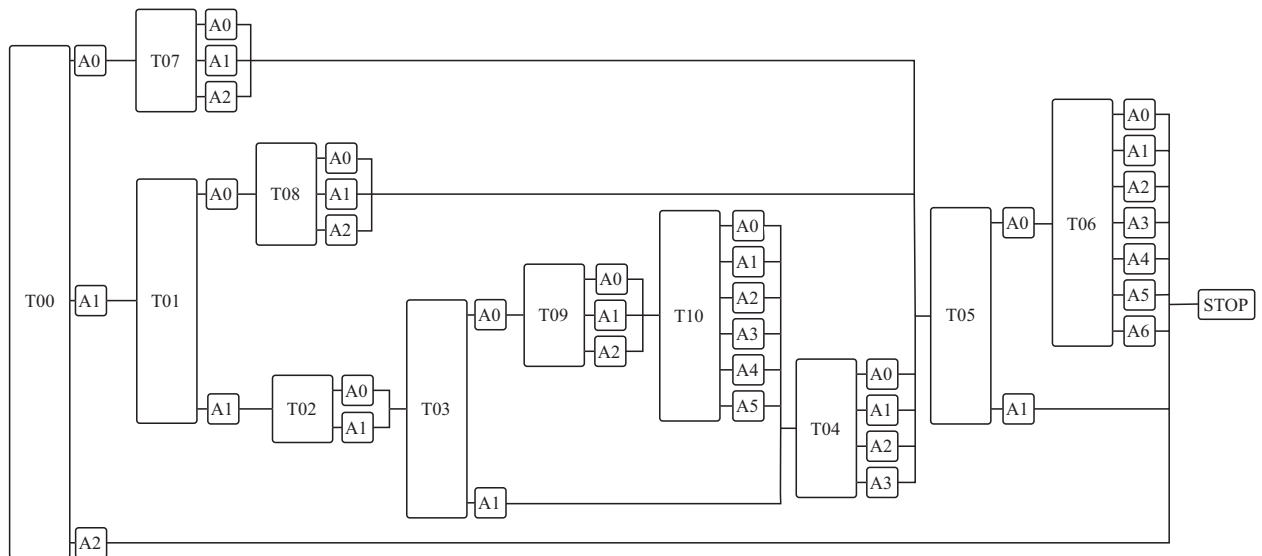
1.2. *Galaxy Zoo* projektas

Galaxy Zoo – mokslinis visuomeninis projektas, kuriame interneto vartotojai gali prisidėti prie galaktikų klasifikavimo. Projektas buvo pradėtas 2007 m., o bėgant laikui buvo išplėtotas į daugiau smulkesnių projektų pagal nagrinėjamus dangaus kūnus. Šiame darbe naudojami duomenys iš šių projektų:

1. *Galaxy Zoo 2* [15] [16] (GZ2) – *Galaxy Zoo* projekto tęsinys. Šiame projekte į klasifikavimą įtrauktas skersių grupavimas, vijų skaičius, elipsinių galaktikų klasifikavimas, centrinio telkinio dydis, neįprasti bruožai ir pan. Projekte naudojamos galaktikų nuotraukos sudarytos iš *SDSS Data Release 7 (SDSS DR7)* [17] duomenų poaibio (pasirenkant maždaug 25% ryškiausių galaktikų). Nuotraukos kuriamos pasitelkiant *gri* filtrų duomenis. Klasifikavimui naudojamų galaktikų raudonojo poslinkio diapazonas: $5 \cdot 10^{-3} < z < 2,5 \cdot 10^{-1}$. Duomenų rinkinį sudaro daugiau kaip 304 000 galaktikų.
2. *Galaxy Zoo Hubble* [18] (GZH) projekte pasirinkimų medis buvo išplėtotas įtraukiant įvairias klases siekiant aprašyti grupuotas galaktikas (angl. *clumpy galaxies*). Projekte naudojami duomenys iš kelių skirtingų duomenų rinkinių, tačiau pagrindą sudaro *Advanced Camera for Surveys General Catalog, ACS-GC* [19]. Duomenų rinkinį sudaro apie 120 000 galaktikų (raudonojo poslinkio mediana $\langle z \rangle = 0.9 \pm 0.6$).
3. *Galaxy Zoo CANDELS* [20] (GZC) – pasirinkimų medis iš esmės nepakitęs nuo naudoto GZH projekte, tačiau sumažintas galimų pasirinkimų skaičius prie klausimų apie centrinį telkinį. Projekte galaktikų fotografijoms naudojami duomenys iš *Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey, CANDELS* [21]. Apie 90 % galaktikų $z \leq 3$. Rinkinį sudaro apie 48 000 galaktikų.
4. *Galaxy Zoo DECaLS* [22] (GZD) – galaktikų nuotraukoms naudoti duomenys iš *Dark Energy Camera Legacy Survey (DECaLS)* [23], klasifikavimo schema tapati GZ2 projekte naudotam pasirinkimų medžiui, bet sumažintas galimų pasirinkimų skaičius galaktikų vijoms, bei pridėtas papildomas klausimas apie matomus galaktikų susijungimo bruožus. Tyrimas buvo padalintas į tris dalis – GZD-1, GZD-2 ir GZD-5. Bendrą rinkinį sudaro apie 314 000 galaktikų.
5. *Galaxy Zoo UKIDSS* [24] (GZU) – projekte naudojami duomenys iš *United Kingdom Infrared Telescope Infrared Deep Sky Survey, UKIDSS* [25] [26]. Galaktikų nuotraukos kuriamos pasitelkiant *yjk* filtrų duomenis. Duomenų rinkinį sudaro daugiau kaip 71 000 galaktikų, kurios taip pat matomos ir GZ2 rinkinyje. Klasifikavimo pasirinkimų medis toks pat kaip ir GZ2 projekte.

Galaxy Zoo 2 ankstesniame projekte sprendžiamą klasifikavimo uždavinį išplėtoja, 2 paveiksle pateiktas projekto klausimų pasirinkimo medis. T_n – klausimo žymėjimas, A_n – galimi atsakymai. Svarbu pastebėti, jog galaktikoms nėra nusakoma tiksli klasė t.y., iš klausimų galima galaktikai priskirti tinkamus bruožus.

Kiekviena galaktika buvo atrenkama atsitiktine tvarka ir pateikiama savanoriams internete talpiama sąsaja. Norint naudotis šia sąsaja, savanoriams buvo iškeltas reikalavimas susikurti paskyrą. Toks sprendimas priimtas su idėja, kad ne kiekvienas balsas gali būti vienodo "svorio". Tokio vartotojų pasirinkimų saugojimo dėka projekto pabaigoje buvo galima atlikti kiekvieno savanorio balso svarumo įvertinimą, todėl atliktas galaktikų klasifikavimas tikslesnis. Šis metodas buvo naudojamas ir originaliajame *Galaxy Zoo* projekte [2].



2 pav. *Galaxy Zoo 2* galaktikų klasifikavimo pasirinkimo medis. Klausimai atsakomi eilės tvarka iš kairės į dešinę, tolydžios linijos indikuoja eigos srautą.

Sudėtingesnis pasirinkimų medis, atsižvelgia į tokius bruožus, kaip:

- skersių grupavimas,
- vijų skaičius,
- vijų glaustumas,
- elipsinių galaktikų klasifikavimas,
- centrinio telkinio dydis,
- neįprasti bruožai.

1 lentelėje pateikiamas projekto klausimynas ir galimi atsakymai. Priskirti jau žinomą klasifikavimo modelį nėra taip paprasta, kadangi pasirinkimo medis turi ~ 1400 galimų atsakymų atsižvelgiant į visas galimas atsakymų variacijas.

Galaxy Zoo 2 projekte buvo atsižvelgta į kelias esmines problemas, siekiant užtikrinti, kad duomenys, surinkti savanoriškai klasifikuojant galaktikų atvaizdus, būtų tinkamai apdoroti ir reprezentatyvūs. Pagrindiniai šio proceso etapai apėmė klausimų, susijusių su raudonojo poslinkio

paklaida (angl. *redshift*), morfologinės klasifikacijos paklaida ir balsavimo duomenų patikimumo užtikrinimu, sprendimą.

1 lentelė. *Galaxy Zoo 2* projekto klausimynas.

$T_n \setminus A_n$	A0	A1	A2	A3	A4	A5	A6
T0: Ar galaktika elipsės formos ir neturi kitų atributų?	Elipsinė	Matomi diskai	Žvaigždė ar kitas dangaus kūnas	-	-	-	-
T1: Ar galaktikos orbitos plokštuma yra lygiagreti regėjimo linijai su Žeme?	Taip	Ne	-	-	-	-	-
T2: Ar matomas skersės?	Taip	Ne	-	-	-	-	-
T3: Ar matomos vijos?	Taip	Ne	-	-	-	-	-
T4: Koks centrinio telkinio ryškumas?	Telkinio nėra	Neryškus	Matomas	Dominuojantis	-	-	-
T5: Ar yra neįprastų bruožų?	Taip	Ne	-	-	-	-	-
T6: Kokie neįprasti bruožai matomi?	Žiedas	Lęšis	Iškreipimai	Netaisyklingos galaktikos bruožai	Kita	Galaktikų susiliejęmas	Dulkių debesys
T7: Koks galaktikos elipsiškumas?	Idealiai apvali	Elipsės formos	Cigaro formos	-	-	-	-
T8: Ar matoma skersė ir kokios ji formos?	Apvali	Kvadratiška	Nėra	-	-	-	-
T9: Ar vijos glaudžios?	Labai glaudžios	Vidutiniškai glaudžios	Visai ne glaudžios	-	-	-	-
T10: Kiek vijų matoma?	1	2	3	4	>Daugiau nei 4	Negaliu atsakyti	-

Kaip minėta anksčiau, kiekvienas balsas buvo palygintas su balsų dalimi $f_r = n_r/n_t$, kur n_r - balsų skaičius atsakymui, n_t - bendras balsų kiekis klausimui, kad nusakyti vartotojo nuoseklumą:[15]:

$$k = \frac{1}{N_r} \sum_{i=1}^{N_r} k_i, \quad (1.2)$$

atsižvelgiant į sumą įmanomų atsakymų N_r klausimui, ir vartotojo nuoseklumą k_i priklausomai nuo pasirinkto atsakymo. Pagal tai buvo paskaičiuotas bendras nuoseklumas vartotojui $\bar{k}$, pagal kurį buvo priskirti svoriai w . Nuoseklumas buvo perskaičiuotas tris kartus, užtikrinant konvergavimą, taip įvedant "pasvertų" balsų vertinimą kiekvienai galaktikai.

Raudonojo poslinkio paklaida atsiranda dėl to, kad didesnio raudonojo poslinkio galaktikos atrodo mažesnės ir neryškos, todėl sunkiau tiksliai klasifikuoti jų morfologines savybes. Ši problema buvo sprendžiama naudojant Petrosiano spindulį (angl. *Petrosian radius*) [27]), kad išdidintų nuotraukas. Tačiau dėl to, galaktikų nuotraukos būna prastesnės skiriamosios gebos ir galaktikų klasifikavimo uždavinys pasidaro gerokai sudėtingesnis tiek savanoriams, tiek ekspertams ar kompiuteriniams algoritmams. Willett, et al. šiam šališkumui ištaisyti pasitelkė metodiką paremtą principu taikytu pirmojo projekto duomenų tvarkyme [28]. Kiekvienos užduoties j neapdorotos balsų dalys f_j buvo pakoreguotos naudojant pataisos koeficientą, pagrįstą galaktikos raudonuoju poslinkiu z . Tiksliau, paklaidos korekcija buvo atliekama remiantis empiriniu santykiu, nustatytu pagal modeliuotus galaktikų, turinčių skirtingas raudonojo poslinkio vertes, vaizdus. Šis santykis

buvo naudojamas balsų dalims pakoreguoti taip, kad jos būtų tokios, kokios būtų, jei galaktika būtų stebėta esant fiksuotam atskaitos raudonajam poslinkiui.

Tačiau, pagal Ross E. Hart, *et al.* [16] buvo išskirtos dvi pagrindinės problemos atsižvelgiant į prieš tai minėtą problemos sprendimą – reguliuojant vidurkių vertes nėra pataisomos konkrečių balsų vertės, o klausimams kurie reikalauja konkrečių atsakymų į prieš tai buvusius klausimus pataisos nėra teisingos ir turi didelius nuokrypius. Taip pat svarbu paminėti, jog tokie klausimai iš esmės visada turi ir gerokai mažesnę balsų kiekį, nes balsuotojas paprasčiausiai šių klausimų nepasiekia, kadangi esant didesniam raudonajam poslinkiui tikėtina, kad ankstesniuose klausimuose bus sudėtinga išskirti reikiamus bruožus. O esant bruožams kuriuos būtų galima vadinti "sunkiai išskiriamais" atitinkamai didėja ir netolygumai duomenų pasiskirstyme konkretaus klausimo atsakymų atžvilgiu. Atsižvelgiant į prieš tai minėtas problemas (ypač gerai matoma su netolygiu duomenų pasiskirstymu klausime T10, 2 pav.), balsų dalies pasiskirstymas buvo modeliuojamas kaip funkcija nuo raudonojo poslinkio z , o galaktikos buvo suskirstytos į atskiras aibes pagal klausimus. Papildomai, duomenys buvo priderinti (angl. *fitted*) pagal tolydžią funkciją, iš kurios surandami geriausi parametrai, atsižvelgiant į raudonąjį poslinkį z , galaktikos dydį R_{50} ir ryškumą M_R . Abu sprendimai palyginti atsižvelgiant į duomenų rinkinį kuris neturi raudonojo poslinkio paklaidos.

Galaxy Zoo Hubble panaudoti keli skirtingi galaktikų nuotraukų rinkiniai: AEGIS [29], GOODS [30], GEMS [31] [32], COSMOS [33] [34]. Svarbu paminėti, kad dalis naudojamų galaktikų buvo papildomai apdorotos pridėdant aktyvų galaktikos branduolį (angl. *active galactic nucleus*) arba įvedant didesnę raudonąjį poslinkį. Antrasis pakeitimas buvo atliktas, kadangi raudonojo poslinkio diapazonas žymiai didesnis, todėl prieš tai taikyti metodai nėra tinkami – tam naudojama papildomai apdorotų galaktikų aibė.

Galaxy Zoo CANDELS projekte naudotos galaktikos pasižymi gerokai žymesniu raudonuoju poslinkiu lyginant su galaktikomis kituose projektuose. Papildomai projekte buvo atlikti paveiksluko gylio (angl. *depth of image*) taisymai, tam atlikti buvo įvedamos papildomos "gilesnės" galaktikų nuotraukos.

Galaxy Zoo DECaLS projekte balsuotojai klasifikavimą atliko galaktikoms iš DECaLS nuotraukų iš SDSS DR8 duomenų rinkinio (angl. *Sloan Digital Sky Survey Data Release 8*, toliau SDSS DR8). Dėl geresnės nuotraukų skiriamosios gebos ir gylio, šiame rinkinyje geriau matyti galaktikų bruožai [22]. Dėl didesnio informacijos kiekio nuotraukose, pasirinkimų medyje buvo įvesti keli pakeitimai (žr. 2):

- T02 klausimas išplėtotas – A0 ("Ne"), A1 ("Vos matomos"), A2 ("Taip")
- T10 klausime panaikintas A5 atsakymas ("Negaliu atsakyti"),
- T04 klausime, centrinio telkinio apibūdinimas apribotas pasirinkimais A0 ("Telkinio nėra"), A1 ("Matomas") ir A2 ("Dominuojantis"),
- T05 klausimas pakeistas nauju ("Ar matomas galaktikų susijungimas ar potvyninės jėgos poveikis?"),
- T06 klausime, atsižvelgiant į T05 pakeisti galimi atsakymai.

Taip pat palyginus su *Galaxy Zoo 2* padidėjo vidutinis balsų skaičius kiekvieno klausimo atžvilgiu. Vartotojams priskiriami svoriai w ir raudonojo poslinkio paklaidos taisymas atliktas pagal metodus naudotus prieš tai aptartame projekte.

Galaxy Zoo UKIDSS projekte galimiems šališkumams pašalinti buvo naudojami tokie pat metodai kaip ir *Galaxy Zoo2* projekte. GZU projekte nagrinėjamos galaktikos turi tiesioginį atitikmenį

GZ2 duomenų rinkinyje, tačiau šiame projekte galaktikų nuotraukos atkuriamos iš duomenų surinktų artimajame IR diapazone, tuo tarpu GZ2 - regimosios šviesos diapazone.

2. Literatūros apžvalga

2.1. Galaktikų klasifikavimas pasitelkiant *ResNet* konvoliucinių tinklų architektūrą

Jia-Ming Dai ir Jizhou Tong straipsnyje "Galaktikų morfologijos klasifikavimas naudojant giliuosius konvoliucinius neuroninius tinklus" [6] pasitelkia *ResNet* (angl. *Residual Neural Network*) V2 architektūrą siekiant atlikti galaktikų morfologijos klasifikavimą.

Autoriai savo modelį pritaikė 28 790 galaktikų atvaizdų imčiai iš *Galaxy Zoo 2* duomenų rinkinio. Tikslas buvo suklasifikuoti galaktikas į penkias klases: idealiai apvalios, elipsinės, cigaro formos, matomos kampu ir spiralinės. Autoriai savo modelį vertino naudodami įvairius rodiklius, tokius kaip tikslumas, preciziškumas, atpažinimas, F_1 , ROC (angl. *Receiver Operating Characteristic*) kreivės AUC (angl. *Area Under Curve*). Rezultatai parodė, kad jų pasiūlytas tinklas pasiekė geriausius klasifikavimo rezultatus tarp tinklų, pranokdamas Dielemano, AlexNet, VGG, Inception ir kitus *ResNet* tinklus. Bendras jų tinklo klasifikavimo tikslumas testavimo rinkinyje buvo 95,21%. Kiekvienos klasės tikslumas: idealiai apvalios – 96,68%; elipsinės – 94,42%; cigaro formos – 58,62%; matomos kampu – 94,36%; spiralinės – 97,69%.

Nuotraukų duomenų rinkinio apdorojimui pasitelkiamas skalės virpesių metodas (angl. *scale jittering*) [35]. Šiuo pagrindu, nuotraukos apkarpomos skirtingais dydžiais. Atliekamas nuotraukų mastelio keitimas, pasukimas kampu, apvertimas. Papildomai, duomenų apdorojimui pritaikomi optiniai iškraipymai (ryškumo, kontrasto, atspalvio reguliavimas).

Ankščiau minėta modelio architektūra pasižymi rizika, jog kai kurie tinklo blokai mokymo metu turės mažai įtakos galutiniam rezultatui. Atsižvelgiant į tai autoriai pasirinko tinklą išplatinti ir sumažinti jo gylį.

2.2. Tikimybinis *Bayesian* morfologinis klasifikavimas naudojant konvoliucinius neuroninius tinklus

Straipsnyje "Galaxy Zoo: probabilistic morphology through Bayesian CNNs and active learning" ("Galaxy Zoo: tikimybinė morfologija naudojant Bayesian CNN ir aktyvųjį mokymąsi"), paskelbtame žurnale "Monthly Notices of the Royal Astronomical Society", pateikiamas kitoks požiūris į galaktikų morfologijos klasifikavimą. Autoriai Mike Walmsley, Lewis Smith, Chris Lintott, *et. al.* siūlo naudoti Bayesian konvoliucinius neuroninius tinklus (CNN) ir aktyvųjį mokymąsi galaktikų vizualinei morfologijai nustatyti [22].

Autorių metodas pagrįstas tikimybine sistema, kuri leidžia CNN mokytis iš galaktikų atvaizdų su neapibrėžtomis klasėmis ir numatyti kiekvienos galimos klasės tikimybę anksčiau neženklinoms galaktikoms. Šis metodas taikomas *Galaxy Zoo* projektui.

Vienas iš svarbiausių tyrimo rezultatų – gerai sukalibruoti Bayesian CNN sukurti posterioriai. Pavyzdžiui, prognozuojant skersėtumą, modelis pasiekė 11,8 proc. aprėpties paklaidas, neviršijančias 0,21 balsų dalies nuokrypį. Dėl šių gerai sukalibruotų posteriorių rezultatai yra patikimi praktiniam naudojimui.

Kitas svarbus rezultatas – tyrime naudotos aktyvaus mokymosi strategijos BALD (angl. *Bayesian Active Learning by Disagreement*) efektyvumas. Autoriai įrodė, kad jų Bayesian CNN mokymui naudojant aktyvųjį mokymąsi reikia iki 35-60 proc. mažiau paženklintų galaktikų, priklausančių nuo klasifikuojamo morfologinio požymio. Šis rezultatas išryškina aktyvaus mokymosi potencialą didinant galaktikų klasifikavimo efektyvumą.

Tyrime daroma išvada, kad sujungus žmogaus ir mašininį intelektą, *Galaxy Zoo* galės klasifikuoti bet kokio įmanomo masto apžvalgas kelių savačių laiko skalėje, pateikdama didžiulius ir išsamius morfologinius katalogus, padedančius atlikti galaktikų evoliucijos tyrimus. Šis darbas yra svarbus indėlis į galaktikų klasifikavimo sritį, parodantis mašininio mokymosi metodų potencialą siekiant pagerinti mūsų supratimą apie galaktikų morfologiją. Tačiau, svarbu atsižvelgti, jog modelis iš esmės prognozuoja vidutinio vartotojo pasirinkimus pildant klausimyną.

2.3. Konvoliucinių neuroninių tinklų galaktikų klasifikavimo palyginimas keičiant klasių kiekį

Mitchell K Cavanagh, Kenji Bekki ir Brent A Groves straipsnyje "Morphological classification of galaxies with deep learning: comparing 3-way and 4-way CNNs" ("Galaktikų morfologinis klasifikavimas taikant gilųjų mokymąsi: 3 ir 4 kryptų CNN palyginimas"), paskelbtame žurnale "Monthly Notices of the Royal Astronomical Society", pristatomas naujas požiūris į galaktikų morfologinį klasifikavimą taikant gilųjų mokymąsi [7].

Autoriai apmoko ir išbando kelias konvoliucinių neuroninių tinklų (CNN) architektūras, skirtas galaktikų morfologijai klasifikuoti pagal 3 klasių (elipsinės, lęšinės ir spiralės) ir 4 klasių (prieš tai minėtas klases papildant netaisyklingomis galaktikomis) schemas, naudodami 14 034 vizualiai klasifikuotų SDSS vaizdų duomenų rinkinį. Darbe naudojamos keturios skirtingos CNN architektūros:

- C1 – dvejetainiam (dvinariam) klasifikavimui skirta konvoliucinio neuroninio tinklo architektūra, išvesties sluoksnyje naudojant 3 arba 4 išvesties mazgus. Modelio architektūra paremta 2020 metais Cavanagh ir Bekki publikuotu klasifikavimo modeliu, kuris buvo skirtas atlikti galaktikų klasifikavimą į spiralines ir skersėtasias spiralines galaktikas [36].
- Diel – konvoliucinis neuroninis tinklas, paremtas Dielemen, *et. al.* pristatyta architektūra [8]. Tinklo architektūrai taip pat buvo pritaikytas Adadelta adaptyvaus mokymosi greičio algoritmas [37], ir pridėtas rinkinio normalizavimo (angl. *batch normalization*) sluoksnis.
- Alex – tinklo architektūra AlexNet pagrindu, taip pat naudojant anksčiau minėtą Adadelta algoritmą ir normalizavimo sluoksnius.
- C2 – visiškai nauja, šiame darbe pristatyta architektūra, naudojanti Adam algoritmą. Tuo pagrindu, jog mokymosi greitis turi būti konstanta, Adadelta algoritmas nebuvo naudojamas.

Visuose architektūrose buvo naudojama ReLU aktyvacijos funkcija ir SoftMax tikimybinė funkcija. Duomenų apdorojimui buvo panaudoti įprasti pasukimo, apvertimo, apkarpymo veiksmai. C2 architektūra pasiekė 83% (3 klasėms) ir 81% (4 klasėms) bendrą klasifikavimo tikslumą.

Autoriai taip pat palygina visų keturių klasių dvinario klasifikavimo tikslumą, parodydami, kad lengviausiai atskiriamos elipsės ir spiralės (> 98% tikslumas), o sunkiausiai – spiralės ir netaisyklingos galaktikos (78% tikslumas). Be to, autoriai sujungė savo CNN klasifikatorius, kad atliktų hierarchinį pavyzdžių klasifikavimą, ir gavo panašų tikslumą (81%) kaip tiesioginis 3 klasių CNN, tačiau gerokai prastesnį tikslumą 4 klasių atveju (65%). Kaip papildomam tyrimui, tinklai buvo pritaikyti nedidelei *Galaxy Zoo* vaizdų imčiai. Gauti 92%, 82% ir 77% tikslumai atitinkamai dvinariam, trijų klasių ir keturių klasių klasifikavimui.

2.4. *Deep Galaxy* gilusis konvoliucinis neuroninis tinklas

Deep Galaxy straipsnyje [9], autoriai naudoja supaprastintą E. Hubble klasifikavimo metodą: galaktikos klasifikuojamos į tris kategorijas: elipsines, spiralines ir netaisyklingas galaktikas. Tinklas sukurtas naudojant 8 sluoksnius, pagrindinis konvoliucinis sluoksnis susideda iš 96 filtrų, bruožų išgavimui. Autoriai tinklą apmokė naudodami 1356 nuotraukų duomenų rinkinį, padalintus į apmokymo, testavimo ir validavimo grupes, santykiškai atitinkamai 60%, 21% ir 19%. Svarbu paminėti, kad nuotraukų pasiskirstymas tarp trijų kategorijų gana daug skiriasi, pavyzdžiui: netaisyklingųjų galaktikų tarp apmokymo paveikslukų yra 2–3 kartus mažiau nei elipsinių ar spiralinių galaktikų. Autorių sukurtas tinklas pasiekė 97.27% tikslumą.

2.5. *Xception* neuroninio tinklo panaudojimas galaktikų klasifikavimui

Khramtsov, Dobrycheva, *et. al.* straipsnyje pristatomas CNN tinklas skirtas klasifikuoti galaktikas į elipsines ir spiralines galaktikas [10].

Darbe buvo naudojamas ~ 86000 galaktikų atvaizdų rinkinys sudarytas iš SDSS DR9 duomenų ir duomenų naudotų praėjusiame autorių tyrime, bei pasitelkiant *Galaxy Zoo* savanorių pateiktus duomenis. Naudojami paveikslai buvo sudėtiniai $g - r - i$ filtrų vaizdai. Klasių pasiskirstymas buvo gana netolygus – elipsinių galaktikų buvo apie 4 kartus daugiau. Modeliui buvo pasirinkta *Xception* architektūra [38], naudojanti vieną filtrą vienam įvesties kanalui, ir taškinę konvoliuciją.

Gauti rezultatai buvo palyginti su rezultatais gautais naudojant SVM mašininio mokymo metodą. Buvo pastebėta, jog elipsinės galaktikos buvo klasifikuojamos prasčiau, darant prielaidą, jog modelis neišskiria spektrinių bruožų, geriau negu SVM, apmokytas fotometriniais duomenimis. Taip pat buvo pastebėta problema dėl galaktikos plokštumos pozicijos, regėjimo linijos atžvilgiu.

2.6. ViT panaudojimas morfologiniam galaktikų klasifikavimui

Straipsyje "Galaxy Morphological Classification with Efficient Vision Transformer" ("Galaktikos morfologinis klasifikavimas su efektyviu regos transformatoriumi") [39] autoriai pristato ViT panaudojimą galaktikų klasifikavimo uždaviniui spręsti.

Tyrime naudojamas *Linformer* modelis – ViT modelio variantas, kuriame dėmesio mechanizmo MHA kompleksiskumas yra tiesinis. Tinklo apmokymui panaudojamas *Galaxy Zoo 2* duomenų rinkinys, išskiriant 8 klases. Nuotraukos papildomai apdorojamos – sumažinama jų skiriamoji geba, nuotraukos apverčiamos ir pasukamos, pritaikomas pikselių normalizavimas kiekvienam kanalui ir duomenų rinkinys normalizuojamas pagal standartinį nuokrypį.

Bendras tinklo tikslumas 80,55%, tačiau tikslumas kiekvienos klasės atžvilgiu žymiai svyruoja (tarp 40% ir 90% naudojant tolygius klasių svorius nuostolių funkcijai, ir tarp 60% ir 87% kai svoriai sureguliuoti). Autoriai tokius žymius skirtumus pagrindžia turimų duomenų pasiskirstymo per visas klases netolygumams, kas sprendžiant galaktikų morfologinio klasifikavimo uždavinį dažnu atveju yra matoma ties konkrečiomis klasėmis. ViT įvertinimui naudotas *ResNet-50* konvoliucinis neuroninis tinklas. Apskritai ViT ir *ResNet* modelių tikslumo vertės panašios ir ViT nėra pranašesnis už konvoliucinį neuroninį tinklą, tačiau analizės metu, autoriai pastebėjo, jog ViT geba geriau klasifikuoti mažesnes ir silpnesnes galaktikas, kurias bendru atveju klasifikuoti sudėtingiau, nes šios nuotraukos triukšmingesnės.

2.7. Galaktikų klasifikavimas pasitelkiant CvT

Straipsnyje "Galaxy morphology classification based on Convolutional vision Transformer (CvT)" ("Galaktikų morfologijos klasifikavimas remiantis konvoliuciniu regos transformatoriumi (CvT)") Jie Cao *et. al.* pristato morfologinių galaktikų klasifikavimo uždavinio sprendimą pasitelkiant CvT [40].

Darbe naudojami *Galaxy Zoo 2*, bei *Galaxy Zoo CANDELS* duomenų rinkiniai, o galaktikos klasifikuojamos į 5 klases. Pasirinkimas naudoti šiuos du duomenų rinkinius buvo padarytas tam, jog tinklas būtų apmokomas tiek prie mažų, tiek prie didelių raudonojo poslinkio verčių. Klasių disbalansui išspręsti buvo įvesta dirbtinių pavyzdžių, pasukant esamas galaktikas keturiais skirtingais kampais. Tyrimui atlikti buvo pasirinktas CvT-13 modelis. Apmokius tinklą klasifikavimo tikslumas virš 98%, tuo tarpu kiti rodikliai (F1, preciziškumas, atkūrimas) – virš 99 %, neskaitant cigaro formos elipsinių galaktikų, kadangi šioje klasėje duomenų kiekis buvo mažesnis. Taip pat autoriai išskiria, jog klaidos šios klasės išskyrimo gali būti nulemtos pačių galaktikų išvaizdos – kadangi galaktika neturi aiškių bruožų ir visuomet yra lygios formos, bei pačio duomenų rinkinio (dėl galimų balsuotojų klaidų). Taip pat, CvT-13 buvo palygintas su CvT-21 ir CvT-W24, tačiau buvo pastebėta, jog modelio klasifikavimo įverčiai skyrėsi nežymiai, tačiau kitiems modeliams reikėjo daugiau resursų. Papildomai buvo atliktas tyrimas palyginant modelio tikslumą kai nuotraukos papildomos skirtingu kiekiu triukšmų. Didinant triukšmų kiekį modelio tikslumas mažėjo, tačiau ne vienu atveju nebuvo žemesnis nei 80 %. Galaktikos su aukštu raudonuoju poslinkiu taip pat buvo klasifikuojamos gerai – bendras tikslumas apie 98%.

2.8. Konvoliucinis neuroninis tinklas galaktikų klasifikavimui

R. Urechiatu ir M. Frincu straipsnyje "Improved Galaxy Morphology Classification with Convolutional Neural Networks" ("Patobulinta galaktikų morfologijos klasifikacija naudojant konvoliucinius neuroninius tinklus") pristatomas CNN tinklas, kuris galaktikas klasifikuoja į penkias galimas klases – apvalias elipsines, elipsines, cigaro formos elipsines, spiralines ir galaktikas matomas kitu kampu (angl. *edge-on*) [41]. CNN tinklas, buvo apmokytas naudojant 10 000 galaktikų paveikslų iš *Galaxy Zoo 2* projekto.

CNN modelio įvestis – $1 \times 64 \times 64$ paveikslai. Architektūra sudaryta iš keturių nuoseklių konvoliucinių blokų, kuriuose filtrų kiekis keičiamas nuo 64 iki 512 (2^n , kur $n = 6, 7, 8, 9$). Tinkle, netiesiškumų įvedimui buvo naudojama ReLU aktyvacijos funkcija, išmetimas nuo antrojo sluoksnio 20%, ir 50% po trečiojo. Optimizavimui naudojamas *Adam* algoritmas, mokymosi greitis $lr = 0.001$, paketų dydis – 16, epochų skaičius – 12 (naudojant ankstyvąjį sustabdymą).

Naudojami paveikslai buvo apdoroti, pirmiausia sumažinant spalvų kanalus (paveikslo gylį) – toliau naudojant nespaltotas nuotraukas. Papildomai atliktas spalvų kanalo normalizavimas [0, 1]. Autoriai pastebėjo, kad normalizavimo operacija, modelio tikslumą pagerino apie 10 %. Prieš pakeičiant paveikslų skiriamąją gebą (jog ji atitiktų modelio įvestį), paveikslai buvo papildomai apkarpyti, taip sumažinant vaizdinius triukšmus (kitus matomus dangaus kūnus). Kadangi naudojamuose duomenyse matomas netolygus pavyzdžių pasiskirstymas tarp apibrėžtų klasių, papildomai buvo atliktas dirbtinių pavyzdžių įvedimas. Turimus paveikslus pasukant pagal laikrodžio rodyklę kas 30° laipsnių ir taip sugeneruojant po dvyliką papildomų pavyzdžių kiekvienam originaliam paveiksliui.

Kitas svarbus veiksnys, galimai darantis įtaką modeliui yra tai, jog net ir papildomai apkarpius paveikslėlius, arčiau esantys kiti dangaus kūnai taip pat gali būti matomi nuotraukose. Atsižvelgiant

į tai, bei į atlikta dirbtinių pavyzdžių generavimą, išlieka tikimybė, jog tinklas išmoks ne galaktikai priklausančius bruožus (t.y. atsižvelgs į nuotraukose matomus triukšmus).

Bendras modelio tikslumas 96.83%, preciziškumas – 96.75%, atkūrimas 96.52%, F1 – 96.72%. Autoriai pabrėžė keletą palyginimų, atkreipdami dėmesį, kad siūlomas CNN modelis pranoko pažangius analogus, tokius kaip ResNet-50 (94,68 % tikslumas), DenseNet (94,73 % tikslumas), Inception (93,8 % tikslumas), EfficientNet (92,89 % tikslumas) ir MobileNet (94,56 % tikslumas), kai visi jie buvo mokomi su tais pačiais hiperparametrais. Papildomai buvo atliktas modelio palyginimas su ankščiau minėtais pavyzdiniais modeliais, palyginant prognozės patikimumą (angl. *prediction confidence*) kurio vertė gali būti tarp 0 ir 1. Buvo pastebėta, jog jų pasiūlytas modelis turėjo didžiausią kiekį atvejų, kai prognozės patikimumas buvo aukščiausias (2465 iš 4511 įvertintų pavyzdžių).

2.9. Galaktikų klasifikavimas naudojant jungtinį *GalaxyZoo* duomenų rinkinį

T. Mūžo darbe "Didelės apimties galaktikų nuotraukų klasifikavimas į spiralines ir elipsines naudojant jungtinį *GalaxyZoo* duomenų rinkinį" (angl. *Large Scale Galaxy Image Classification Into Spiral And Elliptical Using Combined GalaxyZoo Dataset*) galaktikos klasifikuojamos į dvi klases. Autorius naudoja *Galaxy Zoo* pirmojo ir antrojo projektų duomenis (apie 316 000 galaktikų nuotraukų). Šiame darbe galaktikų klasifikavimo uždaviniui pasitelkiamos penkios skirtingos tinklų architektūros, o iš trijų geriausių – sukurtas jungtinis bendras modelis (angl. *ensemble*).

Autoriaus darbe, klasifikavimo uždavinys susitelkia į modelio pritaikomumą plačiam diapazonui galaktikų raudonojo poslinkio ir ryškumo atžvilgiu. Raudonojo poslinkio ribos šiame darbe $z \geq 0.35$. Modelių įvesčiai naudojami 128×128 dydžio paveikslai. Tyrime šiai klasifikavimo problemai buvo pritaikyti ir ištestuoti penki modeliai: *ResNet50*, *InceptionV3*, *InceptionResNetV2*, *Dieleman*, *C2*. Nuotraukos taip pat buvo papildomai apdorotos, ir autorius pastebi jog taikant papildomas transformacijas abiemis klasėms (o ne tik klasei kurios pavyzdžių mažiau) pastebimi geresni rezultatai.

Buvo pastebėta, jog *Dieleman* ir *C2* architektūros, pasiekė apie 95.57% tikslumą ir pralenkė gilesnias modelių architektūras tirtas šiame darbe. Taip pat, atsižvelgiant į palygintų modelių vertinimus, geriausi trys modeliai (*ResNet50*, *Dieleman* ir *C2*) buvo apjungti kartu, naudojant daugumos balso principą. Jungtinio modelio pasiektas tikslumas – 95.70% tikslumą. Taip pat autorius išskyrė modelio vertinimus pagal raudonojo poslinkio vertes:

- $0 \leq z < 0.1$ – 96.58%
- $0.1 \leq z < 0.25$ – 94.56%
- $0.25 \leq z < 0.35$ – 99.35%

buvo pastebėta jog, klasifikavimo klaidos dažniau pastebimos neryškiuose spiralinėse galaktikose ir ryškesnėse elipsinėse galaktikose, o galaktikos su neįprastais bruožais nebuvo pirminė klaidingo klasifikavimo priežastis. Tyrimo metu buvo pastebėta, jog naudojant nuotraukas su vienu spalvos kanalu (angl. *grayscale*) klasifikavimo tikslumas sumažėja apie 0.8 – 1%.

3. Mašininio mokymosi teorijos apžvalga

Šiame skyriuje pristatomi dirbtiniai neuroniniai tinklai (ANN), konvoliuciniai neuroniniai tinklai (CNN), transformatoriai, regos transformatoriai bei konvoliuciniai regos transformatoriai, jų sandara, veikimo principas, naudojami algoritmai, dažniausiai išskylančios problemos bei jų sprendimo būdai, modelio vertinimo metodikos ir parametrai.

3.1. Giliojo mokymo principai

Giliojo mokymo modeliai, yra grindžiami bendromis mokymo ir optimizavimo technikomis, tokiomis kaip aktyvacijos funkcija, reguliarizavimo metodai, optimizavimo technikos ir pan., kurie užtikrina stabilų mokymosi procesą, sumažina perteklinio pritaikymo riziką ir pagerina modelio gebėjimą "suprasti" nematytus duomenis.

Aktyvacijos funkcija naudojama įvesti netiesiškumą, taip įgalinant sudėtingų netiesinių ryšių modeliavimą. Dažniausiai naudojama ReLU aktyvacijos funkcija:

$$f(x) = \max(0, x), \quad (3.1)$$

kurios išvestis bus lygi įvesčiai x , jei $x > 0$, kitu atveju $f(x) = 0$. Esant poreikiui išlaikyti neigiamas įvestis, naudojama *Leaky ReLU* funkcija:

$$f(x) = \max(\alpha x, x), \quad (3.2)$$

kurioje $x \leq 0$ atveju įvestis padauginama iš koeficiento α .

Reguliarizacija naudojama siekiant išvengti perteklinio pritaikymo (angl. *overfitting*). Dažnai naudojamos L1 (*Lasso*) arba L2 (*Ridge*) reguliarizacijos funkcijos, kurios, atitinkamai, nuostolių funkciją (angl. *loss function*) papildo "nuobaudos" nariu proporcingu absoliučioms modelio svorių vertėms arba modelio svorių kvadratui. Lasso reguliarizacija (L1):

$$J(\mathbf{w}) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^d |w_j|, \quad (3.3)$$

kur $\mathbf{w} = (w_1, w_2, \dots, w_d)$, $\hat{y}_i$ – i -toji prognozuojama vertė, y_i – i -toji tikroji vertė, λ – reguliarizacijos parametras. L1 reguliarizacija nuostolių funkcijoje svorius sumažina iki nulio, kas ypač naudinga bruožų parinkimui, kitaip tariant nereikšmingi bruožai pašalinami. Tuo tarpu Ridge reguliarizacija:

$$J(\mathbf{w}) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^d w_j^2 \quad (3.4)$$

Šiuo atveju L2 į nuostolių funkciją įtraukia baudos narį, pagrįstą svorių kvadratinio dydžiu. Didėjant λ nuobaudos svorio koeficientui yra žymi, todėl modelis mažiau prisitaiko prie mokymo duomenų, bet geriau apibendrina nematytus naujus duomenis. Kitaip nei L1 atveju svoriai nesumažinami iki nulio, dėl to modelis gali būti sunkiau paaiškinamas.

Taip pat, siekiant išvengti perteklinio prisitaikymo, plačiai naudojama reguliarizavimo technika – išmetimas (angl. *dropout*). Mokymo proceso metu atsitiktinai "išmetami" tam tikri neuronai, taip užtikrinant, jog modelis perdėtai nesiremtų tam tikru bruožu. Paprasčiausiu atveju, jei išmetimo parametras $p = 0.2$, tai reiškia, kad 80% neuronų bus aktyvūs, o likę 20% – ne. Išmesti neuronai turi nulinę aktyvaciją, o aktyvūs neuronai turi pakoreguotą aktyvaciją, kad išlaikytų bendrą mokymo procesą.

Optimizavimo algoritmai yra naudojami koreguoti modelio parametrus, siekiant sumažinti nuostolių funkciją. Dažnai naudojamas *Adam* algoritmas, kuris apjungia stochastinio optimizavimo (angl. *stochastic optimization*) ir adaptyvaus mokymosi greičio (angl. *adaptive learning rate*) algoritmus. Metodas naudoja eksponentinius svertinius slenkančius vidurkius (angl. *exponentially weighted moving averages*) [42], kad apskaičiuotų pirmąjį pagreitį (gradiento slenkančio vidurkio mažėjimo greitį):

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot \text{grad}(\theta_{t-1}) \quad (3.5)$$

ir antrąjį pagreitį (gradiento kvadrato slenkančio vidurkio mažėjimo greitį):

$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot [\text{grad}(\theta_{t-1})]^2, \quad (3.6)$$

kur β – svorio parametrai, θ – modelio parametrai. Tuomet pagreičiai atnaujinami atsižvelgiant į nuokrypį (angl. *bias*):

$$\begin{aligned} \hat{m}_t &= \frac{m_t}{(1 - \beta_1^t)} \\ \hat{v}_t &= \frac{v_t}{(1 - \beta_2^t)} \end{aligned} \quad (3.7)$$

Galiausiai, atnaujinami modelio parametrai:

$$\theta_t := \theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad (3.8)$$

Standartinis *Adam* optimizatorius, reguliarizaciją taiko nuostolių funkcijoje, ko pasėkoje tai tiesiogiai paveikia mokymosi spartos pokyčius. Norint to išvengti, naudojamas *AdamW* algoritmas, kuris atskiria modelio svorių atnaujinimus nuo modelio parametrų atnaujinimų.

Be adaptivių optimizatorių (pvz.: *Adam*), SGD (angl. *Stochastic Gradient Descent*) su pagreičiu tebėra plačiai naudojamas, ypač konvoliuciniuose neuroniniuose tinkluose. Esminis skirtumas tarp SGD ir GD (angl. *Gradient Descent*), yra tai, jog SGD modelio svorius atnaujinama atsižvelgiant į vieną atsitiktinį mokymo pavyzdį tam tikru laiko momentu, tuo tarpu GD tam naudoja visą mokymosi duomenų rinkinį. Pagreitis esamame žingsnyje prideda dalį praėjusio žingsnio pagreičio, taip sumažinant svyravimus ir padedant modeliui nesustoti ties lokaliu minimumu:

$$v_t = \beta v_{t-1} + (1 - \beta) \nabla L(\theta). \quad (3.9)$$

Norint stabilizuoti mokymosi procesą, naudojami normalizacijos metodai, kurie padeda stabilizuoti gradientus ir pagerinti konvergavimą. Vienas tokių metodų, paketo (angl. *batch*) normalizavimas, skirtas stabilizuoti mokymosi procesą, apskaičiuojant vidurkį kiekvienai įvesčiai (angl. *mini-batch*) $B = \{x_1, \dots, x_n\}$:

$$\text{BN}(x) = \gamma \odot \frac{x - \hat{\mu}_B}{\hat{\sigma}_B} + \beta, \quad (3.10)$$

kur $\hat{\mu}_B$ - B vidurkis, $\hat{\sigma}_B$ - B standartinis nuokrypis, o γ , β – atitinkamai mastelio (angl. *scale*) ir poslinkio parametrai.

Sluoksnio normalizavimas dažniausiai naudojamas Transformatoriaus architektūroje, kur paketo normalizavimas nėra tinkamas dėl sekų ilgio ir pasikeičiančios įvesties struktūros:

$$\text{LN}(x) = \gamma \cdot \frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta. \quad (3.11)$$

Nuostolių funkcijos apibrėžia optimizavimo tikslą ir kiekybiškai įvertina neatitikimą tarp modelio prognozių ir tikrųjų duomenų klasių. Esant netolygiam duomenų pasiskirstymui, įprasta CEL *Cross Entropy Loss* funkcija gali būti šališka klasės kurios pavyzdžių yra daugiausia atžvilgiu. Siekiant sumažinti tokį poveikį, įvedami papildomi svorių koeficientai kiekvienai klasei w_c . Kita svarbi technika naudojama nuostolių funkcijoje – išlyginimas (angl. *label smoothing*), kuri tikrųjų klasių skaitinę reprezentaciją "suprastina" (t.y. dviejų klasių atveju turint klasę 1, jos skaitinė vertė sumažinama ir nedidelė vertė priskiriama klasei 0), taip sumažinant pernelyg užtikrintus modelio spėjimus (mažiau informacijos įsimena, modelis labiau likęs apibendrinti (angl. *generalize*)).

Nors *Cross Entropy Loss* funkcija leidžia įskaityti netolygų duomenų pasiskirstymą tarp klasių, visi pavyzdžiai klasėse yra vertinami vienodai. Šiai problemai spręsti reikalingas papildomas mechanizmas kuris padėtų atkreipti dėmesį į sudėtingus pavyzdžius. Tam dažnu atveju naudojama *Focal Loss* [43]. Ši funkcija papildomai leidžia įvesti koeficientus ne tik netolygiam duomenų pasiskirstymui tarp klasių, bet ir prideda papildomą narį kuris didesnę svarbą suteikia prasčiau klasifikuojamiems pavyzdžiams. *Focal Loss* funkcija iš esmės papildo *Cross Entropy* funkciją papildomu nariu γ :

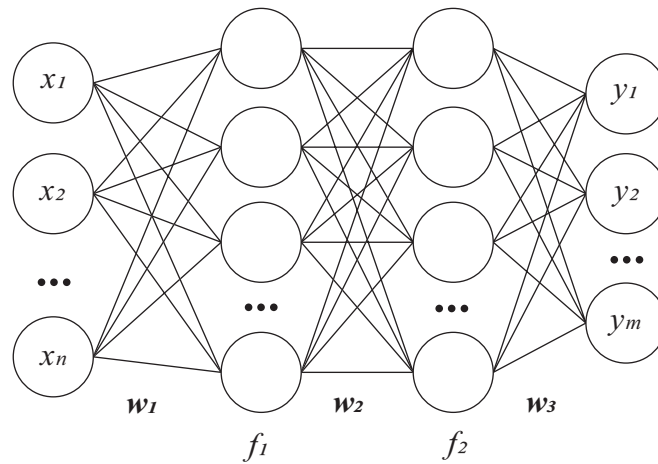
$$\text{FL}(p_t) = -\alpha_t(1 - p_t)^\gamma \log(P_t), \quad (3.12)$$

kur α_t - svorio koeficientas (klasių disbalansui), γ - "fokusavimo" koeficientas. Lygtyje 3.12 $(1 - p_t)^\gamma$ narys sumažina "lengvų" pavyzdžių įtaką.

Apibendrinant, šios bendrosios giliojo mokymo technikos sudaro pamatą įvairioms modelių architektūroms, užtikrinantį efektyvų modelio mokymą.

3.2. Dirbtiniai neuroniniai tinklai

Dirbtiniai neuroniniai tinklai (angl. *Artificial Neural Network*, ANN) – mašininio mokymosi metodas, imituojantis gyvų organizmų smegenyse vykstantį elektroninių signalų perdavimą, kuris padeda apdoroti informaciją. ANN tinklo mažiausias nedalomas matas – dirbtinis neuronas, kuriuos jungiant sudaroma tinklo struktūra (3 pav.). Tinklai gali varijuoti įvesčių, išvesčių, paslėptų sluoksnių kiekiu ar kitais panašiais parametrais.

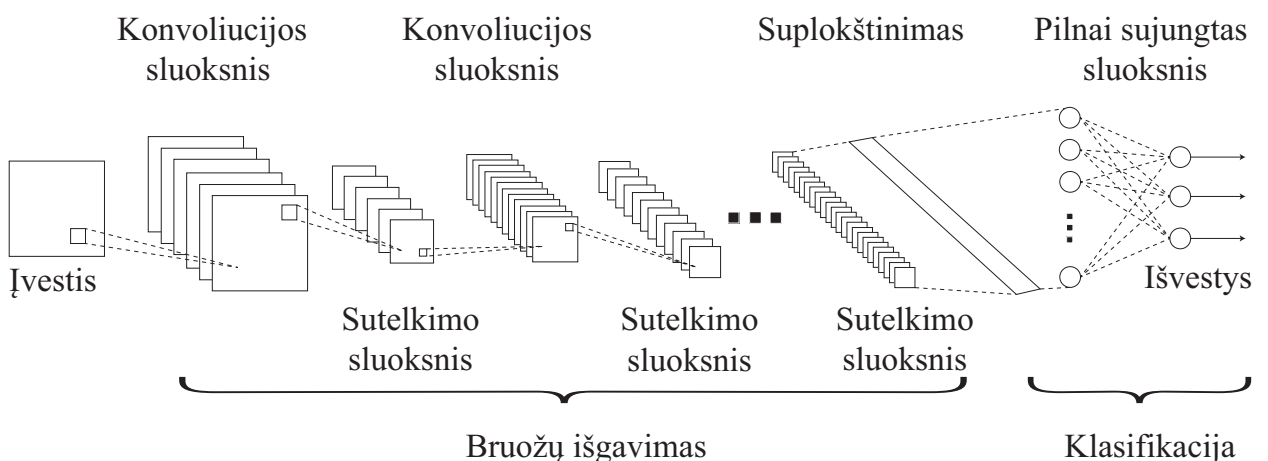


3 pav. Dirbtinio neuroninio tinklo struktūros principinė diagrama, $x_1, \dots, x_n$ – įvestys, $w_1, \dots, w_3$ – svorių koeficientai, $y_1, \dots, y_m$ – išvestys.

3.3. Konvoliuciniai neuroniniai tinklai

Konvoliucinis neuroninis tinklas (angl. *Convolutional Neural Network*, CNN) – giliojo mokymosi tinklo architektūra, dirbtinių neuroninių tinklų (angl. *Artificial Neural Network*, ANN) porūšis. CNN dažniausiai naudojami vaizdų atpažinimui. Skirtingai nei tradicinė ANN architektūra, CNN tinklai, išlaiko 2D struktūrą ir filtro pagalba (angl. *kernel*) sumažina naudojamų bruožų (angl. *features*) kiekį, dėl to skaičiavimai tampa greitesni ir efektyvesni. Įprastai CNN įvestį sudaro trečios (ar aukštesnės) eilės tenzorius, kuris nuosekliai apdorojamas keliais skirtingai etapais (apdorojimo etapas atitinka vieną tinklo sluoksnį) [44].

4 paveiksle pateikiamas apibendrintas CNN architektūros grafinis atvaizdavimas. Tinklą sudaro dvi pagrindinės dalys: bruožų išgavimo ir klasifikacijos. Pirmoji dalis susideda iš n konvoliucinių ir sutelkimo sluoksnių, bruožų išgavimo etapo išvestis suplokštinama ir paduodama į antrąjį etapą – klasifikaciją, kurioje yra pilnai sujungtas sluoksnis (angl. *Fully Connected Layer*).



4 pav. Konvoliucinio neuroninio tinklo architektūros grafinis atvaizdavimas.

Konvoliucinis sluoksnis atlieka konvoliuciją tarp įvesties ir filtro (angl. *kernel*), kuris dažniausiai yra mažesnis už įvesties erdvinius matmenis. Sutelkimo sluoksnis iš esmės atlieka įvesties apibendrinimą pagal pasirinktą metodą (pavyzdžiui, pagal vidurkį ar pagal maksimumą), taip pat atliekant operaciją su įvestimi ir apibrėžtu filtru. Suplokštinimo sluoksnis pertvarko duomenų struktūrą,

pateikdamas gautas vertes vienmatėje struktūroje.

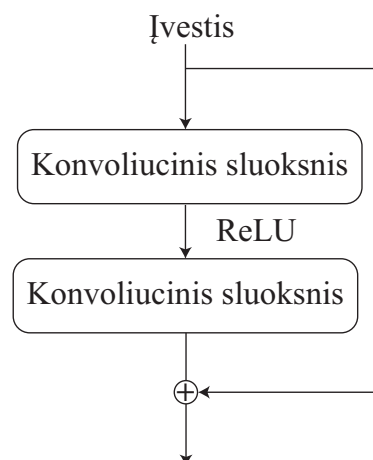
Neuronai konvoliuciniuose sluoksniuose aprašomi trejais kanalais: aukštis, plotis ir gylis. Konvoliuciniame sluoksnyje, neuronai jungiasi tik su tam tikra įvesties dalimi, o pilnai sujungtas sluoksnis naudojamas tik klasifikavimo dalyje. Esminis CNN tinklų architektūros bruožas – filtras, kuris "pereidamas" per pateikta paveikslą sudaro aktyvacijos žemėlapi (angl. *activation map*). Aktyvacijos žemėlapis parodo tam tikrus aptiktus požymius (pirminiuose sluoksniuose aptinkant primityvius bruožus, tuo tarpu vėlesniuose sluoksniuose apjungiant juos ir taip suformuojant kompleksinius bruožus).

Sutelkimo sluoksniai, kuriuose naudojamas filtras reliatyviai mažesnis, aktyvacijos žemėlapius sumažina. Įprastai, šiuose sluoksniuose naudojamos *Max Pooling* arba *Average Pooling* sutelkimo operacijos.

3.4. Likutiniai neuroniniai tinklai

Likutiniai neuroniniai tinklai (angl. *Residual Neural Networks, ResNet*) buvo pristatyti palengvinti gilių tinklų mokymą [45]. Giliuose tinkluose mokymo metu gradientui keliaujant per daug sluoksnių jis nežymiai atnaujinamas žemesniuose sluoksniuose, kas apriboja modelio apmokymą. Ši problema žinoma kaip nykstančio gradiento problema (angl. *Vanishing Gradient Problem*). Taip pat, kai neuroniniai tinklai tampa gilesni, mokymo paklaida didėja, nors pridedama daugiau sluoksnių. Tai prieštarauja logikai, nes teoriškai sluoksnių pridėjimas turėtų padidinti modelio pajėgumą ir pagerinti jo veikimą. Šis reiškinys vadinamas degradacijos problema (angl. *Degradation Problem*) ir atsiranda dėl to, kad labai gilius tinklus sunku optimizuoti. Likutiniai neuroniniai tinklai sprendžia šias problemas, įvedant likutinio mokymosi koncepciją. Užuo tiesiogiai mokęsis norimo atvaizdavimo, tinklas išmoksta likutį (t. y. skirtumą tarp norimos išvesties ir įvesties) ir prideda jį prie įvesties, kad būtų lengviau optimizuoti.

Pagrindinė idėja yra praleidimo jungtis (dar vadinama tapatybės sutrumpinimo jungtimi, angl. *skipped connection*), kuri suteikia galimybę grįžtamojo sklido metu gradientui eiti tiesioginiais keliais. Šios jungtys apeina vieną ar daugiau sluoksnių ir leidžia įvestį pridėti prie konvoliucinio bloko išvesties. Esminis *ResNet* bruožas – likutiniai blokai (angl. *residual block*), kuriuose naudojama praleidimo jungtis vienam ar daugiau sluoksnių. Principinė likutinio bloko diagrama pateikta 5 paveiksle.



5 pav. Likutinio Neuroninio tinklo *ResNet* principinė diagrama.

Įprastai, *ResNet* likutiniuose blokuose yra bent du konvoliuciniai blokai, kuriuose naudojamos

3×3 konvoliucijos, po kiekvieno iš šių sluoksnių seka paketų normalizavimas (angl. *batch normalization*), aktyvacijos funkcija.

3.5. Transformatoriai

2017 m. pristatyta Transformatoriaus (angl. *Transformer*) architektūra, kuri pagrįsta dėmesio mechanizmais [46]. Architektūra sudaryta iš koduotuvo *Encoder* ir dekoderio *Decoder*.

Koduotuvą sudaryta iš N identiškų sluoksnių. Kiekvienas iš šių sluoksnių sudarytas iš:

- MHA (angl. *Multi-Head Attention Mechanism*) – sudaryto iš kelių paralelių dėmesio sluoksnių (angl. *attention layers*) – tokiu būdu vietoje vieno Q , K , V matricių rinkinio, įvestis yra aprašoma keliais atskirais rinkiniais, kurie vėliau apjungiami ir panaudojami gauti galutinėms vertėms.
- FNN (angl. *Feed Forward Network*, *FNN*).

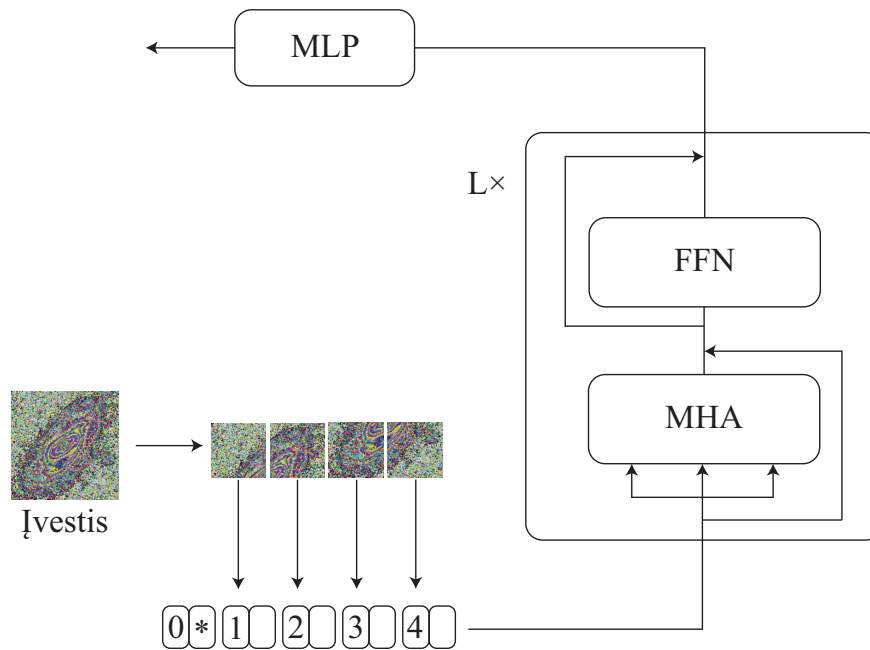
Dekodavimo dalis taip pat sudaryta iš N sluoksnių, kurie susideda iš trijų sluoksnių – du iš jų tokie kaip ir koduotuve (tačiau dėmesio sluoksnyje naudojamas maskavimas taip užtikrinant, jog prognozavimas pozicijai i atsižvelgtų tik į jau žinomas išvestis), o trečiasis – koduotuvo-dekoderio dėmesio sluoksnis.

Tarp prieš tai aprašytų sluoksnių koduotuve ir dekoderiye naudojamos likutinės jungtys (angl. *residual connection*) ir atliekamas normalizavimas. Transformatoriaus architektūroje, prie įvesties pridedama informacija apie žetono (angl. *token*) poziciją.

3.6. Regos transformatoriai

Prieš tai pateikta Transformatoriaus architektūra (3.5 poskyris.) buvo pritaikyta vaizdų klasifikavimo uždaviniui [47] ir vėliau pritaikyta kitiems su vaizdais susijusiems uždaviniais [48], [49], [50]. [47].

Regos transformatoriaus modelio (angl. *Vision Transformer*, *ViT*) įvesčiai paveikslas yra formuojamas į suplokštintų N fragmentų (angl. *patches*) seką, kuri yra apdorojama (angl. *patch embedding*) ir atvaizduojama į D dimensijos erdvę, prijungiant fragmentų pozicijos informaciją (angl. *positional embedding*) ir apdorojamos nuotraukos klasę. Kadangi modelis naudojamas klasifikavimo uždaviniui, ViT architektūroje yra tik koduotuvai. Regos transformatoriaus principinė diagrama pateikiama žemiau – 6 paveiksle. Į koduotuvą pateikiama $[N, D]$ formos vektorių seka.



6 pav. Rego transformatoriaus ViT principinė diagrama.

Dėmesio mechanizmo (angl. *multi-head self-attention mechanism*) įvestis susideda iš d_k dimensijos užklausų $\mathbf{Q}$ ir raktų $\mathbf{K}$, bei d_v dimensijos verčių $\mathbf{V}$. Dėmesio funkcija apskaičiuoja įvesties vektorių svertinę sumą pasiremiant užklausos ir raktų skaliarine sandauga padalinta iš $\sqrt{d_k}$ ir panaudojant *SoftMax* funkciją. Dėmesio funkcija skaičiuojama lygiagrečiai panaudojant kelis dėmesio mechanizmo sluoksnius. Dėmesio mechanizmo išvestis pridedama prie pirminės įvesties (panaudojant likutinį ryšį). Šis rezultatas paduodamas į normalizavimo sluoksnį.

Sekančiame žingsnyje įvestis paduodama į tiesioginio perdavimo pilnai sujungtą tinklą (angl. *position-wise fully connected feed-forward network*). FFN (angl. *feed-forward network*) sluoksnis sudarytas iš dviejų pilnai sujungtų sluoksnių ir *ReLU* aktyvacijos funkcijos viduryje. Kiekviena įvestis apdorojama atskirai.

Kai įvestis pereina per visus koduotuvo sluoksnius, klasifikacijos uždaviniui toliau panaudojama išvesties dalis, atitinkanti CLS žetoną (pridedant pilnai sujungtą sluoksnį klasifikavimo atlikimui).

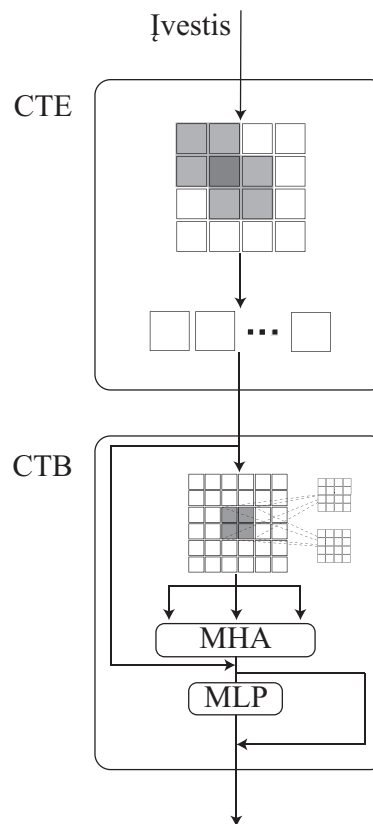
Esminis ViT trūkumas yra tai, jog norint pasiekti konkurencingą našumą, modelis turi būti apmokomas itin dideliu duomenų kiekiu [47]. Šiai problemai spręsti buvo pasiūlytas DeiT (angl. *Data-Efficient Image Transformer, DeiT*) modelis [51], kurio esminis skirtumas yra mokymo strategijoje (detalesnis nuotraukų apdorojimas, išsamesnis hiperparametrų derinimas). Esminis architektūros skirtumas nuo ViT yra papildomo distiliavimo žetono (angl. *distillation token*) įvedimas.

3.7. Konvoliuciniai regos transformatoriai

Kadangi ViT tinkluose iš esmės nėra mechanizmų kurie įvertintu erdvės įtaką, tuo tarpu CNN architektūra geriau išskiria lokalias struktūras 2D erdvėje (šalia esantys pikseliai dažnu atveju turi aukštą koreliaciją), o filtrų pagalba tinklas išmoka vizualinius motyvus atsižvelgiant į erdvės kontekstą, 2021 m. Haiping Wu, *et. al.* pristatė naują regos transformatorių modelį – Konvoliuciniai regos transformatoriai (angl. *Convolutional Vision Transformer, CvT* [52]).

Pirmasis esminis skirtumas tarp CvT ir ViT pastebimas įvesties paruošime. 7 paveiksle pateikiama principinė CvT diagrama. Nuotrauka apdorojama CTE sluoksnio (angl. *Convolutional*

Token Embedding), kuriame atliekama konvoliucijos operacija su persidengiančiais dėmenimis, toliau seka normalizavimo sluoksnis. Toks apdorojimas sumažina žetonų skaičių ir padidina jų plotį (bruožų dimensiją). Taip pat svarbu paminėti, jog CvT nebenaudoja pozicijos įterpimo, kadangi erdvinė informacija išlaikoma dėl konvoliucijos operacijų.



7 pav. Konvoliucinio regos transformatoriaus CvT principinė diagrama.

Toliau išvestis paduodama į CTB sluoksnį (angl. *Convolutional Transformer Block*). Pirmiausia žetonai pertvarkomi į 2D erdvinį tinklą, tuomet seka konvoliucijos operacija (angl. *depth-wise separable convolution*), padalinta į dvi operacijas: erdvinę konvoliuciją (angl. *depth-wise convolution*), kuri atliekama kiekvienam kanalui atskirai, ir taškinę konvoliuciją (angl. *point-wise convolution*). Tuomet išvestis paduodama į MHA, $\mathbf{Q}$, $\mathbf{K}$, $\mathbf{V}$ formatu. Galiausiai, klasifikavimo žetonas (CLS) pridedamas tik paskutiniame etape, o MLP sluoksnis naudoja šio žetono informaciją prognozuojant klasę. Toks modelio realizavimas pagerina erdvinį kontekstą ir sumažina skaičiavimo kaštus, kadangi $\mathbf{K}$ ir $\mathbf{V}$ vektoriai sumažinami prieš MHA.

3.8. Hiperparametrų derinimas

Hiperparametrų derinimas yra svarbus mašininio mokymosi modelių optimizavimo aspektas, turintis tiesioginę įtaką jų našumui ir efektyvumui. Pagrindiniai hiperparametrai apima epochų skaičių, mokymosi greitį, paketo dydį, regularizacijos koeficientus, išmetimo parametrus ir t.t.

Epochų skaičius lemia kiek kartų bus pereita per mokymo duomenų rinkinį. Efektyvus epochų skaičiaus pasirinkimas yra dažnai derinamas su ankstyvo sustabdymo mechanizmu (angl. *early stopping*), kuris sustabdo mokymą, jei nuostolių funkcijos vertė nepagerėja per nustatytą epochų skaičių. Ši strategija apsaugo nuo perteklinio pritaikymo (angl. *overfitting*) ir sumažina nereikalingų

skaičiavimų kiekį, užtikrinant, kad modelis pasiektų optimalų našumą, pernelyg nevilkindamas mokymo proceso.

Mokymosi greitis (angl. *learning rate*) kontroliuoja žingsnio dydį kiekvienoje iteracijoje, judant link nuostolių funkcijos minimumo. Tinkamai parinkta mokymosi sparta subalansuoja mokymo proceso konvergavimo greitį ir mokymo stabilumą. Mokymo greičio derinimui dažnai naudojami tvarkaraščiai (angl. *learning rate schedulers*), pavyzdžiui kosinusoidinis mažinimas (angl. *cosine annealing* arba tiesinis mokymosi greičio didinimas iki tam tikros vertės, o toliau - greičio mažinimas (angl. *linear warmup and decay*).

Paketo dydis (angl. *batch size*) apibrėžia kokį kiekį pavyzdžių pateikti modeliui per vieną iteraciją. Didesnis kiekis stabilizuoja gradientų skaičiavimus, tuo tarpu mažesnis – efektyviau išnaudoja atmintį.

Kitas svarbus parametras – regularizacija. λ_1 (žr. 3.3 skyrių), skatinantis modelio išretinimą (angl. *sparsity*) ir užkertantis kelią pertekliniam pritaikymui. Atsižvelgiant į optimizavimo algoritmą *Adam* galima išskirti kelis pagrindinius hiperparametrus: λ_2 , naudojama *Ridge* regularizacijai (žr. (3.4)), β_1 ir β_2 – nuokrypiui ir dispersijai balansuoti (žr. 3.5, 3.6, 3.7). Be to, plačiai taikomas išmetimo (angl. *dropout*) parametras, kuris atitiktinai išjungia nurodytą dalį neuronų mokymo metu, taip užtikrinant, jog modelis nesirems tik keliais požymiais.

Šių hiperparametrų derinimas yra sudėtingas dėl tarpusavio priklausomybės ir didelės paieškos erdvės. Mašininio mokymosi srityje, tinkamų parametrų radimas dažnai atliekamas eksperimentiniu būdu – parametrus keičiant rankiniu būdu, naudojant atsitiktinę paiešką arba tinklelinę paiešką (angl. *grid search*), ar pasinaudojant Bayesian metodu. Paskutiniu atveju parametrų paieška, skiriasi nuo kitų, prieš tai minėtų metodų, kadangi paieška yra paremta išvestimis iš prieš tai atliktų skaičiavimų (iš esmės sprendžiant prognozavimo uždavinį).

3.9. Vertinimo metodai ir rodikliai

Mašininio mokymosi srityje modelio našumo įvertinimas yra labai svarbus siekiant nustatyti jo veiksmingumą sprendžiant realias problemas. Viena iš pagrindinių klasifikavimo modelių vertinimo priemonių yra klasifikavimo matrica – lentelė, kurioje apibendrinamas klasifikavimo algoritmo veikimas. Klasifikavimo matricos pirmasis stulpelis nurodo prognozuojamas klases, o eilutė nurodo tikrąsias klases, jų sankirtos taške nurodomas kiekis teisingų ir neteisingų prognozių kiekvienai klasei, pavyzdžiui paprasčiausiu binariniu (klasifikuojama ar objektas priklauso klasei ar ne) atveju:

2 lentelė. Klasifikavimo matricos pavyzdys binariniu atveju, TP – teisingų prognozių, jog objektas priklauso klasei skaičius, FP – klaidingų prognozių, jog objektas priklauso klasei skaičius, FN klaidingų prognozių, jog objektas nepriklauso klasei skaičius, TN – teisingų prognozių, jog objektas nepriklauso klasei skaičius.

	1	0
1	TP	FP
0	FN	TN

Iš klasifikavimo matricos galima išvesti keletą našumo rodiklių. Paprasčiausias rodiklis – tikslumas (angl. *accuracy*), kuriuo matuojama teisingai klasifikuotų atvejų dalis palyginus su visais atvejais. Jis apskaičiuojamas kaip teisingų prognozių suma, padalinta iš bendro visų atvejų skaičiaus, paprasčiausiu atveju:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.13)$$

Svarbu pastebėti, jog esant netolygiam objektų pasiskirstymui tarp visų klasių, tikslumo parametras nėra tinkamas. Šiai problemai spręsti įvedami atkūrimo (angl. *recall*) ir preciziškumo (angl. *precision*) parametrai. Preciziškumas nurodo kokia dalis modelio prognozuotų klasės verčių iš tiesų yra tos klasės:

$$Precision = \frac{TP}{TP + FP} \quad (3.14)$$

Atkūrimo rodiklis parodo kokia dalis klasės verčių buvo prognozuota teisingai:

$$Recall = \frac{TP}{TP + FN} \quad (3.15)$$

Norint įskaityti preciziškumo ir atkūrimo rodikliui dažnai naudojamas F1 rodiklis (angl. *F1 score*), kuris apibrėžiamas kaip preciziškumo ir atkūrimo harmoninis vidurkis:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (3.16)$$

Tačiau, dažnu atveju, sprendžiant sudėtingesnes problemas, kur esama daugiau nei viena klasė, ar turimų duomenų kiekis svyruoja priklausomai nuo klasės, F1 rodiklis modelio įvertinimui apskaičiuojamas priklausomai nuo problemos specifikos. Esant daugiaklasei problemai su netolygiu duomenų pasiskirstymu per klases F1:

$$F1_{macro} = \frac{1}{N} \sum_{n=1}^N F1_n, \quad (3.17)$$

kur $n = 1, 2, \dots, N$. Tačiau kai nagrinėjamos klasės nėra vienodai svarbios, tokia metrika nebūtų visiškai tiksli. Tokiu atveju F1 rodiklis gali būti apibrėžiamas kiekvieną klasę įvertinant kaip atskirą binarinę problemą (angl. *one vs all*) ir skaičiuojant atkūrimą ir preciziškumą konkrečiai klasei [53]. Jei visgi svarbus bendras modelio veikimas, F1 skaičiuojamas per visus pavyzdžius:

$$F1_{micro} = 2 \cdot \frac{Precision_{micro} \cdot Recall_{micro}}{Precision_{micro} + Recall_{micro}} \quad (3.18)$$

Taip pat galima vertinimo metrika, kiekvienos klasės įverčiui priskiriant atitinkamą svorio koeficientą pagal klasėje esančių pavyzdžių kiekį (taip atspindint realų duomenų pasiskirstymą ar nagrinėjant klases su skirtinga svarba):

$$F1_{weighted} = \sum_{n=1}^N w_n \cdot F1_n \quad (3.19)$$

Pastaraisiais metais vis didesnis dėmesys skiriamas giluminio mokymosi modelių paaiškinamumui (angl. *explainable artificial intelligence, XAI*), siekiant geriau suprasti, kokia informacija

naudojama priimant sprendimus. Paaiškinamumo metodai leidžia vizualizuoti modelio sprendimų pagrindą ir įvertinti, ar modelis remiasi prasmingais duomenų požymiais, ar atsitiktiniais artefaktais. Skirtingos neuroninių tinklų architektūros reikalauja skirtingų paaiškinamumo metodų. Konvoliuciniams neuroniniams tinklams dažnai taikomi gradientais pagrįsti metodai, tokie kaip *Grad-CAM* (angl. *Gradient-weighted class activation map*). Į tinklą padavus paveikslą, iš paskutiniojo konvoliucinio sluoksnio surenkami bruožų žemėlapiai (angl. *activation maps*) A^k , o iš modelio išvesties paimamas modelio sprendimo *logit* vertė y_c , apskaičiuojami y_c gradientai priklausomai nuo A^k . Tuomet apskaičiuojami koeficientai kiekvienam bruožų žemėlapiui:

$$\alpha_k^C = \frac{1}{Z} \sum_i \sum_j \frac{\delta y_c}{\delta A_{ij}^k} \quad (3.20)$$

Panaudojant koeficientus α_k^C apskaičiuojama visų bruožų žemėlapių suma:

$$L_{Grad-CAM}^C = ReLU\left(\sum_k \alpha_k^C A^k\right) \quad (3.21)$$

ReLU aktyvacijos funkcija naudojama siekiant išlaikyti tik tuos bruožus kurie padeda identifikuoti klasę C .

Transformatorių tipo modeliams – paaiškinamumui naudojamos dėmesio mechanizmais pagrįstos vizualizacijos. Kiekviename transformatoriaus sluoksnyje apskaičiuojamos dėmesio matricų reikšmės, nusakančios, kaip stipriai atskiri žetonai tarpusavyje sąveikauja. Dėmesio išplėtimo (angl. *attention rollout*) metodas leidžia vizualizuoti informacijos srautą nuo įvesties iki galutinio sluoksnio. Užtikrinti, kad teisingai surenkama kaip informacija juda tinklo architektūroje, pirmiausia atsižvelgiama į likutinį ryšį, sluoksnio išvestis V_{l+1} aprašoma:

$$V_{l+1} = (W_{att} + I)V_l, \quad (3.22)$$

kur W_{att} – dėmesio matrica, o I – identiteto matrica. Tuomet normalizavus svorius gaunamas dėmesio svoris įskaitant likutinį ryšį A . Galutinis dėmesio žemėlapis gaunamas rekursyviai dauginant visas dėmesio verčių matricas per visus sluoksnius. Tokiu būdu gaunamas apibendrintas dėmesio žemėlapis, parodantis, kuriems įvesties vaizdo regionams modelis skiria daugiausiai dėmesio galutinės reprezentacijos formavimo metu. Svarbu pabrėžti, kad dėmesio mechanizmais pagrįstos vizualizacijos atspindi informacijos sklaidimą modelio viduje, tačiau nebūtinai tiesiogiai nusako, kurie vaizdo regionai yra kritiškai svarbūs konkrečiam klasifikavimo sprendimui (kadangi traktuoja dėmesio svorius kaip fiksuotus ir neatsižvelgia į gradientus).

4. Darbo metodika

Šiame skyriuje pateikiama darbo metodikos apžvalga, pirmasis poskyris pristato naudotus programinius paketus ir aprašo duomenų rinkinius bei jų analizės ir apdorojimo procesą. Antrajame poskyryje pateikiama informacija apie architektūros modeliavimą. Paskutiniajame poskyryje pristatoma vertinimo metodika ir vertinimo rodikliai.

4.1. Darbo įrankiai ir duomenų rinkiniai

Darbas atliktas pasitelkiant *Python* programavimo kalbą, bei *PyTorch* mašininio mokymosi biblioteką. Duomenų vizualizavimui, apdorojimui ir panašioms funkcijoms atlikti papildomai buvo naudojamos šios bibliotekos: *pandas*, *numpy*, *matplotlib*, *h5py*, *tqdm*, *pillow*, *datasets*, *optuna*, *seaborn*, *pytorch_warmup*, *scikit-learn*, *astropy*. Modelių apmokymai buvo atlikti naudojant sistemą su NVIDIA RTX 3090 vaizdo plokšte (24 GB GDDR6X VRAM, *Ampere* architektūra), AMD Ryzen 5700x (3,4 GHz, 8 fiziniai branduoliai) ir 64GB operatyviosios atminties.

Darbe buvo naudojami duomenys iš šių *Galaxy Zoo* projektų:

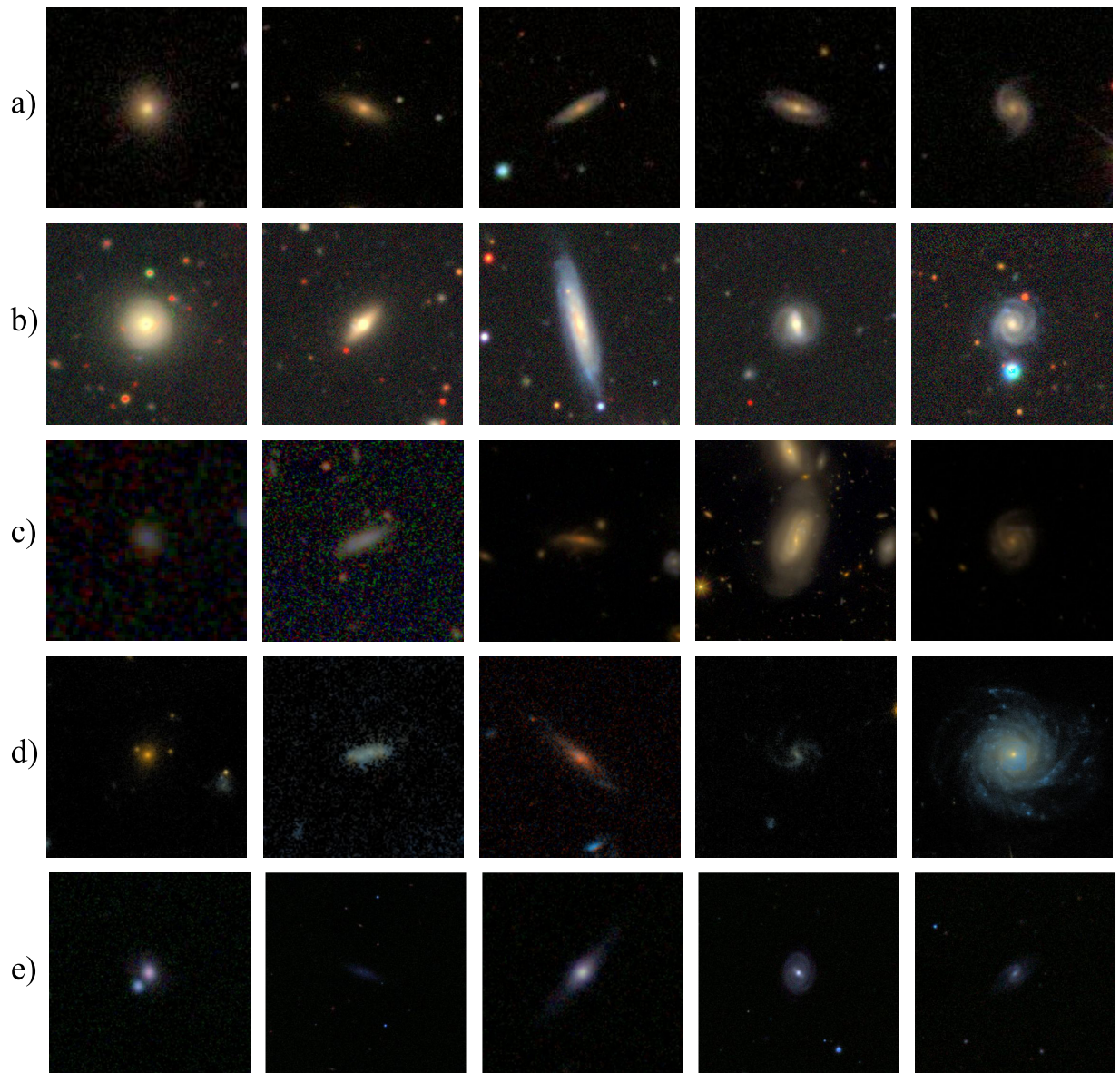
- *Galaxy Zoo 2* (toliau GZ2).
- *Galaxy Zoo Hubble* (toliau GZH).
- *Galaxy Zoo CANDELS* (toliau GZC).
- *Galaxy Zoo DECaLS* (toliau GZD) ¹.
- *Galaxy Zoo UKIDSS* (toliau GZU).

Galaktikų klasifikavimo uždaviniui spręsti, atsižvelgiant į projektų autorių išskirtas problemas (1.2 poskyris), buvo pasirinktos šios galaktikų klasifikavimo klasės (8 paveikslas):

- *C0*: apvali elipsinė galaktika,
- *C1*: cigaro formos elipsinė galaktika,
- *C2*: diskinė galaktika (galaktikos orbitos plokštuma yra lygiagreti regėjimo linijai su Žeme),
- *C3*: skersinė spiralinė galaktika,
- *C4*: spiralinė galaktika

Klasifikavimui į aukščiau pateiktas klases, aktualūs klausimai – T0, T1, T07, T08, T02, T03 (1.2 poskyris, 1 lentelė). Kiekvienai klasei nusakyti klausimai atrenkami pagal klausimyno eigą, t.y. skersinei spiralinei galaktikai identifikuoti reikalingi T0, T1, T02 ir T03 klausimai.

¹Duomenų rinkinys tiesiogiai naudojamas spektroskopinei informacijai, duomenų rinkinio pakeitimai modelio mokymui, pateikiami toliau tekste.



8 pav. Galaktikų naudotų modelių apmokymams pavyzdžiai pagal klases ir duomenų rinkinius. Iš kairės į dešinę stulpeliuose: apvali elipsinė galaktika, cigaro formos elipsinė galaktika, diskinė galaktika, skersinė spiralinė galaktika ir spiralinė galaktika. a) – GZ2 rinkinio pavyzdžiai, b) – G10D pavyzdžiai, c) – GZC pavyzdžiai, d) – GZH pavyzdžiai, e) – GZU pavyzdžiai.

Šiame darbe klasių parinkimui buvo naudojamos [54] išskirtos klasės (naudojama GZ2, GZH, HZC, GZU). GZD projekto galaktikoms buvo naudojamas duomenų rinkinys Galaxy10 DECaLS (toliau G10D) [2], [3], [22], [55], kuris apjungia DR2 ir DECaLS rinkinių nuotraukas, užtikrinant geresnės kokybės vaizdus, ir pasitelkiant *Galaxy Zoo* projekto metu surinktus balsus. Esminiai klasių priskyrimo principai tarp *Galaxy Zoo* ir *Galaxy10* klasifikavimo apibūdinami tolimesnėse pastraipose.

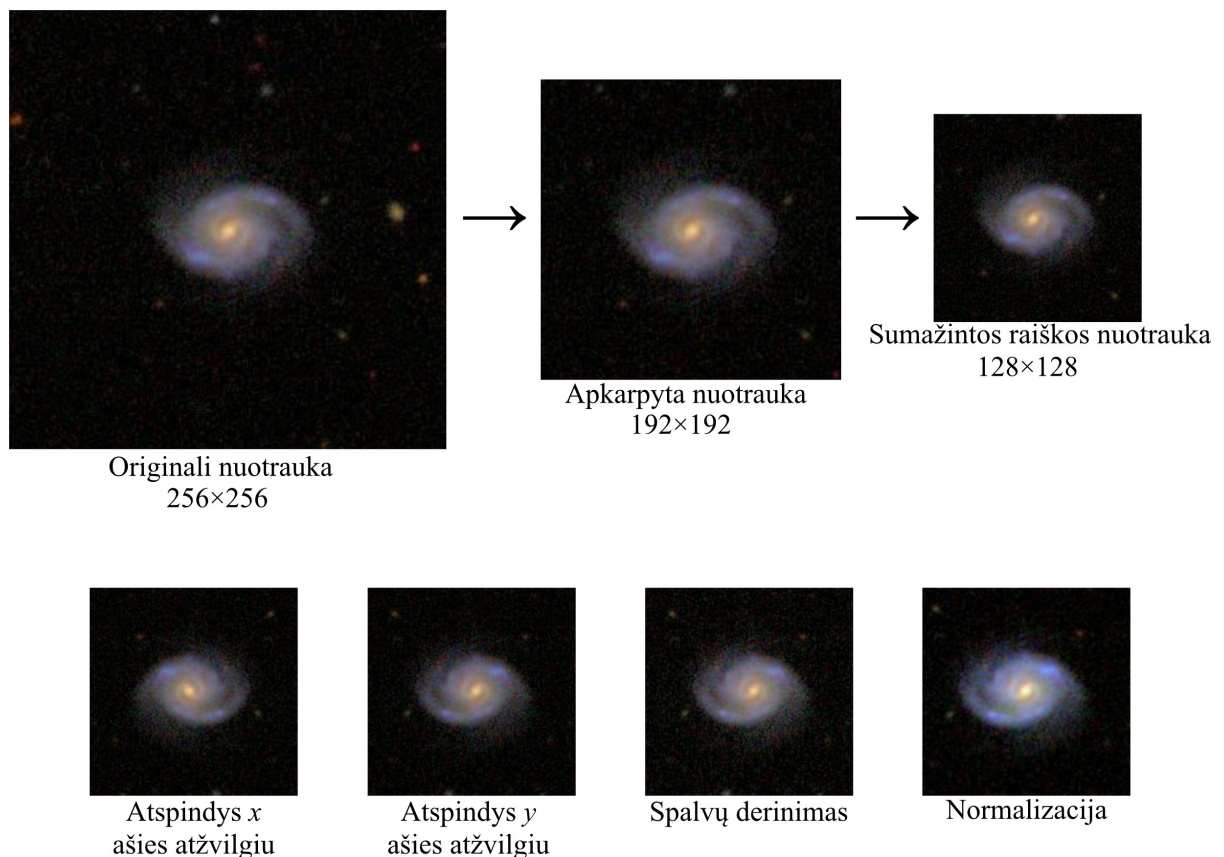
G10D klasifikavimas atliktas galaktikas suskirstant į 10 klasių, tuo tarpu *Galaxy Zoo* sudaro 6 klases. Dauguma klasių sutampa², dėl šios priežasties klasės iš skirtingų duomenų rinkinių buvo apjungtos tokia tvarka:

²T.y. priskirtos klasės yra tapačios arba *Galaxy Zoo* klasė apima dvi ar daugiau G10D išskirtų klasių (pvz.: Spiralinė galaktika apima spiralines galaktikas su glaudžiomis vijomis ir spiralines galaktikas su plačiomis vijomis)

- Spiralinė galaktika: spiralinė galaktika su glaudžiomis vijomis ir spiralinė galaktika su plačiomis vijomis,
- Diskinė galaktika: diskinė galaktika ir diskinė galaktika su skerse,
- Klasė aprašanti galaktikas, kurios neturi nei skersėtumo, nei spiralinių bruožų (tačiau nėra priskiriama elipsinei galaktikai) buvo panaikinta dėl nepakankamo duomenų kiekio ³.

Pačių klasių parinkimas *Galaxy Zoo* duomenų rinkinyje, buvo atliekamas atsižvelgiant į bendrą balsų skaičių ir svorį, t.y. klausime reikiamas atsakymas gavo daugiau negu pusę visų balsų ir balsų kiekis > 30 . G10D duomenų rinkinyje klasės priskyrimas iš esmės priklauso nuo pačios klasės (pvz.: klasės kurioms reikalingas mažas klausimų skaičius, balsų kiekis ir svoris buvo apribojamas griežčiau, tuo tarpu klasėms kurioms reikalingi "sudėtingesni" klausimai, ribojimai buvo reliatyviai lengvesni), tačiau atsižvelgiant į tai, jog G10D rinkinyje klasių kiekis buvo žymiai didesnis (vadinasi, reikalingas didesnis "sudėtingų" klausimų skaičius), šiame darbe autorių parinktos sąlygos buvo laikomos pakankamomis laikyti nustatytą klasę kaip teisingą.

Lentelėje 4 pateikiamas visų darbe naudojamų galaktikų nuotraukų skaičius pagal klases. Pagal išskirtas klases, galaktikų nuotraukos iš duomenų rinkinių buvo apdorotos toliau pateikiama tvarka, o apdorojimo eigos vizualizacija pateikta 9 paveiksle. Nuotraukų apdorojimas buvo parinktas atsižvelgiant į 2 skyriuje nagrinėtus straipsnius, bei atsiremiant į T. Mūžo, A. V. M. Misiūno ir T. Meškausko tyrimą [56], kuris parodė, jog standartinis apdorojimas galaktikų projekcijoms (pvz.: pasukimai, atspindėjimai, priartinimas) gali žymiai pagerinti likutinių tinklų rezultatus.



9 pav. Nuotraukų apdorojimo eigos vizualizacija.

³Modelių mokymo metu buvo pastebėta, jog papildomų dirbtinių duomenų įvedimas siekiant padidinti klasės pavyzdžių kiekį nepasiteisino.

Pirmajame žingsnyje galaktikų nuotraukos buvo apkarpos, siekiant sumažinti foninį triukšmą (t.y. kitus nuotraukose matomus dangaus kūnus), tuomet nuotraukų dydis buvo pakeistas (128×128 raiška). Apkarpos derintos pagal duomenų rinkinius, siekiant užtikrinti, kad galaktikų vaizdas nebūtų apkarpos, tačiau būtų panaikinama kaip įmanoma daugiau nereikalingos informacijos. Įvesčiai į modelius, pateikiamos nuotraukos buvo transformuojamos pasitelkiant atspindėjimą abiejų ašių atžvilgiu, bei spalvų derinimą (ryškumas, kontrastas, sodrumas). Atsižvelgiant į tai, jog kai kurių projektų naudojami nuotraukų rinkiniai sudaryti iš tolimų galaktikų (GZH, GZC) buvo reliatyviai prastos kokybės, o pasukimas atsitiktiniu kampu dažnu atveju sumažintų pačios paveiksle matomos galaktikos raišką, buvo pasirinkta nenaudoti atsitiktinio pasukimo, taip išvengiant kokybės suprastėjimo. Spalvų derinimas, atsižvelgiant į tyrimo sritį, buvo atliekamas siekiant atkartoti galimus realius pasikeitimus (pvz.: įrangos įvedamus triukšmus ir netikslumus). Spalvų derinimui buvo parinktos šios vertės (3 lentelė):

3 lentelė. Spalvų derinimo parametrai.

Parametras	Vertė
Ryškumas	0, 15
Kontrastas	0, 15
Sodrumas	0, 15

Toliau, nuotraukos buvo normalizuojamos pagal viso mokymo rinkinio standartinį nuokrypį σ ir vidurkį μ :

$$\mu = \frac{1}{N} \cdot \sum x_i, \quad (4.1)$$

čia N - nuotraukų skaičius (realizavime operacija atliekama dalimis – atskiriems duomenų paketais (angl. *batches*), gautas vertes susumuojant), o x_i - nuotrauka (operacija atliekama per pikselį),

$$\sigma = \sqrt{\frac{1}{N} \cdot \sum (x_i - \mu)^2}. \quad (4.2)$$

4 lentelė. Nuotraukų pavyzdžių pasiskirstymas pagal klases.

Klasė	Mokymo aibė	Validavimo aibė	Testavimo aibė
<i>C0</i>	78413	9801	9802
<i>C1</i>	27967	3496	3496
<i>C2</i>	8623	1078	1078
<i>C3</i>	10128	1266	1266
<i>C4</i>	28185	3523	3523
Iš viso	153316	19164	19165

Po duomenų apdorojimo nuotraukos buvo suskirstytos į mokymo, validavimo ir testavimo aibes (4 lentelė), tuomet buvo pritaikytos ankščiau pateiktos nuotraukų transformacijos. Duomenų paruošimas *GalaxyCNN* ir *GalaxyCvT* modeliams (žr. 4.2.1 ir 4.2.2 poskyrius) iš esmės vienodas, tačiau mokymo metu, CvT architektūros modeliui buvo pritaikyta papildoma duomenų transformacija *MixUp*. Kadangi šioje transformacijoje nuotraukos perdengiamos viena kita su skirtingu

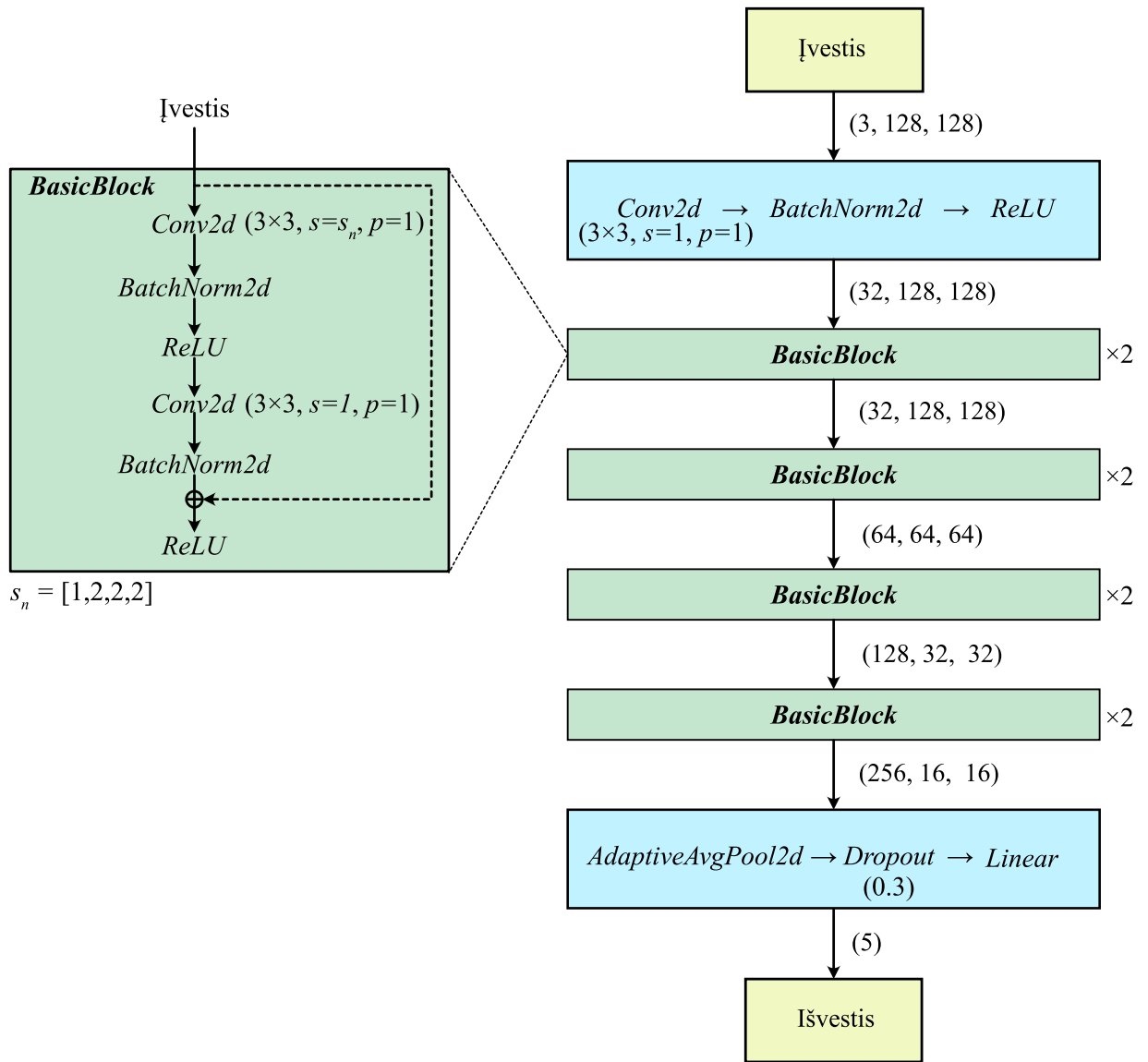
skaidrumu, pateikiamas ne duomenų rinkinys (angl. *dataset*), o toks pat formatas kaip ir naudojama modelio įvesčiai (*data loader*).

4.2. Modelių architektūros

Šiame poskyryje pristatomi, atsižvelgiant į ankstesniame poskyryje aprašytą duomenų rinkinį, sukurti du tinklų modeliai – CNN modelis (pasitelkiant likutines jungtis [45]) – *GalaxyCNN*, bei transformatorių ir konvoliucinių tinklų pagrindu paremtas [52] CvT modelis *GalaxyCvT* – jų architektūra, bei parinkti hiperparametrai. *GalaxyCvT* modelis šiame darbe naudojamas tik palyginamajai analizei, o pagrindinis dėmesys skiriamas CNN pagrįsto *GalaxyCNN* modelio kūrimui ir vertinimui.

4.2.1. *GalaxyCNN* architektūra

Konvoliucinis neuroninis tinklas *GalaxyCNN* buvo sukurtas *ResNet-18* architektūros [45] pagrindu. Pirmasis sluoksnis turi 32 filtrus (3×3), modelis turi 5 išvestis, o paskutiniame sluoksnyje papildomai naudojamas išmetimas (angl. *dropout*). Tinklo architektūros principinė diagrama pateikta 10 paveiksle.



10 pav. GalaxyCNN modelio architektūros principinė diagrama.

Tinklas susideda iš keturių etapų, kuriuos sudaro du *BasicBlock* sluoksniai (kiekvieną iš jų sudaro du konvoliuciniai sluoksniai su likutinėmis jungtimis, po kiekvieno konvoliucinio sluoksnio seka paketo normalizacija ir ReLU aktyvacijos funkcija). Kiekviename etape filtrų skaičius didinamas, pradedant nuo 32 iki 256 ($s = 2$ pirmame konvoliucijos sluoksnyje pradedant nuo antrojo etapo). Po savybių išskyrimo erdvinės dimensijos sumažinamos *Adaptive Average Pooling* pagalba, o prieš pilnai sujungtą sluoksnį naudojamas išmetimas ($d = 0.3$).

Modelio mokymui naudotas paketo dydis (angl. *batch size*) – 512. Pradinis mokymosi greitis $lr = 0.01$, o tolimesnei mokymosi spartai reguliuoti buvo naudojamas *ReduceLROnPlateau* tvarkaraštis, atsižvelgiant į nuostolių funkcijos vertės kitimą:

$$T_d = L_b \cdot (1 - t), \quad (4.3)$$

kur L_b – mažiausia nuostolių vertė, $t = 0.1$ – ribos koeficientas (angl. *factor*). Nuostolių skaičiavimui buvo pasirinkta *Focal Loss* funkcija, o siekiant atsižvelgti į netolygų klasių pasiskirstymą (4 lentelė) klasėms buvo priskirti svoriai. Svoriai apskaičiuoti:

$$w_n = 1 - \frac{N_c}{N_{tot}}, \quad (4.4)$$

kur N_{tot} – visas pavyzdžių kiekis, N_c – pavyzdžių kiekis klasėje. *Focal Loss* realizavimas buvo paremtas [43] autorių viešai prieinamu kodu⁴, jį pritaikant daugiaklasei problemai.

Kadangi *GalaxyCNN* modelis paremtas likutinėmis jungtimis, optimizavimui buvo pasirinktas SGD (angl. *Stochastic Gradient Descent*) algoritmas su pagreičiu ($a = 0.9$) bei $\lambda_{L2} = 0.0001$. Modelio mokymas buvo atliktas 100 epochų, su ankstyvuoju stabdymu (nuostolių vertės pokytis per 7 epochas ne mažesnis nei 10^{-4}).

4.2.2. *GalaxyCvT* architektūra

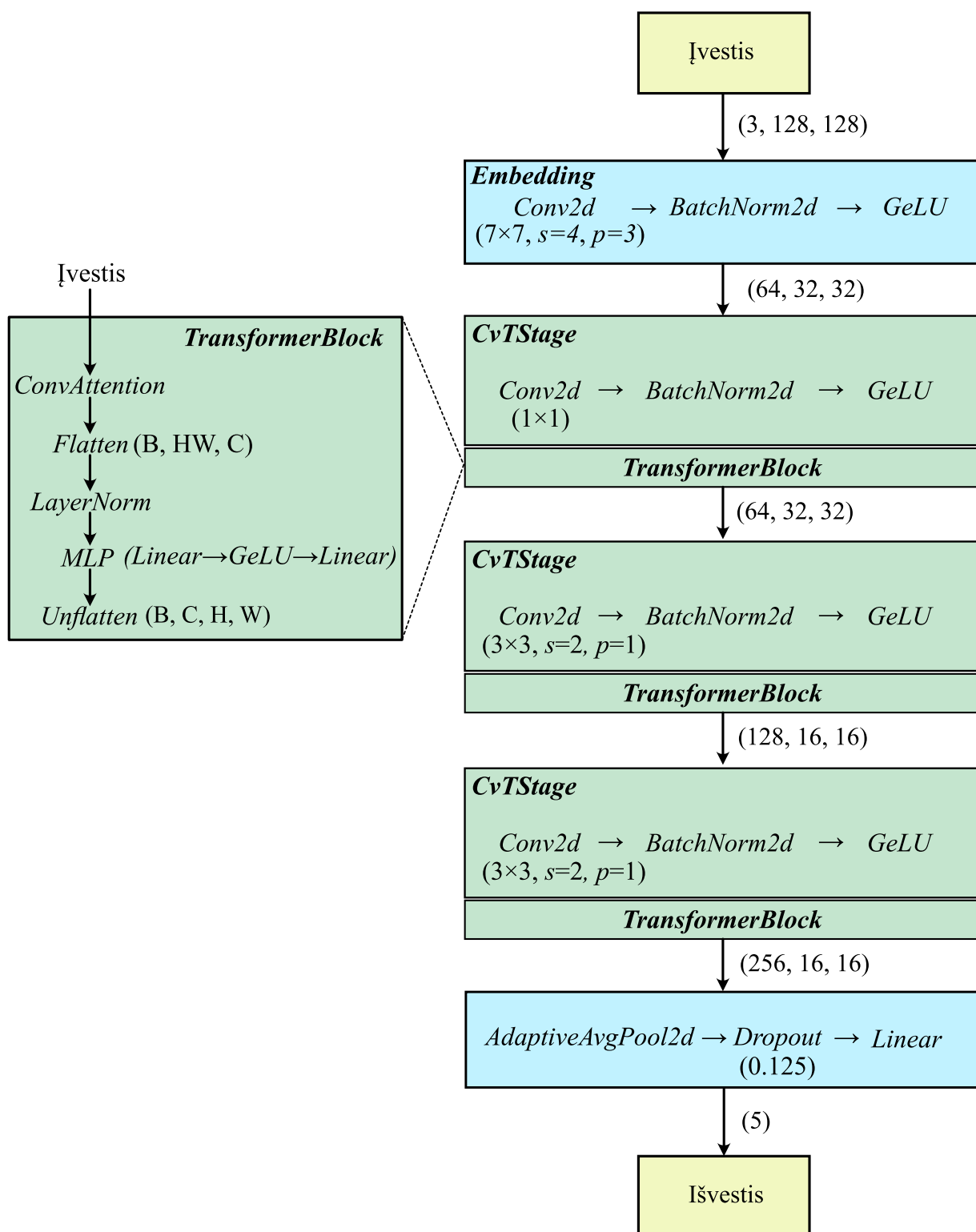
Konvoliucinis regos transformatorius *GalaxyCvT* buvo sukurtas atsižvelgiant į 2021 m. pasiūlytą CvT architektūrą [52]. Modelio įvestis tokia pati kaip ir *GalaxyCNN* – 128×128 paveikslai su 3 spalvų kanalais. Tinklo architektūros principinė diagrama pateikta 11 paveiksle.

Pirmasis įterpimo sluoksnis (*Embedding*) atvaizduoja įvestį į pirminių (primityvių) bruožų seką, kuri toliau nuosekliai apdorojama sekančiuose 3 etapuose, didinant bruožų žemėlapius (filtrus) iki 256. Nuo antrojo etapo atliekama konvoliucija naudojant žingsnį $s = 2$ taip mažinant erdvę (angl. *downsampling*). Dėmesio mechanizmui K, V vertės (naudojant 3×3 konvoliuciją) įterpties sekos elementui (angl. *patch*) pateikiamos su dalimi informacijos iš šalia esančių elementų taip išlaikant lokalų kontekstą. Transformatoriaus blokai (*TransformerBlock*) susideda iš normalizacijos, MHA ir MLP (su GeLU aktyvacijos funkcija) sluoksnių. Po savybių išskyrimo, kaip ir *GalaxyCNN* modelyje, naudojama *Adaptive Average Pooling*, išmetimas ($d = 0.5$).

Modelio mokymui paketo dydis – 1024. Pradinis mokymosi greitis $lr = 0.01$. Mokymosi spartai reguliuoti buvo naudojami du skirtingi tvarkaraščiai – *LinearWarmup* ir *CosineAnnealing-LR*. Šiuo atveju sparta reguliuojama per visus žingsnius (t.y. bendras paketų skaičius per visas epochas). Pirmajam tvarkaraščiui veikimo periodas – 10% visų žingsnių, tuomet pereinama prie antrojo tvarkaraščio.

Nuostolių skaičiavimui, kaip ir *GalaxyCNN* modeliui buvo naudojama *Focal Loss* funkcija. *GalaxyCvT* optimizavimui buvo parinktas *AdamW* algoritmas, šiuo atveju reguliarizacija buvo taikoma tik svoriams, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 1 \cdot 10^{-8}$. Šiuo atveju reguliarizacija taikoma visiems parametrams išskyrus sluoksnio normalizacijos ir nuokrypio (angl. *bias*) parametrus.

⁴<https://github.com/facebookresearch/Detectron>



11 pav. GalaxyCvT modelio architektūros principinė diagrama.

4.3. Modelio vertinimo metodika

Modelių vertinimas buvo skirtas ne tik bendram klasifikavimui įvertinti, bet ir pasiūlytų architektūrų patikimumui, paaiškinamumui ir praktiniam pritaikomumui išanalizuoti. Atsižvelgiant į duomenų rinkinio netolygumą klasių atžvilgiu ir galaktikų klasių morfologinį sudėtingumą, vieno našumo rodiklio nepakanka, kad būtų galima visapusiškai apibūdinti modelio elgseną. Todėl verti-

nis buvo atliekamas keliais lygiais, įskaitant bendrus našumo rodiklius, klasių analizę, kokybinį modelio dėmesio patikrinimą, jautrumą mokymo duomenų dydžiui ir skaičiavimo efektyvumą.

Kiekybinis vertinimas orientuotas į bendrą modelio kokybę bei į modelio veiksmingumą kiekvienai klasei. Kadangi duomenų rinkinyje matomas netolygus pavyzdžių kiekis klasių atžvilgiu, tačiau visos klasės vertinamos kaip vienodai svarbios, pagrindinė bendra modelio vertinimo metrika buvo pasirinkta F1 *macro* metrika (žr. 3.17 formulė). Detalesniam modelių nagrinėjimui papildomai naudojamos atkūrimo, preciziškumo ir F1 metrikos apskaičiuojamos kiekvienai klasei. Klasifikacijos matricos naudojamos vizualizuoti sistemingas klaidas tarp morfologiškai panašių klasių.

Kadangi šiame darbe naudojamas duomenų rinkinys susideda iš kelių skirtingų projektų (žr. 1.2 ir 4.1 poskyrius), modelio veikimas papildomai analizuojamas atskirai kiekvienam duomenų rinkiniui. Tokiu būdu įvertinamas modelio gebėjimas apibendrinti duomenis gautus skirtingomis stebėjimo sąlygomis (pvz.: skirtinga vaizdo skiriamoji geba, triukšmo charakteristikos ir išankstinio apdorojimo procesai). Darbo eigoje kiekvieno projekto duomenų rinkiniui buvo surinkti papildomi spektroskopiniai ir fotometriniai duomenys, tačiau dėl įrangos skirtumų ir tiriamų kūnų nuotolio, panaudoti raudonojo poslinkio, ryškio ar dydžio parametrus tiesioginiam modelių palyginimui nebuvo galima. Todėl tolimesnis modelių vertinimas atsižvelgiant į duomenų kokybę buvo atliktas bendresniu požiūriu – remiantis bendromis kiekvieno duomenų rinkinio iš tam tikro projekto savybėmis. Kiekvienam duomenų rinkiniui pateikiamas klasės F1 vertinimas, siekiant pabrėžti galimus veikimo skirtumus, susijusius su duomenų kokybe ir išgavimo skirtumais.

Modelių mokymui naudotos nuostolių funkcijos nurodytos 4.2.2 ir 4.2.1 poskyriuose. Siekiant analizuoti mokymo stabilumą ir konvergavimo elgseną, visą optimizavimo procesą stebimos mokymo ir validacijos nuostolių kreivės. Šie grafikai suteikia įžvalgų apie modelio mokymosi dinamiką, konvergencijos greitį ir galimai pernelyg didelį pritaikymą. Aptariami konvergencijos elgsenos skirtumai tarp konvoliucinių tinklų ir transformatorių architektūrų, ypač atsižvelgiant į mokymo stabilumą ir jautrumą hiperparametrus.

Įvertinti modelių jautrumą skirtingiems duomenų kiekiams, atliekamas papildomas eksperimentas modelių apmokymui naudojant sumažintus mokymo rinkinio pogrupius. Modeliai mokomi naudojant 25%, 50% ir 100% turimų mokymo duomenų (validavimo ir testavimo rinkiniai lieka nepakitę). Veikimas matuojamas naudojant F1 *macro* metriką. Ši analizė suteikia įžvalgų apie architektūrų efektyvumą priklausomai nuo duomenų kiekio ir jų gebėjimą mokytis esant ribotam duomenų kiekiui.

Dėl esminių architektūrinių skirtumų tarp konvoliucinių neuroninių tinklų ir transformatorių architektūromis paremtų modelių buvo naudojami skirtingi modeliams būdingi paaiškinamumo metodai. *GalaxyCNN* modelio analizei buvo naudojama *Grad-CAM*, siekiant išryškinti erdvinės sritis, kurios daro didžiausią įtaką klasės prognozėms, o *GalaxyCvT* modeliui buvo taikomas dėmesio išskleidimas (angl. *attention rollout*), siekiant vizualizuoti modelio dėmesį per įvesties žetonus ir sluoksnius. Kadangi naudoti metodai iš esmės skirtingi, modelių palyginimas nėra tiesioginis. Vietoje to, paaiškinamumo analizė sutelkta į kokybines savybes – erdvinį fokusą, stabilumą esant nedidelėms įvesties transformacijoms, jautrumą foninei informacijai (triukšmui) ir paaiškinimo modelių nuoseklumą teisingai ir neteisingai klasifikuotuose pavyzdžiuose. Struktūrinis panašumas (angl. *structural similarity, SSIM*) naudojamas tik paaiškinimo stabilumui kiekviename modelyje įvertinti, o ne paaiškinimams skirtingose architektūrose palyginti.

Galiausiai, atliekama skaičiavimo efektyvumo (našumo) analizė. Mokymo laikas per epochą, bendra mokymo trukmė ir klasifikavimo laikas (angl. *inference*) matuojami naudojant tą pačią įrangą.

Rezultatai

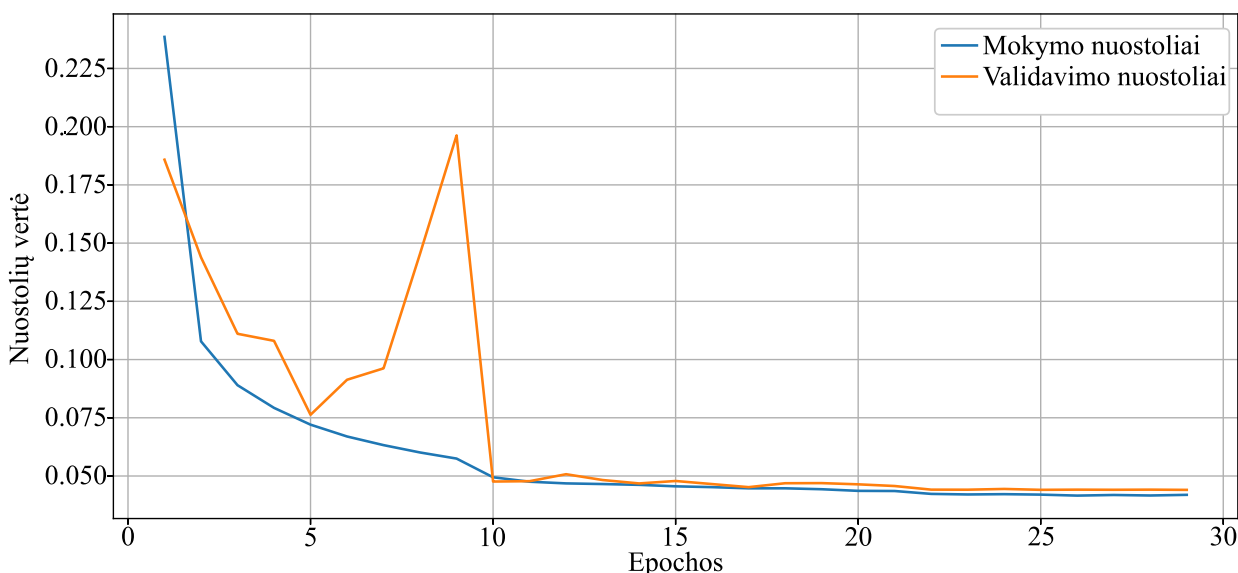
Bendras klasifikacijos vertinimas

Mokymo vertinimas ir nuostolių funkcijos konvergavimas

Abiejuose modeliuose naudota *Focal Loss* nuostolių funkcija. Šis pasirinkimas pagrįstas netolygiu duomenų pasiskirstymu tarp klasių bei empiriniais stebėjimais mokymo metu, kurie parodė, jog C3 klasėje egzistuoja santykinai didelė dalis sunkių“ pavyzdžių, t. y. objektų, kuriuos modeliams sudėtinga klasifikuoti užtikrintai.

Mokymo ir validavimo nuostolių verčių priklausomybė nuo mokymo epochų skaičiaus *GalaxyCNN* ir *GalaxyCvT* modeliams pateikta 12 ir 13 paveiksluose atitinkamai. Galima pastebėti, jog abiejuose modeliuose ankstyvojo stabdymo funkcija suveikė panašiu metu, apie 30 epochą.

GalaxyCNN ir *GalaxyCvT* modelių mokymo bei validavimo nuostolių verčių priklausomybė nuo epochų skaičiaus pateikta atitinkamai 12 ir 13 paveiksluose. Abiem atvejais ankstyvojo stabdymo (*early stopping*) mechanizmas suveikė panašiu metu – maždaug ties 30-ąja epocha, kas rodo stabilų mokymo procesą.

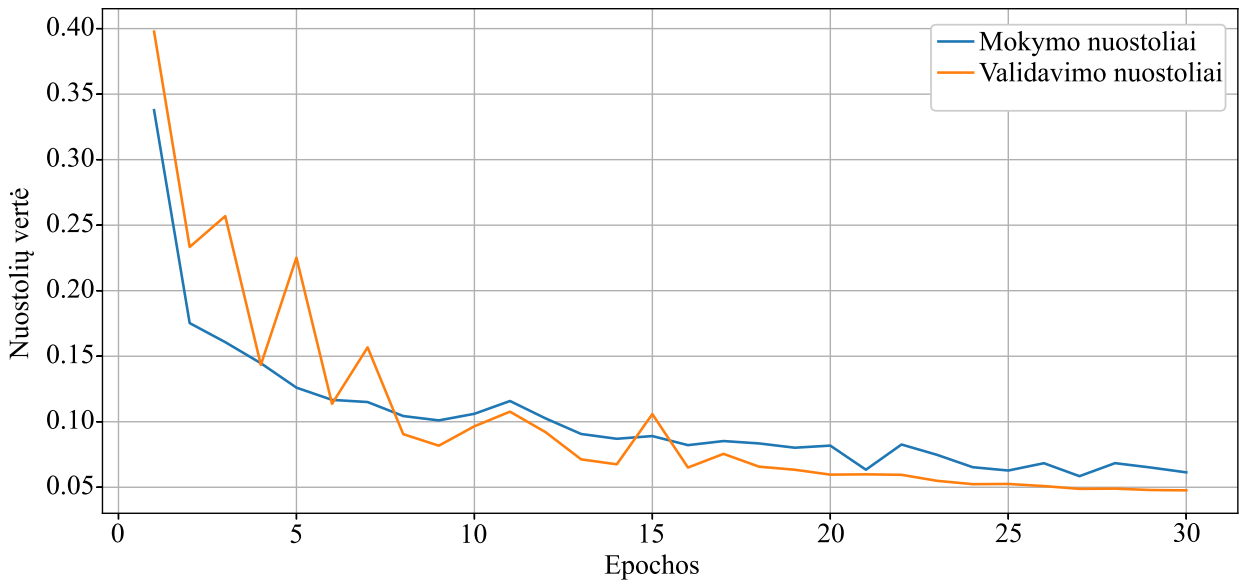


12 pav. *GalaxyCNN* modelio mokymo ir validavimo nuostolių verčių priklausomybė nuo epochų.

GalaxyCNN modelyje aiškus nuostolių verčių konvergavimas stebimas apie 22-ąją epochą. Skirtumas tarp mokymo ir validavimo nuostolių verčių yra nedidelis (apie 0.0021), todėl nėra ryškių modelio prisitaikymo požymių. Ankstyvosiose epochose matomas trumpalaikis validavimo nuostolių padidėjimas gali būti siejamas su *SGD* optimizatoriaus su pagreičiu dinamika, tačiau vėlesnėse epochose nuostolių reikšmės stabilizuojasi tiek mokymo, tiek validavimo duomenų rinkiniuose, kas rodo sėkmingą mokymosi greičio adaptaciją.

GalaxyCvT modelio nuostolių kreivėse pastebimos didesnės fluktuacijos, ypač ankstyvosiose mokymo stadijose. Šis elgesys gali būti paaiškinamas architektūriniais skirtumais tarp konvoliucinių ir transformatorių tipų modelių, kadangi CNN pasižymi stipresniu induktyviu šališkumu lokalių struktūrų atžvilgiu. Tokie architektūriniai ypatumai gali lemti stabilesnį mokymąsi esant sudėtingesniems ar triukšmingesniems validacijos pavyzdžiams. *GalaxyCvT* modelio nuostolių vertės nusistovi

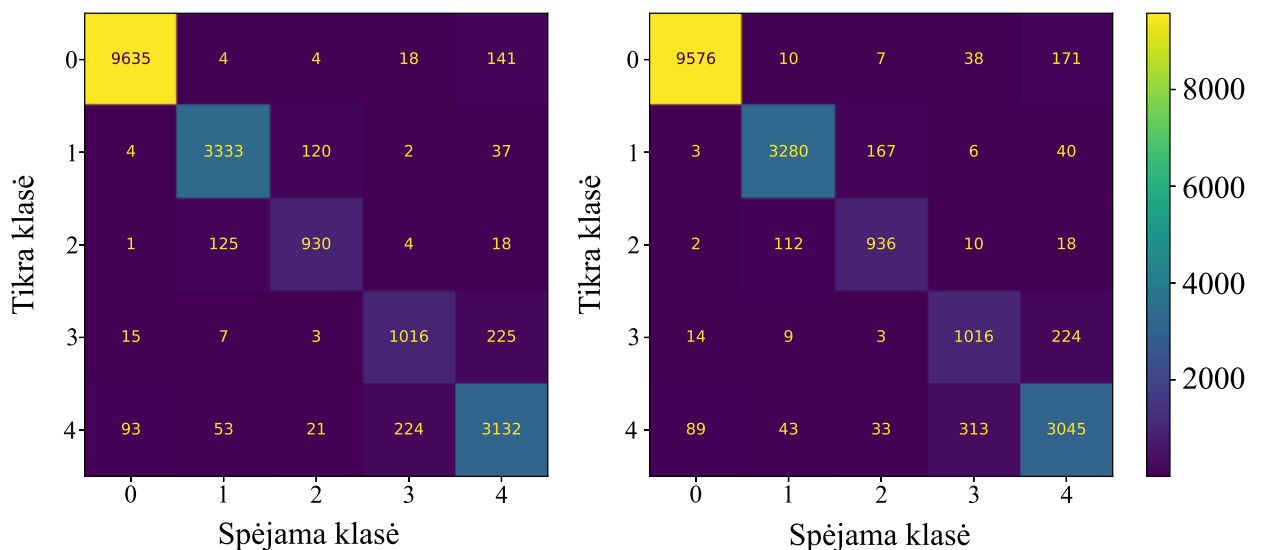
maždaug ties 24-ąja epocha, o galutinės validavimo nuostolių reikšmės išlieka artimos *GalaxyCNN* rezultatams ($\Delta \approx 0.01$).



13 pav. *GalaxyCvT* modelio mokymo ir validavimo nuostolių verčių priklausomybė nuo epochų.

Klasifikavimo kokybės analizė

GalaxyCNN ir *GalaxyCvT* modelių klasifikavimo matricos pateiktos 14 paveiksle. Didžiausias neteisingų klasifikacijų skaičius pastebimas tarp cigaro formos elipsinių (C1) ir šonu stebimų (diskinių) galaktikų (C2), taip pat tarp skersinių spiralinių (C3) ir spiralinių galaktikų (C4). Šiose klasių porose klaidų yra daugiau nei kitais atvejais, tačiau abiem atvejais tai gali būti paaiškinama ryškiais morfologiniais panašumais: cigaro formos elipsinės ir šonu stebimos diskinės galaktikos pasižymi panašia projekcine geometrija, o spiralinių ir skersinių spiralinių galaktikų esminis skirtumas dažniausiai matomas tik centrinėje galaktikos dalyje.



14 pav. Modelių klasifikavimo matricos. Kairėje pusėje *GalaxyCNN*, dešinėje – *GalaxyCvT*.

Papildomas klaidingų klasifikacijų atvejis pastebimas tarp apvalių elipsinių (C0) ir spiralinių (C4) galaktikų. Nors klaidų skaičius šiuo atveju yra mažesnis, šių klasių morfologiniai požymiai nėra tiesiogiai panašūs. Peržiūrėjus klaidingai klasifikuotų paveikslų pasiskirstymą pagal projektų duomenų rinkinius nebuvo pastebėta jokių aiškių tendencijų, tačiau išanalizavus klaidingai klasifikuotus pavyzdžius nustatyta, kad tokios galaktikos dažnai pasižymi netipiniais bruožais, pavyzdžiui, struktūriniais išdarymiais dėl galaktikų sąveikų arba stipriu foniniu triukšmu, kai kiti dangaus objektai yra labai arti galaktikos arba persidengia su ja.

Bendras vertinimo rodiklių palyginimas rodo, kad *GalaxyCNN* modelis ($F1_{macro} = 0.897$) nežymiai pranoksta *GalaxyCvT* ($F1_{macro} = 0.884$). Abu modeliai efektyviai klasifikuoja paprastesnes morfologines klases, tačiau *GalaxyCNN* pranašumas labiau išryškėja sudėtingesnėse klasėse, ypač C3 klasėje ($F1 = 0.80$ lyginant su 0.76; žr. 5 ir 6 lenteles).

5 lentelė. *GalaxyCNN* modelio vertinimo metrikos.

Klasė	Preciziškumas	Atkūrimas	F1
C0	0.9884	0.9830	0.9857
C1	0.9463	0.9534	0.9498
C2	0.8627	0.8627	0.8627
C3	0.8038	0.8025	0.8032
C4	0.8815	0.8890	0.8852
Macro vidurkis	0.8966	0.8981	0.8973
Weighted vidurkis	0.9418	0.9416	0.9417
Tikslumas			0.9416

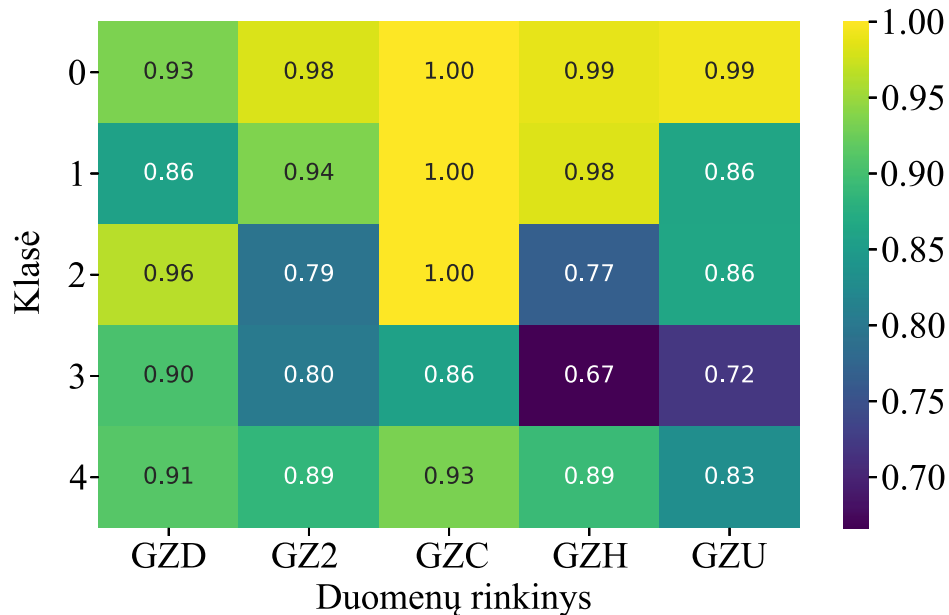
Galima daryti prielaidą, kad pasiūlyta CNN architektūra efektyviau išskiria lokalius morfologinius bruožus nei šiame darbe naudota CvT architektūra. Tačiau svarbu pabrėžti, jog *GalaxyCvT* modelis yra palyginti seklios architektūros, o abu modeliai buvo apmokyti naudojant tą patį duomenų kiekį ir vienodo dydžio įvestis. Papildomi eksperimentai su didesnės raiškos paveikslais (192×192) *GalaxyCvT* atveju parodė nežymų rezultatų pagerėjimą ($F1_{macro} = 0.8914$, $F1_{C3} = 0.7856$), tačiau *GalaxyCNN* rezultatai išliko aukštesni.

6 lentelė. *GalaxyCvT* modelio vertinimo metrikos.

Klasė	Preciziškumas	Atkūrimas	F1
C0	0.9862	0.9847	0.9855
C1	0.9352	0.9582	0.9466
C2	0.8757	0.8367	0.8558
C3	0.7749	0.7504	0.7624
C4	0.8678	0.8720	0.8699
Macro vidurkis	0.8880	0.8804	0.8840
Weighted vidurkis	0.9350	0.9354	0.9351
Tikslumas			0.9354

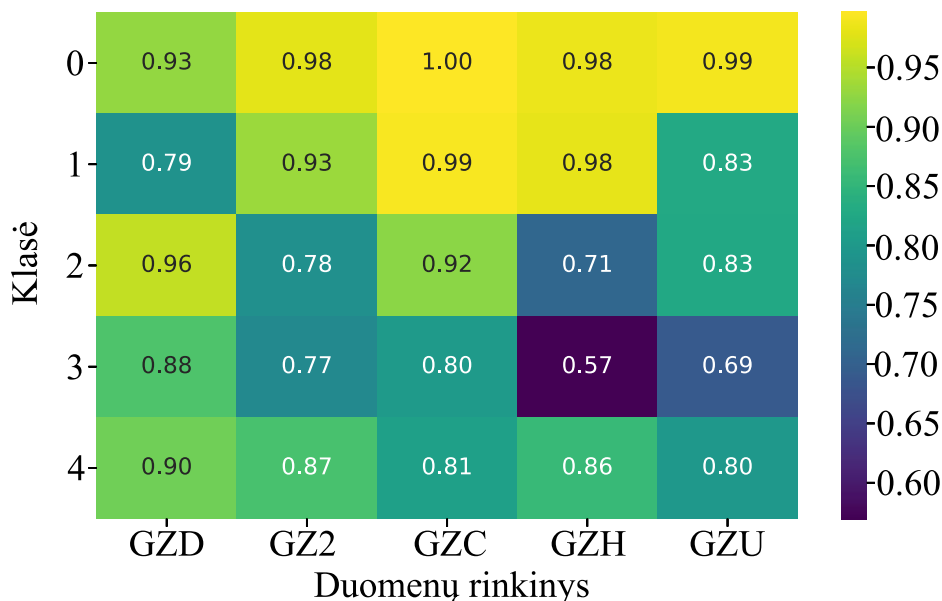
15 ir 16 paveiksluose pateikiamos $F1$ vertės kiekvienai klasei, atsižvelgiant į naudotus duomenų rinkinius. Abiem modeliams prasčiausi rezultatai matomi naudojant *Galaxy Zoo Hubble* (GZH) duomenis. Tai gali būti susiję su tuo, kad GZH tyrime analizuojamos galaktikos pasižymi didesniu raudonuoju poslinkiu ($z \sim 0.9$), todėl matomos galaktikos atrodo neryškios, mažesnės, o šviesos

bangą ilgėja. Prielaidos, jog spiralinių galaktikų kiekis didėjant z vertėms žymiai mažėja atsisakyta, kadangi nauji tyrimai šią prielaidą paneigia [57]. Žemesnės $F1$ vertės C3 klasei taip pat pastebimos GZU duomenų rinkinyje, kas gali būti siejama su tuo, kad GZU fotometriniai duomenys surinkti artimojo infraraudonojo (NIR) spektro diapazone. Skirtingas bangos ilgis lemia pakitusį morfologinių bruožų kontrastą, kas gali apsunkinti tam tikrų bruožų atpažinimą.



15 pav. *GalaxyCNN* klasių $F1$ vertės priklausomybė nuo skirtingų duomenų rinkinių.

Apskritai geriausios vertės matomos su duomenimis iš GZD rinkinio, kurie pasižymi mažomis z vertėmis, o naudojamas teleskopas yra jautresnis raudonos spalvos šviesai. Didesnis šviesos jautrumas taip pat gali nulemti ir prastesnius rezultatus specifinių klasių atžvilgiu, kadangi šiame duomenų rinkinyje galaktikų nuotraukose gali matytis "silpnėsni" bruožai, kurie naudojant kitokią aparatūrą paprasčiausiai nėra matomi.



16 pav. *GalaxyCvT* klasių $F1$ vertės priklausomybė nuo skirtingų duomenų rinkinių.

Bendru atveju rezultatų analizė pagal skirtingus duomenų rinkinius parodo abiejų modelių gebėjimai išskirti galaktikų bruožus plačiame duomenų spektre (apimant platų raudonojo poslinkio diapazoną, skirtingos įrangos parametrų daromą įtaką, bei reliatyviai platų šviesos spektrą). Dauguma klasifikavimo rodiklių išlieka nuosekliai aukšti visuose penkiuose tyrimo duomenų rinkiniuose. Atskiros pastebėtos veikimo problemos, ypač žemesnės $F1$ vertės 3 klasei didelio raudonojo poslinkio GZH tyrime ir NIR šviesos spektre tiriamuose (GZU) duomenyse, pirmiausia yra susijusios su pačių duomenų apribojimais (žemesnė skiriamoji geba ir morfologinių bruožų skirtumai dėl bangos ilgio pasikeitimo).

Papildomai, remiantis 2 skyriuje apžvelgtais darbais, bei sukurtu *GalaxyCvT* modeliu, buvo atliktas *GalaxyCNN* modelio palyginimas, pateiktas 7 lentelėje.

Svarbu pabrėžti, jog lyginamuose darbuose nagrinėjamos klasifikavimo užduotys nėra visiškai tapачios: skiriasi tiek klasių skaičius ir jų apibrėžimas, tiek naudojami duomenų rinkiniai. Be to $F1_{macro}$ rodiklis nebuvo pateiktas visuose nagrinėtuose straipsniuose – darbuose [7] ir [40] ši vertė apskaičiuota remiantis pateiktomis kiekvienos klasės $F1$ reikšmėmis, o [41] atveju – iš nurodytų preciziškumo ir atkūrimo verčių kiekvienai klasei. Dėl šių priežasčių pateiktas palyginimas turėtų būti vertinamas kaip orientacinis ir kontekstinis, o ne kaip tiesioginis skirtingų metodų veikimo reitingavimas.

Trečiame, penktame ir šeštame palyginamuose modeliuose (žr. 7 lentelę) nagrinėjamos klasės, kurios sutampa su šiame darbe išskirtomis klasėmis: apvalios elipsinės, cigaro formos elipsinės, diskinės (stebimos šonu) ir spiralinės galaktikos.

7 lentelė. *GalaxyCNN* modelio palyginimas su kitais modeliais.

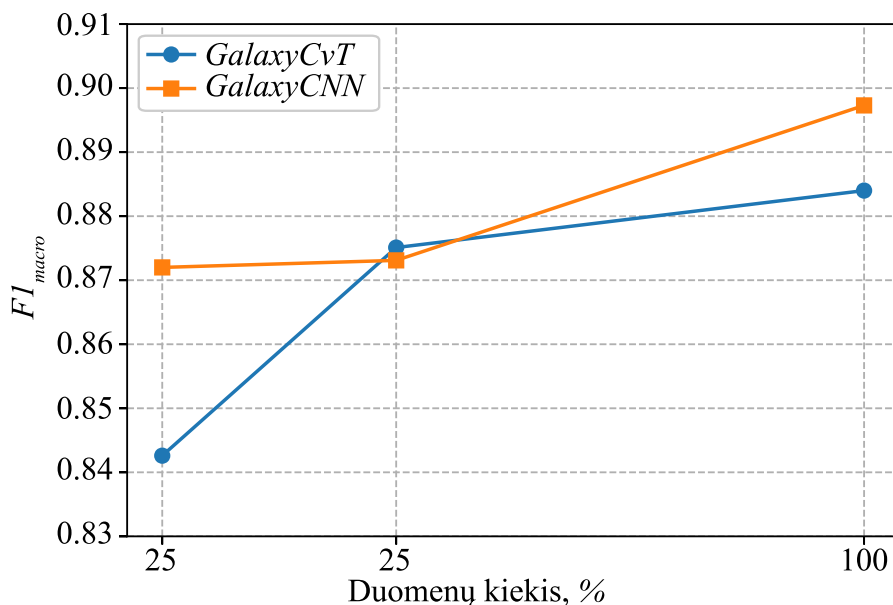
Modelis	Klasių kiekis	$F1_{macro}$	Tikslumas	Duomenų šaltiniai
1. <i>GalaxyCNN</i>	5	0.90	94.2%	GZ2, GZC, GZU, GZD, GZH
2. <i>GalaxyCvT</i>	5	0.88	93.5%	GZ2, GZC, GZU, GZD, GZH
3. [6]	5	0.95	94.68%	GZ2
4. [7]	3, 4	0.78, 0.67	82%, 77%	<i>Nair & Abraham (2010)</i> katalogas
5. [40]	5	0.98	98.8%	GZ2
6. [41]	5	0.97	96.83%	GZ2

Atsižvelgiant į klasifikavimo užduoties skirtumus, [6] tyrime išskirtos klasės sutampančios su klasėmis naudotomis šiame darbe. Buvo pastebėta, jog $F1$ vertės apvalioms elipsinėms ir cigaro formos elipsinėms galaktikoms *GalaxyCNN* pasiekiamos aukštesnės, galimai dėl didesnio šių klasių pavyzdžių kiekio šiame darbe naudotuose duomenyse. Tuo tarpu spiralinių galaktikų klasifikavimui – *GalaxyCNN* $F1$ vertė žemesnė. Tai gali būti siejama su papildomos spiralinių galaktikų klasės išskyrimu šiame darbe (žr. 14 pav.). Lyginant su [40] bei [41] autorių pristatytais modeliais, buvo pastebėta, jog *GalaxyCNN* atveju pastebima ryškesnė klaidinga klasifikacija tarp diskinės (stebimos šonu) ir cigaro formos elipsinės galaktikų klasių. Šis reiškinys tikėtina susijęs su šių klasių morfologiniais panašumais projekcijose — abejais atvejais dominuoja pailga struktūra, o diskinių galaktikų bruožai (pvz., plokštuma ar dulkės) ne visuomet yra aiškiai išskiriami esant triukšmui ar žemai skiriamajai gebai. Be to, šiame darbe naudotos nuotraukos sudarytos iš duomenų apimančių skirtingus stebėjimo instrumentus ir sąlygas, kas galėjo padidinti tarpklasinį variantiškumą ir apsunkinti šių klasių atskyrimą. Atsižvelgiant į tikslumo metriką, buvo pastebėta, jog modelis pasiekia panašią vertę kaip ir [6] pristatyta modelio architektūra.

Jautrumo duomenų kiekiui analizė

Papildomai buvo atlikta modelių jautrumo duomenų kiekiui analizė, kurios metu abu modeliai papildomai buvo apmokyti naudojant sumažintas mokymo duomenų aibes. 17 paveiksle pateikta *GalaxyCNN* ir *GalaxyCvT* modelių $F1_{macro}$ priklausomybė nuo duomenų kiekio.

Naudojant mažiausią mokymo duomenų rinkinį (25% viso mokymo duomenų kiekio) pastebėta, jog CNN architektūra išlaiko aukštesnes $F1_{macro}$ vertes nei konvoliuciniais transformatoriais paremtas modelis. Šie rezultatai atitinka teorinius lūkesčius ir leidžia daryti prielaidą, jog CNN pasižymi stipresniu induktyviniu šališkumu lokalumui, leidžiančiu efektyviau išskirti prasmingus prasmingus vietinius bruožus esant ribotam duomenų kiekiui. Tuo tarpu *GalaxyCvT* modeliui duomenų kiekis per mažas tinkamai išmokyti reikiamus bruožus išvengiant per didelio prisitaikymo.



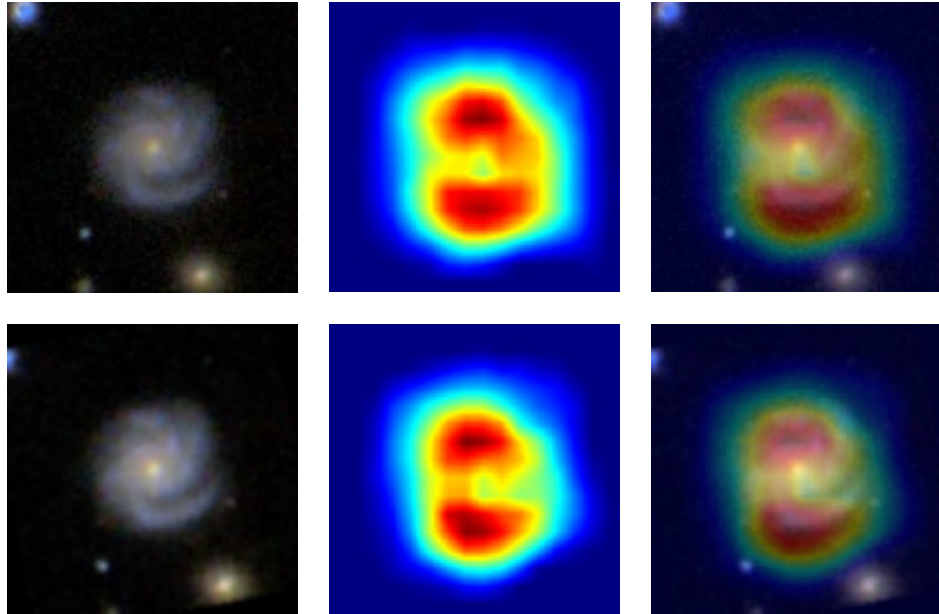
17 pav. *GalaxyCNN* ir *GalaxyCvT* modelių $F1_{macro}$ priklausomybė nuo mokymui naudojamų duomenų kiekio.

Naudojant 50% viso mokymo duomenų kiekio, matoma, jog konvoliuciniais transformatoriais paremtas modelis nežymiai lenkia CNN modelį. Galima daryti prielaidą, jog prie šio duomenų kiekio *GalaxyCvT* veikia efektyviausiai, o padvigubinus duomenų kiekį $F1_{macro}$ vertės pokytis palyginus nežymus. Čia taip pat svarbu atsižvelgti ir į tai, jog pristatyta CvT architektūra yra palyginti sekli. *GalaxyCNN* įverčio pokytis naudojant 50% ir 100% duomenų žymiai ryškesnis ir rodo didesnę priklausomybę nuo duomenų kiekio siekiant maksimalios apibendrinimo kokybės. Tai atskleidžia skirtingą abiejų architektūrų elgseną esant ribotam mokymo duomenų kiekiui ir pabrėžia CNN architektūrų pranašumą mažų duomenų scenarijuose.

Modelių sprendimų vizualinio stabilumo analizė

Pirmajame eksperimente kiekvienam modeliui buvo parinkta teisingai klasifikuota galaktikos nuotrauka. *GalaxyCNN* modeliui, šios nuotraukos atvaizdavimas buvo atliktas naudojant *Grad-CAM*, o *GalaxyCvT* – *Attention Rollout* metodą. Tuomet nuotrauka buvo papildomai transformuojama pritaikant nedidelio laipsnio pasukimą, bei *Gaussian Blur*. Originalių ir transformuotų nuotraukų bei atitinkamų vizualizacijų palyginimas pateiktas 18 ir 19 paveiksluose.

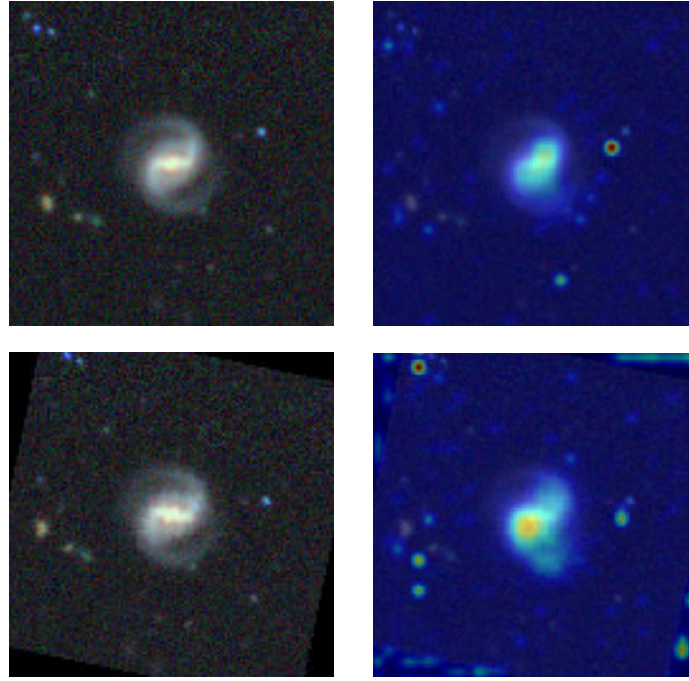
18 paveiksle galima pastebėti, jog *GalaxyCNN* modelis koncentruojasi į sritis kuriose ryškiausiai matomi spiralių bruožai, tuo tarpu centrinis telkinys šiame pavyzdyje nėra dominuojantis sprendimą lemiantis bruožas. Toks susitelkimas atitinka CNN architektūroms būdingą lokalų erdvinį jautrumą.



18 pav. *GalaxyCNN* modelio prognozės *Grad-CAM* vizualizacijos palyginimas (viršutinėje eilėje originalus paveikslas ir jo vizualizacijos, antroje eilėje – transformuotas paveikslas ir jo vizualizacijos). Trečiame stulpelyje *Grad-CAM* vizualizacija perdengta nuotraukos vaizdu.

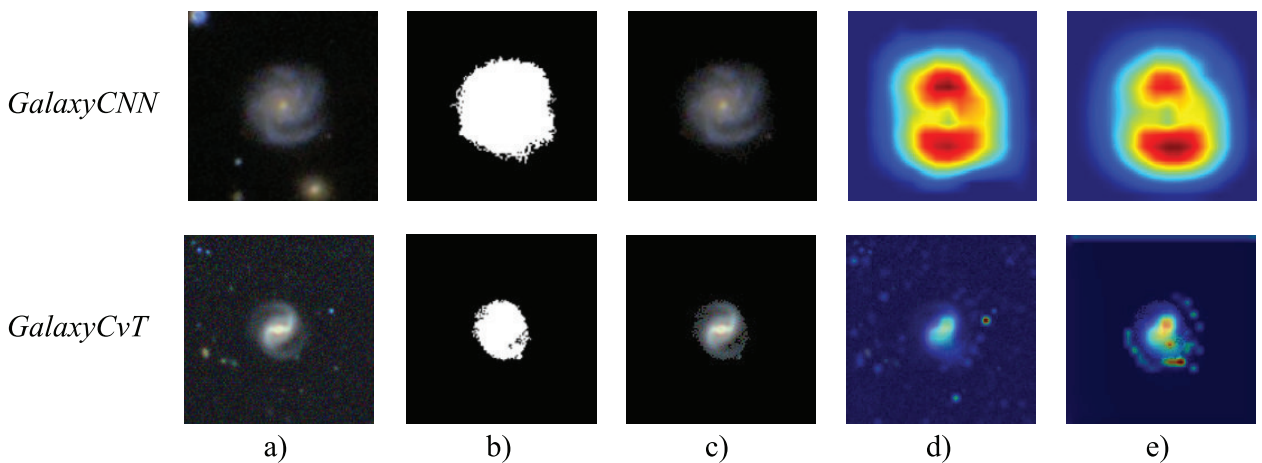
Attention Rollout vizualizacijoje matoma jog modelis nežymiai susitelkia į spiralių motyvą, tačiau taip pat pastebima ir tai, jog skirtingai nei CNN architektūroje taip pat gana ryškiai išskiria ir kitus nuotraukoje esančius objektus. *GalaxyCvT* skirtingai nei CNN atveju, dėmesys nėra aiškiai lokalizuotas tik ties pagrindiniu tiriamu objektu. Taip pat, šiam modeliui papildomai buvo sukurtos vizualizacijos vizualizacijos iš ankstyvųjų ir vėlesniųjų sluoksnių (paveiksluose nepateikiamos): ankstyvuosiuose sluoksniuose išryškėja aiškesni lokalūs bruožai, tuo tarpu gilesniuose sluoksniuose vizualizacijos tampa fragmentiškesnės ir sunkiau interpretuojamos.

GalaxyCNN modelyje susitelkimas į bruožus po nuotraukos transformacijos reikšmingai nepakito, o struktūrinio panašumo indeksas (angl. *Structural Similarity Index Measure*), *SSIM* tarp originalios ir transformuotos vizualizacijos siekė 0.9115, kas rodo aukštą sprendimų vizualizacijų stabilumą. Tuo tarpu *GalaxyCvT* modelio *Attention Rollout* vizualizacijų *SSIM* reikšmė buvo gerokai mažesnė (0.3727), indikuojant didesnę vizualinę jautrumą nedidelėms įvesties transformacijoms.



19 pav. *GalaxyCvT* modelio prognozės *Attention Rollout* vizualizacijos palyginimas (viršutinėje eilėje originalus paveikslas ir jo vizualizacijos, antroje eilėje – transformuotas paveikslas ir jo vizualizacijos).

Toliau buvo analizuojama, kokią įtaką modelių sprendimų vizualizacijoms daro vaizdinis triukšmas. Pavyzdinės nuotraukos, apdorotos pasitelkiant *astropy* biblioteką (panaikinant nuotraukos foną, pikselio vidutinis nuokrypis $n\sigma = 3.0$), bei modelių vizualizacijos pateikiamos 20 paveiksle. *GalaxyCNN* SSIM = 0.9151 lyginant originalaus ir transformuotos paveikslo vizualizacijas, parodo jog modelio sprendimams didžiausią įtaką daro tiriamo objekto sritys, o foninės struktūros turi mažesnę indėlį.



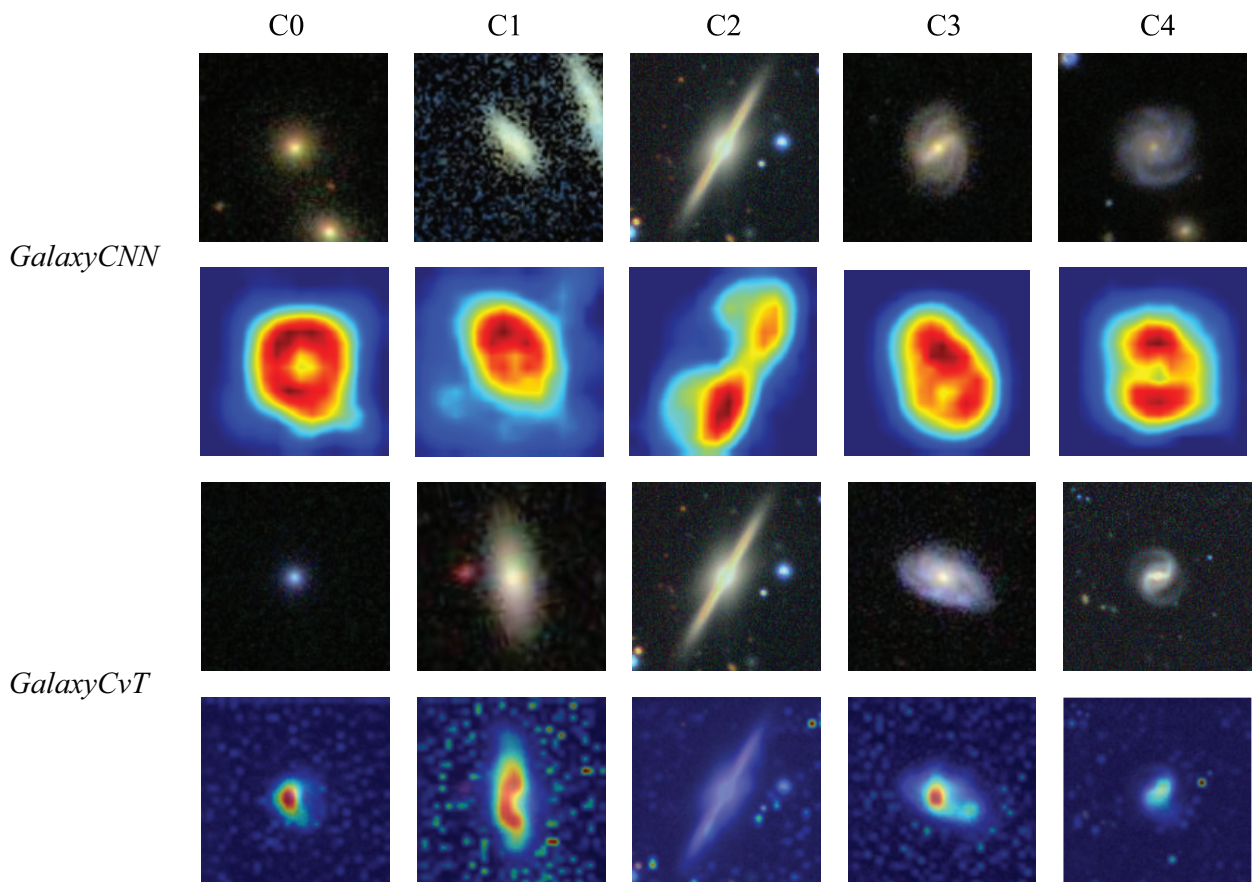
20 pav. Modelių prognozės vizualizacijų palyginimas (a) originalios nuotraukos, b) sukurti nuotraukų filtrai (angl. *mask*), c) transformuotos nuotraukos, d) originalių nuotraukų vizualizacijos, e) transformuotų nuotraukų vizualizacijos).

GalaxyCvT modelio atveju SSIM = 0.3284 lyginant originalaus ir transformuotos paveikslo vizualizacijas, d) ir e) paveiksluose matomi žymesni pasikeitimai – po transformacijos modelis

susitelkia į galaktikos vijų dalį bei išskiria statesnį vijos pasisukimą nuo centrinio telkinio (galimai skersėtumo požymiui išskirti).

Modelių sprendimų vizualizacijos kiekvienai klasei pateikiamos 21 paveiksle. *GalaxyCNN Grad-CAM* vizualizacijose pastebima, jog modelio susitelkimas skiriasi priklausomai nuo klasės, spiralinėms galaktikoms svarbesnės zonos nėra tolygios ir susitelkia į vijas kurios vizualiai ryškesnės, tuo tarpu apvalių elipsinių galaktikų atvejų susitelkimas pasiskirstęs tolygiau per galaktikos kontūrus. Šiuose pavyzdžiuose centrinis telkinys nėra dominuojantis sprendimą lemiantis bruožas.

GalaxyCvT Attention Rollout vizualizacijose matoma, jog triukšmai nuotraukose dažniausiai išskiriami gana stipriai. elipsinių galaktikų (cigaro formos) vizualizacijoje pastebimas didesnis susitelkimas į dangaus kūno formą, tuo tarpu skersinių spiralinių galaktikų pavyzdžiuose dėmesys dažnai koncentruojasi į lokalias sritis, kurių aiškūs morfologiniai skiriamieji bruožai vizualiai nėra akivaizdūs.



21 pav. Modelių prognozės kiekvienai klasei vizualizacijų palyginimas.

Apibendrinant, abiejų modelių sprendimų vizualizacijos atskleidžia skirtingą sprendimų formavimo pobūdį: CNN architektūroje stebimas labiau lokalus ir vizualiai stabilesnis dėmesio pasiskirstymas, tuo tarpu CvT modelyje dėmesys yra labiau išskaidytas ir jautresnis įvesties pokyčiams. Šie rezultatai atitinka nagrinėjamų architektūrų teorinius skirtumus ir pabrėžia skirtingą jų paaiškinamumo savybių elgseną.

Skaičiavimo efektyvumo analizė

Taip pat, buvo atlikta modelių skaičiavimo efektyvumo analizė, kurios metu palyginta mokymo ir prognozavimo sparta. Lentelėje 8 pateikti abiejų modelių vertinimo rodikliai. Mokymo rodikliai

apskaičiuoti iš duomenų surinktų mokymo metu esant įprastoms sąlygoms (naudojant visa mokymo duomenų aibę, ir parametrus nurodytus 4.2 poskyryje).

Prognozavimo našumui vertinti buvo atliktas papildomas tyrimas, kurio metu modeliai gavo testavimo duomenų rinkinį sudarytą iš 300 tenzorių. Sistemos, naudotos tyrimo metu įranga nurodyta 4.1 poskyryje. Kadangi naudotos vaizdo plokštės architektūra buvo tinkama, mokymo metu buvo naudojama AMP (angl. *Automatic Mixed Precision*).

8 lentelė. *GalaxyCNN* ir *GalaxyCvT* modelių skaičiavimo našumo palyginimas.

Rodiklis	<i>GalaxyCNN</i>	<i>GalaxyCvT</i>	Palyginimas
Parametrų skaičius	$2.80 \cdot 10^6$	$3.14 \cdot 10^6$	CvT +12%
Mokymo sparta	2.44 it/s	2.82 it/s	CvT +16%
Mokymo ciklo stabilumas	0.015 it/s	0.117 it/s	CNN stabilesnis
Prognozės trukmės vidurkis	2.36 ms	2.74 ms	CNN –14%
Prognozės trukmės nuokrypis	0.49 ms	0.115 ms	CvT stabilesnis
Prognozės sparta	424 pav./s	365 pav./s	CNN +16%

Nors *GalaxyCNN* parametrų kiekis mažesnis (apie 12%), mokymo sparta yra maždaug 16% mažesnė nei *GalaxyCvT* modelio. Šis skirtumas gali būti paaiškintas architektūriniais skirtumais – CNN tinklai remiasi konvoliucijos operacijomis bei stambiais tarpiniais bruožų žemėlapiais, tuo tarpu CvT architektūroje daugiausia naudojama matricių daugybos operacija, kuri yra labiau optimizuota šiuolaikinėse vaizdo plokštėse.

Nepaisant didesnės *GalaxyCvT* mokymo spartos, prognozavimo trukmė *GalaxyCNN* modelyje yra apie 14% mažesnė. Taip pat pastebima, jog *GalaxyCvT* prognozavimo sparta stabilesnė (mokymo ciklo stabilumas – mokymo spartos vidutinis nuokrypis), tuo tarpu geresnis mokymo stabilumas matomas *GalaxyCNN* tinkle.

Išvados ir rekomendacijos

- Rezultatai parodo, jog pasiūlytas konvoliucinis neuroninis tinklas (*GalaxyCNN*) pasiekia nežymiai, tačiau nuosekliai aukštesnius bendrus klasifikavimo rezultatus nei konvoliuciniu transformatoriumi paremtas *GalaxyCvT* modelis – *GalaxyCNN* $F1_{macro} = 0.897$, *GalaxyCvT* $F1_{macro} = 0.884$. Šis pranašumas ypač ryškus sudėtingesnėse morfologinėse klasėse, kuriose reikalingas detalių lokalių struktūrų atpažinimas.
- Dauguma neteisingų klasifikacijų pastebimos tarp morfologiškai panašių galaktikų klasių, kurių vizualiniai bruožai persidengia net žmogaus vertinime. Tai leidžia manyti, jog modelių klaidos dažniausiai yra nulemtos duomenų morfologinio neapibrėžtumo, o ne architektūrinių trūkumų.
- Jautrumo duomenų kiekiui analizė parodė, jog pasiūlyta CNN architektūra išlaiko aukštesnius klasifikavimo rodiklius esant ribotam mokymo duomenų kiekiui. Tai indikuoja stipresnį CNN induktyvinį šališkumą lokalumui, kuris leidžia efektyviau išmokti morfologinius bruožus mažesnėse duomenų aibėse.
- Nors *GalaxyCvT* modelio veikimas pagerėja didėjant mokymo duomenų kiekiui, nagrinėjamomis sąlygomis jis nepasiekia pasiūlyto CNN modelio bendro klasifikavimo našumo. Tai rodo, jog CvT potencialas, šiame darbe nagrinėtomis sąlygomis, nėra pilnai išnaudotas dėl architektūros seklumo ir ribotų duomenų.
- CNN modelio sprendimų vizualizacijos pasižymi aukštu erdvinio stabilumu esant nedidelioms įvesties transformacijoms bei sumažintam foniniam triukšmui. Tai rodo, jog modelio sprendimai grindžiami nuosekliais, lokaliais morfologiniais bruožais, nepriklausančiais nuo nereikšmingų foninių struktūrų. CvT modelio sprendimų vizualizacijose stebimas labiau išskaidytas dėmesys bei didesnis jautrumas nedideliems įvesties pokyčiams. Dėmesio pasiskirstymas dažnai apima ne tik tiriamą galaktiką, bet ir foninius objektus, kas rodo mažesnį sprendimų erdvinį nuoseklumą.
- Skaičiavimo našumo analizė parodė, jog CvT modelis pasižymi didesne mokymo sparta, tuo tarpu CNN architektūra yra efektyvesnė prognozavimo metu. Šie skirtumai tiesiogiai susiję su architektūrų skaičiavimo pobūdžiu ir turi praktinę reikšmę renkantis modelį realioms taikymo sąlygoms.

Ateities tyrimų planas

Atsižvelgiant į gautus rezultatus, tolimesni tyrimai galėtų būti orientuoti į šias kryptis:

1. **Sudėtingų morfologinių klasių analizė ir architektūros adaptacija** – atsižvelgiant į tai, jog didžiausias klasifikavimo klaidų skaičius stebimas tarp morfologiškai panašių klasių (C1–C2 ir C3–C4), ateities tyrimuose būtų tikslinga nagrinėti architektūrinius ar mokymo strategijų sprendimus, orientuotus į šių klasių atskyrimą, pavyzdžiui taikant daugiapakopį bruožų išgavimą (angl. *multi-scale feature extraction*).
2. **Modelio sprendimų paaiškinamumo analizės išplėtimas** – šiame darbe nustatyti modelio veikimo skirtumai tarp skirtingų duomenų rinkinių (ypač didelio raudonojo poslinkio ir NIR spektro duomenų atveju) rodo poreikį detaliau tirti, ar CNN modeliai išmoksta instrumentams nepriklausomas morfologines reprezentacijas. Ateityje būtų tikslinga tirti sprendimų vizualizacijų stabilumą esant skirtingiems duomenų šaltiniams. Tokie rodikliai kaip SSIM galėtų būti naudojami ne tik interpretacijai, bet ir kaip papildomas modelių palyginimo kriterijus šalia klasifikavimo metrikų.
3. **Fotometrinių ir spektroskopinių savybių analizės išplėtimas** – atsižvelgiant į pastebėtus klasifikavimo skirtumus priklausomai nuo naudojamo duomenų rinkinio, tolimesni tyrimai galėtų būti orientuoti į detalesnę galaktikų fizinių savybių analizę, naudojant prieinamus fotometrinius ir spektroskopinius metaduomenis. Šiuo metu tokia analizė yra apribota, kadangi dalies galaktikų raudonasis poslinkis nustatytas spektroskopiškai, kitų – fotometriškai, o kai kurių naudotų duomenų rinkinių (pvz., *Galaxy Zoo UKIDSS*) išsamūs metaduomenys dar nėra prieinami.

Literatūros šaltiniai

- [1] E. P. Hubble. Extragalactic nebulae. *The Astrophysical Journal*, 64:321, dec 1926. DOI: 10.1086/143018
- [2] Chris J. Lintott, Kevin Schawinski, Anže Slosar, et al. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey*. *Monthly Notices of the Royal Astronomical Society*, 389(3):1179–1189, 09 2008. DOI: 10.1111/j.1365-2966.2008.13689.x
- [3] Chris Lintott, Kevin Schawinski, Steven Bamford, et al. Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies*. *Monthly Notices of the Royal Astronomical Society*, 410(1):166–178, 12 2011. DOI: 10.1111/j.1365-2966.2010.17432.x
- [4] Michael Jordan Raddick, Edward E. Prather, and Colin S. Wallace. Galaxy zoo: Science content knowledge of citizen scientists. *Public Understanding of Science*, 28(6):636–651, may 2019. DOI: 10.1177/0963662519840222
- [5] Jiale Lei and Linghe Kong. *Fundamentals of big data in radio astronomy*, pages 29–58. Elsevier, 2020. DOI: 10.1016/b978-0-12-819084-5.00010-9
- [6] Xiao-Pan Zhu, Jia-Ming Dai, Chun-Jiang Bian, et al. Galaxy morphology classification with deep convolutional neural networks. *Astrophysics and Space Science*, 364(4), April 2019. DOI: 10.1007/s10509-019-3540-1
- [7] Mitchell K Cavanagh, Kenji Bekki, and Brent A Groves. Morphological classification of galaxies with deep learning: comparing 3-way and 4-way CNNs. *Monthly Notices of the Royal Astronomical Society*, 506(1):659–676, 06 2021. DOI: 10.1093/mnras/stab1552
- [8] Sander Dieleman, Kyle W. Willett, and Joni Dambre. Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Monthly Notices of the Royal Astronomical Society*, 450(2):1441–1459, 04 2015. DOI: 10.1093/mnras/stv632
- [9] Nour Eldeen M. Khalifa, Mohamed Hamed N. Taha, Aboul Ella Hassanien, and I. M. Selim. Deep galaxy: Classification of galaxies based on deep convolutional neural networks, 2017. DOI: 10.48550/ARXIV.1709.02245
- [10] V. Khramtsov, D. V. Dobrycheva, M. Yu. Vasylenko, and V. S. Akhmetov. DEEP LEARNING FOR MORPHOLOGICAL CLASSIFICATION OF GALAXIES FROM SDSS. *Odessa Astronomical Publications*, 32(0):21–23, oct 2019. DOI: 10.18524/1810-4215.2019.32.182092
- [11] Edwin Powell Hubble. *The realm of the Nebulae*. Yale University Press, 1982.
- [12] Karen L Masters, Chris J Lintott, Ross E Hart, et al. Galaxy zoo: unwinding the winding problem – observations of spiral bulge prominence and arm pitch angles suggest local spiral galaxies are winding. *Monthly Notices of the Royal Astronomical Society*, 487(2):1808–1820, apr 2019. DOI: 10.1093/mnras/stz1153
- [13] Elmegreen Bruce G. Elmegreen Debra Meloy. Arm classifications for spiral galaxies. *Astrophysical Journal* v.314, p.3, 314:3, 1987. DOI: 10.1086/165034

- [14] Gérard de Vaucouleurs. *Classification and Morphology of External Galaxies*. Springer, Berlin, Heidelberg, 1959. DOI: https://doi.org/10.1007/978-3-642-45932-0_7
- [15] Kyle W. Willett, Chris J. Lintott, Steven P. Bamford, et al. Galaxy zoo 2: detailed morphological classifications for 304 122 galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 435(4):2835–2860, September 2013. DOI: 10.1093/mnras/stt1458
- [16] Ross E. Hart, Steven P. Bamford, Kyle W. Willett, et al. Galaxy Zoo: comparing the demographics of spiral arm number and a new method for correcting redshift bias. *Monthly Notices of the Royal Astronomical Society*, 461(4):3663–3682, 07 2016. DOI: 10.1093/mnras/stw1588
- [17] Kevork N. Abazajian, Jennifer K. Adelman-McCarthy, Marcel A. Agüeros, et al. THE SEVENTH DATA RELEASE OF THE SLOAN DIGITAL SKY SURVEY. *The Astrophysical Journal Supplement Series*, 182(2):543–558, may 2009. DOI: 10.1088/0067-0049/182/2/543
- [18] Kyle W. Willett, Melanie A. Galloway, Steven P. Bamford, et al. Galaxy zoo: morphological classifications for 120 000 galaxies in *HST* legacy imaging. *Monthly Notices of the Royal Astronomical Society*, 464(4):4176–4203, oct 2017. DOI: 10.1093/mnras/stw2568
- [19] Roger L. Griffith, Michael C. Cooper, Jeffrey A. Newman, et al. THE ADVANCED CAMERA FOR SURVEYS GENERAL CATALOG: STRUCTURAL PARAMETERS FOR APPROXIMATELY HALF a MILLION GALAXIES. *The Astrophysical Journal Supplement Series*, 200(1):9, may 2012. DOI: 10.1088/0067-0049/200/1/9
- [20] B. D. Simmons, Chris Lintott, Kyle W. Willett, et al. Galaxy zoo: quantitative visual morphological classifications for 48 000 galaxies from candels. *Monthly Notices of the Royal Astronomical Society*, 464(4):4420–4447, October 2016. DOI: 10.1093/mnras/stw2587
- [21] Norman A. Grogin, Dale D. Kocevski, S. M. Faber, et al. CANDELS: THE COSMIC ASSEMBLY NEAR-INFRARED DEEP EXTRAGALACTIC LEGACY SURVEY. *The Astrophysical Journal Supplement Series*, 197(2):35, dec 2011. DOI: 10.1088/0067-0049/197/2/35
- [22] Mike Walmsley, Chris Lintott, Tobias Géron, et al. Galaxy zoo DECaLS: Detailed visual morphology measurements from volunteers and deep learning for 314 000 galaxies. *Monthly Notices of the Royal Astronomical Society*, 509(3):3966–3988, sep 2021. DOI: 10.1093/mnras/stab2093
- [23] Arjun Dey, David J. Schlegel, Dustin Lang, et al. Overview of the DESI legacy imaging surveys. *The Astronomical Journal*, 157(5):168, apr 2019. DOI: 10.3847/1538-3881/ab089d
- [24] Karen L. Masters, Melanie Galloway, Lucy Fortson, et al. Galaxy zoo: Morphologies based on ukidss nir imaging for 71,052 galaxies. *Research Notes of the AAS*, 8(8):198, aug 2024. DOI: 10.3847/2515-5172/ad6f10
- [25] A. Lawrence, S. J. Warren, O. Almaini, et al. The ukirt infrared deep sky survey (ukidss). *Monthly Notices of the Royal Astronomical Society*, 379(4):1599–1617, August 2007. DOI: 10.1111/j.1365-2966.2007.12040.x

- [26] S. J. Warren, N. C. Hambly, S. Dye, et al. The united kingdom infrared telescope infrared deep sky survey first data release. *Monthly Notices of the Royal Astronomical Society*, 375(1):213–226, 01 2007. DOI: 10.1111/j.1365-2966.2006.11284.x
- [27] Vahe Petrosian. Surface brightness and evolution of galaxies. *The Astrophysical Journal*, 210:L53, December 1976. DOI: 10.1086/182301
- [28] Steven P. Bamford, Robert C. Nichol, Ivan K. Baldry, et al. Galaxy zoo: the dependence of morphology and colour on environment. *Monthly Notices of the Royal Astronomical Society*, 393(4):1324–1352, March 2009. DOI: 10.1111/j.1365-2966.2008.14252.x
- [29] M. Davis, P. Guhathakurta, N. P. Konidaris, et al. The all-wavelength extended groth strip international survey (aegis) data sets. *The Astrophysical Journal*, 660(1):L1–L6, April 2007. DOI: 10.1086/517931
- [30] M. Giavalisco, H. C. Ferguson, A. M. Koekemoer, et al. The great observatories origins deep survey: Initial results from optical and near-infrared imaging. *The Astrophysical Journal*, 600(2):L93–L98, January 2004. DOI: 10.1086/379232
- [31] Hans-Walter Rix, Marco Barden, Steven V. W. Beckwith, et al. Gems: Galaxy evolution from morphologies and seds. *The Astrophysical Journal Supplement Series*, 152(2):163–173, June 2004. DOI: 10.1086/420885
- [32] John A. R. Caldwell, Daniel H. McIntosh, Hans-Walter Rix, et al. Gems survey data and catalog. *The Astrophysical Journal Supplement Series*, 174(1):136–144, January 2008. DOI: 10.1086/521080
- [33] N. Scoville, H. Aussel, M. Brusa, et al. The cosmic evolution survey (cosmos): Overview. *The Astrophysical Journal Supplement Series*, 172(1):1–8, September 2007. DOI: 10.1086/516585
- [34] A. M. Koekemoer, H. Aussel, D. Calzetti, et al. The cosmos survey: Hubble space telescope advanced camera for surveys observations and data processing. *The Astrophysical Journal Supplement Series*, 172(1):196–202, September 2007. DOI: 10.1086/520086
- [35] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *ICLR 2015*, 2015. DOI: <https://doi.org/10.48550/arXiv.1409.1556>
- [36] Cavanagh, M. K. and Bekki, K. Bars formed in galaxy merging and their classification with deep learning. *A&A*, 641:A77, 2020. DOI: 10.1051/0004-6361/202037963
- [37] Matthew D. Zeiler. ADADELTA: an adaptive learning rate method. *CoRR*, abs/1212.5701, 2012. DOI: <https://doi.org/10.48550/arXiv.1212.5701>
- [38] François Chollet. Xception: Deep learning with depthwise separable convolutions. *CoRR*, abs/1610.02357, 2016. DOI: <https://doi.org/10.48550/arXiv.1610.02357>
- [39] Joshua Yao-Yu Lin, Song-Mao Liao, Hung-Jin Huang, et al. Galaxy morphological classification with efficient vision transformer. October 2021. DOI: 10.48550/ARXIV.2110.01024

- [40] Jie Cao, Tingting Xu, Yuhe Deng, et al. Galaxy morphology classification based on convolutional vision transformer (cvt). *Astronomy and Astrophysics*, 683:A42, March 2024. DOI: 10.1051/0004-6361/202348544
- [41] Raul Urechiatu and Marc Frincu. Improved galaxy morphology classification with convolutional neural networks. *Universe*, 10(6), 2024. DOI: 10.3390/universe10060230
- [42] Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola. Dive into deep learning. *CoRR*, abs/2106.11342, 2021. DOI: <https://doi.org/10.48550/arXiv.2106.11342>
- [43] Tsung-Yi Lin, Priya Goyal, Ross Girshick, et al. Focal loss for dense object detection, 2018.
- [44] Keiron O’Shea and Ryan Nash. An introduction to convolutional neural networks. *CoRR*, abs/1511.08458, 2015. DOI: <https://doi.org/10.48550/arXiv.1511.08458>
- [45] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. December 2015. DOI: 10.48550/ARXIV.1512.03385
- [46] Ashish Vaswani, Noam Shazeer, Niki Parmar, et al. Attention is all you need, 2023.
- [47] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, et al. An image is worth 16x16 words: Transformers for image recognition at scale, 2021.
- [48] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, et al. End-to-end object detection with transformers, 2020.
- [49] Hanting Chen, Yunhe Wang, Tianyu Guo, et al. Pre-trained image processing transformer, 2021.
- [50] Huiyu Wang, Yukun Zhu, Hartwig Adam, et al. Max-deeplab: End-to-end panoptic segmentation with mask transformers, 2021.
- [51] Hugo Touvron, Matthieu Cord, Matthijs Douze, et al. Training data-efficient image transformers distillation through attention, 2021.
- [52] Haiping Wu, Bin Xiao, Noel Codella, et al. Cvt: Introducing convolutions to vision transformers. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 22–31. IEEE, October 2021. DOI: 10.1109/iccv48922.2021.00009
- [53] Kanae Takahashi, Kouji Yamamoto, Aya Kuchiba, and Tatsuki Koyama. Confidence interval for micro-averaged F 1 and macro-averaged F 1 scores. *Appl. Intell.*, 52(5):4961–4972, mar 2022. DOI: doi.org/10.1007/s10489-021-02635-5
- [54] Mike Walmsley, Steven P. Bamford, Hugh Dickinson, et al. Galaxy zoo evo: 1 million human-annotated images of galaxies, 2025.
- [55] Mike Walmsley, Tobias Géron, Sandor Kruk, et al. Galaxy zoo desi: Detailed morphology measurements for 8.7m galaxies in the desi legacy imaging surveys. *Monthly Notices of the Royal Astronomical Society*, 526(3):4768–4786, September 2023. DOI: 10.1093/mnras/stad2919

- [56] Tomas Mūžas, Andrius Vytautas Misiukas Misiūnas, and Tadas Meškauskas. Large scale study of binary galaxy image classification and the impact of image augmentation techniques. In Osvaldo Gervasi, Beniamino Murgante, David Taniar, et al., editors, *Computational Science and Its Applications – ICCSA 2023*, pages 402–412, Cham, 2023. Springer Nature Switzerland.
- [57] Vicki Kuhn, Yicheng Guo, Alec Martin, et al. Jwst reveals a surprisingly high fraction of galaxies being spiral-like at $0.5 \leq z \leq 4$, 2024.