



Kauno technologijos universitetas

Informatikos fakultetas

**Artimų kliūčių aptikimo realiuoju laiku iš RGB vaizdų
tyrimas skirtas autonominių bepiločių navigacijai**

Baigiamasis magistro projektas

Eimantas Matuolis

Projekto autorius

Prof. Dr. Andrius Kriščiūnas

Vadovas

Kaunas, 2026



Kauno technologijos universitetas

Informatikos fakultetas

**Artimų kliūčių aptikimo realiuoju laiku iš RGB vaizdų
tyrimas skirtas autonominių bepiločių navigacijai**

Baigiamasis magistro projektas

Dirbtinio intelekto informatika (6211BX007)

Eimantas Matuolis

Projekto autorius

Prof. Dr. Andrius Kriščiūnas

Vadovas

Asist. Dr. Audrius Nečiūnas

Recenzentas

Kaunas, 2026



Kauno technologijos universitetas

Informatikos fakultetas

Eimantas Matuolis

Artimų kliūčių aptikimo realiuoju laiku iš RGB vaizdų tyrimas skirtas autonominių bepiločių navigacijai

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.

Eimantas Matuolis

Patvirtinta elektroniniu būdu

Matuolis Eimantas. Artimų kliūčių aptikimo realiuoju laiku iš RGB vaizdų tyrimas skirtas autonominių bepiločių navigacijai. Magistro baigiamasis projektas / vadovas prof. Andrius Kriščiūnas; Kauno technologijos universitetas, informatikos fakultetas.

Studijų kryptis ir sritis (studijų krypties grupė): Informatikos mokslai, Informatika (B01).

Reikšminiai žodžiai: Objektų vengimas, autonominiai bepiločiai orlaiviai, gylio įvertinimas iš RGB vaizdų, realaus laiko kompiuterinė rega.

Kaunas, 2026. 54 p.

Santrauka

Šiame tyrime yra nagrinėjami monokuliniai gylio įvertinimo metodai, kurie gali būti skirti autonominei bepiločių orlaivių objektų vengimo sistemai realiuoju laiku. Išanalizavus ir eksperimentiškai įvertinus skirtingus gylio įvertinimo modelius, daugiausia dėmesio yra skiriama „Depth Anything V2“. Norint įvertinti kliūčių vengimo algoritmų kokybę, yra nustatomos kelios metrikos kaip atstumo įvertinimo paklaida, susidūrimų dažnis, apdorojimo laikas ir resursų naudojimas realiu laiku. Visos šios metrikos yra įvertintos su trijų tipų modeliais, pagrįstais „Depth Anything V2“: baziniais, paprastai papildomai mokymais su simuliacijoje sudarytais vaizdais ir sukonzentruotais į artimus atstumus, sukurtais eksperimentuose. Eksperimentiniai rezultatai rodo, kad sukonzentruoti artimiems atstumams maži modeliai pasiekia stabilesnes gylio prognozes, mažesnę susidūrimų dažnį ir geresnę tikslumą realiuoju laiku, lyginant su didesniais modeliais, reikalaujančiais daugiau laiko ir kompiuterinių resursų apdorojant vaizdus.

Matuolis Eimantas. Investigation of Real-Time Obstacle Detection from RGB Images for Autonomous UAV Navigation. Master's Final Degree Project / supervisor prof. Andrius Kriščiūnas; Faculty of Informatics, Kaunas University of Technology.

Study field and area (study field group): Computer Science, Informatics (B01).

Keywords: Object avoidance, unmanned aerial vehicles, depth estimation from RGB images, real time computer vision.

Kaunas, 2026. 54 p.

Summary

The study investigates monocular depth estimation methods that can be used for real-time autonomous object avoidance systems for unmanned aerial vehicles. After analyzing and experimenting different depth estimation models, the main focus is on the Depth Anything V2 models. In order to evaluate the quality of obstacle avoidance models, several metrics are determined, such as distance estimation errors, collision rate, processing time and resource usage in real time. All of these metrics were evaluated using three types of models based on Depth Anything V2: base, trained on images created in the simulation, and short range models developed in the experiments. Experimental results show that small models focused on short distances achieve more stable depth predictions, lower collision rates and better accuracy in real time compared to larger models that require more time and computing resources for image processing.

Turinys

Lentelių sąrašas	7
Paveikslų sąrašas	8
Įvadas.....	10
1. Bepiločio orlaivio kompiuterinės vizijos sistemų analizė.....	12
1.1. Bepiločių orlaivių navigacijos analizė naudojant tik RGB vaizdus	14
1.2. Objektų vengimo algoritmų analizė	16
2. Autonominių bepiločių orlaivių kliūčių vengimo eksperimentų metodologija.....	22
2.1. Funkciniai ir nefunkciniai reikalavimai.....	22
2.2. Tyrimo aplinka	23
2.3. Kliūčių aptikimas ir vengimas iš monokulinių RGB vaizdų.....	23
2.4. Modelių, kurie prognozuoja gylio žemėlapius iš monokulinių RGB vaizdų, parinkimas	26
3. Eksperimentinė dalis	35
3.1. Bazinių modelių minimalaus gylio įvertinimo eksperimentai.....	35
3.2. Papildomai mokytų modelių minimalaus gylio įvertinimo eksperimentai.....	37
3.3. Papildomai mokytų modelių artimų pikselių analizė	40
3.4. Papildomai mokytų modelių specifiskai artimiems atstumams eksperimentai	43
3.5. Autonominių bepiločių orlaivių susidūrimų eksperimentai	48
Išvados	51
Dirbtinio intelekto įrankių naudojimas.....	52
Literatūros sąrašas	53

Lentelių sąrašas

1 lentelė. „KITTI VO“ testavimo rezultatai, kur testuojami jų pasiūlytas algoritmas ir kiti vizualinės odometrijos algoritmai [3]	15
2 lentelė. Gauti vertinimo metrikų rezultatai su pasiūlytu metodu ir lyginamais metodais [4]	16
3 lentelė. Susidūrimų prognozavimo tinklo architektūra [8]	20
4 lentelė. Susidūrimų prognozavimo tinklo patvirtinimo rezultatai [8].....	21
5 lentelė. Modelių prognozės su nematytais duomenų rinkiniais. Pirmi trys bandymų rinkiniai yra vidaus aplinkos, o pasuktinai du – lauko aplinkos. Siekdami sąžiningo palyginimo yra pateikiama, jog „ZoeDepth“ modelio rezultatai yra atkuriami autorių aplinkoje [15].	32
6 lentelė. Bazinių modelių metrikos.....	36
7 lentelė. Bazinių modelių metrikos su papildoma sumažinimo per konstantą funkcija	37
8 lentelė. Bazinių ir papildomai mokytų su 250 vaizdais modelių prognozavimo metrikos	38
9 lentelė. Papildomai mokytų modelių gautos metrikos su nedidelio atstumo testavimo rinkiniu...	40
10 lentelė. Papildomai mokytų modelių artimų atstumų eksperimentų rezultatai.	42
11 lentelė. Koncentruotų modelių į artimus atstumus metrinių rezultatų palyginimai	46
12 lentelė. Geriausių papildomai mokytų modelių metrinių rezultatų palyginimai	47
13 lentelė. Susidūrimų dažnis su skirtingais modeliais	49
14 lentelė. Modelių vaizdo apdorojimo ir atminties naudojimo rezultatai	49

Paveikslų sąrašas

1 pav. Bepiločių orlaivių sistemų kompiuterinės vizijos taksonomija [1].....	12
2 pav. Per metus kompiuterinės vizijos publikuojamų straipsnių kiekis [1].....	13
3 pav. Aplinkų kiekis naudojamas eksperimentuojant su kompiuterinės vizijos sistemomis [1].....	13
4 pav. Vizualinės odometrijos pavyzdys, lokalus 3d žemėlapis (a), aptiktos 2D vaizdo ypatybės (b) [2]	14
5 pav. CNN ir RCNN pagrįstos monokulinės vaizdinės odometrijos sistemos architektūra [3]	15
6 pav. Gylio žemėlapiai, kur a - autorių metodas, b - (Godard ir kt., 2019) ir c - (Ma ir Karaman, 2018). [4]	17
7 pav. Bendras siūlomo sprendimo absoliutaus gylio nustatymo metodo procesas [7].....	18
8 pav. Kokybiniai rezultatai kur, „input image“ - įvesties nuotrauka, „GT depth“ - tikras gylis ir „absolute depth map“ - metodo prognozuojamas gylis [7]	19
9 pav. Sistemos diagrama [8]	20
10 pav. „Unreal Engine“ platformoje sukurta eksperimentinė aplinka.....	23
11 pav. Kliūčių vengimas su bepiločiu orlaiviu [11]	24
12 pav. Tankaus prognozavimo transformatoriaus architektūra [12]	27
13 pav. Tankaus prognozavimo transformatoriaus surinkimo ir sujungimo architektūra [12].....	27
14 pav. Gedimų atvejai. Subtilūs santykinio gylio išdėstymo trūkumai arba trūkstamos detalės parodytos žalia spalva [13]	28
15 pav. Gylio prognozavimas, tarp „Monodepth2“ modelio ir kitų modelių [5].....	30
16 pav. Gedimų atvejai, kur „Monodepth“ blogai prognozuoja gylio žemėlapius dėl atspindinčių objektų (viršus) ar sudėtingų formų (apačia) [5]	30
17 pav. „ZoeDepth“ architektūra. RGB vaizdas tiekiamas į „MiDaS“ gylio įvertinimo sistemą prie kurio yra prijungti metrinio gylio įvertinimo moduliai [14].....	31
18 pav. Apmokymo schema. Ištininė linija: pažymėtų vaizdų srautas, punktyrinė linija: nepažymėtų vaizdų srautas. S žymi stiprių pertubacijų (pakitimų) pridėjimą prie nepažymėto vaizdo [15].....	32
19 pav. Tikrų vaizdų (a) ir sintetinių vaizdų (b) gylio žymės ir atitinkamos modelio prognozės (c) [16]	33
20 pav. „Depth Anything V2“ apmokymo schema [16]	34
21 pav. Kairėje - RGB nuotrauką, viduryje - gylio nuotrauka, dešinėje - gylio nuotrauka, mažiausio atstumo apskaičiavimuose naudojami pikseliai pažymėti raudonai	35
22 pav. Grafikas lyginantis tikrą mažiausią atstumą ir „Depth Anything“ bazinių modelių mažiausius prognozuojamus atstumus	36
23 pav. Grafikas lyginantis tikrą mažiausią atstumą ir „Depth Anything“ bazinių modelių mažiausius prognozuojamus atstumus su papildoma funkcija.....	37
24 pav. Grafikas lyginantis tikrą mažiausią atstumą ir „Depth Anything V2“ apmokytų modelių su 250 vaizdais mažiausius prognozuojamus atstumus	38
25 pav. Susidūrimų dažnis su skirtingais slenksčiais.....	39
26 pav. Artimų pikselių procentas per nuotrauką	40
27 pav. Modelio papildomai mokyto su 1000 vaizdų artimo gylio prognozės palyginimas su tikru gyliu.....	41
28 pav. Papildomai mokyto modelių pikselių klaidų pasiskirstymo dėžutės diagrama.....	42
29 pav. Svorio vertė didinant atstumą su eksponentine ir sigmoidine funkcija, kai netolimo gylio parametras - 2, stiprumas - 2	44

30 pav. Pavyzdys, rodantis tikrą gylį ir pikselių svorius su eksponentine funkcija, kai netolimo gylio parametro atstumas yra 2 metrai, o stiprumas - 2.....	44
31 pav. Pavyzdys, rodantis tikrą gylį ir pikselių svorius su eksponentine funkcija, kai netolimo gylio parametro atstumas yra 2 metrai, o stiprumas - 5.....	45
32 pav. Pavyzdys, rodantis tikrą gylį ir pikselių svorius su sigmoidine funkcija, kai netolimo gylio parametro atstumas yra 2 metrai, o stiprumas - 2.....	45
33 pav. Pavyzdys, rodantis tikrą gylį ir pikselių svorius su sigmoidine funkcija, kai netolimo gylio parametro atstumas yra 2 metrai, o stiprumas - 5.....	45
34 pav. Geriausių papildomai mokytų modelių pikselių klaidų pasiskirstymo dėžutės diagrama ...	48
35 pav. Tikro, paprastai papildomai mokyto ir konkrečiai artimiems atstumams modelių vaizdų palyginimai su gylio perdangomis.....	48
36 pav. Mažo tipo paprastai papildomai mokyto modelio sustojimo nuo kliūtis per 1000 eksperimentų grafikas.....	49
37 pav. Mažo tipo koncentruoto modelio į artimus atstumus sustojimo nuo kliūtis per 1000 eksperimentų grafikas.....	50

Įvadas

Projekto naujumas ir aktualumas:

Miškai yra viena iš pačių įvairiausių ir sudėtingiausių ekosistemų Žemėje. Tačiau tyrinėti žemiau miško lajos yra sudėtinga užduotis dėl riboto matomumo, tankios augmenijos ar neprieinamumo. Navigacija su dronu po miško lapija susiduria su įvairiais iššūkiais: ribotas GPS prieinamumas, didelis tankumas, prastas matomumas. Be to, skirtingai nuo struktūrizuotos miesto aplinkos, natūrali miško aplinka kelia žymiai sudėtingesnį algoritmų suvokimo iššūkį. Augalija turi daug netaisyklingų formų, skirtingų mastelių ir netvarkingų fonų, kurie apsunkina patikimą kliūčių suvokimą. Pavyzdžiui, plonos šakos ar iš dalies uždengti kamienai užima tik labai nedidelę regėjimo lauko dalį ir dažnai pasižymi mažu kontrastu fone kas gali lemti praleistus objektus arba nepatikimą miško vietovės interpretavimą. Norint užtikrinti sklandžią navigaciją miške, sistema turi leisti dronui sklandžiai jame judėti. Tai norint įgyvendinti, viena iš pagrindinių sistemų į kurią reikia atsižvelgti yra objektų aptikimas ir vengimas. Norint pasiekti tikslių rezultatų tyrimuose yra atsižvelgiama į objektų vengimo algoritmų sistemos funkcionalumą siekiant užtikrinti bepiločio orlaivio saugumą ir efektyvų nagrinėjimą miške, kur yra daug dinaminių ir statinių kliūčių. Ši sistema turi žvalgytis po aplinką su monokuline kamera ir naudojant iš jos išgautus RGB vaizdus, paversti juos į gylio žemėlapius. Gylio žemėlapiai yra naudojami vietoj objektų aptikimo sistemų, nes vietoj to, kad būtų identifikuojamos ir klasifikuojamos atskiros kliūtys, yra paimama aplinkos geometrinė struktūra. Tai leidžia teikti pikselių lygio erdvinę informaciją, leidžiančią sistemai tiesiogiai prognozuoti apie matomą erdvę ir kliūčių artumą pikselių lygyje. Šiuos žemėlapius vėliau galima naudoti atstumui nustatyti tarp autonominio bepiločio orlaivio ir kliūtis, kaip šaka, medis, kamienas. Taip pat gerai veikia su kliūtimis, kurios yra nepilnai matomos ar labai smulkios, kaip lapai ar pusiau uždengta šaka, nes yra atsižvelgiama į pikselius o ne objektus. Remiantis šia atstumo informacija autonominis bepilotis orlaivis turi gebėti atpažinti netoliese esančias kliūtis natūralioje aplinkoje ir esant reikalui atlikti vengimo manevrus.

Darbo tikslas – ištirti ir patobulinti monokulinio gylio įvertinimo metodus, skirtus autonominei bepiločių orlaivių navigacijai realiuoju laiku aptikti artimas kliūtis naudojant RGB vaizdus.

Darbo uždaviniai:

1. Išanalizuoti monokulinio gylio įvertinimo metodus, taikomus kliūčių vengimui naudojant RGB vaizdus.
2. Ištirti skirtingų gylio įvertinimo modelių privalumus, apribojimus ir pritaikomumą artimo nuotolio kliūčių aptikimui natūralioje aplinkoje.
3. Parinkti ir įvertinti tinkamus modelių variantus, skirtus autonominių bepiločių orlaivių kliūčių vengimui realiuoju laiku.
4. Patobulinti modelių variantus, kad būtų pagerinti artimų atstumų rezultatai.
5. Įvertinti siūlomų metodų tikslumo metrikas, susidūrimų dažnį, apdorojimo laiką ir kompiuterinių resursų naudojimą.

Darbo hipotezė – Autonominio bepiločių orlaivių artimų kliūčių vengimo rezultatai yra pagerinti pritaikius gylio įvertinimo modelius, kurie sutelkti tik į artimus atstumus.

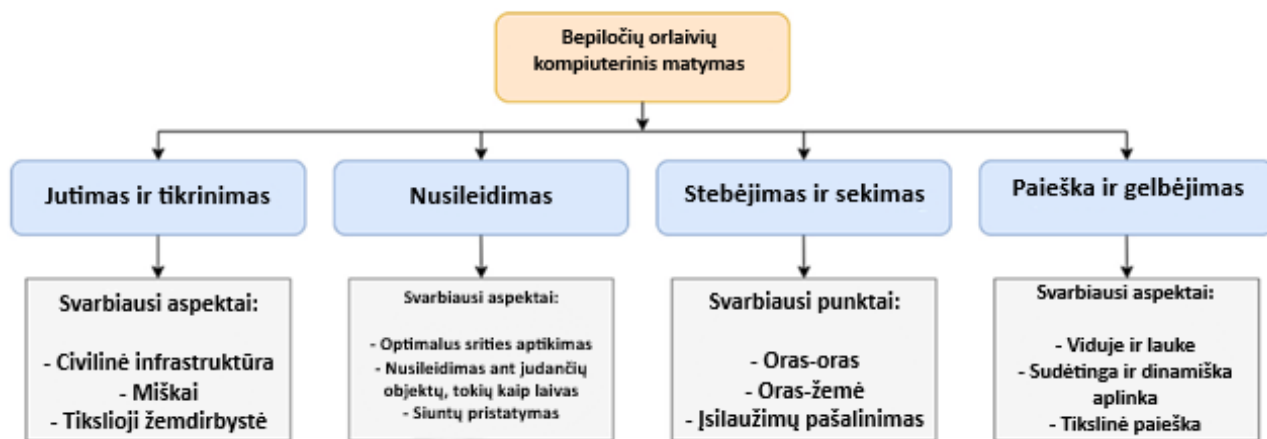
Darbo struktūra:

Darbas susideda iš trijų pagrindinių dalių: bepiločio orlaivio kompiuterinės vizijos sistemų analizės, autonominių bepiločių orlaivių kliūčių vengimo eksperimentų metodologijos ir eksperimentinės dalies. Analizės dalyje pirmiausia apžvelgiami straipsniai susiję su dirbtinio intelekto pasiekimais per paskutinius metus kompiuterinės vizijos srityje. Tada labiau įsigilinama į bepiločių orlaivių navigacijos ir objektų vengimo algoritmų analizę. Eksperimentų metodologijos dalyje, pirmiausia apibrėžiami funkciniai ir nefunkciniai reikalavimai. Tada aprašoma tyrimo aplinka ir naudojamos metrikos. Be viso to, šioje dalyje yra stipriai atsižvelgiama į skirtingus gylio prognozavimo modelius, aprašomi jų privalumai ir trūkumai. Galiausiai eksperimentinė dalis, kurioje yra įvertinami „Depth Anything V2“ modeliai, naudojant atstumo įvertinimo metrikas, apdorojimo laiką, resursų naudojimą ir susidūrimo dažnį realiuoju laiku. Visi šie eksperimentų rezultatai atlikti su baziniais, papildomai mokytais ir koncentruotais modeliais į artimus atstumus. Jie yra aprašomi ir palyginami tarpusavyje. Galiausiai yra pateikiamas geriausias rezultatas autonominių bepiločių orlaivių objektų vengimo užduotims. Darbas užbaigiamas išvadomis, kur yra aptariamoms užsibrėžtos užduotys.

1. Bepiločio orlaivio kompiuterinės vizijos sistemų analizė

Šiame skyriuje nagrinėjami ištirtų projektų teoriniai ir praktiniai aspektai, kuriais yra grindžiamos sistemos pagrindinės funkcijos. Visų pirma yra apžvelgiami tyrimai su bepiločiais orlaiviais naudojant giluminių mokymų pagrįsta kompiuterinio matymo sistemas. Vienas iš pagrindinių tyrimų nagrinėja dirbtinio intelekto pasiekimus nuo 2019 iki 2024 metų, ypač giliojo mokymų pagrindu veikiančiose kompiuterinės regos srityse, kurios padidina bepiločių orlaivių autonomiją, suteikdamos galimybę suvokti, priimti ir duoti sprendimus realiuoju laiku [1]. Autonominės bepiločių orlaivių navigacijos pagrindas yra gebėjimas suvokti ir interpretuoti aplinką. Tai pasiekama naudojant jutiklius ar regėjimo sistemas kaip kameras, įskaitant RGB, RGB-D, šiluminės ir infraraudonųjų spindulių, kurios yra plačiai naudojamos aplinkos duomenims fiksuoti. Gauti rezultatai yra apdorojami naudojant kompiuterinio matymo algoritmus, siekiant išgauti prasmingą informaciją. Pagal tyrimus gilaus mokymų modeliai integruojami į bepiločius orlaivius yra suskirstomi į keturias pagrindines grupes (žr. **1 pav.**).

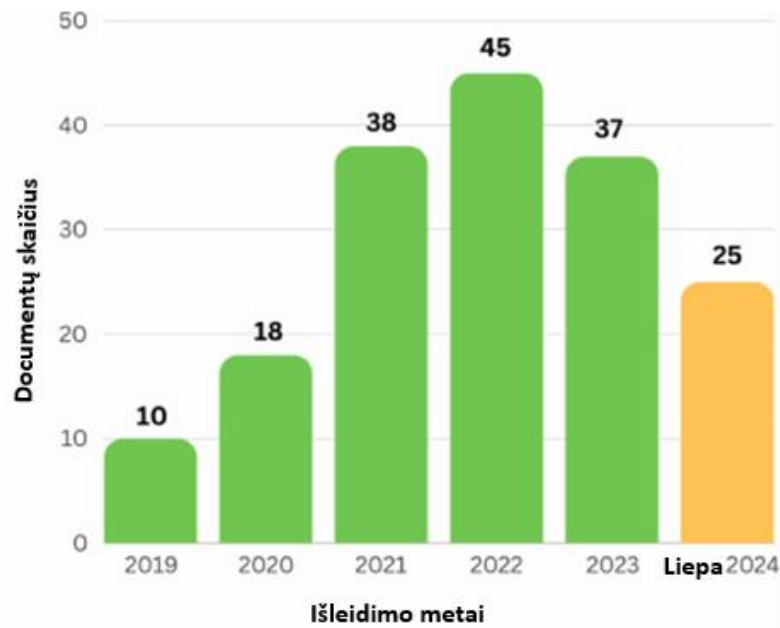
- Jutimas ir tikrinimas
- Nusileidimas (pakilimas)
- Stebėjimas ir sekimas
- Paieška ir gelbėjimas



1 pav. Bepiločių orlaivių sistemų kompiuterinės vizijos taksonomija [1]

Kaip parodyta bepiločių orlaivių sistemų kompiuterinės vizijos taksonomijoje, regėjimu pagrįstus metodus galima plačiai suskirstyti į kelias funkcines sritis. Šios kategorijos atspindi įvairius bepiločių orlaivių platformų veikimo reikalavimus. Tarp šių kategorijų, jutimas ir tikrinimas yra viena iš labiausiai taikomų sričių. Šios srities atvejais yra surenkami vaizdiniai duomenis, kuriuos bepiločiai orlaiviai naudoja užduotims kaip infrastruktūros stebėjimas, pažeidimų aptikimas, aplinkos vertinimas ir objektų vengimas.

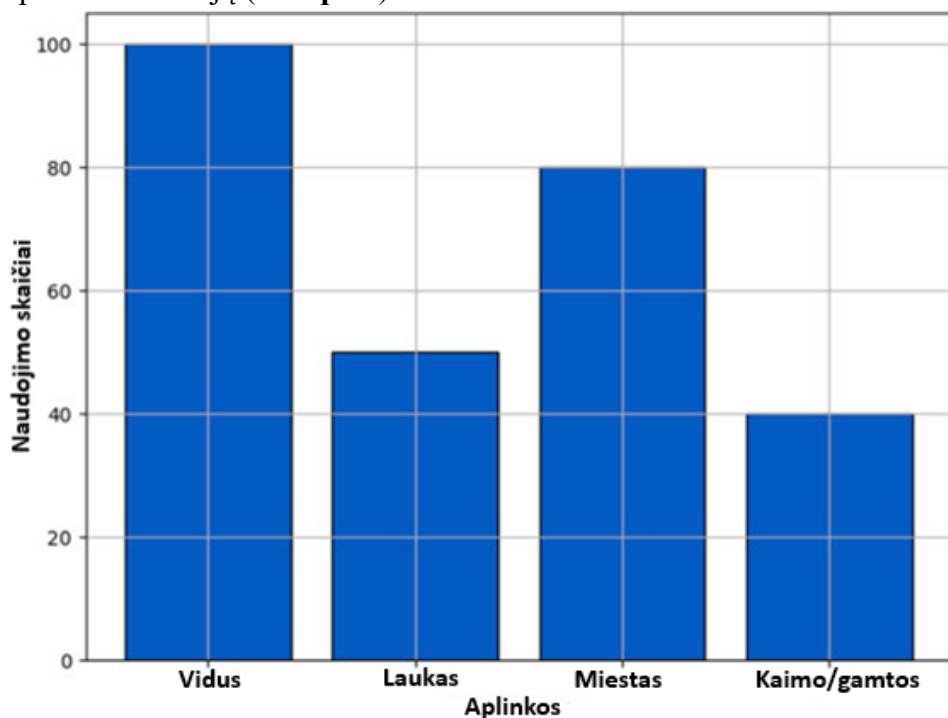
Plati ir auganti kompiuterinės vizijos svarba bepiločių orlaivių sistemose atsispindi ir iš šios srities tyrimų straipsnių kiekiu per praėjusius metus (žr. **2 pav.**). Kaip galima matyti, kompiuterinės regos tyrimų, skirtų bepiločių orlaivių taikymams, publikacijų skaičius didėja kiekvienais metais lyginant nuo 2019 metų. Ši tendencija atspindi tiek sparčią giliojo mokymų metodų pažangą, tiek didėjančią autonominių oro sistemų, galinčių veikti sudėtingoje realaus pasaulio aplinkoje, paklausą.



2 pav. Per metus kompiuterinės vizijos publikuojamų straipsnių kiekis [1]

Pagrindinis veiksnys leidžiantis pasiekti šių kompiuterinės vizijos sistemų yra gilaus mokymosi architektūrų, ypač konvoliucinių neuroninių tinklų modelių naudojimas, kurie leidžia mokytis iš vaizdinių duomenų ir pasiekti reikalingų rezultatų. Šie modeliai yra labai veiksmingi išskiriant vaizdų ypatybes ir atliekant tokias užduotis kaip klasifikavimas, aptikimas ir segmentavimas. Be to, perkėlimo mokymosi metodų taikymas leidžia bepiločių orlaivių sistemoms panaudoti iš anksto apmokytus modelius ir pritaikyti juos naujoms aplinkoms su ribotais mokymo duomenimis, taip pagerinant apibendrinimą ir sumažinant skaičiavimo reikalavimus.

Kalbant apie bepiločių orlaivių autonomija, svarbu paminėti kokiose aplinkose metodai yra mokomi ir vertinami. Tyrimuose yra pabrėžiama, kad eksperimentai turi didelę įvairovę nuo patalpų iki miesto ar natūralios aplinkos scenarijų (žr. 3 pav.).



3 pav. Aplinkų kiekis naudojamas eksperimentuojant su kompiuterinės vizijos sistemomis [1]

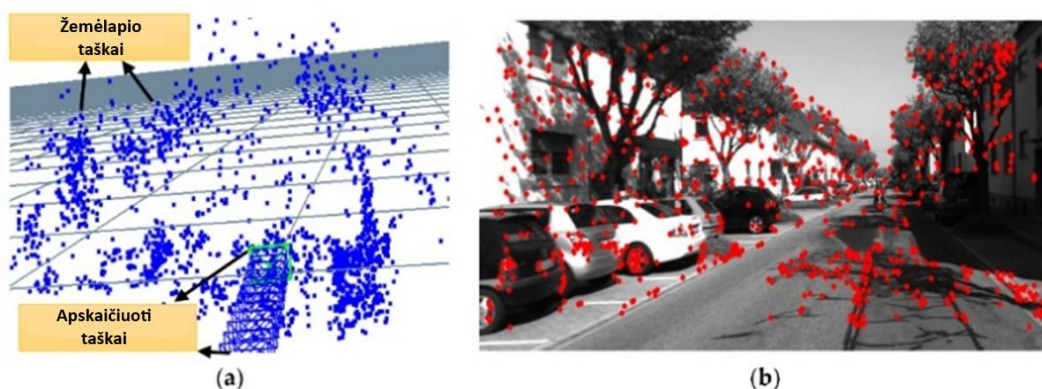
Didelė dalis esamų tyrimų yra atliekami struktūrizuotoje aplinkoje, o mažiau tyrimų nagrinėja natūralias vietas kaip miškai. Tai rodo, kad natūralioje aplinkoje metodų panaudojimas yra labiau ribotas, sudėtingesnis ir turi daugiau papildomų iššūkių, kaip netaisyklinga objektų geometrija ar tanki augmenija, kuri smarkiai paveikia konvoliucinių neuroninių tinklų modelių patikimumą. Kliūtis miškingose vietose, kaip plonos šakos, lapija gali būti prastai matomos, uždengtos ar labai smulkios ir norint, kad bepiločiai orlaiviai suvoktų ir patikimai veiktų tokiose aplinkose, reikia labai stipriai atsižvelgti į šiuos iššūkius.

Šiame tyrime sutelkiamas dėmesys į realaus laiko kliūčių aptikimo ir vengimo tyrimą natūralioje miško aplinkoje, naudojant monokulinius RGB vaizdinius duomenis. Tyrime siekiama spręsti iššūkius, susijusius su miško aplinka, kur patikimam suvokimui trukdo sudėtingi vizualiniai elementai ir mažų ar dviprasmiškų kliūčių buvimas.

1.1. Bepiločių orlaivių navigacijos analizė naudojant tik RGB vaizdus

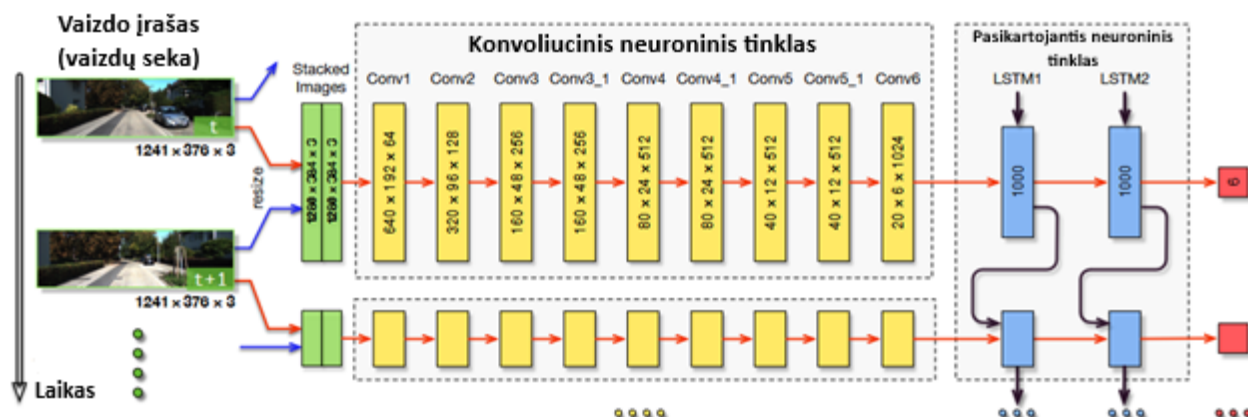
Kalbant apie objektų vengimo sistemas, svarbu paminėti ir platesnes sistemas, kurios sprendžia panašias problemas: bepiločių orlaivių navigacija. Šiais laikais yra daug bepiločių orlaivių navigacinių sistemų, tačiau daugumos jų įgyvendinimas yra labai sudėtingas. Tradiciniai metodai remiasi įvairias jutimo būdais, kaip LiDAR, ultragarsiniai jutikliai ir infraraudonųjų spindulių sistemos. Tačiau šie jutikliai dažnai turi apribojimų, tokių kaip ribotas matymo laukas, jautrumas aplinkos sąlygoms arba padidėję sąnaudų ir apkrovos reikalavimai. Tyrime žadama naudoti tik RGB kameros vaizdus. Tai vaizdinės informacijos šaltiniai, kurie naudojant kartu su giluminio mokymosi ir kompiuterinės vizijos technikomis gali leisti bepiločiams orlaiviams labai tiksliai naršyti sudėtingose aplinkose vengiant įvairių kliūčių. Naudojant tik kameros vaizdus galima supaprastinti bepiločių orlaivių sistemas, sumažinant aparatinės įrangos sudėtingumą ir išlaidas. Yra nedidelis kiekis sprendimų, nagrinėjančių bepiločių orlaivių navigaciją pasinaudojus RGB kameros vaizdais.

Vienas iš duotų sprendimų yra vizualinės odometrijos (angl. visual odometry) naudojimas. Tai metodas, naudojamas įvertinti kameros trajektorijai pagal įvestį, paprastai iš vienos ar kelių kamerų. Jis yra svarbus navigacijai vengiant kliūčių, ypač aplinkoje, kur GPS nepasiekiamas, pavyzdžiui, miškuose, tuneliuose ar tankiose miesto aplinkose. Analizuodamas užfiksuotų vaizdų pokyčius laikui bėgant, vizualinė odometrija apskaičiuoja santykinę kameros judėjimą ir įvertina jo padėtį bei orientaciją. Tai labai svarbu autonominei navigacijai, nes leidžia sistemai suprasti, kur ji yra aplinkoje, ir planuoti tolesnius veiksmus, nepasikliaujant išoriniais signalais kaip matoma iš paveikslėlio (žr. 4 pav.). Pats tikslas yra rasti atitikmenį tarp kiekvieno 3D žemėlapio taško ir aptikto 2D elemento.



4 pav. Vizualinės odometrijos pavyzdys, lokalus 3d žemėlapis (a), aptiktos 2D vaizdo ypatybės (b) [2]

Sprendime naudojamas CNN ir RCNN sistema: DeepVO [3]. CNN išmoksta geometrinės ypatybės iš nuoseklių RGB monokulinių vaizdų, o RNN modeliuoja laiko priklausomybes, kad įvertintų pozas laikui bėgant, spręsdamas nuoseklių duomenų ir judesio dinamikos iššūkį (žr. 5 pav.).



5 pav. CNN ir RCNN pagrįstos monokulinės vaizdinės odometrijos sistemos architektūra [3]

Jų sistema naudoja ilgalaikės-trumpalaikės atmintį (angl. Long Short-Term Memory), tam, kad būtų pagerintas pozos vertinimas ilgose vaizdų sekose. Tinklas apmokomas naudojant sąnaudų funkciją, pagrįstą Euklido atstumo tarp numatytų, tikrų pozicijų ir orientacijų sumažinimu, o praradimo funkcija apima ir padėties, ir orientacijos klaidas. Šio giliu mokymusi pagrįsto metodo pranašumas yra tas, kad jis atlieka pozos įvertinimą nuo galo iki galo (angl. end-to-end manner), todėl nereikia naudoti tradicinių modulių kaip įprastuose vizualinės odometrijos algoritmuose, pavyzdžiui, kameros kalibravimo ar rankinio parametrų derinimo. Eksperimentai atlikti naudojant „KITTI VO“, kuris apima realaus pasaulio duomenis iš vairavimo sekų (žr. 1 lentelė).

1 lentelė. „KITTI VO“ testavimo rezultatai, kur testuojami jų pasiūlytas algoritmas ir kiti vizualinės odometrijos algoritmai [3]

Seq.	DeepVo		VISO2_M		VISO2_s	
	$t_{rel}(\%)$	$r_{rel}(^\circ)$	$t_{rel}(\%)$	$r_{rel}(^\circ)$	$t_{rel}(\%)$	$r_{rel}(^\circ)$
03	8.49	6.89	8.47	8.82	3.21	3.25
04	7.19	6.97	4.69	4.49	2.12	2.12
05	2.62	3.61	19.22	17.58	1.53	1.60
06	5.42	5.82	7.30	6.14	1.48	1.58
07	3.91	4.60	23.61	29.11	1.85	1.91
10	8.11	8.83	41.56	32.99	1.17	1.30
Vidurkis	5.96	6.12	17.48	16.52	1.89	1.96

Testavimams patikrinti naudojamos dvi metrikos: vertimo RMSE poslinkis (angl. translation RMSE drift), kuris paskaičiuoja padėties paklaidą ir rotacinis RMSE (angl. rotational RMSE) matuojantis numatomos kameros ar sistemos orientacijos paklaidą. Jie rodo, kad modelis gali pateikti tikslius vaizdinės odometrijos rezultatus su tikslu mastelio atkūrimu, net ir tokiose aplinkose, kuriose greitas judėjimas (90 km/h), dinamiški objektai ar ribota informacija, kuri keltų daug iššūkių tradiciniams vizualinės odometrijos metodams (kaip monokulinis VISO2). Šis algoritmas parodė ir konkurencingus rezultatus lyginant su stereo VISO2, kuris naudoja pradinio mastelio nuorodas, kaip fiksuotos kameros aukštis, kad santykinė skalė būtų suderinta su realybe (ground truth), nors DeepVO buvo apmokytas tik monokuliniais duomenimis. Rezultatai parodė, kad „DeepVO“ gali pasiekti patikimą pozos įvertinimą, nereikalaujant rankinio skalės derinimo. Tačiau autoriai pamini ir viena iš didelių iššūkių – sunku atkurti absoliutų mastelį aplinkose, kurios labai skiriasi nuo

mokymo duomenų. Šie apribojimai ypač svarbus miško natūralioje aplinkoje, kur vizualinės savybės yra mažiau nuoseklios, todėl autoriai pabrėžia, kad šis metodas nesitikima, jog pakeis klasikinę geometriją pagrįstą metodą, tačiau gali būti perspektyvus papildymas kur tradiciniai metodai susiduria su sunkumais.

Nors vizualinė odometrija daugiau dėmesio skiria judesio įvertinimui, ir nelabai atsižvelgia į netoliese esančias kliūtis, ji atlieka esminį vaidmenį užtikrinant autonominę navigaciją ir nuolatinę trajektorijos įvertinimą. O tai reikalinga norint efektyviai suvokti ir įvertinti tiriamąją aplinką.

1.2. Objektų vengimo algoritmų analizė

Gylio įvertinimas iš monokulinių RGB vaizdų tapo esminiu kompiuterinės vizijos sistemos komponentu, naudojamu bepiločiu orlaiviuose, iš kurios galima išgauti erdvinę informaciją naudojant tik vieną kameros įvestį. Tai vienas iš efektyvesnių būdų, norint įgyvendinti objektų vengimo sistemą naudojant tik RGB kamerą. Pagrindinė informacija, kurią reikia įvertinti yra atstumas tarp esamo bepiločio orlaivio ir kliūtis. Tačiau tai yra sudėtinga užduotis, ypač miško aplinkoje, nes kitaip nei miesto ar uždaros aplinkos, miškas yra tankios augmenijos ir sudėtingos struktūros. Šios savybės sukelia vizualinių ypatybių dviprasmiškumą ir apsunkina atstumo įvertinimo modelių efektyvumą.

Į tai atsižvelgia tyrimas, kurio autoriai pabrėžia, jog net plačiai naudojami duomenų rinkiniai, kaip „KITTI“ ir „NYU Depth V2“ turi tik ribotą skaičių miško vaizdų, todėl prastai veikia miško aplinkoje [4]. Siekdami išspręsti šią problemą autoriai siūlo prižiūrimą gilios mokymosi sistemą, specialiai skirta miško vaizdams. Pirmiausia jie sudarė RGB-D duomenų rinkinį, naudojant antžeminio lazerio skaitytuvą (angl. terrestrial laser scanner), leidžianti generuoti tikslius žemės paviršiaus gylio žemėlapius. Tada šie duomenys apmokomi su „ResNet18“ algoritmu. Norint įvertinti modelį, yra pasitelkiamos vertinimo metrikos. Vidutinė kvadratinė klaida (RMSE) naudojama bendrai prognozės paklaidai matuoti, δ matuoja slenksčio tikslumą, kuris nurodo taškų procentą, kurių nuokrypis nuo pagrindinės tiesos yra mažesnis nei 1,25 karto. Siekdami išspręsti RMSE jautrumą nuokrypiams, yra naudojama vidutinė absoliučioji klaida (MAE), kad patikimiau būtų įvertintas modelio tikslumas. Taip pat buvo lyginamas šis modelis su kitais metodais, vienu kuris naudoja RGB gylio vaizdą (Godard ir kt., 2019)[5], kitas – naudojančių nedidelį gylį (angl. sparse depth) kaip įvestį (Ma ir Karaman, 2018)[6]. Gauti testavimo rezultatai atvaizduojami lentelėje (žr. 2 lentelė).

2 lentelė. Gauti vertinimo metrikų rezultatai su pasiūlytu metodu ir lyginamais metodais [4]

Metodas	Įvesties pavyzdžiai	RMSE[m] (mažiau yra geriau)	MAE[m] (mažiau yra geriau)	$\delta < 1.25$
(Godard ir kt., 2019),	-	4.33	3.58	0.15
(Ma ir Karaman, 2018)	200	3.94	1.72	0.49
Autorių metodas	100	2.34	1.14	0.64
Autorių metodas	500	2.1	0.97	0.7
Autorių metodas	1000	1.97	0.89	0.72

Kaip galima matyti iš rezultatų, pasiūlytas metodas nuolat lenkia kitus, lyginamus metodus ir jo slenksčio tikslumas siekia virš 70%, kas rodo, jog net 70% pikselių atitinka realybę. Iš vaizdinių duomenų, autorių modelis išgauna žymiai tikslesnius rezultatus miško aplinkoje, negu kiti lyginami metodai (žr. 6 pav.). Paveikslėlyje violetinės sritys visuose gylio žemėlapiuose žymi pikselius (medžius), esančius arti kameros, o geltonos sritys iliustruoja tolimesnius pikselius (medžius). Juoda spalva atitinka nulinės vertės pikselius.



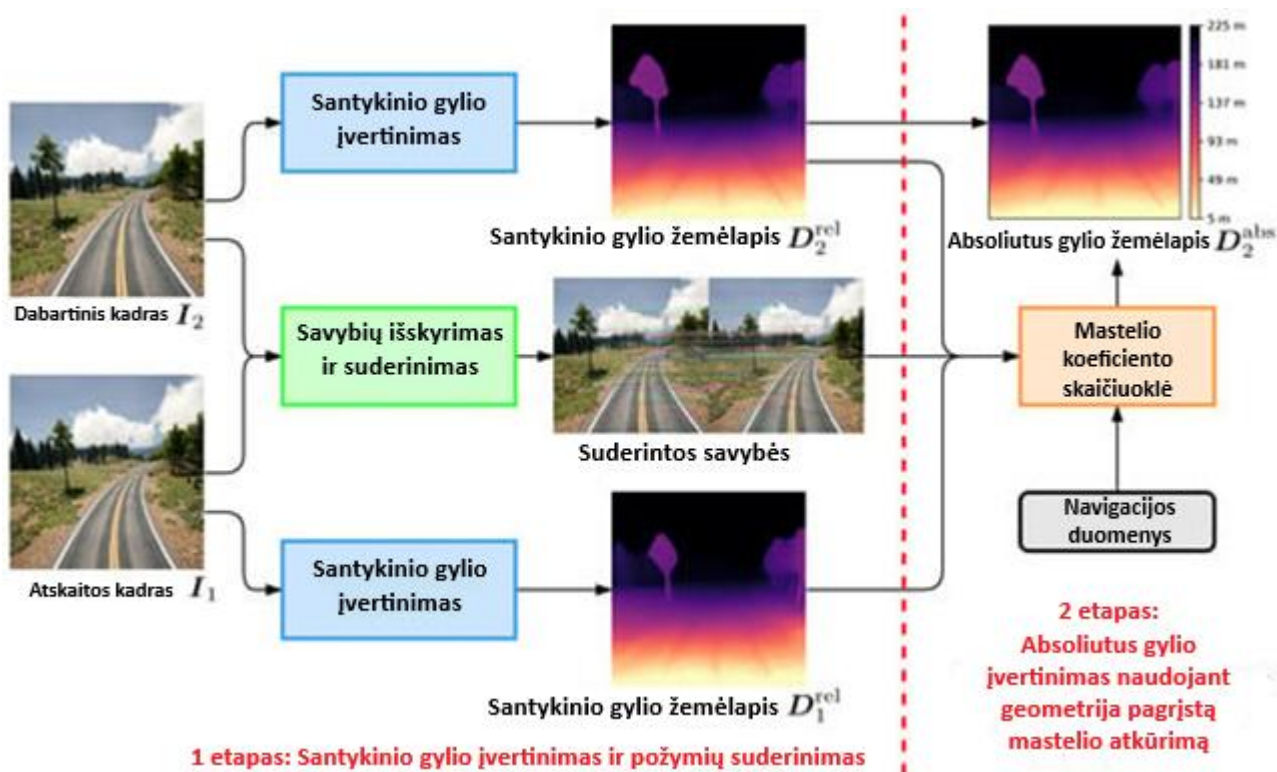
6 pav. Gylio žemėlapiai, kur a - autorių metodas, b - (Godard ir kt., 2019) ir c - (Ma ir Karaman, 2018). [4]

Nepaisant žadančių rezultatų, šis pasiūlytas metodas turi tam tikrų apribojimų. Kartais atsiranda problemų su dideliu kiekiu medžių lajų, lapų, kuriuos sunku sėkmingai išfiltruoti ar objektais kurie yra už didelio atstumo, kurių gylius sunku užfiksuoti dėl per mažo nuotraukos raiškos dydžio.

Kitas tyrimas, kuriame siekiama išspręsti atstumo problemą tarp objekto ir bepiločio orlaivio yra geometrijos pagrįstas mastelio atkūrimo algoritmo metodas [7]. Jame yra aiškiai pabrėžiama problema, kad daugumos esamų metodų apribojimas, išgaunant informaciją iš RGB vaizdų, yra nesugebėjimas numatyti absoliučių gylio verčių dėl mastelio problemos. Kitaip tariant, dauguma taikomų metodų bando įvertinti trimatės struktūros gyli naudojant vieną dvimatį vaizdą ir dėl to kyla geometrinių dviprasmiškumų prognozuojant sudėtingas scenas, tokias kaip uždengimai, tekstūros pokyčiai ar skirtingos apšvietimo sąlygos.

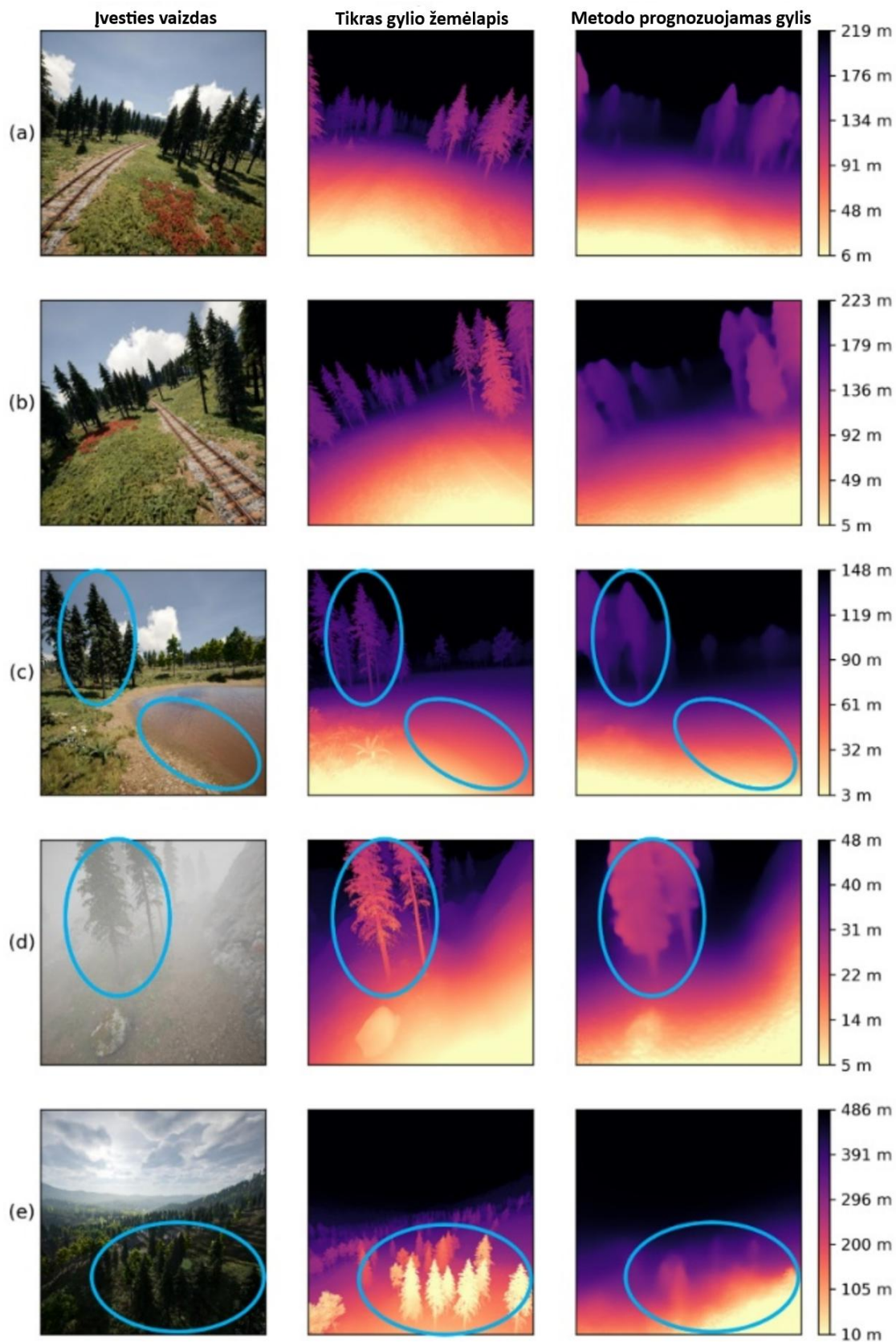
Siekdami išspręsti šią mastelio dviprasmiškumo problemą yra naudojamas geometrijos pagrįstas mastelio atkūrimo metodas, kuris integruoja judėjimo informaciją, kad atkurtų absoliutų gyli. Jis

konvertuoja santykinio gylio įverčius iš neprižiūrimų gilaus mokymosi modelių į absoliutaus gylio žemėlapius. Metodas naudoja požymių atitikmenis tarp vienas po kito einančių vaizdų ir bepiločio orlaivio padėties duomenų, kad apskaičiuotų mastelio koeficientą, kuris jungia santykinį ir absoliutųjį gylį (žr. 7 pav.). Atliekant geometrinės transformacijos procesą yra įtraukiami ir bepiločio orlaivio padėties kampai, taip užtikrinant, kad gylis būtų tiksliai įvertintas skrydžio scenose. Šis metodas yra atsparus navigacijos jutiklių triukšmui, gali būti taikomas skirtingiems bepiločių orlaivių judesiams ir gylio diapazonams, todėl galima gauti patikimą absoliutaus gylio informaciją.



7 pav. Bendras siūlomo sprendimo absoliutaus gylio nustatymo metodo procesas [7]

Iš atliktų kokybinių rezultatų (žr. 8 pav.) galime pamatyti, jog metodas efektyviai generuoja tikslus absoliutaus gylio žemėlapiu iš RGB vaizdų. Jis sukuria vizualiai nuoseklius gylio žemėlapius, kurie glaudžiai sutampa su tikrais duomenimis, net ir sudėtingoje aplinkoje. Taip pat jis sėkmingai tvarko bepiločio orlaivio judesio, jutiklio triukšmo ir gylio diapazonų pokyčius. Sugeneruoti gylio žemėlapiai yra aiškūs ir informatyvūs, o tai yra svarbu atliekant užduotis realiuoju laiku skrydžio scenarijuose. Šis metodas yra paprastas, reikalaujantis tik bepiločio orlaivio su kamera, greitas ir pritaikomas įvairiems gylio diapazonams, gylio variacijoms ar bepiločio orlaivio judėjimo modeliams. Tačiau šis metodas taip pat turi tam tikrų apribojimų. Vienas iš jų yra nesugebėjimas tiksliai įvertinti vandens paviršių (c pavyzdys). Taip pat, netinkamai nuspėja mažyčius objektų kraštus, todėl atsiranda tam tikrų artefaktų, kaip parodyti pavyzdžiuose (c ir d) medžiai. Galiausiai, rezultatai nėra geri rūko ir silpno apšvietimo scenose (c ir e pavyzdžiai), kai vaizdo ypatybės nėra pakankamai aiškios.

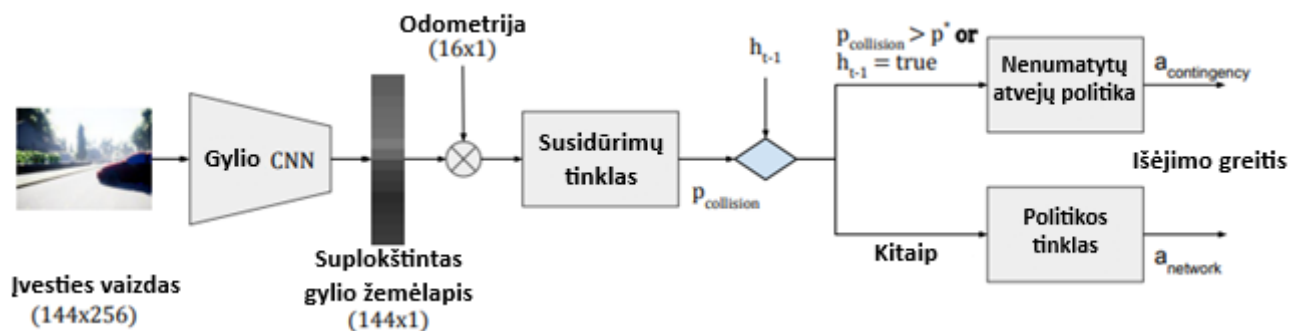


8 pav. Kokybiniai rezultatai kur, „input image“ - įvesties nuotrauka, „GT depth“ - tikras gylys ir „absolute depth map“ - metodo prognozuojamas gylys [7]

Vienam iš tyrimų, kur yra naudojami tik monokuliniai RGB vaizdai objektams vengti, yra sukurtas metodas, kurio architektūra sudaro keli specializuoti komponentai, kur kiekvienas neuroninis tinklas atlieka atskirą užduotį [8]. Konkrečiau sistemą apima:

- Gylio įvertinimo tinklas – iš RGB vaizdų nustato erdvinę struktūrą.
- Susidūrimo prognozavimo tinklas – įvertina neišvengiamo susidūrimo tikimybę.
- Politikos tinklas – generuoja navigacijos komandas.
- Nenumatytų atvejų modulis, kuris atlieka saugų vengimą naudodamas A* arba iš anksto nustatytą ekspertų politiką.

Ši konstrukcija leidžia sistemai atskirti suvokimą, rizikos vertinimą ir kontrolę, tuo pačiu užtikrinant saugumą naudojant papildomą atsarginį mechanizmą. Be viso to, pati sistema iš kelių tarpusavyje sujungtų modulių kurie apdoroja vaizdus ir generuoja navigacijos komandas realiu laiku (žr. 9 pav.).



9 pav. Sistemos diagrama [8]

Pirmiausia yra apdorojama RGB įvesties vaizdai gylio įvertinimo tinkle su CNN, o sumažinto mastelio išvestis suplokštinama į 144x1 ilgio vektorius. Gauti gylio žemėlapiai veikia kaip geometrinė aplinkos abstrakcija ir naudojami kaip pagrindinė įvestis vėlesniems moduliams. Tada gylio žemėlapis perduodamas susidūrimo prognozavimo tinklui, kuris įvertina susidūrimo tikimybę artimiausiu metu. Šis tinklas priima tiek gylio informaciją, tiek dabartinę bepiločio orlaivio judėjimo būseną, todėl gali įvertinti riziką, pagrįstą aplinkos struktūra ir trajektorija. Jei susidūrimo tikimybė yra mažesnė už tikslinę ribą ir nenumatytos situacijos politika šiuo metu nevykdoma, įvesties vektorius perduodamas politikos tinklui, kuris išveda greičio komandą. Priešingu atveju, užklausias nenumatyto atvejo modulis. Aktyvavus nenumatytą situaciją, orlaivis trumpą laiką pilotuojamas, kol jis nustoja veikti. Tada orlaivio valdymas grąžinamas politikos tinklui. Ši konstrukcija leidžia bepiločiui orlaiviui reaguoti į aplinkos pokyčius realiuoju laiku, naudojant gylio žemėlapius kaip pagrindinį kliūčių artumo nustatymą.

Pagrindinis dėmesys skirtas susidūrimo prognozavimo tinklo veikimui. Tai sistema, nustatanti, ar bepilotis orlaivis turi tęsti įprastą veikimą, ar persijungti į saugos mechanizmus remiantis dabartine bepiločio orlaivio būseną. Jo įvestį sudaro gylio žemėlapis ir odometrijos informacija, kurios kartu apibūdina aplinką ir dabartinį bepiločio orlaivio judėjimą. Tinklas apdoroja šią įvestį ir gautą reprezentavimą sujungia su pasirinktu veiksmo indeksu. Šis sujungtas reprezentavimas toliau apdorojamas per papildomus sluoksnius, sukuriant 2×1 išvesties vektorius, atitinkantį nenormalizuotą susidūrimo ir išvengimo tikimybę per kitus d laiko žingsnius. Šie išėjimai konvertuojami į normalizuotas tikimybes naudojant „softmax“ logistinę transformaciją. Visa tinklo architektūra atvaizduojama lentelėje (žr. 3 lentelė).

3 lentelė. Susidūrimų prognozavimo tinklo architektūra [8]

Sluoksnis	Įvesties ilgis	Aktyvacijos funkcija	Išvesties vienetai
Įvestis	10	Leaky ReLU	256
H1	256	Leaky ReLU	256
H2	256	Leaky ReLU	128

H3	128	Leaky ReLU	32
Sujungimas	32+4	Nėra	36
H4	36	Leaky ReLU	32
H5	32	Leaky ReLU	32
H6	32	Leaky ReLU	32
Išvestis	32	Nėra	2

Susidūrimų prognozavimo tinklas apmokomas naudojant prižiūrimą mokymąsi. Mokymo duomenys generuojami paleidus gylio įvertinimo ir politikos tinklus kliūčių ruože 1000 epizodų be susidūrimų prognozavimo tinklo. Šio proceso metu gylio žemėlapių įvertinimai ir odometrijos vektoriai įrašomi kaip įvesties duomenys, o susidūrimų įvykiai gaunami naudojant „AirSim“ simulatoriaus susidūrimo žymę. Žymės sukonstruotos taip, kad laiko žingsnis būtų pažymėtas kaip susidūrimas, jei susidūrimas įvyksta per kitus dešimt kadru. Tinklas apmokomas minimizuojant dvejetainę kryžminės entropijos nuostolį tarp prognozuojamų ir tikrųjų susidūrimo tikimybių. Kadangi mokymo metu duomenų rinkinyje buvo daug daugiau nesusidūrusių kadru pavyzdžių nei susidūrusių kadru – iš viso 295 183 kadrai ir tik 1890 kadru, kuriuose susidūrimai įvyko per kitus dešimt kadru mokymo paketai sudaromi naudojant fiksuotą 25 % susidūrimo pavyzdžių ir 75 % išvengimo pavyzdžių santykį, siekiant užtikrinti stabilų mokymą.

Šio susidūrimų prognozavimo tinklo efektyvumas vertinamas nepriklausomai, naudojant gylio įvertinimo ir politikos tinklų sugeneruotas trajektorijas. Vertinimo duomenų rinkinį sudaro ir susidūrimų, ir išvengimo pavyzdžiai, iš kurių 189 susidūrimais pažymėti kadrai ir maždaug 29 000 išvengimo neturinčių kadru. Siekiant išspręsti klasių disbalansą vertinimo metu, partijos sudaromos taip pat, kaip ir treniravimo paketai. Iš viso išanalizuota 25000 partijų. Rezultatai atvaizduoti lentelėje (žr. 4 lentelė).

4 lentelė. Susidūrimų prognozavimo tinklo patvirtinimo rezultatai [8]

Metrika	Vertė
Tikslumas	0.982
Precizija	0.976
Atkūrimas	0.954
F1 balas	0.963
Kryžminė entropija	0.057
Tikras teigiamų rezultatų dažnis	0.239
Tikras neigiamų rezultatų dažnis	0.743
Klaidingas teigiamų rezultatų dažnis	0.006
Klaidingas neigiamų rezultatų dažnis	0.012

Iš rezultatų galima pamatyti, kad yra aukšti tikslumo rezultatai. Ypač svarbu atsižvelgti, kad tiek klaidingas teigiamų rezultatų dažnis, tiek klaidingas neigiamų rezultatų dažnis yra mažas, o tai rodo, kad tinklas efektyviai atskiria saugias ir didelės rizikos būsenas. Galima sakyti, yra tiksliai prognozuojamos situacijos, kurios gali sukelti susidūrimą per trumpą laiką. Šis mažas klaidų dažnis rodo, kad sistema gali patikimai aptikti galimus susidūrimus, taip palaikydama saugią navigaciją, kai ji yra derinama su A* pagrįsta atsargine strategija.

Visi tyrinėti projektai patvirtina, kad gylio įvertinimas yra būtinas patikimam kliūčių vengimui naudojant tik RGB kameras, tačiau šio tikslo sprendimas gamtinėje aplinkoje išlieka sudėtingas. Šiems apribojimams, kurie trukdo patikimai vengti kliūtis, ir bus skiriama daugiausia dėmesio tolimesnėse tyrimo dalyse. Bus tyrinėjami gylio įvertinimo modeliai, kurie bus panaudojami ir tobulinami, siekiant pagerinti artimo nuotolio suvokimą ir navigacijos patikimumą miško aplinkoje.

2. Autonominių bepiločių orlaivių kliūčių vengimo eksperimentų metodologija

Bepiločių orlaivių sistemos vis dažniau yra taikomos įvairiose aplinkose, sudėtingose gamtinėse vietovėse, kaip miškai, tačiau jų panaudojimas yra ribotas. Tai yra dėl įvairių suvokimo sudėtingumų, sukeltų dėl aplinkos tankios augmenijos, netaisyklingos struktūros ar daug vizualiai neaiškių kliūčių. Šiame tyrime projektuojamoje bepiločių orlaivių sistemoje didžiausias dėmesys skiriamas artimo nuotolio kliūčių vengimo metodui naudojant tik monokulinius RGB vaizdus erdvinei struktūrai nustatyti ir kliūtims aptikti artimu atstumu. Sistema veikia taip, kad nustatytų potencialiai pavojingas kliūtis remiantis vizualine įvestimi ir tokiu būdu greitai priimtų sprendimus apie kliūties atstumą ir jo vengimą. Tai vienas iš pagrindinių žingsnių, norint apsaugoti bepiločius orlaivius veikiančius sudėtingose miško aplinkose. Be viso to, yra aprašomi funkciniai ir nefunkciniai reikalavimai, tyrimo aplinka ir metrikos.

2.1. Funkciniai ir nefunkciniai reikalavimai

Siekiant užtikrinti, kad autonominė bepiločių orlaivių sistema veiksmingai atliktų savo tikslus ir eksperimentus yra būtina aiškiai apibrėžti tiek funkcinius, tiek nefunkcinius reikalavimus.

Funkciniai reikalavimai:

- Sistema turi priimti monokulinius RGB vaizdus realiu laiku ir juos nuolat apdoroti.
- Sistema turi savarankiškai aptikti arti esančias kliūtis ir jas vengti.
- Sistema turi sukurti gylio žemėlapių iš RGB vaizdų, kuris leidžia spręsti apie kliūčių artumą.
- Sistema turi generuoti vengimo veiksmus, kai aptinkamas galimas susidūrimas.
- Sistema turi gebėti veikti miško aplinkoje, su tankiomis ir netaisyklingomis kliūtėmis.

Nefunkciniai reikalavimai:

- Sistema turi veikti realiu laiku.
- Jutiklių triukšmai, uždengimai turi būti apdoroti be gedimų, susidūrimų.
- Sistema turi naudoti nedaug kompiuterinių resursų, kad veiktų įrenginiuose esančiuose bepiločių orlaivių platformose.
- Įgyvendinti modeliai turi veikti tiksliai, ypač neturi būti jokių problemų su objektų susidūrimais.
- Sistema turi remtis tik monokuliniiais RGB vaizdais, vengiant papildomų jutiklių, tokių kaip LiDAR, stereo kameros ar gylio jutikliai.

Kartu šie funkciniai ir nefunkciniai reikalavimai suteikia struktūrizuotą pagrindą projektavimui. Aiškiai apibrėžus šiuos apribojimus, užtikrinama, kad sistema ne tik atliktų būtinas užduotis, bet ir tai darytų tinkamai, efektyviai, patikimai ir prisitaikant prie miško aplinkos.

2.2. Tyrimo aplinka

Šiame tyrime daromi eksperimentiniai tyrimai vienoje simuliacijoje. Aplinka sukuriama pasinaudojant „Unreal Engine“ 3D žaidimų variklį, „Airsim“ simuliacijų sistemą ir „Python programavimo kalbą.

„Unreal Engine“ – galinga realaus laiko 3D kūrimo platforma skirta naudoti žaidimų kūrimui, modeliavimui ar simuliacijoms kurti. Paskutiniu metu plačiai naudojama robotikoje ir kompiuterinės regos tyrimuose foto realistinėms aplinkoms kurti, kadangi turi aukštos kokybės grafikas, realistišką fiziką ir pažangų atvaizdavimo variklį [9]. Eksperimentinę aplinką sudaro miškas su medžiais ir netaisyklingos formos kliūtimis, kurie leidžia įvertinti kliūčių vengimo našumą vizualiai sudėtingose situacijose (žr. **10 pav.**).



10 pav. „Unreal Engine“ platformoje sukurta eksperimentinė aplinka

Simuliacija vykdoma naudojant „Airsim“ platformą. Tai atvirojo kodo simulatorius, kuris yra sukurtas autonominės transporto priemonės eksperimentuoti kaip bepiločiai orlaiviai ar automobiliai. Jis yra sukurtas „Unreal Engine“ pagrindu, kuris siūlo fiziškai ir vizualiai realistiškas simuliacijas [10]. Integruoja didelės raiškos vaizdinį perteikimą su fizikos pagrindu sukurta simuliacija, leisdamas bepiločių orlaivių modeliams nuosekliai sąveikauti su aplinka. Be viso to, „AirSim“ turi didelį kiekį įvairių virtualių jutiklių, kaip RGB kameros, gylio kameros, GPS, kurie gali būti panaudoti apdorojant duomenis realiuoju laiku. Tai naudinga mūsų eksperimentams, kadangi duomenys yra surenkami ir analizuojami per šią platformą naudojant taikomųjų programų programavimo sąsajas su „Python“ programavimo kalba. Tai leidžia sklandžiai integruoti mašininio mokymosi modelius ir juos tyrinėti.

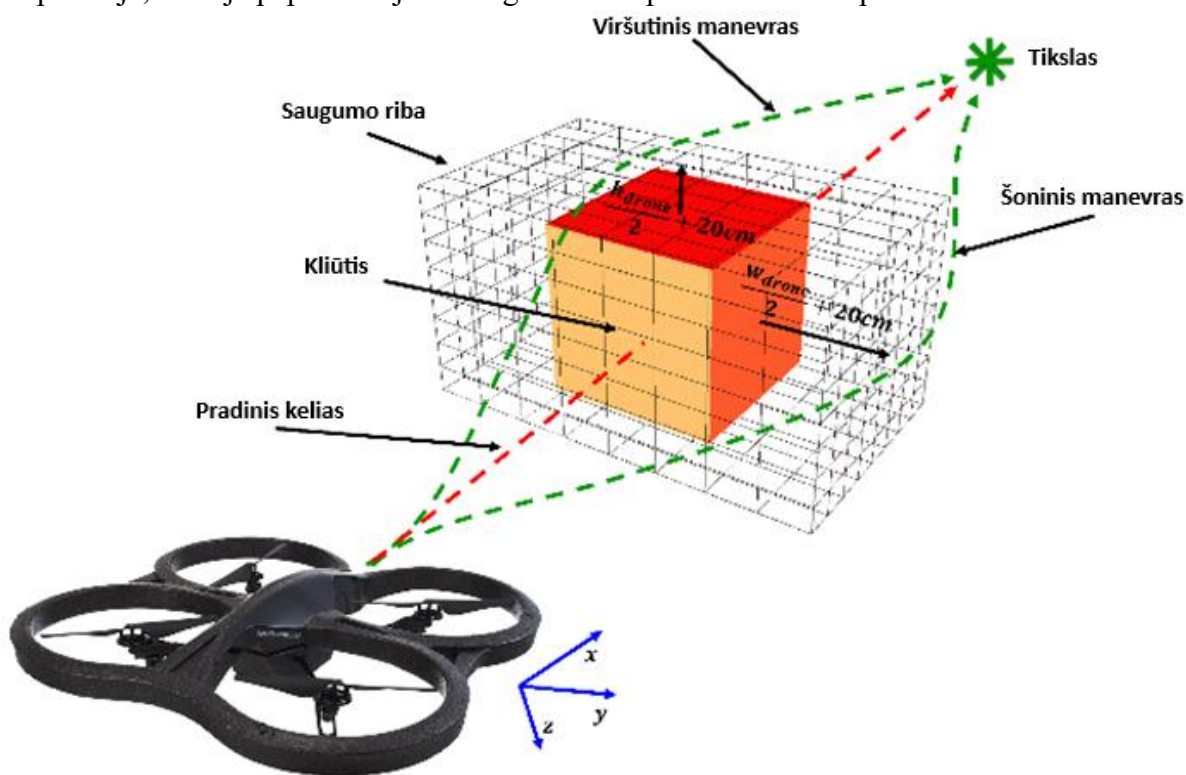
Mašininio mokymo komponentams įdiegti ir naudoti yra pasitelkiama „Python“ programavimo kalba su tokiais bibliotekomis kaip „OpenCV“, „Pytorch“, „Matplotlib“, „transformers“, „NumPy“. Visos šios bibliotekos reikalingos norint efektyviai kurti, mokyti ir vertinti dirbtinio intelekto gilaus mokymosi modelius kliūčių aptikimo ir vengimo užduotims.

2.3. Kliūčių aptikimas ir vengimas iš monokulinių RGB vaizdų

Kliūčių vengimas ir aptikimas yra esminė autonominio bepiločio orlaivio sistemos dalis. Miškingose vietovėse bepiločiai orlaiviai susiduria su daugybe iššūkių, tokių kaip žemai kabanti šakos, kintantis reljefas, tanki lapija ar dinamiškos apšvietimo sąlygos. Jei kliūčių aptikimas yra efektyvus,

tai tada bepilotis orlaivis gali greitai priimti sprendimus realiuoju laiku ir užtikrinti saugų judėjimą miške

Šios dalies tikslas – leisti bepiločiui orlaiviui laiku tiksliai aptikti ir įvertinti atstumą iki netoliese esančių objektų, užtikrinant saugią navigaciją ir ignoruojant nereikšmingą tolimą foną. Pagal tikslą, turėtume gauti sistemą, kuri gauna monokulinę vaizdinę informaciją iš bepiločio orlaivio kameros, ją apdoroja ir tada, panaudojus kliūčių vengimo algoritmą yra aptinkamas atstumas tarp objekto ir bepiločio orlaivio. Jei objektas yra per arti tam tikro nustatyto slenksčio, bepilotis orlaivis turi kliūtį išvengti ir toliau tęsti judėjimą miške iki tol bus ištyrinėta visa pažymėta aplinka (žr. **11 pav.**). Kitaip nei stereo kameros, monokulinės kameros tiesiogiai neapskaičiuoja objektų gylio ir norint tai padaryti reikia pasinaudoti dirbtinio intelekto modeliais. Svarbu paminėti, kad tyrimo metu yra naudojama tik bepiločio orlaivio RGB kamera, kadangi dauguma atveju jos yra žymiai plačiau prieinamos ir lengviau integruojamos, negu LiDAR. Nors kelių jutiklių sistemos gali pateikti tikslesnę informaciją apie gylį, vien tik RGB įvesties duomenys leidžia taikyti lankstesnius ir prieinamesnius sprendimus, ypač aplinkoje, kurioje papildomi jutikliai gali būti nepraktiški arba nepatikimi.



11 pav. Kliūčių vengimas su bepiločiu orlaiviu [11]

Siekiant įvertinti kliūčių vengimo algoritmų kokybę, yra nustatomos kelios metrikos:

- Atstumo įvertinimo paklaida – tikslas yra įvertinti monokulinės kameros suvokimą apie atstumą iki kliūties. Šiame eksperimente yra apskaičiuojamas atstumas iki kliūties su įvairiais algoritmais ir tada rezultatai yra lyginami su realiu atstumu nurodytu simuliacijose. Šiai metrikai apskaičiuoti yra naudojami keli pagrindiniai rodikliai:
 - Vidutinė absoliučioji paklaida (angl. Mean Absolute Error), kuri matuoja vidutinį absoliutų skirtumą tarp prognozuojamo ir tikro atstumo.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

Čia:

n – mėginių, testų skaičius

y_i – tikroji i^{th} vertė

\hat{y}_i – prognozuojama i^{th} vertė

- Vidutinis kvadratinis nuokrypis (angl. Root Mean Squared Error), kuris irgi matuoja vidutinį skirtumą tarp prognozuojamo ir tikro atstumo, tačiau labiau išryškina išskirtines prastas vertes.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

Čia:

n – mėginių, testų skaičius

y_i – tikroji i^{th} vertė

\hat{y}_i – prognozuojama i^{th} vertė

- Likučių standartinis nuokrypis (angl. Standard Deviation of Residuals), įvertina klaidų stabilumą ir sklaidą. Jeigu žemas, reiškia kad prognozės yra stabilios, jei aukštas - nenuspėjamas elgesys.
- Koreliacija, parodo ar prognozės seka ta pačią tendenciją kaip ir faktiniai duomenys. Jei yra 1 – tobulai seka, 0 – nėra visiškai jokios tendencijos, -1 – neigiama tendencija. Ji apskaičiuojama su šia formule:

$$r = \frac{\sum_{i=1}^n (x - m_x)(y - m_y)}{\sqrt{\sum_{i=1}^n (x - m_x)^2 \sum_{i=1}^n (y - m_y)^2}} \quad (3)$$

Čia:

n – mėginių, testų skaičius

x – faktinės vertės

y – modelio prognozuojamos vertės

m_x – faktinių verčių vidurkiai

m_y – modelio prognozuojamų verčių vidurkiai

- Vidutinis absoliutus skirtumas tarp iš eilės einančių imčių – parodo kaip stipriai keičiasi paklaidos tarp gretimų pikselių (atmetant visus minusus). Mažas vienetas parodo, kad klaidos keičiasi sklandžiai, prognozės yra stabilios, o didelis vienetas parodo, kad paklaidos labai svyruoja ir prognozės yra triukšmingos.

$$AvgDeviation = \frac{1}{n-1} \sum_{i=1}^{n-1} |r_i + 1 - r_i| \quad (4)$$

Čia:

n – mėginių, testų skaičius

r_i – paklaida i pozicijoje

Ši paklaida apskaičiuojama pagal formulę:

$$r_i = x_i - y_i \quad (5)$$

Čia:

r_i – paklaida i pozicijoje

x_i – faktinė vertė i pozicijoje

y_i – modelio prognozuojama vertė i pozicijoje

- Susidūrimų dažnis – tikslas yra autonominiui bepiločiui orlaiviui efektyviai aptikti ir išvengti kliūčių skrydžio metu. Rodiklis tiesiog rodo susidūrimų skaičių per bandomąjį skrydį. Susidūrimas – bet kokia kliūtis kurią autonominis bepilotis orlaivis paliečia skrisdamas. Šis eksperimentas svarbus, nes kuo mažesnis susidūrimų skaičius, tuo sistema yra patvaresnė ir saugesnė.
- Apdorojimo laikas – tikslas yra įvertinti, kaip greitai sistema apdoroja gautą informaciją. Matuojama imant laiko vidurkį iš didelio kiekio vaizdų, per kurį yra apdorojami vaizdai.
- Resursų naudojimas – eksperimentas nurodo kiek ir kokių resursų yra išnaudojama kliūčių vengimo veikimo metu. Kaip jau minėta prieš tai, yra būtina, kad sistema veiktų nedaug kompiuterinių resursų turinčiose sistemose, todėl įvairūs procesorio, disko, kompiuterio atminties ir grafikos apdorojimu įrenginių resursų naudojimo rodikliai turi būti kiek įmanoma mažesni.

Visi šie rodikliai yra tiriami simuliacijoje, kuri atvaizduoja miško aplinką. Rodyklių rezultatai gaunami saugant prognozuojamus gylio žemėlapius ir tikrus, arba vykdant realaus laiko eksperimentus, kurių metu autonominis bepilotis orlaivis klaidžioja po mišką ir vengia kliūčių. Eksperimentiniai tyrimai tarp skirtingai patobulintų modelių yra atliekami su tais pačiais vaizdais ar simuliacijų pozicijomis, norint užtikrinti sąžiningus palyginimus.

Be viso to, yra dvi pagrindinės prognozuojamų gylio žemėlapių rūšys, kurias naudoja įvairūs modeliai:

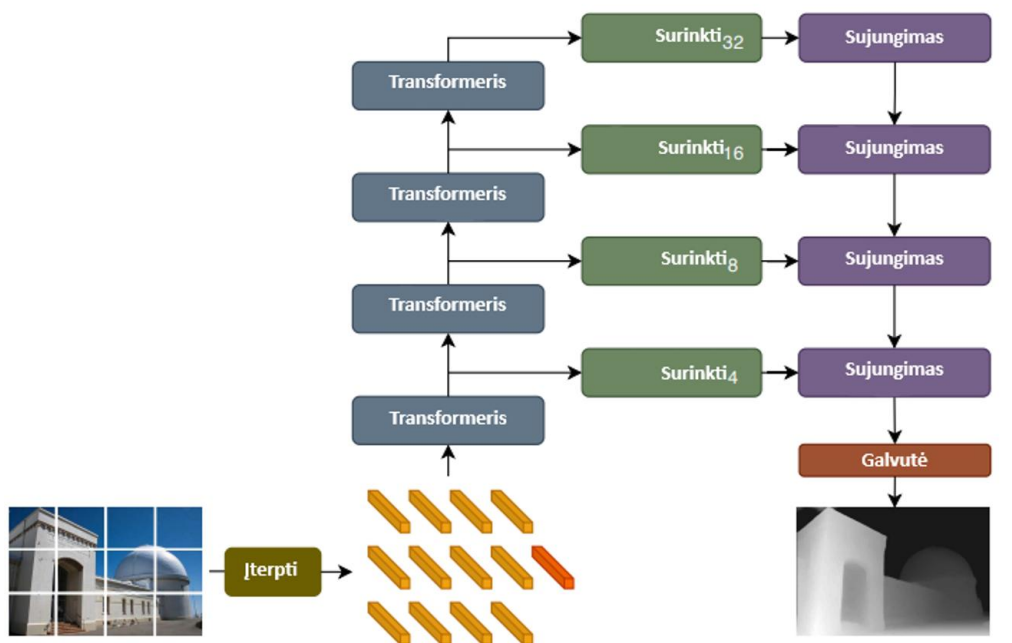
- Santykinio gylio žemėlapiai, kur objektai yra arčiau ar toliau vienas kito, tačiau jie neturi tikrų skaičiavimo metrikų, kaip metras. Jų pagrindinė esmė yra tiesiog reikšmingai ir tvarkingai iš eilės sudėlioti objektus ištrauktus iš RGB vaizdo.
- Metriniai gylio žemėlapiai – nurodo faktinį atstumą nuo kameros iki kiekvieno pikselio realiais vienetais kaip metrais. Tai reiškia, kad jie palaiko tikrus fizinius atstumus ir iškaliko nuoseklų mastelį visuose vaizduose

Tyrimo atveju, metriniai atstumai yra labai svarbūs objektų vengimui, kadangi reikia žinoti ne tik kuris objektas yra arčiau, bet ir koks atstumas iki to objekto yra realiais vienetais, kadangi nuo to priklausys, ar kliūtis yra arčiau iki bepiločio orlaivio negu nustatytas slenkstis. Tik tokiu atveju galima užtikrinti saugia navigaciją po miško aplinką.

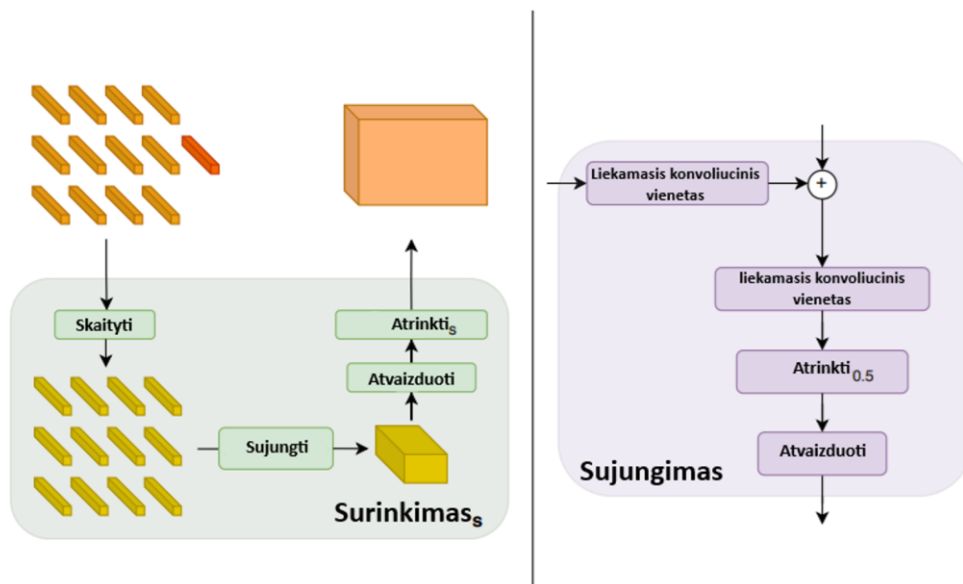
2.4. Modelių, kurie prognozuoja gylio žemėlapius iš monokulinių RGB vaizdų, parinkimas

Norint išvengti kliūčių, reikia kad bepiločio orlaivio sistema gerai suprastų supančią aplinką. Šiai užduočiai išspręsti su monokuliniais RGB vaizdais dažniausiai pasitelkiami metodai, kurie leidžia įvertinti monokulinį gylį. Skirtingai nei stereo ar LiDAR pagrįsti metodai, monokulinis gylio įvertinimas neužtikrina tiesioginių geometrinių matavimų, todėl tai sudėtinga problema, kuri labai priklauso nuo išmokyto modelių. Šiais laikais, yra keli tokių metodų.

Vienas iš jų – tankus prognozavimo transformatorius (angl. dense prediction transformer), kuris tradicinius konvoliucinius neuroninius tinklus pakeičia transformerių pagrindu [12]. Pagrindinė metodo idėja - panaudoti transformatoriais pagrįstas reprezentacijas tankioms prognozavimo užduotims, tokioms kaip gylio įvertinimas ir semantinė segmentacija (žr. **12 pav.**, **13 pav.**).



12 pav. Tankaus prognozavimo transformatoriaus architektūra [12]



13 pav. Tankaus prognozavimo transformatoriaus surinkimo ir sujungimo architektūra [12]

Architektūroje parodyta, kaip įvesties vaizdas yra transformuojamas į žetonus (oranžinė spalva) išskiriant nepersidengiančius fragmentus ir atliekant jų suplokštinto atvaizdavimo linijinę projekciją, arba naudojant „ResNet-50“ požymių ištraukiklį. Vaizdo įterpimas papildomas poziciniu įterpimu ir pridedamas nuo fragmento nepriklausomas nuskaitymo žetonas (raudonas). Žetonai praeina per kelis transformavimo etapus ir tada iš skirtingų etapų vėl visi žetonai yra surenkami į vaizdą primenantį atvaizdavimą su keliomis skiriamosiomis gebomis (žalia spalva). Suliejimo moduliai (violetiniai) palaipsniui sujungia ir padidina atvaizdavimų rodiklį, kad sugeneruotų tikslią prognozę. Dėl šio metodo transformerių struktūros, jis sugeba pateikti labiau globaliai nuosekliai ir detalesnes prognozes, palyginant su visiškai konvoliuciniais tinklais, nes jie apdoroja informaciją santykinai didele skiriamąja geba, įtraukdami globalų kontekstą. Tai pagerina gylio įvertinimo užduočių našumą net iki 28 %, palyginant su ankstesniais pažangiaisiais konvoliuciniais metodais. Tačiau nepaisant šių privalumų, šis metodas turi vieną didelį apribojimą, susijusį su šiuo tyrimu. Šis metodas numato

tik santykinį gylį, o tai riboja jo gebėjimą pateikti tikslų metrinį atstumo įvertinimą, reikalingą bepiločio orlaivio kliūčių vengimui.

Kita modelio architektūra skirta gylio įvertinimui iš monokulinių vaizdų „MiDaS“ [13]. Ji išplečia tankaus transformatorių prognozavimo modelių idėjas įtraukdama transformatoriais pagrįstus tinklus, siekiant pagerinti bendrą scenos supratimą. Tai leidžia prognozuoti tiek artimų, tiek tolimų nuotolių objektus, todėl gaunami vizualiai aiškūs ir nuoseklūs gylio žemėlapiai. Pagrindinis dėmesys šios architektūros atveju yra patikimas kryžminių duomenų rinkinių apibendrinimas. Kitaip tariant, vietoj to, kad būtų optimizuojamas našumas viename duomenų rinkinyje, modelis yra apmokomas su daug skirtingų duomenų rinkinių kaip „ReDWeb“, „DIML Indoor“, „3D Movies“, „MegaDepth“, „WSVD“, „DIW“, „ETH3D“, „Sintel“, „KITTI“, „NYUDv2“, „TUM-RGBD“, naudojant mastelio ir poslinkio invariantinę nuostolių funkciją (angl. scale and shift-invariant loss function), kuri leidžia mokytis iš įvairių duomenų rinkinių, nereikalaujant nuoseklių faktinių gylio skalių. Dėl to pasižymi gerais tyrimų rezultatais įvairiose vidaus ir lauko aplinkos scenarijuose.

Tačiau šiame metode, kaip ir tankus prognozavimo transformatoriuje yra naudojamas santykinis gylio įvertinimas. Tai reiškia, kad modelis gali tiksliai prognozuoti scenos struktūrą ir nurodyti, kuris objektas yra arčiau ar toliau, bet jis nepateikia metrinio gylio įvertinimo, kas suteikia daug problemų norint tiksliai įvertinti atstumą iki netoliese esančių kliūčių ir saugiai jų vengti su bepiločiu orlaiviu. Be viso to, tyrimu metu yra nustatytos kelios pagrindinės klaidos ir gedimų atvejai (žr. 14 pav.).



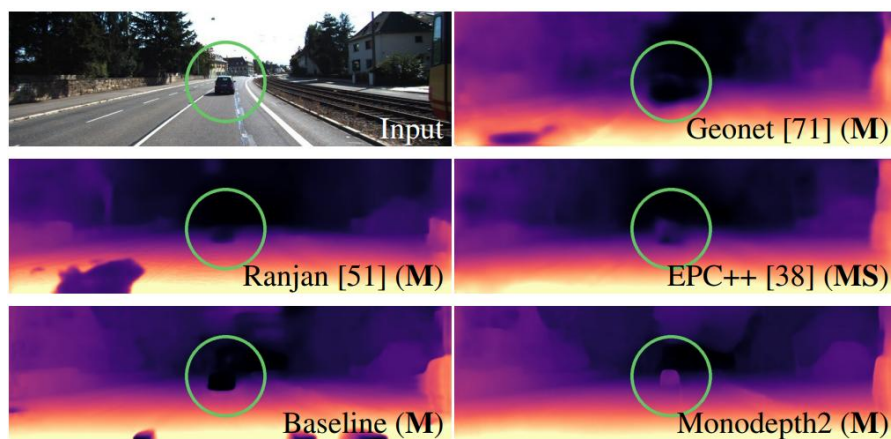
14 pav. Gedimų atvejai. Subtilūs santykinio gylio išdėstymo trūkumai arba trūkstamos detalės paryškintos žalia spalva [13]

Vieni iš pagrindinių, tai kai vaizdas pasuktas 90 laipsnių, kaip matoma viršutiniame kairiajame pavyzdyje, modeliui nepavyksta atkurti įžeminimo plokštumos. Be to, kaip matoma viršutiniame dešiniajame pavyzdyje, modelis natūraliai turi paklaidą - apatinės vaizdo dalys yra arčiau kameros nei viršutinės dalys. Kita problema matoma antroje eilutėje, kur paveikslai, nuotraukos, veidrodžiai dažnai yra neatpažįstami kaip tokie, ir tinklas įvertiną gylį, kuris pavaizduotas ant veidrodžio, o ne pagal patį veidrodį. Tačiau tai minimalios problemos, ir bepiločių orlaivių objektų vengimo miške atvejų yra įmanomi patobulinimai, ar beveik visiškai jie neegzistuoja. Tačiau yra didesnės problemos parodytos likusiose eilutėse, kur ryškūs kraštai gali sukelti haliucinacijas, gylio netolygumus, plonos struktūros gali būti nepastebėtos, o santykinis gylio išdėstymas tarp nesusijusių objektų kai kuriais atvejais gali būti netikslus. Rezultatai būna neryškūs fono srityse, o tai galima paaiškinti ribota įvesties vaizdų skiriamąja geba ir netobulu fono atpažinimu. Tokios haliucinacijos problemos ir neturėjimas metrinio gylio įvertinimo labai trukdo bepiločių orlaivių kliūčių vengimui, kadangi gali per arti ar per toli sustoti nuo esamos kliūties arba jei tai labai plona šaka, jos išvis neatpažinti, kas sukeltų daug susidūrimų.

„Monodepth2“ yra kitas gylio įvertinimo modelis. Skirtingai nei nuo prieš tai minėtų modelių, šis netaiko prižiūravimo mokymosi. Vietoj to, kad pasikliautų tikrojo gylio anotacijomis „Monodepth2“ taiko savarankišką prižiūravimo mokymosi strategiją, kai gylis apmokomas netiesiogiai iš monokulinių vaizdo įrašų arba stereo sekų, atpažindamas scenos geometriją ir žinomą ar numatomą kameros judesį. Jo mokymo tikslas – turint tikslinį vaizdą ir šaltinio vaizdus išmokyti numatyti objektų gylį vaizde, sumažindamas foto metrinę paklaidą (angl. photometric error) tarp tikslinio vaizdo ir jo sintezuotų vaizdų. Pats modelis naudoja „U-Net“ stiliaus kodavimo-dekodavimo architektūrą su „ResNet-18“ pagrindu. Su šiuo modeliu yra pristatytos trys pagrindinės naujovės, kurios pagerina gylio prognozavimo kokybę [5]:

- Automatinis maskavimas (angl. Auto-Masking), kuris užmaskuoja tarp kadru nesikeičiančius pikselius. Tai padeda sutelkti mokymąsi į galiojančius signalus ir išvengti tinklo painiavos su pikseliais kurie pažeidžia judėjimo prielaidas.
- Visos skiriamosios gebos kelių mastelių atranka (angl. Full-Resolution Multi-Scale Sampling), kuris nurodo, kad gylis yra prognozuojamas keliais masteliais, tačiau perprojektavimo paklaidos visada apskaičiuojamos esant visai įvesties skiriamajai gebai. Tai padeda išvengti problemų kaip tekstūrų artefaktai ar skylės tekstūros regionuose, kurios dažnai pasitaiko skaičiuojant nuostolius esant mažai skiriamajai gebai.
- Minimalus perprojektavimo nuostolis (angl. Minimum Reprojection Loss), skirtas kai modelis rekonstruoja tikslinį vaizdą iš kelių šaltinių vaizdų. Ir vietoj to, kad imtų vidurkinę paklaidą, šis metodas parenka minimalią paklaidą vienam pikseliui. Taip išvengiama baudų, kai yra uždengtos sritys ar visuose kadruose nematomi pikseliai. Pagrindinė idėja yra ta, kad pasitikima tik geriausiai atitinkančių kiekvieno pikselio vaizdo nuobauda.

Be šių naujovių yra patobulinti keli ankstesni metodai, įskaitant patikimesnes nuostolių funkcijas ir uždengimų tvarkymą, kurie pagerina mokymo stabilumą ir gylio kokybę. Tai pasireiškia ir rezultatuose, kuriuose yra pasiekiami pažangiausių rezultatų, lyginant su ankstesniais metodais. Rezultatai ypač pasimato lyginant su kitais metodais, kur „Monodepth2“ prognozuoja daugiau detalių ir objektų, negu anksčiau tyrinėti metodai, net įskaitant tokius, kurie specifiskai yra kuriami, kad kurtu gylio žemėlapius su judančiais objektais (žr. **15 pav.**). Paveikslėlyje galima pamatyti, kad mašina yra prastai prognozuojama visose kituose modeliuose išskyrus „Monodepth2“.



15 pav. Gylio prognozavimas, tarp „Monodepth2“ modelio ir kitų modelių [5]

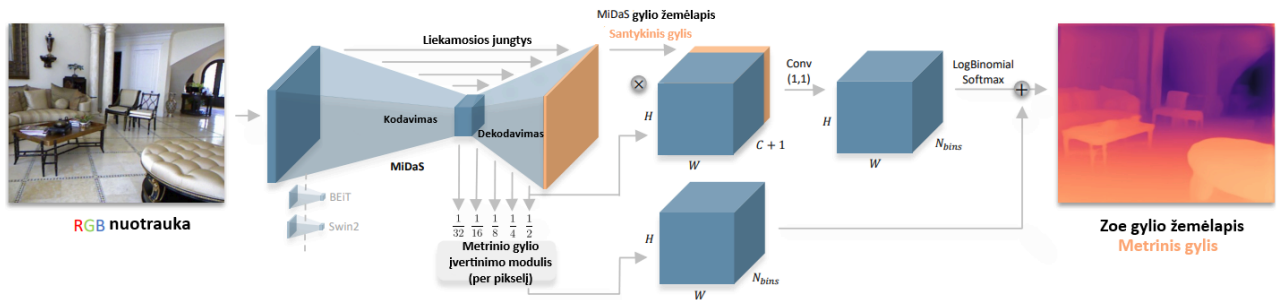
Nepaisant šių pažangumų gylio prognozavimo srityje, šis metodas turi keletą apribojimų. Pirmiausia, kaip ir visi praeiti metodai, šis metodas neturi metrinio gylio įvertinimo. Tai kyla iš to, kad modelis apmokytas su savarankišką prižiūrimo mokymosi strategija ir dėl to negali nustatyti mastelio be papildomų technikų. Taip pat šis metodas nesugeba išmokti ir prognozuoti gerų gylių iškraipytais, atspindintais ar spalvomis prisotintiniems vaizdams, kurie gali pakenkti bepiločio orlaivių kliūčių vengimą (žr. **16 pav.**).



16 pav. Gedimų atvejai, kur „Monodepth“ blogai prognozuoja gylio žemėlapius dėl atspindinčių objektų (viršus) ar sudėtingų formų (apačia) [5]

Kitas modelis yra „ZoeDepth“, kuris išsprendžia esminį ankstesnių monokulinių gylio įvertinimo metodų apribojimą – metrinio gylio prognozavimą [14]. Nors praeiti modeliai, kaip „MiDaS“ ir „Monodepth2“ pateikia santykinai tikslius gylio įverčius, dėl šio metrinio gylio prognozavimo trūkumo, šie metodai negali būti pritaikyti bepiločių orlaivių vengimų užduotims nuo tam tikrų atstumų, nenaudojant daug papildomų permokymų ar struktūros pakeitimų. „ZoeDepth“ turi dvi pagrindines sistemas (žr. **17 pav.**):

- Kodavimo-dekodavimo architektūros sistema, skirta santykiniam gylio įvertinimui, naudojant standartinę mokymo schemą. Modelis remiasi stipriu santykinio gylio sistemos pagrindu, kaip „MiDaS“ ir su šia sistema yra iš anksto apmokoma 12 duomenų rinkinių, kurie lemia gerą santykinio gylio prognozavimą.
- Prie kodavimo-dekodavimo įrenginio architektūros yra pridedama metrinio gylio įvertinimo moduliai, kurie yra tiksliai derinami su metriniais gyliais iš dviejų duomenų rinkinių. Pridėjus šiuos modulius, modelis gali išmokti metrinį gylį, tuo pačiu pasinaudodamas santykinio gylio išankstinio mokymo privalumais.

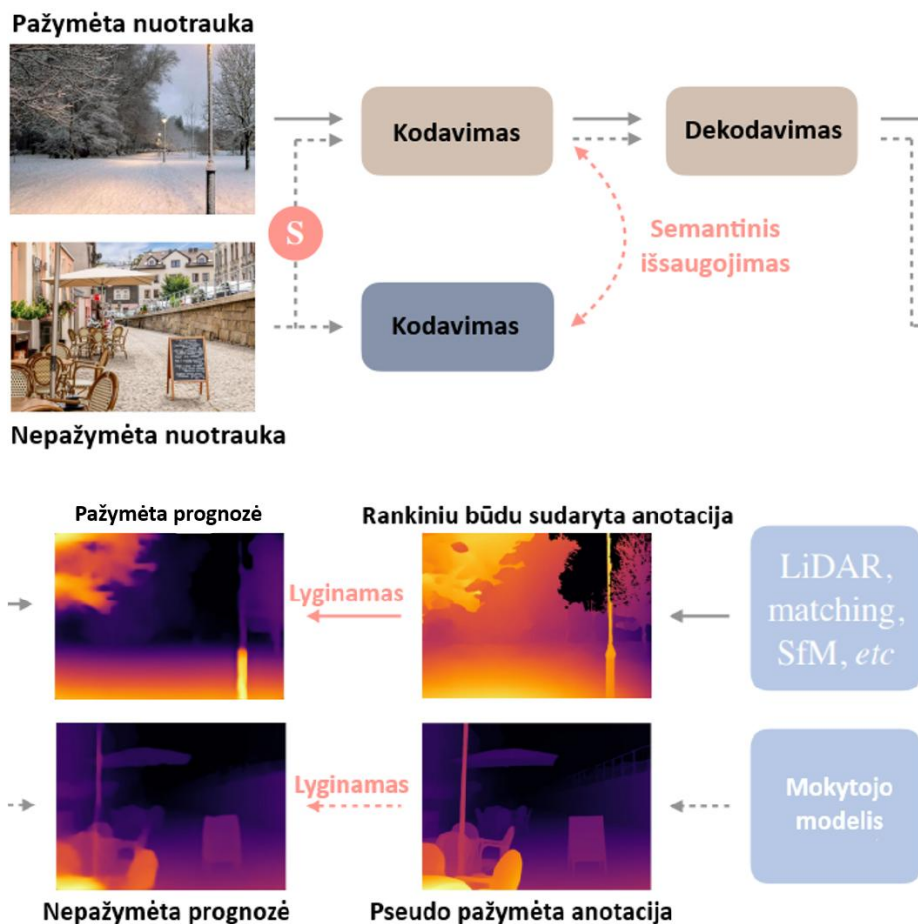


17 pav. „ZoeDepth“ architektūra. RGB vaizdas tiekiamas į „MiDaS“ gylio įvertinimo sistemą prie kurios yra prijungti metrinio gylio įvertinimo moduliai [14]

Vienas iš didžiausių šio metodų privalumų yra nulinio kadro gylio įvertinimas (angl. zero-shot depth estimation) kurio tikslas yra numatyti vaizdų gylį nematytoje aplinkoje, nereikalaujant konkrečiai užduočiai skirtu duomenų rinkinio. Kaip minėta tyrime, net apmokius su vidaus ir lauko domenais kartu, „ZoeDepth“ demonstruoja žymiai geresnius rezultatus tik vidaus ar tik lauko aplinkose negu ankstesni moderniausi modeliai, kurie specifiškai yra apmokyti tik su vienais iš dviejų duomenų rinkinių.

Apskirtai, „ZoeDepth“ modelis yra labai svarbus pasiekimas gylio įvertinimo iš monokulinių vaizdų srityje. Jis sujungia santykinio ir metrinio gylio prognozavimus, kas yra labai aktualu tokiose sistemose kaip bepiločių orlaivių kliūčių vengimas, kur tikslus atstumo įvertinimas yra būtinas. Tačiau šis metodas turi pora problemų – modelio sudėtingumas. Pridėjus metrinio gylio įvertinimo modulius, modelis tampa pakankamai sudėtingas. Todėl tiek apmokymai, tiek jo panaudojimas reikalauja daug kompiuterinių resursų, kurių dažniausiai nėra bepiločiuose orlaiviuose. Dėl to gali iškilti problemų su modelio reagavimu ir kliūčių atsitrenkimu, kas yra didelė problema šio tyrimo atveju.

Kita modelio architektūra, kuri pristato monokulinių gylio įvertinimo metodą yra „Depth Anything“. Šiam modeliui daugiausia dėmesio yra skiriama gebėjimui gerai dirbti su naujais, nematytais duomenimis, su kuriais jis nebuvo apmokytas, naudojant didelio masto nepažymėtus duomenis [15]. Skirtingai nuo anksčiau paminėtų metodų, kurie daugiausiai rėmėsi sužymėtais duomenų rinkiniais, šis metodas yra apmokomas su 1.5 milijonų sužymėtų nuotraukų ir daugiau nei 62 milijonais nesužymėtų nuotraukų. Tai yra daroma, nes sužymėtų duomenų rinkinių įvairovė ir jų kiekis yra ribotas. Todėl, tyrimo metu yra nutarta sujungti pažymėtus ir nepažymėtus vaizdus, kad apmokytų modelį, galintį apskaičiuoti gylį be papildomo tikslinimo. Šis metodas remiasi mokytojo ir mokinio sistema, kurios metu mokytojo modelis, kuris yra „MiDaS“ santykinio gylio stiliaus metodas, apmokomas naudojant pažymėtus gylio duomenis. Tada, šis naudojamas generuoti gylio prognozes didelei nepažymėtų vaizdų kolekcijai, efektyviai sukuriant pseudo pažymėtas gylio anotacijas. Šie pseudo pažymėti vaizdai sujungiami su originaliu pažymėtu duomenų rinkiniu ir tada yra apmokomas mokinio modelis (žr. **18 pav.**).



18 pav. Apmokymo schema. Ištinis linija: pažymėtų vaizdų srautas, punktyrinė linija: nepažymėtų vaizdų srautas. S žymi stiprių pertubacijų (pakitimų) pridėjimą prie nepažymėto vaizdo [15]

Po šių apmokymų, „Depth Anything“ modelis išplėstas, kad būtų galima nustatyti metrinio gylio įvertinimą. Tai daroma naudojant „ZoeDepth“ kodo bazę, tiesiog pakeičiant originalų „MiDaS“ kodavimo įrenginį (angl. encoder) į stipresnį „Depth Anything“ ir kelis hiperparametrus. Tada šis apmokomas su vidaus „NYUv2“ ir lauko „KITTI“ duomenų rinkiniais. Rezultatai yra palyginami su „ZoeDepth“ modeliu lentelėje, kurioje abu modeliai prognozuoja gylio žemėlapius su naujais, nematytais duomenimis (žr. **5 lentelė**). Ten yra dvi metrikos:

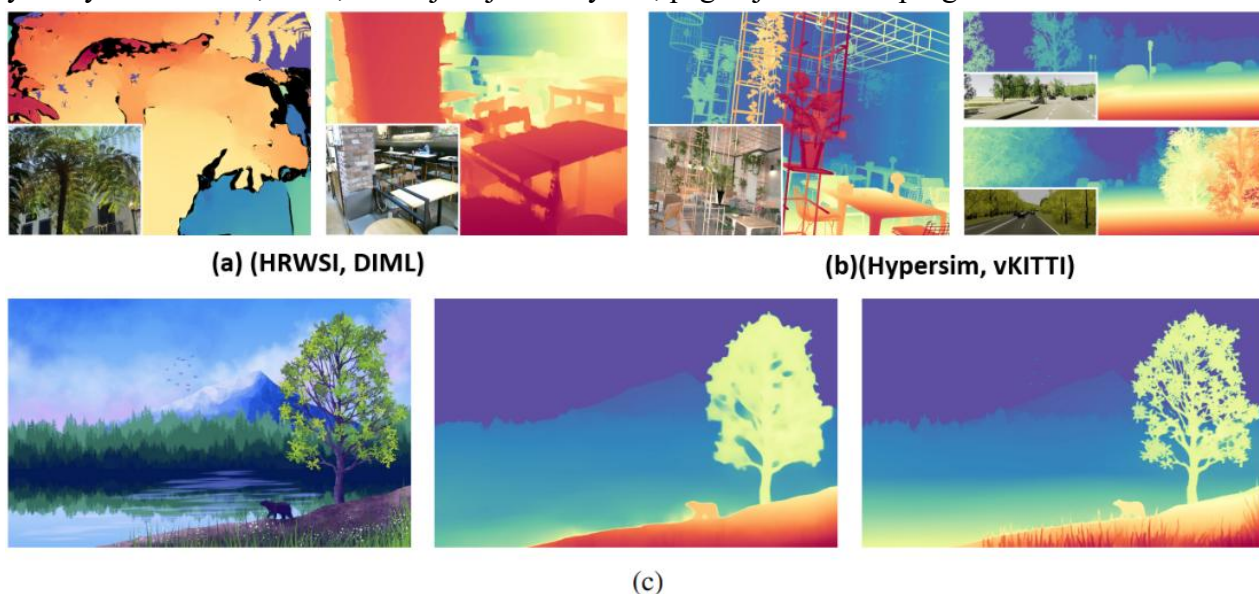
- Absoliuti santykinė paklaida (AbsRel) rodanti kaip toli prognozuojamas gylis yra nuo tikrosios vertės (mažiau yra geriau)
- Delta1 tikslumas ($\delta 1$) – pikselių procentinė dalis, kada prognozės numatytas gylis yra mažesnis už 1.25 karto ir didesnis nei 0.75 karto (daugiau yra geriau).

5 lentelė. Modelių prognozės su nematytais duomenų rinkiniais. Pirmi trys bandymų rinkiniai yra vidaus aplinkos, o pasuktinai du – lauko aplinkos. Siekdami sąžiningo palyginimo yra pateikiama, jog „ZoeDepth“ modelio rezultatai yra atkuriami autorių aplinkoje [15].

Metodas	SUN RGB-D		iBims-1		HyperSim		Virtual KITTI2		DIODE Outdoor	
	AbsRel	$\delta 1$	AbsRel	$\delta 1$	AbsRel	$\delta 1$	AbsRel	$\delta 1$	AbsRel	$\delta 1$
ZoeDepth	0.520	0.545	0.169	0.656	0.407	0.302	0.106	0.844	0.814	0.237
Depth Anything	0.500	0.660	0.150	0.714	0.363	0.361	0.085	0.913	0.794	0.288

Kaip galima pamatyti iš rezultatų, visuose atvejuose, kur „ZoeDepth“ ir „Depth Anything“ prognozuoja gylio žemėlapius su nematytais duomenimis, „Depth Anything“ visada išgauna geresnius, tikslesnius rezultatus.

Naujesnis ir patobulintas ankstesnis modelis „Depth Anything V2“ pristato kitokia mokymo strategiją, kuri pagerina tiek prognozavimo kokybę, tiek kompiuterinių resursų efektyvumą. Kitaip nei ankstesniame modelyje, didelis dėmesys yra skiriamas sintetiniams duomenis, kadangi yra teigiama, jog aukštos kokybės sintetiniai duomenų rinkiniai teikia tikslesnes ir nuoseklesnes gylio anotacijas, palyginti su realaus pasaulio žymėtais duomenimis [16]. Skirtingai nuo realaus pasaulio duomenų rinkinių, kuriuose gylio žymės gali būti triukšmingos, neišsamios arba nenuoseklios dėl jutiklių apribojimų, sintetiniai duomenys siūlo visiškai tikslius gylio žemėlapius, sugeneruotus tiesiai iš imituotos aplinkos. Šis skirtumas pavaizduotas paveiksle (žr. **19 pav.**), kuriame sintetinės gylio žymos yra tikslesnės, todėl, naudojant jas mokymui, pagerėja modelio prognozės.



19 pav. Tikrų vaizdų (a) ir sintetinių vaizdų (b) gylio žymės ir atitinkamos modelio prognozės (c) [16]

Tačiau sintetinių duomenų naudojimas iškelia problemų, kaip domeno atotrūkis tarp sintetinių ir realių vaizdų. Tai reiškia, kad vien sintetiniams duomenimis apmokyti modeliai gali nebūti tinkami naudoti realaus pasaulio aplinkų gylio prognozavimui. Ši problema yra sprendžiama su nauju mokymo procesu (žr. **20 pav.**). Pirmiausia, didelio našumo mokytojo modelis apmokomas vien tik su sintetiniams vaizdais, leidžiant jam mokytis iš tikslios ir be triukšmo gylio priežiūros. Antra, šis mokytojo modelis naudojamas pseudo gylio žymėms generuoti didelio kiekio nepažymėtiems realiams vaizdams, efektyviai perkeliant žinias iš sintetinės srities į realaus pasaulio duomenis. Galiausiai, studentų modeliai apmokomi vien tik šiais pseudo pažymėtais realiais vaizdais, leidžiant jiems pasiekti stiprų apibendrinimą, tiesiogiai nesiremiant sintetiniams duomenimis paskutiniame mokymo etape.



20 pav. „Depth Anything V2“ apmokymo schema [16]

Dėl šio naujo mokymo proceso, ir kitų patobulinimų, kurie leidžia gauti patikimesnius ir detalesnius gylio žemėlapius, šis metodas demonstruoja puikius rezultatus keliuose eksperimentuose, pranokdamas ankstesnius metodus tiek su santykinėmis gylio žemėlapių prognozėmis, tiek su metriniais gylio žemėlapių prognozėmis.

Apskritai, „Depth Anything V2“ šiuo atveju yra pagrindinis kandidatas monokulinių gylio įvertinimo ir tyrimo srityje. Jis išsprendžia daugumą senesnių metodų, kaip „MiDaS“, „Monodepth2“ problemas dėl tik santykinio gylio žemėlapių prognozių ar prastų rezultatų sudėtingų formų aplinkose, kas yra labai svarbu, norint išgauti tikslius gylio žemėlapius miško aplinkoje. Be viso to, turi žymiai įvairesnį mokymo procesą, negu „ZoeDepth“, todėl pagerėja jo prognozės nematytose aplinkose. Bei šis metodas, pagal autorius, yra net daugiau nei 10 kartų greitesnis, negu kiti difuzijomis pagrįsti gylio įvertinimo modeliai, kas suteikia galimybę greitai reaguoti į kliūtis su autonominiu bepiločiu orlaiviu, realiu laiku.

3. Eksperimentinė dalis

Eksperimentinėje dalyje siekiama įvertinti magistro darbo eigoje suprojektuotus bepiločio orlaivio sistemų funkcionalumus. Eksperimentai susideda iš duomenų rinkimo, apdorojimo ir įvairių algoritmų bei modelių testavimų, apmokymų kurie yra susiję su kliūčių aptikimu ir vengimu. Eksperimentai atliekami naudojant simuliacijoje sukurtais duomenimis, o jų našumas matuojamas naudojant anksčiau paminėtus vertinimo rodiklius.

Visiems simuliacinėje aplinkoje atliktiems eksperimentams naudojami tokie kompiuteriniai resursai:

- Procesorius – Intel Core I7-14650HX
- Vaizdo plokštė – Nvidia GeForce RTX 4070 TI SUPER
- Atmintis – 64 GB

Pagrindinis modelis, naudojamas gylio žemėlapiams prognozuoti iš monokulinių RGB vaizdų yra „Depth Anything V2“. Šio gylio žemėlapio su metriniai atstumais patobulinimas turi tris skirtingus modelius:

- Depth-Anything-V2-Metric-Outdoor-Large-hf
- Depth-Anything-V2-Metric-Outdoor-Base-hf
- Depth-Anything-V2-Metric-Outdoor-Small-hf

Su visais šiais trimis modeliais yra eksperimentuojama šioje dalyje.

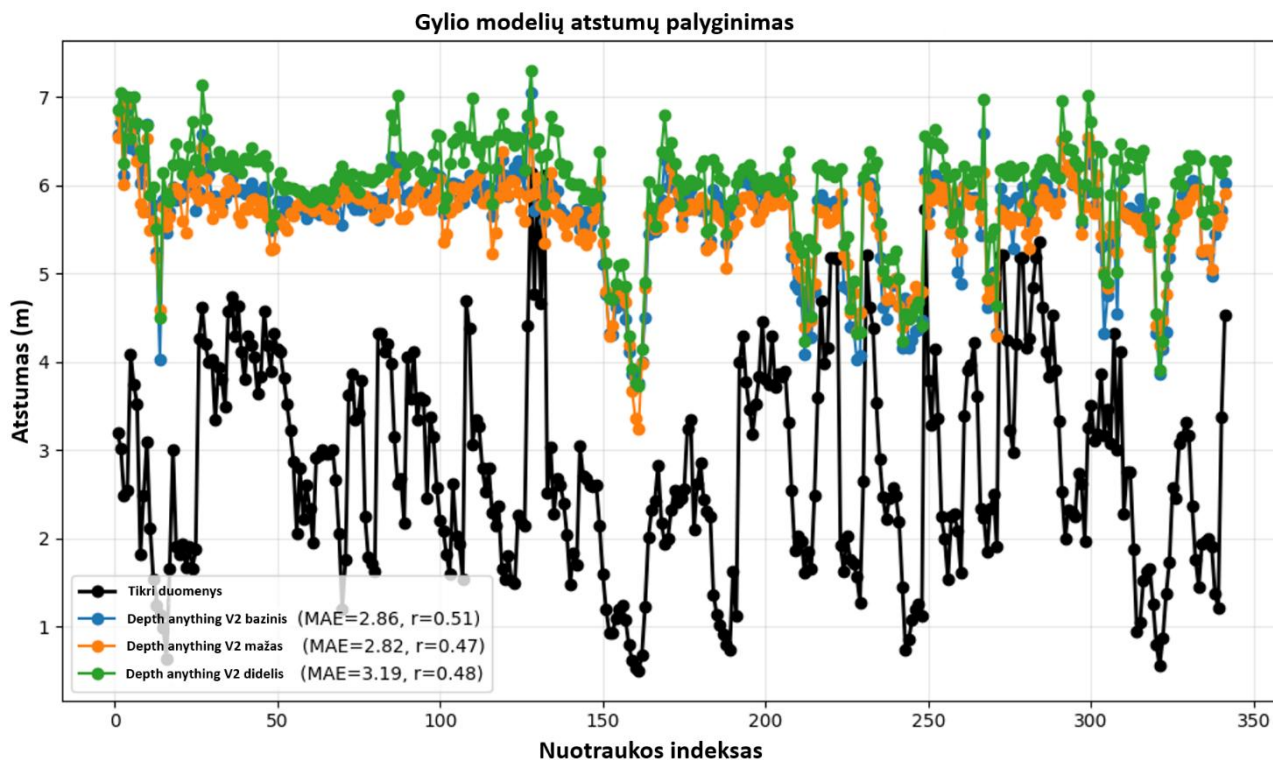
3.1. Bazinių modelių minimalaus gylio įvertinimo eksperimentai

Šiems eksperimentams atlikti yra pasitelkiamos visos metrikos atstumo įvertinimo paklaidai apskaičiuoti. Eksperimentuose yra ištestuojamas minimalus atstumas kiek su baziniais „Depth Anything V2“ modeliais, kiek su modeliais sumažintais per konstantą, kiek su papildomai mokytais simuliacijoje surinktais duomenimis. Mažiausias atstumas apskaičiuojamas paimant vidurkį iš 5% artimiausių pikselių nuotraukoje (žr. **21 pav.**), kadangi imant vieną artimiausią pikselį gali iškilti išimčių ir neatitikimų.



21 pav. Kairėje - RGB nuotrauką, viduryje - gylio nuotrauka, dešinėje - gylio nuotrauka, mažiausio atstumo apskaičiavimuose naudojami pikseliai pažymėti raudonai

Šių modelių eksperimentai vyksta dienos apšvietimo aplinkoje: ir išviso naudojama 341 monokulinių vaizdų. Iš pradžių ištestuojami „Depth Anything“ šeimos baziniai modeliai ir nubraižomi grafikai, kurie parodo mažiausio atstumo tikrus rezultatus, ir juos sulygina su prognozuojamais modelių rezultatais (žr. 22 pav.).



22 pav. Grafikas lyginantis tikrą mažiausią atstumą ir „Depth Anything“ bazinių modelių mažiausius prognozuojamus atstumus

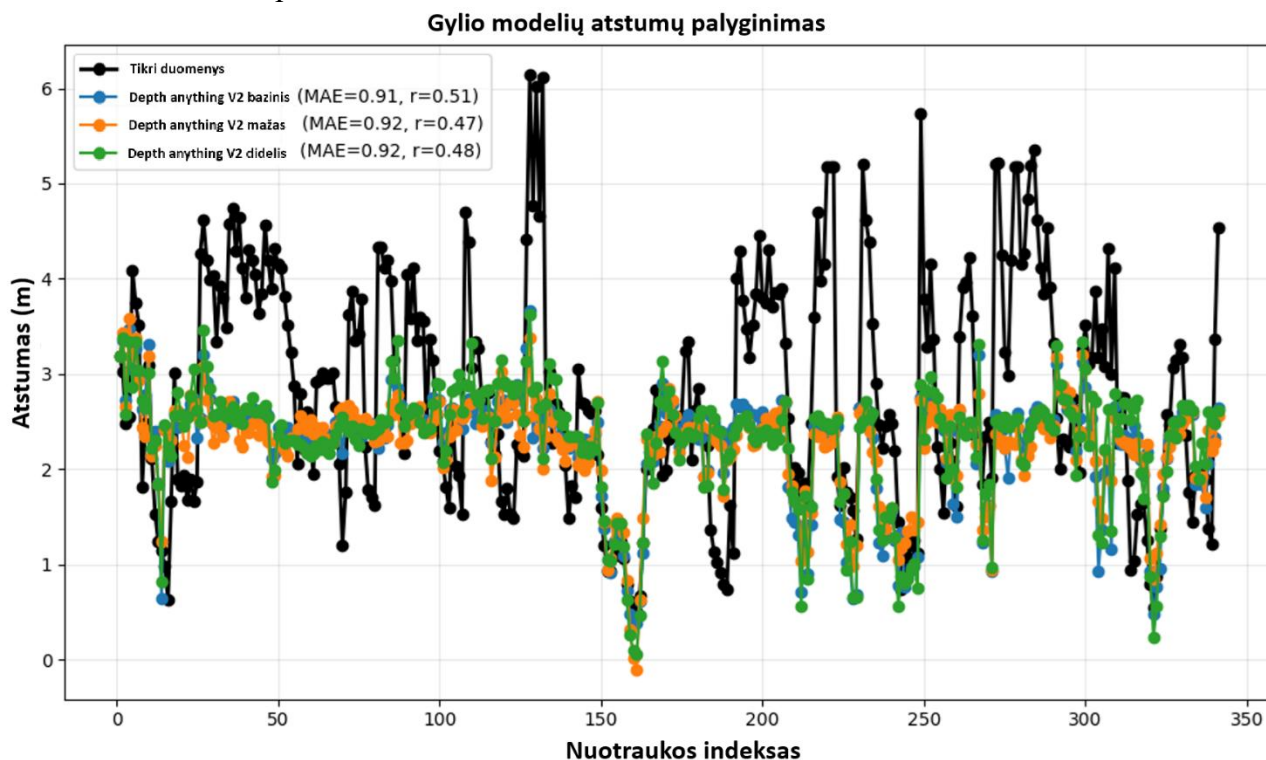
Kaip galime pamatyti iš rezultatų šis modelis visada prognozuoja daug didesnius artimus atstumus, negu jie iš tiesų yra. Atsižvelgus į lentelės rezultatus (žr. 6 lentelė) galima pamatyti, kad visi modeliai pasižymi MAE nuo 2.816 iki 3.187 metrų atstumo. Tai yra blogai, žinant kad tikslas šios eksperimentinės dalies yra vengti labai artimų objektų, todėl skirtumas tarp tikrų ir prognozuojamų atstumų turi būti labai mažas. Taip pat galime pastebėti, kad padidėjęs bazinio modelio dydis specifiškai su „Depth Anything“ modeliais nepagerina našumo, net ir atvirkščiai lemia šiek tiek didesnes paklaidas bei rodo sumažėjusį nuoseklumą.

6 lentelė. Bazinių modelių metrikos

Modelis	MAE	RMSE	STD	Koreliacija	AvgDev
Depth-Anything-V2-Metric-Outdoor-Base-hf	2.86	3.036	1.027	0.509	0.508
Depth-Anything-V2-Metric-Outdoor-Small-hf	2.816	3.004	1.056	0.467	0.529
Depth-Anything-V2-Metric-Outdoor-Large-hf	3.187	3.353	1.047	0.479	0.537

Nors ir „Depth Anything“ modelių metrikos yra prastos, vizualiai jie atrodo turintys panašią formą į faktinius duomenis. Todėl visiems šiems modeliams yra pritaikoma papildoma funkcija – visi jie yra sumažinami per konstantą. Ši konstanta yra sudaryta atimant skirtumą tarp pirmo vaizdo faktinio mažiausio atstumo ir modelio prognozuojamo atstumo visiems vaizdams. Gauti prognozuojami rezultatai abiejų bazinių modelių atvaizduojami grafike (žr. 23 pav.). Iš grafikų, kuriuose modeliai prognozuoja artimiausius atstumus ir tada yra sumažinami per konstantą galime pamatyti, kad „Depth

Anything“ modeliai pateikia prognozes, kurios vizualiai sklandžios, turi gera svyravimą lyginant su tikru atstumu, tačiau dažnai šie modeliai nepakankamai įvertina didelius atstumus ar pervertina mažus, o tai vis tiek rodo neprisitaikiusi metrinį mastelį. Labai didelių skirtumų vizualiai tarp didelio ir mažo modelio nėra pastebėta.



23 pav. Grafikas lyginantis tikrą mažiausią atstumą ir „Depth Anything“ bazinių modelių mažiausius prognozuojamus atstumus su papildoma funkcija

Atsižvelgiant į metrikas, tik MAE ir RMSE pasikeičia, nes kitos metrikos yra nekintamos pastoviams poslinkiams (žr. 7 lentelė).

7 lentelė. Bazinių modelių metrikos su papildoma sumažinimo per konstantą funkcija

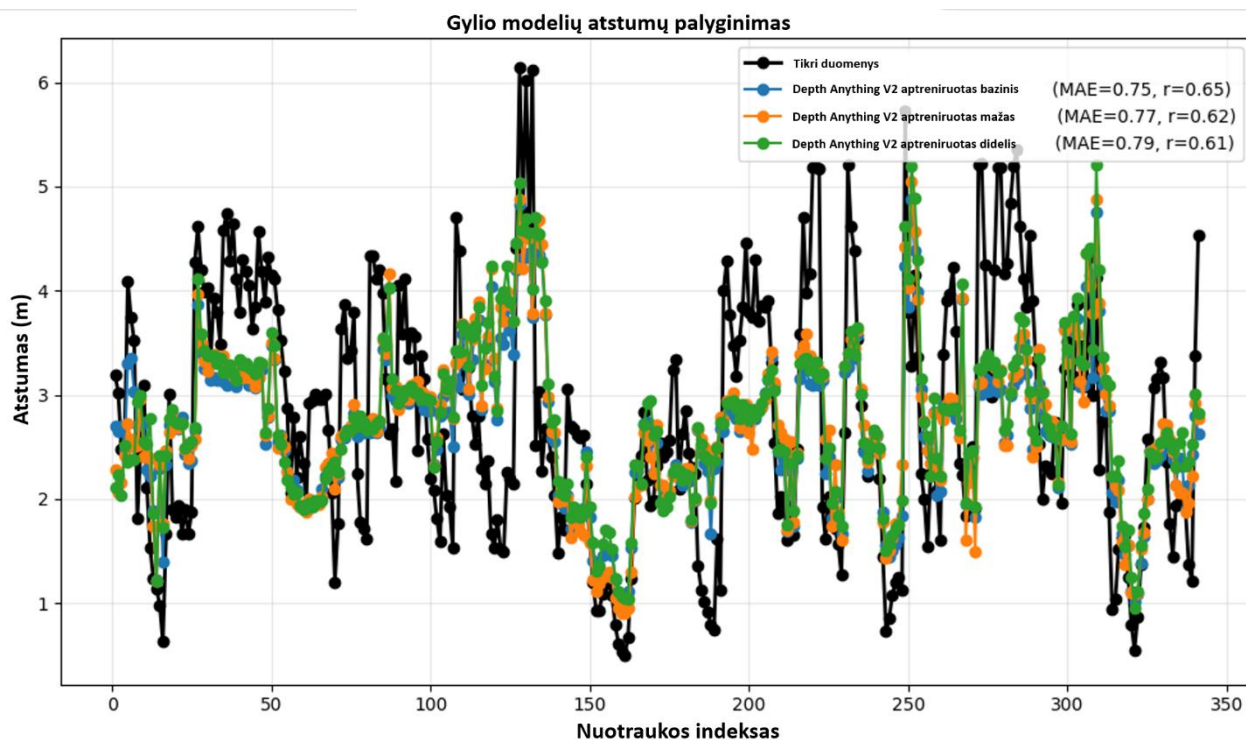
Modelis	MAE	RMSE	STD	Koreliacija	AvgDev
Depth-Anything-V2-Metric-Outdoor-Base-hf	0.912	1.155	1.027	0.509	0.508
Depth-Anything-V2-Metric-Outdoor-Small-hf	0.92	1.186	1.056	0.467	0.529
Depth-Anything-V2-Metric-Outdoor-Large-hf	0.919	1.155	1.047	0.479	0.537

3.2. Papildomai mokytų modelių minimalaus gylio įvertinimo eksperimentai

Tolimesniuose eksperimentuose „Depth Anything V2“ modeliai yra apmokomi su vaizdais sudarytais iš simuliacinių eksperimentų. Apmokymai vyksta su keturiais duomenų rinkiniais:

- 250 vaizdais
- 500 vaizdais
- 750 vaizdais
- 1000 vaizdais

Visi monokuliniai vaizdai yra 518x518 rezoliucijos su gylio žemėlapiais ir metriniais atstumais. Aptreniravus modelius, ir atlikus eksperimentus yra nubraižomi grafikai (žr. **24 pav.**). Vien atvaizdus modelio, apmokyto su 250 vaizdais grafika, lyginanti mažiausius atstumus matoma, kad „Depth Anything V2“ modelių artimiausio atstumo prognozavimas yra panašus, į tų pačių modelių bazinius rezultatus su sumažinimo per konstantą rezultatus, tik papildomai šie apmokyti modeliai dar geriau prisitaiko prie aukštų ir žemų metrių.



24 pav. Grafikas lyginantis tikrą mažiausią atstumą ir „Depth Anything V2“ apmokytų modelių su 250 vaizdais mažiausius prognozuojamus atstumus

Taip pat yra palyginamos šių metodų gautų metrikų rezultatai (žr. **8 lentelė**). Kaip matome, net lyginant su baziniais modeliais, kurie turėjo papildoma sumažinimo per konstantą funkciją, visi papildomai mokytieji modeliai su 250 vaizdais, kurie sudaryti simuliacijoje, gauna žymiai geresnius rezultatus.

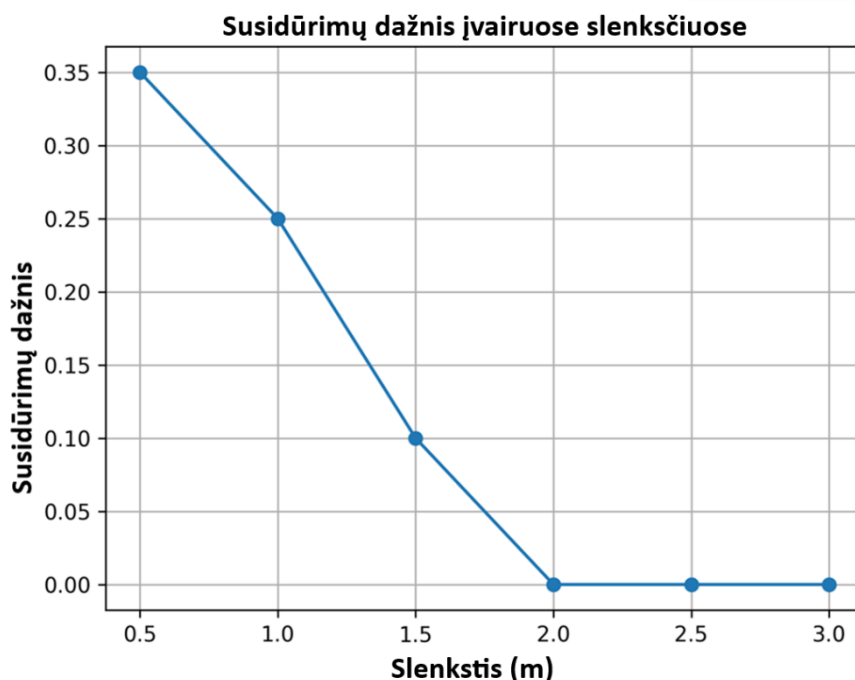
8 lentelė. Bazinių ir papildomai mokytų su 250 vaizdais modelių prognozavimo metrikos

Modelis	Tipas	MAE	RMSE	STD	Koreliacija	AvgDev
Depth-Anything-V2-Metric-Outdoor-Base-hf	Bazinis	2.86	3.036	1.027	0.509	0.508
Depth-Anything-V2-Metric-Outdoor-Small-hf	Bazinis	2.816	3.004	1.056	0.467	0.529
Depth-Anything-V2-Metric-Outdoor-Large-hf	Bazinis	3.187	3.353	1.047	0.479	0.537
Depth-Anything-V2-Metric-Outdoor-Base-hf	Per konstanta	0.912	1.155	1.027	0.509	0.508
Depth-Anything-V2-Metric-Outdoor-Small-hf	Per konstanta	0.92	1.186	1.056	0.467	0.529
Depth-Anything-V2-Metric-Outdoor-Large-hf	Per konstanta	0.919	1.155	1.047	0.479	0.537
Depth-Anything-V2-Metric-Outdoor-Base-hf	250	0.748	0.921	0.912	0.646	0.518
Depth-Anything-V2-Metric-Outdoor-Small-hf	250	0.774	0.942	0.939	0.617	0.532
Depth-Anything-V2-Metric-Outdoor-Large-hf	250	0.79	0.949	0.949	0.607	0.526

Jie pasižymi mažesniu MAE IR RMSE, kurie parodo, kad vidutiniškai nuokrypis tarp realių ir prognozuojamų duomenų yra žymiai mažesnis negu bazinių modelių prognozuojamų duomenų. MAE vidutiniškai geresnis negu bazinis modelis per daugiau nei 3.8 karto ir 1.19 karto geresnis negu

naudojant bazinį modelį su sumažinimo per konstantą funkciją. Papildomai mokyto modelio RMSE pagerėja daugiau nei 3.33 karto lyginant su baziniu modeliu ir 1.24 karto lyginant su baziniu modeliu su sumažinimo per konstantą funkciją. Tai tik įrodo, kad papildomai apmokytas modelis su simuliaciniais kameros vaizdais geriau nustato atstumus tarp esančių objektų negu bazinis modelis ir sumažina labai didelius nuokrypius. Taip pat kitos metrikos irgi truputi geresnės, mažesnis standartinis nuokrypis, kuris nurodo, kad prognozės yra vidutiniškai geresnės, pastovesnės, geresnė koreliacija, kuri pasako, kad prognozuojami atstumai atitinka tikrosios padėties tendenciją. Vidutinis nuokrypis yra panašus, kas parodo kad paklaidos vidutiniškai vis tiek svyruoja gan stipriai. Galima teikti, kad papildomai mokytų modelių rezultatai su specifiniais simuliacinės kameros duomenimis yra geresni.

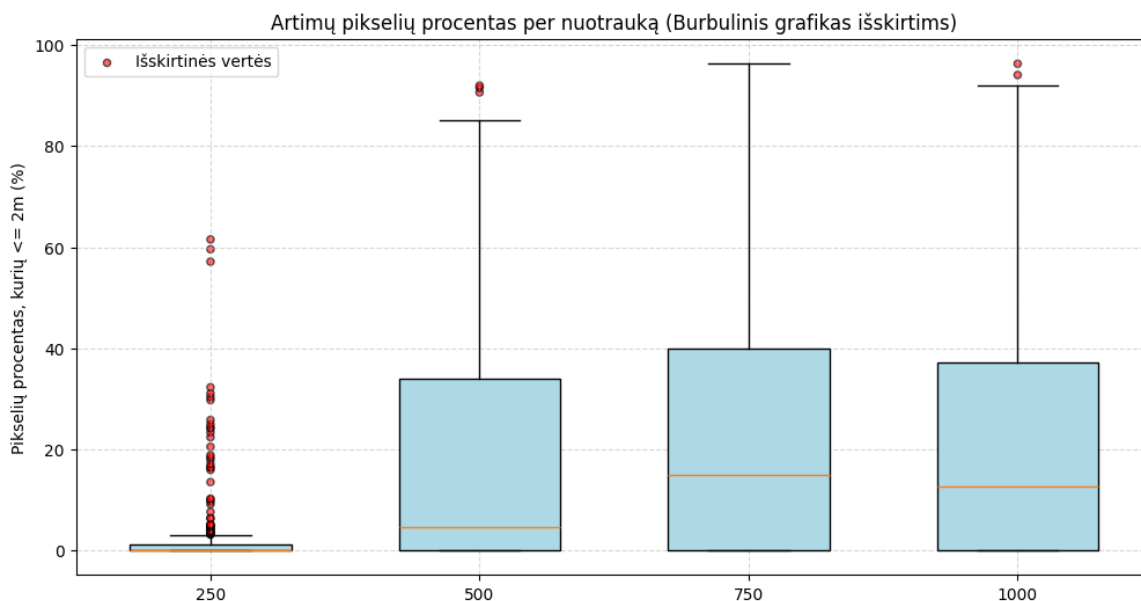
Tačiau šie padaryti eksperimentai parodo tik mažiausius atstumus su nuotraukomis, kurių mažiausi atstumai gali būti nuo 0 iki 10 metrų. Tai nepatvirtina mūsų nusistatyto tikslo - vengti per arti esančių kliūčių skrendant pastoviu greičiu. Norint tai padaryti, pirmiausia reikia nusistatyti pastovų greitį. Šiuo atveju, tai yra 3.5 m/s, kuris yra nustatomas, kaip autonominio bepiločio greitis simuliacijoje. Tada yra sukuriama 1000 eksperimentų, kurių metu autonominis bepilotis orlaivis yra paleidžiamas į mišką iš skirtingų pozicijų. Jo tikslas yra priartėti prie kliūtis per nustatytą atstumo slenkstį (turi būti bent 2% pikselių arčiau negu slenkstis) ir tos kliūtis vengti naudojant įmontuotas gylio prognozavimo sistemas. Jei kliūtis yra neišvengiama ir įvyksta susidūrimas, reiškiasi slenkstis yra per arti. Nustatomi keli slenksčiai – 0.5 metro, 1.0 metras, 1.5 metro, 2.0 metrai, 2.5 metrai, 3.0 metrai. Susidūrimų dažnis naudojant skirtingus slenksčius yra atvaizduojamas grafike (žr. **25 pav.**).



25 pav. Susidūrimų dažnis su skirtingais slenksčiais

Kaip matome iš grafiko, susidūrimai pradingsta visiškai, kai autonominis bepilotis orlaivis pradeda stabdyti esant kliūčiai ties dviem metrais. Turint šią metriką, yra susidaromas nedidelio atstumo duomenų rinkinys. Jis sudarytas iš 200 vaizdų, kurių minimalūs atstumai yra tarp 0 ir 2 metrų. Tai padaroma, nes bepiločiui orlaiviui nerūpi, kaip tiksliai yra išmatuojami tolimesni medžiai, krūmai, kalnai, jam svarbiausia yra patikimai aptikti netoliese esančias kliūtis ir jų vengti. Tada šis naujai sukurtas duomenų rinkinys yra testuojamas su modeliais, apmokytais su 250, 500, 750 ir 1000

vaizdais. Šiuose duomenų rinkiniuose visos nuotraukos turi pikselių arčiau negu du metrai, tačiau jų pasiskirstymas yra įvairus (žr. **26 pav.**). Kaip matome, surinktame duomenų rinkinyje, mažiausiai artimų pikselių turi duomenų rinkinys sudarytas iš 250 vaizdų, o daugiausiai turi iš 750 vaizdų.



26 pav. Artimų pikselių procentas per nuotrauką

Tačiau artimų pikselių artumas ne viską lemia. Kaip matoma iš rezultatų atvaizduojamų lentelėje (žr. **9 lentelė**), geriausiai veikia modeliai, kurie apmokyti su daugiausiai vaizdų. Jų beveik visos metrikos yra geriausios, kas rodo, kad modelių tikslumas priklauso ir ne tik nuo tiek duomenų rinkinyje yra artimų pikselių, bet ir nuo duomenų kiekio.

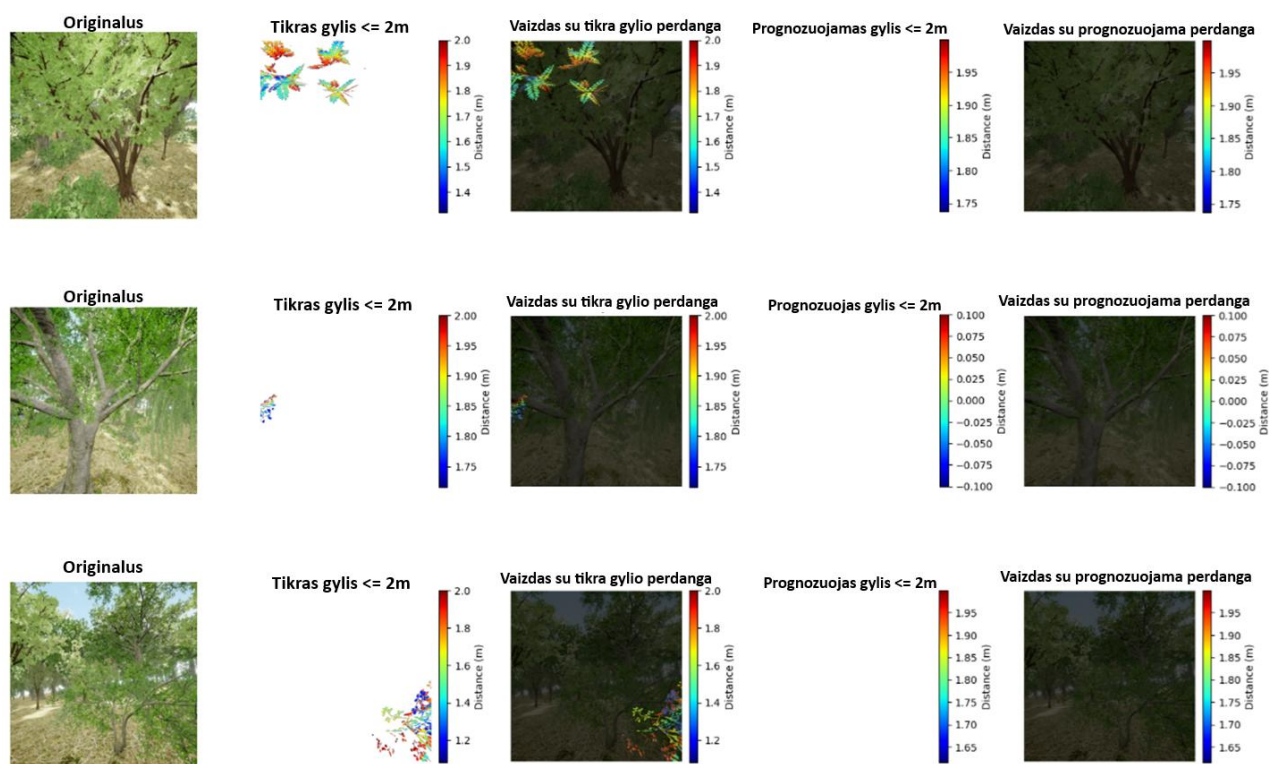
9 lentelė. Papildomai mokytų modelių gautos metrikos su nedidelio atstumo testavimo rinkiniu

Modelis	Tipas	MAE	RMSE	STD	Koreliacija	AvgDev
Depth-Anything-V2-Metric-Outdoor-Base-hf	250	0.387	0.481	0.37	0.637	0.256
Depth-Anything-V2-Metric-Outdoor-Small-hf	250	0.403	0.509	0.377	0.607	0.235
Depth-Anything-V2-Metric-Outdoor-Large-hf	250	0.367	0.44	0.302	0.732	0.217
Depth-Anything-V2-Metric-Outdoor-Base-hf	500	0.17	0.259	0.258	0.811	0.183
Depth-Anything-V2-Metric-Outdoor-Small-hf	500	0.189	0.302	0.302	0.77	0.199
Depth-Anything-V2-Metric-Outdoor-Large-hf	500	0.151	0.209	0.207	0.87	0.16
Depth-Anything-V2-Metric-Outdoor-Base-hf	750	0.146	0.224	0.22	0.862	0.165
Depth-Anything-V2-Metric-Outdoor-Small-hf	750	0.156	0.244	0.241	0.848	0.163
Depth-Anything-V2-Metric-Outdoor-Large-hf	750	0.136	0.197	0.19	0.893	0.147
Depth-Anything-V2-Metric-Outdoor-Base-hf	1000	0.142	0.212	0.206	0.88	0.151
Depth-Anything-V2-Metric-Outdoor-Small-hf	1000	0.15	0.24	0.238	0.854	0.159
Depth-Anything-V2-Metric-Outdoor-Large-hf	1000	0.134	0.201	0.198	0.889	0.145

3.3. Papildomai mokytų modelių artimų pikselių analizė

Anksčiau daryti tyrimai parodo, kaip papildomai mokytas modelis su papildomais duomenimis gauna žymiai tikslesnius ir geresnius rezultatus. Tačiau šie tyrimai neapibendrina labai svarbių problemų: ko prognozuojami modeliai nemato. Kadangi tyrimai turi būti pritaikyti miškingose vietose, kur yra labai daug sudėtingų formų, reikia atsižvelgti į kliūtis, kurių šie metodai prognozuodami neįžvelgia ar neaptinka. Problemas galima pastebėti vizualiai pažvelgus į kai kurių vaizdų prognozuojamus gylio

žemėlapius (žr. 27 pav.). Galima pamatyti, kad modelio prognozės neatitinka tikro gylio žemėlapio: visiškai neprognozuoja ar praleidžia kai kurias šakas ir lapelius. Tai yra didelė problema autonominių bepiločių orlaivių sistemai, kadangi visos šios smulkios kliūtis gali užkirsti kelią judėjimui.



27 pav. Modelio papildomai mokyto su 1000 vaizdų artimo gylio prognozės palyginimas su tikru gyliu

Norint specifiskai atsižvelgti į šias problemas, yra sudaromas dar vienas 200 vaizdų artimu atstumų rinkinys, kuris turi daug įvairių smulkių objektų, kaip plonos šakos ar smulkūs lapeliai. Norint labiau atsižvelgti į artimų pikselių prognozes, yra atliekami papildomi eksperimentai su artimais atstumais. Prie to paties susidaromos papildomos metrikos, padedančios geriau prognozuoti rezultatus:

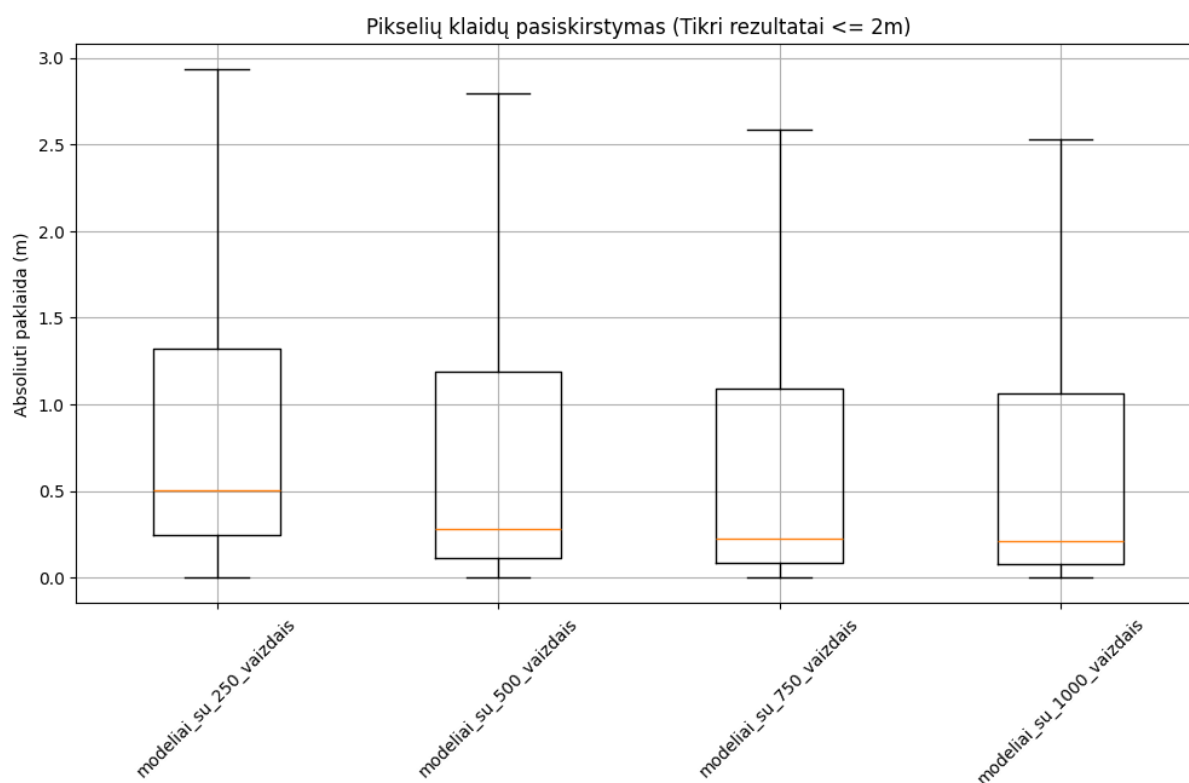
- Jautrumas (angl. recall) – matuoja modelio teisingai aptiktų artimo nuotolio pikselių procentinę dalį. Šiuo atveju pikselių arčiau 2 metrų. Kuo didesnė vertė, tuo daugiau pikselių yra aptinkama.
- Persidengimo koeficientas (angl. intersection over union) – rodo prognozuojamų pikselių regionų ir tikrųjų pikselių regionų persidengimą, kurių atstumas arčiau negu 2 metrai. Didesnė vertė nurodo tikslų ir gerą persidengimą.
- F1 balas (angl. F1 score) – rodo, ar modelis išlaiko gera pusiausvyrą tarp kliūčių aptikimo ir per didelio klaidingų teigiamų rezultatų (angl. false positives), kurie gali lemti nestabilių objektų vengimą. Kuo didesnė vertė, tuo labiau yra palaikoma pusiausvyrą
- Vidutinio absoliutaus atstumo klaida (angl. mean absolute distance error) – matuoja vidutinį absoliutų skirtumą tarp numatytų ir tikrojo gylio verčių artimiems pikseliams (iki 2 metrų). Kuo mažesnė paklaida, tuo geresnis atstumo suvokimas.

Šie keturi rodikliai sudaro patikimą vertinimo sistemą, kuri nurodo kaip gerai modelis prognozuoja artimus atstumus. Pasitelkiant šiomis metrikomis ir nauju duomenų rinkiniu yra atliekami papildomi eksperimentai su papildomai mokytais modeliais. Rezultatai yra pateikiami lentelėje (žr. **10 lentelė**).

10 lentelė. Papildomai mokytų modelių artimų atstumų eksperimentų rezultatai.

Modelis	Jautrumas (%)	Persidengimo koeficientas (%)	F1 balas (%)	Vidutinio absoliutaus atstumo paklaida (m)
Modelis papildomai mokytas su 250 vaizdais	49.4	42.1	59.2	0.165
Modelis papildomai mokytas su 500 vaizdais	61.4	51.3	67.8	0.125
Modelis papildomai mokytas su 750 vaizdais	64.7	53.8	69.9	0.108
Modelis papildomai mokytas su 1000 vaizdais	64.6	54.2	70.3	0.103

Kaip galima pastebėti dauguma atveju tik truputi negu puse pikselių arčiau 2 metrų yra aptinkami bei persidengimo koeficientas vidutiniškas ar net silpnas, žiūrint į modelius derintus su 250 vaizdais. F1 balas taip pat vidutiniškas, kas reiškia, kad yra per daug klaidų ar klaidingų aliarmų. Vidutinio absoliutaus atstumo paklaida vidutiniškai yra 0.125 metro. Šis vienetas yra priimtinas, tačiau būtų gerai jį pagerinti. Taip pat yra susidaromi kiekvieno metodo pikselių klaidų pasiskirstymo dėžutės diagrama (angl. boxplot), kurioje galima pamatyti, kad modelių klaidų sklenda tampa siauresnė apmokant vis su didesniu kiekiu vaizdų, tačiau didelės paklaidos vis dar egzistuoja (žr. **28 pav.**). Nors ir vidutiniškai, geriausių modelių vidutinė paklaida yra nuo 0.1 iki 1.1 metro atsiranda išskirtinių verčių, kur modeliai prastai prognozuoja artimus atstumus, su iki 2.5 metro paklaida. Tai labai nestabilu tokioms sistemoms kaip autonominių bepiločių orlaivių objektų vengimas.



28 pav. Papildomai mokytų modelių pikselių klaidų pasiskirstymo dėžutės diagrama

3.4. Papildomai mokytų modelių specifiskai artimiems atstumams eksperimentai

Siekiant pagerinti papildomai mokytų modelių rezultatus yra susidaromi nauji treniravimo metodai, kurie turi specifines praradimo ir svorių funkcijas. Iš viso yra dvi svorių funkcijos ir trys naujos praradimo funkcijos. Svorių funkcijos yra tokios:

- Eksponentinė funkcija:

$$weight = e^{-\frac{strenght}{near} * target} \quad (6)$$

Čia:

weight – gylis atstumo svoris

strenght – mastelio koeficientas (stiprumas)

near – netolimo gylis atstumas (metrais)

target – realaus vaizdo gylis

- Sigmoidinė funkcija:

$$weight = strenght * \frac{1}{1 + e^{-3(near - target)}} \quad (7)$$

Čia:

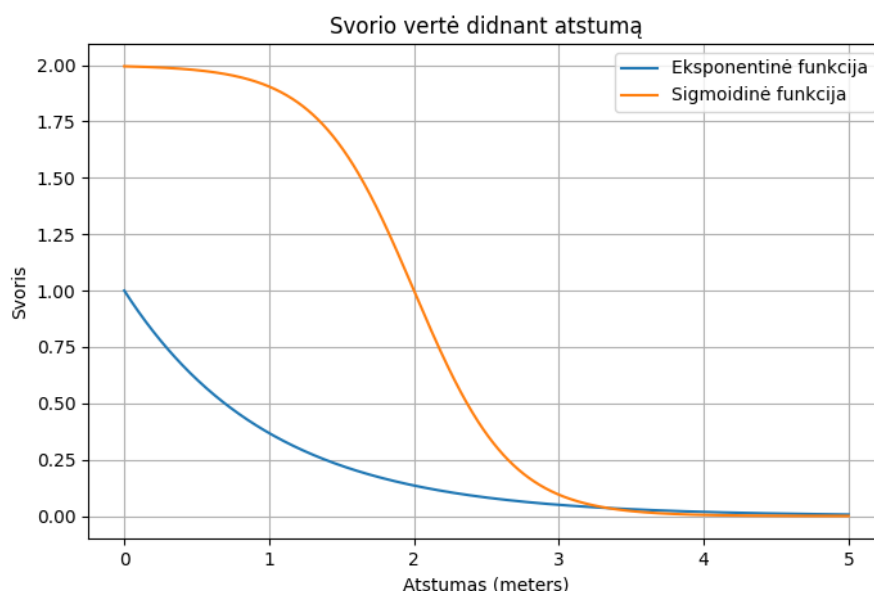
weight – gylis atstumo svoris

strenght – mastelio koeficientas (stiprumas)

near – netolimo gylis atstumas (metrais)

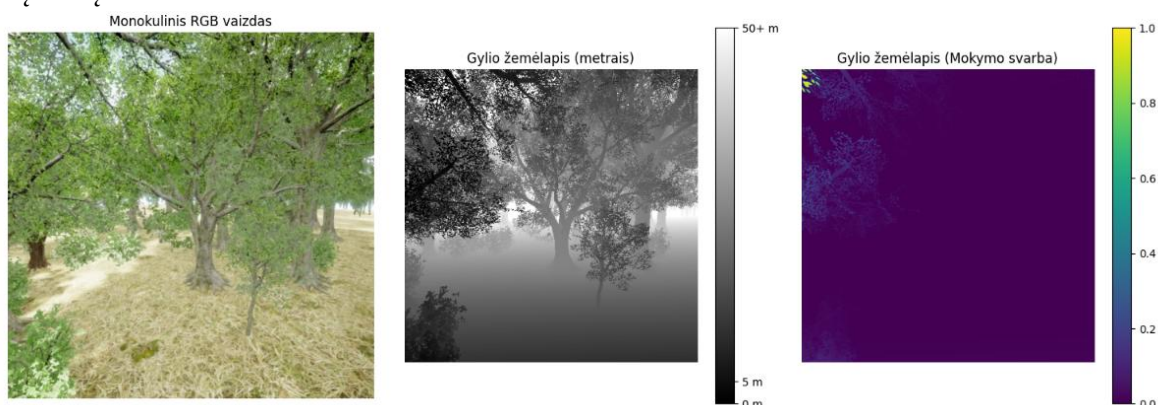
target – realaus vaizdo gylis

Skirtumas tarp jų tas, kad sigmoidinė funkcija teikia didelę reikšmę objektams, esantiems arti pasirinkto atstumo ir toliau greitai ignoruoja viską, kas yra toliau. Eksponentinė funkcija palaipsniui mažina svarbą didėjant atstumui, todėl tiek artimi, tiek truputi tolimi pikseliai prisideda prie skaičiavimų (žr. 29 pav.).

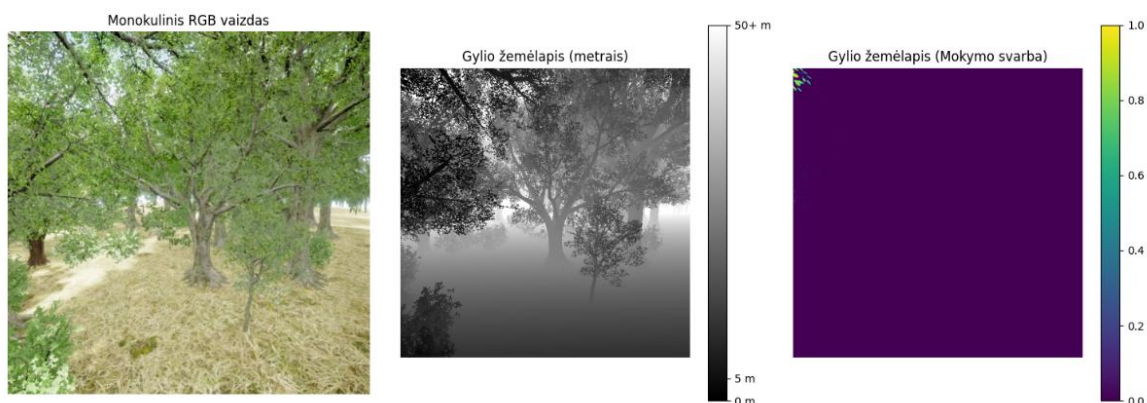


29 pav. Svorio vertė didinant atstumą su eksponentine ir sigmoidine funkcija, kai netolimo gylio parametras - 2, stiprumas - 2

Pagal šias formules yra nustatoma, kad artimesni gylio žemėlapių pikseliai turėtų didesnę įtaką mokymo rezultatui negu tolimi pikseliai, kurie yra nereikalingi šio tyrimo atžvilgiu. Šiuo atveju pagrindinis vienetas, kontroliuojantis, kaip stipriai yra atmetami tolimesni pikseliai yra stiprumas. Tai galima pamatyti iš vaizdų (žr. **30 pav.**, **31 pav.**) kuriuose matoma, kad su eksponentinę funkciją didinant stiprumo konstantą atsiranda statesnis eksponentinis kritimas ir mažiau tolimesnių pikselių įgauna svorį.

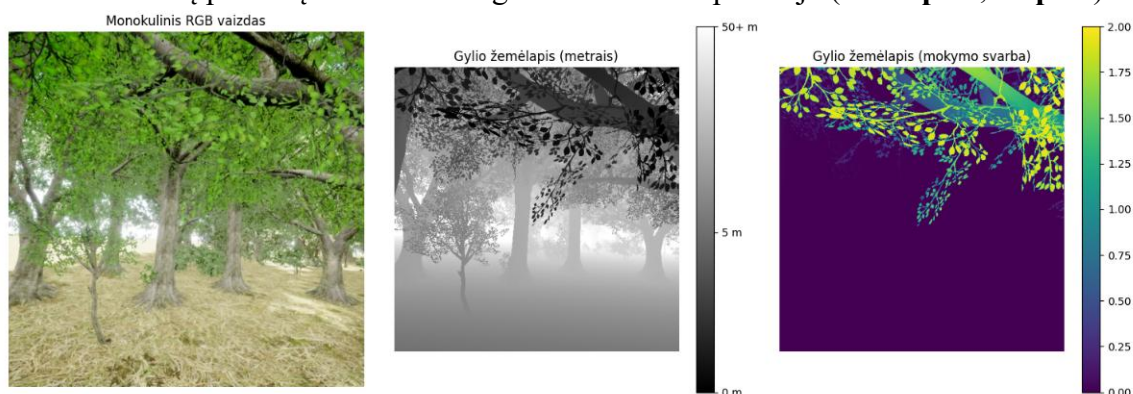


30 pav. Pavyzdys, rodantis tikrą gylį ir pikselių svorius su eksponentine funkcija, kai netolimo gylio parametro atstumas yra 2 metrai, o stiprumas - 2

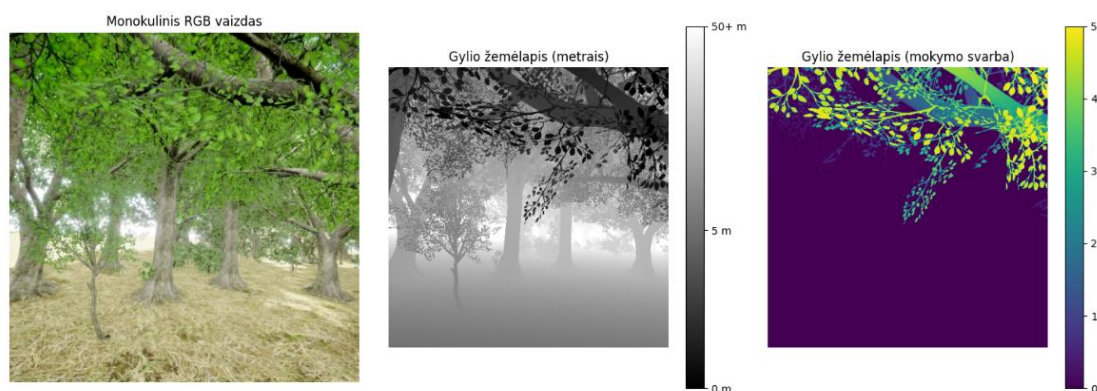


31 pav. Pavyzdys, rodantis tikrą gylį ir pikselių svorius su eksponentine funkcija, kai netolimo gylio parametro atstumas yra 2 metrai, o stiprumas - 5

Sigmoidinė funkcija veikia šiek tiek kitaip. Didinant stiprumo konstantą ji praktiškai nedidina ir nemažina tolimesnių pikselių kiekio. Tiesiog svorio reikšmė padidėja (žr. **32 pav.**, **33 pav.**).



32 pav. Pavyzdys, rodantis tikrą gylį ir pikselių svorius su sigmoidine funkcija, kai netolimo gylio parametro atstumas yra 2 metrai, o stiprumas - 2



33 pav. Pavyzdys, rodantis tikrą gylį ir pikselių svorius su sigmoidine funkcija, kai netolimo gylio parametro atstumas yra 2 metrai, o stiprumas - 5

Kadangi konkrečios geresnės svorio vertės nėra, yra patikrinama su kokia stiprumo konstanta modeliai gauna geriausius rezultatus, išbandant kelis kintamuosius – 2, 3, 4, 5.

Kaip ir minėta prieš tai, prie svorio funkcijų įvyksta pakeitimų ir su praradimo funkcijomis. Atsiranda tris naujos funkcijos:

- Subalansuota praradimo funkcija, kur stipriai atsižvelgia į kriterijus, svorių funkcijas, bet taip pat suteikia mažus svorius tolimiems atstumams, net jei svorių funkcija to neleidžia.

$$loss = loss_{SiLog} + 0.1 * \frac{1}{n} \sum_{i=1}^n w_i |y_i - x_i| + 0.05 * \frac{1}{n} \sum_{i=1}^n |y_i - x_i| \quad (8)$$

- Absoliuti artimo praradimo funkcija – kuri fokusuojasi tik į artimus atstumus pagal kriterijus ir svorių funkciją.

$$loss = loss_{SiLog} + 0.1 * \frac{1}{n} \sum_{i=1}^n w_i |y_i - x_i| \quad (9)$$

- Artimo normalizuota praradimo funkcija – panašiai kaip ir absoliuti artimo praradimo funkcija, fokusuojasi tik į artimus atstumus, tačiau dar labiau užtikrina, kad artimos paklaidos turėtų žymiai didesnę svorį, o tolimos būtų dar mažiau svarbios

$$loss = loss_{SiLog} + 0.1 * \frac{1}{n} \sum_{i=1}^n w_i \frac{|y_i - x_i|}{\max(x_i, \epsilon)} \quad (10)$$

Visos funkcijos čia turi šiuos žymėjimus:

$loss_{SiLog}$ – logaritminė praradimo funkcija

n – pikselių skaičius

x_i – faktinė vertė i pozicijoje

y_i – prognozuojama vertė i pozicijoje

w_i – svorio vertė i pozicijoje

Sudėjus visas svorių ir praradimo funkcijų kombinacijas su skirtingais svoriais gaunami 24 skirtingi modeliai, kurie yra apmokomi šiuo atveju su 500 vaizdais. Artimų atstumų metrikų rezultatai atvaizduojami lentelėje (žr. **11 lentelė**). Modeliai pateikiami formatu: stiprumo konstanta, praradimo funkcija, svorių funkcija.

11 lentelė. Koncentruotų modelių į artimus atstumus metrinių rezultatų palyginimai

Modelis	Jautrumas (%)	Persidengimo koeficientas (%)	F1 balas (%)	Vidutinio absoliutaus atstumo paklaida (m)
2_subalansuotas_eksponentinis	71.4	54.3	70.4	0.121
2_absoliutus_eksponentinis	75.4	55.3	71.2	0.117
2_normalizuotas_eksponentinis	75.3	55.2	71.2	0.114
2_subalansuotas_sigmoidinis	78.4	54.6	70.6	0.118
2_absoliutus_sigmoidinis	81.2	53.6	69.8	0.122
2_normalizuotas_sigmoidinis	81.4	53.6	69.8	0.124
3_subalansuotas_eksponentinis	76.3	54.0	70.1	0.110
3_absoliutus_eksponentinis	78.6	53.8	70.0	0.113
3_normalizuotas_eksponentinis	79.2	54.3	70.4	0.113

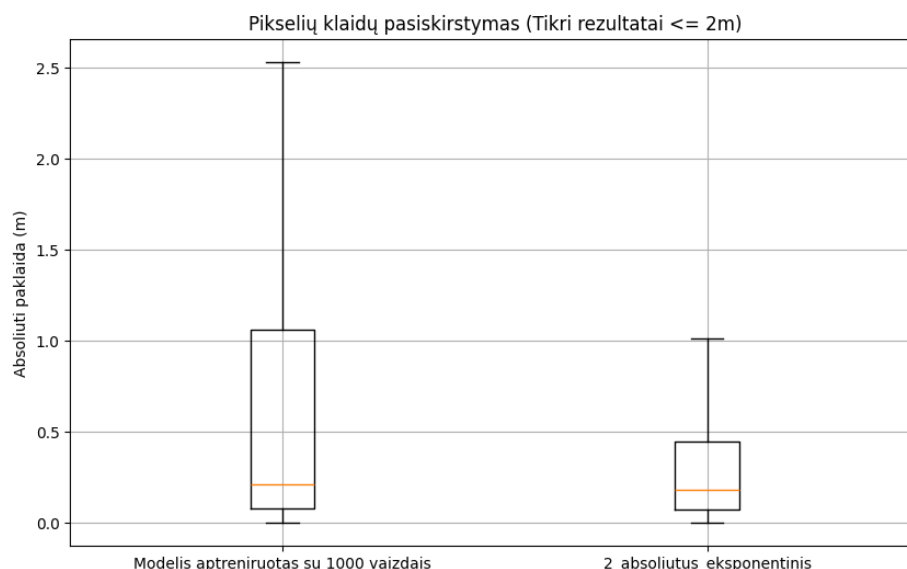
3_subalansuotas_sigmoidinis	78.2	54.7	70.7	0.117
3_absoliutus_sigmoidinis	81.0	53.5	69.7	0.126
3_normalizuotas_sigmoidinis	81.5	53.2	69.5	0.124
4_subalansuotas_eksponentinis	78.0	53.5	69.7	0.111
4_absoliutus_eksponentinis	81.8	52.3	68.7	0.108
4_normalizuotas_eksponentinis	83.1	52.4	68.8	0.107
4_subalansuotas_sigmoidinis	78.7	54.6	70.6	0.114
4_absoliutus_sigmoidinis	81.4	53.7	69.9	0.123
4_normalizuotas_sigmoidinis	82.7	53.1	69.4	0.118
5_subalansuotas_eksponentinis	80.0	52.9	69.2	0.109
5_absoliutus_eksponentinis	84.7	50.1	66.8	0.111
5_normalizuotas_eksponentinis	85.0	50.1	66.7	0.111
5_subalansuotas_sigmoidinis	78.5	54.4	70.5	0.117
5_absoliutus_sigmoidinis	81.4	53.6	69.8	0.124
5_normalizuotas_sigmoidinis	83.3	52.5	68.9	0.118

Pastebėta, kad modeliai apmokyti su 5 stiprumo konstanta, turi labai didelį jautrumą. Ypač „5_normalizuotas_eksponentinis“ modelis, kurio jautrumas siekia net 85%, kas reiškia kad jis nepraleidžia jokių artimų pikselių ir aptinka kliūtis. Be to, jo vidutinio absoliutaus atstumo paklaida taip pat yra žema, kas rodo, kad modelio prognozių atstumai lyginant su tikrais yra pakankamai artimos. Tačiau lyginant su kitais modeliais, jis turi žemą persidengimo koeficientą ir F1 balą, kas rodo reikšmingą klaidingai teigiamų prognozavimų padidėjimą. Tai galima pastebėti ir kituose modeliuose, kurie turi labai aukštą jautrumą, tačiau jų kitos metrikos sumažėja. Jie yra labai agresyvūs pikselių prognozėms, ir saugumo atžvilgiu yra neblogai – aptinka daug objektų, sumažina rizika atsitrekti į kliūtį, tačiau jie taip pat daro daug klaidų ir aptinka per daug pikselių, triukšmo, kurio tiesiog nėra. Tai gali trukdyti autonominio bepiločio orlaivio judėjimui miške. Priešingai, „2_absoliutus_eksponentinis“ modelis užtikrina geriausią persidengimo koeficientą ir F1 balą. Jo jautrumas, taip pat yra aukštas – 75.4%, kas užtikrina gerą saugumo ir aptikimo kokybės pusiausvyrą. Todėl jį verta naudoti tolimesniuose tyrimuose. Lyginant šio modelio metrikas su baziniu papildomai mokytu geriausiu modeliu „Modelis papildomai mokytas su 1000 vaizdais“, galime pamatyti išties žymiai geresnius rezultatus (žr. **12 lentelė**). Neskaitant vidutinio absoliutaus atstumo paklaidos, visos metrikos yra aukštesnės, ypač jautrumas. Kas parodo, kad patobulintas modelis artimiems atstumams ir apmokytas tik su 500 vaizdais pranoksta modelį, paprastai papildomai mokytą su dvigubai daugiau duomenų.

12 lentelė. Geriausių papildomai mokytų modelių metrinių rezultatų palyginimai

Modelis	Jautrumas (%)	Persidengimo koeficientas (%)	F1 balas (%)	Vidutinio absoliutaus atstumo paklaida (m)
2_absoliutus_eksponentinis	75.4	55.3	71.2	0.117
Modelis papildomai mokytas su 1000 vaizdais	64.6	54.2	70.3	0.103

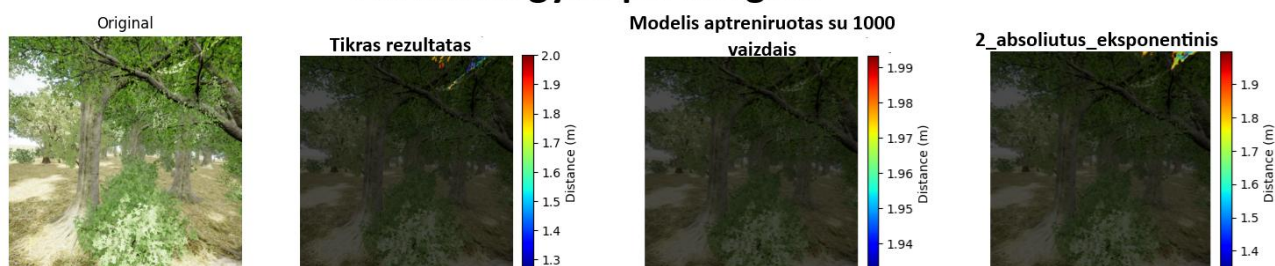
Be to, žiūrint į šių modelių pikselių klaidų pasiskirstymo dėžutės diagrama, taip pat galime pamatyti, kad paprastai apmokytas modelis su 1000 vaizdais pasižymi žymiai platesniu paklaidų pasiskirstymu, negu patobulintas modelis artimiems atstumams (žr. **34 pav.**). Taip pat, apmokyto modelio kraštutinės vertės viršija 2.5 metro, tuo tarpu artimų atstumo modelio kraštutinės vertės maždaug yra 1 m. Tai rodo didesnę patikimumą ir stabilumą vertinant artimų pikselių atstumus, kurie yra labai svarbūs patikimam kliūčių vengimui autonominių bepiločių orlaivių sistemoje.



34 pav. Geriausių papildomai mokyty modelių pikselių klaidų pasiskirstymo dėžutės diagrama

Iš vizualiai atvaizduotų gylio žemėlapių taip pat galima pamatyti, jog konkrečiai papildomai mokyti modeliai artimiems atstumams kai kurias atvejais aptinka kliūtis, kurių paprastai papildomai mokytas modelis nemato, ir rezultatai yra tikslesni lyginant su tikro gylio žemėlapiu (žr. 35 pav.).

Vaizdai su gylio perdangom



35 pav. Tikro, paprastai papildomai mokyto ir konkrečiai artimiems atstumams modelių vaizdų palyginimai su gylio perdangomis

3.5. Autonominių bepiločių orlaivių susidūrimų eksperimentai

Nors pikselių vertinimo metrikos suteikia svarbių įžvalgų apie gylio įvertinimo ir kliūčių aptikimo tikslumą, iš kurių galime pamatyti stiprius pagerėjimus su modeliais, kurie specifiskai pritaikyti artimiems atstumams, jie neatspindi sistemos veikimo realiu skrydžio metu. Dėl to yra atliekami papildomi eksperimentai pagrįsti susidūrimų procentu. Šiems eksperimentams yra naudojami visi trys „Depth Anything V2“ tipų modeliai – mažas, bazinis, didelis. Ir parenkami prieš tai du geriausiai papildomai mokyti metodai – vienas, kuris yra paprastai derintas su 1000 vaizdų, kitas – koncentruotas į artimus atstumus su 1000 vaizdų. Eksperimentai vykdomi „Unreal Engine“ sukurtoje miško aplinkoje. Iš viso vykdomi tie patys 1000 eksperimentų, kaip ir prieš tai, kai buvo nustatomas slenkstis, per kurį autonominis bepilotis orlaivis vengė susidūrimų. Autonominio bepiločio greitis – 3.5 m/s ir slenkstis – 2 metrai. Gauti rezultatai atvaizduojami lentelėje (žr. 13 lentelė).

13 lentelė. Susidūrimų dažnis su skirtingais modeliais

Modelio tipas	Susidūrimų dažnis (%) su paprastai derintu modeliu	Susidūrimų dažnis (%) su koncentruotu modeliu į artimus atstumus
Mažas	7	4.7
Bazinis	7.3	6.2
Didelis	9.6	7

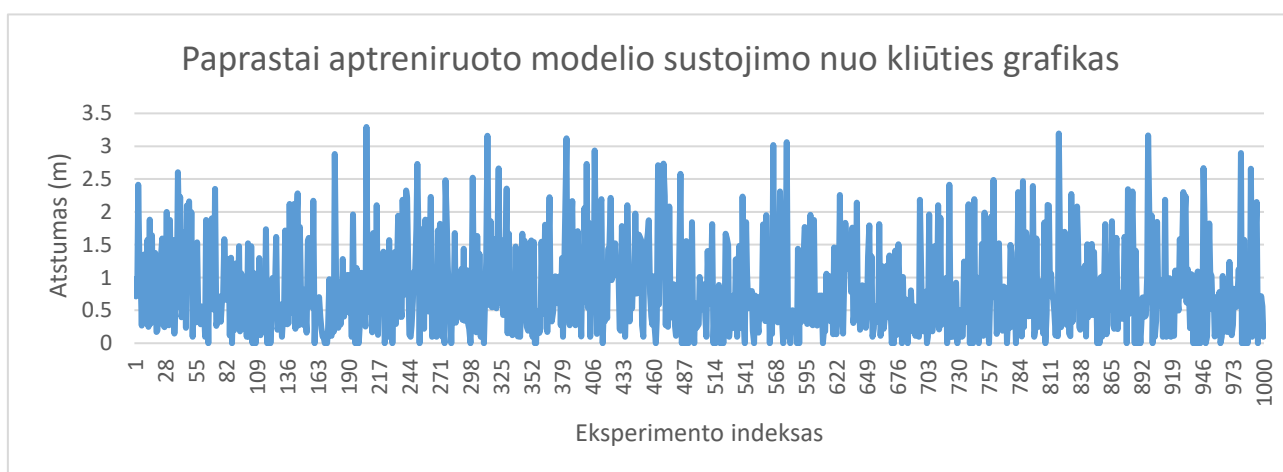
Kaip galima pamatyti, visi autonominio bepiločio orlaivio metodai, kurie buvo sukcentruoti į artimus atstumus naudojant eksponentinę svorių funkciją ir absoliučią praradimo funkciją, gavo itin geresnius rezultatus visais atvejais. Tai reiškia kad koncentruotas mokymas sumažina susidūrimų dažnį visuose modelių dydžiuose. Geriausių modelių, šiuo atveju mažo tipo, susidūrimų dažnis sumažėja nuo 7% iki 4.7%. Tai apytiksliai 1.49 karto pagerėjimas. Baziniai ir dideli modeliai taip pat pasižymi kliūčių vengimo pagerėjimu, tačiau našumas išlieka prastesnis nei mažo modelio. Tai prieštarauja prieš tai darytiems eksperimentams, kuriuose geriausius rezultatus gavo dideli modeliai. Tai galima paaiškinti tuo, kad vykdant eksperimentus realiu laiku, reikia atsižvelgti į resursų naudojimo ir apdorojimo laiko metrikas (žr. 14 lentelė).

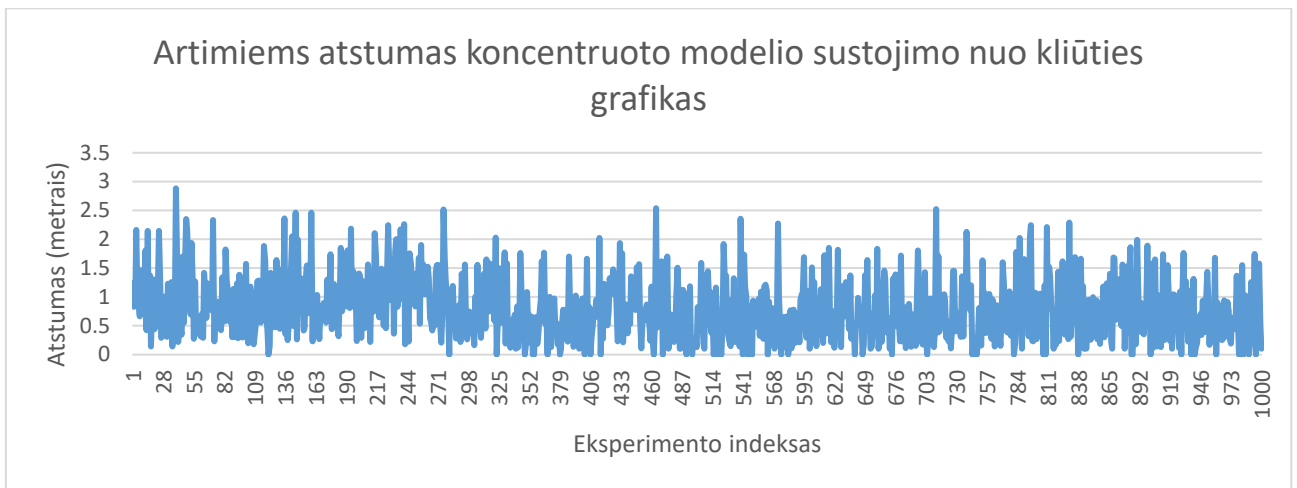
14 lentelė. Modelių vaizdo apdorojimo ir atminties naudojimo rezultatai

Modelio tipas	Vaizdo apdorojimo laikas (sekundėmis)	GPU vaizdo atminties naudojimas (GB)
Mažas	0.03	0.09
Bazinis	0.04	0.37
Didelis	0.11	1.25

Šie rezultatai būtent ir patvirtina efektyvumo realioju laiku prieštaravimą. Nors ir dideli modeliai apdoroja vaizdus geriau, jis beveik užtrunka 4 kartus ilgiau negu mažas modelis apdorojant vieną vaizdą. Dėl šio sulėtėjimo, automatiškai sulėtėja ir kliūčių vengimo reakcijos, kadangi autonominis bepilotis orlaivis pajuda keturis kartus toliau kai apdoroja vaizdus. Automatiškai, susidūrimų dažnis irgi padidėja, nes modelis nesugeba laiku ir greitai apdoroti vaizdų. Be to, didelis modelis naudoja žymiai daugiau resursų, kas gali lemti dar prastesnį ir lėtesnį vaizdų apdorojimą naudojant silpnus kompiuterinius resursus.

Mažiausio tipo modeliams yra sudaromi ir sustojimo grafikai, kurie parodo ties kokių atstumu nuo artimiausios kliūties bepilotis orlaivis sustoja (žr. 36 pav., 37 pav.).

**36 pav.** Mažo tipo paprastai papildomai mokyto modelio sustojimo nuo kliūties per 1000 eksperimentų grafikas



37 pav. Mažo tipo koncentruoto modelio į artimus atstumus sustojimo nuo kliūtis per 1000 eksperimentų grafikas

Iš grafikų galima pastebėti, paprastai apmokytas modelis kai kuriais atvejais sustoja prieš kliūtį esant virš trijų metrų. Tokių atvejų nėra su koncentruotu modeliu į artimus atstumus. Visuose eksperimentuose jis sustoja esant kliūčiai arčiau negu trys metrai. Tai tik dar patvirtina, praeitą teiginį apie paprastai derinto modelio platesnių paklaidų pasiskirstymą. Miško aplinkoje, esant daug įvairių smulkių objektų, gali atsirasti problemų, kur pavojaus pervertinimas ir per ankstyvas stabdymas sumažina navigacijos efektyvumą, švaisto laiką ir apriboja autonominio bepiločio orlaivio gebėjimą skristi siaurose erdvėse.

Visi šie susidūrimų eksperimentai įrodo, kad norint užtikrinti saugų ir efektyvų bepiločio orlaivio kliūčių vengimo veikimą reikia specifiskai atsižvelgti ir patobulinti modelį priderinti jį prie artimų atstumų, norint, kad atliktu efektyvius skaičiavimus. Be viso to, turi būti ir maža delta bei nedidelis kompiuterinių resursų naudojimas kad vaizdai būtų apdorojami greitai ir autonominio bepiločio reakcija į aplinką nesukeltų susidūrimo problemų.

Išvados

1. Monokulinio gylio įvertinimo metodų analizė parodė, kad šiuolaikiniai metodai gali teikti prasmingą informaciją autonominių bepiločių orlaivių kliūčių vengimui, be jokių papildomų jutiklių kaip LiDAR, GPS ar stereo kameros, kurios tiesiogiai apskaičiuoja atstumą nuo objektų. Tyrimuose taip pat paminėta, kad transformeriais pagrįstos sistemos gauna geresnius rezultatus, nei ankstesni konvoliuciniai metodai. Be viso to, iš naujesnių modelių specifiskai išsiskiria metodai, galintys įvertinti ne tik santykinį, bet ir metrinį atstumą, kuris yra labai reikšmingas objektų aptikimo ir vengimo uždaviniuose.
2. Skirtingų gylio įvertinimo modelių tyrimas parodė, kad artimo atstumo kliūčių suvokimas išlieka sudėtinga problema, ypač miško aplinkoje, kurioje pasitaiko smulkių, netaisyklingų kliūčių. Susidaro daug problemų, kada modelis neaptinka mažų objektų, ar tiesiog suklysta - tolimus objektus paverčia artimais ar atvirkščiai. Ypač naudojant santykinio gylio įvertinimo metodus, galima pastebėti, kaip nestandartinius objektus sujungia į vieną to paties lygio objektą. Naudojant metrinis gylio įvertinimo modelius gaunamos tinkamesnės ir tikslesnės prognozės, tačiau šios irgi klysta ir neaptinka smulkių objektų.
3. Parinkti „Depth Anything v2“ modelio variantai parodė, kad jie gali įvertinti netoliese esančių kliūčių atstumus, naudojant tik RGB vaizdus, tačiau norint išgauti tikslesnių rezultatų, juos būtina apmokyti su papildomais kameros duomenimis, kitaip metriniai atstumų rezultatai gali visiškai neatitikti tikrų. Yra analizuojami kelių dydžių modeliai ir iš teorinės, ne realaus laiko, eksperimentų sudėtingesni modeliai gauna tikslesnes metrikas.
4. Yra sukuriami koncentruoti modeliai į artimus atstumus. Jie naudoja kelias svorių ir praradimo funkcijas su skirtingais stiprumo rodikliais. Dauguma modifikuotų modelių parodė geresnes atstumų metrikas, mažesnę paklaidų pasiskirstymą ir geresnę nuoseklumą vertinant atstumus tarp netoliese esančių kliūčių. Labiausiai išsiskyrė metodas, papildomai mokytas su eksponentine svorių funkcija ir su absoliučia artimo praradimo funkcija naudojant dviejų kintamojo stiprumą, kuris pasižymi itin aukštomis persidengimo koeficiento ir F1 balo metrikomis, turint gerus ir visus kitus rezultatus. Šio modelio trys variantai: mažas, bazinis, didelis yra apmokomi su 1000 vaizdais ir lyginami su paprastai apmokytu modeliu tolimesniuose realaus laiko eksperimentuose.
5. Įvairios metrikos ir realaus laiko eksperimentai patvirtino, kad bepiločių orlaivių kliūčių vengimas realiuoju laiku priklauso ne tik nuo prognozavimo tikslumo, bet ir nuo apdorojimo greičio be skaičiavimo efektyvumo. Didesniems modeliams reikia daugiau apdorojimo laiko ir kompiuterinių resursų, o tai neigiamai paveikia objektų vengimą realiuoju laiku ir padidina susidūrimų dažnį. Mažesni modeliai galintys greičiau reaguoti į aplinką išgauna geresnius rezultatus ir labiau tinka autonominėms bepiločių orlaivių sistemoms realiuoju laiku dėl savo efektyvesnės architektūros.

Dirbtinio intelekto įrankių naudojimas

Eksperimentiniais tyrimo tikslams pasiekti naudojami „Depth Anything V2“ dirbtinio intelekto modeliai. Programinio kodo formatavimui ir sudėtingesnių frazių vertimui naudojamas „ChatGPT“ dirbtinio intelekto įrankis.

Literatūros sąrašas

1. KATKURI, A.V.R., et al. Autonomous UAV Navigation using Deep Learning-Based Computer Vision Frameworks: A Systematic Literature Review. *Array (New York)*, Sep, 2024, vol. 23. pp. 100361. Available from: <https://dx.doi.org/10.1016/j.array.2024.100361> CrossRef. ISSN 2590-0056. DOI 10.1016/j.array.2024.100361.
2. ALADEM, M. and RAWASHDEH, S.A. Lightweight Visual Odometry for Autonomous Mobile Robots. *Sensors (Basel, Switzerland)*, Aug 28, 2018, vol. 18, no. 9. pp. 2837. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/30154311> PubMed. ISSN 1424-8220. DOI 10.3390/s18092837.
3. Sen Wang, CLARK, R., Hongkai Wen and TRIGONI, N. *DeepVO: Towards End-to-End Visual Odometry with Deep Recurrent Convolutional Neural Networks*. IEEE, May 2017 Available from: <https://ieeexplore.ieee.org/document/7989236> DOI 10.1109/ICRA.2017.7989236.
4. HRISTOVA, H., ABEGG, M., FISCHER, C. and REHUSH, N. Monocular Depth Estimation in Forest Environments. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.*, May 30, 2022, vol. XLIII-B2-2022. pp. 1017–1023. Available from: <https://isprs-archives.copernicus.org/articles/XLIII-B2-2022/1017/2022/> CrossRef. ISSN 2194-9034. DOI 10.5194/isprs-archives-XLIII-B2-2022-1017-2022.
5. GODARD, C., AODHA, O.M., FIRMAN, M. and BROSTOW, G. *Digging into Self-Supervised Monocular Depth Estimation*. Piscataway: IEEE, Oct 2019 Available from: <https://ieeexplore.ieee.org/document/9009796> DOI 10.1109/ICCV.2019.00393.
6. Fangchang Ma and KARAMAN, S. *Sparse-to-Dense: Depth Prediction from Sparse Depth Samples and a Single Image*. IEEE, May 2018 Available from: <https://ieeexplore.ieee.org/document/8460184> CiNII Research of NDL. DOI 10.1109/ICRA.2018.8460184.
7. ZHANG, C., WENG, X., CAO, Y. and DING, M. Monocular Absolute Depth Estimation from Motion for Small Unmanned Aerial Vehicles by Geometry-Based Scale Recovery. *Sensors (Basel, Switzerland)*, Jul 13, 2024, vol. 24, no. 14. pp. 4541. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/39065938> PubMed. ISSN 1424-8220. DOI 10.3390/s24144541.
8. HATCH, K., MERN, J. and KOCHENDERFER, M. *Obstacle Avoidance using a Monocular Camera*. Ithaca: Cornell University Library, arXiv.org. Jan 25, 2021 Available from: <https://www.proquest.com/docview/2467250169> Publicly Available Content Database. DOI 10.48550/arxiv.2012.01608.
9. CASAO, S., et al. *A Framework for Fast Prototyping of Photo-Realistic Environments with Multiple Pedestrians*. Piscataway: IEEE, May 29, 2023 Available from: <https://ieeexplore.ieee.org/document/10160586> DOI 10.1109/ICRA48891.2023.10160586.
10. SHAH, S., DEY, D., LOVETT, C. and KAPOOR, A. AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles, May 15, 2017. Available from: <https://arxiv.org/abs/1705.05065> DOI 10.48550/arxiv.1705.05065.
11. AL-KAFF, A., et al. Obstacle Detection and Avoidance System Based on Monocular Camera and Size Expansion Algorithm for UAVs. *Sensors (Basel, Switzerland)*, May 7, 2017, vol. 17, no. 5.

pp. 1061. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28481277> CrossRef. ISSN 1424-8220. DOI 10.3390/s17051061.

12. RANFTL, R., BOCHKOVSKIY, A. and KOLTUN, V. *Vision Transformers for Dense Prediction*. Piscataway: IEEE, Oct 2021 Available from: <https://ieeexplore.ieee.org/document/9711226> DOI 10.1109/ICCV48922.2021.01196.

13. RANFTL, R., et al. Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Mar 1, 2022, vol. 44, no. 3. pp. 1623–1637. Available from: <https://ieeexplore.ieee.org/document/9178977> CiNII Research of NDL. ISSN 0162-8828. DOI 10.1109/TPAMI.2020.3019967.

14. BHAT, S.F., et al. ZoeDepth: Zero-Shot Transfer by Combining Relative and Metric Depth, Feb 23, 2023. Available from: <https://arxiv.org/abs/2302.12288> DOI 10.48550/arxiv.2302.12288.

15. YANG, L., et al. *Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data*. IEEE, Jun 16, 2024 Available from: <https://ieeexplore.ieee.org/document/10657693> CiNII Research of NDL. DOI 10.1109/CVPR52733.2024.00987.

16. YANG, L., et al. Depth Anything V2, Jun 13, 2024. Available from: <https://arxiv.org/abs/2406.09414> DOI 10.48550/arxiv.2406.09414.