



Kauno technologijos universitetas
Matematikos ir gamtos mokslų fakultetas

Skurdo rizikos lygio Lietuvoje nustatymas iš satelitinių nakties vaizdų

Baigiamasis magistro studijų projektas

Roberta Budnikaitė

Projekto autorė

Prof. dr. Evaldas Vaičiukynas

Vadovas

Doc. dr. Meda Andrijauskienė

Vadovė

Kaunas, 2024



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Skurdo rizikos lygio Lietuvoje nustatymas iš satelitinių nakties vaizdų

Baigiamasis magistro studijų projektas

Didžiųjų verslo duomenų analitika (6213AX001)

Roberta Budnikaitė

Projekto autorė

Prof. dr. Evaldas Vaičiukynas

Vadovas

Doc. dr. Meda Andrijauskienė

Vadovė

Doc. dr. Tomas Iešmantas

Recenzentas

Doc. dr. Raminta Vaitiekūnienė

Recenzentė

Kaunas, 2024



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Roberta Budnikaitė

Skurdo rizikos lygio Lietuvoje nustatymas iš satelitinių nakties vaizdų

Akademinio sąžiningumo deklaracija

Patvirtinu, kad:

1. baigiamąjį projektą parengiau savarankiškai ir sąžiningai, nepažeisdama(s) kitų asmenų autoriaus ar kitų teisių, laikydamasi(s) Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo nuostatų, Kauno technologijos universiteto (toliau – Universitetas) intelektinės nuosavybės valdymo ir perdavimo nuostatų bei Universiteto akademinės etikos kodekse nustatytų etikos reikalavimų;
2. baigiamajame projekte visi pateikti duomenys ir tyrimų rezultatai yra teisingi ir gauti teisėtai, nei viena šio projekto dalis nėra plagijuota nuo jokių spausdintinių ar elektroninių šaltinių, visos baigiamojo projekto tekste pateiktos citatos ir nuorodos yra nurodytos literatūros sąrašė;
3. įstatymų nenumatytų piniginių sumų už baigiamąjį projektą ar jo dalis niekam nesu mokėjęs (-usi);
4. suprantu, kad išaiškėjus nesąžiningumo ar kitų asmenų teisių pažeidimo faktui, man bus taikomos akademinės nuobaudos pagal Universitete galiojančią tvarką ir būsiu pašalinta(s) iš Universiteto, o baigiamasis projektas gali būti pateiktas Akademinės etikos ir procedūrų kontrolieriaus tarnybai nagrinėjant galimą akademinės etikos pažeidimą.

Roberta Budnikaitė

Patvirtinta elektroniniu būdu

Budnikaitė, Roberta. Skurdo rizikos lygio Lietuvoje nustatymas iš satelitinių nakties vaizdų. Magistro studijų baigiamasis projektas / vadovas prof. dr. Evaldas Vaičiukynas, vadovė doc. dr. Meda Andrijauskienė; Kauno technologijos universitetas, Matematikos ir gamtos mokslų fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Taikomoji matematika (Matematikos mokslai).

Reikšminiai žodžiai: konvoliuciniai neuroniniai tinklai, *Elastic Net* regresija, skurdo rizikos lygis, *Black Marble*, nakties šviesos.

Kaunas, 2024. 81 p.

Santrauka

Tiksli ir patikima skurdo statistika yra vienas svarbiausių faktorių, siekiant parengti efektyvią socialinę politiką, mažinti pajamų nelygybę bei atotrūkį tarp turtingiausiųjų ir skurdžiausiųjų gyventojų. Visgi, įprasti socialinių ir ekonominių duomenų gavimo metodai dažnai užima daug laiko bei išteklių ir yra brangūs. Nors šiuo metu Lietuvos oficialiosios statistikos portalas siūlo eksperimentinę statistiką, apimančią metinius skurdo rizikos rodiklius, prieiga prie agreguotų skurdo rizikos lygio duomenų galėtų padėti geriau suprasti skurdo dinamiką ir tendencijas.

Giliojo mokymosi architektūros vis dažniau taikomos įvairioms kompiuterinės regos problemoms spręsti. Konvoliuciniai neuroniniai tinklai (CNN) naudojami ir palydovinių vaizdų analizėje, kur agreguotas nakties šviesos intensyvumas atskleidžia socialinės ir ekonominės veiklos lygius. Ankstesni tyrimai parodė, kad nuotolinio stebėjimo metodai yra tinkamas būdas ir skurdo lygiui įvertinti.

Šiame tyrime naudojamas NASA „Black Marble“ produktų rinkinys, teikiantis koreguotus aukštos kokybės NPP-VIIRS vaizdus. Šio produktų rinkinio algoritmai geba atpažinti ir pašalinti atmosferos, reljefo, mėnulio šviesos atspindžių, sniego, augmenijos bei šilumos efektus, veikiančius mūsų matomus šviesas, ir sukurti nuoseklius kasdienius žmogaus sukurtų šviesų duomenis.

Eksperimento metu iš NASA LAADS DAAC archyvų gaunami mėnesiniai NTL vaizdai, apimantys 2013–2023 m. periodą. Jie paruošiami CNN mokymui: interpoliuojami, normalizuojami ir vėliau supjaustomi į 56×56 , 112×112 ir 224×224 pikselių gardeles. Siekiant padidinti mokymosi imtį, vaizdai taip pat perstumiami 50 % x bei y ašimi. Tokiu būdu vaizdų skaičius padidinamas atitinkamai iki 54 780, 15 840 bei 5 016. Papildomai taikyta požymių inžinerija – įtraukiami požymiai apie savivaldybės vietą vaizde bei praeitų metų skurdo rodiklius. Eksperimente pakartotinai naudojamos 6 CNN architektūros (*NASNet*, *ResNet-50*, *ResNet-101*, *DenseNet-101*, *DenseNet-201* ir *EfficientNet-B0*) bei mokomi skirtingi modeliai – *Elastic Net* regresija ir atsitiktiniai miškai.

Geriausi rezultatai, prognozuojant mėnesinius skurdo rizikos lygio rodiklius, buvo gauti naudojant *ResNet-50* vektorizavimo metodą, *Elastic Net* modelį mokant su mažiausiais nepersidengiančiais 56×56 dydžio vaizdais, įtraukiant papildomus požymius. Čia testavimo $R^2 = 0,510$, MDAE – 1,960, o MAPE – 9,811.

Budnikaitė, Roberta. Estimation of Poverty Rate in Lithuania from Night-time Light Imagery. Master's Final Degree Project / supervisor prof. dr., Evaldas Vaičiukynas, supervisor doc. dr., Meda Andrijauskienė; Faculty of Mathematics and Natural Sciences, Kaunas University of Technology.

Study field and area (study field group): Applied Mathematics (Mathematical Sciences).

Keywords: convolutional neural networks, Elastic Net regression, poverty level, Black Marble, nighttime lights.

Kaunas, 2024. 81.

Summary

Accurate and reliable poverty statistics are essential for policy makers to develop effective programs and interventions to reduce income inequality and improve social assistance. However, conventional methods of acquiring socioeconomic data are often time-consuming, resource-intensive, and expensive. Specifically, the Lithuanian Official Statistics Portal offers only annual indicators of at-risk-of-poverty. Access to monthly data could lead to a deeper understanding of poverty dynamics and facilitate forecasting.

Convolutional deep learning architectures are increasingly applied to various computer vision problems, including satellite imagery, where aggregated nightlight luminosity is of particular importance as a proxy of socioeconomic activities. Previous studies have shown that remote sensing techniques offer a viable method for assessing poverty levels. In particular, the utilization of convolutional neural networks (CNNs) to analyze satellite imagery has demonstrated potential in forecasting the luminosity of nighttime lights, which can then be used to determine the underlying poverty levels.

This study leverages the NASA Black Marble product suite, which offers cloud-free satellite imagery retrieved from Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. These Night-time Light (NTL) images are corrected for various factors including atmospheric conditions, terrain, lunar BRDF effects, thermal and straylight interferences. Thus, they effectively capture the presence of human activity through artificial lighting.

During the experiment, monthly NTL images spanning from 2013 to 2023 are retrieved from NASA LAADS DAAC archives and prepared for CNN training: interpolated, normalized and subsequently cut into 56×56 , 112×112 and 224×224 pixel tiles. Next, data augmentation is applied by shifting the NTLs by 50 % in both x and y directions which effectively generates additional images and expands the training set respectively to 54780, 15840 and 5016 images. Experiment reuses 6 convolutional pre-trained models (NASNet, ResNet50, ResNet101, DenseNet101, DenseNet201, and EfficientNetb0) to obtain image embeddings, which are further employed in elastic net type regularized regression and random forest models. The tested solution, which relies solely on nighttime satellite data, shows potential for accurate and more nuanced forecasts of poverty. The best model for monthly poverty predictions, incorporating ResNet50 CNN architecture and 56×56 size tiles, yielded test $R^2 - 0.510$, MDAE – 1.960 and MAPE – 9.811.

Turinys

Lentelių sąrašas.....	8
Paveikslų sąrašas	9
Santrumpų ir terminų sąrašas.....	12
Įvadas.....	13
1. Literatūros apžvalga.....	14
1.1. Skurdo samprata	14
1.2. Piniginio skurdo rodikliai	16
1.3. Skurdo matavimas Lietuvoje	18
1.4. Satelitiniai nakties vaizdai kaip priemonė skurdo lygiui nustatyti	21
1.5. Literatūros analizės apibendrinimas	27
2. Tyrimo metodai.....	29
2.1. Konvoliuciniai neuroniniai tinklai.....	29
2.1.1. Konvoliuciniai sluoksniai	29
2.1.2. Sutelkimo sluoksniai	30
2.1.3. Visiškai sujungtas sluoksnis	31
2.1.4. Konvoliucinių neuroninių tinklų architektūros	31
2.1.4.1. ResNet-101 ir ResNet-50.....	31
2.1.4.3. NASNetMobile	32
2.1.4.4. EfficientNet-B0	32
2.1.4.5. DenseNet-201 ir DenseNet-121.....	33
2.2. Regresija	33
2.2.1. Ridge regresija	33
2.2.2. LASSO regresija.....	34
2.2.3. Elastic Net regresija.....	35
2.3. Atsitiktiniai miškai	35
2.4. Modelio tikslumo įvertinimo parametrai.....	36
2.4.1. Determinacijos koeficientas	36
2.4.2. Paklaidos.....	36
2.5. Laiko eilučių prognozavimo metodai	37
2.5.1. ETS	38
2.5.2. Theta	38
2.5.3. ARIMA.....	38
2.5.4. Naïve.....	39
2.5.5. Tiesinė regresija.....	39
2.5.6. Prophet.....	40
2.6. Satelitiniai nakties vaizdai	40
2.6.1. DMSP-OLS ir Suomi NPP-VIIRS vaizdų palyginimas	40
2.6.2. Koreguoti Suomi NPP-VIIRS vaizdai – NASA „Black Marble“	43
2.7. Duomenų paruošimas	45
2.7.1. Vaizdų gavimas	45
2.7.2. Vaizdų paruošimas eksperimentams	47
2.7.3. Požymių inžinerija.....	51
3. Tyrimo eiga ir rezultatai	52

3.1. Nakties šviesų žvalgomoji analizė.....	52
3.2. Skurdo rizikos lygio žvalgomoji analizė	53
3.3. Požymių vektorių gavimas	55
3.4. Duomenų paruošimas regresijai	57
3.5. Rezultatai	59
3.5.1. Elastic Net modelių rezultatai	60
3.5.2. Atsitiktinių miškų modelių rezultatai	69
3.6. Geriausių rezultatų palyginimas su laiko eilučių prognozėmis	72
3.7. Prognozės 2024-iesiems	73
3.8. Apibendrinimas ir įžvalgos.....	74
Išvados	76
Literatūros sąrašas	77
Priedai.....	82
1 priedas. Lietuvos skurdo rizikos lygis (eksperimentinė statistika).....	82
2 priedas. Eksperimentinių skurdo rodiklių statistika	85
3 priedas. <i>Elastic Net</i> modelių, mokytų be papildomų X, rezultatai	87
4 priedas. <i>Elastic Net</i> modelio, mokyto su <i>DenseNet-201</i> vektorizavimo būdu gautais požymiais, spėjimų rezultatai. Prognozavimas atliktas iš 224 x 224 dydžio vaizdų su 456 gardelių testavimo imtimi (10 prasčiausių metinių prognozių)	89
5 priedas. Skirtingais metodais sudarytų modelių palyginimas su faktinėmis reikšmėmis.....	89
6 priedas. <i>Elastic Net</i> modelių, mokytų su papildomais X, rezultatai.....	91
7 priedas. <i>Elastic Net</i> modelio, mokyto su <i>ResNet-50</i> vektorizavimo būdu gautais požymiais ir papildomais X, mėnesinės prognozės.....	92
8 priedas. Atsitiktinių miškų modelių, mokytų be papildomų X, rezultatai.....	95
9 priedas. Atsitiktinių miškų modelių, mokytų su papildomais X, rezultatai	96
10 priedas. Visų savivaldybių prognozės 2024 m. sausio–balandžio mėn., gautos taikant geriausią prognozavimo metodą	96

Lentelių sąrašas

1 lentelė. Kiekybinių ir kokybinių metodikų charakteristikos [4]	14
2 lentelė. Tarptautinių organizacijų (Pasaulio Banko, OECD) naudojami skurdo rodikliai [12, 13] 18	
3 lentelė. Eksperimentinės ir oficialiosios statistikos skirtumai [18].....	20
4 lentelė. Sintezės lentelė - panašūs tyrimai.....	26
5 lentelė. DMSP-OLS ir NPP-VIIRS palyginimas [55, 56, 57, 58].....	42
6 lentelė. „R“ programos paketai, reikalingi NASA „Black Marble“ duomenims parsisiųsti.....	46
7 lentelė. Mokymo ir testavimo imčių parametrai	59
8 lentelė. <i>Elastic Net</i> modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais, metinių spėjimų rezultatai (testavimo imtis – $2\,472\,56 \times 56$ dydžio vaizdai).....	60
9 lentelė. <i>Elastic Net</i> modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais, metinių spėjimų rezultatai (testavimo imtis – $2\,472\,56 \times 56$ dydžio vaizdai).....	61
10 lentelė. <i>Elastic Net</i> modelio, mokyto su <i>DenseNet-201</i> vektorizavimo būdu gautais požymiais, spėjimų rezultatai (10 prasčiausių prognozių).....	63
11 lentelė. <i>Elastic Net</i> modelio, mokyto su <i>DenseNet-201</i> vektorizavimo būdu gautais požymiais, spėjimų rezultatai (10 geriausių prognozių)	64
12 lentelė. <i>Elastic Net</i> modelio, mokyto su <i>DenseNet-201</i> vektorizavimo būdu gautais požymiais, spėjimų rezultatai (didžiųjų miestų prognozės).....	64
13 lentelė. <i>Elastic Net</i> modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais bei papildomais X, mėnesinių spėjimų rezultatai (mokymo ir testavimo imtys be persidengiančių vaizdų)	65
14 lentelė. <i>Elastic Net</i> modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais bei papildomais X, metinių spėjimų rezultatai (mokymo ir testavimo imtys be persidengiančių vaizdų)	65
15 lentelė. <i>Elastic Net</i> modelio, mokyto su <i>ResNet-50</i> vektorizavimo būdu gautais požymiais ir papildomais X, spėjimų rezultatai (10 prasčiausių prognozių)	68
16 lentelė. <i>Elastic Net</i> modelio, mokyto su <i>ResNet-50</i> vektorizavimo būdu gautais požymiais ir papildomais X, spėjimų rezultatai (10 geriausių prognozių).....	68
17 lentelė. <i>Elastic Net</i> modelio, mokyto su <i>ResNet-50</i> vektorizavimo būdu gautais požymiais ir papildomais X, spėjimų rezultatai (didžiųjų miestų rezultatai)	69
18 lentelė. Atsitiktinių miškų modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais, mėnesinių spėjimų rezultatai (testavimo imtis – $2\,472\,56 \times 56$ dydžio vaizdai).....	69
19 lentelė. Atsitiktinių miškų modelio, mokyto su <i>DenseNet-121</i> vektorizavimo būdu gautais požymiais, spėjimų rezultatai (didžiųjų miestų rezultatai)	71
20 lentelė. Atsitiktinių miškų modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais ir papildomais X, mėnesinių spėjimų rezultatai (testavimo imtis – $2\,472\,56 \times 56$ dydžio vaizdai) .	72
21 lentelė. Atsitiktinių miškų modelio, mokyto su <i>NASNetMobile</i> vektorizavimo būdu gautais požymiais ir papildomais X, spėjimų rezultatai (didžiųjų miestų rezultatai)	72
22 lentelė. <i>Elastic Net</i> modelių, mokytų su <i>ResNet-50</i> (su papildomais X) bei <i>DenseNet-201</i> (be papildomų X) vektorizavimo būdais gautais požymiais, palyginimas su kitais metodais gautais rezultatais.....	72

Paveikslų sąrašas

1 pav.	Pagrindiniai skurdo matavimo būdai.....	16
2 pav.	Skurdo rizikos lygis Europos Sąjungoje, 2022 m. [17].....	19
3 pav.	Skurdo rizikos lygio kaita Lietuvoje [17].....	20
4 pav.	B. Babenko ir kitų eksperimento rezultatai: „Digital Globe“ ir „Planet Imagery“ palyginimas su 10 % patvirtinimo imtimi [21]	22
5 pav.	D. Chen ir kitų eksperimente naudotos satelitinės nuotraukos pavyzdys – NTL vaizdas Mvanzoje, Tanzanijoje, perdengtas su miesto žemėlapiu (vidutinis pikselių skaičius: 28) [22]	22
6 pav.	A. Perez'o ir kitų eksperimento rezultatai: vidutiniai prognozių rezultatai už šalies ribų. Rezultatai gauti mokant modelius su 4 šalių duomenimis bei prognozuojant penktos šalies duomenis. Agreguotas likutinis determinacijos koeficientas nurodo kvadratinę koreliaciją tarp liekanų iš NTL prognozių ir liekanų iš modelių prognozių, agreguotų tarp visų penkių šalių [23]	23
7 pav.	Peng Wu ir kitų tyrimo eigos srauto diagrama [25]	24
8 pav.	Peng Wu ir kitų tyrimo rezultatai: Prognozuoto BVP, tenkančio vienam gyventojui bei tikrojo BVP, tenkančio vienam gyventojui regresijos modelio determinacijos koeficientas (R^2) [25]	24
9 pav.	S. R. Putri ir kitų tyrimo rezultatai: skurdo prognozės rezultatų sklaidos diagrama [26]	25
10 pav.	L. Kondmann'o ir X. Xiang Zhu tyrimo rezultatai: skurdo lygio prognozės rezultatai Ruandoje su 2005 m. duomenimis [27].....	26
11 pav.	Konvoliucinis sluoksnis, siekiant išskirti požymius, transformuoja įvesties vaizdą. Šioje transformacijoje vaizdas sujungiamas su branduoliu (arba filtru) [31]	30
12 pav.	Maksimalaus ir vidurkio sutelkimo procesai [32].....	30
13 pav.	<i>Softmax</i> veikimo iliustracija [33]	31
14 pav.	<i>ResNet-34</i> , <i>ResNet-50</i> ir <i>ResNet-101</i> architektūrų struktūros [35].....	32
15 pav.	Scheminė NASNet paieškos sistema [39]	32
16 pav.	<i>EfficientNet-B0</i> modelio veikimo principas [37].....	33
17 pav.	<i>DenseNet</i> modelio su trimis tankiai sujungtų sąryšių blokais veikimo principas [41]	33
18 pav.	Atsitiktinių miškų veikimo principas [42].....	36
19 pav.	Los Andželo NTL vaizdai. Kairėje: NPP-VIIRS vaizdas. Dešinėje: DMSP-OLS vaizdai [51]	41
20 pav.	Saulei sinchroninė (poliarinė) orbita – palydovų skriejimo trajektorija [59].....	43
21 pav.	NASA „Black Marble“ vaizdų apdorojimo procesas [56]	44
22 pav.	Žemės NTL iš NASA „Black Marble“ projekto [54].....	45
23 pav.	„R“ kodo, leidžiančio parsisiųsti NASA „Black Marble“ duomenis, iškarpa.....	46
24 pav.	2023 m. sausio mėnesio Kauno ir Vilniaus miestus apimantys vaizdai. Kairėje – „NearNadir_Composite_Snow_Free“ vaizdas, dešinėje – „OffNadir_Composite_Snow_Free“	47
25 pav.	Neapdoroti 2023 m. Lietuvos NTL vaizdai iš NASA „Black Marble“ archyvų.....	48
26 pav.	2023 m. rastriniai Lietuvos NTL vaizdai su interpoliuotomis reikšmėmis.....	48
27 pav.	Kairėje pusėje – 2023 m. sausio mėn. Lietuvos NTL vaizdas, išsaugotas GeoTIFF formatu. Dešinėje – vaizdas su savivaldybių ribomis	49
28 pav.	Viršuje – GeoTIFF formatu išsaugoti Lietuvos vaizdai (2023 m. sausio, gegužės bei spalio mėn). Per vidurį – to paties laikotarpio vaizdai, išsaugoti „ESRI Ascii“ formatu bei atidaryti per „QGIS“ programą. Apačioje – to paties laikotarpio vaizdai, apdoroti naudojant „QGIS“ programą.....	50

29 pav. 2013 m. lapkričio mėnesio Kauno nuotraukos: kairėje – neapdorotas NTL vaizdas, centre – koreguotas NTL vaizdas sujungtas su Lietuvos reljefo žemėlapiu, dešinėje – koreguotas NTL vaizdas, naudotas CNN mokymui	50
30 pav. 2016 m. sausio mėnesio Vilniaus nuotraukos: kairėje – neapdorotas NTL vaizdas, centre – koreguotas NTL vaizdas sujungtas su Lietuvos reljefo žemėlapiu, dešinėje – koreguotas NTL vaizdas, naudotas CNN mokymui	51
31 pav. Nakties šviesų intensyvumo dinamika pagal apskritis (2013–2023 m.)	52
32 pav. Skurdo rizikos lygio duomenų pagal savivaldybes 2013–2023 m. iškarpa (iš Oficialiosios statistikos portalo [18])	53
33 pav. Skurdo rizikos lygio pokyčiai Lietuvoje pagal savivaldybes 2013–2023 m.	54
34 pav. Skurdo rizikos lygio pokyčiai didžiųjų ir mažųjų Lietuvos miestų savivaldybėse 2013–2023 m. Punktyrinė linija žymi mažųjų Lietuvos miestų savivaldybes, ištisinė linija – didžiųjų.	54
35 pav. Skurdo rizikos lygio statistika didžiųjų ir mažųjų Lietuvos miestų savivaldybėse 2013–2023 m.	55
36 pav. 2023 m. sausio mėn. Lietuvos NTL vaizdas, sukarpytas į 56×56 dydžio gardeles	55
37 pav. 2023 m. sausio mėn. Lietuvos NTL vaizdas, sukarpytas į 112×112 dydžio gardeles	56
38 pav. 2023 m. sausio mėn. Lietuvos NTL vaizdas, sukarpytas į 224×224 dydžio gardeles	56
39 pav. Naudojamų CNN architektūrų parametrai	57
40 pav. CNN modelių parametrų nustatymas	57
41 pav. Savivaldybių pikselių skaičius sukarpytuose vaizduose (pateikiama lentelės iškarpa, sudaryta iš 224×224 vaizdų duomenų)	58
42 pav. Svertiniai skurdo rizikos lygio rodikliai kiekviename vaizde (pateikiama lentelės iškarpos pavyzdys, sudarytas iš 224×224 vaizdų duomenų)	59
43 pav. <i>Elastic Net</i> modelio, mokyto su <i>DenseNet-201</i> vektorizavimo būdu gautais požymiais, liekanų grafikas.....	61
44 pav. <i>Elastic Net</i> modelio, mokyto su <i>DenseNet-201</i> vektorizavimo būdu gautais požymiais, liekanų pasiskirstymo bei Q-Q grafikai.....	62
45 pav. <i>Elastic Net</i> modelio, mokyto su <i>DenseNet-201</i> vektorizavimo būdu gautais požymiais, Bland-Altman'o grafikas	63
46 pav. Iš kairės į dešinę – Švenčionių rajono, Šiaulių miesto, Neringos, Druskininkų bei Utenos rajono miesto savivaldybių (arba didžiausių jų dalių) gardelės. 56×56 dydžio vaizdų pavyzdys...	64
47 pav. Iš kairės į dešinę – Šiaulių miesto, Vilniaus miesto, Panevėžio miesto, Kauno miesto bei Klaipėdos miesto savivaldybių (arba didžiausių jų dalių) gardelės. 56×56 dydžio vaizdų pavyzdys	65
48 pav. <i>Elastic Net</i> modelio, mokyto su <i>ResNet-50</i> vektorizavimo būdu gautais požymiais bei papildomais X, liekanų grafikas	66
49 pav. <i>Elastic Net</i> modelio, mokyto su <i>ResNet-50</i> vektorizavimo būdu gautais požymiais bei papildomais X, liekanų pasiskirstymo bei Q-Q grafikai	67
50 pav. <i>Elastic Elastic Net</i> modelio, mokyto su <i>ResNet-50</i> vektorizavimo būdu gautais požymiais bei papildomais X, Bland-Altman'o grafikas.....	67
51 pav. Atsitiktinių miškų modelio, mokyto su <i>DenseNet-121</i> vektorizavimo būdu gautais požymiais, liekanų grafikas.....	70
52 pav. Atsitiktinių miškų modelio, mokyto su <i>DenseNet-121</i> vektorizavimo būdu gautais požymiais, liekanų pasiskirstymo bei Q-Q grafikai.....	70

53 pav. Atsitiktinių miškų modelio, mokyto su <i>DenseNet-121</i> vektorizavimo būdu gautais požymiais, Bland-Altman'o grafikas	71
54 pav. Didžiųjų miestų skurdo rodiklio prognozės 2024 m. sausio–balandžio mėn.....	73

Santrumpų ir terminų sąrašas

CNN (*Convolutional Neural Networks*) – konvoliuciniai neuroniniai tinklai.

DMSP-OLS (angl. *Defense Meteorological Satellite Program's Operational Linescan System*) – JAV Gynybinės meteorologinės palydovinės programos operatyvinė linijinio skenavimo sistema.

DVNL (angl. *DMSP-like Nighttime Lights Derived from VIIRS Nighttime lights*)

EOG (angl. *Earth Observation Group*) – Žemės stebėjimo grupė.

Gardelė – kvadratinis vaizdo gabalėlis.

GeoJSON (angl. *Geographic JavaScript Object Notation*) – failo formatas, skirtas geografinėms ypatybėms su jų neerdviniais atributais pavaizduoti.

GeoTIFF (angl. *Geographic Tag Image File Format*) – failo formatas, naudojamas erdviniams rastriniams duomenims saugoti. Tokio tipo failuose saugomos geografinės atskaitos koordinatės [1].

JPSS (angl. *Joint Polar Satellite System*) – poliarinės orbitos palydovinė sistema, NASA ir NOAA organizacijų jungtinė programa. Pirmasis paleistas palydovas – Suomi NPP.

LAADS DAAC (angl. *Level-1 and Atmosphere Archive and Distribution System Distributed Active Archive Center*) – NASA atmosferos mokslo duomenų archyvas.

NOAA (angl. *National Oceanic and Atmospheric Administration*) – JAV Nacionalinė vandenynų ir atmosferos administracija.

NTL (angl. *night-time light*) – nakties šviesos.

RGB – spalvų maišymo schema, kai naudojamos trys juostos – raudona (angl. *red*), žalia (angl. *green*) ir mėlyna (angl. *blue*).

Suomi NPP / SNPP (angl. *Suomi National Polar-orbiting Partnership*) – Suomijos nacionalinė poliarinė orbitos partnerystė. Taip pat taip vadinasi orų palydovas, valdomas NOAA bei NASA.

VIIRS (angl. *Visible Infrared Imaging Radiometer Suite*) – instrumentas, surenkantis regimuosius ir infraraudonuosius vaizdus bei globalius sausumos, atmosferos, kriosferos ir vandenynų stebėjimus. Šis instrumentas šiuo metu yra Suomi NPP ir NOAA-20 palydovuose.

Ivadas

Lietuva jau daugelį metų yra viena iš didžiausių skurdo rizikos rodiklius turinčių Europos Sąjungos šalių. Naudojamų skurdo duomenų problema yra mažas granuliarumas, t. y., visi oficialiai pateikiami skurdo duomenys yra metiniai. Metiniai duomenys turi daug trūkumų: neatskleidžia trumpalaikių skurdo pokyčių (neleidžia nustatyti skurdo tendencijų ir aptikti svarbių pokyčių reikiamu metu), taip pat tokie duomenys gali būti nepakankami norint įvertinti ilgalaikį skurdo lygio kitimą – skurdo priežasčių ir pasekmių ryšį per ilgesnį laikotarpį. Mėnesiniai duomenys, kita vertus, suteikia daugiau detalumo apie skurdo lygį ir jo pokyčius. Analizuojant mėnesinį skurdo lygį, galima nustatyti rodiklio sezoniškumą ar tendencijas, geriau suprasti skurdo priežastis bei dinamiką, taip pat turint mėnesinius duomenis galima efektyviau planuoti politikos priemonės ir programas skurdui mažinti.

Pastaraisiais metais vis daugiau diskutuojama apie alternatyvius metodus skurdo duomenims gauti – nuolat daugėja mokslinės literatūros apie mašininio mokymo ir satelitinių nakties vaizdų apjungimą agreguotiems skurdo duomenims surinkti bei prognozuoti. Vis didėjanti palydovinių vaizdų raiška bei santykinai lengvas ir viešas prieinamumas internete padaro jį potencialiu resursu šiam tikslui įgyvendinti. Giliojo mokymosi architektūra, sėkmingai taikoma įvairioms kompiuterinio matymo problemoms spręsti, pasitarnauja ir palydovinių vaizdų segmentacijos bei klasifikavimo užduotims, pvz., reljefo padalijimui į klases, tokias kaip žemė, vanduo, augalija ir pastatai. Neseniai, satelitinės nuotraukos pradėjo būti naudojamos ir statistiniam rodikliams, tokiems kaip skurdas ar pajamos, prognozuoti [2].

Šio darbo tikslas yra detalesnė skurdo lygio stebėseną ir prognostiką, pasitelkiant satelitinius nakties vaizdus.

Darbo tikslui įgyvendinti keliami šie uždaviniai:

1. apžvelgti atliktus panašaus pobūdžio mokslinius darbus;
2. pasirinkti konvoliucines architektūras vaizdams;
3. pasirinkti ir paruošti vaizdų rinkinius eksperimentams;
4. pritaikyti pasirinktas CNN architektūras vaizdams vektorizuoti;
5. sudaryti mašininio mokymosi modelius bei įvertinti jų tikslumą;
6. palyginti geriausio modelio rezultatus su laiko eilučių modelių prognozėmis;
7. prognozuoti skurdo rodiklius 2024 m. sausio–balandžio mėnesiams;
8. apibendrinti tyrimo rezultatus ir pateikti įžvalgas.

1. Literatūros apžvalga

Literatūros apžvalgos dalyje aptariamos skirtingos skurdo sampratos (objektyvi ir subjektyvi), metodikos (kiekybinė ir kokybinė) bei skurdo rodiklių rūšys (piniginiai ir nepiniginiai). Plačiau nagrinėjami piniginiai skurdo rodikliai, jų apibrėžimai ir skaičiavimai, taip pat jų trūkumai ir apribojimai. Aptariamos Lietuvoje naudojamos skurdo statistikos (eksperimentinė ir oficialioji) bei atskiri skurdo rodikliai. Galiausiai, apžvelgiami panašaus pobūdžio moksliniai darbai, tiriantys alternatyvius metodus skurdui nustatyti.

1.1. Skurdo samprata

Skurdas yra sudėtingas socialinis reiškiny, įvairiose pasaulio vietose apibrėžiamas ir skaičiuojamas skirtingai. Tyrėjai sutinka, kad skurdas gali būti skaidomas į dvi pagrindines kategorijas – objektyvų ir subjektyvų skurdą. Objektyvi (gerovės) perspektyva įtraukia normatyvinius požiūrius į tai, iš ko susideda skurdas bei ką galima padaryti dėl to, kad gyventojai iš jo išbristų. Subjektyvi perspektyva didesnę dėmesį skiria gyventojų preferencijoms – kiek jie patys vertina prekes ir paslaugas (t. y., žvelgia į individualią naudą) [3]. Tai dažnai pastebima besivystančiose pasaulio šalyse, kur kitokia metodika yra sunkiau prieinama – ten skurdui nustatyti dažniausiai pasitelkiamos namų ūkių apklausos, kurios gali būti tiek kiekybinės, tiek kokybinės.

Dar 1997-aisiais H. White ir S. Carvalho atliko kiekybinės ir kokybinės skurdo nustatymo metodikos lyginamąją analizę bei nustatė, jog šias metodikas galima integruoti, pavyzdžiui, kokybinę metodiką naudoti, siekiant patobulinti skurdo rodiklių struktūrą [4]. Pagrindinis skirtumas tarp kiekybiniais ir kokybiniais būdais gaunamų duomenų yra tas, jog kiekybiniai duomenys gali būti agreguojami, tuo tarpu kokybiniai – dažniausiai ne [3].

1 lentelė. Kiekybinių ir kokybinių metodikų charakteristikos [4]

Aspektas	Kiekybinė metodika	Kokybinė metodika
Skurdo apibrėžimas	Žmonės yra laikomi neturtingais, jei jų pragyvenimo lygis nukrenta žemiau skurdo ribos, t. y. pajamos (arba vartojimas), asocijuojamos su minimaliu priimtiniu maisto ar kitų būtinybių lygiu, nukrenta žemiau nustatytos ribos.	Neturtingi žmonės apibūdina, ką reiškia skurdas. Gali būti įtraukiami ir kiti kriterijai, ne tik pajamos ir/ar vartojimas.
Filosofinis pagrindas	Pozityvi paradigma: egzistuoja viena realybė (Chung'as, 1996).	Pozityvios paradigmos atmetimas: yra kelios realybės formos, todėl beprasmiška bandyti apibrėžti tik vieną (Chung'as 1996).
Skurdo nustatymas	Pasitelkiant išorinius tyrėjus.	Pasitelkiant interaktyvų vidinį-išorinį procesą, kuriame dalyvauja atsakingi asmenys bei dalyviai.
Kintamųjų, kurių duomenys renkami, pobūdis	Kiekybiškai įvertinami, pvz., namų ūkio išlaidos maistui, nedarbo lygis.	Požiūrio kintamieji, atspindintys pozicijas, preferencijas bei prioritetus (Moser'as, 1996); panašūs atsakymai, susiję su kiekvienu kintamuoju, gali būti sumuojami skaičiais, tačiau pačių kintamųjų negalima kiekybiškai įvertinti.
Interviu formatas	Struktūrizuotas, formalus, iš anksto parengtas klausimynas.	Atviro pobūdžio, pusiau struktūruotas, interaktyvus.

Atrankos imtis	Tikimybinė.	Tikslinė.
Atrankos paklaida	Mažesnė atrankos paklaida, tačiau galimos kitos klaidos.	Didesnė atrankos paklaida, tačiau sumažina kitų klaidų tikimybę.
Atrankos dydis	2000–8000 namų ūkių (pagal LSMS1): 2000–5000 namų ūkių).	1–1000 asmenų ar bendruomenių.
Geografinė aprėptis	Plati: įprastai nacionalinio lygmens.	Maža: įprastai apimami keli regionai, ar atrinktos bendruomenės.
Vidutinis laikas	LSMS: maždaug 2,5 metų aukščiausios kokybės apklausai atlikti, kuomet norima padengti visų metų šalies statistiką (1 metai skiriami planavimui; vieni pačiai apklausai atlikti, 6 mėnesiai pirminei analitikos fazei bei santraukai, duomenų dokumentavimui bei kitų analizių paruošimui). Planavimo procesas gali būti sutrumpintas, jei (i) pajėgumas yra labai didelis; ir/arba (ii) yra pasiryžimas eiti į kompromisus dėl kokybės. Taip pat apklausos atlikimo procesas gali būti sumažinamas nuo 1 metų iki, pavyzdžiui, 3 mėnesių, jei sutinkama atsisakyti apklausomis padengti pilnus metus. Dauguma apklausų eina į kompromisą dėl vieno, ar kito aspekto, todėl vidutinis laikas yra beveik visada trumpesnis. Prioritetinė apklausa: 7 mėnesiai.	Nuo 6 iki 9 mėnesių skurdo vertinimui; apie 4 mėnesiai naudos gavėjo vertinimui. Kai kuriose situacijose, kokybinis požiūris gali užimti daug laiko dėl apklausėjų apmokymo, taip pat ilgo proceso, klasifikuojant kokybines išvadas į analitines kategorijas.
Statistinė analizė	Statistinė analizė yra svarbi metodikos dalis.	Statistinė analizė teikia mažai naudos. Naudojama trianguliacija, kuomet vienu metu yra pasitelkiami keli skirtingi šaltiniai bei informacijos gavimo ir interpretavimo būdai. Tikimasi, kad informacija, surinkta iš skirtingų šaltinių, atskleis tam tikrą atsakymų dėsningumą. Taip pat gali būti naudojama sisteminė turinio analizė ir laipsniškas duomenų kaupimas pagal namų ūkio, grupės, kaimo, rajono ir nacionalinio lygmenų tematikas.

¹ (angl. *Living Standards Measurement Survey*) – Pasaulio Banko gyvenimo standartų matavimo apklausa.

Ekonomistai dažniausiai renkasi naudoti kiekybinį, kitaip tariant, statistinį požiūrį į skurdą, kadangi agreguoti gyventojų grupes pagal individualius požiūrius į gerovę yra sudėtinga, taip pat galima gauti prieštarigus vertinimus.

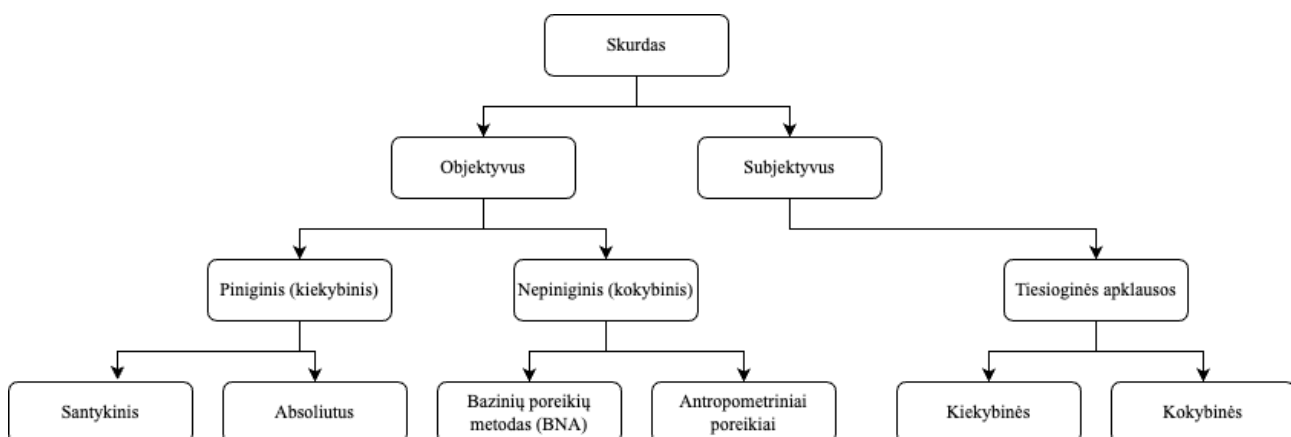
Visas skurdas, tiek objektyvus ir subjektyvus (taip pat kiekybinis ir kokybinis), yra analizuojamas per nepritekliaus konceptą. Jį galima išskirti į fiziologinio ir sociologinio grupes: fiziologinis nepritekliaus apibūdinamas per pajamų ir pagrindinių poreikių patenkinimo prizmę, tuo tarpu sociologinis nepritekliaus apima struktūrinę nelygybę, su kuria ypač susiduria gyventojai iš mažiau privilegijuotų socialinių grupių. Skurdo mažinimo strategijos kuriamos kombinuojant abu šiuos požiūrius – jomis

siekiama padidinti žemiau skurdo ribos gyvenančių žmonių pajamas, patenkinti pagrindinius jų poreikius bei suteikti prieigą prie sveikatos priežiūros ir švietimo [3].

Mokslinėje literatūroje dažniausiai išskiriamos dar dvi skurdo rūšys – ekonominis (kitai – piniginis) ir nepiniginis skurdas. Piniginio skurdo analizė grindžiama kiekybiniais skaičiavimo metodais, t. y., gyventojų pajamomis, užimtumu, ar išlaidomis [5], tuo tarpu nepiniginis skurdas atsižvelgia į socialinius, demografinius ir kitus svarbius aspektus, pavyzdžiui, gyventojų prieigą prie sveikatos priežiūros paslaugų, švietimo, materialinių išteklių, socialinės pagalbos, taip pat kultūros ir laisvalaikio veiklų [6].

1.2. Piniginio skurdo rodikliai

Pasak R. Lok-Dessallien (1999) [3], dėl multidimensinio skurdo pobūdžio, taip pat problemos kompleksiskumo, piniginis skurdas dar gali būti dalomas į absoliutų bei santykinį skurdą. Absoliutaus skurdo terminas apibūdina gyvenimą žemiau minimumo – socialiai priimtinių gyvenimo sąlygų, kurios nustatomos pagal prieigą prie maisto bei kitų būtiniausių prekių. Santykinis skurdas, kita vertus, paprastai matuojamas pajamų kvantiliais ir deciliais, lygina žemiausius gyventojų segmentus su aukštesne klase. Absoliutaus ir santykinio skurdo tendencijos gali judėti priešingomis kryptimis. Pavyzdžiui, jei sumažėja atotrūkis tarp aukštesnių ir žemesnių gyventojų sluoksnių dėl pastarojo sluoksnio gerovės sumažėjimo, bet tuo pat metu daugiau namų ūkių krenta žemiau absoliučios skurdo ribos, tuomet santykinio skurdo rodiklis gali kristi, o absoliutaus – augti.



1 pav. Pagrindiniai skurdo matavimo būdai

Ekonominio bendradarbiavimo ir plėtros organizacija (angl. *Organisation for Economic Co-operation and Development*, OECD) skurdo lygiui matuoti naudoja santykinį skurdo lygį, kuris skaičiuojamas, remiantis pajamų nelygybe – bendrosiomis vidutinėmis namų ūkių disponuojamomis pajamomis. Skurdo lygis nustatomas pagal tai, kiek žmonių patenka žemiau skurdo ribos. Žmonės laikomi, kaip gyvenantys žemiau skurdo lygio, kai jų bendros vidutinės namų ūkių disponuojamos pajamos yra žemiau 50 % šalies medianos. Visgi, santykinio pajamų slenksčio naudojimas reiškia, jog turtingesnės šalys turi aukštesnes skurdo ribas – čia skurdo išvengimas reiškia galimybę gauti prekes ir paslaugas, kurios laikomos įprastomis bet kurioje šalies vietoje [7].

Skurdo lygis pagal amžiaus grupes skaičiuojamas pagal visų gyventojų pajamų medianą. Santykinio skurdo pokyčius, susijusius su dabartinėmis vidutinėmis pajamomis, recesijos metu gali būti sunku interpretuoti. Esant tokiai situacijai, kai visų namų ūkių pajamos krenta, tačiau visų gyventojų pajamų

apačioje jos krenta mažiau nei viduryje, santykinis skurdas mažes. Todėl norint papildyti santykinio pajamų skurdo vaizdą, reikia absoliutaus skurdo rodiklius susieti su ankstesniu gyvenimo lygiu [7]. Tokios skurdo ribos yra vadinamos įtvirtintomis (angl. *anchored*) bei yra santykinės ir absoliučios tuo pat metu, kadangi santykinė riba yra pirmiausia apskaičiuojama vienam laikotarpiui, tuomet užfiksuojama, ir toliau naudojama kaip absoliuti. Įtvirtinta skurdo riba fiksuoja skurdo pokyčius, netiesiogiai veikdama gyvenimo lygio konstantos pokyčius. Tačiau kadangi santykinio skurdo sąvoka siekiama užfiksuoti ir socialinę įtrauktį, problema, kaip interpretuoti nesikeičiančios santykinio skurdo ribos reikšmę, išlieka aktuali [8].

Pasaulio Bankas didesnę dėmesį skurdui ėmė skirti tik 1970-aisiais, kuomet apie skurdo mažinimą pradėta kalbėti kaip apie pagrindinį šalių vystymosi katalizatorių [9]. Nors prieš tai skurdo terminas ir buvo minimas įvairiuose Pasaulio Banko dokumentuose, tai niekada nebuvo oficialus institucinis terminas. 1990-aisiais Pasaulio Bankas monetarizavo skurdo terminą (patį svarbiausią vystymosi indeksą), nustatydamas tarptautinę skurdo ribą ties 1 JAV doleriu per dieną. Dabartiniais laikais Pasaulio Bankas tapo vienu svarbiausių globalių „kovotojų“ su skurdu, siekiančiu užbaigti ypač didelį skurdą bei padidinti bendrąją gerovę [10]. Šio tikslo progresui stebėti, Pasaulio Bankas yra įvedęs keturių kategorijų indikatorius: skurdo lygio pagal nacionalines skurdo ribas, skurdo lygio pagal tarptautines skurdo ribas bei bendrą gerovę [11].

Skurdo lygio pagal nacionalines skurdo ribas rodiklis remiasi konkrečios šalies duomenimis. Valstybė gali turėti vieną unikalią nacionalinę skurdo ribą, atskiras ribas kaimo ir miesto regionams arba atskiras ribas skirtingoms geografinėms teritorijoms. Tokiu būdu atsižvelgiama į pragyvenimo išlaidų skirtumus skirtingose šalies vietovėse, taip pat skirtumus mitybos ir vartojimo įpročiuose. Dėl šios priežasties nacionalinės skurdo ribos yra tokie rodikliai, kurie atitinka kiekvienos šalies konkrečias ekonomines ir socialines sąlygas. Beveik visos nacionalinės skurdo ribos yra pagrįstos maisto krepšelio kaina, atsižvelgiant į tai, ko reikia adekvačiai mitybai, gerai sveikatai ir aktyvumui palaikyti. Nacionalinės skurdo ribos yra koreguojamos prie infliacijos – tai užtikrina duomenų tikslumą bei leidžia palyginti skurdo ribos pokyčius laike. Kadangi mityba ir vartojimo krepšeliai taip pat keičiasi per laiką, šalys, besiremiamos naujausiais apklausų duomenimis, periodiškai perskaičiuoja savo skurdo ribas [11].

Pasaulio Banko duomenimis, skurdo lygio pagal tarptautines skurdo ribas rodikliai nėra tokie tikslūs. Susiduriama tiek su konceptualiomis, tiek su praktinėmis problemomis, kadangi šalys apibrėžia ir matuoja skurdą skirtingai, atlikti nuoseklius ir tikslingus palyginimus yra sudėtinga. Nacionalinės skurdo ribos turi didesnę perkamąją galią turtingose šalyse nei skurdžiose. Dėl šios priežasties skirtingose šalyse gyvenančių žmonių gerovę galima palyginti tik pakoreguojant perkamosios galios skirtumus pagal atitinkamas valiutas [11].

Nelygybė ir bendra gerovė yra platesnis konceptas nei skurdas, kadangi apima ne tik gyventojus žemiau skurdo ribos, bet visą populiaciją. Nelygybė dažniausiai matuojama Gini koeficientu, gaunamu iš Lorenzo kreivės, kuri klasifikuoja populiaciją pagal skurdo lygį (nuo skurdžiausios iki turtingiausios), ir rodo kumuliatyvią populiacijos dalį horizontaliojoje ašyje bei kumuliatyvią išlaidų arba pajamų dalį vertikaliojoje ašyje. Bendra gerovė, kita vertus, matuojama standartiniu metodu – vartojimu ir/arba pajamomis. Bendros gerovės duomenų augimo rodikliai skaičiuojami kaip vidutinis metinis vartojimo arba pajamų augimo tempas per penkerių metų laikotarpį. Nors Pasaulio Bankas pripažįsta, kad gerovės matas turi daug skirtingų aspektų, sprendimas matuoti bendrą gerovę pagal

vartojimą ar pajamas buvo motyvuotas poreikiu turėti tokį rodiklį, kuris būtų lengvai suprantamas, paaiškinamas ir apskaičiuojamas [11].

Šiuo metu Pasaulio Bankas ir OECD pateikia duomenis apie 5 skirtingus skurdo rodiklius: gyventojus, gyvenančius žemiau skurdo ribos, skurdo rizikos lygį, skurdo rizikos gylį, skurdo sunkumą bei Watts'o indeksą.

2 lentelė. Tarptautinių organizacijų (Pasaulio Banko, OECD) naudojami skurdo rodikliai [12, 13]

Skurdo rodiklis	Organizacija	Apibrėžimas
Gyventojai, gyvenantys žemiau skurdo ribos	Pasaulio Bankas	Gyventojų, turinčių mažesnę nei 2017 m. perkamosios galios paritetą (PGP) per dieną skaičius (milijonais), gaunamas padauginant skurdo lygį iš gyventojų skaičiaus. Kadangi PGP keičiasi, šis rodiklis negali būti lyginamas su ankstesnių metų rezultatais.
Skurdo rizikos lygis	Pasaulio Bankas, OECD	Gyventojų, gyvenančių namų ūkiuose, kuriuose vartojimas ar pajamos, tenkančios žmogui, yra žemiau skurdo ribos (pagal 2017 m. tarptautines kainas), skaičius procentais. Kadangi PGP keičiasi, šis rodiklis negali būti lyginamas su ankstesnių metų rezultatais.
Skurdo rizikos gylis	Pasaulio Bankas, OECD	Skurdo rizikos gylis per dieną (2017 m. PGP) yra vidutinis pajamų arba vartojimo atitolimas nuo skurdo ribos per dieną (išreikštas procentais nuo skurdo ribos). Šis matavimo būdas atspindi skurdo rizikos gylį bei paplitimą. Kadangi PGP keičiasi, šis rodiklis negali būti lyginamas su ankstesnių metų rezultatais.
Skurdo sunkumas	Pasaulio Bankas	Skurdo gylio kvadratas. Kadangi PGP keičiasi, šis rodiklis negali būti lyginamas su ankstesnių metų rezultatais.
Watts'o indeksas	Pasaulio Bankas	Skurdo rodiklis, atsižvelgiantis į pajamų pasiskirstymą. Indeksas apskaičiuojamas skurdo rizikos lygį dalinant iš skurstančiųjų pajamų augimo tempo.

Nors skurdo rodiklių įvairovė yra didelė, duomenys jiems apskaičiuoti yra dažnai sunkiai prieinami. Tiksliai išmatuoti pajamas, ypač besivystančiose šalyse, yra sudėtinga. Kadangi vartojimo duomenys nėra visada pasiekiami, skurdo rodikliai, apimantys pajamas ir pragyvenimo lygį, gali būti netikslūs. Skurdo nustatymas apklausų metodu taip pat nėra efektyvus – apklausų rezultatai gali būti nepalyginami dėl atlikimo laiko skirtumų, imties dydžio, ar interviu kokybės. Problemos gali kilti ir lyginant skirtingo išsivystymo lygio šalis, nes skiriasi santykinė ne rinkos prekių vartojimo svarba [7].

Kai kurie atrinkti namų ūkiai gali atsisakyti dalyvauti apklausose, arba per interviu vizitus nebūti namuose. Taip pat skirtingas pajamas gaunantys namų ūkiai gali būti skirtingai įsitraukę į dalyvavimą apklausose: turtingesni namų ūkiai gali nenorėti švaistyti laiko, arba būti susirūpinę savo privatumu, tuo tarpu skurdesni namų ūkiai gali būti nepakankamai atstovaujami, jei neturi pastovios gyvenamosios vietos arba yra klajokliai (t. y., fiziškai, ar socialiai izoliuoti). Jei šie duomenys nėra koreguojami, tai gali pakenkti tiek skurdo, tiek nelygybės matavimams [7].

1.3. Skurdo matavimas Lietuvoje

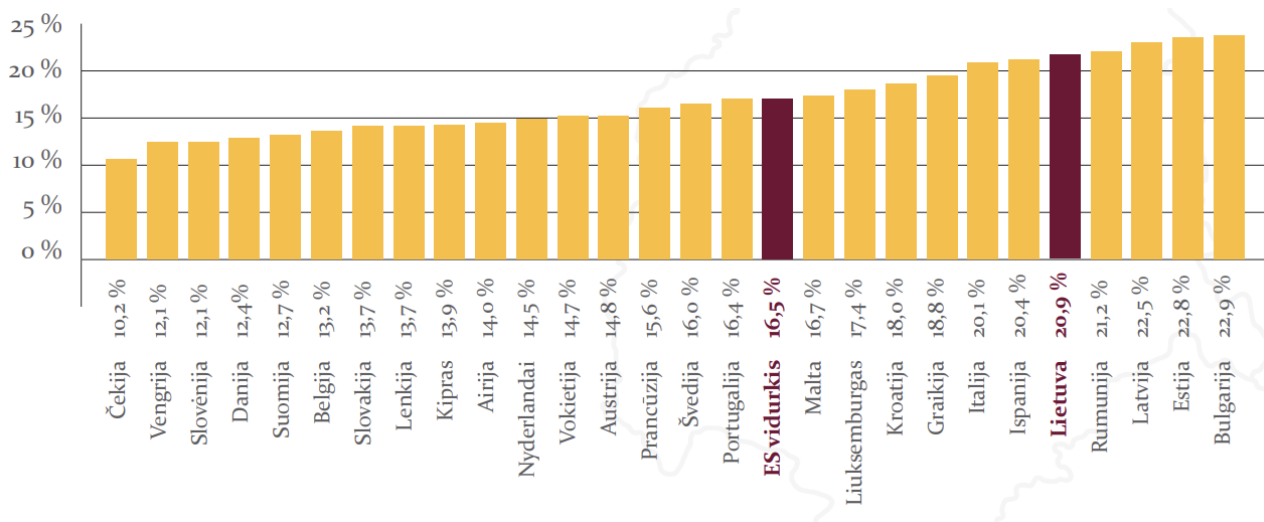
Lietuvoje pagrindiniai skurdo rodikliai yra (1) skurdo rizikos riba, (2) skurdo rizikos lygis, (3) absoliutaus skurdo lygis, (4) absoliutaus skurdo riba, (5) GINI koeficientas bei (6) pajamų pasiskirstymo koeficientas. Lietuvos Respublikos socialinės apsaugos ir darbo ministerija skurdo

rizikos ribą apibūdina kaip pajamų dydį, už kurį mažesnes pinigines ekvivalentines disponuojamas pajamas gaunantys asmenys priskiriami prie skurstančių. Pagal oficialią metodiką, skurdo rizikos riba laikoma 60 % gyventojų ekvivalentinių disponuojamų pajamų medianos. Skurdo rizikos lygis, tuo tarpu, yra asmenų dalis, kurių ekvivalentinės pinigines disponuojamosios pajamos yra mažesnės už skurdo rizikos ribą [14].

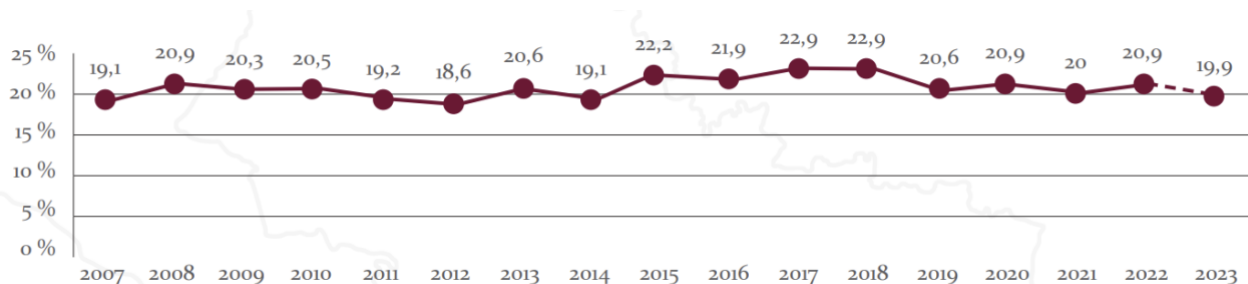
Absoliutaus skurdo lygis LR Socialinės apsaugos ir darbo ministerijos apibūdinamas kaip asmenų, kurių ekvivalentinės disponuojamosios pajamos mažesnės už absoliutaus skurdo ribą, dalis. Absoliutaus skurdo riba yra minimalių vartojimo poreikių dydis, reikalingas minimaliems asmens maisto ir ne maisto (prekių bei paslaugų) poreikiams patenkinti. Ši riba kasmet perskaičiuojama pagal maisto ir ne maisto prekių ir paslaugų kainų pokyčius [14].

GINI (pajamų nelygybės) koeficientas yra rodiklis, kurio reikšmė svyruoja nuo 0 (reiškiančio absoliučią lygybę, kai visi šalies namų ūkiai gauna vienodas pajamas) iki 100 (reiškiančio absoliučią nelygybę, kai vienas namų ūkis uždirba visas šalies pajamas). Pajamų pasiskirstymo koeficientas (S80/S20) parodo pajamų santykį tarp 20 % didžiausias pajamas gaunančių ir 20 % mažiausias pajamas gaunančių gyventojų [14].

„Eurostat“ duomenimis, asmenų, kuriems gresia skurdas arba socialinė atskirtis rodiklis Lietuvoje 2021 m. siekė 20 %, o 2022 m. pakilo iki 20,9 % [15] – nors pastaruoju metu šalies ekonominiai rodikliai gerėja, vis dar smarkiai viršijame Europos Sąjungos skurdo rizikos vidurkį [16] (2 pav.). Valstybės duomenų agentūra prognozuoja, kad 2023 m. skurdo rizikos lygis (vertinant 2022 m. šalies gyventojų pajamas) turėtų siekti apie 19,9 % (3 pav.). Tai yra 1 procentiniu punktu mažiau nei ankstesniais metais. Nors tikslų duomenų dar nėra, tokios tendencijos rodo, kad infliacijos pasekmių švelninimo priemonių taikymas, atlyginimų ir socialinių išmokų didinimas padėjo suvaldyti sudėtingą 2022 m. situaciją ir turėjo teigiamos įtakos skurdo lygio mažėjimui šalyje [17].



2 pav. Skurdo rizikos lygis Europos Sąjungoje, 2022 m. [17]



3 pav. Skurdo rizikos lygio kaita Lietuvoje [17]

Tapus Europos Sąjungos nare, Lietuva skurdo lygį pradėjo skaičiuoti santykinio skurdo apskaičiavimo būdu, kuris grindžiamas asmens gyvenimo lygio, apibrėžto santykiu su vidutiniu gyvenimo lygiu šalyje, matavimu [16]. Šiuo matavimo būdu apskaičiuojami šie rodikliai:

1. skurdo rizikos lygis;
2. skurdo ir socialinės atskirties rizikos lygis.

Lietuvoje oficialiosios statistikos portalas taip pat teikia eksperimentinę statistiką apie trijų rūšių skurdo rodiklius: skurdo rizikos ribą, absoliutaus skurdo lygį bei skurdo rizikos lygį [18]. Šie duomenys gali būti agreguojami pagal namų ūkio tipą, amžių, laikotarpį bei administracinę teritoriją (apskritį arba savivaldybę). Eksperimentinė statistika apima alternatyvius statistikos skaičiavimo būdus, kurie dar nėra galutinai išbandyti, ištestuoti, suderinti metodologiškai, patvirtinti kaip oficialioji statistika ir įtraukti į Oficialiosios statistikos programą. Rodikliai gali būti rengiami administracinių duomenų pagrindu ir pagrįsti matematiniais skaičiavimais, prognozėmis paremta statistika [18]. Visgi, eksperimentinė statistika dažnai yra tikslesnė nei oficialioji, kadangi naudoja naujas skaičiavimo metodikas ir alternatyvius būdus duomenims gauti.

3 lentelė. Eksperimentinės ir oficialiosios statistikos skirtumai [18]

Eksperimentinė statistika	Oficialioji statistika
Patvirtinto kalendoriaus kol kas nėra, rengiama pagal poreikį, galimybes.	Paskelbiama pagal patvirtintą Oficialiosios statistikos kalendorių.
Administraciniai šaltiniai, matematiniai prognozavimo metodai, naudojami alternatyvūs informaciniai šaltiniai.	Naudojami administraciniai šaltiniai, statistiniai tyrimai.
Neįtraukta į Oficialiosios statistikos programą.	Įtraukta į Oficialiosios statistikos programą.
Ne visi rodikliai palyginami su kitų šalių duomenimis.	Suderinta ir dažnai palyginama su kitomis šalimis metodika
Metodikos dar kuriamos, testuojamos.	Patvirtintos metodikos.

Eksperimentinė statistika taip pat gali būti operatyvesnė nei įprasta (oficialioji), nes joje naudojami ir prognostiniai metodai – dar neturint administracinių duomenų, jie pagal ankstesnių metų rodiklius prognozuojami. Pavyzdžiui, eksperimentiniai skurdo rodikliai paskelbiami beveik metais anksčiau nei įprasta skurdo statistika. Šiai statistikai parengti naudojami prognozavimo modeliai, sudaryti iš kelerių paskutinių metų pajamų ir gyvenimo sąlygų tyrimo duomenų, ir turima atitinkama administracinių šaltinių informacija [18].

Nors eksperimentiniai skurdo rodikliai yra pakankamai patikimi, jie neišsprendžia svarbios su skurdo matavimu susijusios problemos – duomenų periodiškumo. Lietuvoje tiek oficialūs, tiek

eksperimentiniai skurdo yra teikiami metiniu dažnumu. Mėnesinių duomenų turėjimas leistų įsigilinti į skurdo kaitą, geriau informuotų politikos formuotojus ir tyrėjus apie socialines ir ekonomines sąlygas realiu laiku, taip pat būtų naudingas priedas prie metinių skurdo įverčių [19]. Dėl šių priežasčių ieškoma naujų metodų skurdo rodikliams gauti.

1.4. Satelitiniai nakties vaizdai kaip priemonė skurdo lygiui nustatyti

Vis dažniau skurdui nustatyti pasitelkiami alternatyvūs būdai – satelitiniai nakties vaizdai ir mašininis mokymasis. Mokslininkai naudoja šį metodą nustatyti skurdui ten, kur nėra daug patikimų skurdo duomenų. Nors skurdas gali būti apskaičiuojamas iš ekonominių duomenų, tokių kaip namų ūkio gerovė ir turtas, tačiau jie surenkami per namų ūkio apklausas, kurios yra brangios ir užima daug laiko.

Nauji skurdo nustatymo būdai, pasitelkiant satelitinius nakties vaizdus, pradėti naudoti santykinai neseniai. N. Jean'as ir kiti [20] teigia, jog besivystančiame pasaulyje yra sunku kurti skurdą mažinančias politikos priemones, kadangi vis dar trūksta patikimų duomenų apie ekonominį pragyvenimo lygį. Siekdami išspręsti šią problemą, 2016 metais autoriai pirmieji pademonstravo tikslų ir nebrangų metodą vartojimo išlaidoms ir gerovei nustatyti iš didelės raiškos satelitinių vaizdų. Naudodami apklausų ir palydovų duomenis iš penkių Afrikos šalių – Nigerijos, Tanzanijos, Ugandos, Malavio ir Ruandos – jie parodė, kaip CNN gali būti išmokytas atpažinti vaizdo požymius, kurie gali paaiškinti iki 75 % namų ūkių gerovės variacijos vietiniu lygmeniu. N. Jean'as šiam tyrimui atlikti naudojo *ResNet-34* architektūros tipą su dienos ir nakties satelitiniais duomenimis. Rezultatai atskleidė, jog sukurtas modelis buvo pajėgus semantiškai identifikuoti reikšmingus požymius, pavyzdžiui, gyvenamuosius regionus, greitkelius, vandens telkinius, žemės ūkio plotus, ir kita [20].

Kiek vėliau B. Babenko ir kiti [21] šį mokslinį darbą papildė įtraukdami žemės naudojimo duomenis, gautus iš „Planet“ ir „Digital Globe“ platformų teikiamų satelitinių vaizdų. Autoriai CNN pasitelkė skurdui vertinti tiesiogiai iš didelės ir vidutinės raiškos palydovinių vaizdų. Jie naudojo 3–5 m² ir 50 cm² erdvinės raiškos vaizdus, kurie padengia 2 mln. km² Meksikos ploto. Tyrimui atlikti buvo išbandytos kelios architektūros – VGG bei *GoogleNet* – ir galiausiai pasirinkta pastaroji. Buvo nustatyta, jog (1) geriausi modeliai, kurie įtraukia palydovais nustatytą žemės naudojimą kaip prognozavimo priemonę, paaiškina maždaug 57 % skurdo variacijos patvirtinimo imtyje (apimančioje 10 % savivaldybių); (2) naudojant pilną duomenų rinkinį (t.y. visas savivaldybes) aiškinamoji galia sumažėja iki 44 %; (3) prognozuojamas skurdas, pasitelkiant vien CNN, paaiškina apie 47 % skurdo variacijos patvirtinimo imtyje bei 37 % visose savivaldybėse. (4) miesto zonose pastebimas nedidelis rezultatų pagerėjimas vietoje „Planet“ vaizdų, naudojant „Digital Globe“ – su „Digital Globe“ skurdo variacija paaiškinama 61 %, tuo tarpu „Planet“ – 54 % (4 pav.). Autoriai padaro išvadą, jog CNN gali būti pilnai pritaikomi skurdui prognozuoti iš satelitinių vaizdų [21].

Sample	R^2 CNN Predictions using Digital Globe imagery	R^2 CNN Predictions using Planet imagery	# municipalities
Urban areas	0.61	0.54	58

Table 2: CNN Predictions In and Out of Sample

Validation	Sample	R^2 CNN Predictions	R^2 Landcover	R^2 Both	Areas
10% MCS-ENIGH Validation	All	0.47	0.49	0.57	109
	Urban	0.54	0.52	0.64	58
	Rural	0.47	0.49	0.64	51
all MCS-ENIGH areas	All	0.37	0.37	0.44	1115
	Urban	0.34	0.31	0.4	619
	Rural	0.38	0.34	0.44	496
non MCS-ENIGH areas	All	0.15	0.23	0.28	2834
	Urban	0.06	0.19	0.21	944
	Rural	0.22	0.25	0.31	1890

4 pav. B. Babenko ir kitų eksperimento rezultatai: „Digital Globe“ ir „Planet Imagery“ palyginimas su 10 % patvirtinimo imtimi [21]

D. Chen’as ir kiti [22] taip pat atliko panašaus pobūdžio tyrimą, tačiau naudojo kiek kitus duomenų šaltinius. Skurdo duomenys buvo gauti iš LSMS apklausų, tuo tarpu NTL duomenys – iš NGDC EOG. Konkrečiai, NTL duomenys buvo paimti iš Gynybos meteorologinių palydovų programos (DMSP), kuri teikia vaizdus 30 kampo minučių raiška (apie 1,0 km), ir buvo naudojami kaip miesto šviesumo matas naktį (5 pav.) [22].



5 pav. D. Chen ir kitų eksperimente naudotos satelitinės nuotraukos pavyzdys – NTL vaizdas Mvanzoje, Tanzanijoje, perdengtas su miesto žemėlapiu (vidutinis pikselių skaičius: 28) [22]

Tuomet CNN mokymas buvo išskaidytas į tris dalis – pirmajame etape CNN tinklas nustatė savo svorį RGB kanaluose, naudodamas modelį, kuris iš pradžių buvo mokytas su *ImageNet* – duomenų baze, leidžiančia pirmiesiems sluoksniams beveik nedelsiant įgyti gebėjimą aptikti kraštus (angl. *edge detection*) ir atlikti kitas panašias grubias užduotis. Antrajame etape tinklas buvo mokomas prognozuoti NTL intensyvumą iš dienos vaizdų, tuo pačiu mokantis funkcijų, kurios yra naudingos numatant skurdą. Paskutiniame etape, siekiant prognozuoti galutinį gerovės balą, tarpinė tinklo išvestis buvo perkeliama į atskirą regresijos modelį [22].

Gavus vaizdų požymius, kiekvienas vektorinis vaizdo atvaizdas buvo pritaikytas kryžminės regresijos modeliui, siekiant prognozuoti galutinius gerovės rodiklius kiekvienam laiko periodui. Nors ir pavyko gauti panašius kryžminės prieigos rezultatus, autoriai teigia, jog vis dar trūksta

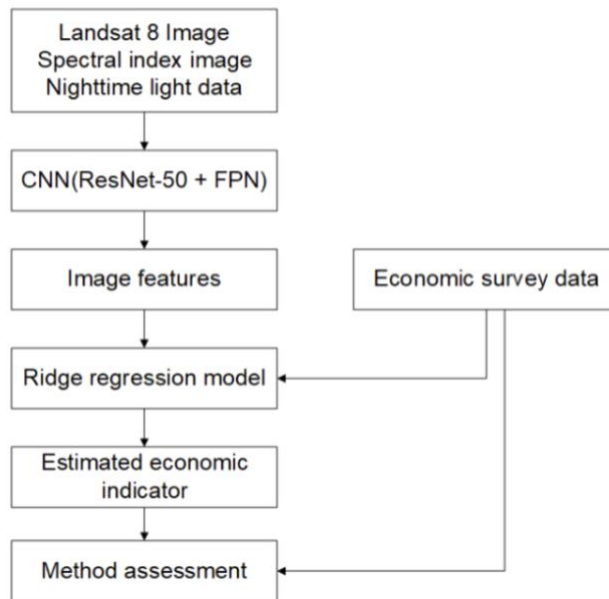
galimybių prognozuoti skurdo lygių skirtumus per laiką. Jų manymu, alternatyvūs skurdo rodiklių nustatymo būdai, pavyzdžiui, žemės ūkio produktyvumo rodikliai per derlių arba išėigą, tenkančią vienam darbuotojui, galėtų padėti atrasti naujus ir efektyvesnius metodus gerovės pokyčiams skaičiuoti [22].

A. Perez'as ir kiti [23] taip pat naudojo satelitinius nakties vaizdus skurdui nustatyti. Autoriai mokė CNN modelius su nemokamomis ir viešai prieinamomis daugiaspektrėmis satelitinėmis Afrikos žemyno dienos nuotraukomis, gautomis iš palydovo „Landsat 7“, kuris Žemės vaizdus pradėjo rinkti prieš daugiau nei du dešimtmečius. Tyrimas parodė, jog nepaisant prastos nuotraukų raiškos, sudaryti modeliai sugebėjo pasiekti ankstesnius standartus viršijantį tikslumą. Modeliams sudaryti buvo naudojamos kelios architektūros – 18 ir 34 sluoksnių *ResNet* ir VGG-F. Skurdui prognozuoti taip pat buvo testuojami keli modeliai – *Ridge* regresija bei GBT (angl. *gradient-boosted trees*). Rezultatai parodė, jog satelitiniais vaizdais pagrįstas prognozavimas leidžia nustatyti gerovę šalyse, kur kažkiek skurdo duomenų jau egzistuoja, tačiau gali susidurti su sunkumais ekstrapoliuojant už šalies ribų (6 pav.) [23].

Model	Mean Train r^2	Mean Test r^2	Aggregate Residual r^2
Nightlights / GBT	0.63	0.66	1.0
VGG-F, RGB / ridge	0.71	0.64	0.69
VGG-F, 9 Band / ridge	0.68	0.63	0.70
ResNet-18, 9 Band / ridge	0.69	0.64	0.73
ResNet-34, 9 Band / ridge	0.70	0.65	0.74
Jean et al. [8]	0.53	0.54	0.76

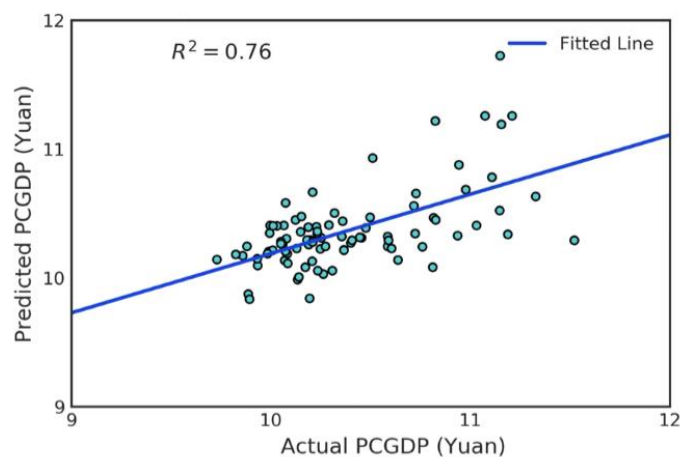
6 pav. A. Perez'o ir kitų eksperimento rezultatai: vidutiniai prognozių rezultatai už šalies ribų. Rezultatai gauti mokant modelius su 4 šalių duomenimis bei prognozuojant penktos šalies duomenis. Agreguotas likutinis determinacijos koeficientas nurodo kvadratinę koreliaciją tarp liekanų iš NTL prognozių ir liekanų iš modelių prognozių, agreguotų tarp visų penkių šalių [23]

Peng Wu ir Yumin Tan [25] taip pat aprašė CNN taikymo galimybes regioniniam skurdui prognozuoti. Kitaip nei kituose darbuose, autoriai klasifikavimo užduotims atlikti naudojo ne tik iš anksto mokytus giliojo mokymosi modelius (šiuo atveju *ResNet-50*), bet ir požymių piramidės tinklą FPN (angl. *Feature Pyramid Network*). Šis tinklas, pirmąkart pasiūlytas Lin ir bendraautorų 2017 m. [24], veikia piramidės principu – leidžia identifikuoti skirtingo masto objektus, todėl pagerina objektų aptikimo efektyvumą. Kaip CNN architektūros įvestis buvo naudojamos „Landsat 8“ dirbtinio palydovo nuotraukos, tuo tarpu išvestis – NTL intensyvumas, normalizuotas augmenijos skirtumo indeksas (angl. *normalized difference vegetation index*, NDVI), modifikuotas normalizuotas vandens skirtumo indeksas (angl. *modified normalized difference water index*, MNDWI) ir normalizuotas užstatymo skirtumo indeksas (angl. *normalized difference built-up index*, NDBI) rodikliai. Išmokius modelį, iš paskutinio CNN globalaus vidurkio telkimo (angl. *global average pooling*) sluoksnio buvo ištraukiami 256-dimensiniai požymių vektoriai, iš keturių klasifikavimo užduočių sujungiami į 1024-dimensinius požymių vektorių failus. Galiausiai, turimi ekonominiai rodikliai ir vaizdų požymiai buvo panaudoti *Ridge* regresijos modeliui mokyti. Viso eksperimento eiga pateikiama žemiau [25].



7 pav. Peng Wu ir kitų tyrimo eigos srauto diagrama [25]

Atlikus eksperimentą buvo nustatyta, jog CNN ištraukti vaizdo požymiai gali paaiškinti regiono BVP (bendrojo vidaus produkto), tenkančio vienam gyventojui, variaciją iki 76 % (8 pav.). Autorių teigimu, lyginant su kitais metodais, šio metodo tikslumas gana aukštas ir gali būti naudojamas regioniniam skurdui nustatyti [25].



8 pav. Peng Wu ir kitų tyrimo rezultatai: Prognozuoto BVP, tenkančio vienam gyventojui bei tikrojo BVP, tenkančio vienam gyventojui regresijos modelio determinacijos koeficientas (R^2) [25]

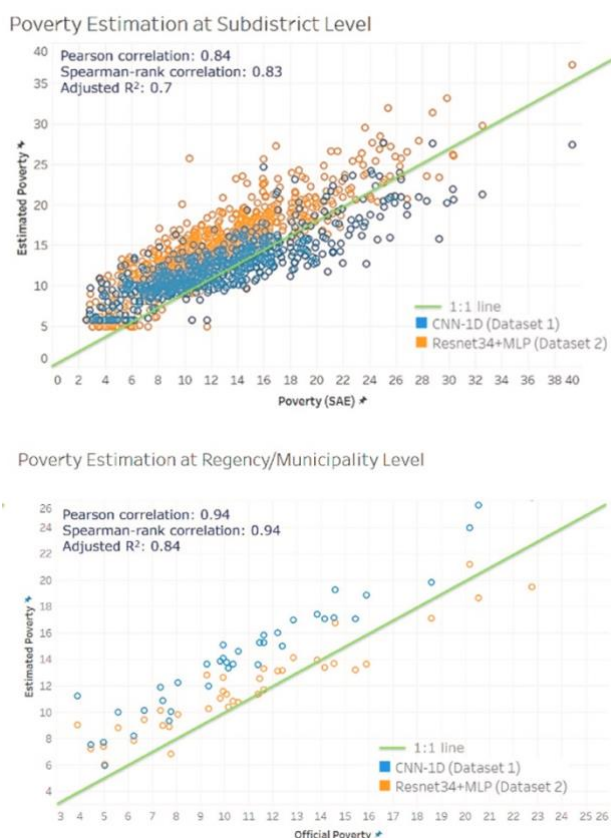
S. R. Putri ir kiti [26] pateikė granuliarnesnę požiūrį į skurdo lygio nustatymą – autoriai aprašė modelį, sukurtą naudojant kelių šaltinių palydovinius nakties vaizdus bei geografinius taškus, reikalingus požymių vektoriams ištraukti [26].

Pirma, NTL duomenys buvo naudojami kaip skurdo rodikliai – būdas atpažinti ekonominę veiklą; NDVI, kaip būdas atpažinti kaimo bei mažiau apgyvendintas vietas pagal augaliją; BUI (angl. *built-up index*) – užstatymo indeksas, kaip būdas atpažinti miesto zonas pagal pastatų pasiskirstymą; NDWI (angl. *normalized difference water index*) – normalizuotas vandens skirtumo indeksas, kaip

būdas atpažinti žemės dangą; LST (angl. *land surface temperature*) – žemės paviršiaus temperatūra, kaip būdas atpažinti miesto zonas pagal paviršiaus temperatūrą; anglies monoksidas (CO), natrio dioksidas (NO₂), ir sieros dioksidas (SO₂), kaip būdas atpažinti ekonomines veiklas pagal užterštumą; lankytinos vietos tankumas, kaip būdas nustatyti vietovės pasiekiamumą [26].

Antra, modelis buvo sukonstruotas iš dienos daugiajuosčių ir NTL intensyvumą fiksuojančių nuotraukų, naudojant giliojo mokymosi *ResNet-34* architektūrą [26].

Autoriai palygino skirtingų mašininio mokymosi – atraminių vektorių regresijos (angl. *support vector regression*, SVR), sprendimų medžių regresijos (angl. *decision tree regression*, DTR) bei atsitiktinių miškų regresijos (angl. *random forest regression*, RFR) algoritmus su giliojo mokymosi (daugiasluoksnio perceptrono (angl. *multilayer perceptron*, MLP) ir viendimensinio CNN (CNN-1D)) algoritmais. Remiantis modelių vertinimu, CNN-1D modelis buvo pasirinktas kaip geriausias pirmame scenarijuje, tuo tarpu antrame scenarijuje geriausi modeliai buvo *ResNet-34* bei MLP. Šie modeliai buvo toliau naudojami skurdui, su 1,5 km erdvine raiška, prognozuoti. Remiantis oficialiais skurdo duomenimis, CNN-1D modelis buvo pasirinktas kaip geriausias modelis su 1,95 RMSE bei 0,84 koreguotu determinacijos koeficientu (R^2) rajonų lygmeniu (9 pav.). Taip pat buvo vizualiai identifikuota, jog aukštas skurdo lygis paprastai nustatomas retai apgyvendintose vietovėse, apsuptose negyvenama žemės ūkio paskirties žeme. Ir priešingai – žemi skurdo rodikliai priskiriami labiau apgyvendintuose regionuose su patogiu susisiekimu [26].



9 pav. S. R. Putri ir kitų tyrimo rezultatai: skurdo prognozės rezultatų sklaidos diagrama [26]

L. Kondmann’as ir X. Xiang Zhu [27] iškėlė kiek kitokį tikslą – apžvelgti, ar mokymosi perkėlimu (angl. *transfer learning*) grįstas skurdo nustatymas gali būti naudojamas skurdo pokyčiams laike

fiksuoti. Autoriai sprendė klausimą, ar mokymosi perkėlimu strategija gali išmatuoti absoliučius vidutinės lokalsios gerovės pokyčius Ruandoje iš demografinių ir sveikatos tyrimų programos (angl. *The Demographic and Health Surveys*, DHS) ir „Landsat 7“ palydovo vaizdų. Buvo nustatyta, jog šiuo metodu nepavyko aptikti minėtų reikšmingų pokyčių šalyje nuo 2005 iki 2015 metų. Tai reiškia, jog nors su realiais skurdo duomenimis mokyti CNN modeliai ir gali nustatyti skurdą iš satelitinių vaizdų, prognozuoti skurdą į ateitį yra per anksti [27].

	Mean Training R^2 2005	Test R^2 2010	Test R^2 2015
Transfer Learning & ResNet50	0.47	0.69	0.57
Mean Nightlights per cluster	0.52	0.74	0.61
ResNet50	0.55	0.43	0.36

Training Scores are averaged over 10 runs of 10-fold cross validation.
Test scores are calculated on the full sample.

10 pav. L. Kondmann'o ir X. Xiang Zhu tyrimo rezultatai: skurdo lygio prognozės rezultatai Ruandoje su 2005 m. duomenimis [27]

Pastebima, kad daugiausia atliktų tyrimų naudojo *ResNet* – liekamųjų ryšių CNN architektūrą. Autoriai teikia teikia pirmenybę būtent šiam CNN modeliui, kadangi *ResNet* yra viena efektyviausių neuroninių tinklų architektūrų, padedanti palaikyti mažą klaidų lygį giliau tinkle. *ResNet*, naudodama liekamuosius blokus, taip pat praleidžia jungtis, o tai leidžia informacijai tekėti iš vieno sluoksnio į kitą, praleidžiant tarp jų esančius sluoksnius [28].

Daugiausia tyrimų buvo atlikta su naktiniais Afrikos vaizdais (Etiopijos, Nigerijos, Malio, Malavio, Ruandos, Ugandos, Tanzanijos), taip pat Meksikos, Kinijos, ar Indonezijos satelitinėmis nuotraukomis. Pastebima, jog tyrimuose dažniausiai naudoti namų ūkio apklausos rezultatai, kadangi kitokie duomenų tipai neprieinami arba jų trūksta. Lentelėje žemiau galima pamatyti susistemintą informaciją apie tai, kokie skurdo rodikliai, architektūros bei modeliai buvo naudojami šiems eksperimentams atlikti. Taip pat kokie yra gautų modelių rezultatai – determinacijos koeficientas R^2 .

4 lentelė. Sintezės lentelė - panašūs tyrimai

Autorius	Metai	Valstybė	Skurdo rodiklis	Architektūra	Modelis	R^2
Jean, Neal, et al	2016	Etiopija Nigerija Malis Malavis	Išlaidų duomenys (iš LSMS), turto indeksas (iš DHS)	<i>ResNet-34</i>	<i>Ridge</i> regresija	0,55-0,75
Babenko, Boris, et al	2017	Meksika	Namų ūkio apklausos rezultatai (MCS-ENIGH)	<i>GoogleNet</i>	Nespecifikuota	0,54-0,61
Chen, Derek	2017	Uganda Nigerija Tanzanija Malavis	Namų ūkio apklausos rezultatai (LSMS)	<i>ResNet-50</i>	<i>Ridge</i> regresija	0,45-0,65
Perez, Anthony, et al	2017	Ruanda Uganda Nigerija Tanzanija Malavis	Namų ūkio apklausos rezultatai (DHS)	<i>ResNet-18</i> , <i>ResNet-34</i> , VGG-F	<i>Ridge</i> regresija, GBT	0,63-0,66

Autorius	Metai	Valstybė	Skurdo rodiklis	Architektūra	Modelis	R ²
Wu, Peng, ir Yumin Tan	2019	Kinija	PCGDP (vienam gyventojui tenkantis BVP)	<i>ResNet-50</i>	<i>Ridge</i> regresija	0,76
Kondmann, Lukas, ir Xiao Xiang Zhu	2020	Ruanda Uganda Nigerija Tanzanija Malavis	Namų ūkio apklausos rezultatai (DHS)	<i>ResNet-50</i>	<i>Ridge</i> regresija	0,36-0,43
Putri, Salwa Rizqina, et al	2023	Indonezija (Rytų Java)	Metiniai Indonezijos Nacionalinės Socioekonominės Apklausos duomenys	<i>ResNet-34</i>	SVR, DTR, RFR, MLP, CNN-1D	0,84

1.5. Literatūros analizės apibendrinimas

Taigi, skurdas yra giliai įsišaknijusi problema, su kuria, kurdamos į skurdo mažinimą orientuotas strategijas ir programas, kovoja įvairios tarptautinės organizacijos. Siekiant suprasti, ar priemonės skurdui mažinti veikia, svarbu stebėti skurdo lygio pokyčius. Yra daug skirtingų skurdo matavimo rodiklių, tačiau dažnai jie yra netikslūs arba nepakankamai informatyvūs: nors skurdas dažniausiai matuojamas reguliariai, dėl laiko ir pinigų išteklių, jis matuojamas retai.

Lietuvoje naudojami kelių tipų skurdo rodikliai. Nuo įstojimo į Europos Sąjungą Lietuvoje skurdo lygiui apskaičiuoti imtas naudoti santykinio skurdo apskaičiavimo metodas, kuris apibrėžiamas kaip asmenų, kurių ekvivalentinės piniginės disponuojamosios pajamos mažesnės už skurdo rizikos ribą, dalis [29]. Šis rodiklis teikiamas tiek kaip oficialioji statistika, tiek kaip eksperimentinė. Eksperimentinė statistika, nors ir nėra galutinai išbandyta, ištestuota bei suderinta metodologiškai, yra dažnai tikslesnė nei oficialioji. Eksperimentiniai skurdo rodikliai taip pat paskelbiami beveik metais anksčiau nei įprasta skurdo statistika. Visgi, visi skelbiami rodikliai yra metiniai – agreguotų duomenų trūkumas neleidžia įsigilinti į skurdo kaitą ir nesuteikia informacijos apie skurdą mažinančių priemonių veiksmingumą.

Satelitiniai vaizdai, kaip alternatyvus būdas skurdo lygiui matuoti, pirmąkart panaudoti palyginti neseniai – vos 2016 metais. Nuo to laiko, šioje srityje buvo atlikta ir daugiau tyrimų. Pagrindinė nagrinėtų tyrimų geografinė aprėptis – Afrikos bei Azijos šalys, ar atskiri regionai, o naudojami skurdo rodikliai – apklausų duomenys. Yra tyrimų, naudojančių NPP-VIIRS satelitinius duomenis (N. Jean’as, P. Wu, S. R. Putri), tačiau daug tyrimų naudoja ir prastos raiškos DMSP-OLS vaizdus (D. Chen’as, A. Perez’as, L. Kondmann’as), ar ne nakties, o dienos nuotraukas (B. Babenko). Nors rezultatai varijuoja, tačiau autoriai sutinka, jog CNN turi daug potencialo ir gali būti naudojami skurdo lygiui nustatyti iš satelitinių nakties vaizdų.

Šiame darbe nuspręsta prognozuoti skurdą iš Lietuvos satelitinių nakties vaizdų, mokant modelius su skirtingais CNN metodais gautais požymių vektoriais bei eksperimentiniais skurdo rizikos lygio rodikliais. Projekte išsikeltas tikslas – detalesnė skurdo lygio stebėsena ir prognostika, pasitelkiant satelitinius nakties vaizdus. Šiam tikslui pasiekti iškeliama tokie uždaviniai:

1. apžvelgti atliktus panašaus pobūdžio mokslinius darbus;

2. pasirinkti konvoliucines architektūras vaizdams;
3. pasirinkti ir paruošti vaizdų rinkinius eksperimentams;
4. pritaikyti pasirinktas CNN architektūras vaizdams vektorizuoti;
5. sudaryti mašininio mokymosi modelius bei įvertinti jų tikslumą;
6. palyginti geriausio modelio rezultatus su laiko eilučių modelių prognozėmis;
7. prognozuoti skurdo rodiklius 2024 m. sausio–balandžio mėnesiams;
8. apibendrinti tyrimo rezultatus ir pateikti išvagas.

2. Tyrimo metodai

Tyrimo metodų dalyje glaustai aptariami pagrindiniai naudojamų metodikų aspektai – paaiškinamas CNN veikimo principas, aptariamose eksperimente naudojamos CNN architektūros bei apžvelgiami satelitinių nakties vaizdų šaltinių tipai – jų privalumai bei trūkumai.

Aptariami skirtingų tipų modeliai: regresijos (*Ridge*, *LASSO* bei *Elastic Net*) bei atsitiktiniai miškai. Aprašomi metodai, modelių tikslumui įvertinti: determinacijos koeficientas (R^2) bei paklaidos (RMSE, MAPE, MDAE, SMAPE, MAAPE). Aptariami skirtingi laiko eilučių prognozavimo metodai (ETS, *Theta*, ARIMA, *Naïve*, tiesinė regresija, *Prophet*).

Taip pat aprašomi ir palyginami iš skirtingų palydovinių jutiklių gaunami satelitiniai nakties vaizdai. Pabrėžiami koreguotų NASA „Black Marble“ NTL vaizdų privalumai. Paaiškinama, kokiais būdais gauti bei apdoroti Lietuvos satelitiniai vaizdai.

Galiausiai aprašoma atlikta požymių inžinerija.

2.1. Konvoliuciniai neuroniniai tinklai

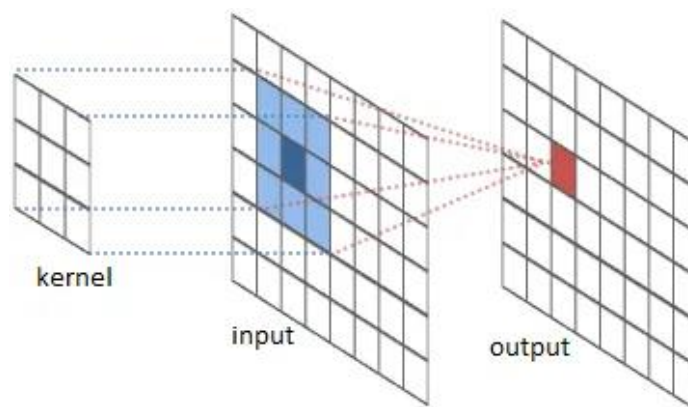
CNN – mašininio mokymosi pogrupis, vienas iš giluminio mokymosi algoritmų. CNN susideda iš mazgų sluoksnių, kuriuose yra įvesties sluoksnis, vienas ar daugiau paslėptų sluoksnių bei išvesties sluoksnis. Kiekvienas mazgas jungiasi prie kito mazgo, ir turi susietą svorį bei slenkstį. Jei kurio nors atskiro mazgo išvestis viršija nurodytą slenkstinę reikšmę, tas mazgas suaktyvinamas ir siunčia duomenis į kitą tinklo sluoksnį [30].

CNN išsiskiria iš kitų neuroninių tinklų tuo, jog naudodami vaizdo, kalbos ar garso signalų įvestis, yra našesni. Jie turi tris pagrindinius sluoksnių tipus, kurie yra:

1. Konvoliucinis sluoksnis;
2. Telkimo (angl. *pooling*) sluoksnis;
3. Visiškai sujungtas (angl. *fully-connected* (FC)) sluoksnis, kitaip – klasifikavimo „galva“.

2.1.1. Konvoliuciniai sluoksniai

Konvoliucinis sluoksnis yra pagrindinis CNN kūrimo blokas, kuriame atliekama daugiausia skaičiavimų. Tam reikia kelių komponentų: įvesties duomenų, filtro ir funkcijų žemėlapių. Taip pat reikalingas funkcijų detektorius, dar žinomas kaip branduolys arba filtras. Šis procesas vadinamas konvoliucija [30].



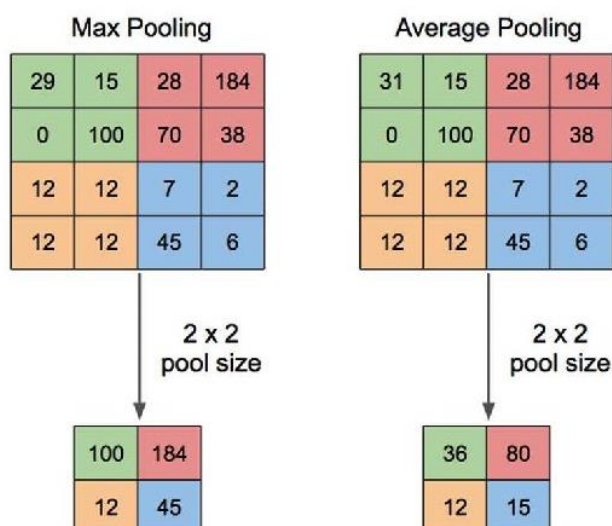
11 pav. Konvoliucinis sluoksnis, siekiant išskirti požymius, transformuoja įvesties vaizdą. Šioje transformacijoje vaizdas sujungiamas su branduoliu (arba filtru) [31]

2.1.2. Sutelkimo sluoksniai

Sutelkimo sluoksniai yra dar žinomi kaip sumažinimo sluoksniai (angl. *downsampling*). Šie sluoksniai, sumažindami įvesties parametrų skaičių, atlieka dimensionalumo mažinimą. Panašiai kaip konvoliuciniuose sluoksniuose, čia sutelkimo operacija taip pat „nubraukia“ filtrą per visą įvestį, tačiau kitaip nei konvoliuciniuose sluoksniuose, filtras neturi jokių svorių. Vietoje to, branduolys pritaiko agregavimo funkciją reikšmėms priimamajame lauke, užpildydamas išvesties masyvą.

Yra du pagrindiniai telkimo tipai [30]:

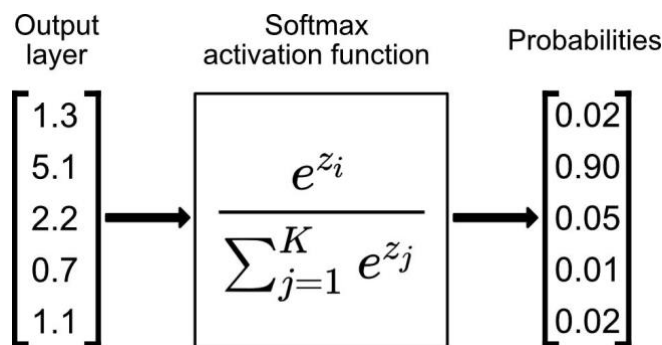
1. Maksimalus sutelkimas: kai filtras juda per įvestį, jis pasirenka didžiausios vertės pikselį, išsiunčiamą į išvesties masyvą. Šis metodas yra populiariausias.
2. Vidurkio sutelkimas: kai filtras juda per įvestį, jis pasirenka suskaičiuoja verčių vidurkį, išsiunčiamą į išvesties masyvą.



12 pav. Maksimalaus ir vidurkio sutelkimo procesai [32]

2.1.3. Visiškai sujungtas sluoksnis

Visiškai sujungtame sluoksnyje kiekvienas išvesties sluoksnio mazgas tiesiogiai jungiasi prie ankstesnio sluoksnio mazgo. Šis sluoksnis atlieka klasifikavimo užduotį pagal savybes, išskirtas per ankstesnius sluoksnius ir skirtingus jų filtrus. Nors konvoliuciniai ir telkimo sluoksniai dažniausiai naudoja *ReLU* (glodintos tiesinės funkcijos) funkcijas, FC sluoksniai paprastai naudoja *softmax* aktyvinimo funkciją [30].



13 pav. *Softmax* veikimo iliustracija [33]

2.1.4. Konvoliucinių neuroninių tinklų architektūros

Projekto tyrimams atlikti buvo naudotos šešios giliųjų CNN architektūros: *ResNet-101*, *ResNet-50*, *NASNetMobile*, *EfficientNet-B0*, *DenseNet-201* bei *DenseNet-121*. Šie modeliai pasirinkti dėl jų prieinamumo Keras atvirojo kodo bibliotekoje, taip pat dėl skirtingo sluoksnių ir parametrų kiekio. Tolesniuose skyreliuose bus trumpai aptariamose naudojamų modelių architektūros bei jų ypatumai.

2.1.4.1. ResNet-101 ir ResNet-50

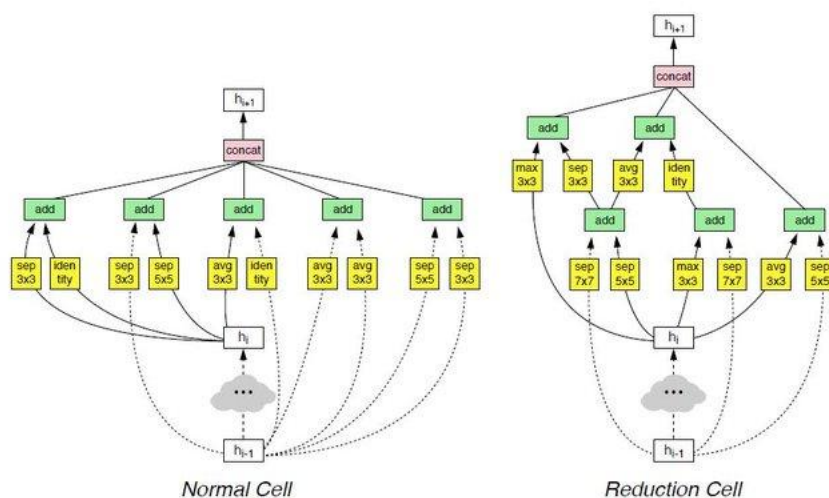
ResNet yra tokia neuroninių tinklų architektūra, kuri naudoja liekamuosius ryšius nykstančiojo gradiento (angl. *vanishing gradient*) problemai spręsti. Šioje architektūroje kiekvienas sluoksnis naudoja ne tik ankstesnio sluoksnio išvestis, bet ir originalius nepalietus duomenis požymiams gauti [34]. Tokiu būdu modelis lengviau optimizuoja savo parametrus. Naudojant šiuos ryšius galima mokyti gilesnius, daugiau sluoksnių turinčius neuroninius tinklus. *ResNet* pavadinime dažniausiai nurodoma, kiek sluoksnių turi tinklas, pvz., šiuo atveju *ResNet-50* bei *ResNet-101* turi atitinkamai 50 ir 101 konvoliucinį sluoksnį. *ResNet* struktūra pateikiama iliustracijoje žemiau (14 pav.) [35].

layer name	34-layer	50-layer	101-layer
conv1	$7 \times 7, 64, \text{stride } 2$		
	$3 \times 3 \text{ max pool, stride } 2$		
conv2_x	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
conv4_x	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$
conv5_x	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	average pool, 2048-d fc		

14 pav. ResNet-34, ResNet-50 ir ResNet-101 architektūrų struktūros [35]

2.1.4.3. NASNetMobile

NASNet (angl. *Neural Architecture Search Network*) – CNN, kuris buvo rastas per neuroninės architektūros paiešką. Pagrindiniai NASNet komponentai yra dviejų tipų: normalusis ir mažinantysis (angl. *reduction*). Normalusis elementas apibrėžia požymių vektorių žemėlapi, o mažinantysis – sumažina požymių žemėlapi du kartus [36].

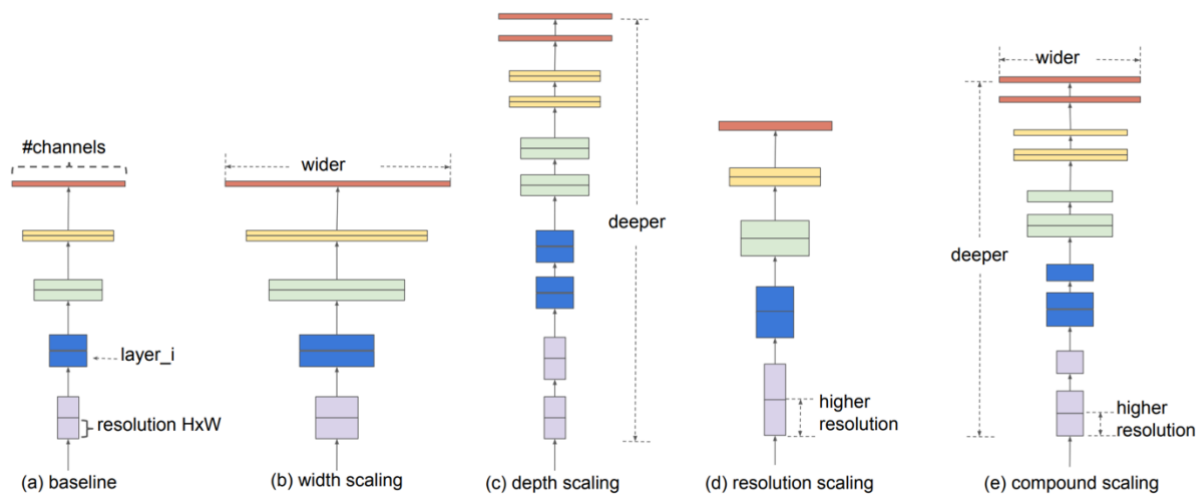


15 pav. Scheminė NASNet paieškos sistema [39]

2.1.4.4. EfficientNet-B0

EfficientNet yra tokia CNN architektūra, kuri, skirtingai nuo įprastos praktikos, tolygiai išplečia visas gylio/pločio/raiškos dimensijas, naudodama sudedamąjį koeficientą (angl. *compound coefficient*) [37]. Šis metodas remiasi tuo, jog jei įvesties nuotrauka yra didesnė, tuomet tinklui reikia daugiau sluoksnių laukams padidinti bei daugiau kanalų smulkesnėms detalėms didesniuose vaizduose užfiksuoti [37].

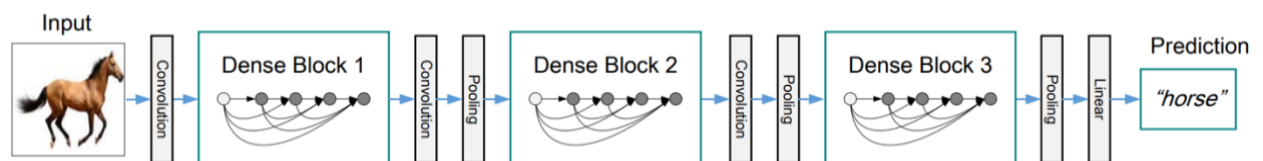
EfficientNet-B0 bazė yra grindžiama apverstomis liekamųjų MobileNetV2 blokų kliūtėmis (16 pav.).



16 pav. *EfficientNet-B0* modelio veikimo principas [37]

2.1.4.5. DenseNet-201 ir DenseNet-121

DenseNet yra toks CNN tipas, kuris panaudoja tankius sąryšius tarp sluoksnių, pro tankiai sujungtų sąryšių blokus (angl. *dense blocks*). Čia visi sluoksniai (su sutampančiais požymių žemėlapių dydžiais) yra tiesiogiai sujungiami vienas su kitu. Pakartotinis gautų požymių naudojimas pagerina 31 gradientų sklaidimą tinkle bei leidžia sumažinti konvoliucinių filtrų skaičių, todėl tokio tipo tinklai greičiau mokosi [38].



17 pav. *DenseNet* modelio su trimis tankiai sujungtų sąryšių blokais veikimo principas [41]

2.2. Regresija

Regresinė analizė yra statistinis metodas, leidžiantis tyrinėti ryšius tarp vieno ar daugiau kintamųjų. Daugelyje sričių, įskaitant ekonometriją, chemiją ir inžineriją, vieni dažniausiai naudojamų regresijos tipų yra *Ridge*, *LASSO* bei *Elastic Net*.

2.2.1. Ridge regresija

Ridge regresija (taip pat žinoma, kaip *L2* reguliarizacija) yra daugkartinės regresijos (angl. *multiple-regression*) modelių koeficientų įvertinimo metodas tokiuose scenarijuose, kai nepriklausomi kintamieji stipriai koreliuoja. *Ridge* regresija yra vienas iš reguliarizacijos tipų linijiniams regresijos modeliams. Reguliarizacija, kaip statistinis metodas, yra skirtas sumažinti klaidoms, gaunamoms dėl mokymo duomenų dalies permokymo, tuo tarpu *Ridge* regresija pakoreguoja regresinės analizės multikolinearumą. Tai yra naudinga tuomet, kai vystomi mašininio mokymosi modeliai turi didelį kiekį parametrų, ypač jei tie parametrai taip pat turi didelius svorius [39].

Ridge regresija modifikuoja mažiausių kvadratų metodą (angl. *Ordinary Least Squares*, OLS) apskaičiuodama koeficientus, kurie atsižvelgia į potencialiai koreliuojančius prediktorius (angl. *predictors*). Ji pakoreguoja didelio reikšmingumo koeficientus įvedant reguliarizacijos sąlygą (dažnai vadinamą „baudos“ sąlyga) į RSS (čia RSS – liekamoji paklaidų kvadratų suma, angl. *Residual Sum of Squares*) funkciją. Ši baudos sąlyga yra modelio koeficientų kvadratų suma. Ji yra išreiškiama tokia formule (1) [39]:

$$L2 = ||B||^2 = B_1^2 + B_2^2 + \dots + B_n^2. \quad (1)$$

čia $L2$ – baudos sąlyga;

B – regresijos koeficientas.

$L2$ baudos sąlyga tuomet yra įterpiama į RSS funkcijos pabaigą, taip sukuriant naują formulę, vadinamą *Ridge* regresijos koeficientu (angl. *estimator*). Jame šios sąlygos poveikį modeliui valdo derinimo parametras λ (lambda):

$$RSS_{L2} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda \sum_{j=1}^P B_j^2. \quad (2)$$

čia RSS – liekamoji paklaidų kvadratų suma;

n – stebėjimų skaičius;

Y_i – faktinės priklausomojo kintamojo reikšmės taške i ;

$\hat{Y}_i$ – prognozuojamos priklausomojo kintamojo reikšmės taške i ;

λ – derinimo parametras lambda;

P – nepriklausomųjų kintamųjų skaičius;

B – regresijos koeficientas.

2.2.2. LASSO regresija

LASSO (angl. *Least Absolute Shrinkage and Selection Operator*) regresija (taip pat žinoma, kaip $L1$ reguliarizacija) yra kitas metodas, padedantis spręsti permokymo bei optimistinio šališkumo (angl. *optimism bias*, modelio veikimo pervertinimo) problemas.

LASSO regresija siekia identifikuoti kintamuosius ir atitinkamus regresijos koeficientus, kurie minimizuoja modelio prognozavimo klaidas. Tai pasiekama įvedant modelio parametų apribojimus, kurie mažina regresijos koeficientus link nulio, t. y. padaro regresijos koeficientų absoliučių reikšmių sumą mažesnę nei fiksuota reikšmė (λ):

$$L1 = |B_1| + |B_2| + \dots + |B_n|. \quad (3)$$

čia $L1$ – baudos sąlyga;

B – regresijos koeficientas.

$$RSS_{L1} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda \sum_{j=1}^P |B_j|. \quad (4)$$

Trumpai tariant, apribojamas modelio kompleksiskumas. Kintamieji su nuliniu regresijos koeficientu tuomet iš modelio išimami [39].

Taigi, LASSO regresija sumažina nepriklausomų kintamųjų kiekį, tuo tarpu *Ridge* regresija sumažina nepriklausomų kintamųjų svorį.

2.2.3. Elastic Net regresija

Elastic Net regresija (elastiniai tinklai) yra metodas, sujungiantis LASSO ir *Ridge* regresijas bei išsprendžiantis jų apribojimus (5).

$$L_{Elastic\ Net} = \frac{\sum_{i=1}^n (y_i - x_i' \hat{\beta})^2}{2n} + \lambda \left(\frac{1-\alpha}{2} \sum_{j=1}^k \hat{\beta}_j^2 + \lambda \alpha |\hat{\beta}_j| \right). \quad (5)$$

$$RSS_{Elastic\ Net} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda_1 \sum_{j=1}^P |B_j| + \lambda_2 \sum_{j=1}^P B_j^2. \quad (6)$$

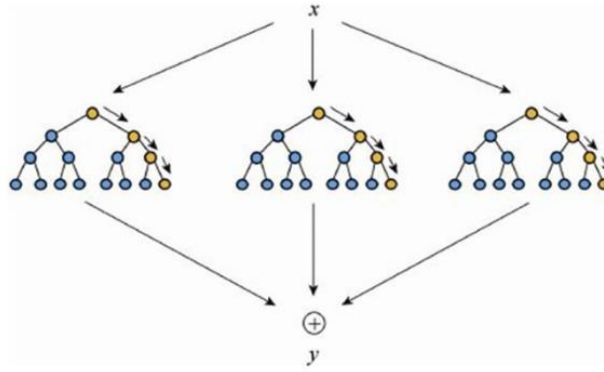
Šioje funkcijoje, naujas parametras α yra „maišymo“ parametras, kuris išbalansuoja minėtuosius du metodus. Jei α yra 0, tuomet grįžtama į *Ridge* regresiją, jei $1 - \alpha$ į LASSO (5) [40].

Šio regularizacijos tipo privalumai yra tai, kad ją galima naudoti net tuomet, kai kintamųjų skaičius yra didesnis nei stebėjimų. Elastiniai tinklai apjungia LASSO ir *Ridge* privalumus: dėl *Ridge* regularizacijos, elastinio tinklo koeficientas gali susitvarkyti su koreliacijomis tarp predikatorių geriau nei LASSO, tuo tarpu dėl L1 regularizacijos sumažėja nepriklausomų kintamųjų kiekis. Tiesa, LASSO šališkumo problema egzistuoja ir *Elastic Net* metode [40].

2.3. Atsitiktiniai miškai

Atsitiktiniai miškai yra mašininio mokymosi algoritmų kombinacija, naudojama klasifikavimo ir regresijos uždaviniuose. L. Breiman'as [41] teigia, jog atsitiktiniai miškai yra priskiriami ansamblių metodui (angl. *ensemble method*), o jiems „užauginti“ dažnai sugeneruojami atsitiktiniai vektoriai, prižiūrintys kiekvieno ansamblio medžio augimą.

Atsitiktinių miškų veikimo principas grįstas tuo, jog kiekvienas ansamblio medis yra sudarytas iš mokymo imties duomenų dalies (angl. *bootstrap sample*), iš kurios 1/3 atskiriama, kaip testavimo dalis. Ši testavo dalis dar vadinama mokymosi procese nedalyvavusia testavimo duomenų imtimi (angl. *out-of-bag sample*, OOB). Dar vienas atsitiktinumo veiksnys įterpiamas per požymių kaupimą (angl. *feature bagging*), kuris suteikia daugiau įvairovės bei sumažina koreliaciją tarp sprendimų medžių. Tuomet OOB yra naudojamas kryžminei validacijai atlikti [41].



18 pav. Atsitiktinių miškų veikimo principas [42]

Taigi, nors sprendimų medžiai yra dažni prižiūrimo mokymosi algoritmai, jiems neretai būdingas šališkumas ir pervertinimas. Visgi, kai daug sprendimų medžių suformuoja ansamblį atsitiktinių miškų algoritme, tuomet gaunami tikslesni rezultatai, ypač tada, kai atskiri medžiai nekoreliuoja vienas su kitu.

2.4. Modelio tikslumo įvertinimo parametrai

Siekiant įvertinti modelių veikimą ir palyginti skirtingais CNN būdais gautus rezultatus, pasirinkti šie metodai: determinacijos koeficientas bei skirtingi paklaidų (tikslumo) matai.

Matematinėse formulėse bus naudojami šie raidiniai žymėjimai:

- Y_i – faktinės priklausomojo kintamojo reikšmės taške i ;
- $\hat{Y}_i$ – prognozuojamos priklausomojo kintamojo reikšmės taške i ;
- $\bar{Y}$ – priklausomo kintamojo vidurkis;
- n – stebėjimų skaičius.

2.4.1. Determinacijos koeficientas

Determinacijos koeficientas, arba R^2 (angl. *R-Squared*). Šis statistinis matas yra naudojamas tiesinės regresijos modelio tikslumui įvertinti, t. y., nustatyti, kokią prognozuojamo kintamojo variacijos proporciją paaiškina regresijos modelis. Determinacijos koeficientas apskaičiuojamas, kaip koreliacijos tarp stebimų reikšmių Y bei prognozuojamų reikšmių $\hat{Y}$ (klasikinės Pirsono koreliacijos kvadratas):

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}. \quad (7)$$

2.4.2. Paklaidos

Šiame darbe taip pat bus lyginamos kelių tipų paklaidos:

Vidutinis kvadratinis nuokrypis (angl. *Root Mean Squared Error*, RMSE). Šis tikslumo matas matuoja vidutinį skirtumą tarp prognozuojamų ir tikrųjų reikšmių (8). Matematiškai tai yra standartinis liekanų nuokrypis [43].

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{Y}_i - Y_i)^2}{n}}. \quad (8)$$

Vidutinė absoliutinė procentinė paklaida (angl. *Mean Absolute Percentage Error*, MAPE). Šis matas apskaičiuojamas, kaip p (čia p – tikrųjų ir prognozuojamų reikšmių skirtumas, padalintas iš tikrųjų reikšmių) vidurkis (9) [43].

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{Y_i - \hat{Y}_i}{Y_i} \right). \quad (9)$$

Vidutinė absoliutinė paklaida (angl. *Median Absolute Error*, MDAE). Šis paklaida apskaičiuojama, kaip absoliutinio skirtumo tarp prognozuojamų ir tikrųjų reikšmių mediana (10) [44].

$$MDAE = \text{median}(|\hat{Y}_i - Y_i|). \quad (10)$$

Vidutinė simetrinė absoliutinė procentinė paklaida (angl. *Symmetric Mean Absolute Percentage Error*, SMAPE). Kadangi MAPE turi trūkumų (šio rodiklio rezultatai reikšmingai paveikiami, kai Y yra 0), M3 prognozavimo varžybose buvo pasiūlytas alternatyvus klaidų skaičiavimo metodas SMAPE (11). Visgi, kadangi matas yra nestabilus (jo reikšmė gali būti neigiama), šio skaičiavimo būdo naudoti nerekomenduojama [43]. Nors eksperimente SMAPE ir įtraukiama į modelių vertinimą, dėl minėtų priežasčių, renkant geriausią modelį, į ją neatsižvelgiama.

$$SMAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|Y_i - \hat{Y}_i|}{(|Y_i| + |\hat{Y}_i|)/2} \right). \quad (11)$$

Vidutinė arktangento absoliutinė procentinė paklaida (angl. *Mean Arctangent Absolute Percentage Error*, MAAPE). Kaip ir SMAPE, šis matas buvo taip pat pasiūlytas M3 prognozavimo varžybose. MAAPE skaičiavimas naudojamas procentinėms paklaidoms transformuoti į ribotą diapazoną nuo $-\frac{\pi}{2}$ iki $\frac{\pi}{2}$ [45].

$$MAAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \arctan \left(\left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \right). \quad (12)$$

2.5. Laiko eilučių prognozavimo metodai

Siekiant įvertinti geriausio modelio veikimą, prognozės iš satelitinių nakties vaizdų taip pat lyginamos ir su laiko eilučių metodais gautomis prognozėmis. Darbe naudojami 6 skirtingi laiko eilučių metodai: ETS, *Theta*, ARIMA, *Naïve*, tiesinė regresija bei *Prophet*.

2.5.1. ETS

Eksponentinio glotninimo (angl. Exponential Smoothing, ETS) metodas yra grįstas praeities stebėjimų svertiniu vidurkiu, kuris bėgant laikui eksponentiškai mažėja, t. y., naujesni stebėjimai turi didesnę svorį nei ankstesni. Dėl to prognozės yra „protingesnės“, lyginant su įprastais laiko eilučių metodais [46].

Darbe naudojami du ETS prognozavimo būdai – ETS (A, N, N) bei ETS (M, N, N). Čia A, N, N yra:

- A – suminė paklaida (angl. *additive error*);
- N – be trendo (angl. *no trend*);
- N – be sezoniškumo (angl. *no seasonality*).

M, N, N:

- M – dauginamoji paklaida (angl. *multiplicative error*);
- N – be trendo (angl. *no trend*);
- N – be sezoniškumo (angl. *no seasonality*).

2.5.2. Theta

Theta modelis yra vienmatis prognozavimo metodas, pagrįstas laiko eilučių išlinkio (angl. *curvature*) modifikavimu, naudojant koeficientą *Theta* ($\theta \in \mathbb{R}$). Dėl šios modifikacijos atsiranda naujos eilutės, turinčios originalių laiko eilučių vidurkį ir nuolydį. Kai *Theta* koeficientas yra $0 \leq \theta < 1$, tuomet pastebima kreivės defliacija bei galima identifikuoti ilgalaikius trendus. Ekstremaliais atvejais, kai $\theta = 0$, laiko eilutės transformuojasi į tiesinės regresijos tiesę. Kai $\theta = 1$, gaunamos originalios laiko eilutės. Jei *Theta* pakyla virš 1, tuomet laiko eilutės laikomos išsiplėtusiomis bei galima identifikuoti trumpalaikę elgseną [47].

Originalus *Theta* modelis, sukuriantis *Theta* liniją $Z(\theta)$, gaunamas pagal tokią formulę:

$$\nabla^2 Z_t(\theta) = \theta \nabla^2 Y_t, \quad t = 3, 4, \dots, n. \quad (13)$$

čia ∇ – skirtumo operatorius.

2.5.3. ARIMA

Integruotas autoregresinis slenkančio vidurkio modelis (ARIMA) yra generalizuotas autoregresinis slenkančio vidurkio modelis (ARMA), apjungiantis autoregresinį (AR) ir slenkančio vidurkio (MA) procesus bei sukuriantis sudėtinį laiko eilučių modelį. Pagrindiniai ARIMA (p, d, q) modelio elementai yra [48]:

- AR: autoregresija. Regresinis modelis, naudojantis stebėjimo ir vėluojančių (angl. *lagged*) stebėjimų priklausomybes (p);
- I: integracija. Integracija, paverčianti laiko eilutes stacionariomis, matuojant stebėjimų skirtumus skirtingu laiku (d);
- MA: slenkantis vidurkis. Metodas, kuriuo atsižvelgiama į stebėjimų ir liekanų paklaidų priklausomybę (q).

Paprasta AR (p) modelio forma gali būti užrašoma kaip tiesinis procesas [48]:

$$x_t = c + \sum_{i=1}^p \varphi_i x_{t-1} + \epsilon_t. \quad (14)$$

čia x_t – stacionarusis kintamasis;

c – konstanta;

φ_i – autokoreliacijos koeficientai 1, 2, ..., p vėlavimuose;

ϵ_t – baltasis triukšmas su nuliniu vidurkiu ir dispersija σ_ϵ^2 ;

t – laiko momentas.

MA (q) modelio procesas gali būti išreiškiamas tokia formule [48]:

$$x_t = \mu + \sum_{i=0}^q \theta_i \epsilon_{t-i}. \quad (15)$$

čia $\mu - x_t$ prielaida;

$\theta_i t$ – svoriai, taikomi esamoms ir ankstesnėms stochastinėms laiko eilučių reikšmėms.

Tuomet ARIMA (p, q) formulė gaunama sujungus AR (p) bei MA (q) [48]:

$$x_t = c + \sum_{i=1}^p \varphi_i x_{t-1} + \epsilon_t + \sum_{i=0}^q \theta_i \epsilon_{t-i}. \quad (16)$$

Kur $\varphi_i \neq 0$ ir $\sigma_\epsilon^2 > 0$.

2.5.4. Naïve

Naïve metodas yra vienas paprasčiausių prognozavimo metodų. Šiuo metodu prognozės reikšmė yra laikoma lygia naujausiai tikrajai reikšmei (17).

$$\hat{Y}_i = Y_i - 1. \quad (17)$$

Statistinis modelis, kuriuo grindžiamas Naïve yra vadinamas „Random Walk“ bei užrašomas tokia formule:

$$Y_i = Y_{i-1} + \epsilon_i. \quad (18)$$

čia ϵ_i kintamumas turi įtakos duomenų kaitos greičiui: kuo jis didesnis, tuo sparčiau keisis reikšmės.

2.5.5. Tiesinė regresija

Tiesinis laiko eilučių metodas yra paprasta tiesinė regresija, kuri daro prielaidą, jog santykis tarp prognozuojamų ir stebimų reikšmių yra tiesinis [49]:

$$Y_t = X_t \beta + \epsilon_t. \quad (19)$$

Čia β – regresijos koeficientas;
 X – nepriklausomas kintamasis.

Šis metodas paremtas prielaidomis, jog:

- Regresijos funkcija koeficientų ir paklaidų atžvilgiu yra tiesinė;
- Modelis yra nestochastiškas;
- Paklaidų vidurkis lygus nuliui (nulinis vidurkis);
- Paklaidų dispersija yra pastovi (ne heteroskedastiškumas);
- Paklaidos neautokoreliuoja.

2.5.6. Prophet

Prophet yra „R“ ir „Python“ programavimo kalbose įdiegta prognozavimo procedūra. Šis metodas remiasi: (1) A. C. Harvey ir S. Peters'o pasiūlyta laiko eilučių dekompozicija [50], kuomet netiesiniams trendams priskiriami metiniai, savaitiniai ir dieniniai sezoniškumo efektai, ir (2) T. Hastie ir R. Tibshirani [51] pasiūlytu apibendrintu adityviu modeliu (angl. *generalized additive model*, GAM).

Toks metodas geriausiai veikia su laiko eilutėmis, kurios pasižymi stipriu sezoniškumu ir bent keliais istorinių duomenų periodais. *Prophet* efektyviai tvarko trūkstamas reikšmes bei trendo pokyčius, taip pat sėkmingai valdo išskirtis [52].

2.6. Satelitiniai nakties vaizdai

Egzistuoja kelių tipų satelitinių NTL duomenys: kombinuoti NTL duomenys iš DMSP-OLS bei naujos kartos kombinuoti NTL duomenys iš NPP-VIIRS, gaunami iš Suomi NPP – bendro NASA ir NOAA palydovo. Neseniai, siekdama išnaudoti visą NPP-VIIRS potencialą, NASA sukūrė naują standartizuotą NTL produktų rinkinį, atitinkantį dabartines NTL technologijas – NASA „Black Marble“ produktų rinkinį („VNP46“, „VJ146“).

Šiame skyriuje lyginamos senosios (OLS) ir naujosios (VIIRS) palydovinių jutiklių technologijos bei pagrindžiama, kodėl pasirinkta naudoti būtent NASA „Black Marble“ vaizdų rinkinį. Taip pat paaiškinama, koku būdu buvo gauti Lietuvos satelitiniai nakties vaizdai, apžvelgiamas jų paruošimas analizei bei pateikiamos svarbiausios žvalgomosios NTL duomenų analizės įžvalgos.

2.6.1. DMSP-OLS ir Suomi NPP-VIIRS vaizdų palyginimas

Nuo 1970-ųjų, DMSP pradėjo naudoti tokius palydovinius jutiklius, kurie gali atpažinti regimąją ir artimąją infraraudonąją spinduliuotę (angl. *visible and near infrared*, VNIR). Globalias dienos ir nakties šviesų Žemės nuotraukas DMSP-OLS gauna iš dviejų spektrinių juostų – regimosios juostos (angl. *visible*, VIS) ir šiluminės infraraudonosios juostos (angl. *thermal infrared*, TIR). Naktinis VIS dažnių praeinamumas apima VNIR spektro dalį (0,5–0,9 μm). VIS juostos signalas naktį sustiprinamas naudojant fotodaugintuvo vamzdelį (angl. *photomultiplier tube*, PMT), o tai leidžia aptikti silpnus VNIR spinduliuotės šaltinius. Nors PMT sistema buvo originaliai įdiegta debesims aptikti nakties metu, nenumatytas NTL intensyvėjimo atsiradimas leido aptikti ir miesto šviesas, dujų fakes bei gaisrus [53].

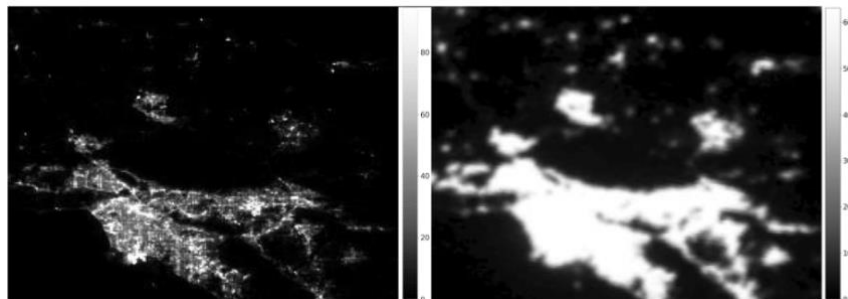
Nuo 1997 m., EOG inicijavo DMSP-OLS vaizdų sujungimą globaliems NTL žemėlapiams gauti. Kadangi turimi metiniai duomenys tęsiasi nuo 1992 iki 2013 m., DMSP NTL galima laikyti ilgiausia naktiniam nuotoliniam žmonių veiklai stebėti skirta duomenų serija [53].

Dėl jutiklio jautrumo naktį, DMSP-OLS duomenys gali būti naudojami kaip įrankis įvairioms VNIR emisijoms aptikti. Ilgalaikių kasdieninių duomenų su vidutine erdvine raiška (pvz.: 1 km) prieinamumas suteikia galimybę tirti įvairius globalaus, nacionalinio ir regioninio lygmens mokslinius klausimus. Nors DMSP-OLS privalumai yra plačiai žinomi ir pripažįstami, taip pat buvo identifikuoti ir keli trūkumai:

1. Erdvinė raiška yra santykinai žema;
2. Palydove trūksta kalibravimo įrenginio (pasikliaujama prieš paleidimą atliktu kalibravimu);
3. Nėra tarp-palydovinio kalibravimo;
4. Trūksta įrašų apie antenos stiprinimo pokyčius skrydžio metu;
5. Ribotas 6 bitų skaitmeninių skaičių kvantifikavimas;
6. Didelis šviesos sodrumas (angl. *saturation*) tankiai apgyvendintuose miestų centruose, esant standartiniam didelio antenos stiprinimo nustatymui (19 pav.);
7. Žydėjimo efektas (angl. *blooming effect*, kitaip – apšviesto ploto pervertinimas), atsirandantis dėl grubios erdvinės duomenų raiškos ir šviesos atspindžio iš gretimų teritorijų (pvz.: vandens telkinių) (19 pav.);
8. Ribota prieiga prie aukštos kokybės kasdieninių OLS duomenų [54].

Dėl šių priežasčių galima teigti, jog DMSP-OLS nakties nuotraukos yra nenuoseklios, ne itin kokybiškos ir pakankamai žemos raiškos. Taip pat šios nuotraukos prieš naudojimą reikalauja didelio apdirbimo. Nors anksčiau DMSP-OLS jutikliai buvo sėkmingai naudojami, jie kliaujasi senomis technologijomis, turinčiomis žemesnę raišką nei reikalauja dabartiniai moksliniai tyrimai.

EOG NTL žemėlapis, sudarytas iš DMSP-OLS duomenų, pirmąkart pasirodė 1994 m. [53]. Naujausios kartos Žemės stebėjimo palydovai (priklausantys JPSS) buvo paleisti 2011 m., o kartu su jais atsirado ir naujas neklasifikuotų palydovinių Žemės nakties vaizdų šaltinis [56]. SNPP ir NOAA palydovuose buvę NPP-VIIRS jutikliai parodė stulbinančius patobulėjimus žemo šviesos intensyvumo nuotraukose (lyginant su DMSP-OLS) (19 pav.). EOG pasinaudojo šia technologine pažanga, kad galėtų teikti aukščiausios kokybės globalius naktinių šviesų produktus [55].



19 pav. Los Andželo NTL vaizdai. Kairėje: NPP-VIIRS vaizdas. Dešinėje: DMSP-OLS vaizdai [51]

Suomi NPP-VIIRS duomenys yra naudojami kaip pagrindinis duomenų rinkinys didžiojoje daugumoje ekonominių tyrimų, naudojančių aukštos raiškos NTL. Kadangi tai yra iš esmės jau

apdirbtas produktas, tokius duomenis paprasta naudoti – užtenka atsisiųsti. Jis nereikalauja papildomų žingsnių, pavyzdžiui, nereikia:

- atsižvelgti į debesų dangą;
- kontroliuoti palydovo kampą;
- atsižvelgti į saulės spindulius [55].

Duomenys taip pat lengvai prieinami per „Google Earth Engine“ – debesų technologijos pagrindu veikiančia geografinių duomenų analizės platformą, kuri leidžia vartotojams vizualizuoti ir analizuoti Žemės palydovinius vaizdus. Vaizdus taip pat teikia EOG bei „Payne Institute for Public Policy“ – institutas, priklausantis „Colorado School of Mines“ (CSM) universitetui.

5 lentelė. DMSP-OLS ir NPP-VIIRS palyginimas [55, 56, 57, 58]

Kintamasis	DMSP-OLS	NPP-VIIRS
Savininkas / vystytojas	JAV Oro pajėgos	NASA – NOAA JPSS
Duomenų laikotarpis	1992–2013	2012–dabar
Orbita	Saulei sinchroninė (poliarinė) orbita: 850 km aukštis, 98,8 laipsnių nuolydis, 102 minutės	Saulei sinchroninė (poliarinė) orbita: 827 km aukštis, 98,7 laipsnių nuolydis, 102 minutes
Aprėptis	3000 km	3000 km
Pusiaujo kirtimo laikas	~8:30–9:30 ir 20:30–21:30 (vietiniu laiku)	~01:30 ir ~13:30 (vietiniu laiku)
Žemo šviesos intensyvumo nuotraukų juostos pralaidumas	Panchromatinis nuo 0,5 iki 0,9 um	Panchromatinis nuo 0,5 iki 0,9 um
Laiko raiška	12 h	12 h
Erdvinė raiška	2.7 km × 2.7 km	742 × 742 m
Bangos ilgio diapazonas	400–1 100 nanometrai	500–900 nanometrai
Papildomos spektrinės juostos	Šiluminė infraraudonųjų spindulių (10 um)	21 papildoma juosta nuo 0,4 iki 13 um
Kvantifikavimas	6 bitai	14 bitų
Sodrumas / saturacija	Dažna miestų zonose	Nėra
Žemo šviesos intensyvumo nuotraukų aptikimo limitas	~ 5 × 10 ⁻¹⁰ W/cm ² /sr	~ 2 × 10 ⁻¹¹ W/cm ² /sr
Kalibravimas	Žemo šviesos intensyvumo nuotraukų juostai – jokio.	Saulės difuzorius, naudojamas dienos duomenims kalibruoti. Kalibravimas išplėstas iki žemo šviesos intensyvumo nuotraukų režimo, naudojant duomenis, surinktus palei saulės terminatorių.
Tęstinumas	Paskutiniai palydovai, fiksuojantys naktines šviesas, skrido poliarinėmis orbitomis (20 pav.). DMSP NTL duomenys nebeprieinami viešai nuo 2013 m.	JPSS paleido tris palydovus - Suomi-NPP, NOAA-20 ir NOAA-21; planuojama paleisti JPSS-3 (2027 m.) ir JPSS-4 (2032 m.). Abu skris poliarinėmis orbitomis (20 pav.).



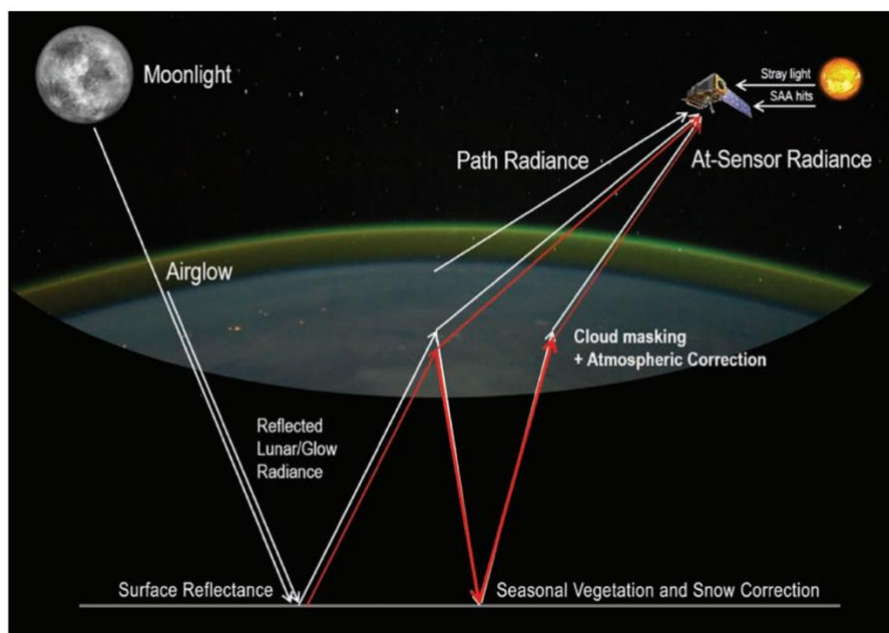
20 pav. Saulei sinchroninė (poliarinė) orbita – palydovų skriejimo trajektorija [59]

2.6.2. Koreguoti Suomi NPP-VIIRS vaizdai – NASA „Black Marble“

NASA „Black Marble“ yra NASA projektas, kuris teikia pakoreguotus NPP-VIIRS vaizdus. Šio produktų rinkinio algoritmai geba atpažinti ir pašalinti atmosferos, reljefo, mėnulio šviesos atspindžių, sniego, augmenijos bei šilumos efektus, veikiančius mūsų matomus šviesas, ir sukurti nuoseklius kasdienius žmogaus sukurtų šviesų duomenis [56, 60].

Pakoreguotas NTL spinduliavimas leidžia kiekybiškai aptikti ir analizuoti kasdieninius, sezoninius ir metinius svyravimus, dėl tokių patobulinimų taip pat sumažėja foninis triukšmas. Pagrindiniai atlikti keitimai, lyginant su EOG-CSM teikiamais vaizdais, apima:

- mėnulio spinduliuotės modeliavimą sprendžiant netiesinius fazės ir libracijos pokyčius;
- BRDF funkciją (angl. *bidirectional reflectance distribution function*) mėnulio šviesos atspindžių efektams ištaisyti;
- geometrinį-optinį ir atspindėjimo nuo lajos dangos modeliavimą, siekiant atsižvelgti į sezoninius NTL svyravimus;
- laiko spragų užpildymą, siekiant sumažinti pasikartojančią duomenų trūkumo problemą [60].



21 pav. NASA „Black Marble“ vaizdų apdorojimo procesas [56]

Nepaisant stulbinančio NTL vaizdų kokybės pagerinimo, NASA „Black Marble“ duomenys turi du reikšmingus trūkumus.

Pirma, NASA „Black Marble“ algoritmas nebando taisyti saulės šviesa užterštų pikselių, o juos „išmeta“. Palyginimui, NPP-VIIRS duomenis teikiančios EOG ir CSM organizacijos turi būdą, kaip saulės šviesa užterštus pikselius pakoreguoti. Dėl šios priežasties, NASA „Black Marble“ turi daug trūkstamų reikšmių kiekvienų metų birželio-rugpjūčio mėnesiais [55].

Antra, NASA „Black Marble“ duomenis yra sunku pasiekti. Neapdoroti NASA „Black Marble“ produktai yra prieinami per NASA LAADS DAAC sistemą, kuri leidžia vartotojams parsisiųsti kasdieninius, mėnesinius bei metinius NTL duomenis nuo 2012 m. iki dabar. Tačiau NASA „Black Marble“ duomenų konvertavimas į analizei paruoštą formatą reikalauja papildomų žingsnių, įskaitant atskirų vektorinių išsklotinių (angl. *vector tiles*) parsisiuntimą ir sugretinimą, konvertavimą į GeoTIFF (arba kitos geografinės informacijos sistemos) formatą bei žemos kokybės pikselių (pvz.: paveiktų debesų dangos) pašalinimą. Daugelis tyrimų taip pat reikalauja agreguoti NTL duomenis į didesnius erdvinius vienetus, pvz. siekiant fiksuoti vidutinius NTL pokyčius laikui bėgant [61].

NASA „Black Marble“ duomenų gavimui pagreitinti, Pasaulio Bankas (DIME (angl. *Development Impact Evaluation*) ir duomenų laboratorija (angl. *Data Lab*)) 2023 m. pabaigoje sukūrė „R“ bei „Python“ paketus: *blackmarbler* (pasiekiamu per „R“ programavimo kalbą) bei *blackmarblepy* (pasiekiamu per „Python“ programavimo kalbą). Šie paketai siūlo funkcijas, leidžiančias parsisiųsti duomenis tiesiai iš NASA LAADS DAAC. Jie taip pat yra ypač naudingi teikiant mėnesinių ar metinių agreguotų NTL duomenų užklausas, ar siekiant gauti duomenis tiesiai iš NASA „Black Marble“ archyvų [61].

Abu paketai siūlo dvi pagrindines funkcijas: (1) *bm_raster*, kuri leidžia parsisiųsti rastrinius NTL duomenis, pritaikytus konkrečiam dominančiam regionui, ir (2) *bm_extract*, kuri iš agreguotų NTL

duomenų sudaro lenteles, leidžiančias analizuoti duomenis konkrečiame regione nurodytais laiko intervalais [61].

NASA „Black Marble“ algoritmai NTL duomenis pavertė galingu įrankiu ekonomikos tendencijoms ir raidos rodikliams sekti. Išnaudodama naujų NASA „Black Marble“ „R“ ir „Python“ paketų galimybes bei specializuotos NTL analizės privalumus, Pasaulio banko duomenų laboratorija geba padėti operacijoms Gazoje, Sirijoje, Mianmare ir Libane. Šie paketai taip pat papildo tradicinius metodus bei leidžia fiksuoti ir analizuoti vietinę ekonominę veiklą, urbanizacijos tendencijas, energijos vartojimo modelius, elektros energijos tiekimo sutrikimus, gamtinių katastrofų poveikį, konfliktus ir kt. [61].



22 pav. Žemės NTL iš NASA „Black Marble“ projekto [54]

Taigi, nors tiek NASA „Black Marble“, tiek EOG/CSM teikia to paties satelito Suomi NPP nuotraukas, NASA „Black Marble“ duomenys yra tikslesni – EOG/CSM teikiami nakties vaizdai yra gerokai šviesesni žiemą, kuomet mažiau augalijos ir daugiau sniego. NASA „Black Marble“ su tuo susitvarko įvertindama, kiek šviesos ateina iš kosmoso (žvaigždžių, mėnulio) bei kiek atsispindi nuo žemės paviršiaus, taip pat aptinka ir atsižvelgia į sniegą. Dėl šių priežasčių NASA „Black Marble“ duomenys ir bus naudojami šiame tyrime.

2.7. Duomenų paruošimas

Šiame poskyryje aprašomas Lietuvos satelitinių nakties vaizdų gavimas iš NASA „Black Marble“ projekto bei jų apdorojimas: trūkstamų reikšmių interpoliavimas, vaizdų normalizavimas ir požymių inžinerija.

2.7.1. Vaizdų gavimas

Analizėje naudojami NASA „Black Marble“ satelitiniai nakties vaizdai buvo pasiekti per NASA LAADS DAAC platformą. Užsiregistravus bei prisijungus prie svetainės buvo sugeneruotas prieigos raktas (angl. *token*). Šis raktas toliau naudotas prisijungti prie NASA „Black Marble“ duomenų bazės per aplikacijų programavimo sąsają (angl. *Application Programming Interface*, API).

Duomenys buvo gauti naudojant „R“ programavimo kalbą bei šiuos „R“ paketus: *sf*, *terra*, *blackmarbler*, *raster*, *exactextractr* ir *lubridate* (6 lentelė).

6 lentelė. „R“ programos paketai, reikalingi NASA „Black Marble“ duomenims parsisiųsti

Paketo pavadinimas	Paskirtis
<i>sf</i>	Palengvina darbą su paprastosiomis funkcijomis (angl. <i>simple features</i>) palengvinti, tai taip pat yra standartizuotas būdas vektoriniams erdviniams duomenims užkoduoti. Naudojamas su <i>GDAL</i> paketu leidžia nuskaityti ir rašyti duomenis, su <i>GEOS</i> paketu leidžia atlikti geometrines operacijas, o su <i>PROJ</i> paketu leidžia konvertuoti projekcijas ir transformuoti nulines atskaitas. Pagal numatytuosius nustatymus naudoja <i>s2</i> paketą sferinės geometrijos operacijoms su elipsoidinėmis (ilg./pl.) koordinatėmis atlikti [62].
<i>terra</i>	Paketas skirtas vektorinių (taškai, linijos ir poligonai) ir rastrinių (tinklis) erdvių duomenų analizei atlikti. Metodai vektoriniams duomenims apima tokias geometrines operacijas kaip susikirtimas (angl. <i>intersect</i>) ir buferis (angl. <i>buffer</i>). Metodai rastriniams duomenims apima vietines, židinio, globalias, zonines ir geometrines operacijas. Prognozavimo ir interpoliavimo metodai palengvina regresijos tipo (interpoliavimo, mašininio mokymosi) modelių naudojimą erdviniam (įskaitant palydovinio nuotolinio stebėjimo duomenis) prognozavimui. Palaikomas labai didelių failų apdorojimas [63].
<i>blackmarbler</i>	Leidžia parsisiųsti geografinius NTL duomenis iš NASA „Black Marble“ projekto, automatizuoja vaizdų atsisiuntimo iš NASA LAADS DAAC procesą [64].
<i>raster</i>	Leidžia nuskaityti, rašyti, valdyti bei analizuoti erdvinis duomenis [65]. Leidžia išsisaugoti failus pasirinktu formatu: <i>raster</i> (. <i>grd</i>), <i>ascii</i> (. <i>asc</i>), <i>SAGA</i> (. <i>sdat</i>), <i>IDRISI</i> (. <i>rst</i>), <i>CDF</i> (. <i>nc</i>), <i>GTiff</i> (. <i>tif</i>), <i>ENVI</i> (. <i>envi</i>), <i>EHdr</i> (. <i>bil</i>) arba <i>HFA</i> (. <i>img</i>) [66].
<i>exactextractr</i>	Leidžia greitai ir tiksliai sutrumpinti rastrines reikšmes pagal poligonines teritorijas („zoning statistika“) [67].
<i>lubridate</i>	Palengvina darbą su datomis ir laiku [68].

Pirmiausia buvo parsisiųsti Lietuvos geografiniai duomenys. Tai galima atlikti keliais būdais: pasiekti duomenis per *GADMTTools* paketą (naudojant *gadm* funkciją), arba įkeliant parsisiųstą geografinių Lietuvos duomenų failą iš internetinių šaltinių. Išbandžius pirmąjį būdą, pastebėta, jog parsisiųstuose antro lygio (t. y. savivaldybių lygmens) duomenyse trūksta 12 savivaldybių: Elektrėnų, Kalvarijos, Kazlų Rūdos, Pagėgių, Prienų rajono, Rietavo, Šakių rajono, Šalčininkų rajono, Šilalės rajono, Širvintų rajono, Skuodo rajono bei Zarasų rajono. Dėl šios priežasties buvo nuspręsta parsisiųsti duomenis antruoju būdu, „shapefile“ (.*shp*) formatu. Lietuvos žemėlapis su savivaldybių ribomis buvo pasiektas per atvirą geografinių duomenų bazę „Open Street Map“ (OSM).

Apibrėžus dominantę regioną (Lietuvos ribas), buvo parsisiųsti NASA „Black Marble“ duomenys (23 pav.).

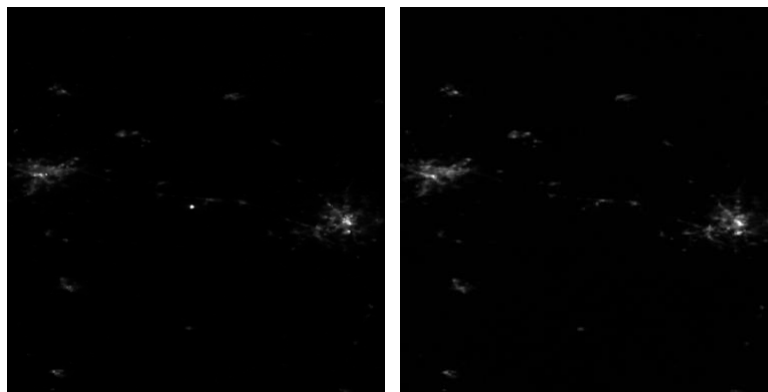
```
black_marble_data <- bm_raster(roi_sf = shapefile,
                              product_id = "VNP46A3",
                              date = seq.Date(from = ymd("2023-01-01"), to = ymd("2023-12-01"), by = "month"),
                              bearer = bearer,
                              variable = c("OffNadir_Composite_Snow_Free"),
                              quality_flag_rm = 0)
```

23 pav. „R“ kodo, leidžiančio parsisiųsti NASA „Black Marble“ duomenis, iškarpa

Tam naudota *bm_raster* funkcija bei nustatyti parametrai [60]:

1. *roi_sf* – dominančio regiono (angl. *region of interest*, ROI) poligoniniai duomenys. Privalo būti pateikiami pasaulinės koordinatės sistemos WGS84 (EPSG:4326) formatu.

2. *product_id* – produkto identifikacijos kodas. Galimi keturi variantai: neapdoroti dieniniai („VNP46A1“), koreguoti dieniniai („VNP46A2“), mėnesiniai („VNP46A3“) bei metiniai („VNP46A4“). Šiame tyrime naudoti mėnesiniai duomenys.
3. *date* – rastrinių duomenų data. Įvedus datą, sukuriamas „SpatRaster“ objektas. Įvedus kelias datas, sukuriamas „SpatRaster“ objektas su keliomis juostomis (viena juosta per datą). Pasirinktas laikotarpis – nuo 2013 iki 2023 m.
4. *bearer* – NASA LAADS DAAC prieigos raktas.
5. *variable* – kintamasis, naudojamas rastrui kurti. Jei kintamasis nenurodomas, tuomet parenkami šie numatytieji kintamieji:
 - a. „VNP46A1“ produktui – „DNB_At_Sensor_Radiance_500m“;
 - b. „VNP46A2“ produktui – „Gap_Filled_DNB_BRDF-Corrected_NTL“;
 - c. „VNP46A3“ ir „VNP46A4“ produktui – „NearNadir_Composite_Snow_Free“.
 - d. Šiuo atveju kintamasis buvo nurodytas, kaip „OffNadir_Composite_Snow_Free“. Toks kintamasis pasirinktas todėl, jog nuotraukos, atliktos „Off Nadir“ kampu buvo vizualiai ryškesnės bei neturėjo nepageidautinų defektų (24 pav.).
6. *quality_flag_rm* – pikselių kokybės žymėjimas, leidžiantis pašalinti žemos kokybės reikšmes (numatytasis yra *NULL*). Mėnesiniams („VNP46A3“) duomenims pateikiami trys pasirinkimo variantai:
 - a. 0: geros kokybės. Duomenys laikomi geros kokybės, kai stebėjimų skaičius didesnis nei 3.
 - b. 1: prastos kokybės. Duomenys laikomi prastos kokybės, kai stebėjimų skaičius mažesnis arba lygus 3.
 - c. 2: Duomenys su pagal istorinius duomenis užpildytomis NTL reikšmėmis.
 Šiuo atveju buvo pasirinkta naudoti tik geros kokybės duomenis.



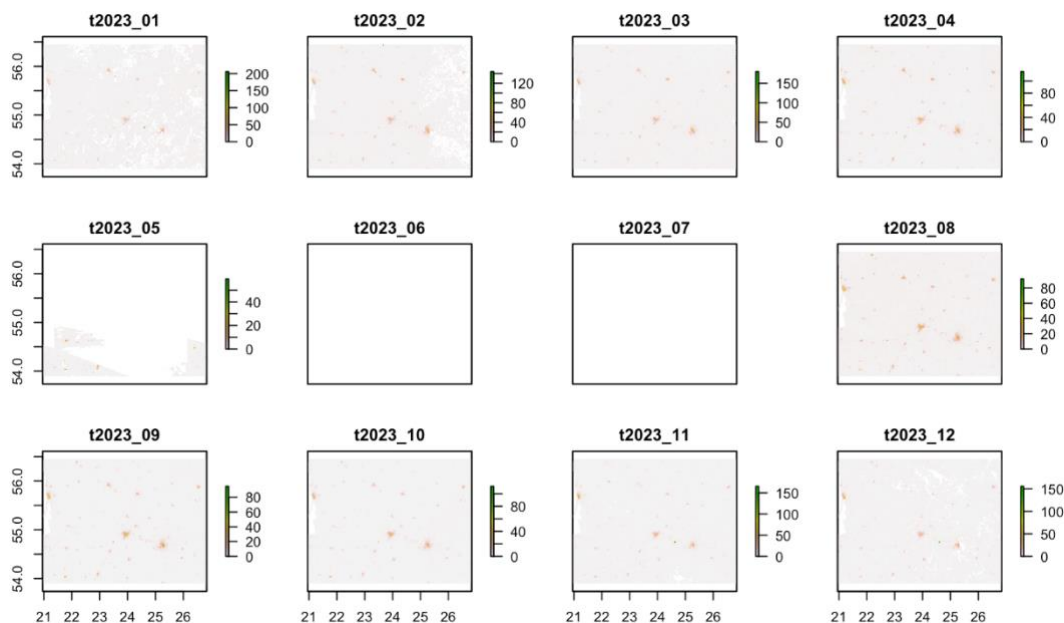
24 pav. 2023 m. sausio mėnesio Kauno ir Vilniaus miestus apimantys vaizdai. Kairėje – „NearNadir_Composite_Snow_Free“ vaizdas, dešinėje – „OffNadir_Composite_Snow_Free“

2.7.2. Vaizdų paruošimas eksperimentams

Nors „R“ *blackmarbler* paketas ir palengvina satelitinių nakties vaizdų parsisiuntimą iš NASA LAADS DAAC archyvų, parsiųsti duomenys, prieš pradedant darbą su giliojo mokymosi metodais, reikalauja papildomo paruošimo.

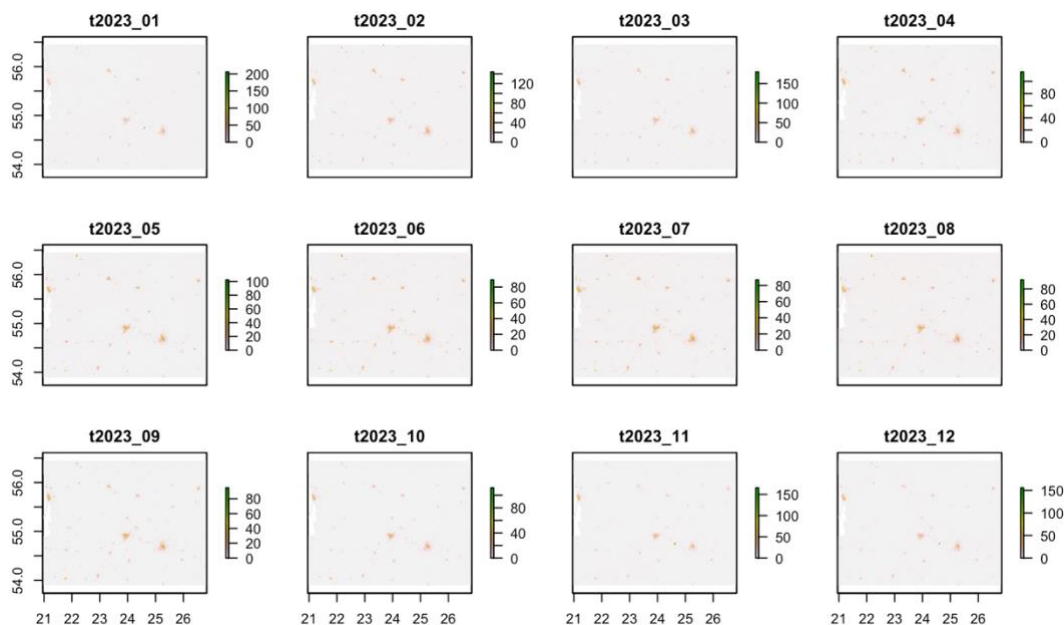
Kadangi NASA „Black Marble“ algoritmas nebando taisyti saulės šviesa užterštų pikselių, o juos „išmeta“, Lietuvos vaizduose yra matomas pikselių trūkumas kiekvienų metų dalyje gegužės mėnesio

bei visuose birželio ir liepos mėnesiuose (25 pav.). Šis trūkumas gali turėti įtakos giliojo mokymosi modelių tikslumui, todėl svarbu duomenis apdoroti – įrašyti trūkstamas reikšmes.



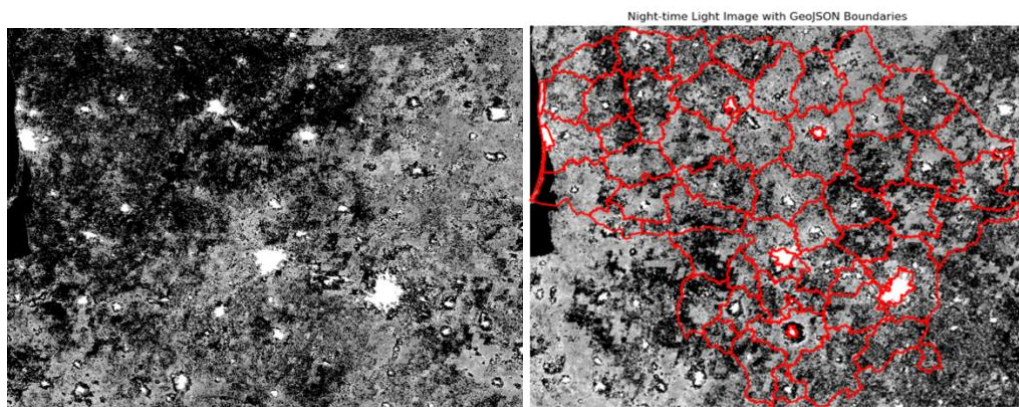
25 pav. Neapdoroti 2023 m. Lietuvos NTL vaizdai iš NASA „Black Marble“ archyvų

Reikšmių įrašymas buvo atliktas naudojant funkciją *approxNA*. Ši *raster* paketo funkcija naudoja statistinę funkciją *approx*, kuri, interpoliuodama tarp sluoksnių, nustato tuščių (N/A) laukelių reikšmes (26 pav.). Sluoksniai laikomi esantys vienodu atstumu, nebent panaudojamas nepriklausomas kintamas dydis *z* arba *getZ* funkcija grąžina reikšmes. Tokiu atveju grąžintos reikšmės yra naudojamos atstumui tarp sluoksnių nustatyti.



26 pav. 2023 m. rastriniai Lietuvos NTL vaizdai su interpoliuotomis reikšmėmis

Interpoliavus visų metų (2013–2023 m.) ir visų mėnesių (sausio–gruodžio) reikšmes, duomenys buvo išsaugoti kaip nuotraukos GeoTIFF formatu (27 pav.).

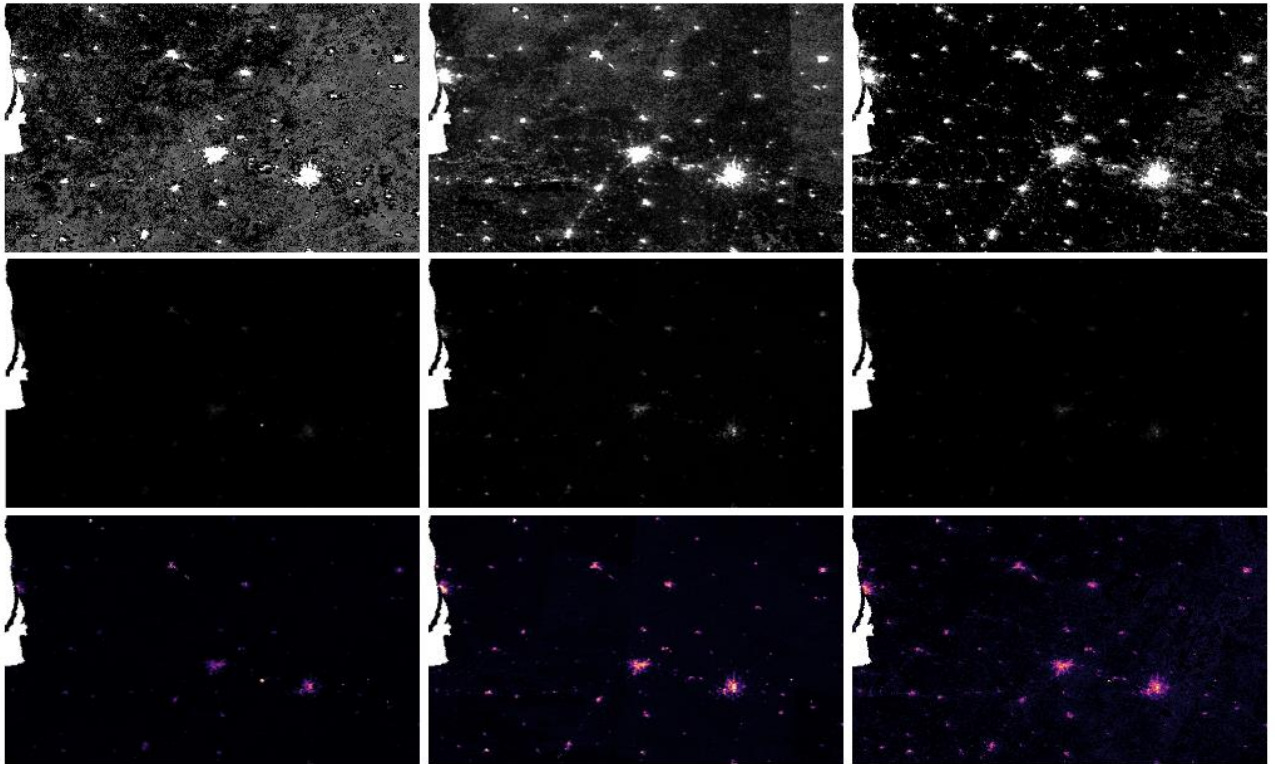


27 pav. Kairėje pusėje – 2023 m. sausio mėn. Lietuvos NTL vaizdas, išsaugotas GeoTIFF formatu. Dešinėje – vaizdas su savivaldybių ribomis

Toks saugojimo būdas leido išlaikyti geografinius metaduomenis – kiekvieno nuotraukos pikselio faktines koordinates. Visgi, kadangi didelėje dalyje nuotraukų buvo pastebėtas triukšmas (28 pav.), buvo nuspręsta ieškoti kitų būdų nuotraukoms išsaugoti.

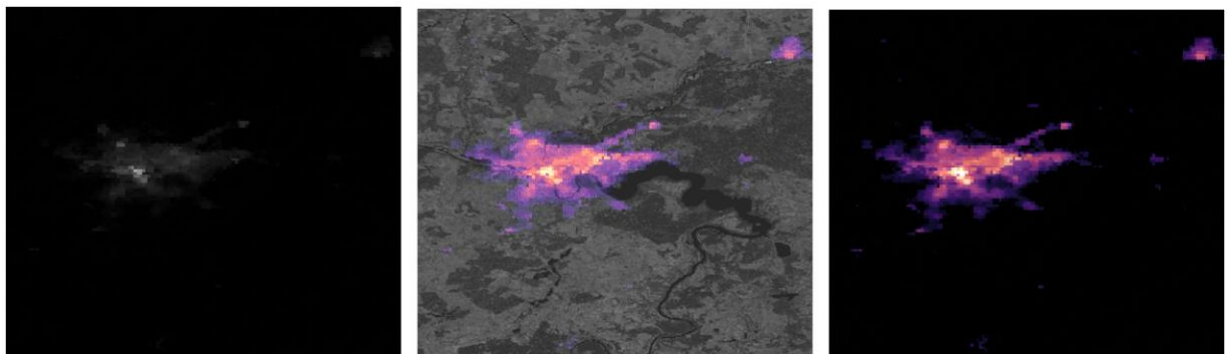
Pasirinkta alternatyva – parsisiųsti nuotraukas ne GeoTIFF, bet „ESRI Ascii“ formatu bei jas atidaryti ir apdoroti per „QGIS“ geografinės informacijos sistemos programinę įrangą. Ši programa leido pakeisti bei suvienodinti visų nuotraukų savybes. Spalvų juostos buvo subraižytos pagal paletę „Magma“, kurios spalvų rampoje varijuoja juoda, mėlyna, violetinė, raudona, oranžinė bei geltona spalvos. Kiekvienos nuotraukos pikselių reikšmės buvo suklasifikuotos minėtų spalvų skalėje. Klasifikavimas buvo individualiai pritaikyti kiekvienai nuotraukai: atsižvelgiant į spalvų rampą bei skirtingų pikselių klasių kiekį, šis metodas sugeneravo tarpinius spalvų žemėlapių įrašus pagal klasių ribas. Kiekviena spalva buvo susieta su reikšme bei interpoliuota iš verčių diapazono pagal klasifikavimo būdą. Tuomet skaliariniams duomenų rinkinio elementams (pikseliams) buvo priskiriama spalva, atsižvelgiant į jų klasę.

Šis metodas leido gauti vienodos kokybės vaizdus (28 pav.) bei panaikinti prieš tai matytą triukšmą.

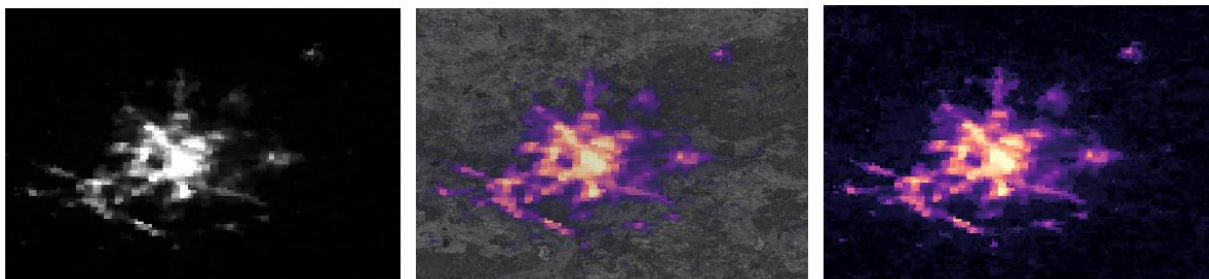


28 pav. Viršuje – GeoTIFF formatu išsaugoti Lietuvos vaizdai (2023 m. sausio, gegužės bei spalio mėn). Per vidurį – to paties laikotarpio vaizdai, išsaugoti „ESRI Ascii“ formatu bei atidaryti per „QGIS“ programą. Apačioje – to paties laikotarpio vaizdai, apdoroti naudojant „QGIS“ programą

Kadangi kiekvienam pikseliui, priklausomai nuo jo šviesumo, spalva buvo priskirta pagal „Magma“ spalvų paletę, vaizdai buvo ne tik normalizuoti, tačiau toks spalvų pritaikymas taip pat išryškino ir tamsius pikselius (29, 30 pav.). Tai atlikta siekiant pagerinti modelių tikslumą.



29 pav. 2013 m. lapkričio mėnesio Kauno nuotraukos: kairėje – neapdorotas NTL vaizdas, centre – koreguotas NTL vaizdas sujungtas su Lietuvos reljefo žemėlapiu, dešinėje – koreguotas NTL vaizdas, naudotas CNN mokymui



30 pav. 2016 m. sausio mėnesio Vilniaus nuotraukos: kairėje – neapdorotas NTL vaizdas, centre – koreguotas NTL vaizdas sujungtas su Lietuvos reljefo žemėlapiu, dešinėje – koreguotas NTL vaizdas, naudotas CNN mokymui

2.7.3. Požymių inžinerija

Požymių inžinerija yra efektyvus būdas modelių veikimui pagerinti. Pagrindiniai požymių inžinerijos komponentai yra [69]:

1. Požymių transformacija: naujų požymių sukūrimas iš egzistuojančių požymių. Dažniausiai tai pasiekti galima sudarant matematinius žemėlapius;
2. Požymių generavimas: požymių sukūrimas ne požymių transformacijos metodais. Pavyzdžiui, požymių arba pikselių gavimas iš vaizdų;
3. Požymių pasirinkimas: požymių pasirinkimas iš didelio požymių rinkinio;
4. Požymių analizė ir vertinimas: konceptai, metodai ir priemonės požymių naudai įvertinti. Šis komponentas dažnai naudojamas kartu su požymių pasirinkimu.

Šiame darbe buvo naudojamas ne tik požymių generavimas (požymių vektorių gavimas iš vaizdų), tačiau ir požymių transformacija. Buvo sudaryti papildomi modeliai su 181 nauju požymiu – prie kiekvieno mėnesinio vaizdo požymių vektorių pridėti:

- 60 stulpelių su kiekvienos savivaldybės pikselių skaičiumi vaizde, padalintu iš visų vaizde esančių pikselių skaičiaus (jei savivaldybės vaizde nėra, priskiriamas 0);
- 60 stulpelių su kiekvienos savivaldybės praėjusių metų skurdo rizikos rodikliu. Metinės reikšmės atkartojamos kiekvienam mėnesiui bei standartizuojamos skalėje nuo 0 iki 1;
- 60 stulpelių su kiekvienos savivaldybės pikselių procentinės dalies ir skurdo rodiklio sandauga;
- 1 stulpelis su visų savivaldybių pikselių procentinės dalies ir skurdo rodiklio sandaugų suma.

Visų naujai pridėtų požymių intervalas – [0; 1].

3. Tyrimo eiga ir rezultatai

Šiame skyriuje pateikiama trumpa nakties šviesų ir statistinė skurdo rizikos lygio Lietuvoje apžvalga, apžvelgiamas vaizdų rinkinių giliajam mokymuisi paruošimas: (1) vaizdų sukarpymas į vienodo dydžio gardeles: 56×56 , 112×112 , 224×224 , (2) apskaičiavimas, kiek pikselių kiekviena savivaldybė užima kiekvienoje gardelėje, (3) kiekvienos gardelės svertinio skurdo vidurkio apskaičiavimas.

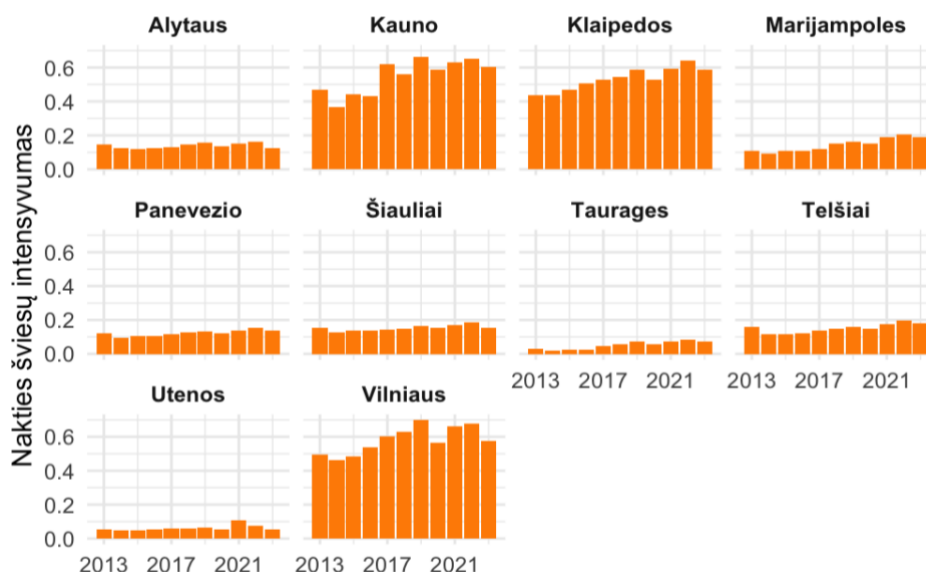
Toliau aprašoma, kaip pasirinktos CNN architektūros naudojamos vaizdams vektorizuoti. Sudaromi *Elastic Net* regresijos ir atsitiktinių miškų modeliai, įvertinamas jų veikimas bei prognozuojami mėnesiniai skurdo rodikliai 4-iems 2024 m. mėnesiams į priekį.

Galiausiai, geriausio vektorizavimo būdo bei modelio rezultatai palyginami su laiko eilučių prognozavimo metodu gautais rezultatais. Rezultatai apibendrinami.

3.1. Nakties šviesų žvalgomoji analizė

Parsisiuntus Lietuvos NTL duomenis, buvo atlikta žvalgomoji jų analizė. Didieji miestai (Vilnius, Kaunas ir Klaipėda), lyginant su kaimo regionais, rodo ir didesnę NTL intensyvumą. Tai galima paaiškinti didesniu gyventojų skaičiumi, labiau išvystyta infrastruktūra ir gausesniu dirbtinio apšvietimo šaltinių kiekiu. Taip pat pramoninėse zonose, uostuose ir miestų centruose NTL intensyvumas paprastai būna didesnis nei retai apgyvendintose kaimo vietovėse su minimaliu dirbtiniu apšvietimu. Klaipėdos uosto teritorijoje NTL gali suintensyvėti ir dėl vykstančios jūrinės veiklos.

2023 metais Vilniuje vidutinis NTL intensyvumas buvo apie 0,58, Kaune – 0,60, o Klaipėdoje – 0,59. Regionų centrai, pavyzdžiui, Telšiai (Žemaitija) ir Alytus (Dzūkija), pasižymėjo žemesniu šviesų intensyvumu – atitinkamai 0,18 ir 0,12. Tokie dideli skirtumai tarp regionų ir miestų šviesų intensyvumo išlieka pastovūs, be didelių nuokrypių (31 pav.).



31 pav. Nakties šviesų intensyvumo dinamika pagal apskritis (2013–2023 m.)

3.2. Skurdo rizikos lygio žvalgomoji analizė

Tyrimui atlikti buvo naudojami dviejų tipų duomenys: skurdo rizikos lygis (eksperimentinė statistika) iš Oficialiosios statistikos portalo (1 priedas, 32 pav.) bei satelitiniai NASA „Black Marble“ nakties vaizdai iš NASA LAADS DAAC archyvų. Naudoti duomenys yra vieši. Pasirinktas analizės laikotarpis – nuo 2013 iki 2023 metų.

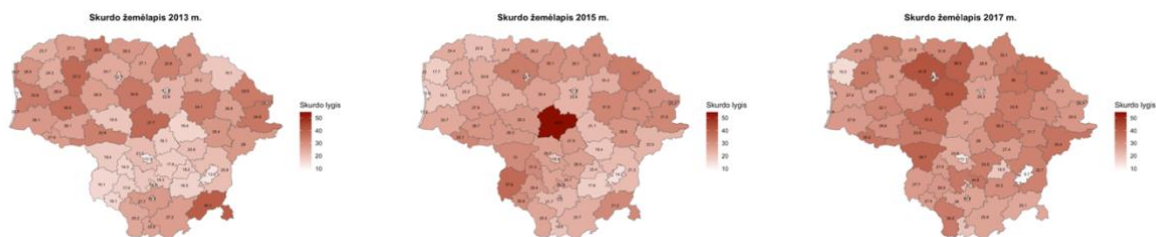
Skurdo rizikos lygis (eksperimentinė statistika)

	Skurdo rizikos lygis (eksperimentinė statistika) proc. ^{1,2}													
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Elektrėnų sav. ³	19,4	18,5	14,0	18,2	25,1	20,4	16,1	18,9	24,1	20,0	17,4	16,2	17,4	16,7
Šalčininkų r. sav.	19,4	30,2	30,7	40,1	28,0	31,2	45,6	25,1	30,8	22,5	26,0	20,1	21,0	21,8
Širvintų r. sav.	20,8	28,0	31,3	20,6	17,4	19,4	27,3	27,4	25,8	17,4	15,7	22,8	21,0	19,3
Švenčionių r. sav.	24,4	16,7	13,8	28,0	24,0	22,5	41,4	35,4	40,7	27,9	34,2	18,8	22,0	20,1
Trakų r. sav.	19,0	13,7	16,6	16,9	16,6	17,6	19,1	30,2	26,0	17,5	30,1	18,5	23,8	19,1
Ukmergės r. sav.	22,3	19,7	17,0	16,6	15,8	21,1	32,1	35,2	31,6	32,5	26,9	26,1	24,1	22,0
Vilniaus m. sav.	15,1	12,8	12,0	13,6	12,4	14,2	9,0	9,7	12,6	13,4	14,7	11,5	13,9	13,1
Vilniaus r. sav.	25,8	29,3	19,6	20,9	15,3	21,3	30,3	32,7	27,2	19,2	19,1	14,8	14,3	15,9
Alytaus m. sav.	23,9	16,4	16,7	18,3	17,4	19,0	13,9	19,9	20,7	27,3	26,4	19,7	20,8	20,3

32 pav. Skurdo rizikos lygio duomenų pagal savivaldybes 2013–2023 m. iškarpa (iš Oficialiosios statistikos portalo [18])

Lyginant skurdo rizikos lygio duomenis savivaldybių lygmenimi, pastebima didelė variacija (2 priedas). Reikšmingiausi rodiklio skirtumai matomi tarp didžiųjų miestų savivaldybių bei regionų. Žemiausias skurdo lygis 2023 m. fiksuojamas didžiuosiuose miestuose ir jų rajonuose, taip pat Neringos bei Elektrėnų savivaldybėse (svyruoja nuo 13,1 iki 18,8 %). Didžiausias – Kelmės rajono, Skuodo rajono, Pagėgių, Biržų rajono ir Jurbarko rajono savivaldybėse (28,7–31,3 %).

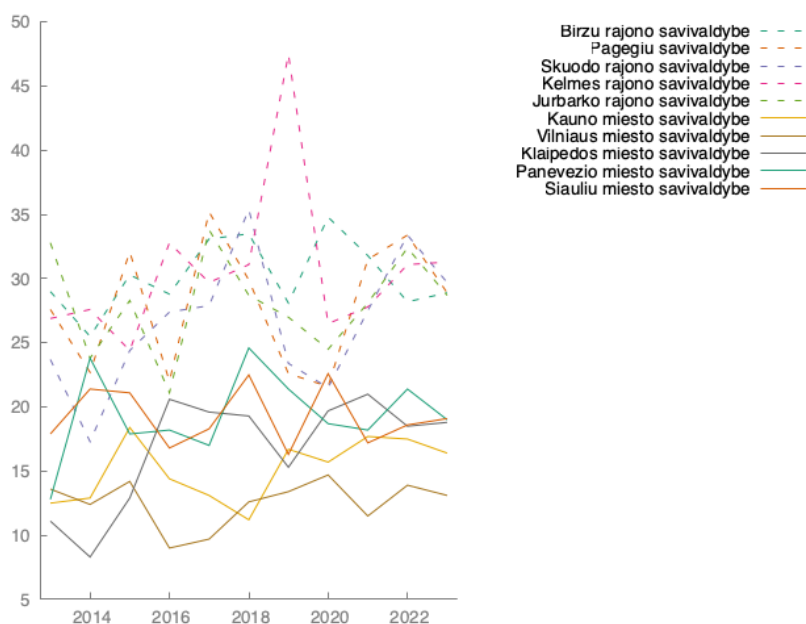
Mažiausias skurdo lygis buvo pasiektas Vilniaus miesto savivaldybėje 2016 (9 %) ir 2017 (9,7 %) metais. Didžiausias – 2015 m. Kėdainių rajone, kur skurdo rizikos lygis siekė net 54,5 %. Pastebima, jog laikui bėgant atotrūkis tarp mažiausios ir didžiausios absoliučios skurdo lygio reikšmės mažėja. Pavyzdžiui, jei 2013 m. tarpas tarp mažiausio (Klaipėdos miestas, 11,11 %) ir didžiausio (Šalčininkų rajonas, 40,1 %) skurdo rizikos lygio buvo apie 28,99 %, 2014 m. 35,7 %, o 2015 m. 41,6 %, tai 2023 m. jis sumažėjo iki 18,2 % (atitinkamai 21,2 % 2022 m. ir 20,5 % 2021 m.).





33 pav. Skurdo rizikos lygio pokyčiai Lietuvoje pagal savivaldybes 2013–2023 m.

Didžiuosiuose Lietuvos miestuose (Vilnius, Kaunas, Klaipėda, Panevėžys, Šiauliai) skurdo rodikliai per analizuojamą laikotarpį neperkopė 25 % ribos, tuo tarpu mažųjų miestų (šiuo atveju pasirinktos 5 didžiausių skurdo rizikos lygį 2023 m. turinčios savivaldybės) skurdo rizikos lygis varijavo nuo 17,3 iki net 47,5 % (34, 35 pav.).



34 pav. Skurdo rizikos lygio pokyčiai didžiųjų ir mažųjų Lietuvos miestų savivaldybėse 2013–2023 m.
Punktyrinė linija žymi mažųjų Lietuvos miestų savivaldybes, ištisinė linija – didžiųjų

Didžiųjų Lietuvos miestų savivaldybėse taip pat nepastebimas toks didelis nuokrypis nuo vidurkio (1,8–4,3), kaip mažųjų (2,8–6,1) (35 pav.). Tiesa, tokiam dideliame kintamumei įtakos gali turėti ir gyventojų skaičiaus pokyčiai – mažųjų miestų savivaldybėse vienas gyventojas turi didesnę statistinę svorį dėl mažesnės gyventojų populiacijos.

	Mean	Median	S.D.	Min	Max
Kelmesrajonosavi~	30.61	29.70	6.148	24.40	47.50
Skuodorajonosavi~	26.51	27.40	5.219	17.30	35.40
Pagegiusavivaldy~	27.95	28.90	4.988	21.60	35.20
Birzurajonosaviv~	30.18	29.00	2.804	25.50	34.80
Jurbarkorajonosa~	28.11	28.30	3.943	21.10	33.80
Vilniausmiestosa~	12.55	13.10	1.821	9.000	14.70
Kaunomiestosaviv~	15.14	15.70	2.435	11.20	18.40
Klaipėdosmiestos~	16.83	18.80	4.286	8.300	21.00
Paneveziomiestos~	19.36	18.70	3.309	12.80	24.60
šiauliumiestosav~	19.25	18.60	2.278	16.30	22.60

35 pav. Skurdo rizikos lygio statistika didžiųjų ir mažųjų Lietuvos miestų savivaldybėse 2013–2023 m.

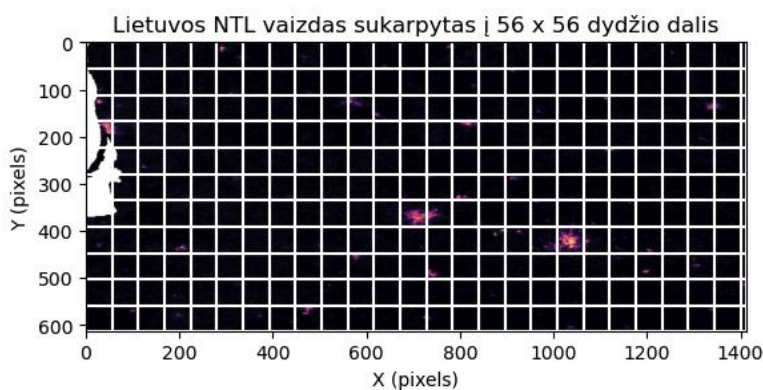
Analizuojant skurdo rizikos lygio pokyčius Lietuvoje per nurodytą laikotarpį, buvo pastebėta skurdo lygio mažėjimo tendencija daugumoje šalies savivaldybių. Lyginant su analizuojamo laikotarpio pradžia 2013–2016 m., skurdo rodiklis daugumoje Lietuvos savivaldybių yra reikšmingai kritęs.

Didžiuosiuose miestuose, tokiuose kaip Vilnius, Kaunas ir Klaipėda, skurdo rizikos lygis yra žemesnis nei mažųjų miestų savivaldybėse. Taip pat galima pastebėti didėjančią apatinę skurdo ribą bei mažėjančią apatinę – tai rodo skurdo rizikos lygio atotrūkio tarp didžiųjų ir mažųjų miestų savivaldybių mažėjimą.

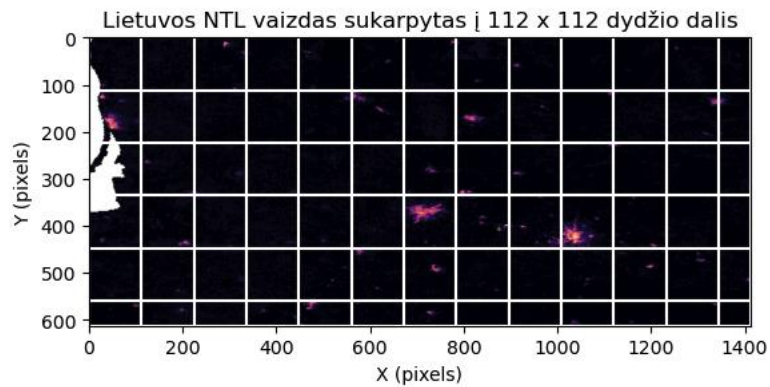
3.3. Požymių vektorių gavimas

Kadangi naudojamų architektūrų numatytasis įvesties dydis yra 224×224 , Lietuvos vaizdai buvo sukarpomi į vienodo aukščio ir pločio gardeles. Siekiant įvertinti, kaip didesnis ar mažesnis savivaldybių plotas vaizde keičia rezultatus, 132 Lietuvos NTL nuotraukas pasirinkta sukarpyti trimis būdais:

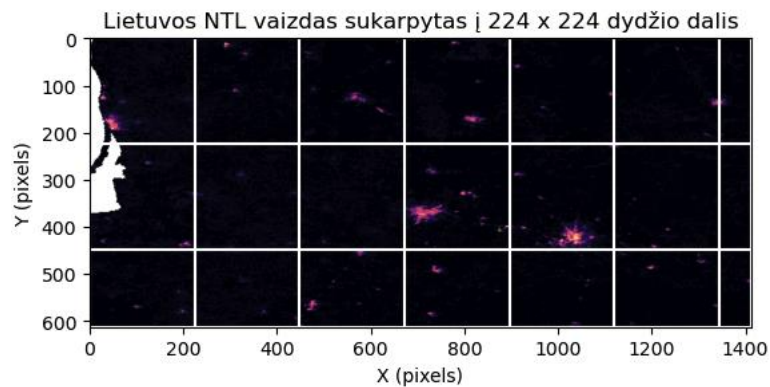
1. Į 56×56 dydžio gardeles (36 pav.). Kiekvienam mėnesiui tokių gardelių susidarė po 264, tačiau iš jų paliktos 206 – tik tos, kuriose buvo bent kiek Lietuvos teritorijos. Iš viso gautų failų skaičius – 27 192.
2. Į 112×112 dydžio gardeles (37 pav.). Kiekvienam mėnesiui tokių gardelių susidarė po 78, tačiau iš jų paliktos 59 – taip pat tik tos, kuriose buvo bent kiek Lietuvos teritorijos. Iš viso gautų failų skaičius – 7 788.
3. Į 224×224 dydžio gardeles (38 pav.). Kadangi visų nuotraukų pradinis dydis yra 1412×613 , kiekvienam mėnesiui tokių gardelių susidarė po 21, tačiau iš jų paliktos 19 – tik tos, kuriose buvo bent kiek Lietuvos teritorijos. Iš viso gautų failų skaičius – 2 508.



36 pav. 2023 m. sausio mėn. Lietuvos NTL vaizdas, sukarpytas į 56×56 dydžio gardeles



37 pav. 2023 m. sausio mėn. Lietuvos NTL vaizdas, sukarpytas į 112×112 dydžio gardeles



38 pav. 2023 m. sausio mėn. Lietuvos NTL vaizdas, sukarpytas į 224×224 dydžio gardeles

Siekiant dar labiau padidinti mokymosi imties dydį ir pagerinti modelių mokymąsi, originalios nuotraukos buvo karpomos dar kartą. Šio karpymo metu plotas persidengė – kiekviena gardelė buvo paslinkta x ir y ašimi 50 % (atitinkamai 28, 56 ir 112 pikseliais). Tokiu būdu vaizdų skaičius buvo padidintas: pirmuoju variantu iki 54 780, antruoju iki 15 840, o trečiuoju – iki 5 016. Pažymėtina, jog mažesni nei 224×224 dydžio vaizdai, atliekant vektorizavimą su skirtingomis architektūromis, buvo „ištempiami“ iki numatytosios įvesties dydžio.

Požymių vektoriams iš satelitinių nakties vaizdų gauti buvo naudojamos CNN architektūros: *ResNet-101*, *ResNet-50*, *NASNetMobile*, *EfficientNet-B0*, *DenseNet-201* bei *DenseNet-121*. 39 pav. matomi šių architektūrų parametrai – gylis (sluoksnių skaičius), įvesties dydis, paskutinio globalaus vidurkio telkimo sluoksnio pavadinimas, parametrų kiekis bei išvestis.

Architecture	Depth	Input Size (HxW)	Layer	Params	Output
NASNet	389	(224, 224)	global_average_pooling2d	5326716	1056
resnet101	209	(224, 224)	avg_pool	44707176	2048
densenet201	402	(224, 224)	avg_pool	20242984	1920
resnet50	107	(224, 224)	avg_pool	25636712	2048
efficientnetb0	132	(224, 224)	avg_pool	5330571	1280
densenet121	242	(224, 224)	avg_pool	8062504	1024

39 pav. Naudojamų CNN architektūrų parametrai

Požymių vektorizacija apėmė vaizdų transformavimą į skaitinius požymių vektorius, naudotus modelių mokyme ir testavime. Kadangi naudojami CNN modeliai parsiončiami iš *TensorFlow* bibliotekos per *Keras* giliojo mokymosi API, jų svoriai buvo iš anksto sukonfigūruoti. Šie modeliai taip pat buvo iš anksto išmokyti su *ImageNet* duomenų rinkiniais [70].

Svarbu paminėti, jog numatytoji CNN modelių įvestis yra „224, 224, 3“, kur „224, 224“ reiškia vaizdo dydį (rezoliuciją), o trejetas – RGB spalvų sistemą. Kadangi sateliniai nakties vaizdai yra spalvoti, šie įvesties parametrai yra nekeičiami, tačiau vis tiek nurodomi. Taip pat modelių telkimo sluoksniai nustatomi į vidurkio telkimą (40 pav.).

```
(input_shape=(224, 224, 3), include_top=False, weights='imagenet', pooling='avg')
```

40 pav. CNN modelių parametrų nustatymas

Tai reiškia, kad paskutinio konvoliucinio sluoksnio išvesčiai taikomas globalus vidurkio telkimas, o visų sluoksnių svoriai „užšaldomi“. Galiausiai, gauti vaizdų požymių vektoriai, kartu su nuotraukų pavadinimais, išsaugomi kaip duomenų lentelės .csv formato failuose. Kiekvieno CNN modelio .csv failo stulpelių skaičius sutampa su išvesties skaičiumi, matomu 39 pav. Modelio požymių vektorių failai yra sukuriami kiekvienam mėnesiui (iš viso – 132).

3.4. Duomenų paruošimas regresijai

Turint metinius kiekvienos savivaldybės skurdo rizikos lygio rodiklius bei satelitinį nakties vaizdų gardeles, toliau buvo apskaičiuojamas kiekvienoje vaizdo gardelėje esančių savivaldybių pikselių skaičius.

Pirmiausia buvo įkeltos Lietuvos savivaldybių ribos iš „GeoJSON“ failo, pasiekto per OSM platformą. Tuomet, naudojant transformacijos matricą, buvo apskaičiuotos kiekvieno nuotraukos pikselio ilgumos ir platumos koordinatės, vėliau priskirtos konkrečioms savivaldybėms. Šie skaičiavimai buvo atlikti visiems persidengiantiems vaizdams (41 pav.).

Image	tile_4_0.tif	tile_4_1.tif	tile_6_1.tif	tile_4_2.tif	tile_6_0.tif	tile_2_0.tif	tile_2_1.tif	tile_0_1.tif
Silutės rajono savivaldybė	0	0	0	0	0	0	0	12674
Širvintų rajono savivaldybė	0	6613	0	0	0	0	0	0
Skuodo rajono savivaldybė	0	0	0	0	0	0	0	0
Švenčionių rajono savivaldybė	0	224	861	0	0	0	0	0
Tauragės rajono savivaldybė	0	0	0	0	0	0	0	1
Telšių rajono savivaldybė	0	0	0	0	0	0	0	0
Trakų rajono savivaldybė	0	2499	0	3948	0	0	0	0
Ukmergės rajono savivaldybė	0	6785	0	0	0	0	0	0
Utenos rajono savivaldybė	1888	1687	0	0	0	0	0	0
Varėnos rajono savivaldybė	0	0	0	4154	0	0	0	0
Vilkaviškio rajono savivaldybė	0	0	0	0	0	0	4785	0
Vilniaus miesto savivaldybė	0	3164	0	61	0	0	0	0
Vilniaus rajono savivaldybė	0	11441	0	2872	0	0	0	0
Visagino savivaldybė	0	0	0	0	85	0	0	0
Zarasų rajono savivaldybė	0	0	0	0	332	0	0	0
	39314	50176	1979	21632	610	45247	49848	15230

41 pav. Savivaldybių pikselių skaičius sukarpytuose vaizduose (pateikiama lentelės iškarpa, sudaryta iš 224 × 224 vaizdų duomenų)

Turint pikselių skaičių kiekvienoje gardelėje, buvo apskaičiuotas kiekvieno vaizdo metinis svertinis skurdo vidurkis – savivaldybės pikselių skaičių vaizde padalinamas iš visų savivaldybių pikselių vaizde sumos.

Papildomai apskaičiuoti visų savivaldybių pikselių skaičių vaizde buvo reikalinga todėl, jog kai kurie vaizdai apėmė ne tik Lietuvos teritoriją, tačiau ir užsienio – taip į svertinius skurdo rodiklius buvo įtraukti tik aktualūs, Lietuvos ribas žymintys pikseliai.

Savivaldybės skurdo lygio rodikliai kiekviename vaizde (stulpeliai) buvo susumuoti. Tokiu būdu nustatyta, koks skurdo rizikos lygis patenka į kiekvieną iškirptą gardelę (42 pav.).

Image	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
tile_4_0.tif	26,70	29,35	29,95	28,78	34,45	32,25	27,72	30,78	29,09	27,34	26,64
tile_4_1.tif	22,51	21,38	23,25	27,10	29,99	29,18	23,88	23,25	22,60	21,30	20,42
tile_6_1.tif	31,90	25,75	27,58	32,76	31,50	36,46	27,05	32,73	24,22	25,67	23,83
tile_4_2.tif	30,76	24,57	25,92	35,48	27,15	30,43	22,74	25,21	19,04	21,06	20,85
tile_6_0.tif	33,21	27,61	30,07	26,49	27,72	35,44	28,92	33,07	30,71	29,24	27,05
tile_2_0.tif	28,55	25,32	28,55	38,30	36,21	29,92	31,77	24,46	25,39	26,90	25,94
tile_2_1.tif	21,86	20,41	31,42	31,85	33,06	31,26	29,63	25,92	26,62	28,02	25,22
tile_0_1.tif	28,13	24,58	24,84	27,13	28,31	22,40	18,36	17,30	27,25	26,38	25,61
tile_0_0.tif	28,17	20,42	20,73	18,18	25,24	26,37	18,91	18,93	22,00	23,39	22,45
tile_2_2.tif	20,84	24,66	28,55	31,36	30,19	36,72	30,45	26,93	24,42	28,04	26,29
tile_5_2.tif	35,15	24,73	28,65	41,66	27,06	29,87	21,65	24,22	18,73	19,27	20,28
tile_5_0.tif	28,92	29,43	30,92	28,30	30,35	35,45	29,33	33,56	29,39	28,68	26,35
tile_5_1.tif	29,43	25,18	26,53	32,79	32,24	35,71	26,54	31,11	23,44	23,92	22,31
tile_1_1.tif	28,45	22,91	29,08	24,66	32,21	29,92	24,43	22,99	27,65	29,28	26,98
tile_3_2.tif	25,59	27,59	22,60	27,89	26,58	33,78	27,66	23,75	19,95	24,14	23,48
tile_1_0.tif	30,90	22,80	24,12	25,93	30,60	29,88	25,13	21,17	26,72	27,15	25,70
tile_1_2.tif	16,10	18,50	37,60	32,40	27,70	37,10	28,60	35,00	31,60	35,10	28,10
tile_3_0.tif	27,97	25,31	27,01	30,18	31,49	27,04	25,81	28,04	25,59	26,40	24,61
tile_3_1.tif	22,63	21,38	31,01	29,43	28,96	23,53	22,10	23,65	21,22	23,89	21,71
tile_shifted_4_0.tif	30,00	26,67	29,82	28,17	31,84	30,35	26,02	34,79	30,59	28,48	27,89
tile_shifted_6_2.tif	28,00	24,00	22,50	41,40	35,40	40,70	27,90	34,20	18,80	22,00	20,10
tile_shifted_4_1.tif	24,33	24,77	26,02	27,68	31,29	28,29	25,85	26,43	26,03	25,17	23,76
tile_shifted_6_1.tif	32,99	26,68	29,05	29,58	29,65	35,98	27,87	33,05	27,42	27,67	25,59
tile_shifted_4_2.tif	22,61	21,82	22,81	26,86	27,99	27,58	22,89	23,29	19,42	22,27	20,79
tile_shifted_2_0.tif	30,57	24,82	26,56	32,87	33,03	27,83	27,09	23,38	25,74	25,34	24,77
tile_shifted_2_1.tif	27,86	24,41	27,59	28,70	33,29	29,11	31,23	23,89	26,71	28,81	27,57
tile_shifted_0_1.tif	29,09	21,74	20,26	19,65	24,52	20,78	17,53	17,35	22,15	22,01	21,63
tile_shifted_2_2.tif	17,07	19,37	33,84	36,69	32,02	37,25	31,95	31,45	29,59	31,61	26,73
tile_shifted_0_0.tif	27,83	23,54	19,60	19,86	18,37	25,86	15,00	18,11	23,72	24,60	23,40
tile_shifted_5_2.tif	26,06	19,55	23,34	33,55	28,89	28,57	20,72	22,50	16,59	17,16	17,81
tile_shifted_5_0.tif	19,37	33,67	33,61	33,08	36,12	33,99	29,85	33,92	25,19	27,42	25,88
tile_shifted_5_1.tif	27,64	29,45	29,47	29,55	32,54	34,60	28,43	30,84	28,01	26,37	24,63
tile_shifted_1_1.tif	29,83	22,46	25,09	21,72	29,58	26,83	19,71	19,48	25,26	25,77	24,51
tile_shifted_1_0.tif	26,77	19,73	22,63	24,16	29,40	33,29	22,77	20,09	26,91	28,12	25,77
tile_shifted_3_2.tif	21,36	21,97	25,99	29,58	28,77	30,96	25,96	24,06	22,30	24,19	23,03
tile_shifted_3_0.tif	27,94	23,84	29,59	38,44	35,50	30,71	31,50	26,39	24,66	26,68	24,90
tile_shifted_1_2.tif	27,60	22,70	32,10	22,00	35,20	29,90	22,60	21,60	31,50	33,40	28,90
tile_shifted_3_1.tif	29,71	23,62	33,50	34,95	33,23	27,92	26,15	23,99	24,19	27,18	24,78

42 pav. Svertiniai skurdo rizikos lygio rodikliai kiekviename vaizde (pateikiama lentelės iškarpos pavyzdys, sudarytas iš 224×224 vaizdų duomenų)

Galiausiai, prieš sudarant modelius, duomenys buvo padalinti į mokymo (2013–2022 m.) ir testavimo (2023 m.) rinkinius (7 lentelė). Mokymo rinkinys (X_{train} , Y_{train}) buvo naudojamas modeliui mokytis, o testavimo rinkinys (X_{test} , Y_{test}) – modeliui vertinti.

7 lentelė. Mokymo ir testavimo imčių parametrai

Vaizdų pobūdis	Vaizdų dydis	Mokymo imtis (X_{train})	Testavimo imtis (X_{test})
Nepersidengiantys	56×56	24 720	2 472
	112×112	7 080	708
	224×224	2 280	228
Persidengiantys	56×56	49 800	4 980
	112×112	14 400	1 440
	224×224	4 560	456

3.5. Rezultatai

Šiame skyriuje pateikiami eksperimento rezultatai: (1) *Elastic Net* ir (2) atsitiktinių miškų modelių, sudarytų naudojant skirtingas CNN architektūras rezultatai ir jų įvertinimas. Gauti rezultatai aptariami dvejopai: naudojant vien tik vaizdų požymių vektorius bei į modelius įtraukiant ir papildomą 181 požymį.

3.5.1. Elastic Net modelių rezultatai

Atlikus duomenų paruošimą buvo inicijuota *Elastic Net CV* regresija, kuri pati išrenka geriausią alfa reikšmę, o parametrų derinimui naudoja 5 dalių kryžminę validaciją. Tokiu būdu buvo gaunamos mėnesinės kiekvienos architektūros prognozės kiekvienai gardelei. Šie rezultatai buvo išsaugojami .csv formato failuose, kuriuose buvo atliekamas gardelių lygio prognozių konvertavimas į savivaldybių reikšmes. Tai buvo atliekama sudauginant kiekvienos savivaldybės pikselių skaičių su svertiniu atitinkamo vaizdo skurdo rizikos lygio rodikliu bei padalinant gautą skaičių iš savivaldybės pikselių skaičiaus sumos. Taip gautos mėnesinės 2023 m. skurdo prognozės kiekvienai savivaldybei. Metinės savivaldybių prognozės apskaičiuotos, kaip visų mėnesių prognozuotų skurdo rodiklių vidurkis. Skirtingų CNN architektūrų rezultatus mėnesinėms bei metinėms prognozėms pateikiami 3 priede.

Geriausieji rezultatai, prognozuojant mėnesinius skurdo rizikos lygio rodiklius, buvo gauti naudojant *DenseNet-201* vektorizavimo metodą, *Elastic Net CV* modelį mokant su mažiausiais 56×56 dydžio vaizdais, o testavimą atliekant be persidengiančių vaizdų (8 lentelė). Šio modelio optimali alfa reikšmė – 0,01, mokymo determinacijos koeficientas R^2 – 0,136, testavimo determinacijos koeficientas R^2 – -0,145, MDAE – 2,622, o MAPE – 14,461. Šis rezultatas reiškia, jog prognozių klydimo mediana yra 2,62 vieneto, o vidutinė klaida – 14,461 %.

Pastebima, jog kai naudojami didesnio dydžio įvesties vaizdai (224×224), geresni rezultatai gaunami testuojant su persidengiančiais vaizdais (t. y., padidinant testavimo imtį). Tuo tarpu, kai naudojami mažesnio dydžio įvesties vaizdai (56×56), geresni rezultatai gaunami testuojant be persidengiančių vaizdų. Tai galima paaiškinti tuo, kad naudojant į mažesnes gardeles sukarpytas nuotraukas, testavimo imtis yra pakankamai didelė (2 472). Tokiais atvejais papildomi vaizdai rezultatus pablogina.

8 lentelė. *Elastic Net* modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais, metinių spėjimų rezultatai (testavimo imtis – 2 472 56×56 dydžio vaizdai)

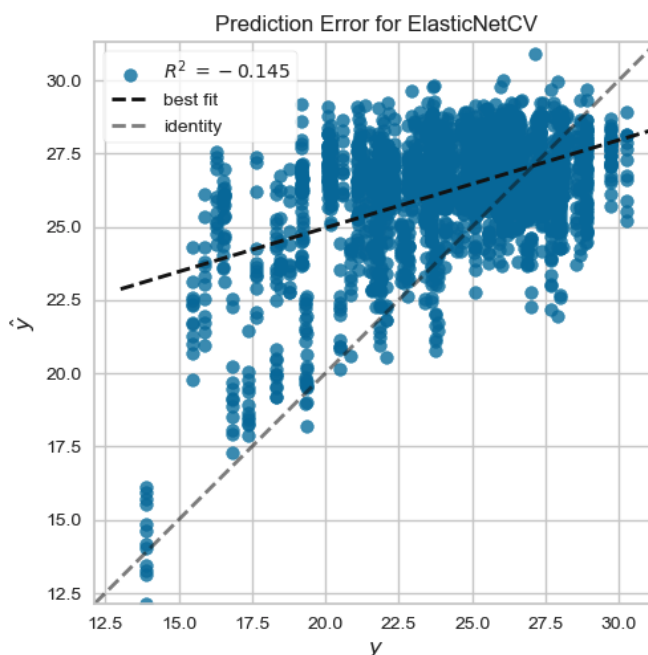
CNN architektūra	Optimali alfa reikšmė	Mokymo R^2	Testavimo R^2	RMSE	MDAE	MAPE	SMAPE	MAAPE
<i>ResNet-50</i>	0,1	0,099	-0,186	3,880	2,740	15,328	13,665	14,878
<i>ResNet-101</i>	0,1	0,097	-0,230	3,964	2,637	15,767	14,007	15,292
<i>NASNetMobile</i>	0,01	0,110	-0,250	4,033	2,738	16,072	14,228	15,560
<i>EfficientNet-B0</i>	0,01	0,100	-0,231	4,010	2,616	15,861	14,049	15,364
<i>DenseNet-121</i>	0,1	0,087	-0,151	3,981	2,709	15,743	13,948	15,227
<i>DenseNet-201</i>	0,01	0,136	-0,145	3,714	2,622	14,461	12,997	14,091

Geriausieji rezultatai, prognozuojant metinius skurdo rizikos lygio rodiklius, buvo gauti su taip pat naudojant *DenseNet-201* vektorizavimo metodą, *Elastic Net CV* modelį mokant su mažiausiais 56×56 dydžio vaizdais, o testavimą atliekant be persidengiančių vaizdų. Pastebima, jog geresni rezultatai gauti prognozuojant metinius, o ne mėnesinius skurdo rizikos lygio rodiklius. Šiuo atveju optimali alfa reikšmė – 0,01, mokymo R^2 – 0,136, testavimo R^2 – -0,145, MDAE – 2,335, o MAPE – 14,224 (9 lentelė).

9 lentelė. *Elastic Net* modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais, metinių spėjimų rezultatai (testavimo imtis – $2\,472\,56 \times 56$ dydžio vaizdai)

CNN architektūra	Optimali alfa reikšmė	Mokymo R^2	Testavimo R^2	RMSE	MDAE	MAPE	SMAPE	MAAPE
<i>ResNet-50</i>	0,1	0,099	-0,186	3,824	2,570	15,188	13,547	14,751
<i>ResNet-101</i>	0,1	0,097	-0,230	3,938	2,680	15,739	13,991	15,271
<i>NASNetMobile</i>	0,01	0,110	-0,250	3,981	2,585	15,974	14,152	15,476
<i>EfficientNet-B0</i>	0,01	0,100	-0,231	3,971	2,470	15,815	14,026	15,332
<i>DenseNet-121</i>	0,1	0,087	-0,151	3,930	2,630	15,638	13,866	15,138
<i>DenseNet-201</i>	0,01	0,136	-0,145	3,623	2,335	14,224	12,799	13,877

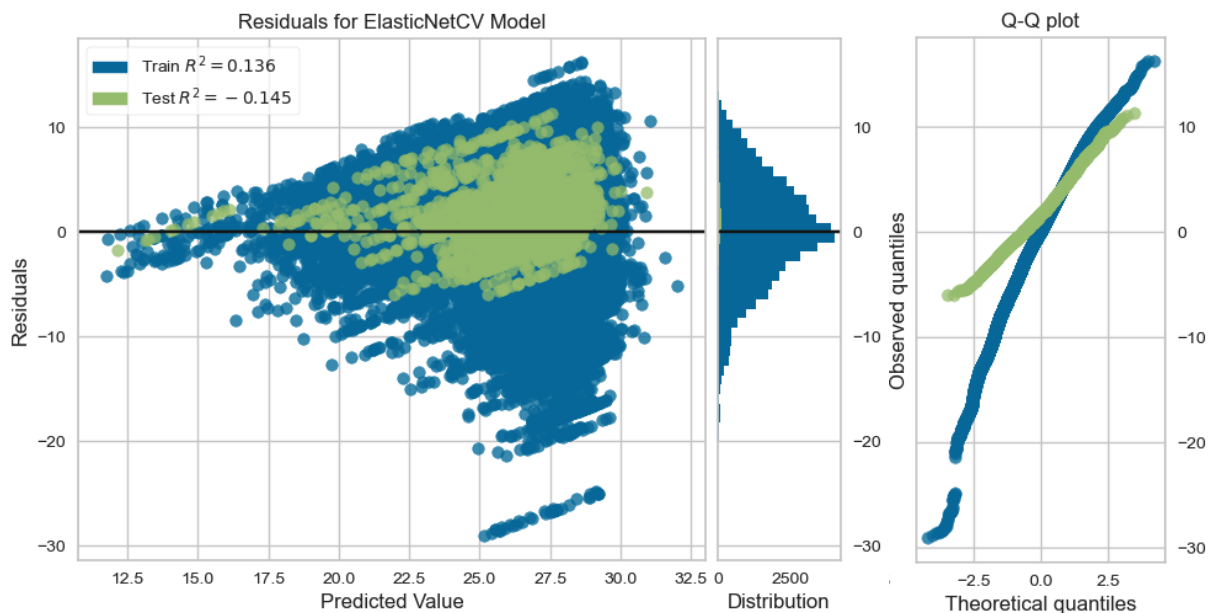
Skirtingų CNN rezultatai, gauti su 56×56 dydžio nuotraukomis, tarpusavyje skiriasi nedaug (lyginant metines ar mėnesines prognozes, testuojant su persidengiančiais ar nepersidengiančiais vaizdais). Visgi, didžiausi skirtumai pastebimi šiuos skaičiavimus kartojant su skirtingo dydžio vaizdais: prasčiausi rezultatai gaunami su 224×224 , geriausi – su 56×56 . Tai galima paaiškinti tuo, jog atliekant dideliame 224×224 vaizde esančio skurdo rodiklio konvertavimą į savivaldybės lygmenį, prognozės suprastėja. Kuo daugiau skirtingų savivaldybių ploto telpa viename vaizde, tuo rezultatai prastesni. Dalinant vaizdus į mažesnes gardeles, rezultatai gerėja.



43 pav. *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, liekanų grafikas

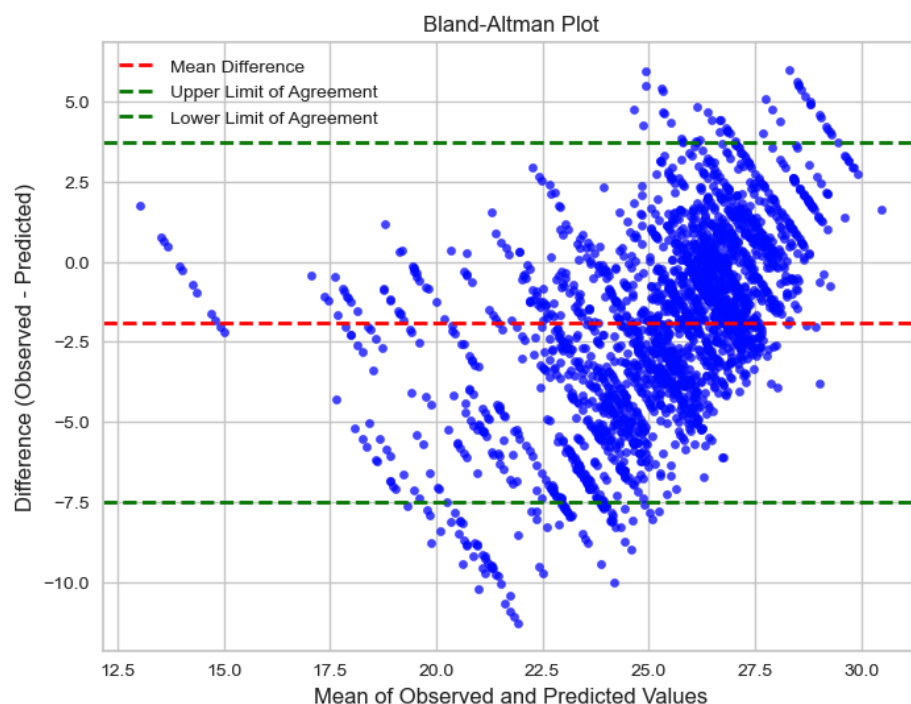
Elastic Net modelio, sudaryto naudojant *DenseNet-201* vektorizavimo būdu gautus požymius, liekanų grafike (43 pav.) galima pastebėti, jog liekanos yra nėra pasiskirsčiusios tolygiai, ar atsitiktinai – daugiausia jų yra susikoncentravusios viename plote. Šis klasterizavimas gali reikšti, kad modelis nepakankamai gerai atspindi nepriklausomų ir priklausomų kintamųjų ryšį. Taip pat gali reikšti, jog modelis gali turėti nepageidautinų trendų (prognozės yra linkusios sistemingai nukrypti nuo stebimų

reikšmių). Modelis pervertina, ar nuvertina priklausomą kintamąjį pastoviu prognozuojamų reikšmių skaičiumi. Kita vertus, tokie rezultatai taip pat gali indikuoti, jog netiesiniai modeliai šioms duomenims galėtų būti tinkamesni.



44 pav. *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, liekanų pasiskirstymo bei Q-Q grafikai

Elastic Net modelio, sudaryto naudojant *DenseNet-201* vektorizavimo būdu gautus požymius, liekanų pasiskirstymo grafike (44 pav.) matoma, jog testavimo imties liekanos išsidėsčiusios -10–10 intervale. Q-Q grafikas, rodantis ar liekanos yra pasiskirsčiusios normaliai, ar turi kitokį pasiskirstymą, indikuoja, jog testavimo liekanos pasiskirsčiusios normaliuoju skirstiniu, tuo tarpu testavimo liekanų kvantilės turi nedidelę dešiniąją asimetriją.



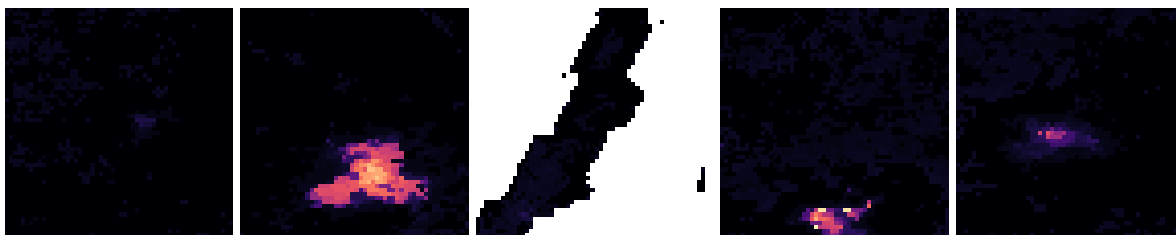
45 pav. *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, Bland-Altman'o grafikas

Bland-Altman'o grafikas rodo (45 pav.), jog dalis prognozuojamų reikšmių yra didesnės nei faktinės, ypač kai faktinės reikšmės yra mažos – modelis rezultatus pervertina. Tai taip pat matoma ir prognozuotuose rezultatuose. Pastebėta, jog su didesnių (112×112 ir 224×224) dydžių vaizdais sudaryti modeliai prasčiau spėjo didžiųjų miestų savivaldybių skurdo rodiklius. Tai galima paaiškinti tuo, jog didelėse gardelėse sutelpa daugiau savivaldybių. Tai ypač apsunkina skurdo prognozavimą tokiose savivaldybėse, kurios yra mažos bei turi žemus skurdo rodiklius, tačiau jas supa didelės aukštu skurdo rizikos lygiu pasižyminčios savivaldybės. Pavyzdžiui, svertinis skurdo rodiklis gardelėje, kurioje sutelpa Vilniaus miesto ir Šalčininkų savivaldybės, išsikreipia. Skurdo rodikliai visu analizuojamu laikotarpiu reikšmingai skyrėsi – 2016 m. Šalčininkų savivaldybėje skurdas siekė net virš 45 %, tuo tarpu Vilniaus miesto savivaldybėje jis buvo vos 9 %. Tokiais atvejais skurdo lygis Vilniuje prognozuojamas didesnis, o Šalčininkuose – mažesnis. Abejais atvejais prognozių tikslumas mažėja. Panaši situacija pastebima ir su kitais regionais, pavyzdžiui, Neringos ir Šilutės rajono savivaldybėmis. Tokie (prasčiausi) rezultatai, modelį mokant su dideliais (224×224) vaizdais, matomi 4 priede. Tiesa, ateityje mažėjant skurdo lygio atotrūkiui tarp savivaldybių, galima tikėtis geresnio modelių veikimo.

10 lentelė. *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, spėjimų rezultatai (10 prasčiausių prognozių)

Savivaldybė	Skurdo rodiklis (%)	Prognozė (%)	$ \Delta $
Švenčionių rajono savivaldybė	20,1	28,37	8,27
Šiaulių miesto savivaldybė	19,1	26,47	7,37
Neringos savivaldybė	15,5	21,92	6,42
Druskininkų savivaldybė	20,1	26,45	6,35
Utenos rajono savivaldybė	21,4	27,32	5,92
Vilniaus rajono savivaldybė	15,9	21,68	5,78
Kauno rajono savivaldybė	17,8	23,19	5,39
Šiaulių rajono savivaldybė	22,7	27,87	5,17
Skuodo rajono savivaldybė	29,7	24,86	4,84
Pakruojo rajono savivaldybė	24,0	28,79	4,79

Kaip minėta, pastebima, jog net ir prasčiausios prognozės yra geresnės, prognozuojant su mažesniais vaizdais (10 lentelė) nei su didesniais (4 priedas). Čia prasčiausiai prognozuojami Švenčionių rajono, Šiaulių miesto, Neringos, Druskininkų bei Utenos rajono savivaldybių skurdo rizikos lygio rodikliai, tačiau jie vis tiek geresni už gautus prognozuojant su 224×224 dydžio vaizdais. Prastesnes prognozes taip pat galima paaiškinti ir tuo, jog šios trys savivaldybės yra pakankamai tamsios (46 pav.), tačiau jų skurdo rodikliai ne itin aukšti. Galima daryti prielaidą, jog tokiais atvejais NTL skurdo lygį aiškina prasčiau.



46 pav. Iš kairės į dešinę – Švenčionių rajono, Šiaulių miesto, Neringos, Druskininkų bei Utenos rajono miesto savivaldybių (arba didžiausių jų dalių) gardelės. 56×56 dydžio vaizdų pavyzdys

Rezultatai, artimiausi tikrajam skurdo rodikliui, buvo gauti prognozuojant Prienų rajono, Kretingos rajono, Biržų rajono, Alytaus rajono bei Raseinių rajono savivaldybių skurdo rizikos lygio rodiklius (11 lentelė).

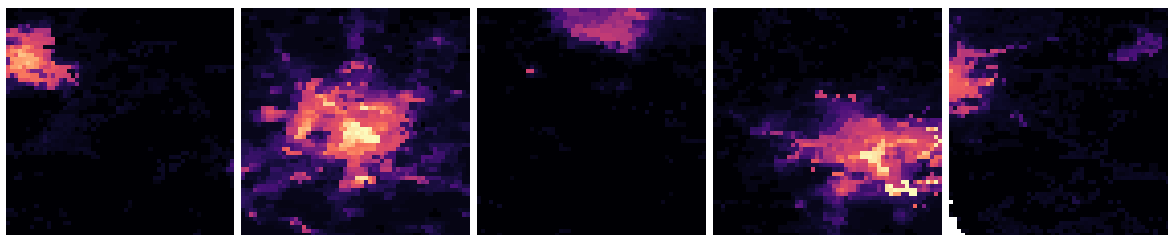
11 lentelė. *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, spėjimų rezultatai (10 geriausių prognozių)

Savivaldybė	Skurdo rodiklis (%)	Prognozė (%)	$ \Delta $
Prienų rajono savivaldybė	23,9	23,83	0,07
Kretingos rajono savivaldybė	21,6	21,46	0,14
Biržų rajono savivaldybė	28,9	29,11	0,21
Alytaus rajono savivaldybė	25,0	25,24	0,24
Raseinių rajono savivaldybė	26,3	26,56	0,26
Mažeikių rajono savivaldybė	25,1	25,77	0,67
Kupiškio rajono savivaldybė	26,6	27,28	0,68
Palangos miesto savivaldybė	21,9	21,08	0,82
Telšių rajono savivaldybė	25,0	25,86	0,86
Lazdijų rajono savivaldybė	27,3	28,23	0,93

Didžiųjų miestų skurdo rodiklių prognozės matomos 12 lentelėje. Pastebima, jog Klaipėdos miesto savivaldybės skurdo rodiklis prognozuojamas su mažiausia paklaida, o Vilniaus miesto ir Šiaulių miesto savivaldybių rodikliai prognozuojami prasčiau. Nors šių miestų NTL vaizdai buvo sąlyginai šviesūs (47 pav.), rezultatai nėra reikšmingai geresni už kitoms savivaldybėms gautas prognozes. Visų savivaldybių prognozes galima pamatyti 5 priede.

12 lentelė. *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, spėjimų rezultatai (didžiųjų miestų prognozės)

Savivaldybė	Skurdo rodiklis (%)	Prognozė (%)	$ \Delta $
Šiaulių miesto savivaldybė	19,1	26,47	7,37
Vilniaus miesto savivaldybė	13,1	16,99	3,89
Panevėžio miesto savivaldybė	19,0	22,55	3,55
Kauno miesto savivaldybė	16,4	19,60	3,20
Klaipėdos miesto savivaldybė	18,8	19,93	1,13



47 pav. Iš kairės į dešinę – Šiaulių miesto, Vilniaus miesto, Panevėžio miesto, Kauno miesto bei Klaipėdos miesto savivaldybių (arba didžiausių jų dalių) gardelės. 56×56 dydžio vaizdų pavyzdys

Elastic Net modelių rezultatai su papildomais požymiais

Geriausi rezultatai, prognozuojant mėnesinius skurdo rizikos lygio rodiklius, naudojant *Elastic Net* regresiją bei požymių žemėlapių imtis papildant naujai įtrauktais 181 požymiais, buvo gauti naudojant *ResNet-50* architektūrą. Buvo pastebėta, kad rezultatai pagerėjo iš regresijos išimant persidengiančius vaizdus ne tik iš testavimo, tačiau ir mokymo imties (13 lentelė, 6 priedas). Tai galima aiškinti tuo, jog pridėjus informaciją apie savivaldybių dydžius, vietą bei kiekvienos savivaldybės praėjusių metų skurdo rodiklius, papildomi vaizdai tapo nebereikalingi. Kita vertus, sudarant modelius (aprašytus skyriuose 3.5.1. ir 3.5.2.) be papildomų 181 požymių, persidengiantys vaizdai mokymo imtyje buvo naudingi.

13 lentelė. *Elastic Net* modelių, mokyto su skirtingais vektorizavimo būdais gautais požymiais bei papildomais X, mėnesinių spėjimų rezultatai (mokymo ir testavimo imtys be persidengiančių vaizdų)

CNN architektūra	L1 santykis	Optimali alfa reikšmė	Mokymo R^2	Testavimo R^2	RMSE	MDAE	MAPE	SMAPE	MAAPE
ResNet-50	0,7	0,0001	0,469	0,510	2,675	1,877	9,962	9,321	9,832
ResNet-101	0,5	0,0001	0,478	0,476	2,718	1,859	10,127	9,465	9,993
NASNetMobile	0,5	0,0001	0,471	0,466	2,814	1,978	10,629	9,878	10,479
EfficientNet-B0	0,5	0,0001	0,466	0,455	2,859	1,897	10,825	10,033	10,664
DenseNet-121	0,5	0,0001	0,473	0,480	2,738	1,919	10,229	9,568	10,092
DenseNet-201	0,7	0,0001	0,468	0,503	2,722	1,939	10,267	9,588	10,130

Geriausi rezultatai, prognozuojant metinius skurdo rizikos lygio rodiklius, buvo gauti taip pat be persidengiančių vaizdų tiek mokymo, tiek testavimo imtyse, ir naudojant *ResNet-50* vektorizavimo metodą (14 lentelė). Šiuo atveju optimali alfa reikšmė – 0,0001, mokymo R^2 – 0,469, testavimo R^2 – 0,510, MDAE – 1,960, o MAPE – 9,811 (16 lentelė).

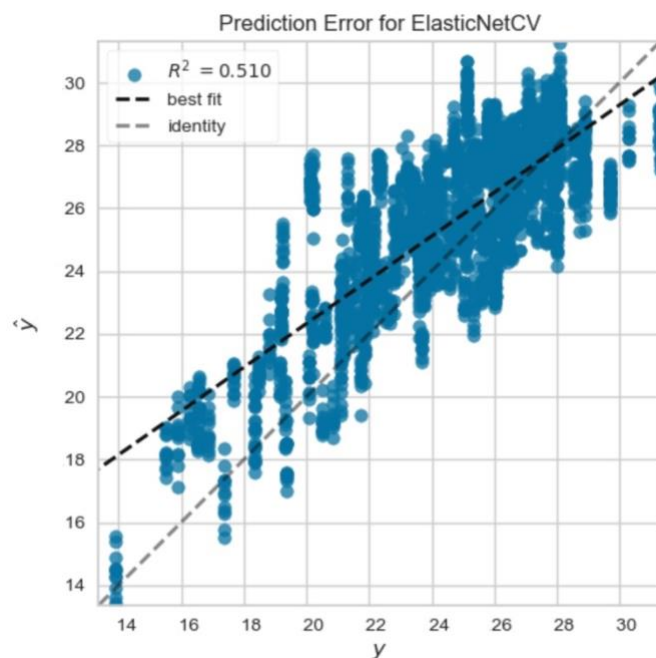
14 lentelė. *Elastic Net* modelių, mokyto su skirtingais vektorizavimo būdais gautais požymiais bei papildomais X, metinių spėjimų rezultatai (mokymo ir testavimo imtys be persidengiančių vaizdų)

CNN architektūra	L1 santykis	Optimali alfa reikšmė	Mokymo R^2	Testavimo R^2	RMSE	MDAE	MAPE	SMAPE	MAAPE
ResNet-50	0,7	0,0001	0,469	0,510	2,636	1,960	9,811	9,179	9,685
ResNet-101	0,5	0,0001	0,478	0,476	2,688	1,855	10,045	9,390	9,915
NASNetMobile	0,5	0,0001	0,471	0,466	2,795	1,915	10,568	9,823	10,421

EfficientNet-B0	0,5	0,0001	0,466	0,455	2,849	1,885	10,780	9,991	10,621
DenseNet-121	0,5	0,0001	0,473	0,480	2,706	1,940	10,083	9,431	9,951
DenseNet-201	0,7	0,0001	0,468	0,503	2,692	1,945	10,136	9,464	10,003

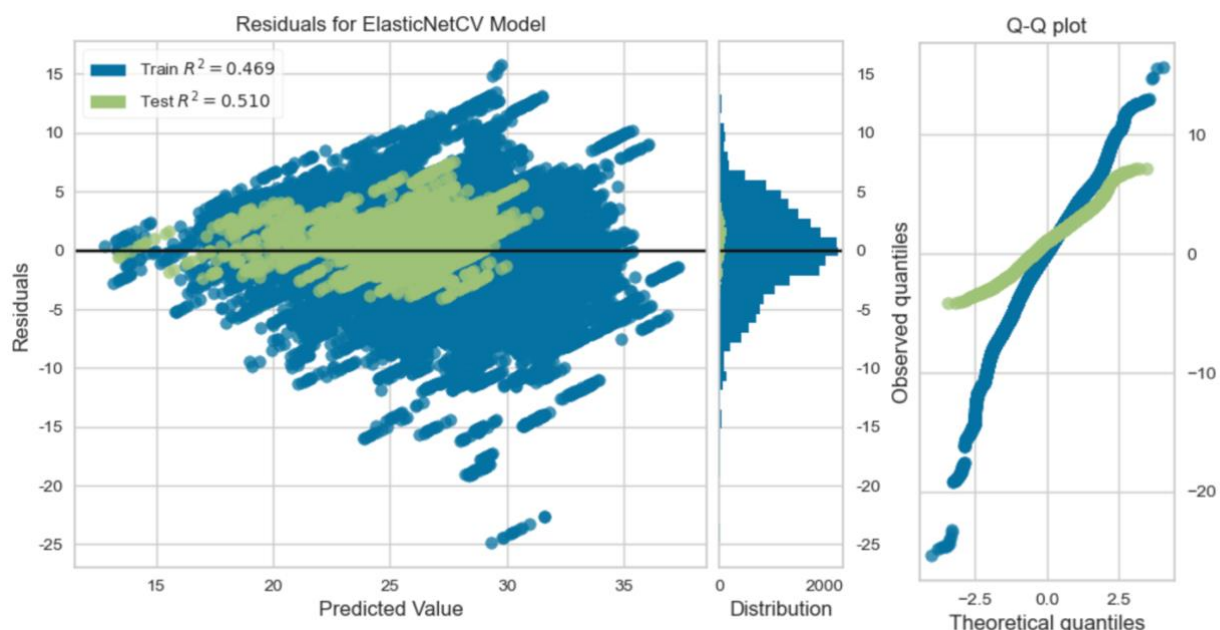
Šie rezultatai buvo gauti naudojant LASSO ir *Ridge* reguliarizacijos kombinaciją. Taikant *ResNet-50* ir *DenseNet-201* požymių išgavimo metodus, geriausias L1 santykis – 0,7. Tai reiškia, jog L1 (LASSO) reguliarizacija šiose regresijose buvo naudingesnė nei L2 (*Ridge*). Naudojant likusias architektūras šis santykis 0,5 – lygus abiejų reguliarizacijų mišinys.

Elastic Net modelio, mokyto su *ResNet-50* vektorizavimo būdu gautais požymiais bei papildomais požymiais, liekanų grafike (48 pav.) galima pastebėti, jog liekanos yra plačiau išsiskirsčiusios, lyginant su geriausiu *Elastic Net* modeliu, mokytu be papildomų požymių. Tai reiškia, jog modelis geriau įsimeina mokymo imtį bei tiksliau prognozuoja nematytas reikšmes. Į regresiją įtraukus papildomą 181 požymį, modelis geba paaiškinti 51 % priklausomojo kintamojo reikšmių.



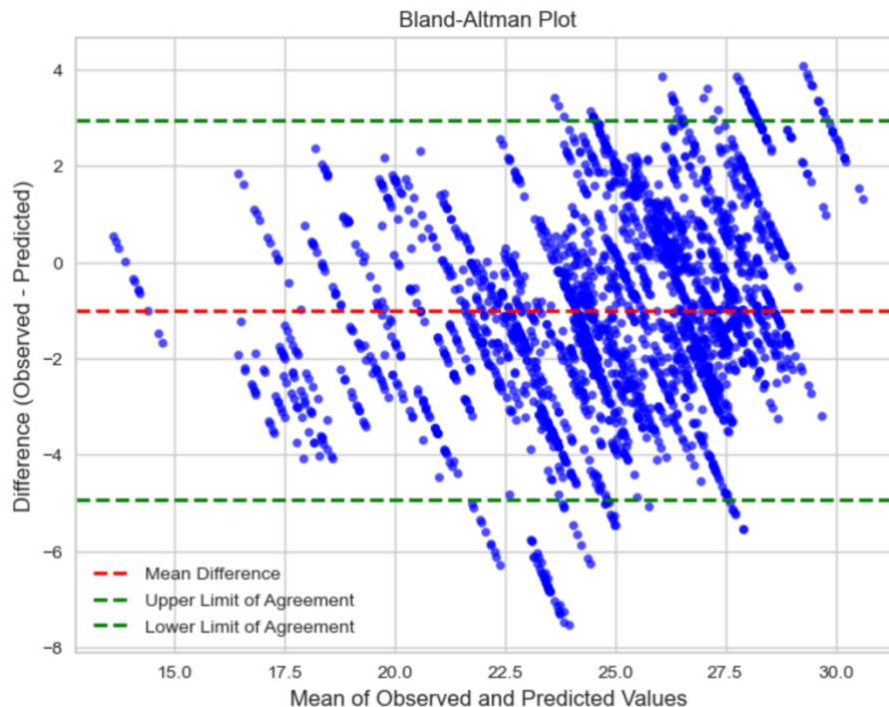
48 pav. *Elastic Net* modelio, mokyto su *ResNet-50* vektorizavimo būdu gautais požymiais bei papildomais X, liekanų grafikas

Šio modelio liekanų pasiskirstymo grafike (49 pav.) matoma, jog testavimo imties liekanos pasiskirsčiusios beveik dvigubai mažesniame intervale (-5-7). Tai reiškia, jog šis modelis stebimas reikšmes nuspėja geriau.



49 pav. Elastic Net modelio, mokyto su ResNet-50 vektorizavimo būdu gautais požymiais bei papildomais X, liekanų pasiskirstymo bei Q-Q grafikai

Bland-Altman'o grafikas (50 pav.) rodo, jog taip pat nebeliko ir tokių didelio reikšmių pervertinimo – nebepastebima reikšmių koncentracija kairiajame apatiniame kampe. Tai parodo, jog šis modelis nėra linkęs priskirti tokių didelių prognozuojamų reikšmių žemus skurdo rodiklius turinčioms savivaldybėms.



50 pav. Elastic Elastic Net modelio, mokyto su ResNet-50 vektorizavimo būdu gautais požymiais bei papildomais X, Bland-Altman'o grafikas

Prasčiausiai prognozuojamos savivaldybės kartojasi (15 lentelė) – čia didžiausiai nuo tikrųjų reikšmių nutolusios prognozės pastebimos tose pačiose savivaldybėse, kaip ir rezultatus prognozuojant iš modelių, mokytų neįtraukiant papildomų požymių. Visas mėnesines prognozes galima pamatyti 7 priede.

15 lentelė. *Elastic Net* modelio, mokyto su *ResNet-50* vektorizavimo būdu gautais požymiais ir papildomais X, spėjimų rezultatai (10 prasčiausių prognozių)

Savivaldybė	Skurdo rodiklis (%)	Prognozė (%)	Δ
Švenčionių rajono savivaldybė	20,1	26,35	6,25
Šiaulių miesto savivaldybė	19,1	25,14	6,04
Neringos savivaldybė	15,5	21,52	6,02
Skuodo rajono savivaldybė	29,7	24,88	4,82
Utenos rajono savivaldybė	21,4	26,02	4,62
Druskininkų savivaldybė	20,1	24,67	4,57
Kauno rajono savivaldybė	17,8	21,84	4,04
Birštono savivaldybė	19,8	23,74	3,94
Elektrėnų savivaldybė	16,7	20,56	3,86
Šiaulių rajono savivaldybė	22,7	26,42	3,72

Tiksliausiai prognozuojamos savivaldybės taip pat didžiaja dalimi sutampa. Visgi, žiūrint į istorinius duomenis, skurdo rodikliai šiose savivaldybėse svyravo ne mažiau nei prognozuojamose prasčiausiai, todėl ženkliai geresnes prognozes lemti galėjo ir kiti faktoriai – kaimynių savivaldybių skurdo rodikliai (kitais tariant, savivaldybės vieta Lietuvoje), ir/ar NTL intensyvumas.

16 lentelė. *Elastic Net* modelio, mokyto su *ResNet-50* vektorizavimo būdu gautais požymiais ir papildomais X, spėjimų rezultatai (10 geriausių prognozių)

Savivaldybė	Skurdo rodiklis (%)	Prognozė (%)	Δ
Pasvalio rajono savivaldybė	26,2	26,17	0,03
Prienų rajono savivaldybė	23,9	23,75	0,15
Klaipėdos miesto savivaldybė	18,8	18,64	0,16
Telšių rajono savivaldybė	25,0	25,17	0,17
Lazdijų rajono savivaldybė	27,3	27,13	0,17
Mažeikių rajono savivaldybė	25,1	25,32	0,22
Kupiškio rajono savivaldybė	26,6	26,33	0,27
Raseinių rajono savivaldybė	26,3	25,95	0,35
Alytaus rajono savivaldybė	25,0	24,63	0,37
Tauragės rajono savivaldybė	25,3	25,75	0,45

Lyginant su prieš tai aptartais rezultatais, didžiųjų miestų prognozės (17 lentelė) taip pat pagerėjo, tačiau Šiaulių miesto prognozės išlieka ženkliai didesnės nei tikrasis šios savivaldybės skurdo rizikos lygio rodiklis.

17 lentelė. *Elastic Net* modelio, mokyto su *ResNet-50* vektorizavimo būdu gautais požymiais ir papildomais X, spėjimų rezultatai (didžiųjų miestų rezultatai)

Savivaldybė	Skurdo rodiklis (%)	Progozė (%)	Δ (Pokytis, lyginant su <i>Elastic Net</i> be papildomų X)
Šiaulių miesto savivaldybė	19,1	25,14	6,04 (-1,33)
Vilniaus miesto savivaldybė	13,1	15,71	2,61 (-1,28)
Panevėžio miesto savivaldybė	19,0	22,15	3,15 (-0,40)
Kauno miesto savivaldybė	16,4	17,49	1,09 (-2,11)
Klaipėdos miesto savivaldybė	18,8	18,64	0,16 (-0,97)

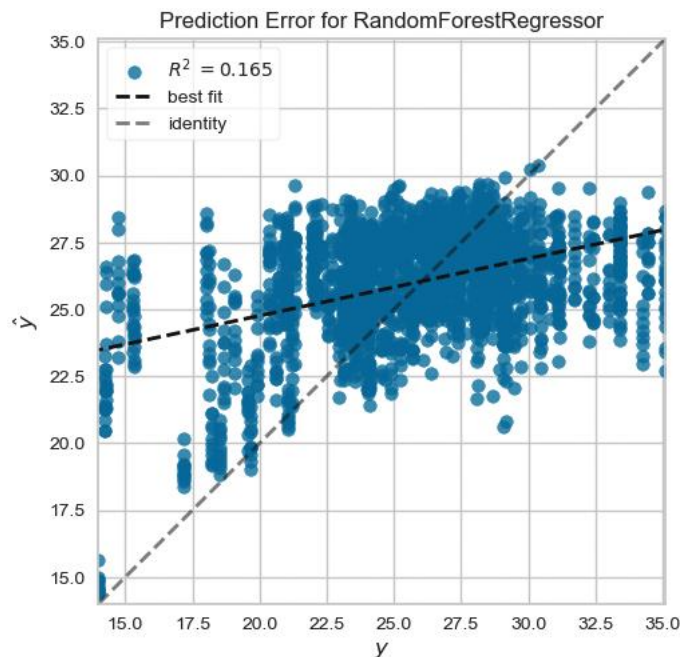
3.5.2. Atsitiktinių miškų modelių rezultatai

Progozuojant skurdo rizikos lygio rodiklius su atsitiktinių miškų modeliais (mokant su mažiausiais 56×56 dydžio vaizdais, o testavimą atliekant be persidengiančių vaizdų), geriausi rezultatai mėnesiniams ir metiniams duomenims gauti su *DenseNet-121* vektorizavimo metodu gautais požymių vektoriais (18 lentelė) (8 priedas).

18 lentelė. Atsitiktinių miškų modelių, mokyto su skirtingais vektorizavimo būdais gautais požymiais, mėnesinių spėjimų rezultatai (testavimo imtis – $2\,472\,56 \times 56$ dydžio vaizdai)

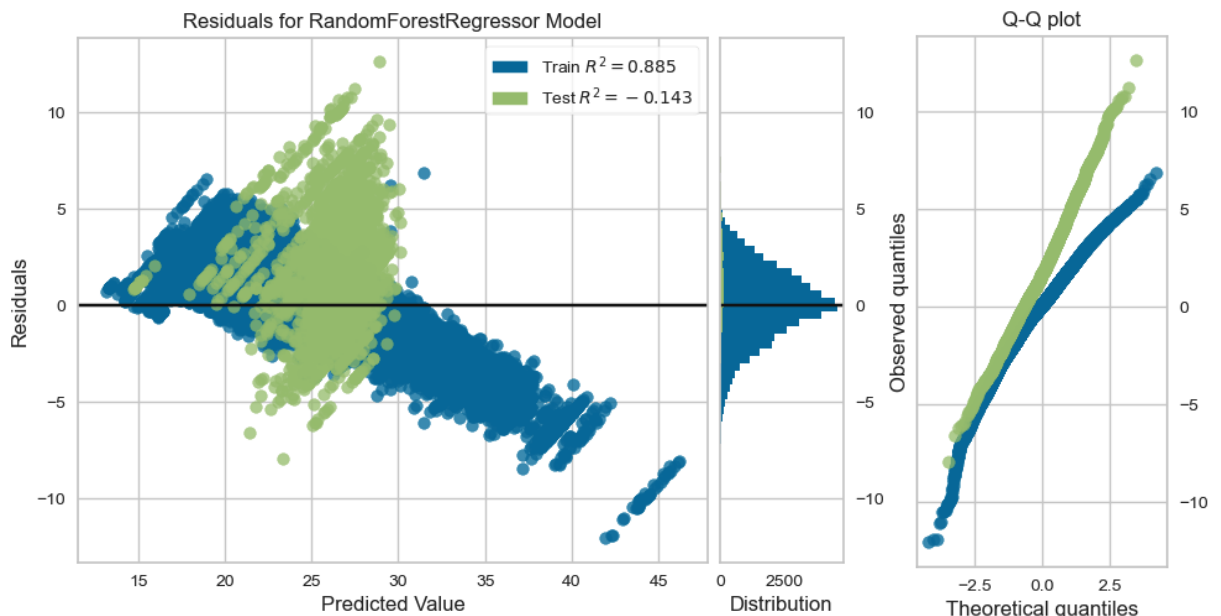
CNN architektūra	Mokymo R^2	Testavimo R^2	RMSE	MDAE	MAPE	SMAPE	MAAPE
<i>ResNet-50</i>	0,885	-0,197	3,901	2,770	15,528	13,851	15,097
<i>ResNet-101</i>	0,885	-0,236	3,977	2,734	15,857	14,091	15,402
<i>NASNetMobile</i>	0,882	-0,283	4,106	2,848	16,431	14,526	15,914
<i>EfficientNet-B0</i>	0,887	-0,197	3,891	2,765	15,481	13,807	15,055
<i>DenseNet-121</i>	0,885	-0,143	3,769	2,689	14,787	13,278	14,399
<i>DenseNet-201</i>	0,887	-0,138	3,785	2,775	14,906	13,386	14,523

Atsitiktinių miškų modelio, mokyto su *DenseNet-121* vektorizavimo būdu gautais požymiais, liekanų grafike (51 pav.) galima pastebėti, jog liekanų pasiskirstymas yra panašus į *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, be papildomų X (43 pav.). Jos taip pat yra labiau susikoncentravusios viename plote bei modelis yra linkęs pervertinti, ar nuvertinti priklausomą kintamąjį pastoviu prognozuojamų reikšmių skaičiumi.



51 pav. Atsitiktinių miškų modelio, mokyto su *DenseNet-121* vektorizavimo būdu gautais požymiais, liekanų grafikas

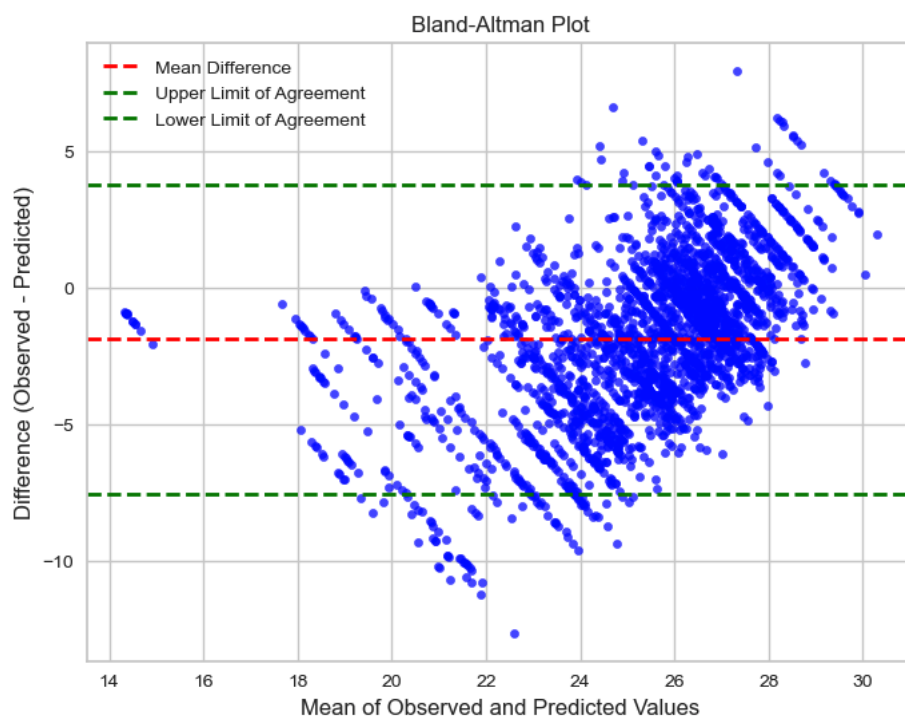
Liekanų pasiskirstymo grafike (52 pav.) matoma, jog testavimo imties liekanos, lyginant su *Elastic Net* regresijos liekanų intervalu, yra išsidėsčiusios didesniame intervale: apie -13–13. Mokymo R^2 yra daug didesnis nei *Elastic Net* – net 0,885, o testavimo R^2 netgi neigiamas. Q-Q grafikas indikuoja, jog testavimo liekanos pasiskirsčiusios normaliuoju skirstiniu.



52 pav. Atsitiktinių miškų modelio, mokyto su *DenseNet-121* vektorizavimo būdu gautais požymiais, liekanų pasiskirstymo bei Q-Q grafikai

Bland-Altman'o grafikas (53 pav.) yra beveik identiškas *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, be papildomų X (45 pav.), todėl galima teigti, jog dalis

atsitiktinių miškų metodu prognozuojamų reikšmių yra taip pat didesnės nei faktinės, ypač kai faktinės reikšmės yra mažos. Todėl atsitiktinių miškų modelis yra taip pat linkęs rezultatus pervertinti.



53 pav. Atsitiktinių miškų modelio, mokyto su *DenseNet-121* vektorizavimo būdu gautais požymiais, Bland-Altman'o grafikas

Skurdo rodiklių prognozės didiesiems miestams, naudojant atsitiktinių miškų metodą yra šiek tiek prastesnės (19 lentelė). Pastebima, jog rezultatai yra blogesni nei *Elastic Net* – geriau prognozuojamas tik Šiaulių miesto savivaldybės skurdo rodiklis.

19 lentelė. Atsitiktinių miškų modelio, mokyto su *DenseNet-121* vektorizavimo būdu gautais požymiais, spėjimų rezultatai (didžiųjų miestų rezultatai)

Savivaldybė	Skurdo rodiklis (%)	Prognozė (%)	$ \Delta $
Šiaulių miesto savivaldybė	19,1	24,03	4,93
Vilniaus miesto savivaldybė	13,1	17,07	3,97
Panevėžio miesto savivaldybė	19,0	23,84	4,84
Kauno miesto savivaldybė	16,4	19,60	3,20
Klaipėdos miesto savivaldybė	18,8	20,46	1,66

Atsitiktinių miškų modelių rezultatai su papildomais požymiais

Į modelio mokymo ir testavimo imtis įtraukus papildomą 181 požymį bei nustačius parametrus šiems požymiams sureikšminti (*always.split.variables*), rezultatai pagerėjo (20 lentelė, 9 priedas) – čia geriausius rezultatus parodė atsitiktinių miškų modelis, mokytas su *NASNetMobile* vektorizavimo būdu gautais požymiais. Šio modelio testavimo determinacijos koeficientas R^2 – 0,378, MDAE – 2,055, o MAPE – 10,940.

20 lentelė. Atsitiktinių miškų modelių, mokytų su skirtingais vektorizavimo būdais gautais požymiais ir papildomais X, mėnesinių spėjimų rezultatai (testavimo imtis – $2\,472\,56 \times 56$ dydžio vaizdai)

CNN architektūra	Testavimo R^2	RMSE	MDAE	MAPE	SMAPE	MAAPE
<i>ResNet-50</i>	0,342	2,940	2,135	11,164	10,445	11,002
<i>ResNet-101</i>	0,371	2,889	2,062	11,014	10,365	10,862
<i>NASNetMobile</i>	0,378	2,887	2,055	10,940	10,287	10,785
<i>EfficientNet-B0</i>	0,377	2,887	2,060	11,008	10,356	10,858
<i>DenseNet-121</i>	0,378	2,903	2,137	11,089	10,372	10,931
<i>DenseNet-201</i>	0,339	2,945	2,109	11,174	10,539	11,020

Lyginant su atsitiktinių miškų rezultatais, gautais naudojant vien vaizdų požymių vektorius, nežymiai pagerėjo Šiaulių miesto ir Klaipėdos miesto savivaldybių prognozės, gana ženkliai pagerėjo Vilniaus miesto, tačiau Kauno ir Panevėžio – suprastėjo. Tai reiškia, jog nors vidutinės klaidos (ar jų mediana) ir geresnės, visgi, reikšmingo rezultatų pagerėjimo didžiųjų miestų prognozėse, įtraukus naujus požymius, nėra. Visos prognozės matomos 5 priede.

21 lentelė. Atsitiktinių miškų modelio, mokyto su *NASNetMobile* vektorizavimo būdu gautais požymiais ir papildomais X, spėjimų rezultatai (didžiųjų miestų rezultatai)

Savivaldybė	Skurdo rodiklis (%)	Prognozė (%)	$ \Delta $ (Pokytis, lyginant su atsitiktiniais miškais be papildomų X)
Šiaulių miesto savivaldybė	19,1	23,87	4,77 (-0,04)
Vilniaus miesto savivaldybė	13,1	15,00	1,90 (-2,44)
Panevėžio miesto savivaldybė	19,0	25,61	6,61 (+1,69)
Kauno miesto savivaldybė	16,4	21,36	4,96 (+1,55)
Klaipėdos miesto savivaldybė	18,8	20,06	1,26 (-0,47)

Dėl šių priežasčių galima teigti, jog atsitiktinių miškų metodas skurdo lygiui iš satelitinių nakties vaizdų nepasiteisino – *Elastic Net* reguliarizavimo metodais sudaryti modeliai, su papildomais požymiais ar be jų, skurdo lygį prognozavo tiksliau.

3.6. Geriausių rezultatų palyginimas su laiko eilučių prognozėmis

Siekiant įvertinti geriausių *Elastic Net* modelių rezultatus, prognozavimas buvo atliktas ir kitais metodais – laiko eilutėmis (22 lentelė). Visi laiko eilučių metodai pasirodė prasčiau nei mūsų siūlomas metodas. Taip pat laiko eilučių metodai turi trūkumų: lyginant su prognozavimu iš NTL taikant *DenseNet-121* ir *ResNet-50* architektūras bei *Elastic Net* regresiją, laiko eilučių metodai prognozes atlieka tik metiniams duomenims ir negali nustatyti skurdo rizikos lygio mėnesiniu agreguotumu.

22 lentelė. *Elastic Net* modelių, mokytų su *ResNet-50* (su papildomais X) bei *DenseNet-201* (be papildomų X) vektorizavimo būdais gautais požymiais, palyginimas su kitais metodais gautais rezultatais

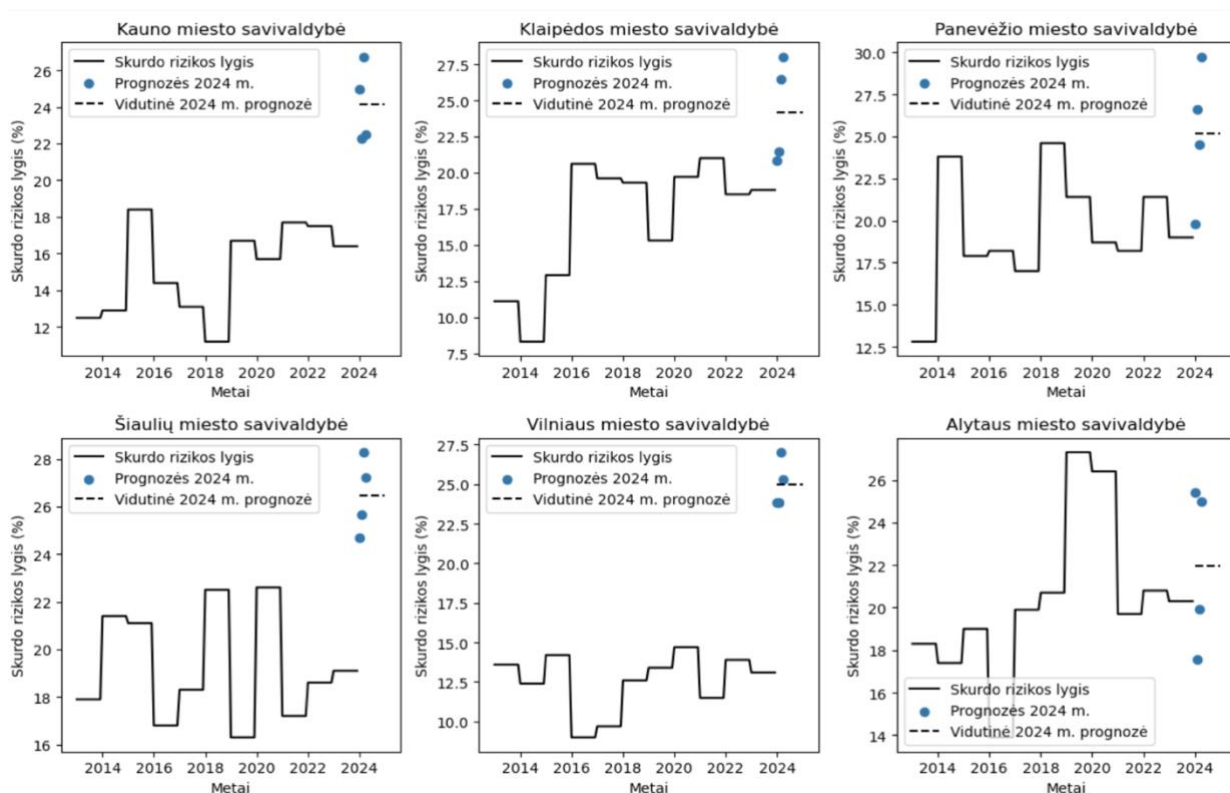
Prognozavimo metodas	RMSE	MDAE	MAPE	SMAPE	MAAPE
Mūsų su papildomais X	2,636	1,960	9,811	9,179	9,685

Mūsų be papildomų X	3,623	2,335	14,224	12,799	13,877
ETS	4,528	2,161	15,126	14,084	14,361
<i>Theta</i>	4,702	2,335	15,591	14,704	14,947
ARIMA	4,556	2,620	15,776	14,569	15,014
<i>Naïve</i>	4,918	2,400	15,834	14,427	14,712
TSLM	5,033	3,290	17,716	16,447	16,996
<i>Prophet</i>	5,109	3,311	17,730	16,508	17,052

Geriausio modelio, sudaryto naudojant *Elastic Net* regresiją (į modelio mokymo ir testavimo imtis įtraukiant papildomą 181 požymį bei požymius, gautus *ResNet-50* vektorizavimo būdu), MAPE yra net 35% mažesnė, lyginant su geriausiu laiko eilučių metodu gautais rezultatais (ETS). Tuo tarpu nuo prasčiausio laiko eilučių modelio (*Prophet*) mūsų siūlomas rezultatas skiriasi net 45%.

3.7. Prognozės 2024-iesiems

Didžiųjų miestų prognozės 2024-iesiems nėra itin pozityvios – *Elastic Net* modelis, mokytas su *ResNet-50* vektorizavimo būdu gautais požymiais bei papildomais požymiais (žiūrėti geriausius rezultatus 13 ir 14 lentelėse), prognozuoja skurdo rizikos lygio rodiklio kilimą (52 pav.). Pažymėtina, jog Alytaus miesto savivaldybės skurdo rodiklis 2024 m. sausio ir vasario mėnesiais prognozuojamas mažesnis nei 2022–2023 m., tuo tarpu vidutinė 4 mėnesių prognozė yra šiek tiek didesnė. Visgi, matoma vidutinė skurdo tendencija gali kisti, kadangi prognozėms gauti buvo naudojami tik pirmųjų keturių mėnesių vaizdų duomenys. Visų savivaldybių prognozės 2024 m. pateikiamos 10 priede.



54 pav. Didžiųjų miestų skurdo rodiklio prognozės 2024 m. sausio–balandžio mėn.

3.8. Apibendrinimas ir išvalgos

Visi vertinimo kriterijai (RMSE, MDAE, MAPE, SMAPE, MAAPE) buvo reikšmingai geresni naudojant mažesnio dydžio įvesties vaizdus (56×56) – turint labai didelę mokymo imtį (54 780), tačiau testuojant su mažesne (2 472), t. y., be persidengiančių vaizdų. Įvedus papildomus 181 požymius, geresni rezultatai pastebėti modelį mokant ir testuojant su nepersidengiančiais vaizdais. Metinių rodiklių prognozės yra šiek tiek tikslesnės nei prognozuojant mėnesiniu agreguotumu.

Galima teigti, jog geriausi mėnesiniai rezultatai (pagal MDAE bei MAPE) buvo pasiekti naudojant *Elastic Net* regresiją: modelį mokant su *ResNet-50* vektorizavimo būdu gautais požymių vektoriais bei įtraukiant papildomus X. Čia L1 santykis buvo 0,7, optimali alfa reikšmė – 0,0001, mokymo R^2 – 0,469, testavimo R^2 – 0,510, RMSE – 2,636, MDAE – 1,960, MAPE – 9,811, SMAPE – 9,179, o MAAPE – 9,685.

Įtraukdama LASSO baudos sąlygą, *Elastic Net* kai kuriuos koeficientus nustato iki 0 (atrenka svarbiausius kintamuosius), tuo tarpu įtraukdama *Ridge* – sutvarko multikolinearumo apribojimus. Tokiu būdu, parinkdama optimalų L1 santykį, *Elastic Net* reguliarizacija sprendžia minėtus apribojimus. Taikant *ResNet-50* ir *DenseNet-201* vektorizavimo metodus, L1 santykis – 0,7. Tai reiškia, jog L1 (LASSO) reguliarizacija šiose regresijose buvo naudingesnė nei L2 (*Ridge*). Naudojant likusias architektūras (*Resnet-101*, *NASNetMobile*, *EfficientNet-B0* ir *DenseNet-201*) geriausi rezultatai gauti tuomet, kai L1 santykis 0,5 – lygus abiejų reguliarizacijų mišinys.

Atsitiktinių miškų algoritmai pasirodė prasčiau nei *Elastic Net* – čia modelių klaidos didesnės tiek įtraukiant papildomus požymius, tiek jų neįtraukiant. Geriausi atsitiktinių miškų rezultatai pasiekti su *NASNet* vektorizavimo metodu – su papildomais požymiais, didele mokymo imtimi (54 780), tačiau testuojant be persidengiančių vaizdų.

Pastebima, kad *Elastic Net* regresijos modelis, mokytas su *ResNet-50* metodu gautais požymių vektoriais (žiūrėti geriausius rezultatus 13 ir 14 lentelėse), pasirodė geriau ir už tradicinius laiko eilučių prognozavimo metodus, tokius kaip ETS, *Theta*, ARIMA, *Naïve*, TSLM ar *Prophet*. Taip pat laiko eilučių metodai yra tinkami tik metinėms reikšmėms prognozuoti ir negali nustatyti skurdo rizikos lygio mėnesiniu agreguotumu.

Skirtingais CNN metodais gaunami rezultatai skiriasi nereikšmingai dėl Lietuvos skurdo rizikos lygio specifikos. Buvo pastebėta tendencija kai kurioms savivaldybėms priskirti didesnes reikšmes nei yra iš tikrųjų. Pavyzdžiui, visi modeliai yra linkę pervertinti Šiaulių miesto rodiklius. Galima teigti, jog tokie neatitikimai galimi dėl įvairių priežasčių. Šiuo atveju čia didesnėms skurdo rodiklio prognozėms įtakos gali turėti šios savivaldybės geografinė vieta – kadangi į 56×56 dydžio vaizdą sutelpa ne tik miesto, tačiau ir rajono savivaldybė (turinti istoriškai aukštesnius skurdo rodiklius), modelis didesnę skurdo rodiklį priskiria ir pačiam miestui. Taip pat, žvelgiant į prasčiausias prognozes, pastebėtas ir kitų savivaldybių pasikartojimas – abejais mašininio mokymosi metodais (*Elastic Net* bei atsitiktiniais miškais, mokant su papildomais požymiais, ar be jų) prasčiausiai prognozuojamos buvo Kauno rajono, Neringos, Švenčionių rajono ir Elektrėnų savivaldybės. Kauno rajono savivaldybės rodiklius gali iškreipti į vaizdus patenkančios besiribojančios savivaldybės – pavyzdžiui, kaimyninė Kėdainių rajono savivaldybė visu analizuojamu laikotarpiu turėjo aukštus skurdo rodiklius (ypač 2015 m., kai skurdas siekė net 54,5 %), tuo tarpu pats Kauno rajonas – kur kas mažesnius (tais pačiais

2015 m. Kauno rajono skurdo rodiklis buvo kone perpus mažesnis – 26,7 %). Neringos savivaldybė, nors vaizduose daugiausia patenka kartu su Klaipėdos miesto ir rajono savivaldybėmis, kurios analizuojamu laikotarpiu turėjo mažus ar vidutinio dydžio skurdo rodiklius, gali būti prognozuojama prasčiau dėl kelių priežasčių – pavyzdžiui, šią savivaldybę apimančiuose vaizduose šviesų beveik nėra, tačiau skurdo rodiklis vienas mažiausių visoje Lietuvoje. Prie prastesnio Švenčionių rajono savivaldybės rodiklių prognozavimo gali prisidėti ne tik juodi nakties vaizdai, tačiau ir svyruojantis skurdo rodiklis – ši savivaldybė turi vieną didžiausių standartinių nuokrypių – 7,74 (2 priedas), o jos skurdo rodikliai varijavo nuo 18,8 iki 41,4 %, todėl modeliams prognozuoti skurdo rodiklius buvo sudėtingiau. Nors Elektrėnų savivaldybės skurdo rodikliuose dideli svyravimai nepastebėti, tačiau dauguma ją supančių savivaldybių turi aukštesnius skurdo rodiklius – dėl kontrastingų skurdo rizikos lygio atotrūkių tarp savivaldybių, kai kurių savivaldybių rodiklius modeliai prognozuoja prasčiau.

Visgi, didesnių reikšmių priskyrimą ir pakankamai aukštas 2024 m. skurdo lygio prognozes didiesiems miestams galima aiškinti ne tik savivaldybių dydžiu, lokacija, ankstesnių metų skurdo rodikliais, NTL intensyvumu, ar visų šių veiksnių kombinacija, tačiau ir kitomis priežastimis. 2023 m. tyrimai rodo, kad daugiau nei 400 tūkst. žmonių Lietuvoje yra „nematomi“ [71] – jie nepatenka į skurstančių asmenų statistiką dėl per mažų pajamų, išsilavinimo, ar staigiai pablogėjusios finansinės ir socialinės situacijos [71]. Kadangi tokių žmonių skurdo rodikliai dažnai neįtraukia, tikrasis Lietuvos skurdo lygis iš tiesų gali būti gerokai didesnis. Būtent tokią tendenciją dalyje savivaldybių rodo ir šiame tyrime sudaryti modeliai.

Apibendrinant, CNN metodai yra tinkami vaizdams vektorizuoti, tačiau prognozuojant mėnesinius skurdo rizikos lygio rodiklius yra linkę pervertinti kai kurių savivaldybių reikšmes. Už tikrąsias reikšmes didesnės prognozės galimos ir dėl oficialiųjų skurdo rodiklių apribojimų – šie, manoma, realybėje yra gerokai didesni nei pateikiamoje statistikoje. Tai fiksuoja ir sudaryti modeliai. CNN metodai (ypač *ResNet-50*) turi daug potencialo skurdui tikslesnėms skurdo rizikos lygio prognozėms ateityje – kuomet skurdo rizikos lygio rodikliai tarp savivaldybių susilygins, ar bent neturės tokio kontrastingo atotrūkio.

Išvados

1. Santykiniai skurdo rodikliai yra aktualesni už absoliučius ir eksperimentinė skurdo rizikos statistika yra viena jautriausių ir tiksliausių. Satelitiniai nakties vaizdai gali praversti stebint įvairius ekonominius šalies aspektus, tame tarpe ir skurdą ar pajamų nelygybę. Artimiausi tyrimai nagrinėjo Afrikos, Pietų Amerikos ir Azijos šalis ir jų namų ūkių apklausas. Autoriai naudojo įvairias CNN architektūras, ypač *ResNet-50*, ir su jomis vektorizuotiems vaizdams taikė reguliarizuotą (Ridge tipo) regresiją.
2. Darbe pasirinktos 6 *ImageNet* vaizdų klasifikavimo uždaviniui paruoštos CNN architektūros (nuėmus klasifikavimo „galvą“ išvestyje gaunamas fiksuoto ilgio vektorius): *ResNet-50* ir *ResNet-101* (2048), *NASNetMobile* (1056), *EfficientNet-B0* (1280), *DenseNet-201* (1920) bei *DenseNet-121* (1024). Šių architektūrų pagalba kiekvienai vaizdo gardelei buvo gaunamas požymių vektorius.
3. Nakties vaizdams gauti pasirinktas NASA „Black Marble“ produktas, pasižymintis aukštos kokybės koreguotais satelitiniais vaizdais (Suomi-NPP VIIRS). Rastriniai 2013–2023 m. vaizdai parsisiųsti naudojant R paketą *blackmarbler*, trūkstami pikseliai interpoliuoti. Panaudojant „QGIS“ įrankį atliktas normalizavimas ir priskirta „Magma“ spalvų paletė.
4. CNN architektūros buvo pritaikytos vaizdams vektorizuoti: vaizdai sukarpyti į 56, 112 bei 224 dydžio gardeles. Siekiant padidinti mokymosi imtį, vaizdai taip pat buvo perstumti 50 % x bei y ašimi. Tokiu būdu vaizdų skaičius padidėjo atitinkamai iki 54 780, 15 840 bei 5 016.
5. Eksperimento metu buvo sudaryti *Elastic Net* regresijos bei atsitiktinių miškų modeliai – buvo naudojamos skirtingo dydžio įvestys bei skirtingais CNN gauti požymių vektoriai. Skurdo rizikos rodikliai prognozuoti 2023-iesiems metams. Geriausi rezultatai, prognozuojant mėnesinius skurdo rodiklius, buvo gauti atliekant vektorizavimą su *ResNet-50*, *Elastic Net* modelį mokant su mažiausio dydžio nepersidengiančiais vaizdais bei įvedant papildomus požymius. Čia testavimo $R^2 = 0,510$, RMSE – 2,636, MDAE – 1,960, MAPE – 9,811.
6. 2024-tųjų pirmų 4 mėnesių skurdo rizikos prognozės didiesiems miestams gana pesimistiškos – geriausias modelis prognozuoja skurdo rizikos lygio kilimą. Visgi, matoma vidutinė skurdo tendencija gali kisti, nes prognozėms gauti buvo naudojami tik pirmųjų 4 mėnesių vaizdai.
7. Lyginant geriausio vektorizavimo būdo prognozes su laiko eilučių modelių rezultatais nustatyta, jog tiksliausi *Elastic Net* modeliai pasirodė geriau nei tradiciniai laiko eilučių prognozavimo metodai, tokie kaip ETS, *Theta*, ARIMA, *Naïve*, TSLM ar *Prophet*. Laiko eilučių metodai yra tinkami tik metinėms reikšmėms prognozuoti ir negali nustatyti skurdo rizikos lygio mėnesiniu agreguotumu.
8. Geriausi rezultatai gauti su nepersidengiančiais 56×56 dydžio vaizdais tada, kai į modelių mokymo imtis buvo įtraukiami papildomi požymiai apie savivaldybių vietą, dydį bei ankstesnius skurdo rodiklius. Rezultatai rodo, jog CNN metodai yra tinkami mėnesiniams Lietuvos skurdo rizikos lygio rodikliams nustatyti, tačiau reikia įvertinti tai, jog sudaryti modeliai yra linkę kai kurioms savivaldybėms priskirti didesnes reikšmes nei pateikiamos oficialioje statistikoje. Visgi, manoma, jog tikrasis Lietuvos skurdo rizikos rodiklis yra daug didesnis nei oficialus – būtent tai dalyje savivaldybių fiksuoja ir sudarytas modelis. Todėl galima kelti prielaidą, jog pasiūlytas modelis išvelgia daugiau nei pateikia oficiali skurdo rizikos lygio statistika. CNN metodai (ypač *ResNet-50*) turi daug potencialo tikslesnėms skurdo prognozėms ateityje – kuomet skurdo rizikos lygio rodikliai tarp savivaldybių susilygins, ar bent neturės tokio kontrastingo atotrūkio.

Literatūros sąrašas

1. Nacionalinė žemės tarnyba prie Aplinkos ministerijos. *Geografinių Informacinių Sistemų Pagrindai*. 2008 Prieiga per: https://www.geoportal.lt/geoportal/documents/11958/13305/GII-01_mokomoji_knyga.pdf/bfd136c9-8647-42be-bbe0-692b3ba9a192 DOI geoportal.lt.
2. PANDEY, S., AGARWAL, T. and KRISHNAN, N.C. Multi-Task Deep Learning for Predicting Poverty from Satellite Images. *Proceedings of the AAAI Conference on Artificial Intelligence* Vol. 32. no. 1.2018 DOI 10.1609/aaai.v32i1.11416.
3. LOK-DESSALLIEN, R. Review of Poverty Concepts and Indicators. 2000. Prieiga per: [https://ftp.unpad.ac.id/orari/library/library-ref-ind/ref-ind-1/application/poverty-reduction/Poverty/Review of Poverty Concepts.pdf](https://ftp.unpad.ac.id/orari/library/library-ref-ind/ref-ind-1/application/poverty-reduction/Poverty/Review%20of%20Poverty%20Concepts.pdf).
4. H. WHITE and S. CARVALHO. *Combining the Quantitative and Qualitative Approaches to Poverty Measurement and Analysis: The Practice and the Potential*. 1818H Street, N.W.Washington, D.C.20433,U.S.A.: The World Bank. 1997.
5. Instituto Nacional De Estadística. *The Presentation of a Range of Methods to Obtain Measures of Poverty*. Madrid, España: 2007 Prieiga per: https://www.ine.es/en/daco/daco42/sociales/pobreza_en.pdf.
6. ZABARAUSKAITĖ, R. Skurdo Matavimas Ir Mažinimo Kryptys Lietuvoje. 2008 ISSN 978-9955-28-371-3.
7. OECD. *Society at a Glance 2019*. 2019 Prieiga per: <https://www.oecd.org/social/society-at-a-glance-19991290.htm>.
8. FÖRSTER, M. and D'ERCOLE, M. The OECD Approach to Measuring Income Distribution and Poverty. 2012. Prieiga per: https://www.researchgate.net/publication/287677872_The_OECD_Approach_to_Measuring_Income_Distribution_and_Poverty DOI 10.1093/acprof:oso/9780199860586.003.0002.
9. KONKEL, R. The Monetization of Global Poverty: The Concept of Poverty in World Bank History, 1944–90. *Journal of Global History*2014, vol. 9.2. pp. 276-300 DOI 10.1017/S1740022814000072.
10. World Bank. *About the World Bank*. Prieiga per: <https://www.worldbank.org/en/about>.
11. The World Bank. *Poverty and Inequality*. Prieiga per: <https://datatopics.worldbank.org/world-development-indicators/themes/poverty-and-inequality.html>.
12. OECD. *Poverty Rate*. 2024 Prieiga per: <https://data.oecd.org/inequality/poverty-rate.htm#indicator-chart> DOI 10.1787/0fe1315d-en.
13. World Bank. *Poverty and Inequality Platform*. 2024. Prieiga per: <https://pip.worldbank.org/#>.
14. Lietuvos Respublikos Socialinės apsaugos ir darbo ministerija. *Skurdo Ir Pajamų Nelygybės Prognozės*. 2022 Prieiga per: https://socmin.lrv.lt/uploads/socmin/documents/files/pdf/Trumprastis_Skurdas_nelygybe_Final.pdf.

15. Eurostat. *Persons at Risk of Poverty Or Social Exclusion by Age and Sex*. 2023 Prieiga per: https://ec.europa.eu/eurostat/databrowser/view/ilc_peps01n/default/table?lang=en/.
16. Nacionalinio skurdo mažinimo organizacijų tinklas. *Skurdas Ir Socialinė Atskirtis 2021*. Vilnius: 2021 Prieiga per: <https://www.eapn.eu/wp-content/uploads/2021/10/eapn-Leidiny-PW-Lithuania-Final-5258.pdf>.
17. Nacionalinis skurdo mažinimo organizacijų tinklas. *Skurdas Ir Socialinė Atskirtis 2023*. 2023, no. 1 (5). Prieiga per: <https://www.smtinklas.lt/wp-content/uploads/2023/10/Skurdas-ir-socialin%C4%97-atskirtis-Lietuvoje-2023.pdf> ISSN 2783-8080.
18. OSP. *Eksperimentinė Statistika. Skurdo Rodikliai*. 2023 Prieiga per: <https://osp.stat.gov.lt/eksperimentine-statistika/skurdo-rodikliai>.
19. PAROLIN, Z., et al. Estimating Monthly Poverty Rates in the United States . *Journal of Policy Analysis and Management* 2022, vol. 41, no. 4. pp. 975-1270. Prieiga per: <https://onlinelibrary.wiley.com/doi/full/10.1002/pam.22403> ISSN 0276-8739. DOI 10.1002/pam.22403.
20. JEAN, N., et al. Combining Satellite Imagery and Machine Learning to Predict Poverty. 2016, vol. 353, no. 6301. pp. 790-794 DOI 10.1126/science.aaf7894.
21. BABENKO, B., et al. Poverty Mapping using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, with an Application in Mexico. 2017 DOI 10.48550/arXiv.1711.06323.
22. D. CHEN. *Temporal Poverty Prediction using Satellite Imagery*. 2017 Prieiga per: <http://vision.stanford.edu/teaching/cs231n/reports/2017/pdfs/552.pdf>.
23. PEREZ, A., et al. Poverty Prediction with Public Landsat 7 Satellite Imagery and Machine Learning. 2017 DOI 10.48550/arXiv.1711.03654.
24. LIN, T., et al. Feature Pyramid Networks for Object Detection. 2017. Prieiga per: <https://arxiv.org/abs/1612.03144> DOI 10.48550/arXiv.1612.03144.
25. WU, P. and TAN, Y. Estimation of Poverty Based on Remote Sensing Image and Convolutional Neural Network. 2019, no. 8. pp. 89-98. Prieiga per: <https://www.scirp.org/journal/paperinformation?paperid=97205> DOI 10.4236/ars.2019.84006.
26. PUTRI, S. R., WIJAYANTO, A. W. and PRAMANA, S. Multi-Source Satellite Imagery and Point of Interest Data for Poverty Mapping in East Java, Indonesia: Machine Learning and Deep Learning Approaches. *Remote Sensing Applications: Society and Environment* 2023, no. 29 DOI 10.1016/j.rsase.2022.100889.
27. KONDMANN, L. and ZHU, X. X. Measuring Changes in Poverty with Deep Learning and Satellite Imagery. 2020. Prieiga per: https://www.researchgate.net/publication/345762555_Measuring_Changes_in_Poverty_with_Deep_Learning_and_Satellite_Imagery.
28. WU, D., et al. Skip Connections Matter: On the Transferability of Adversarial Examples Generated with ResNets. 2020 DOI 10.48550/arXiv.2002.05990.

29. Lietuvos Statistikos Departamentas. *Skurdo Statistinių Rodiklių Skaičiavimo Metodika*. 2019
Prieiga per: https://osp.stat.gov.lt/documents/10180/576798/Skurdo_rodikliu_metodika_2016.pdf.
30. AGHDAM, H. H. and HERAVI, E. J. Convolutional Neural Networks. Guide to Convolutional Neural Networks. 2017, no. 10.978-973. pp. 85–130 DOI 10.1007/978-3-319-57550-6_3.
31. KALOGEITON, V. Localizing Spatially and Temporally Objects and Actions in Videos. 2017.
Prieiga per: https://www.researchgate.net/publication/322262996_Localizing_spatially_and_temporally_objects_and_actions_in_videos.
32. YANI, M., BUDHI IRAWAN, S., Si, MT and CASI SETININGSIH, S., MT. Application of Transfer Learning using Convolutional Neural Network Method for Early Detection of Terry's Nail. *Journal of Physics Conference Series* 2019 ISSN 1201(1):012052. DOI 10.1088/1742-6596/1201/1/012052.
33. NABIYEV, N. and MALEKZADEH, S. Anomalous Sound Localization Estimation. 2021, no. CC BY-NC-SA 4.0 DOI 10.13140/RG.2.2.25949.95201.
34. TOLIUSIS, R. Akies Dugno Nuotraukų Segmentavimas Naudojant Dirbtinius Neuroninius Tinklus. *Vilniaus Universitetas Duomenų Mokslo Ir Skaitmeninių Technologijų Institutas* 2019.
Prieiga per: https://www.mii.lt/files/doc/lt/doktorantura/ataskaitine_konferencija/t007_ma_2019_toliusis.pdf
ISSN DMSTI-DS-T007-19-10.
35. WU, H., XIN, M. and FANG, W. Multi-Level Feature Network with Multi-Loss for Person Re-Identification. 2019, vol. 7. Prieiga per: <https://ieeexplore.ieee.org/document/8755991> DOI 10.1109/ACCESS.2019.2927052.
36. ADDAGARLA, K., CHAKRAVARTHI, G. K. and ANITHA, P. Real Time Multi-Scale Facial Mask Detection and Classification using Deep Transfer Learning Techniques . *International Journal of Advanced Trends in Computer Science and Engineering* 2020, vol. 9, no. 4. Prieiga per: https://www.researchgate.net/publication/344252628_Real_Time_Multi-Scale_Facial_Mask_Detection_and_Classification_Using_Deep_Transfer_Learning_Techniques
ISSN 2278-3091. DOI 10.30534/ijatcse/2020/33942020.
37. TAN, M. and LE, Q.V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *International Conference on Machine Learning*, 2019 2019, no. 5. Prieiga per: <https://arxiv.org/abs/1905.11946v5> DOI 10.48550/arXiv.1905.11946.
38. HUANG, G., LIU, Z., MAATEN, L.v.d. and WEINBERGER, K.Q. Densely Connected Convolutional Networks. *Cvpr 2017* 2018. Prieiga per: <https://arxiv.org/abs/1608.06993v5> DOI 10.48550/arXiv.1608.06993.
39. OWEN, A.B. A Robust Hybrid of Lasso and Ridge Regression. 2007, vol. 443. Prieiga per: <https://artowen.su.domains/reports/hhu.pdf> DOI 10.1090/conm/443/08555.
40. NARISSETTY, N.N. Bayesian Model Selection for High-Dimensional Data. *Handbook of Statistics* 2020, vol. 43. pp. 207-248. Prieiga per: <https://www.sciencedirect.com/science/article/pii/S0169716119300380> DOI 10.1016/bs.host.2019.08.001.

41. L. BREIMAN. *Random Forests*. 2001 Prieiga per: stat.berkeley.edu.
42. KHAN, M. Y., QAYOOM, A. and NIZAMI, M. S. Automated Prediction of Good Dictionary EXamples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning, and Word Embedding-Based Deep Learning Techniques. 2021 DOI 10.1155/2021/2553199.
43. HYNDMAN, R. J. and ATHANASOPOULOS, G. *Forecasting: Principles and Practice*. 2014. pp. 291 ISSN 0987507109.
44. RDocumentation. *Mdae: Median Absolute Error*. Prieiga per: rdocumentation.org.
45. KIM, S. and KIM, H. A New Metric of Absolute Percentage Error for Intermittent Demand Forecasts. 2016, vol. 32, no. 3. pp. 669-679 DOI 10.1016/j.ijforecast.2015.12.003.
46. JAIN, G. and MALLICK, B. A Study of Time Series Models ARIMA and ETS. 2017. Prieiga per: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2898968 DOI 10.2139/ssrn.2898968.
47. DUDEK, G. Short-Term Load Forecasting using Theta Method. 2018. Prieiga per: https://www.e3s-conferences.org/articles/e3sconf/pdf/2019/10/e3sconf_pe2019_01004.pdf DOI 10.1051/e3sconf /20198401004.
48. NAMINI, S. S., TAVAKOLI, N. and NAMIN, A. S. A Comparison of ARIMA and LSTM in Forecasting Time Series. *Ieee* 2018. Prieiga per: <https://ieeexplore.ieee.org/abstract/document/8614252> DOI 10.1109/ICMLA.2018.00227.
49. OSTROM, C. W. *Time Series Analysis—Regression Techniques*. SAGE Publications 1990. pp. 7-15 DOI 10.1177/089443939501600217.
50. HARVEY, A. C., PETERS, S., Estimation procedures for structural time series models. *Journal of forecasting*. 1990, vol. 9.2, pp. 89-108 DOI 10.1002/for.3980090203
51. HASTIE, T., TIBSHIRANI, R., Generalized additive models: some applications. *Journal of the American Statistical Association*. 1987, vol. 82, pp. 371-386 DOI 10.1214/ss/1177013604
52. RDocumentation. *Prophet (Version 1.0)*. 2021 Prieiga per: <https://www.rdocumentation.org/packages/prophet/versions/1.0> DOI rdocumentation.org.
53. NECHAEV, D., et al. Cross-Sensor Nighttime Lights Image Calibration for DMSP/OLS and SNPP/VIIRS with Residual U-Net. 2021, vol. 13, no. 24. Prieiga per: <https://www.mdpi.com/2072-4292/13/24/5026> DOI 10.3390/rs13245026.
54. HUANG, Q., et al. Application of DMSP/OLS Nighttime Light Images: A Meta-Analysis and a Systematic Literature Review. 2014, vol. 6, no. 8. Prieiga per: <https://www.mdpi.com/2072-4292/6/8/6844> DOI 10.3390/rs6086844.
55. ELVIDGE, C. D., ZHIZHIN, M., BAUGH, K. and HSU, F. Why VIIRS Data are Superior to DMSP for Mapping Nighttime Lights. *Asia-Pacific Advanced Network* 35. 2013, vol. 35, pp. 62. Prieiga per: <https://eogdata.mines.edu/products/vnl/>.
56. NASA. *Earth at Night*. 2019 Prieiga per: nasa.gov.

57. LI, X., ZHOU, Y., ZHAO, M. and ZHAO, X. A Harmonized Global Nighttime Light Dataset 1992–2018. *Scientific Data* 2020 DOI 10.1038/s41597-020-0510-y.
58. NASA. *Joint Polar Satellite System (JPSS)*. Prieiga per: <https://science.nasa.gov/mission/jpss/>.
59. GISGeography. *Polar Orbit Vs Sun Synchronous Orbit*. 2024 Prieiga per: gisgeography.com.
60. ROMÁN, M. O., et al. Black Marble User Guide Version 1.2. 2021. Prieiga per: https://viirsland.gsfc.nasa.gov/PDF/BlackMarbleUserGuide_v1.2_20210421.pdf.
61. MARTY, R. and VICENTE, G.S. Illuminating Insights: Harnessing NASA's Black Marble with R and Python Packages. 2024. Prieiga per: <https://blogs.worldbank.org/en/opendata/illuminating-insights-harnessing-nasas-black-marble-r-and-python-packages>.
62. PEBESMA, E., et al. *Sf: Simple Features for R*. 2024 Prieiga per: cran.r-project.org.
63. HIJMANS, R. J., et al. *Terra: Spatial Data Analysis*. 2024 Prieiga per: cran.r-project.org.
64. MARTY, R., et al. *Blackmarbler: Black Marble Data and Statistics*. 2024 Prieiga per: cran.r-project.org.
65. HIJMANS, R. J., et al. *Raster: Geographic Data Analysis and Modeling*. 2023 Prieiga per: cran.r-project.org/.
66. RDocumentation. *writeRaster: Write Raster Data to a File*. Prieiga per: rdocumentation.org.
67. BASTON, D. *Exactextractr: Fast Extraction from Raster Datasets using Polygons*. 2023 Prieiga per: cran.r-project.org.
68. SPINU, V., et al. *Lubridate: Make Dealing with Dates a Little Easier*. 2023 Prieiga per: cran.r-project.org.
69. GUOZHU, D. and LIU, H. Feature Engineering for Machine Learning and Data Analytics. 2018 DOI <https://doi.org/10.1201/9781315181080>.
70. Keras. *Keras Applications*. Prieiga per: keras.io.
71. JANAUSKAITĖ, I. ir KUČINSKAS, K. *Vargstantys, bet nematomi – penktadalis Lietuvos gyventojų skursta, bet tarnyboms jie nematomi*. 2023 Prieiga per: <https://www.lrt.lt/naujienos/lietuvoje/2/2009446/vargstantys-bet-nematomi-penktadalis-lietuvos-gyventoju-skursta-bet-tarnyboms-jie-nematomi>

Priedai

1 priedas. Lietuvos skurdo rizikos lygis (eksperimentinė statistika)

Savivaldybė	2013 (%)	2014 (%)	2015 (%)	2016 (%)	2017 (%)	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)
Akmenės rajono savivaldybė	35,60	24,70	24,40	29,10	27,80	24,00	33,30	27,50	24,60	24,70	25,90
Alytaus miesto savivaldybė	18,30	17,40	19,00	13,90	19,90	20,70	27,30	26,40	19,70	20,80	20,30
Alytaus rajono savivaldybė	27,70	22,10	21,70	28,40	26,00	37,30	27,00	25,70	20,50	25,60	25,00
Anykščių rajono savivaldybė	34,10	33,70	31,90	30,50	33,90	37,00	29,90	28,10	29,30	26,50	26,60
Birštono savivaldybė	18,40	17,60	26,40	21,80	21,00	17,10	18,80	19,60	19,40	20,70	19,80
Biržų rajono savivaldybė	29,00	25,50	30,30	28,80	33,10	33,50	28,10	34,80	31,80	28,20	28,90
Druskininkų savivaldybė	25,80	44,00	19,90	23,60	20,00	26,20	23,20	20,30	17,10	19,80	20,10
Elektrėnų savivaldybė	18,20	25,10	20,40	16,10	18,90	24,10	20,00	17,40	16,20	17,40	16,70
Ignalinos rajono savivaldybė	34,90	27,10	31,50	26,10	28,50	33,20	26,40	31,60	28,40	28,50	26,70
Jonavos rajono savivaldybė	18,10	23,00	27,90	20,20	29,00	25,80	22,10	29,70	20,10	23,40	22,50
Joniškio rajono savivaldybė	28,50	24,20	28,20	39,20	31,60	31,00	34,10	25,10	26,20	28,10	26,00
Jurbarko rajono savivaldybė	32,80	23,80	28,30	21,10	33,80	28,70	27,00	24,50	28,10	32,40	28,70
Kaišiadorių rajono savivaldybė	17,90	17,80	26,50	28,20	33,50	18,00	24,80	21,60	19,80	23,70	21,30
Kalvarijos savivaldybė	16,10	19,00	30,60	29,00	27,40	30,40	26,70	32,10	27,40	29,10	28,00
Kauno miesto savivaldybė	12,50	12,90	18,40	14,40	13,10	11,20	16,70	15,70	17,70	17,50	16,40
Kauno rajono savivaldybė	21,10	16,10	26,70	30,10	24,90	21,60	14,10	23,40	20,60	17,00	17,80
Kazlų Rūdos savivaldybė	16,50	20,70	31,30	31,60	27,90	36,20	30,60	28,10	28,20	27,50	24,00

Kėdainių rajono savivaldybė	37,70	29,30	54,50	42,60	27,00	25,40	21,70	22,70	21,20	29,80	24,70
Kelmės rajono savivaldybė	26,90	27,60	24,40	32,80	29,70	31,10	47,50	26,50	27,80	31,10	31,30
Klaipėdos miesto savivaldybė	11,10	8,30	12,90	20,60	19,60	19,30	15,30	19,70	21,00	18,50	18,80
Klaipėdos rajono savivaldybė	33,60	19,70	19,10	13,10	27,40	18,90	20,20	17,40	18,20	19,60	19,30
Kretingos rajono savivaldybė	28,90	25,60	17,70	17,50	15,30	23,70	12,40	17,20	23,00	22,00	21,60
Kupiškio rajono savivaldybė	25,20	24,90	24,20	23,80	36,00	25,50	23,90	27,00	30,90	28,30	26,60
Lazdijų rajono savivaldybė	25,20	32,10	25,50	36,40	34,20	39,90	34,10	22,60	21,20	28,20	27,30
Marijampolės savivaldybė	17,60	21,50	28,40	24,30	28,50	36,10	29,90	25,10	24,00	22,30	22,10
Mažeikių rajono savivaldybė	27,10	19,80	20,50	26,60	33,00	34,00	28,20	18,50	31,40	28,10	25,10
Molėtų rajono savivaldybė	28,40	30,90	28,80	28,80	31,70	35,70	30,40	29,90	31,40	27,40	25,60
Neringos savivaldybė	20,70	15,70	17,20	20,40	21,90	20,90	16,60	18,60	17,80	16,60	15,50
Pagėgių savivaldybė	27,60	22,70	32,10	22,00	35,20	29,90	22,60	21,60	31,50	33,40	28,90
Pakruojo rajono savivaldybė	27,10	21,00	30,10	38,90	39,50	32,70	33,40	26,80	22,50	25,60	24,00
Palangos miesto savivaldybė	29,70	22,60	23,00	19,80	19,10	18,80	14,70	16,80	19,70	23,30	21,90
Panevėžio miesto savivaldybė	12,80	23,80	17,90	18,20	17,00	24,60	21,40	18,70	18,20	21,40	19,00
Panevėžio rajono savivaldybė	22,90	26,40	22,80	24,10	26,30	22,10	21,60	23,90	23,60	23,40	22,20
Pasvalio rajono savivaldybė	32,90	28,60	29,10	27,00	28,90	24,50	22,00	35,40	28,60	29,10	26,20
Plungės rajono savivaldybė	24,30	18,70	24,20	15,40	34,10	36,20	22,20	21,80	20,00	21,10	21,40

Prienų rajono savivaldybė	19,30	20,70	25,70	28,80	34,80	23,90	24,20	21,10	23,10	27,10	23,90
Radviliškio rajono savivaldybė	34,60	21,50	26,40	40,70	40,90	34,10	31,60	23,30	27,60	30,10	27,40
Raseinių rajono savivaldybė	19,50	18,10	28,30	24,80	37,40	26,40	32,60	23,20	26,30	26,90	26,30
Rietavo savivaldybė	28,60	20,50	22,20	20,70	26,50	30,10	19,40	19,20	26,80	24,20	22,70
Rokiškio rajono savivaldybė	19,10	33,90	33,70	33,20	36,20	34,00	29,90	33,90	25,00	27,40	25,80
Šakių rajono savivaldybė	18,40	20,30	31,00	45,00	38,70	39,70	37,60	27,70	28,40	29,10	25,10
Šalčininkų rajono savivaldybė	40,10	28,00	31,20	45,60	25,10	30,80	22,50	26,00	20,10	21,00	21,80
Šiaulių miesto savivaldybė	17,90	21,40	21,10	16,80	18,30	22,50	16,30	22,60	17,20	18,60	19,10
Šiaulių rajono savivaldybė	24,70	28,90	33,70	43,10	41,90	27,30	21,80	22,60	23,20	22,90	22,70
Šilalės rajono savivaldybė	35,60	22,80	27,90	21,80	32,10	23,40	19,80	19,20	26,40	30,40	27,40
Šilutės rajono savivaldybė	28,10	25,40	24,70	28,70	27,90	21,90	17,90	16,80	27,70	26,40	26,00
Širvintų rajono savivaldybė	20,60	17,40	19,40	27,30	27,40	25,80	17,40	15,70	22,80	21,00	19,30
Skuodo rajono savivaldybė	23,70	17,30	24,40	27,40	27,90	35,40	23,40	21,50	27,50	33,40	29,70
Švenčionių rajono savivaldybė	28,00	24,00	22,50	41,40	35,40	40,70	27,90	34,20	18,80	22,00	20,10
Tauragės rajono savivaldybė	26,10	23,30	28,70	25,00	29,80	32,60	22,60	23,00	26,40	25,50	25,30
Telšių rajono savivaldybė	37,30	24,10	24,60	25,10	28,00	27,90	14,40	19,90	25,50	26,60	25,00
Trakų rajono savivaldybė	16,90	16,60	17,60	19,10	30,20	26,00	17,50	30,10	18,50	23,80	19,10
Ukmergės rajono savivaldybė	16,60	15,80	21,10	32,10	35,20	31,60	32,50	26,90	26,10	24,10	22,00
Utenos rajono savivaldybė	26,80	27,70	30,10	23,80	30,70	30,40	25,00	27,70	28,30	23,50	21,40

Varėnos rajono savivaldybė	27,20	30,00	23,70	29,20	25,90	36,20	30,90	22,90	19,90	23,40	23,60
Vilkaviškio rajono savivaldybė	16,10	18,50	37,60	32,40	27,70	37,10	28,60	35,00	31,60	35,10	28,10
Vilniaus miesto savivaldybė	13,60	12,40	14,20	9,00	9,70	12,60	13,40	14,70	11,50	13,90	13,10
Vilniaus rajono savivaldybė	20,90	15,30	21,30	30,30	32,70	27,20	19,20	19,10	14,80	14,30	15,90
Visagino savivaldybė	26,70	27,60	28,30	24,20	26,40	30,90	28,50	27,70	30,90	24,80	24,50
Zarasų rajono savivaldybė	33,90	27,90	29,70	27,30	27,60	37,90	30,50	35,30	32,00	30,80	27,90

2 priedas. Eksperimentinių skurdo rodiklių statistika

Savivaldybė	Vidurkis	Mediana	Standartinis nuokrypis	Maksimumas	Minimumas
Akmenės rajono savivaldybė	27,42	25,90	3,70	35,60	24,00
Alytaus miesto savivaldybė	20,34	19,90	3,60	27,30	13,90
Alytaus rajono savivaldybė	26,09	25,70	4,29	37,30	20,50
Anykščių rajono savivaldybė	31,05	30,50	3,23	37,00	26,50
Birštono savivaldybė	20,05	19,60	2,42	26,40	17,10
Biržų rajono savivaldybė	30,18	29,00	2,67	34,80	25,50
Druskininkų savivaldybė	23,64	20,30	6,97	44,00	17,10
Elektrėnų savivaldybė	19,14	18,20	2,91	25,10	16,10
Ignalinos rajono savivaldybė	29,35	28,50	2,84	34,90	26,10
Jonavos rajono savivaldybė	23,80	23,00	3,66	29,70	18,10
Joniškio rajono savivaldybė	29,29	28,20	4,23	39,20	24,20
Jurbarko rajono savivaldybė	28,11	28,30	3,76	33,80	21,10
Kaišiadorių rajono savivaldybė	23,01	21,60	4,73	33,50	17,80
Kalvarijos savivaldybė	26,89	28,00	4,70	32,10	16,10
Kauno miesto savivaldybė	15,14	15,70	2,32	18,40	11,20
Kauno rajono savivaldybė	21,22	21,10	4,61	30,10	14,10
Kazlų Rūdos savivaldybė	27,51	28,10	5,20	36,20	16,50
Kėdainių rajono savivaldybė	30,60	27,00	9,88	54,50	21,20
Kelmės rajono savivaldybė	30,61	29,70	5,86	47,50	24,40
Klaipėdos miesto savivaldybė	16,83	18,80	4,09	21,00	8,30
Klaipėdos rajono savivaldybė	20,59	19,30	5,19	33,60	13,10
Kretingos rajono savivaldybė	20,45	21,60	4,64	28,90	12,40
Kupiškio rajono savivaldybė	26,94	25,50	3,51	36,00	23,80

Lazdijų rajono savivaldybė	29,70	28,20	5,74	39,90	21,20
Marijampolės savivaldybė	25,44	24,30	4,80	36,10	17,60
Mažeikių rajono savivaldybė	26,57	27,10	5,00	34,00	18,50
Molėtų rajono savivaldybė	29,91	29,90	2,52	35,70	25,60
Neringos savivaldybė	18,35	17,80	2,17	21,90	15,50
Pagėgių savivaldybė	27,95	28,90	4,76	35,20	21,60
Pakruojo rajono savivaldybė	29,24	27,10	5,98	39,50	21,00
Palangos miesto savivaldybė	20,85	19,80	3,78	29,70	14,70
Panevėžio miesto savivaldybė	19,36	18,70	3,15	24,60	12,80
Panevėžio rajono savivaldybė	23,57	23,40	1,50	26,40	21,60
Pasvalio rajono savivaldybė	28,39	28,60	3,48	35,40	22,00
Plungės rajono savivaldybė	23,58	21,80	5,96	36,20	15,40
Prienų rajono savivaldybė	24,78	23,90	4,13	34,80	19,30
Radviliškio rajono savivaldybė	30,75	30,10	6,11	40,90	21,50
Raseinių rajono savivaldybė	26,35	26,30	5,16	37,40	18,10
Rietavo savivaldybė	23,72	22,70	3,62	30,10	19,20
Rokiškio rajono savivaldybė	30,19	33,20	5,01	36,20	19,10
Šakių rajono savivaldybė	31,00	29,10	8,00	45,00	18,40
Šalčininkų rajono savivaldybė	28,38	26,00	7,78	45,60	20,10
Šiaulių miesto savivaldybė	19,25	18,60	2,17	22,60	16,30
Šiaulių rajono savivaldybė	28,44	24,70	7,44	43,10	21,80
Šilalės rajono savivaldybė	26,07	26,40	4,99	35,60	19,20
Šilutės rajono savivaldybė	24,68	26,00	3,91	28,70	16,80
Širvintų rajono savivaldybė	21,28	20,60	3,88	27,40	15,70
Skuodo rajono savivaldybė	26,51	27,40	4,98	35,40	17,30
Švenčionių rajono savivaldybė	28,64	27,90	7,74	41,40	18,80
Tauragės rajono savivaldybė	26,21	25,50	2,93	32,60	22,60
Telšių rajono savivaldybė	25,31	25,10	5,31	37,30	14,40
Trakų rajono savivaldybė	21,40	19,10	4,97	30,20	16,60
Ukmergės rajono savivaldybė	25,82	26,10	6,27	35,20	15,80
Utenos rajono savivaldybė	26,85	27,70	2,94	30,70	21,40
Varėnos rajono savivaldybė	26,63	25,90	4,43	36,20	19,90
Vilkaviškio rajono savivaldybė	29,80	31,60	6,76	37,60	16,10
Vilniaus miesto savivaldybė	12,55	13,10	1,74	14,70	9,00
Vilniaus rajono savivaldybė	21,00	19,20	6,11	32,70	14,30
Visagino savivaldybė	27,32	27,60	2,20	30,90	24,20
Zarasų rajono savivaldybė	30,98	30,50	3,33	37,90	27,30

3 priedas. *Elastic Net* modelių, mokytų be papildomų X, rezultatai

CNN architektūra	Optimali alfa reikšmė	Mokymo R^2	Testavimo R^2	RMSE	MDAE	MAPE	SMAPE	MAAPE
Metiniai spėjimai. Testavimas su 4980 56 x 56 dydžio gardelėmis								
<i>ResNet-50</i>	0,1	0,099	-0,193	3,842	2,425	15,309	13,618	14,851
<i>ResNet-101</i>	0,1	0,097	-0,242	3,968	2,695	15,856	14,042	15,358
<i>NASNetMobile</i>	0,01	0,110	-0,232	3,951	2,480	15,874	14,073	15,385
<i>EfficientNet-B0</i>	0,01	0,100	-0,244	3,998	2,510	15,952	14,106	15,438
<i>DenseNet-121</i>	0,1	0,087	-0,176	3,956	2,750	15,807	13,986	15,280
<i>DenseNet-201</i>	0,01	0,136	-0,140	3,654	2,575	14,391	12,904	14,021
Mėnesiniai spėjimai. Testavimas su 4980 56 x 56 dydžio gardelėmis								
<i>ResNet-50</i>	0,1	0,099	-0,193	3,894	2,661	15,433	13,723	14,963
<i>ResNet-101</i>	0,1	0,097	-0,242	3,993	2,613	15,884	14,060	15,379
<i>NASNetMobile</i>	0,01	0,110	-0,232	3,997	2,722	15,954	14,134	15,453
<i>EfficientNet-B0</i>	0,01	0,100	-0,244	4,035	2,667	15,998	14,133	15,472
<i>DenseNet-121</i>	0,1	0,087	-0,176	4,004	2,661	15,888	14,046	15,345
<i>DenseNet-201</i>	0,01	0,136	-0,140	3,741	2,576	14,623	13,100	14,230
Metiniai spėjimai. Testavimas su 2472 56 x 56 dydžio gardelėmis								
<i>ResNet-50</i>	0,1	0,099	-0,186	3,824	2,570	15,188	13,547	14,751
<i>ResNet-101</i>	0,1	0,097	-0,230	3,938	2,680	15,739	13,991	15,271
<i>NASNetMobile</i>	0,01	0,110	-0,250	3,981	2,585	15,974	14,152	15,476
<i>EfficientNet-B0</i>	0,01	0,100	-0,231	3,971	2,470	15,815	14,026	15,332
<i>DenseNet-121</i>	0,1	0,087	-0,151	3,930	2,630	15,638	13,866	15,138
<i>DenseNet-201</i>	0,01	0,136	-0,145	3,623	2,335	14,224	12,799	13,877
Mėnesiniai spėjimai. Testavimas su 2472 56 x 56 dydžio gardelėmis								
<i>ResNet-50</i>	0,1	0,099	-0,186	3,880	2,740	15,328	13,665	14,878
<i>ResNet-101</i>	0,1	0,097	-0,230	3,964	2,637	15,767	14,007	15,292
<i>NASNetMobile</i>	0,01	0,110	-0,250	4,033	2,738	16,072	14,228	15,560
<i>EfficientNet-B0</i>	0,01	0,100	-0,231	4,010	2,616	15,861	14,049	15,364
<i>DenseNet-121</i>	0,1	0,087	-0,151	3,981	2,709	15,743	13,948	15,227
<i>DenseNet-201</i>	0,01	0,136	-0,145	3,714	2,622	14,461	12,997	14,091
Metiniai spėjimai. Testavimas su 1440 112 x 112 dydžio gardelėmis								
<i>ResNet-50</i>	0,1	0,156	-0,113	3,732	2,425	14,858	13,278	14,466
<i>ResNet-101</i>	0,1	0,154	-0,237	3,935	2,225	15,731	13,939	15,255
<i>NASNetMobile</i>	0,1	0,091	-0,431	4,515	2,670	18,216	15,725	17,376
<i>EfficientNet-B0</i>	0,01	0,165	-0,256	4,151	2,610	16,664	14,609	16,065
<i>DenseNet-121</i>	0,1	0,123	-0,355	4,299	2,505	17,233	15,002	16,548

<i>DenseNet-201</i>	0,1	0,137	-0,288	4,196	2,620	16,766	14,671	16,134
Mėnesiniai spėjimai. Testavimas su 1440 112 x 112 dydžio gardelėmis								
<i>ResNet-50</i>	0,1	0,156	-0,113	3,804	2,570	14,971	13,356	14,554
<i>ResNet-101</i>	0,1	0,154	-0,237	3,978	2,600	15,806	13,995	15,316
<i>NASNetMobile</i>	0,1	0,091	-0,431	4,533	2,795	18,239	15,740	17,392
<i>EfficientNet-B0</i>	0,01	0,165	-0,256	4,195	2,805	16,772	14,698	16,161
<i>DenseNet-121</i>	0,1	0,123	-0,355	4,333	2,740	17,305	15,058	16,609
<i>DenseNet-201</i>	0,1	0,137	-0,288	4,242	2,750	16,871	14,753	16,219
Metiniai spėjimai. Testavimas su 708 112 x 112 dydžio gardelėmis								
<i>ResNet-50</i>	0,1	0,156	-0,204	3,730	2,340	14,895	13,327	14,498
<i>ResNet-101</i>	0,1	0,154	-0,383	3,906	2,285	15,621	13,875	15,157
<i>NASNetMobile</i>	0,1	0,091	-0,586	4,595	3,110	18,543	15,989	17,661
<i>EfficientNet-B0</i>	0,01	0,165	-0,427	4,234	2,800	16,997	14,874	16,360
<i>DenseNet-121</i>	0,1	0,123	-0,468	4,229	2,535	16,866	14,730	16,213
<i>DenseNet-201</i>	0,1	0,137	-0,442	4,207	2,630	16,682	14,595	16,052
Mėnesiniai spėjimai. Testavimas su 708 112 x 112 dydžio gardelėmis								
<i>ResNet-50</i>	0,1	0,156	-0,204	3,809	2,480	14,994	13,388	14,570
<i>ResNet-101</i>	0,1	0,154	-0,383	3,959	2,710	15,706	13,937	15,227
<i>NASNetMobile</i>	0,1	0,091	-0,586	4,618	3,165	18,595	16,033	17,707
<i>EfficientNet-B0</i>	0,01	0,165	-0,427	4,288	2,855	17,072	14,926	16,420
<i>DenseNet-121</i>	0,1	0,123	-0,468	4,270	2,790	16,935	14,780	16,269
<i>DenseNet-201</i>	0,1	0,137	-0,442	4,267	2,675	16,775	14,655	16,115
Metiniai spėjimai. Testavimas su 456 224 x 224 dydžio gardelėmis								
<i>ResNet-50</i>	0,3	0,172	-0,640	4,407	2,955	17,754	15,375	16,964
<i>ResNet-101</i>	0,3	0,190	-0,540	4,178	2,590	16,704	14,581	16,016
<i>NASNetMobile</i>	0,1	0,199	-0,718	4,501	2,755	17,963	15,491	17,082
<i>EfficientNet-B0</i>	0,1	0,158	-0,557	4,305	2,700	17,208	14,958	16,447
<i>DenseNet-121</i>	0,1	0,213	-0,601	4,271	2,940	17,133	14,934	16,426
<i>DenseNet-201</i>	0,1	0,260	-0,390	3,841	2,135	14,887	13,097	14,324
Mėnesiniai spėjimai. Testavimas su 456 224 x 224 dydžio gardelėmis								
<i>ResNet-50</i>	0,3	0,172	-0,640	4,440	2,990	17,800	15,405	16,998
<i>ResNet-101</i>	0,3	0,190	-0,540	4,243	2,735	16,836	14,688	16,129
<i>NASNetMobile</i>	0,1	0,199	-0,718	4,538	2,865	18,064	15,577	17,170
<i>EfficientNet-B0</i>	0,1	0,158	-0,557	4,336	2,825	17,273	15,011	16,503
<i>DenseNet-121</i>	0,1	0,213	-0,601	4,343	2,925	17,306	15,075	16,575
<i>DenseNet-201</i>	0,1	0,260	-0,390	3,963	2,350	15,249	13,399	14,642
Metiniai spėjimai. Testavimas su 228 224 x 224 dydžio gardelėmis								
<i>ResNet-50</i>	0,3	0,172	-0,932	4,716	3,170	18,919	16,245	17,967
<i>ResNet-101</i>	0,3	0,190	-0,659	4,327	2,780	17,245	15,042	16,512

<i>NASNetMobile</i>	0,1	0,199	-0,901	4,603	2,945	18,155	15,626	17,245
<i>EfficientNet-B0</i>	0,1	0,158	-0,726	4,451	2,670	17,610	15,247	16,772
<i>DenseNet-121</i>	0,1	0,213	-0,832	4,503	2,995	17,885	15,475	17,064
<i>DenseNet-201</i>	0,1	0,260	-0,502	3,933	2,210	14,975	13,188	14,385
Mėnesiniai spėjimai. Testavimas su 228 224 x 224 dydžio gardelėmis								
<i>ResNet-50</i>	0,3	0,172	-0,932	4,746	3,190	19,007	16,317	18,044
<i>ResNet-101</i>	0,3	0,190	-0,659	4,407	2,935	17,478	15,245	16,720
<i>NASNetMobile</i>	0,1	0,199	-0,901	4,659	2,980	18,362	15,808	17,433
<i>EfficientNet-B0</i>	0,1	0,158	-0,726	4,489	2,675	17,743	15,365	16,893
<i>DenseNet-121</i>	0,1	0,213	-0,832	4,594	3,040	18,181	15,728	17,329
<i>DenseNet-201</i>	0,1	0,260	-0,502	4,071	2,430	15,569	13,724	14,933

4 priedas. *Elastic Net* modelio, mokyto su *DenseNet-201* vektorizavimo būdu gautais požymiais, spėjimų rezultatai. Prognozavimas atliktas iš 224 x 224 dydžio vaizdų su 456 gardelių testavimo imtimi (10 prасčiausių metinių prognozių)

Savivaldybė	Skurdo rodiklis (%)	Prognozė (%)	MDAE
Vilniaus miesto savivaldybė	13,1	23,26	10,16
Vilniaus rajono savivaldybė	15,9	23,80	7,90
Neringos savivaldybė	15,5	23,01	7,51
Kauno miesto savivaldybė	16,4	23,19	6,79
Elektrėnų savivaldybė	16,7	23,45	6,75
Šiaulių miesto savivaldybė	19,1	26,74	7,64
Kauno rajono savivaldybė	17,8	24,61	6,81
Švenčionių rajono savivaldybė	20,1	26,95	6,85
Panevėžio miesto savivaldybė	19	25,32	6,32
Trakų rajono savivaldybė	19,1	24,50	5,40

5 priedas. Skirtingais metodais sudarytų modelių palyginimas su faktinėmis reikšmėmis

Savivaldybė	Oficialus skurdo rodiklis iš OSP	2023 prognozė, gauta su <i>DenseNet-201</i> CNN architektūra bei taikant <i>Elastic Net</i> CV regresijos metodą	2023 prognozė, gauta su <i>ResNet-50</i> CNN architektūra ir papildomais X bei taikant <i>Elastic Net</i> CV regresijos metodą	2023 prognozė, gauta su <i>DenseNet-121</i> CNN architektūra bei taikant atsitiktinių miškų metodą	2023 prognozė, gauta su <i>NASNetMobile</i> CNN architektūra ir papildomais X bei taikant atsitiktinių miškų metodą
Akmenės rajono savivaldybė	25,90	26,55	25,35	26,61	24,88
Alytaus miesto savivaldybė	20,30	22,41	23,66	23,98	22,06

Alytaus rajono savivaldybė	25,00	25,76	24,63	26,16	23,52
Anykščių rajono savivaldybė	26,60	26,95	27,65	26,99	28,63
Birštono savivaldybė	19,80	23,83	23,74	24,74	22,8
Biržų rajono savivaldybė	28,90	27,01	27,85	27,19	29,28
Druskininkų savivaldybė	20,10	25,92	24,67	26,12	22,13
Elektrėnų savivaldybė	16,70	23,90	20,56	23,88	21,15
Ignalinos rajono savivaldybė	26,70	26,59	27,37	26,56	27,32
Jonavos rajono savivaldybė	22,50	24,43	23,18	24,39	23,47
Joniškio rajono savivaldybė	26,00	27,08	27,41	27,11	24,77
Jurbarko rajono savivaldybė	28,70	26,53	26,76	26,87	25,49
Kaišiadorių rajono savivaldybė	21,30	25,10	22,11	25,02	22,74
Kalvarijos savivaldybė	28,00	24,94	25,98	25,25	27,45
Kauno miesto savivaldybė	16,40	19,52	17,49	19,6	21,36
Kauno rajono savivaldybė	17,80	23,13	21,84	23,27	23,26
Kazlų Rūdos savivaldybė	24,00	25,94	25,57	25,79	25,37
Kėdainių rajono savivaldybė	24,70	25,76	26,5	25,9	28,92
Kelmės rajono savivaldybė	31,30	26,92	27,86	27,02	26,59
Klaipėdos miesto savivaldybė	18,80	20,28	18,64	20,46	20,06
Klaipėdos rajono savivaldybė	19,30	24,24	20,01	24,44	20,99
Kretingos rajono savivaldybė	21,60	25,98	20,94	26,04	22,84
Kupiškio rajono savivaldybė	26,60	26,72	26,33	27,01	28,45
Lazdijų rajono savivaldybė	27,30	26,62	27,13	26,59	28,56
Marijampolės savivaldybė	22,10	25,04	24,66	24,82	24,63
Mažeikių rajono savivaldybė	25,10	26,41	25,32	26,4	21,26
Molėtų rajono savivaldybė	25,60	26,26	26,91	26,51	27,72
Neringos savivaldybė	15,50	23,03	21,52	23,45	22,51
Pagėgių savivaldybė	28,90	25,14	26,34	25,85	24,37
Pakruojo rajono savivaldybė	24,00	26,63	27,3	26,9	23,93
Palangos miesto savivaldybė	21,90	22,70	19,8	22,95	22,31
Panevėžio miesto savivaldybė	19,00	22,88	22,15	23,84	23,92
Panevėžio rajono savivaldybė	22,20	25,91	24,18	26,26	25,61
Pasvalio rajono savivaldybė	26,20	26,60	26,17	26,96	27,16
Plungės rajono savivaldybė	21,40	27,04	23,11	27,06	22,98
Prienų rajono savivaldybė	23,90	24,97	23,75	25,28	23,59
Radviliškio rajono savivaldybė	27,40	26,55	27,96	26,55	23,78
Raseinių rajono savivaldybė	26,30	26,61	25,95	26,9	23,95

Rietavo savivaldybė	22,70	27,27	23,93	27,15	23,02
Rokiškio rajono savivaldybė	25,80	27,15	28,18	26,96	27,85
Šakių rajono savivaldybė	25,10	26,51	28,75	26,8	28,43
Šalčininkų rajono savivaldybė	21,80	26,08	24,39	26,25	21,48
Šiaulių miesto savivaldybė	19,10	23,43	25,14	24,03	23,87
Šiaulių rajono savivaldybė	22,70	26,12	26,42	26,33	24,73
Šilalės rajono savivaldybė	27,40	27,24	25,32	27,28	24,03
Šilutės rajono savivaldybė	26,00	26,21	23,56	26,7	24,16
Širvintų rajono savivaldybė	19,30	25,73	21,3	25,8	21,29
Skuodo rajono savivaldybė	29,70	27,55	24,88	27,5	24,48
Švenčionių rajono savivaldybė	20,10	26,66	26,35	26,81	24,19
Tauragės rajono savivaldybė	25,30	26,62	25,75	26,78	24,13
Telšių rajono savivaldybė	25,00	27,04	25,17	27,06	22,75
Trakų rajono savivaldybė	19,10	25,00	21,6	25,17	22,79
Ukmergės rajono savivaldybė	22,00	26,49	24,51	26,47	24,95
Utenos rajono savivaldybė	21,40	26,34	26,02	26,42	25,80
Varėnos rajono savivaldybė	23,60	26,59	24,44	26,78	23,62
Vilkaviškio rajono savivaldybė	28,10	26,17	28,94	26,21	30,78
Vilniaus miesto savivaldybė	13,10	16,08	15,71	17,07	15,00
Vilniaus rajono savivaldybė	15,90	23,40	19,62	23,84	17,35
Visagino savivaldybė	24,50	24,71	27,67	25,2	29,48
Zarasų rajono savivaldybė	27,90	26,35	28,39	26,52	29,83

6 priedas. *Elastic Net* modelių, mokytų su papildomais X, rezultatai

CNN architektūra	L1	Optimali alfa reikšmė	Mokymo R ²	Testavimo R ²	RMSE	MDAE	MAPE	SMAPE	MAAPE
Metiniai spėjimai. Testuota su 2472 56 x 56 dydžio gardelėmis									
<i>ResNet-50</i>	0,5	0,0001	0,474	0,475	2,696	1,920	10,006	9,345	9,877
<i>ResNet-101</i>	0,5	0,0001	0,474	0,456	2,788	1,820	10,385	9,657	10,241
<i>NASNetMobile</i>	0,5	0,0001	0,470	0,440	2,893	1,885	10,989	10,159	10,821
<i>EfficientNet-B0</i>	0,7	0,0001	0,464	0,477	2,846	1,860	10,829	10,029	10,668
<i>DenseNet-121</i>	0,5	0,0001	0,473	0,468	2,755	1,870	10,297	9,596	10,157
<i>DenseNet-201</i>	0,5	0,0001	0,475	0,464	2,718	1,945	10,233	9,550	10,102
Mėnesiniai spėjimai. Testuota su 2472 56 x 56 dydžio gardelėmis									
<i>ResNet-50</i>	0,5	0,0001	0,474	0,475	2,733	1,915	10,156	9,486	10,022
<i>ResNet-101</i>	0,5	0,0001	0,474	0,456	2,817	1,894	10,498	9,762	10,350

<i>NASNetMobile</i>	0,5	0,0001	0,470	0,440	2,903	1,987	11,036	10,204	10,867
<i>EfficientNet-B0</i>	0,7	0,0001	0,464	0,477	2,854	1,903	10,869	10,066	10,706
<i>DenseNet-121</i>	0,5	0,0001	0,473	0,468	2,785	1,923	10,413	9,704	10,268
<i>DenseNet-201</i>	0,5	0,0001	0,475	0,464	2,746	1,932	10,348	9,659	10,214
Metiniai spėjimai. Mokyta ir testuota naudojant nepersidengiančius vaizdus									
<i>ResNet-50</i>	0,7	0,0001	0,469	0,510	2,636	1,960	9,811	9,179	9,685
<i>ResNet-101</i>	0,5	0,0001	0,478	0,476	2,688	1,855	10,045	9,390	9,915
<i>NASNetMobile</i>	0,5	0,0001	0,471	0,466	2,795	1,915	10,568	9,823	10,421
<i>EfficientNet-B0</i>	0,5	0,0001	0,466	0,455	2,849	1,885	10,780	9,991	10,621
<i>DenseNet-121</i>	0,5	0,0001	0,473	0,480	2,706	1,940	10,083	9,431	9,951
<i>DenseNet-201</i>	0,7	0,0001	0,468	0,503	2,692	1,945	10,136	9,464	10,003
Mėnesiniai spėjimai. Mokyta ir testuota naudojant nepersidengiančius vaizdus									
<i>ResNet-50</i>	0,7	0,0001	0,469	0,510	2,675	1,877	9,962	9,321	9,832
<i>ResNet-101</i>	0,5	0,0001	0,478	0,476	2,718	1,859	10,127	9,465	9,993
<i>NASNetMobile</i>	0,5	0,0001	0,471	0,466	2,814	1,978	10,629	9,878	10,479
<i>EfficientNet-B0</i>	0,5	0,0001	0,466	0,455	2,859	1,897	10,825	10,033	10,664
<i>DenseNet-121</i>	0,5	0,0001	0,473	0,480	2,738	1,919	10,229	9,568	10,092
<i>DenseNet-201</i>	0,7	0,0001	0,468	0,503	2,722	1,939	10,267	9,588	10,130

7 priedas. *Elastic Net* modelio, mokyto su *ResNet-50* vektorizavimo būdu gautais požymiais ir papildomais X, mėnesinės prognozės

Savivaldybė	2023-01 (%)	2023-02 (%)	2023-03 (%)	2023-04 (%)	2023-05 (%)	2023-06 (%)	2023-07 (%)	2023-08 (%)	2023-09 (%)	2023-10 (%)	2023-11 (%)	2023-12 (%)
Akmenės rajono savivaldybė	25,57	25,71	24,91	25,59	25,94	25,29	24,42	24,46	25,33	25,50	25,65	25,83
Alytaus miesto savivaldybė	24,04	23,98	24,00	23,24	22,94	23,70	23,37	23,11	23,59	24,09	23,54	24,32
Alytaus rajono savivaldybė	24,63	24,81	24,76	24,34	24,38	24,57	24,21	23,93	24,51	25,17	25,01	25,21
Anykščių rajono savivaldybė	27,19	28,22	28,03	26,89	26,81	27,41	27,55	27,73	27,76	28,07	27,99	28,11
Birštono savivaldybė	24,07	23,89	23,77	23,17	23,25	23,82	23,36	23,15	23,73	24,09	24,00	24,61
Biržų rajono savivaldybė	27,93	27,89	28,39	27,17	26,94	27,30	27,97	27,96	28,15	28,31	28,10	28,10
Druskininkų savivaldybė	25,05	25,40	24,81	23,97	24,59	25,04	24,58	23,61	24,70	25,08	24,58	24,61
Elektrėnų savivaldybė	20,99	20,34	20,46	20,27	20,23	20,49	20,40	20,59	20,47	20,76	20,88	20,81
Ignalinos rajono savivaldybė	27,89	27,64	27,28	27,01	26,69	27,06	26,90	26,86	27,42	27,93	27,69	28,10

Jonavos rajono savivaldybė	23,63	23,23	22,99	22,50	22,77	23,24	22,98	22,57	23,31	23,60	23,63	23,70
Joniškio rajono savivaldybė	27,53	27,60	27,45	27,29	28,05	27,44	26,64	26,47	27,51	27,96	27,52	27,48
Jurbarko rajono savivaldybė	26,63	27,36	26,50	26,62	26,71	26,64	26,45	26,27	26,49	26,93	27,02	27,44
Kaišiadorių rajono savivaldybė	22,39	21,65	22,01	21,71	21,60	22,15	22,07	21,91	22,33	22,42	22,55	22,47
Kalvarijos savivaldybė	26,66	25,89	25,59	25,73	25,64	25,72	25,57	25,70	25,92	26,59	26,06	26,67
Kauno miesto savivaldybė	18,82	17,71	17,26	16,95	16,59	17,09	17,38	16,23	17,68	17,71	18,46	18,01
Kauno rajono savivaldybė	22,42	21,96	21,51	21,32	21,46	21,69	21,47	21,13	21,77	21,90	22,69	22,74
Kazlų Rūdos savivaldybė	25,67	25,79	25,47	25,36	25,81	25,53	25,27	25,03	25,34	25,59	25,81	26,15
Kėdainių rajono savivaldybė	26,79	26,66	26,12	25,72	26,18	26,49	26,28	26,03	26,70	27,11	26,94	27,01
Kelmės rajono savivaldybė	28,14	28,35	28,24	27,52	27,46	27,30	27,10	27,17	27,91	28,83	28,08	28,22
Klaipėdos miesto savivaldybė	19,40	19,08	18,89	18,03	18,00	18,03	18,73	18,21	18,26	18,90	18,33	19,80
Klaipėdos rajono savivaldybė	20,44	20,49	20,15	20,01	19,86	19,55	19,40	19,30	19,97	20,17	20,41	20,38
Kretingos rajono savivaldybė	21,27	21,25	20,45	21,00	21,08	20,72	20,51	20,33	20,98	21,30	21,47	20,94
Kupiškio rajono savivaldybė	26,32	26,32	26,77	25,39	25,16	25,40	26,58	27,02	26,65	26,99	26,47	26,91
Lazdijų rajono savivaldybė	27,09	27,28	27,30	26,85	27,33	27,21	26,90	26,53	26,55	27,84	27,26	27,36
Marijampolės savivaldybė	25,25	24,80	24,51	24,43	24,47	24,02	24,15	24,18	24,66	25,09	24,96	25,34
Mažeikių rajono savivaldybė	25,77	25,25	25,01	25,40	25,82	25,12	24,62	24,45	25,48	25,73	25,57	25,69
Molėtų rajono savivaldybė	26,64	27,30	26,91	26,50	26,57	26,75	26,69	26,62	26,92	27,27	27,36	27,41
Neringos savivaldybė	21,76	21,92	21,53	21,32	21,35	21,38	21,52	21,15	21,40	21,80	21,19	21,94
Pagėgių savivaldybė	26,51	26,82	25,42	26,39	27,21	26,56	25,94	25,48	26,64	26,68	26,21	26,17
Pakruojo rajono savivaldybė	27,61	27,73	27,41	26,79	26,92	27,18	26,69	26,93	27,46	27,92	27,13	27,87
Palangos miesto savivaldybė	19,76	19,97	19,59	19,52	19,72	19,82	20,02	19,68	19,09	20,93	19,63	19,91

Panevėžio miesto savivaldybė	22,67	20,84	22,20	21,82	21,61	21,68	22,00	22,52	21,76	22,53	23,23	22,94
Panevėžio rajono savivaldybė	24,39	24,02	24,35	23,46	23,40	23,77	24,34	24,30	24,29	24,81	24,51	24,55
Pasvalio rajono savivaldybė	26,43	26,27	26,54	25,33	25,22	25,46	26,37	26,54	26,35	26,79	26,29	26,42
Plungės rajono savivaldybė	23,45	22,75	23,04	23,23	23,55	22,57	22,47	22,34	23,36	23,51	23,73	23,28
Prienų rajono savivaldybė	23,93	23,92	23,37	23,22	23,58	23,94	23,31	23,38	23,73	24,02	24,26	24,40
Radviliškio rajono savivaldybė	27,94	28,24	28,10	27,67	27,52	27,87	27,40	27,55	27,89	28,70	28,30	28,36
Raseinių rajono savivaldybė	26,17	26,52	25,54	25,61	25,65	25,53	25,55	25,50	25,71	26,53	26,36	26,69
Rietavo savivaldybė	23,97	24,19	23,66	24,07	24,25	23,28	23,06	23,07	24,51	24,43	24,37	24,33
Rokiškio rajono savivaldybė	28,05	28,60	28,38	27,69	27,46	27,70	27,67	27,86	28,35	28,90	28,89	28,58
Šakių rajono savivaldybė	28,92	29,19	28,47	28,67	28,76	28,50	28,11	28,11	28,49	29,18	29,11	29,53
Šalčininkų rajono savivaldybė	24,42	24,52	24,00	24,04	24,16	24,39	24,27	24,27	24,32	24,69	24,92	24,72
Šiaulių miesto savivaldybė	24,50	24,87	25,44	24,84	24,85	25,05	24,10	24,48	25,41	25,47	26,39	26,24
Šiaulių rajono savivaldybė	26,66	26,66	26,83	26,33	26,66	26,19	25,63	25,36	26,37	26,74	26,76	26,84
Šilalės rajono savivaldybė	25,07	25,73	24,95	25,57	25,72	24,73	24,48	24,48	25,71	25,68	25,74	25,92
Šilutės rajono savivaldybė	23,58	23,98	23,27	23,79	23,77	23,14	23,18	23,07	23,68	23,87	23,75	23,63
Širvintų rajono savivaldybė	21,92	21,51	21,50	20,72	20,36	21,12	20,67	20,83	21,42	21,97	21,79	21,76
Skuodo rajono savivaldybė	25,19	24,82	24,65	25,12	25,19	24,39	24,22	24,33	25,17	25,00	25,46	24,99
Švenčionių rajono savivaldybė	26,55	26,81	26,24	26,05	25,83	26,08	26,17	25,97	25,99	26,75	26,81	26,90
Tauragės rajono savivaldybė	25,65	26,19	25,65	25,80	26,01	25,44	25,12	24,93	25,87	26,11	26,06	26,17
Telšių rajono savivaldybė	25,18	25,50	24,76	25,50	25,55	24,59	24,42	24,33	25,58	25,60	25,44	25,59
Trakų rajono savivaldybė	22,08	21,35	21,15	21,26	21,50	21,55	21,56	21,43	21,30	21,91	22,21	21,93

Ukmergės rajono savivaldybė	24,73	24,92	24,59	23,94	24,07	24,38	24,12	24,10	24,54	25,09	24,79	24,87
Utenos rajono savivaldybė	25,91	26,66	26,21	25,46	25,48	25,89	25,38	25,24	26,21	26,53	26,71	26,59
Varėnos rajono savivaldybė	24,65	24,48	24,50	23,90	24,23	24,50	23,97	23,98	24,25	24,99	25,03	24,81
Vilkaviškio rajono savivaldybė	29,33	28,94	28,33	28,81	28,97	28,92	28,48	28,48	28,71	29,52	29,28	29,56
Vilniaus miesto savivaldybė	16,68	15,28	15,78	15,64	14,96	15,42	15,03	15,52	15,69	16,01	16,74	15,83
Vilniaus rajono savivaldybė	20,28	19,78	19,24	18,91	19,05	19,48	19,33	19,26	19,61	19,78	20,37	20,33
Visagino savivaldybė	27,69	28,31	28,28	27,48	26,97	27,04	26,55	26,53	27,17	28,29	29,27	28,48
Zarasų rajono savivaldybė	28,47	28,92	28,57	28,14	27,78	27,97	27,79	27,71	28,31	28,97	28,92	29,15

8 priedas. Atsitiktinių miškų modelių, mokytų be papildomų X, rezultatai

CNN architektūra	Mokymo R ²	Testavimo R ²	RMSE	MDAE	MAPE	SMAPE	MAAPE
Metiniai spėjimai. Testuota su 2472 56 x 56 dydžio gardelėmis. Taikytas <i>max_features</i> ="sqrt" parametras							
<i>ResNet-50</i>	0,886	0,210	3,852	2,500	15,348	13,691	14,929
<i>ResNet-101</i>	0,885	0,197	3,724	2,260	14,786	13,247	14,400
<i>NASNetMobile</i>	0,882	0,149	4,033	2,760	16,214	14,368	15,721
<i>EfficientNet-B0</i>	0,887	0,210	3,695	2,430	14,675	13,150	14,301
<i>DenseNet-121</i>	0,886	0,165	3,523	2,400	13,802	12,462	13,474
<i>DenseNet-201</i>	0,887	0,183	3,544	2,510	13,942	12,587	13,616
Mėnesiniai spėjimai. Testuota su 2472 56 x 56 dydžio gardelėmis. Taikytas <i>max_features</i> ="sqrt" parametras							
<i>ResNet-50</i>	0,886	0,210	3,890	2,700	15,444	13,772	15,016
<i>ResNet-101</i>	0,885	0,197	3,795	2,700	15,020	13,458	14,619
<i>NASNetMobile</i>	0,882	0,149	4,015	2,830	16,028	14,227	15,543
<i>EfficientNet-B0</i>	0,887	0,210	3,770	2,700	14,889	13,338	14,499
<i>DenseNet-121</i>	0,886	0,165	3,661	2,670	14,272	12,891	13,918
<i>DenseNet-201</i>	0,887	0,183	3,656	2,685	14,281	12,894	13,934
Metiniai spėjimai. Testuota su 2472 56 x 56 dydžio gardelėmis. Taikytas <i>max_features</i> ="log" parametras							
<i>ResNet-50</i>	0,883	0,185	3,821	2,305	15,202	13,563	14,778
<i>ResNet-101</i>	0,884	0,201	3,902	2,490	15,592	13,878	15,149
<i>NASNetMobile</i>	0,881	0,142	4,102	2,795	16,544	14,608	16,011
<i>EfficientNet-B0</i>	0,886	0,199	3,753	2,450	14,945	13,363	14,548
<i>DenseNet-121</i>	0,885	0,163	3,578	2,485	14,081	12,688	13,733
<i>DenseNet-201</i>	0,885	0,180	3,618	2,535	14,304	12,873	13,952

Mėnesiniai spėjimai. Testuota su 2472 56 x 56 dydžio gardelėmis. Taikytas <i>max_features="log"</i> parametras							
<i>ResNet-50</i>	0,883	0,185	3,876	2,660	15,379	13,719	14,941
<i>ResNet-101</i>	0,884	0,201	3,930	2,645	15,659	13,935	15,210
<i>NASNetMobile</i>	0,881	0,142	4,062	2,815	16,276	14,407	15,763
<i>EfficientNet-B0</i>	0,886	0,199	3,817	2,620	15,122	13,519	14,712
<i>DenseNet-121</i>	0,885	0,163	3,704	2,665	14,491	13,059	14,118
<i>DenseNet-201</i>	0,885	0,180	3,715	2,695	14,598	13,140	14,227

9 priedas. Atsitiktinių miškų modelių, mokytų su papildomais X, rezultatai

CNN architektūra	Testavimo R ²	RMSE	MDAE	MAPE	SMAPE	MAAPE
Metiniai spėjimai. Testuota su 2472 56 x 56 dydžio gardelėmis						
<i>ResNet-50</i>	0,342	2,938	2,105	11,164	10,447	11,003
<i>ResNet-101</i>	0,371	2,888	2,085	11,015	10,367	10,863
<i>NASNetMobile</i>	0,378	2,884	2,060	10,938	10,287	10,784
<i>EfficientNet-B0</i>	0,377	2,885	2,035	11,010	10,358	10,860
<i>DenseNet-121</i>	0,378	2,899	2,145	11,083	10,368	10,925
<i>DenseNet-201</i>	0,339	2,942	2,145	11,166	10,532	11,012
Mėnesiniai spėjimai. Testuota su 2472 56 x 56 dydžio gardelėmis						
<i>ResNet-50</i>	0,342	2,940	2,135	11,164	10,445	11,002
<i>ResNet-101</i>	0,371	2,889	2,062	11,014	10,365	10,862
<i>NASNetMobile</i>	0,378	2,887	2,055	10,940	10,287	10,785
<i>EfficientNet-B0</i>	0,377	2,887	2,060	11,008	10,356	10,858
<i>DenseNet-121</i>	0,378	2,903	2,137	11,089	10,372	10,931
<i>DenseNet-201</i>	0,339	2,945	2,109	11,174	10,539	11,020

10 priedas. Visų savivaldybių prognozės 2024 m. sausio–balandžio mėn., gautos taikant geriausią prognozavimo metodą

Savivaldybė	2024-01	2024-02	2024-03	2024-04
Akmenės rajono savivaldybė	24,93	25,89	23,24	24,29
Alytaus miesto savivaldybė	25,39	17,57	19,94	24,98
Alytaus rajono savivaldybė	23,11	23,20	23,53	27,61
Anykščių rajono savivaldybė	24,28	26,39	23,10	25,33
Birštono savivaldybė	25,15	20,94	21,78	25,58
Biržų rajono savivaldybė	25,76	25,36	26,08	24,79
Druskininkų savivaldybė	24,82	24,52	23,77	27,43
Elektrėnų savivaldybė	25,96	23,12	25,2	25,03
Ignalinos rajono savivaldybė	22,90	22,48	26,39	23,79

Jonavos rajono savivaldybė	25,87	23,16	26,90	25,00
Joniškio rajono savivaldybė	23,71	23,66	27,33	25,10
Jurbarko rajono savivaldybė	26,18	24,43	24,41	27,10
Kaišiadorių rajono savivaldybė	25,76	24,18	25,69	26,26
Kalvarijos savivaldybė	21,45	27,33	21,35	27,51
Kauno miesto savivaldybė	24,95	22,27	26,74	22,51
Kauno rajono savivaldybė	25,93	23,12	26,09	25,76
Kazlų Rūdos savivaldybė	26,16	23,48	26,77	29,19
Kėdainių rajono savivaldybė	24,43	23,97	27,67	23,48
Kelmės rajono savivaldybė	24,49	23,05	27,24	24,16
Klaipėdos miesto savivaldybė	20,82	21,48	26,44	28,03
Klaipėdos rajono savivaldybė	26,44	24,08	25,66	25,40
Kretingos rajono savivaldybė	24,66	22,99	23,78	25,99
Kupiškio rajono savivaldybė	22,10	25,43	23,59	25,17
Lazdijų rajono savivaldybė	24,80	26,13	22,82	25,23
Marijampolės savivaldybė	24,70	24,96	24,23	27,80
Mažeikių rajono savivaldybė	23,28	26,10	27,16	26,38
Molėtų rajono savivaldybė	23,55	22,78	25,57	23,77
Neringos savivaldybė	22,69	24,73	26,49	28,79
Pagėgių savivaldybė	21,67	23,42	25,93	26,03
Pakruojo rajono savivaldybė	24,76	24,22	23,29	26,14
Palangos miesto savivaldybė	22,38	20,77	25,16	26,82
Panevėžio miesto savivaldybė	19,80	26,60	24,53	29,75
Panevėžio rajono savivaldybė	23,17	26,04	26,82	26,21
Pasvalio rajono savivaldybė	24,38	26,07	25,84	26,53
Plungės rajono savivaldybė	25,55	25,91	25,99	25,86
Prienų rajono savivaldybė	25,70	23,36	26,00	26,95
Radviliškio rajono savivaldybė	24,81	22,71	26,50	26,05
Raseinių rajono savivaldybė	24,74	23,48	26,41	26,35
Rietavo savivaldybė	25,28	24,08	26,42	23,17
Rokiškio rajono savivaldybė	24,21	23,38	26,41	26,16
Šakių rajono savivaldybė	24,97	24,70	23,31	26,24
Šalčininkų rajono savivaldybė	23,92	24,70	24,12	25,07

Šiaulių miesto savivaldybė	24,71	25,67	28,28	27,22
Šiaulių rajono savivaldybė	24,95	23,91	26,47	25,88
Šilalės rajono savivaldybė	25,80	23,15	26,37	24,50
Šilutės rajono savivaldybė	24,95	24,60	27,89	26,61
Širvintų rajono savivaldybė	23,26	23,91	24,88	23,25
Skuodo rajono savivaldybė	25,26	25,57	25,09	23,98
Švenčionių rajono savivaldybė	22,68	23,38	25,52	24,21
Tauragės rajono savivaldybė	24,39	24,39	24,61	25,41
Telšių rajono savivaldybė	24,55	24,28	25,35	24,54
Trakų rajono savivaldybė	25,64	20,96	23,07	23,42
Ukmergės rajono savivaldybė	23,57	25,52	26,39	26,38
Utenos rajono savivaldybė	23,47	23,93	26,04	25,21
Varėnos rajono savivaldybė	23,60	23,52	26,38	27,39
Vilkaviškio rajono savivaldybė	22,16	24,15	24,93	26,83
Vilniaus miesto savivaldybė	23,85	23,86	27,02	25,31
Vilniaus rajono savivaldybė	25,62	23,69	26,11	25,09
Visagino savivaldybė	26,01	22,95	27,74	19,06
Zarasų rajono savivaldybė	25,85	24,95	26,07	24,04