



Kauno technologijos universitetas

Elektros ir elektronikos fakultetas

**Erdvinių objektų aptikimo ir jų padėties nustatymo naudojant
RGB-D kamerą metodų palyginimas**

Baigiamasis magistro projektas

Tautvydas Žulys

Projekto autorius

lekt. Gintautas Narvydas

Vadovas

Kaunas, 2022



Kauno technologijos universitetas

Elektros ir elektronikos fakultetas

Erdvinių objektų aptikimo ir jų padėties nustatymo naudojant RGB-D kamerą metodų palyginimas

Baigiamasis magistro projektas

Valdymo technologijos (6211EX014)

Tautvydas Žulys

Projekto autorius

lekt. Gintautas Narvydas

Vadovas

doc. dr. Virginijus Baranauskas

Recenzentas

Kaunas, 2022



Kauno technologijos universitetas

Elektros ir elektronikos fakultetas

Tautvydas Žulys

Erdvinių objektų aptikimo ir jų padėties nustatymo naudojant RGB-D kamerą metodų palyginimas

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.

Tautvydas Žulys

Patvirtinta elektroniniu būdu

Žulys, Tautvydas. Erdvinių objektų aptikimo ir jų padėties nustatymo naudojant RGB-D kamerą metodų palyginimas. Magistro baigiamasis projektas / vadovas lekt. Gintautas Narvydas; Kauno technologijos universitetas, Elektros ir Elektronikos fakultetas.

Studijų kryptis ir sritis: Elektronikos inžinerija, inžinerijos mokslai.

Reikšminiai žodžiai: objektų atpažinimas, gylio kamera, aptikimo metodų palyginimas.

Kaunas, 2022. 52 p.

Santrauka

Šiame darbe testuojami keturi skirtingi objektų aptikimo metodai atliekant bandymus su erdviniais objektais. Parinkti objektai skiriasi forma, dydžiu, spalva. Pradžioje sudaroma mokymo vaizdų imtis, augmentacijos pagalba ji padidinama. Augmentacijos metu iš esamų vaizdų sugeneruojami nauji vaizdai pridodant filtrus, triukšmus, pasukimus. Vertinti metodų efektyvumui atskirai kiekvienam objektui sudaroma 100 vaizdų imtis su natūralia šviesa, bei 100 vaizdų imtis su dirbtine šviesa ir šalutiniais objektais. Aptikimo metodams įvertinti ir palyginti pasirenkami šie kriterijai: vidutinis aptikimo tikslumas, vidutinė kvadratinė paklaida, aptiktų objektų skaičius procentais, vieno kadro apdorojimo laikas. Kriterijų skaitinėms vertėms gauti naudojamos aptikimo metodų sugeneruojamos koordinatės, nurodančios objektų pozicijas paveiksle. Jos lyginamos su testavimo vaizduose esančių objektų pozicijomis, kurios naudojamos kaip etaloninės vertės. Objektų koordinatėms pikseliais kadre konvertuoti į realias pritaikyti du būdai. Vienas būdas paremtas gylio kamera, kitas RGB kameros vaizdu bei proporcijų skaičiavimu. Abiejų metodų paklaidų vertės yra artimos ir mažesnės už 10 milimetrų. Informatyvesnis ir paprastesnis taikyti yra pozicijų konvertavimo būdas paremtas gylio kameros duomenimis. Ištestavus visus aptikimo metodus įvertinami suminiai rezultatai, apžvelgiami grafikai rodantys aptikimo kriterijų progresą testavimo vaizduose.

Žulys, Tautvydas. Comparison of spatial objects detection and positioning methods using RGB-D camera. Master's project / supervisor lekt. Gintautas Narvydas; Faculty of Electrical and Electronics Engineering, Kaunas University of Technology.

Study field and area: Electronics engineering, engineering sciences.

Keywords: object detection, depth camera, comparison of detection methods.

Kaunas, 2022. 52 p.

Summary

In this work is tested and compared four different object detection methods. Objects selected for testing has different shape, color, size. At the beginning training images is collected. To increase size of training data, augmentation is used. One hundred images samples are collected for each test object. Sample images can be split in two categories one with natural light and no additional objects. Second with artificial light and a few additional objects in testing background. Comparison criteria to evaluate methods are average precision, mean square error, average detection time, percent of successfully recognized objects. Criteria values calculated from the object detection methods results in pixel coordinates. All these coordinates are compared with ground truth object positions in testing images. Two different methods used to convert pixels coordinates to object positions in real world measurements. One method is based on depth camera data, another on RGB image, markers, and proportions. Both methods have similar error values. Calculation based on depth camera is simpler to use and gives some additional info about object. All results from object detection methods are summarized and compared at the last chapter.

Turinys

| | |
|--|-----------|
| Lentelių sąrašas | 8 |
| Paveikslų sąrašas | 9 |
| Santrumpų ir terminų sąrašas | 10 |
| Įvadas..... | 11 |
| 1. RGB-D kameros duomenų apdorojimo bei objektų aptikimo metodų analizė | 12 |
| 1.1. Gylio kameros | 12 |
| 1.2. ToF kameros veikimas | 12 |
| 1.3. ToF ir RGB-D kameros | 13 |
| 1.4. Gylio duomenų filtravimo metodai | 14 |
| 1.5. Aptikimo metodų tipai ir veikimas | 17 |
| 1.6. HOG su SVM klasifikatoriumi..... | 17 |
| 1.7. Faster R-CNN..... | 18 |
| 1.7.1. InceptionV2 tinklo savybės | 20 |
| 1.8. YOLO (You Look Only Once)..... | 21 |
| 1.9. SSD (Single Shot Detector)..... | 21 |
| 1.9.1. MobileNetV1 tinklas | 22 |
| 1.10. Aptikimo rezultatų kokybinis įvertinimas | 23 |
| 2. Tyrimo įrangą ir pasiruošimas..... | 25 |
| 2.1. Kameros specifikacijos..... | 25 |
| 2.2. Įrangos specifikacijos | 26 |
| 2.3. Gylio ir RGB kamerų testavimas | 26 |
| 2.4. Ieškomi objektai | 27 |
| 2.5. Duomenų augmentacija | 28 |
| 2.6. Vertinimo kriterijų apskaičiavimas | 28 |
| 3. Algoritmų realizavimas ir eksperimentai..... | 29 |
| 3.1. HOG su SVM klasifikatoriumi..... | 29 |
| 3.1.1. Metodo aprašymas..... | 29 |
| 3.1.2. Metodo parengimas bei apmokymas | 31 |
| 3.1.3. Metodo testavimas su vaizdų rinkiniais | 31 |
| 3.2. Faster R-CNN..... | 33 |
| 3.2.1. Metodo parengimas bei apmokymas | 33 |
| 3.2.2. Metodo bandymai su iš anksto neparengtu modeliu | 34 |
| 3.2.3. Metodo bandymai su parengtu modeliu | 34 |
| 3.2.4. Metodo testavimas su vaizdų rinkiniais | 36 |
| 3.3. YOLOv5..... | 37 |
| 3.3.1. Metodo parengimas bei apmokymas | 37 |
| 3.3.2. Metodo testavimo su vaizdų rinkiniais..... | 38 |
| 3.4. SSD..... | 39 |
| 3.4.1. Metodo parengimas bei apmokymas | 39 |
| 3.4.2. Metodo testavimas su vaizdų rinkiniais | 40 |
| 4. Objekto padėties radimas naudojant Kinect kamerą | 42 |
| 4.1. Objektų pozicijos radimas | 42 |

| | |
|--|-----------|
| 4.2. Objektų koordinacių nustatymas naudojant gylio duomenimis | 43 |
| 4.3. Objektų koordinacių nustatymas naudojantis RGB vaizdu..... | 44 |
| 4.4. Objektų padėties nustatymo metodų palyginimas..... | 46 |
| 5. Aptikimo metodų rezultatų palyginimas..... | 47 |
| Rezultatai ir išvados | 50 |
| Literatūros sąrašas | 51 |
| Priedai..... | 53 |
| 1 Tyrimo metu gauti vaizdai..... | 53 |
| 2 Pirminių bandymų rezultatų lentelė..... | 53 |
| 3 Suminiai grafikai | 54 |

Lentelių sąrašas

| | |
|--|----|
| 2.1 lentelė. Kinect v2 kameros techninės specifikacijos..... | 25 |
| 2.2 lentelė. Naudojamos įrangos specifikacijos | 26 |
| 2.3 lentelė. Ieškomi objektai ir jų dimensijos | 27 |
| 3.1 lentelė. HOG ir SVM metodo apmokymo imtis bei rezultatai..... | 31 |
| 3.2 lentelė. HOG ir SVM testavimo rezultatai (natūrali šviesa ir tiriamieji objektai) | 31 |
| 3.3 lentelė. HOG ir SVM testavimo rezultatai (dirbtinė šviesa ir šalutiniai objektai) | 32 |
| 3.4 lentelė. Metodo apmokymo rezultatai | 35 |
| 3.5 lentelė. Testavimo rezultatai (natūrali šviesa ir tiriamieji objektai)..... | 36 |
| 3.6 lentelė. Testavimo rezultatai (dirbtinė šviesa ir šalutiniai objektai) | 36 |
| 3.7 lentelė. Mokymo rezultatai..... | 38 |
| 3.8 lentelė. Testavimo rezultatai (natūrali šviesa ir tiriamieji objektai)..... | 38 |
| 3.9 lentelė. Testavimo rezultatai (dirbtinė šviesa ir šalutiniai objektai) | 38 |
| 3.10 lentelė. Metodo rengimo rezultatai | 40 |
| 3.11 lentelė. Testavimo rezultatai (natūrali šviesa ir tiriamieji objektai)..... | 40 |
| 3.12 lentelė. Testavimo rezultatai (dirbtinė šviesa ir šalutiniai objektai) | 41 |
| 4.1 lentelė. Išmatuotų ir apskaičiuotų pozicijų vertės | 44 |
| 4.2 lentelė. Išmatuotų ir apskaičiuotų pozicijų vertės | 46 |
| 5.1 lentelė. Suminė lentelė (natūrali šviesa ir tiriamieji objektai)..... | 47 |
| 5.2 lentelė. Suminė lentelė (dirbtinė šviesa ir šalutiniai objektai) | 47 |

Paveikslų sąrašas

| | |
|--|----|
| 1.1 pav. Atstumo matavimas naudojant fazės pokyčio įvertinimą [1]..... | 12 |
| 1.2 pav. Kairėje amplitudinis iškraipymas, dešinėje skirtingų atspindžių iškraipyti duomenys [1]. | 13 |
| 1.3 pav. Atlikti kamerų bandymai lyginant RGB-D ir ToF kameras [2]..... | 14 |
| 1.4 pav. Gylio žemėlapių apdorojimo rezultatai pasitelkus skirtingus filtrus [3]..... | 15 |
| 1.5 pav. Tyrimas įvertinant apšvietumo įtaką kairėje ir spalvų įtaką matavimams dešinėje [4]..... | 16 |
| 1.6 pav. Bandymai atlikti su skirtingomis medžagomis [4]..... | 16 |
| 1.7 pav. Analizuojamas vaizdas sudalintas 8x8 matricomis kairėje ir skaitinės matricos vertės dešinėje [6]..... | 17 |
| 1.8 pav. a) tiriamas vaizdas, b) gautas vektorių žemėlapis, c) pikseliai suskirstyti į vektorių histogramas pagal kampus [5]..... | 18 |
| 1.9 pav. R-CNN algoritmo veikimo etapai, atliekant objekto paiešką [7]..... | 19 |
| 1.10 pav. Konvoliucinių sluoksnių pasiskirstymas tarp modulių [7]..... | 20 |
| 1.11 pav. InceptionV2 tinklo struktūra [9]..... | 20 |
| 1.12 pav. YOLO atliekama vaizdo analizė [11]..... | 21 |
| 1.13 pav. SSD tinklo schema [12]..... | 22 |
| 1.14 pav. Sluoksniuose naudojamų konvoliucinių metodų palyginimas[13]..... | 23 |
| 1.15 pav. Bendras sankirtos plotas (<i>IOU</i>) [15]..... | 24 |
| 2.1 pav. Kinect V2 kamera..... | 25 |
| 2.2 pav. Kameros pateikiamas gylio vaizdas dešinėje ir standartinis vaizdas kairėje..... | 25 |
| 2.3 pav. RGB ir IR kameros duomenų atitikimas..... | 26 |
| 2.4 pav. Tyrimo metu gauta paklaidos priklausomybė nuo atstumo..... | 27 |
| 2.5 pav. Augmentacijos taikymas duomenų rinkiniui praplėsti..... | 28 |
| 3.1 pav. Objekto paieškos algoritmas naudojant HOG ir SVM..... | 30 |
| 3.2 pav. Natūralaus apšvietimo įtaka aptikimo metodui..... | 32 |
| 3.3 pav. Butelio paieškos rezultatai naudojantis HOG ir SVM metodu..... | 32 |
| 3.4 pav. Faster R-CNN, YOLO, SSD metodų supaprastintas vaizdo analizės algoritmas..... | 33 |
| 3.5 pav. Faster R-CNN rezultatai su neaprengtu modeliu..... | 34 |
| 3.6 pav. Rezultatai naudojant parengta Faster R-CNN InceptionV2 modelį..... | 34 |
| 3.7 pav. Modelio mokymosi regresijos pavyzdžiai, plastiko kairėje, bendras visų objektų dešinėje..... | 35 |
| 3.8 pav. Klaidingai atpažįstami objektai..... | 35 |
| 3.9 pav. Faster R-CNN rezultatai kairėje natūrali šviesa, dešinėje dirbtinė ir šalutiniai objektai.... | 37 |
| 3.10 pav. Faster R-CNN rezultatai kairėje natūrali šviesa, dešinėje dirbtinė šviesa ir šalutiniai objektai..... | 37 |
| 3.11 pav. YOLO rezultatai kairėje natūrali šviesa, dešinėje dirbtinė ir šalutiniai objektai..... | 39 |
| 3.12 pav. YOLO rezultatai kairėje natūrali šviesa, dešinėje dirbtinė ir šalutiniai objektai..... | 39 |
| 3.13 pav. Paklaidos kitimas augant mokymo iteracijoms: kairėje plastiko ritė, dešinėje butelis..... | 40 |
| 3.14 pav. Plastiko ričių aptikimo rezultatai kairėje natūrali šviesa, dešinėje dirbtinė šviesa ir šalutiniai objektai..... | 41 |
| 4.1 pav. Bandymų sistema..... | 42 |
| 4.2 pav. Objekto pozicijos patikslinimas naudojantis gylio kameros duomenimis..... | 42 |
| 4.3 pav. Gylio kameros matomas vaizdas su atskaitos bei ieškomais taškais..... | 43 |
| 4.4 pav. Skaičiavimui naudojamų taškų bei ieškomų pozicijų schema..... | 44 |
| 5.1 pav. Varžtų aptikimo bendras grafikas (natūrali šviesa ir ieškomi objektai)..... | 48 |
| 5.2 pav. Plastiko aptikimo bendras grafikas (dirbtinė šviesa ir šalutiniai objektai)..... | 49 |

Santrumpų ir terminų sąrašas

Santrumpos:

CNN (angl. *Convolutional neural network*) – konvoliucinis neuroninis tinklas.

YOLO (angl. *You look only once*) – išvertus pažodžiui „tu pažiūrėsi tik kartą“ tai aptikimo metodo pavadinimas reiškiantis, kad atliekant paieška vaizdas skenuojamas tik kartą.

SSD (angl. *Single shot detector*) – išvertus pažodžiui „vieno šūvio detektorius“ tai aptikimo metodas naudojantis vienkartinį kadro skenavimą.

R-CNN (angl. *Region based convolutional neural networks*) – regionais paremtas konvoliucinis neuroninis tinklas.

CUDA (angl. *Compute unified device architecture*) – programinė įranga leidžianti naudoti grafinius procesorius duomenų apdorojimui.

RGB (angl. *Red Green Blue*) – spalvinis kodas sudarytas iš spalvų einančių tokia seka raudona, žalia, mėlyna.

RGB-D (angl. *RGB-Depth*) – vaizdinė informacija su gylio duomenimis.

IR (angl. *Infrared*) – infraraudonieji spinduliai.

CMOS (angl. *complementary metal-oxide-semiconductor*) – puslaidininkinis metalo oksido komponentas.

TOF (angl. *Time of flight*) – gylio kameros, naudojančios šviesos spindulio sklidimo laiką atstumui iki objekto nustatyti.

MSE (angl. *Mean square error*) – vidutinė kvadratinė paklaida.

AP (angl. *Average precision*) – vidutinis tikslumas.

MAP (angl. *Mean average precision*) – vidutinio tikslumo vidurkis, visiems tirtiems objektams bendrai.

FC (angl. *Fully connected layer*) – pilnai sujungtas sluoksnis.

CDT (angl. *common distance transform*) – bendroji gylio transformacija.

JBF (angl. *Joint Bilateral Filter*) – jungtinis dvišalis filtras.

PF-SVM (angl. *Particle Filter-Support Vector Machine*) – dalelių filtras su atraminių vektorių mašina.

ROI (angl. *Region of interest*) – telkimo sluoksnio tipas, kuris atlieka maksimumų apjungimą.

RPN (angl. *Region proposal network*) – tinklas aptinkantis objektų pozicijas bei nurodantys objekto klasės tikėtinumą.

COCO (angl. *Common objects in context*) – įvairių bendrinių duomenų rinkinių talpykla.

NMS (angl. *Non maximum suppression*) – kandidatų filtravimo būdas.

IOU (angl. *Intersection over union*) – bendras sankirtos plotas.

HOG (angl. *Histogram of oriented gradients*) – orientuotų gradientų histograma.

SVM (angl. *Support vector machine*) – atraminių vektorių klasifikatorius.

Terminai:

Konvoliucija – tai procesas, kurio metu atliekant operacijas su pikseliais išgaunami svoriniai koeficientai. Šis būdas turi panašumą su matematikoje naudojama konvoliucija.

Įvadas

Gylio kameros vis plačiau taikomos praktikoje ne tik pramogų sferoje ar autonominiame vairavime bet ir sprendžiant gamybinius uždavinius. Šios kameros naudojamos objektų stebėjimui ir sekimui kontrolės sistemose. Patobulėjusios kamerų gamybos technologijos leido išgauti tikslesnius duomenis, todėl išaugo praktinės pritaikymo galimybės. Kameros duomenims apdoroti ir skirtingiems uždaviniams spręsti naudojami įvairūs metodai ir algoritmai. Vieni iš jų reikalauja didelio duomenų apdorojimo našumo, kiti remiasi parengtais ir apmokytais modeliais. Kai kurie naudojami metodai tobulėja atlikdami užduotį ir mokydami. Siekiant palyginti skirtingus metodus atliekančius tą pačią užduotį ir rasti pranašumas bei trūkumus pasirenkamas objekto identifikavimo uždavinys. Šis uždavinys dažnai sprendžiamas gamyboje pavyzdžiui rūšiuojant ant konvejerio važiuojančius objektus ar bandant manipulatoriumi paimti konkretų objektą. Pasitelkiant gylio duomenis galima nustatyti objekto poziciją, o esant reikalui rasti judėjimo kryptį bei greitį. Tokios kameros naudojamos sudarant žemėlapius ir sprendžiant lokalizacijos uždavinius. Visos šios taikymo galimybės paremtos identifikavimo uždavinio sprendimu. Šiame darbe siekiama palyginti aptikimo metodus ir rasti minimalias parengimo bei realizavimo sąlygas reikalingas identifikuoti objektą. Dažnu atveju vaizdais grįsta logika yra sudėtingai realizuojama ir reikalaujanti laiko bei resursų atlikti apdorojimo operaciją. Siekiant spėti apdoroti vaizdus realiu laiku išauga duomenų apdorojimo įrangos kaštai. Didesnio populiarumo susilaukia nauji modeliai ir metodai, kurie geba veikti naudodami tik sistemos procesorių. Tyrimo metu pasirenkami lyginti metodai turėtų būti grįsti skirtinga logika bei metodika. Taip galima išvelgti pagrindinius privalumus ir trūkumus, bei įvertinti galimus praktinio taikymo atvejus. Objektų aptikimo metodai realizuojami ir išbandomi realiomis sąlygomis palyginant jų veiksmingumą. Tyrimo metu naudojama Microsoft Kinect v2 kamera. Ši kamera turi išstobulintą ir aprašyta bazinę biblioteką darbui su transliuojamais duomenimis. Kadangi kamera buvo skirta veikti, kartu su žaidimų konsolėmis, todėl duomenims atliekamas pirminis apdorojimas. Tyrimo metu duomenims apdoroti naudojamas kompiuteris su Microsoft Windows operacine sistema, nes Kinect v2 kamera naudoja *.NET* struktūrą. Tiriamiems atpažinimo metodams keliamas vienas tikslas aptikti objektą ir nustatyti jo poziciją kadre. Gavus objekto koordinates bei tipą galima atlikti tolimesnius veiksmus. Esant reikalui paskaičiuoti jo poziciją, posūkio kampą ar judėjimo trajektoriją. Apdoroti duomenys gali būti pateikiami tolimesnei įrangai, kuri atlieka operacija su matomu objektu. Tokio veiksmo pavyzdys būtų gaminių rūšiavimas ar nukrovimas. Visiems tiriamiems metodams taikomos vienodos bandymų sąlygos. Rezultatams vertinti sukuriamas testavimo programa, kuri įvertina aptikimo stabilumą bei palygina objektų pozicijas paveiksle su jau nustatytomis pozicijomis.

Tikslas: Įgyvendinti ir iširti keturis objektų aptikimo metodus bei palyginti jų tikslumą ir objektų paieškos trukmę esant vienodoms sąlygoms.

Uždaviniai:

1. Realizuoti erdvinių objektų aptikimo metodus.
2. Surinkti apmokymo ir testavimo vaizdų imtis.
3. Atlikti metodų bandymus su erdviniais objektais.
4. Įgyvendinti bei išbandyti pozicijos nustatymo metodus.
5. Pasirinkti kriterijus ir palyginti skirtingų aptikimo metodų rezultatus.

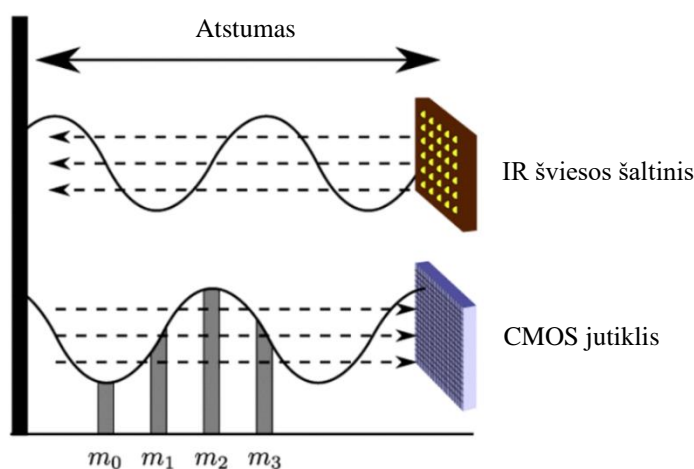
1. RGB-D kameros duomenų apdorojimo bei objektų aptikimo metodų analizė

1.1. Gylio kameros

Kameros pramonėje naudojamos įvairioms gamybinėms problemoms spręsti, jų taikymą skatina neuroninių tinklų tobulėjimas bei didėjantis vaizdų apdorojimo sistemų tikslumas. RGB-D ir ToF kameros atsirado sąlyginai neseniai, bet didėjantis jų patikimumas bei matavimo tikslumas atveria naujas taikymo galimybes. Kameros duomenų tikslumas dažnai būna priklausomas nuo aplinkos sąlygų tai viena iš priežasčių, kodėl jos vis dar retai sutinkamos pramonėje. Tačiau įveikus šią kliūtį galima atlikti daugelį sudėtingų užduočių. Kameros suteikia trimatę erdvinę objekto matricą su daugeliu matavimo taškų, tokia informacija naudojant su RGB vaizdu galima spręsti erdvinius uždavinius. Siekiant suprasti kameros taikymo galimybes bei ribas analizuojamas jų veikimo principas.

1.2. ToF kameros veikimas

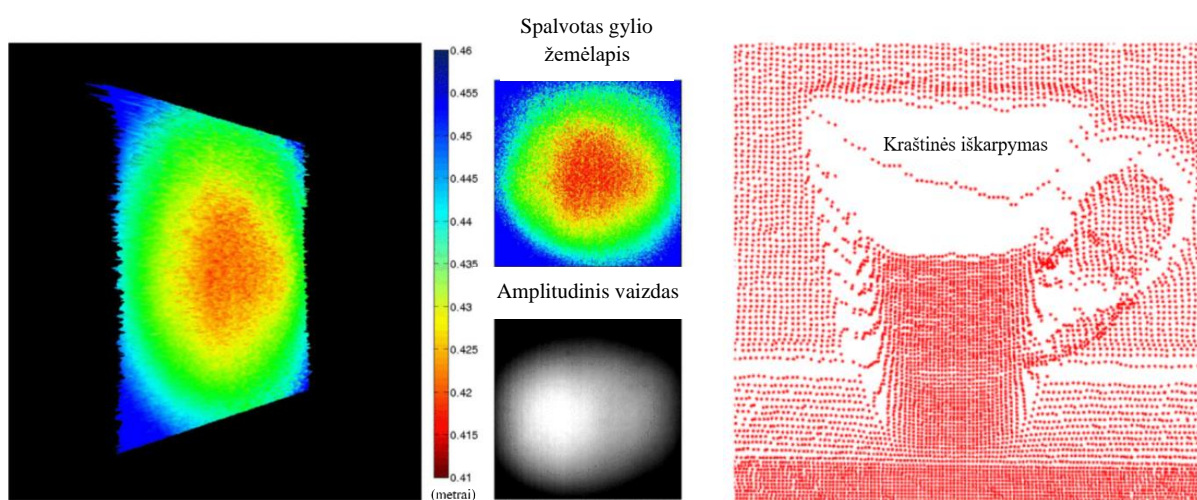
ToF kameros veikia išspinduliuodamos žinomo bangos ilgio ir dažnio šviesos spindulius, dažniausiai tai IR (infraraudonieji) spinduliai [1]. Atsispindėję spinduliai grįžta ir yra surenkama CMOS (puslaidininkis elektronikos komponentas) veikimo principu paremtu jutikliu. Tada įvertinimas praėjęs laikas bei skirtumas tarp fazių ir apskaičiuojamas kelias, kurį nukeliavo išspinduliuotas spindulys. Spindulio kelią padalinus pusiau gaunamas atstumas iki objekto, nuo kurio atsispindėjo spindulys (1.1 pav.). Kadangi kamera dirba su šviesos spinduliais, tikslumui įtakos turi optinės kameros savybės. Kameros lęšis įtakoja matavimo pakitimus einant nuo kadro centro į kraštus, tiksliausi matavimai atliekami kameros optiniame centre tolstant nuo jo didėja lęšio įtaka matavimų tikslumui. Kameros lęšiams dažnai būdingas kelių tipų iškreipimo modelis. Egzistuoja keli skirtingi lęšio poveikio modeliai. Duomenų iškraipymo modeliai yra statinės, smeigtuko, tangetinis. Praktikoje dažniausiai sutinkamas mišrus iškraipymas turintis po dalį požymių būdingų atskiriems iškreipimo tipams. Visi šie efektai stipriausiai juntami kadro kraštuose, šiose zonose matavimai turi didžiausias paklaidas. Tokie iškraipymai taisomi programiškai, nes neegzistuoja idealūs lęšiai, kurie neįtakoja kadro duomenų. Ne visos kameros turi integruotus filtras ir korekcijas tokioms duomenų pataisymo operacijoms atlikti.



1.1 pav. Atstumo matavimas naudojant fazės pokyčio įvertinimą [1]

Kameroms įtaka daro ir aplinkos sąlygos, poveikis gali būti nuo klaidingai išmatuoto atstumo iki pikselių, kuriuose nėra duomenų. Ne visos kameras pavyksta pagaminti idealiai todėl atsiranda

gamybinės paklaidos, kurios taip pat turi įtakos tikslumui [1]. Gylio iškraipymas atsiranda dėl sklaidžiamos infraraudonos šviesos, kuri praktiškai negali sklįsti taip, kaip planuota todėl atsiranda sinusoidės tipo iškraipymai dėl moduliacijos proceso neatitikimų. Paklaidos rezultatuose atsiranda dėl CMOS jutiklio netobulumo, generuojamas silicio kokybės kitimas matricoje įtakoja pikselių rezultatus, todėl netgi šalia esantys pikseliai gali turėti keliskart besiskiriančias vertes. Amplitudiniai iškraipymai atsiranda dėl ne vienodai atspindimų spindulių, centre tiesiai prieš šviesos šaltinį jie atspindi geriau nei kraštuose (1.2 pav.). Gylio duomenų tikslumas priklauso nuo atsispindėjusių spindulių stiprumo, kuomet didesnė amplitudė tuo tikslesnius matavimų rezultatus galima gauti. Šviesos sklaidos efektas atsirandantis dėl daugelio šviesos atspindžių, tai įtakoja matavimų kokybę, nes skirtingai paveikia matricos pikselius vienoje atsiranda grįžtančių spindulių stiprio padidėjimas kituose sumažėjimas (1.2 pav.). Esant fotografuojamo objekto judesiui atsiranda išblukimo efektas kaip ir įprastose kamerose, daugelyje ToF kamerų sumuojami keli kadrai todėl šis efektas dažnai matomas sustiprėjęs. Kameros nėra tobulos, todėl reikia vengti žinomų triukšmų šaltinių jei įmanoma arba juos atpažinti ir pašalinti programiškai.

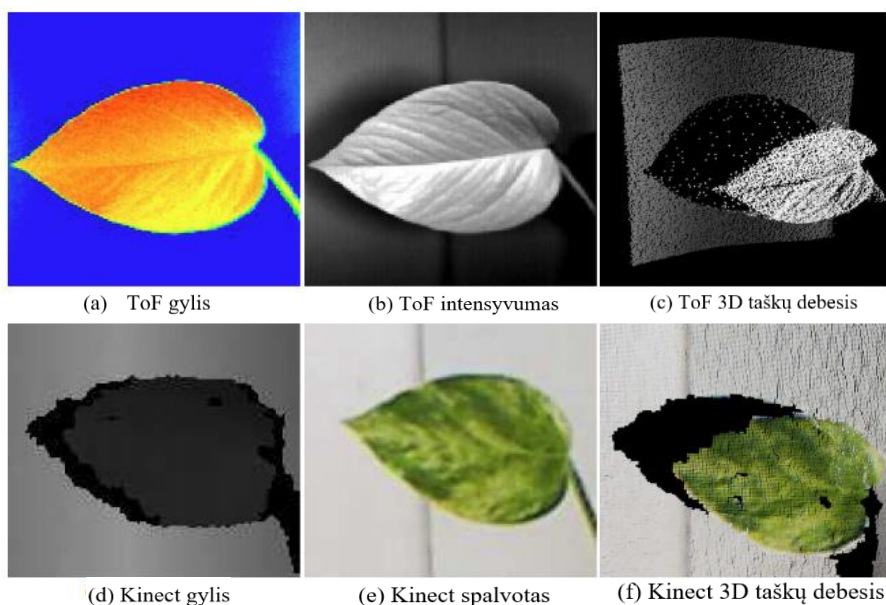


1.2 pav. Kairėje amplitudinis iškraipymas, dešinėje skirtingų atspindžių iškraipyti duomenys [1]

1.3. ToF ir RGB-D kameros

Gylio informacija registruojančios kameros skirias pagal išspinduliuojamų spindulių tipą bei koncentraciją. Vienose kamerose naudojamas IR spindulių blyksnis (Kinect Azure, Kinect v2), o kitose (RealSense 435i) spinduliuojami taškiniai IR spinduliai. Kameros CMOS jutiklis taip pat įtakoja gylio žemėlapių kokybę ir patikimumą. Vieni jutikliai skirti stebėti aplinkai, o kiti atlikti atstumo matavimo operacijas. Dauguma naujesnių kamerų geba fiksuoti mums įprastą spalvotą vaizdą bei gylio žemėlapių toks formatas vadinamas RGB-D. Gauti duomenys vėliau apjungiami ir taip gaunama išsamus ir daugiau informacijos teikiantis aplinkos vaizdas. Tokie tipo kameros pradžioje buvo skirtos žmonių stebėjimui ir sekimui erdvėje. Vėliau buvo pradėtos naudoti ir kitose srityse dėl greitaiveikos ir didelio duomenų kiekio apie aplinkos objektus lyginant su įprastomis ToF ar RGB kameromis. Kameros tikslumo įvertinimas gamybiniais uždaviniais spręsti ir robotų pozicijoms nustatyti apžvelgiamas [2] šaltinyje. Tyrimo metu lyginamos *Microsoft Kinect* – RGB-D ir *PMD CamCube 3*, *Mesa Swissranger 4K* – ToF kameros. RGB-D kameros įprastai turi didesnę raišką lyginant su ToF kameromis, RGB-D rezoliucija – 640x480, 240x940 bei naujausių kamerų siekia 1080x1920, kai tuo tarpu įprastos ToF kamerų rezoliucijos yra apie 200x200 ir naujesnėse kamerose siekia 640x480. ToF kameros leidžia arčiau priartėti prie matuojamo objekto apie 0,3m atstumu gaunami neiškreipti duomenys, tuo tarpu RGB-D kameroms reikalingas bent 0,7 metro

atstumas siekiant užfiksuoti matavimų rezultatus. Abejuose kamerų tipuose artėjant prie objekto kamerų matavimo tikslumas išsikreipia, šiai problemai spręsti kamerose taikomas automatinis fokusavimas tik ne visos kameros turi šią opciją. Abejose kamerose vyraujantys trikdžiai yra gylio žemėlapių iškraipymai bei triukšmai. Vienas didelis pranašumas, kurį siūlo RGB-D kameros tai spalvoti gylio kadrai, suteikiantys daugiau informacijos, jais naudojantis galima efektyviai užpildyti skylės, filtruoti kraštines, atpažinti objektus. Kameros gali veikti plačiame apšvietimo spektre, tačiau daugelis kamerų prasčiau veikia lauko sąlygomis esant saulės šviesai, kuri skleidžia paltų spindulių spektrą ir iškreipia kamerų spinduliuojamus spindulius. Vienas iš bandymų matomas (1.3 pav.) siekiant palyginti kamerų detalumą ir tikslumą objektai fotografuojami tomis pačiomis sąlygomis ir palyginami rezultatai. Kitame bandyme kameros turėjo atpažinti objektą, ir pagal gautas koordinates atvažiuoti ir paimti objektą. Robotas priartėjęs prie reikalingo pakelti daikto atlikdavo kelias nuotraukas ir iš jų buvo paskaičiuojamos galutinės koordinatės. Galiniame etape įvertinus kamerų rezultatus ToF kameros yra atsparesnės natūraliai šviesai bei lengviau kalibruojamos lyginant su RGB-D *Kinect* kamera. RGB-D kameros suteikia daugiau informacijos todėl lengviau atpažinti ieškomą objektą bei filtruoti gylio duomenis, nes pasitelkiami ir RGB kameros duomenys. Šaltinyje apžvelgiamos pirmosios RGB-D kamerų versijos,



1.3 pav. Atlikti kamerų bandymai lyginant RGB-D ir ToF kameras [2]

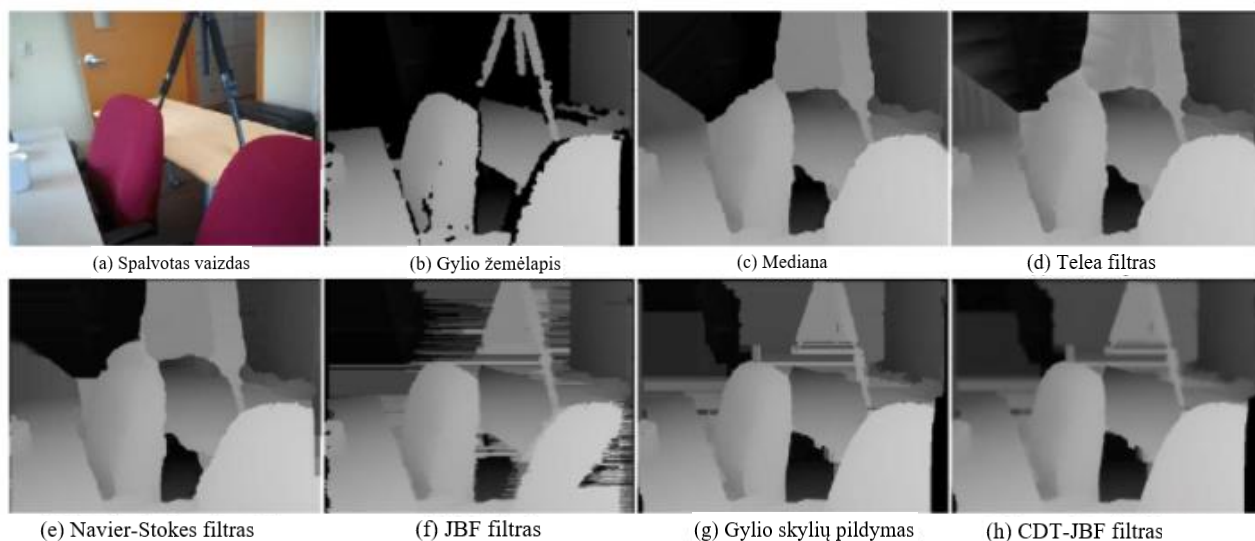
Dideliam duomenų kiekiui apdoroti reikalingi efektingi algoritmai, nors šiuolaikiniai procesoriai geba atlikti daugelį lygiagrečių skaičiavimų, per mažą laiko tarpą, tačiau beatodairiškai naudojami resursai neduoda naudos. Siekiat gauti patikimus rezultatus atliekamas pradinių duomenų filtravimas ir taisymas.

1.4. Gylio duomenų filtravimo metodai

Kamerų pateikiami gylio duomenys nėra idealus ir tolygūs. Matavimų matricoje pasitaiko įvairių trikdžių: pikselių be duomenų – skylių, netolygių matavimų, šuolių ir kitų duomenų iškraipymo būdų. Trikdžiai apsunkina darbą su gylio duomenimis bei mažina aptikimo algoritmų tikslumą. Egzistuoja keletas skirtingų filtravimo metodų. Dažnai vieno gylio filtro nepakanka žemėlapių pataisymui, skirtingoms problemoms spręsti reikalingi skirtingi filtrai [3]. Galima įvardinti tris pagrindines kadro duomenų iškreipimo problemas 1) Pikselių sumaišymas, 2) prarasti pikselių duomenys, 3)

aptriukšminti pikseliai. Kameros duomenų iškraipymą ar praradimą įtakoja blizgūs arba didelio kontrasto paviršiai, kurie išsklaido IR spindulius arba juos sugeria. Tokie iškraipymai dažnai taisomi pačių kamerų apdorojimo operacijose, nes iškraipymai yra per dideli norint toliau dirbti su duomenimis. Tai pat dažnai prisideda ir optiniai iškraipymai atsirandantys dėl netiesiškai apspindėto IR spindulio ar pačios kameros lęšio įtakos spinduliams. Prarastiems ar iškreiptiems duomenims atstatyti pasitelkiami adaptyvūs algoritmai. Įprasti matematiniai filtrai tokie kaip medianos, vidurkio ne visada duoda norimą rezultatą.

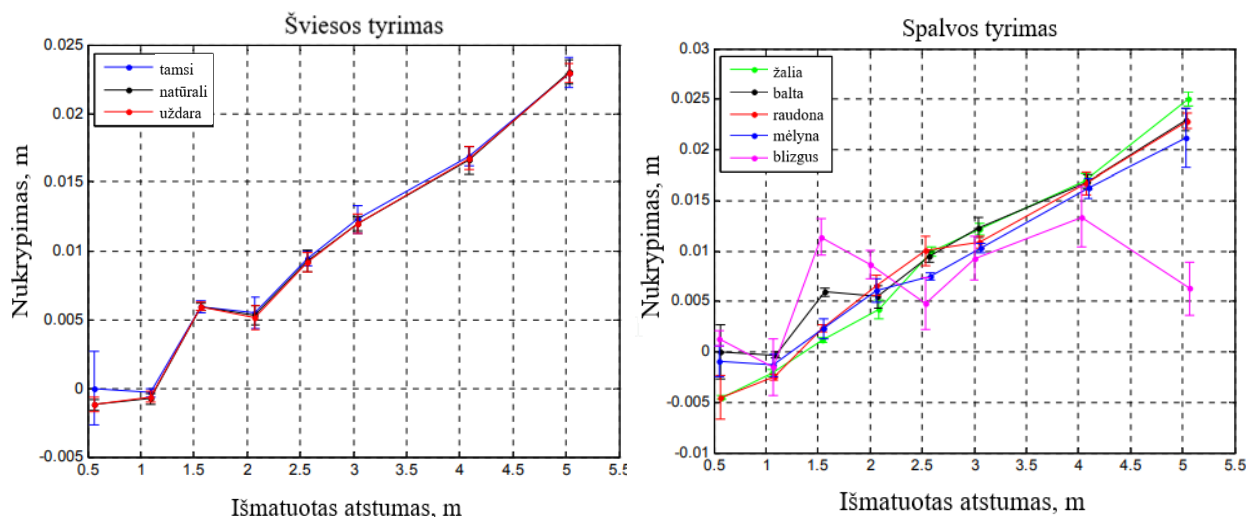
Šaltinio autoriai remdamiesi esamomis žiniomis kuria filtravimo algoritmą paremta nauja struktūra. Skylių atsirandančios kontūruose arba objektų kraštuose dažna problema, kai nėra pikselių nusakančių tikslią objekto ribą, sudėtinga atkurti tikslias ribas. Skylių pildymas remiamasi ne įprastu medianos filtru, o filtravimu skirtingomis kryptimis atstatant šalia esančius pikselius. Šalinant triukšmus pasitelkiami aplinkiniai plotai ir ieškoma anomalijų pikselių. Vienas pikselis nebūdingas tam tikram plotui koreguojamas. Sukūrus algoritmus atliekami jų bandymai įvairiomis sąlygomis su skirtingais objektais. Palyginus žinomus filtrus medianos, Telea, Navier-Stokes metodus su sukurtais naujais skylių filtravimo metodais matomi aiškesni objektų kontūrai, bei stabilūs perėjimai (1.4 pav.). Taip pat pastebimas sumažėjęs triukšmų kiekis lyginant su originaliu vaizdu ir kitais metodais. Šiame darbe buvo pateiktas naujas metodas, kuriuo naudojantis galima atstatyti prarastų pikselių vertes bei sumažinti triukšmo lygį kadre.



1.4 pav. Gylio žemėlapio apdorojimo rezultatai pasitelkus skirtingus filtrus [3]

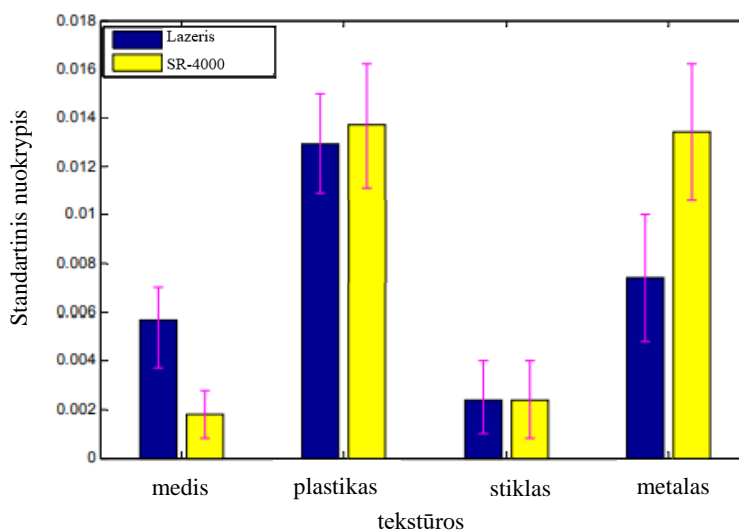
Iškreiptiems pikseliams aptikti buvo naudojama bendroji gylio transformacija – CDT. Triukšmams mažinti naudojama ta pati CDT skirtingomis kryptimis, filtruojant vaizdą. Eksperimentiniai rezultatai parodė, kad siūlomas gylio skylių užpildymo būdas (1.4 pav. g) ir CDT-JBF filtras (1.4 pav. h) teikia tikslesnę informaciją nei medianos ar jungtinis dvišalis filtras – JBF filtras.

ToF kameras veikiantys aplinkos veiksniai sukelia skirtingus rezultatus galiniame vaizde. Juos ištyrus galima jų išvengti arba juos kompensuoti programinėmis priemonėmis jei jų neįmanoma pašalinti fiziškai [4]. Tyrimu metu atliekami keli skirtingai bandymai ir palyginami rezultatai. Pradžioje palyginama aplinkos apšviestumo įtaka matavimo tikslumui. Bandymu metu aplinka buvo tamsi, su natūralia saulės šviesa, uždara su dirbtine šviesa. Gauti rezultatai visais atvejais gana artimi (1.5 pav.) grafikas kairėje. Rezultatai šiek tiek prastesni esant tamsiai aplinkai, visų atvejų matavimo tikslumas mažėja augant matavimo distancijai. Sekančiame bandyme, kurio rezultatai matomi (1.5 pav.) dešinėje, ištyrus spalvos įtaką tikslumui pastebėta jos blizgūs paviršiai įtakoja anomalinius rezultatus, ir taip pat augant atstumui tikslumas mažėja.



1.5 pav. Tyrimas įvertinant apšviestumo įtaką kairėje ir spalvų įtaką matvimams dešinėje [4]

Atlikus bandymus su skirtingomis medžiagomis ABS plastikumu, plieniu, medžiu ir stiklu (1.6 pav.) matoma jog labiausiai rezultatai nukrypa su plastikumu ir metalu. Apibendrinus rezultatus teigiama jog didėjant atstumui matavimo paklaida didėja visais atvejais. Skirtingos spalvos ir tankio paviršiai iškreipia matavimus, o blizgūs paviršiai sukelia nenuspėjamus rezultatus. Apšviestumas taip pat daro įtaką matavimo kokybei, nors poveikis nėra labai ryškus. Problemoms spręsti galima pasitelkti filtrą paremtą PF-SVM. Filtras parenka geriausius rezultatus davusius parametrus ir pritaiko paklaidoms minimizuoti. Panaudojus tokį filtrą pavyko sumažinti paklaidas nuo 8.6 mm iki 4.6 mm kai matavimo atstumas kito 0.5–5 metrų ribose.



1.6 pav. Bandymai atlikti su skirtingomis medžiagomis [4]

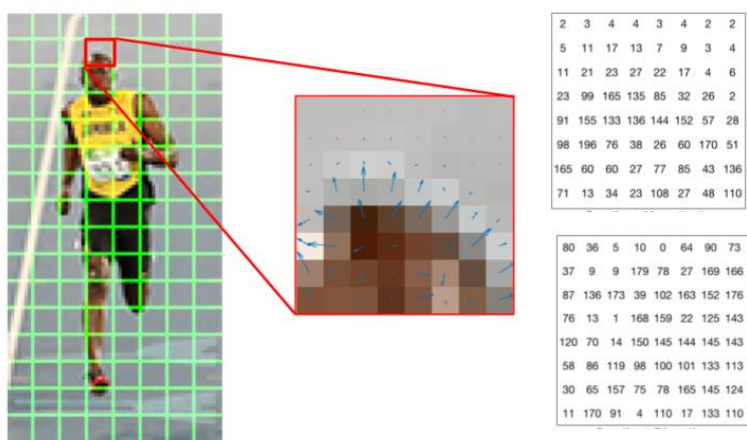
1.5. Aptikimo metodų tipai ir veikimas

Atpažinimo ir identifikavimo realizavimas pasinaudojus kompiuterinę regą atveria daugelį galimybių. Turint tiksliai ir be klaidų veikiančią sistemą galima iš dalies pakeisti darbuotoją. Tačiau objektų aptikimas vis dar nėra patikimas nors ir egzistuoja keletas aptikimo strategijų. Viena populiariausia ir labiausiai augančių aptikimo sričių yra paremta neuroniniais tinklais. Jei priskiriami tokie metodai kaip YOLO, Faster R-CNN, SSD, ir keletas kitų. Seniau naudojami metodai buvo paremti panašumų ir savybių paieška. Vienas patikimesnių metodų naudojamų prieš neuroninių tinklų paplitimą buvo vektorių atrankos mašina arba vektorių klasifikatorius – SVM, kuriam duomenys pateikiami naudojant orientuotų gradientų histogramas – HOG. Jiems naudoti reikėjo mažiau resursų, bet aptikimo ir tikslumo rodikliai buvo mažesni, nei šiuolaikiniu metodu.

1.6. HOG su SVM klasifikatoriumi

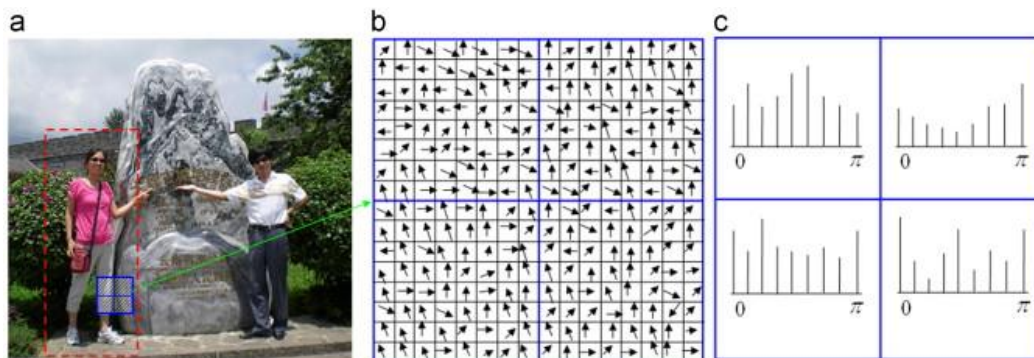
HOG ir SVM naudojami kartu, tokio tipo algoritmai nėra sudėtingi, tačiau jų aptikimo galimybės nėra tekinančios daugelį uždavinių. Metodo veikimas nuo nuotraukų analizės iki klasifikatoriaus apmokymo pateikiamas [5] šaltinyje. Šis kombinuotas metodas šaltinyje naudojamas žmonių atpažinimui. Kadangi metodą galima suskaidyti į dvi dalis pirma vadinama HOG, antroji SVM. Naudojant HOG metodą išgaunamos vaizdo savybės, pikselių segmentams. Tokių būdu išlaikant tą patį informacijos kiekį, sumažinamas apdorojimo duomenų kiekis ir paspartinamas aptikimas.

Prieš sudarant vektorinę vaizdo žemėlapi pasirenkamas vaizdo dydis. Dažnu atveju vaizdas būna 64x64 pikselių arba 128x64 pikselių dydžio. Tokio dydžio pilnai užtenka išgauti tikslinga informacija apie objektą (1.8 pav. a). Kadangi autoriaus tiriami objektai, tai yra žmonės, geriau telpa į stačiakampio matrica, todėl naudojama 128x64 pikselių matrica. Generuojant vektorinę histogramą pradžioje iš vaizdo gaunami kontūrai. Gavus juodai baltą paveikslą su objektų kontūrais, vaizdas skaidomas į 8x8 dydžio pikselių segmentus, kuriuose kiekvienam pikseliui paskaičiuojama vektoriaus kryptis bei dydis (1.7 pav.). Gavus dvi skaitines matricas, duomenys sugrupuojami į 9 grupes, kas 20 posūkio laipsnių, įsirašant dydžio vertes (1.8 pav. c). Taip susumavus visus pikselius gaunama histograma, nusakanti vektorių kampų vertes kas 20 laipsnių. Sekančiame etape paimama didžiausia histogramos vertė ir priskiriame visai 8x8 pikselių matricai (1.8 pav. b). Taip slenkama ir kartojama per visą vaizdą kol sudaromas pilnas vektorinis žemėlapis.



1.7 pav. Analizuojamas vaizdas sudalintas 8x8 matricomis kairėje ir skaitinės matricos vertės dešinėje [6]

Turint vektorinę vaizdo analizę, buvo pasitelkiamas klasifikatorius atpažinti objektams. SVM algoritmo tikslas surasti tokia erdvę susidedančia iš N -dimensijų (N – savybių arba parametru skaičius), kuri aiškiai klasifikuoja duomenų taškus. Stengiamasi rasti tokią sritį, kurioje tarp grupių yra didžiausia atskirtis. SVM erdvėje gaunamos dvi klasės, kurioms paieškos metu bandoma priskirti objektus. Objektas patekęs į klasės teritoriją užskaitomas kaip tos klasės objektas. Objektai, kurie nepatenka į klasės zonas priskiriami prie artimiausio sprendinio klasės. Todėl SVM metodui realizuoti yra reikalingi pozityvus ir negatyvus pavyzdžiai. Pozityvus sudaryti iš ieškomo objekto vaizdų, negatyvūs sudaryti iš šalutinių objektų paveikslų.



1.8 pav. a) tiriamas vaizdas, b) gautas vektorių žemėlapis, c) pikseliai suskirstyti į vektorių histogramas pagal kampus [5]

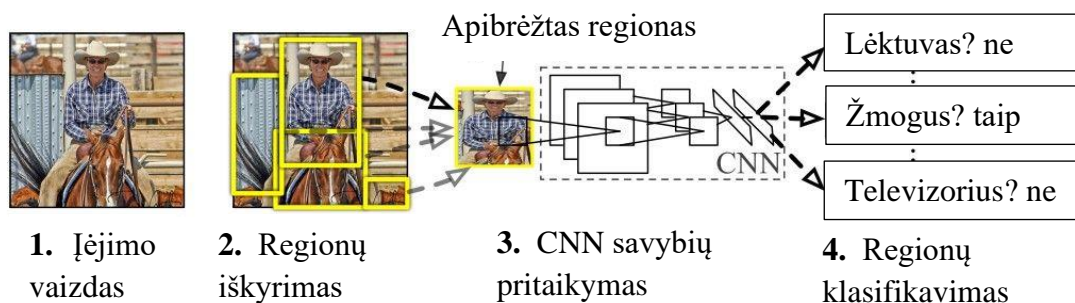
Ieškoti objekto vaizde galima naudojant HOG ir ieškant didžiausiu procentu sutampančio ploto vaizde. Tačiau pasitelkus SVM gaunami tikslesni rezultatai, bei sutrumpinamas aptikimo laikas. Apmokytas modelis slenkamas vaizde ieškant atitikmens ar sutapimo viršijančio pasirinktą slenkstinę vertę. Radus pasikartojančia sutapimų zoną galima teigti jog aptikome ieškomą objektą, ir tai nėra triukšmai ar atsitiktinumas. Slenkamai matricai dažnai pritaikomas dydžio keitimas norint praplėsti imtį bei rasti artimesnius ar tolimesnius objektus. Metodas nėra greitas ir galintis veikti realiu metu, jam taikomi spartinimo metodai dažnu atveju sumažina aptikimo tikslumą. Aprašytame darbe pasitelkiami du spartinimo būdai ir išbandomi su skirtingais duomenų apmokymo rinkiniais. Bendriniame duomenų rinkinyje pateikiama 12 tūkst. pozityvių bei beveik 3 tūkst. negatyvių apmokymo pavyzdžių. Apmokius metodą aptikimo greitis lenkia CNN, o panaudotos spartinimo priemonės nedaro įtakos tikslumui. Antru atveju naudojama apie 3 tūkst. pozityvių bei 5 tūkst. negatyvių pavyzdžių. Šiuo atveju kamera stebi žmones iš viršaus todėl neatsiranda persidieгимų. Tyrimo metu autoriaus gauti rezultatai su HOG ir SVM metodu yra artimi CNN metodui tikslumu ir greitaveikai. HOG ir SVM metodas nėra sparčiausias, tačiau veikia patikimai ir tiksliai, todėl galima identifikuoti ieškomus objektus vaizde užtikrintai. Realaus laiko operacijos praktiškai neįmanomos dėl didelio kiekio skaičiavimų atliekamų paieškos metu. Bandymuose atliktuose su nuotraukomis metodas veikia stabiliai bei randa ieškomus objektus, tyrimo metu tai buvo žmonės. Pateikiamų vaizdų raiška nėra didelė todėl tai yra viena iš priežasčių kodėl algoritmas veikia sparčiai ir tiksliai. Tyrimo metu naudoti vaizdai yra 320x480 arba 640x480 dydžio matricos.

1.7. Faster R-CNN

Konvoliuciniai neuroniniai tinklai naudojami objektų aptikimui nebuvo spartūs bei tikslūs, todėl turėjo ribotas praktinio taikymo galimybes. Vėliau atsiradus R-CNN metodui aptikimo kokybinei rodikliams pagerėjo. Lūžis šiame metode įvyko atsiradus Fast R-CNN metodui kuris, leido sparčiai

aptikti objektus bei buvo vienas tikslesnių metodų. Sekančiame žingsnyje jį patobulinus gautas Faster R-CNN aptikimo būdas, kuris buvo dar greitesnis už Fast R-CNN.

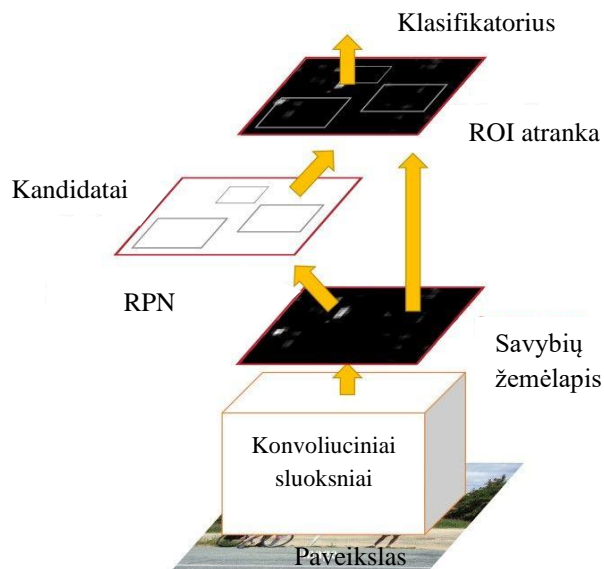
R-CNN – regionais paremtas konvoliucinis neuroninis tinklas. Metodas gali įsiminti iki 80 skirtingų objektų, esančių vaizde [8]. Šis tinklas išgauna savybes remdamasis konvoliuciniu neuroniniu tinklu. R-CNN susideda iš trijų pagrindinių dalių (1.9 pav.). Pirmoje dalyje nuskaitomas analizuojamas vaizdas. Antroje dalyje sugeneruojama apie 2 tūkst. galimų objektų, pasinaudojus selektyvios atrankos algoritmu. Trečiame etape pakeičiamas vaizdo dydis ir išgaunami požymių vektoriai iš kiekvieno regiono naudojant konvoliucinius tinklus - CNN. Ketvirta dalis naudoja apmokytą SVM algoritmą tam kad suklasifikuotų pateiktus regionų bandinius, į objektus ir foną. Metodas taip pat turi keletą trūkumų. Išgautos mokomų objektų savybės užima nuo kelių iki keliasdešimt gigabaitų vietos. Metode naudojamas požymių išskyrimo algoritmas užtrunka ilgai, bet to negali būti pritaikytas konkrečiai problemai spręsti. Kiekvienas regionas analizuojamas atskirai, todėl šio metodo praktiškai neįmanoma taikyti realaus laiko operacijose tiesiogiai apdorojant vaizdą.



1.9 pav. R-CNN algoritmo veikimo etapai, atliekant objekto paieška [7]

Fast R-CNN yra R-CNN pagrindu veikiantis metodas, gautas patobulinus vaizdo duomenų apdorojimo algoritmą. Modelis susideda iš vieno apdorojimo etapo kai R-CNN naudojami trys. Pateiktas analizei vaizdas išanalizuojamas ir atgal gražinama informacija su objektų pozicija bei priskirtomis klasėmis. Savybių žemėlapis gautas po paskutinio konvoliucinio sluoksnio perduodamas aktualių regionų analizei – ROI. Kitaip pasakius ROI analizatorius padalina paieškos regionus į celes, suformuodamas tinklėlį. Maksimumo analizė pritaikoma kiekvienai celei tinklėlyje, ir gražina vieną vertę. Visos tinklėlio vertės atitinka savybių arba kitaip vadinamą svorių žemėlapi. Gauti svoriai toliau naudojami vaizdo analizei. Normalizuotas eksponentinės funkcijos sluoksnis naudojamas nustatyti tikėtiniausią klasę, kuriai priklauso objektas. FC – pilnai sujungtas sluoksnis, naudojamas atlikti prognoze ir apskaičiuoti tikslia objekto poziciją, kuri pažymima vaizde.

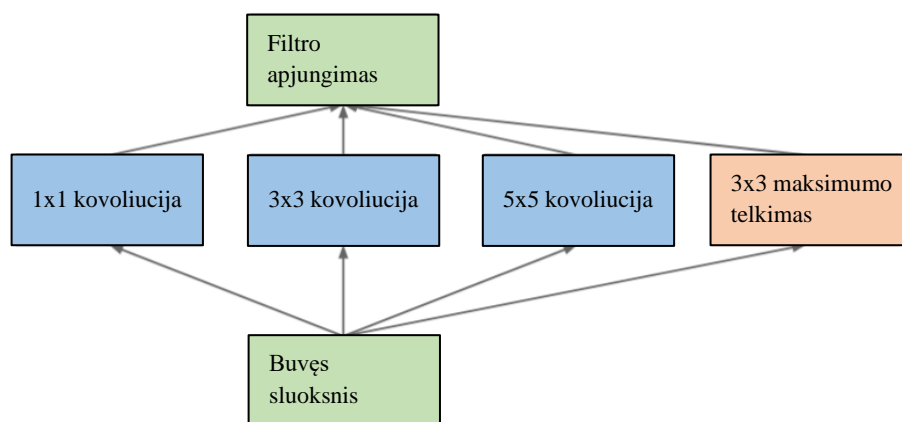
Faster R-CNN gautas pagreitus ir patobulinus Fast R-CNN metodą. Aptikimas tapo spartesnis panaudojus RPN – regionų parinkimo tinklą. RPN pateikia vaizdo dalis, kuriose galimai yra ieškomi objektai taip paspartindamas vaizdo analizę. Objektu aptikimą galima skirstyti į kelis etapus. Pradžioje RPN sugeneruoja galimas vaizdo dalis su objektais. Iš kiekvienos vaizdo dalies naudojant ROI ištraukimas vektorių savybių žemėlapis. Gauti vektorių rinkiniai suklasifikuojami pasitelkus R-CNN. Galiausiai įvertinus duomenis gražinamas pažymėtas rastas objektas, su didžiausia klasifikavimo klasės sutapimo verte. Metodo veikimas grafiškai atvaizduotas (1.10 pav.). Jame matomas duomenų šakojimas į du srautus.



1.10 pav. Konvoliucinių sluoksnių pasiskirstymas tarp modulių [7]

1.7.1. InceptionV2 tinklo savybės

Objektai paveiksluose dažnai būna įvairaus dydžio. Objektui esant arti jis užima didelę dalį nuotraukos, jam tostant tik mažesnę fragmentą. Ieškomo objekto dydis kinta su kiekvienu vaizdu. Todėl parinkti tinkamą matricos branduolio dydį tampa sunku. Didelis matricos branduolys geriau pasiteisina esant globaliai informacijai, o mažas branduolys esant lokaliai išdėstytiems duomenims. Sudėtingi tinklai turintys daug sluoksnių yra linkę persimokinti. Taip pat atlikti dideliais kovoliucines operacijas užima daug laiko bei reikalauja didesniu techninių resursų. Šioms problemoms spręsti pasitelkiami kelių dydžių filtrai, naudojami tame pačiame lygyje. Taip neuroninis tinklas tampa platesnis bet ne gilesnis (1.11 pav.), tai reiškia jog nepridedamas papildomas sluoksnis, o tiesiog prasiplečia jau naudojami sluoksniai. Tokiu pagrindu buvo sukurta *inception* tinklo struktūra, kuri atlieka kovoliucija su 3 skirtingo dydžio filtrais (1x1, 3x3, 5x5). Be to atliekamas maksimumo radimas kiekvienam vaizdo regionui ir gauti duomenys siunčiami į tolimesnius tinklo sluoksnius. Taip gaunamas spartesnis patobulintas tinklo modelis, veikiantis platesniame dydžiu spektre.

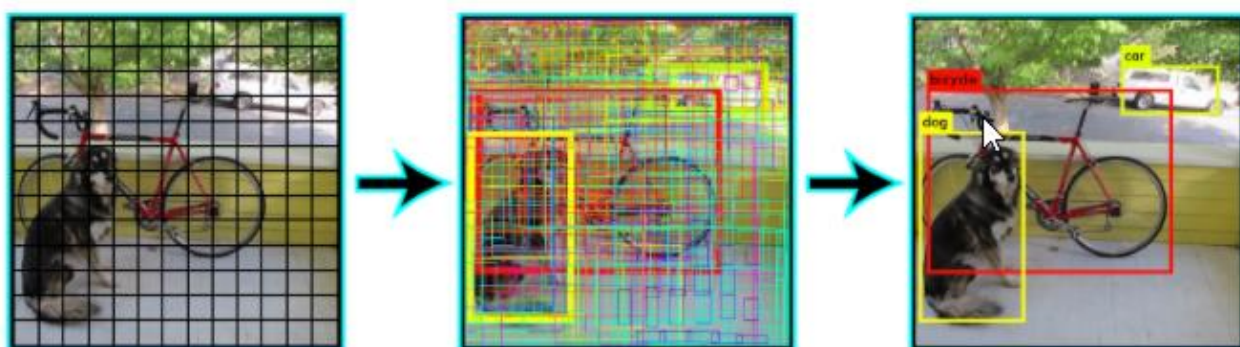


1.11 pav. InceptionV2 tinklo struktūra [9]

1.8. YOLO (You Look Only Once)

YOLO atsirado po Faster R-CNN lenkdamas savo pirmąją greitaveiką. Šis metodas bene plačiausiai taikomas praktikoje ir yra nuolat tobulinamas. YOLO neatlieka daugkartinio vaizdo skenavimo, atpažinimas atliekamas iškart, analizuojant vaizdą. YOLO dažnai taikomas realaus laiko vaizdo analizėms bei objektui aptikimui [10]. Tačiau esant mažiems objektams arba objektams grupėse, aptikimo kokybė prastėja. Lyginant su Faster R-CNN jis nėra toks tikslus ir rečiau teisingai identifikuoja objektus, bet egzistuoja keletas YOLO kaip ir Faster R-CNN modelių, kurie yra patobulinti ir skiriasi nuo originalių metodų.

YOLO atlikdamas objektų paiešką vaizde pradžioje padalina paveikslą į mažesnes dalis, tarsi tinklelį. Kiekviena tinklelio dalis yra $S \times S$ dydžio matrica, ir yra atsakinga už objekto indentifikavimą, kuris patenka į matricos vidų. Atitinkamai šis tinklelis apibrėžia galimas aptikimo koordinatas, kurios sutampa su tinklelio celėmis bei tikimybę jog objektas bus langelyje. Tai padeda sumažinti skaičiavimų skaičių, nes aptikimas ir aptižinimas yra apdorojami tinklelio celėmis. Tačiau šis metodas turi ir neigiamą pusę. Langeliuose esančios objekto dalys priskiriamos tam pačia tipui, taip kuriami dublikatai (1.12 pav.). YOLO pasitelkia NMS taip parinkdamas tinkamiausias objekto ribas ir pašalindamas dublikatus. Panaudojus šį metodą visi kvadratai turintys mažesnes tikimybinės vertes, jog jame yra ieškomas objektas atmetami. YOLO slenkstinę ribą apsprendžia imdamas didžiausią rastą vertę. Jei artimos vertės sutraukiamos į grupę, kuri nusako objekto poziciją. Taip slenkamasi per visus plotus vaizde kol atliekama pilna analizė. YOLO architektūra įprastai sudaro 24 konvoliuciniai sluoksniai, su dviem apjungiančiais sluoksniais gale.



1.12 pav. YOLO atliekama vaizdo analizė [11]

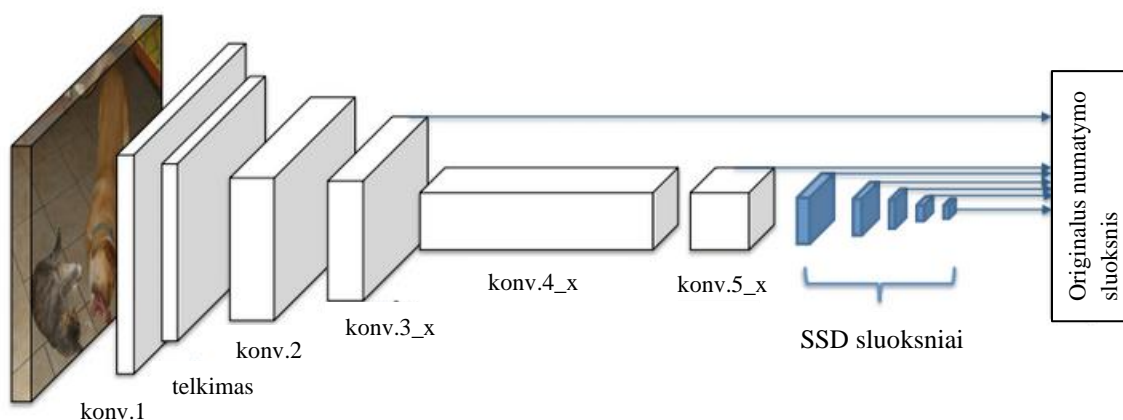
1.9. SSD (Single Shot Detector)

SSD kaip ir YOLO objektų aptikimui naudoja vienkartinį paveikslo skenavimą. Ši savybė užkoduota pavadinimo santrumpoje SSD, ką pažodžiui išvertus iš anglų kalbos gautume vieno šūvio detektorius. Tokie modeliai dažniau sutinkami realaus laiko aplikacijose, kur reikalinga didelė sparta. Naudojamas vienetinis kadro skenavimą leidžia sumažinti paieškos laiką, bet atsiliepia aptikimo tikslumui. Siekiant pagerinti rezultatus SSD modelio struktūra galima modifikuoti [12]. Metodas paremtas konvoliuciniu tinklu, kuris sukuria fiksuoto dydžio langelių rinkinį su svoriniais klasifikavimo koeficientais, kad tame langelyje yra viena iš ieškomų objektų klasių. Po to atliekamas NMS žingsnis, galutinai pasakyti objektų klases.

Pirminiai tinklo sluoksniai yra pagrįsti standartine architektūra, naudojama aukštos kokybės vaizdų klasifikavimui, kuris vadinamas baziniu tinklu (1.13 pav.). Tada prie tinklo prijungiama pagalbinė

struktūra, kad būtų galima realizuoti pagrindines funkcijas. 1) Bazinio tinklo gale pridamas kovoliucinis sluoksnis. Šio sluoksnio matricos dydis stabiliai mažėja, taip leisdamas identifikuoti skirtingų dydžių objektus. Tai suteikia pranašumą lyginant su YOLO, kuris geba veikti vieno dydžio mastelyje.

2) Toliau įtrempiama prognozių struktūra. Kiekvienas pridėtas funkcinis sluoksnis gali sukurti fiksuotą numatytų aptikti objektų rinkinį. Savybių sluoksnyje, kurio dydis $m \times n$ su p kanalų skaičiumi naudojamas $3 \times 3 \times p$ nedidelis branduolys, kuris kiekvienai kategorijai sugeneruoja svorinį koeficientą ir rasto ploto poslinkio vertę. Visame matricos $m \times n$ plote pritaikomas tas pats branduolys, gražina išeigos vertes. Aptikimo stačiakampio poslinkiai išmatuojami reliatyviai nuo numatytojo stačiakampio, visame vaizdo žemėlapyje. YOLO architektūra naudoja tarpinį visiškai sujungtą sluoksnį, o ne konvoliucinį filtrą, šiai operacijai atlikti. 3) Aptikimo laukai ir santykiniai koeficientai. Kiekvienas gautas gaubiantis stačiakampis susiejamas su svorių žemėlapiu dalimi. Gauti stačiakampiai objektų žemėlapij perdengia taip, kad kiekvieno langelio padėtis atitinkamo atskaitinio langelio atžvilgiu būtų fiksuota. Visuose objektų žemėlapiuose numatomi poslinkiai ir objektų klasių svoriniai balai, kurie rodo kurios klasės egzempliorius yra tikėtiniausias gaubiančiame plote. Sujungus visus pridėtus bei bazinius sluoksnius gaunama nauja tinklo struktūra. Pridėti patobulinimai leidžia tiksliau parodyti objektų dėžučių mastelį bei poziciją.

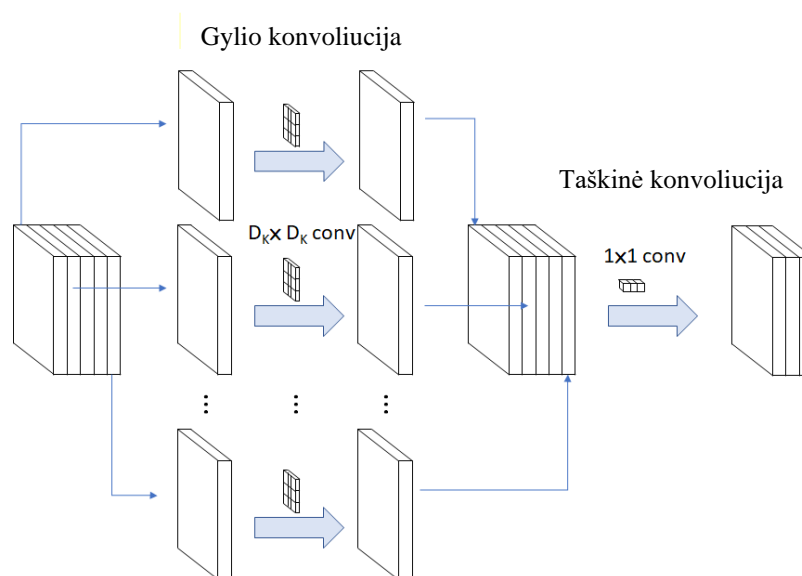


1.13 pav. SSD tinklo schema [12]

1.9.1. MobileNetV1 tinklas

SSD aptikimo modelis papildomas *MobileNetV1* tinklo struktūra. Toks medelio patobulinimas leidžia praplėsti aptikimo galimybes ir pasižymi naudingomis savybėmis [13]. Šiuo atveju *MobileNetV1* leidžia sumažinti tinklo struktūrą, o tai padaro šį modelį naudotiną mobiliuose ir integruotose aplikacijose. Kartu su modeliu naudojami parametrai gali būti nesunkiai pamedifikuoti pagal poreikį. Pagrindinei tinklo parametrai yra pločio koeficientas α ir raiškos daugiklis ρ . Tinklas pasižymi keletu išskirtinių savybių. Tinkle naudojama gylio sluoksniu kovoliucija pereinanti į taškinę (1.14 pav.). Gylio kovoliucija yra kanalu $DK \times DK$ erdvinė kovoliucija. Taškinė kovoliucija yra 1×1 , skirta pakeisti dimensijas. Kai $DK \times DK$ yra 3×3 , pereinant ant 1×1 dydžio matricos galima sumažinti skaičiavimus iki 9 kartų, nežymiai sumažėjus tikslumui. Tokiu būdu sumažėja reikalingas pridėtinių parametru skaičius, lyginant su originaliu *Conv MobileNet* tinklu, jų kiekis sumažėja 9 kartus tačiau tikslumas nukenčia tik apie 1%. Pločio koeficientas α leidžia išgauti paprastesnį modelį ir valdyti kanalų skaičių arba kanalų gylį. Rezoliucijos daugiklis ρ kontroliuoja įvesties vaizdo skiriamąją gebą, sumažindamas apdorojimų duomenų kiekį. Apibendrinus visus pakeitimus ir palyginus naują ir

originalų modelių teigiama jog išgaunamas panašus našumas kaip ir naudojant moderniausius metodus. Tačiau tinklas yra daug mažesnis, dėl pagrindinės savybės – skaidomos gylio kovoliucijos.



1.14 pav. Sluoksniuose naudojamų kovoliucinių metodų palyginimas[13]

1.10. Aptikimo rezultatų kokybinis įvertinimas

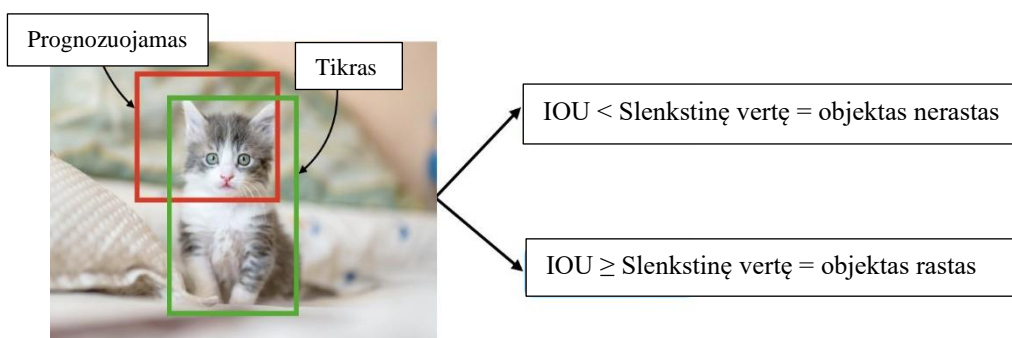
Aptikimo metodų efektyvumą bei tikslumą nėra paprasta įvertinti. Gauti rezultatai matomi tik vaizduose, o pozicijas nurodo koordinatės. Vienas iš metodų įvertinti naujai kuriamiems metodams yra COCO testai [14]. Ši etaloninė sistema naudojama naujai sukurtų aptikimo metodų palyginimui. Sistema pateikia duomenų rinkinį pagal, kurį algoritmai mokosi ir vėliau pateikiami vaizdai su ieškomais objektais. Objektų koordinatės yra žinomos, gavus rezultatus sistema įvertina kiek laiko užtruko objekto aptikimas ir koku tikslumu nustatyta objekto pozicija vaizde. Visiems metodams tirti naudojama ta pati apsimokymo bei testavimo duomenų imtis. Taip galima adekvačiai palyginti metodus tarpusavyje, atlikus keletą testų iš eilės. Dalinai metodų tikslumą įtakoja sistemos našumas, todėl visi testai atliekami tojo pačioje platformoje. Sulyginus parametrus visiems metodams galima palyginti jų skirtumus, atmetant aplinkinius veiksnius.

Aptikimo tikslumui vaizde nustatyti naudojamas bendras sankirtos plotas – IOU, kuris kokybiškai įvertina rastų koordinatinių tikslumą [15]. Sankirtos ploto įvertinimas, vienas iš dažniausiai naudojamų kriterijų rezultatų tikslumui įvertinti. Norint rasti IOU vertę metodo aptikti gaubiantys stačiakampiai ir jų koordinatės lyginamos su tiksliais objekto poziciją nusakančiomis koordinatėmis. Pirmame etape apskaičiuojamas bendras abiejų stačiakampių persidengimo plotas. Sekančiame etape apskaičiuojamas bendras abiejų stačiakampių suminis plotas. Galiausiai persidengimo plotas padalinamas iš bendro abiejų gaubiančių stačiakampių ploto (1.3) formulė. Gauti rezultatai parodo kaip arti tikslų objekto koordinatinių buvo algoritmas. Naudojant slenkstinę vertę galima įvertinti ar pavyko aptikti ieškoma objektą (1.16 pav.). Jei stačiakampiai idealiai sutapo gaunama 1 vertė, kuo gauta vertė artimesnė 0 tuo mažesnis sutapimo plotas, jei gauta vertė lygi 0, reiškia jog nepavyko nieko aptikti.

Bendrai visų ieškomu objektų aptikimo tikslumui nusakyti naudojama vidutinio tikslumo vertė – AP [16]. Ši vertė parodo kaip tiksliai ir koku procentu iš visų ieškomų objektų pavyko sėkmingai identifikuoti bandinius. Šis vertinamo metodas yra dvimatis ir galutinę vertę pateikia kaip bendrą

dviejų kriterijų suminę vertę. Vidutinį tikslumą sudaro dvi dedamosios viena iš jų yra tikslumas kita rastų metodo ir visų egzistuojančių ieškomų objektų santykis.

Ši vertė auga didėjant rastų objektui skaičiui, ir lieka artima 0 jei aptiktų objektų kiekis yra mažas. Tikslumas apskaičiuojamas pasinaudojus (1.2) formulę, o aptikimo santykinis koeficientas (1.1) formule. Gautos vertės sumuojamos ir susideda į bendrą vieno objekto aptikimo tikslumą AP. Ši vertė apskaičiuojama gavus ploto sumą kadangi yra dvi dedamosios. Egzistuoja du AP gavimo metodai vienu grafikas sudalinamas į 11 verčių, nuo 0 iki 1 žingsniu 0,1. Kiekviename taške tikslumo vertė padauginama iš žingsnio vertės ir sumuojama. Skaičiuoti plotą esantį po kreive galima pasitelkti integralą arba kaip siūloma [17] šaltinyje naudoti interpoliacija (1.4) ir (1.5) formulės. Aptikimo metodui su skirtingais tirtais objektais galima paskaičiuoti vidutinę tikslumo vertę MAP pagal (1.6) formulę, kuri parodo aptikimo vidurkį visoms tirtoms objektų grupėms.



1.15 pav. Bendras sankirtos plotas (IOU) [15]

$$R = \frac{TP}{TP + FP} \quad (1.1)$$

čia R – aptikimo santykinis koeficientas; TP – tikras teigiamas; FP – klaidingas teigiamas.

$$P = \frac{TP}{TP + FN} \quad (1.2)$$

čia FN – klaidingas neigiamas, P – tikslumas.

$$\text{Sankirtos plotas (IOU)} = \frac{\text{Persidengimo plotas}}{\text{Bendras plotas}} \quad (1.3)$$

$$AP = \frac{1}{N} \sum_{r(0,0\dots1,0)} P_{interp}(R) \quad (1.4)$$

čia N – objektų skaičius, P_{interp} – gauta interpoliuota tikslumo vertė.

$$P(R) = \max P(R) \quad (1.5)$$

čia $\max p$ – didžiausios tikslumo vertės.

$$MAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (1.6)$$

2. Tyrimo įrangą ir pasiruošimas

2.1. Kameros specifikacijos

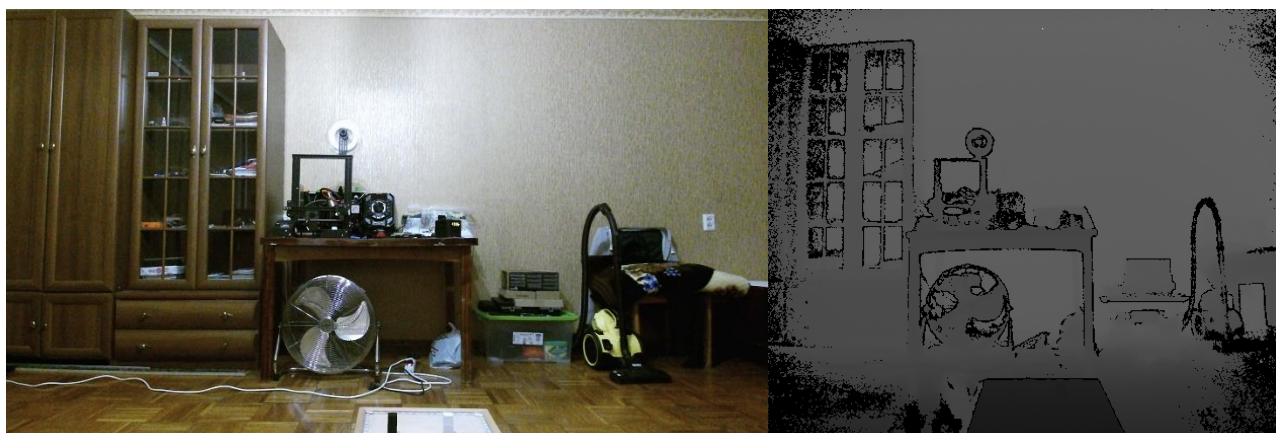
Tyrimo metu naudojama *Kinect v2* gylio kamera (2.1 pav.). Kameros specifikacijos pateikiamos 2.1 lentelėje [1]. Ši kamera IR spindulius skleidžia kaip trumpus šviesos blyksnius todėl atkuriami tikslesni ir detalesni gylio žemėlapiai. Kameros pateikiamas gylio duomenų bei spalvotas RGB vaizdas matomas (2.2 pav.). Kamera sukalibruota ja pagaminus tai reiškia, kad gylio ir RGB duomenys yra susieti ir galimas jų konvertavimas. Kadangi kameros gamintojos yra Microsoft kamera turi bibliotekų paketą skirtą Microsoft Windows operacinei sistemai. Programuoti pasirinkta *Python* kalba bei *JetBrains PyCharm* programavimo aplinka dėl patogaus bibliotekų importavimo. Be to visi tiriami aptikimo metodai realizuojami *Python* programavimo kalba, todėl lengviau suderinti ir išbandyti tiriamus metodus su kameros transliuojamais duomenimis.

2.1 lentelė. Kinect v2 kameros techninės specifikacijos

| | |
|---|---------------------------------------|
| Infraraudonųjų spindulių (IR) kameros rezoliucija | 512 × 424 pikseliai |
| Matymo kampas (IR kamera) | 70 × 60 laipsnių |
| RGB kameros rezoliucija | 1920 × 1080 pikseliai |
| Matymo kampas (RGB kamera) | 84,1 × 53,8 laipsnių |
| Kadrų skaičius per sekundę | 30 kadrų |
| Veikimo atstumas | Nuo 1,0 iki 4,5 metrų |
| Objekto dydis pikseliais nuo atstumo | 1,4mm iš 0,5m ir 12mm iš 4,5m atstumo |



2.1 pav. Kinect V2 kamera



2.2 pav. Kameros pateikiamas gylio vaizdas dešinėje ir standartinis vaizdas kairėje

2.2. Įrangos specifikacijos

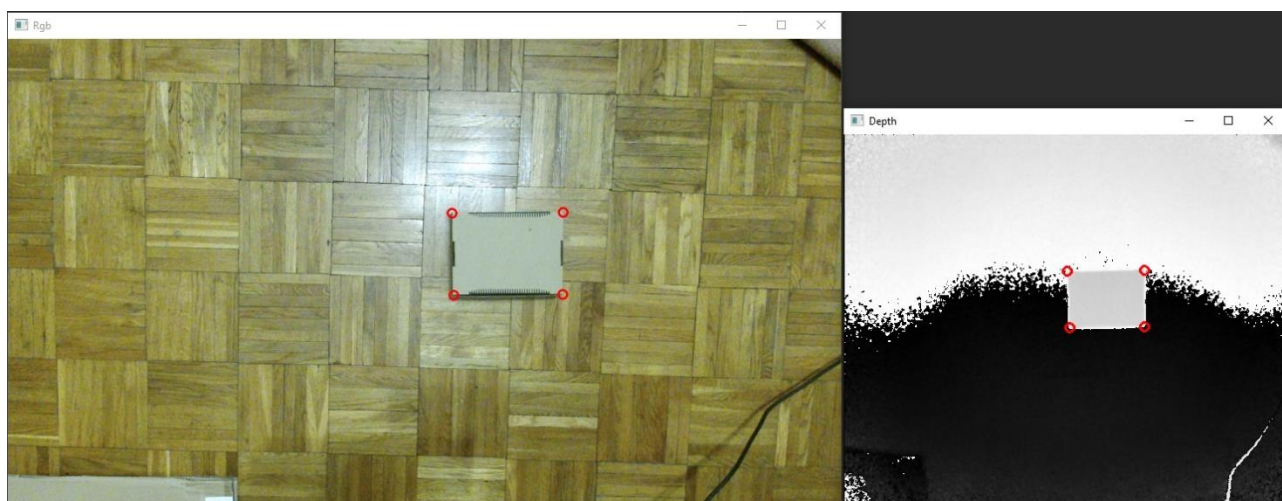
Tyrimo metu gauti rezultatai gali skirtis priklausomai nuo naudojamų įrangos našumų. Žemiau lentelėje 2.2 pateikiami naudotos įrangos parametrai.

2.2 lentelė. Naudojamos įrangos specifikacijos

| | |
|-----------------------|------------------------|
| Procesorius | Intel i7-5700HQ 2,7Ghz |
| Operatyvioji atmintis | 16GB 1600Mhz |
| Vaizdo plokštė | Nvidia GeForce GTX950M |
| Kietasis diskas | SSD Kingston 1Tb |

2.3. Gylio ir RGB kamerų testavimas

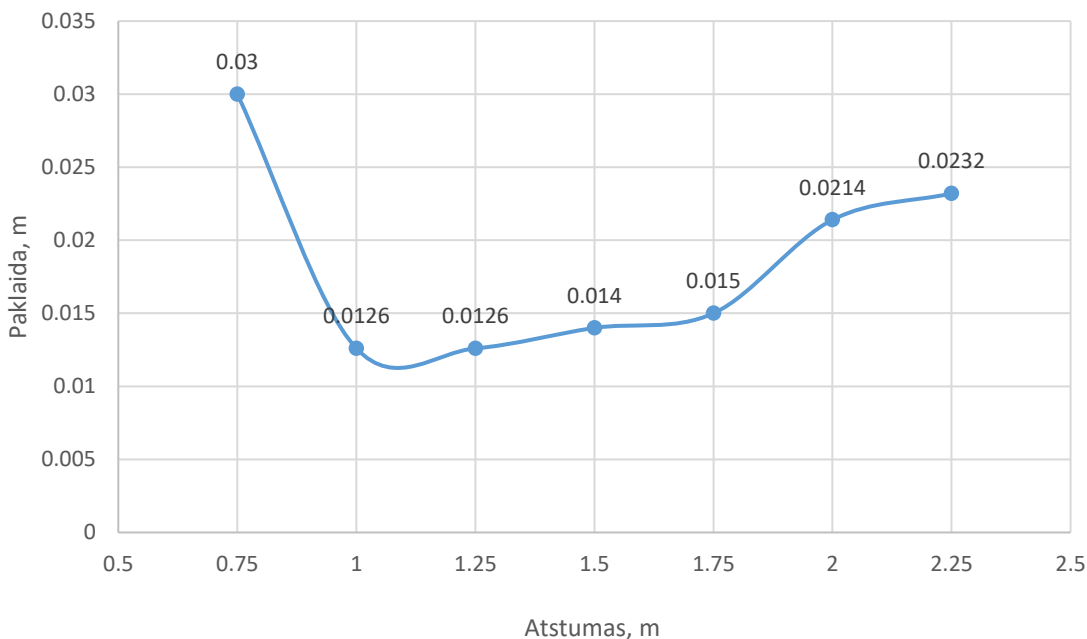
Kameros pateikiamas IR vaizdas yra nespalvotas RGB vaizdo atitikmuo, kuriame taip pat galima vykdyti objekto paiešką. Kamera geba perskaičiuoti RGB vaizde rastos koordinatės poziciją, į atitinkamą poziciją gylio vaizde. RGB ir IR kamerų suderinimo įvertinimas matomas (2.3 pav.), kur RGB pasirinkti objekto kampai atitinka gylio vaizde matomo objekto taškus. Pasinaudojus šia funkcija galima nustatyti objekto lokacija vaizde ir sutapatinti rezultatus RGB ir gylio matricose.



2.3 pav. RGB ir IR kamerų duomenų atitikimas

Pradžioje atliekami testai nustatant kameros duomenų tikslumą bei paklaidas. Atlikti bandymai padeda išsiaiškinti galimas problemas bei pasirinkti dėl techninių ribojimų galinčius neveikti metodus. Kameros naudojamas lęšis išgaubia vaizdą, todėl naudojami kompensavimo algoritmai, kurie perskaičiuoja pikselių pozicijas į tikslias ir atlieka duomenų korekcija. Toks efektas gerai matomas žiūrint į plokštumą, užfiksavus gylio matricą ir atvaizdavus ją matoma išbaubta plokštuma. Taip yra todėl, nes tolygiai nuo kameros lęšio optinio centro matuojamas atstumas iškreipiamas labiau bei stiprėja lęšio poveikis duomenims. Kameroje naudojami pataisymo metodai dalinai kompensuoja šį efektą todėl, papildomos priemonės nenaudojamos. Atliekamas kameros atstumo iki objekto įvertinimas. Tyrimo metu pasitelkiamas baltas objektas, kuris gerai atspindi IR spindulius. Keičiant atstumą iki objekto registruojami gylio rodmenys. Kiekviename taške atlikta po 8 matavimus, ir apskaičiuotas vidutinis atstumas. Grafikas pateikiamas (2.4 pav.) iš jo galima matyti jog pirmame matavimo taške kai dar nepatenkama į rekomenduojamą matavimo diapazoną pateikiamą 2.1 lentelėje, todėl paklaidos vertė stipriai išauga. Pasiekus 1 metro atstumą nuo kameros paklaida






pasiekia mažiausią vertę. Toliau didėjant atstumui tarp kameros ir taikinio paklaidos didėjimas yra proporcingas padidėjusiam atstumui. Todėl paklaidos neturėtų sukelti nenumatytų rezultatų ar stiprių šuolinių trikdžių. Bandymus su objektais galima atlikti 1 – 1,75 metrų atstumu, nes tada paklaida išlaiko stabilų kitimą. Darbo metu bandymai bus atliekami su kamera įtvirtina 1 metro aukštyje virš pagrindo.



2.4 pav. Tyrimo metu gauta paklaidos priklausomybė nuo atstumo

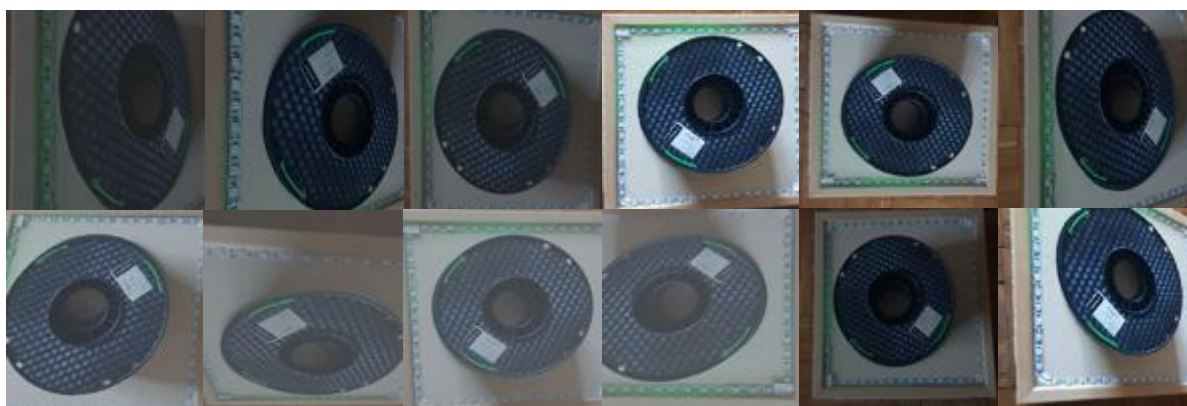
2.4. Ieškomi objektai

2.3 lentelė. Ieškomi objektai ir jų dimensijos

| | Butelis | Ventiliatorius | Replės | Varžtas | Plastiko ritė |
|--------------------------|---|---|--|---|---|
| Objekto nuotrauka |  |  |  |  |  |
| Dimensijos, mm | Ilgis: nuo 250 iki 300 Diametras: nuo 80 iki 88 | Ilgis: nuo 92 iki 140 Plotis: nuo 92 iki 140 Aukštis: nuo 15 iki 25 | Ilgis: nuo 120 iki 170 Plotis: 28 iki 110 Aukštis: nuo 10 iki 18 | Ilgis: nuo 25 iki 110 Diametras: nuo 6 iki 14 | Diametras: 200 Aukštis: 70 |

2.5. Duomenų augmentacija

Turint ribotą kiekį mokymo duomenų, sunku tinkamai parengti modelį aptikimo uždaviniui. Kuriant duomenų rinkinį savo pasirinktiems objektams, nepavyksta rasti internete egzistuojančių duomenų rinkinių. Apmokymo vaizdus tenka rinkti iš pavienių vaizdų internete bei fotografuoti ieškomus objektus. Dėl per mažo apmokymo duomenų kiekio tinklas įsimena konkrečius vaizdus, o ne jų savybes. Siekiant padidinti mokymo imtį naudojama vaizdų augmentacija. Tai leidžia iš turimo vaizdų kiekio papildomai sugeneruoti kelis kartus didesnę vaizdų imtį. Augmentacijos metu atliekamas įvairios vaizdų modifikavimo operacijos: apvertimas horizontaliai ir vertikalčiai, pasukimas į šonus įvairiais kampais, vaizdų aptriuškinimas, kontrasto keitimas, šviesumo keitimas ir keletas kitų operacijų. Šios operacijos pagerina apmokymo kokybę ir sukuria tokias sąlygas, kurias sunku išgauti kaip paminėtas šviesumo pakeitimas, ar aptriuškinimas. Augmentacijos metu sugeneruoti nauji vaizdai pateikiami (2.5 pav.).



2.5 pav. Augmentacijos taikymas duomenų rinkiniui praplėsti.

2.6. Vertinimo kriterijų apskaičiavimas

Darbe objektų aptikimo tikslumui įvertinti naudojama vidutinė kvadratinė paklaida MSE, kuri apskaičiuojama pagal 2.1 formulę. Vidutinis aptikimo laikas randamas pasinaudojus 2.2 formule. Aptiktų objektų procentas gaunamas apskaičiavus 2.3 išraišką. Kitų kriterijų vertės gaunamos naudojantis 1.10 skyrelio 1.1 – 1.6 formulėmis.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.1)$$

čia \hat{Y} – tikra vertė; Y – prognozuojama vertė, n – tiriamų verčių skaičius.

$$\text{Vidutinis aptikimo laikas} = \frac{\text{Bendra aptikimo laikų suma}}{\text{Vaizdų skaičiaus}} \quad (2.2)$$

$$\text{Aptiktų objektų procentas} = 100\% \frac{\text{Rasti objektai}}{\text{Visi bandymų imtyje esantys objektai}} \quad (2.3)$$

3. Algoritmų realizavimas ir eksperimentai

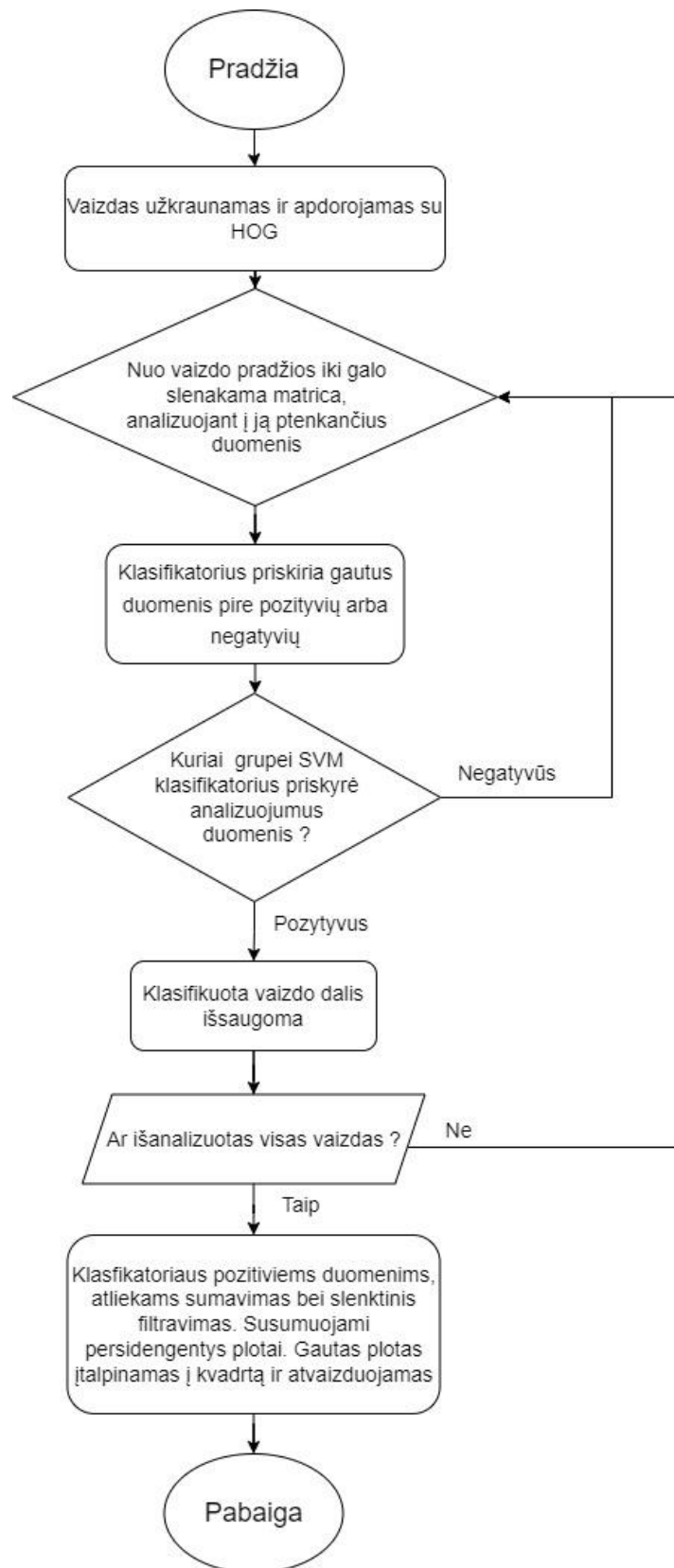
Šioje dalyje išbandomi visi pasirinkti tirti aptikimo metodai. Metodu tikslumui nustatyti bus pasitelkti vaizdų rinkiniai su žinomomis objektų koordinatėmis bei tipais. Vaizdų rinkinį apdorojus vienu iš aptikimo metodų sugeneruojama vaizdinė bei tekstinė išėiga. Tekstiniame rezultatu faile yra objektai aprašyti koordinatėmis ir priskirti grupėms. Gautus duomenis galima palyginti su tiksliais objektų koordinatėmis ir įvertinti skirtumus. Gauti rezultatai pateikiami žemiau pateikiamose (3.1 – 3.12) lentelėse. Testavimui naudojami du vaizdų rinkiniai viename natūrali dienos šviesa ir išdėlioti tik ieškomi objektai. Antrame rinkinyje dirbtinė šviesa bei vaizduose atsiradę šalutiniai objektai. Abiem bandymo vaizdų rinkiniams sudaryti buvo naudojama ta pati kameros pozicija bei atstumas iki pagrindo. Grindinio blizgumui sumažinti pasitelkiamas kartono lakštas, tai gerai atitinka realybėje dažnai sutinkamą pagrindą, ant kurio būna išdėlioti produktai. Visi gauti rezultatai yra priklausomi nuo pradinės apmokymo duomenų imties, bei turimos įrangos našumo. Tikslas yra įvertinti ar su turima įranga galima identifikuoti objektą, naudojant minimalią apmokymo imtį. Ne visada pavyksta surinkti maksimalų kiekį vaizdų sprendžiant gamybinį uždavinį, todėl adekvatu žinoti koks kiekis duomenų garantuotų objekto identifikavimą. Kaip svarbiausi kriterijai kiekvieno bandymo metu įvertinamas vieno vaizdo apdorojimo laikas, vidutinis tikslumas AP, vidutinė kvadratinė paklaida MSE (2.1) formulė ir teisingai identifikuotų objektų procentas (2.3) formulė. Lentelėse pateikiamas laiko vidurkis gautas pagal (2.2) formulę. Vidutinis tikslumas apskaičiuojamas pagal teorijoje pateiktas (1.3 – 1.4) formules, tai yra standartinis aptikimo metodų kokybės įvertinimo rodiklis. Vidutinė kvadratinė paklaida leidžia įvertinti kaip tiksliai aptinkami objektai, kuo ji artimesnė 0 tuo, mažesnis nuokrypis nuo tikros objekto pozicijos. Neradus objekto MSE prisumuojama 1. Radus paskaičiuojama paklaidos dydis, per kuria skyrėsi pažymėtas ir rastas stačiakampiai. Stačiakampiams sutapus prisumuojama 0 vertė. Objekto aptikimo tikslumas įvertinamas pasinaudojus bendru sankirtos plotu IOU, kuris apskaičiuojama pagal (1.3) formulę. Gauta vertė lyginimą su pasirinkta slenkstine verte, ją viršijus objektas užskaitomas kaip atpažintas, kitu atveju kaip neatpažintas. Tyrimo metu slenkstinis koeficientas yra 0,5. Įvertinimo operacija kartojama su visais gautais objektų duomenimis. Toliau atliekami eksperimentai su keturiais skirtingais aptikimo metodais: HOG ir SVM, Faster R-CNN, YOLO, SSD.

3.1. HOG su SVM klasifikatoriumi

3.1.1. Metodo aprašymas

Tiriamas metodas nėra naujas, jam realizuoti pasitelkiami išbandyti bei veikiantys pradiniai modeliai. Metodas realizuojamas *Python* programavimo aplinkoje, naudojant *OpenCv*, *Numpy* bei kitas darbu su vaizdas ir matricomis skirtas bibliotekas. Atlikus klasifikatoriaus apmokymą vaizduose ieškoma objektų požymių. Paieškai pasitelkiamas slenkantis stačiakampis ir analizuojama jame esanti vaizdinė informacija. Turint veikiantį algoritmą bei ištestavus jį su jau paruoštais duomenų rinkiniais galima pereiti prie pasirinktų tiriamųjų objektų. Mokymo imties dydis taip pat apsprendžia galutinio metodo tikslumą. Siekiant turėti tikslų ir visapusiškai taikoma modelį rekomenduojama surinkti bent 8000 vaizdų. Kadangi modelis yra taikomas žinomomis sąlygomis tokia didelė imtis nėra naudojama. Gaunami rezultatai yra priklausomi nuo koeficientų, kurie nulemia aptikimo tikslumą. Šiuos koeficientus tenka suderinti pagal ieškomą objektą, tai yra vienas iš šio modelio trukumų. Klasifikatoriaus apmokymas yra priklausomas nuo pateikiamų vaizdų kiekio. Klasifikatoriui blogai įsisavinus duomenis ir atlikus automatinius testus gaunamas aptikimo procentas. Jei gauti rezultatai netenkina užsibrėžtos ribos, klasifikatorius iš naujo atlieka vaizdų grupių klasifikavimą. Gavus naują

suskirstymą į grupes, vėl tikrinimą ar po testų gautos aptikimo vertės yra 90% ir didesnės. Pasiekus ribą klasifikatoriaus parametrai išsaugomi ir naudojami vaizdų analizėje. Algoritmo vaizdo analizės schema pateikiama (3.1 pav.).



3.1 pav. Objekto paieškos algoritmas naudojant HOG ir SVM

Pradžioje užkraunamas vaizdas, bei atliekamas HOG savybių išskyrimas. Gavus savybes pradeda viso vaizdo analizę slenkant per jį pasirinkto dydžio matricai. Matricoje esantys duomenys perduodami SVM klasifikatoriui. Jei klasifikatorius sugeba atpažinti vaizdo fragmentą, bei priskirti pozityviai objektų grupei vaizdo dalis įsimenama. Taip keliaujama per visa vaizdą, kelis kartus, keičiant matricos pradinę poziciją. Atlikus visą vaizdo skenavimą atliekamas gautu rezultatų apdorojimas, rasti plotai sumuojami ir ieškomą sutapimo bei persidengimo sričių. Sutankėjimo srityse yra didžiausias kiekis pozityviai klasei priskirtų matricos dalių. Susumavus rezultatus pozityvios dalys nufiltruojamos su slenkstine verte ir likę plotai įtalpinami į stačiakampius, bei atvaizduojami.

3.1.2. Metodo parengimas bei apmokymas

Ieškomi objektai yra skirtingi dydžiu, forma, spalva ir kitomis savybėmis. Todėl kiekvienam iš jų atliekamas parametrų derinimas, siekiant aptikti objektą. Derinimo metu nustatomos paieškos zonos, slenkančios matricos parametrai, poslinkio žingsnis ir persidengimo ploto slenkstinė riba. Visi šie parametrai įtakojo aptikimo tikslumą. Toliau pateikiami rezultatai gauti kamerasi esant 1 metro aukštyje. Kartono lapas panaudotas sumažinti šviesos atspindi, kuris galimai iškreipia aptikimo rezultatus. Apmokymui naudoti vaizdai iš interneto, bei padaryti tyrimo aplinkoje su kamera. Prieš naudojant vaizdus mokymui jiems pritaikoma augmentacija, taip padidinant vaizdų imtį. Po augmentacijos pakoreguojamas vaizdų dydis į 64x64 dydžio matricą. Klasifikatoriui reikalingas didelis vaizdų kiekis, kuriame yra tik mokomasis objektas. Kitas rinkinys su negatyviai vaizdais, tai vaizdų imtis kurioje nėra ieškomo objekto tik aplinkos vaizdai. Tiriamu atveju negatyvių vaizdų rinkinį sudaro medinių grindų, parketo, kilimo, kartono lakšto ir keletu kitu būdai būdingų vaizdų imtis. Klasifikatoriui apmokyti naudojamų vaizdų kiekis, apmokymo laikai bei SVM galutinis įsivertinimo procentas pateikiami 3.1 lentelėje.

3.1 lentelė. HOG ir SVM metodo apmokymo imtis bei rezultatai

| Ieškomas objektas | SVM apsimokymo tikslumas, % | Laikas, s | Mokymosi imtis |
|-------------------|-----------------------------|-----------|-----------------|
| Plastiko ritė | 94,886 | 31,852 | 2400pos/2200neg |
| Ventiliatorius | 95,754 | 29,754 | 2400pos/2200neg |
| Butelis | 96,126 | 32,457 | 2400pos/2200neg |
| Varžtas | 92,234 | 35,241 | 2400pos/2200neg |
| Replės | 95,187 | 33,147 | 2400pos/2200neg |

3.1.3. Metodo testavimas su vaizdų rinkiniais

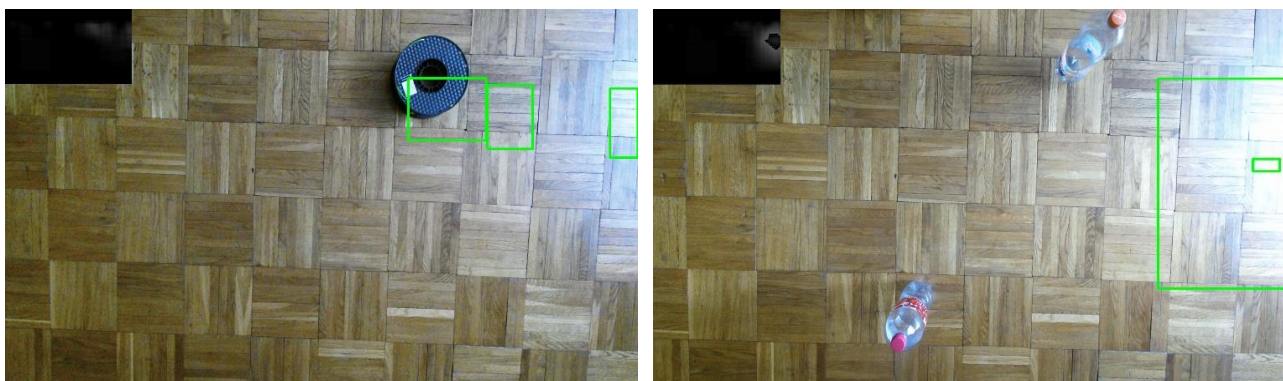
3.2 lentelė. HOG ir SVM testavimo rezultatai (natūrali šviesa ir tiriamieji objektai)

| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas | Aptiktų/visų objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|---------------------------|-----------------------------|--|
| Plastikas | 100 | 0,0783 | 0,825 | 0,781 | 24/174 | 13,5 |
| Butelis | 100 | 0,443 | 0,866 | 0,675 | 32/257 | 12,4 |
| Varžtas | 100 | 0,0481 | 0,954 | 0,421 | 10/174 | 5,7 |
| Ventiliatorius | 100 | 0,036 | 0,899 | 0,519 | 40/204 | 9,8 |
| Replės | 100 | 0,05 | 0,911 | 0,843 | 8/170 | 4,7 |

3.3 lentelė. HOG ir SVM testavimo rezultatai (dirbtinė šviesa ir šalutiniai objektai)

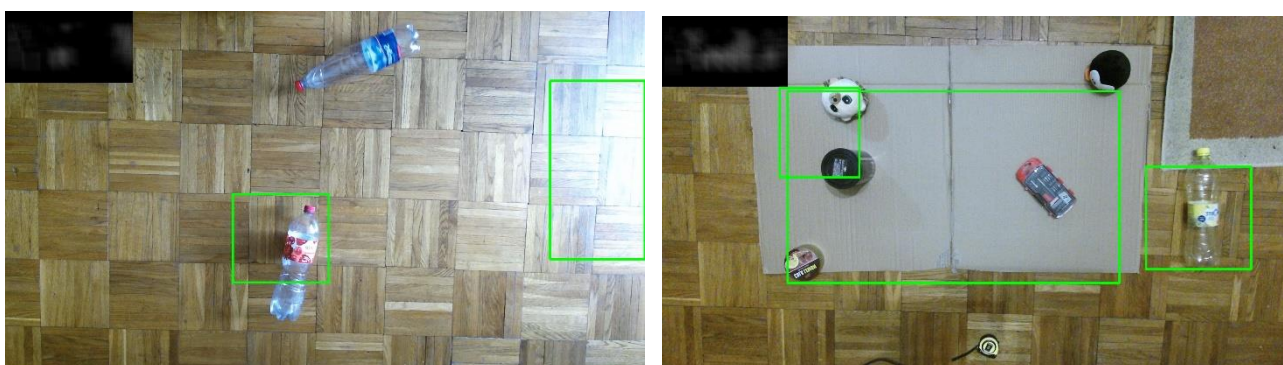
| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas, s | Aptiktų/visų objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|------------------------------|-----------------------------|--|
| Plastikas | 100 | 0,0347 | 0,895 | 0,824 | 10/174 | 5,7 |
| Butelis | 100 | 0,0124 | 0,926 | 0,756 | 13/257 | 5,0 |
| Varžtas | 100 | 0,0234 | 0,984 | 0,563 | 6/174 | 3,4 |
| Ventiliatorius | 100 | 0,072 | 0,954 | 0,682 | 9/204 | 4,4 |
| Replės | 100 | 0,024 | 0,987 | 0,921 | 5/170 | 2,9 |

Atlikus bandymus su tiriamais vaizdais pastebėta jog aptikimo tikslumą iškreipia atsiradę atspindžiai. Ryškias vietas metodas atpažįsta kaip ieškomą objektą tai matoma (3.2 pav.). Tokios rastos zonos sumažina metodo tikslumą. Šiuo būdu sunkiai sekėsi rasi objektus esant vien ieškomiems objektams. Prisdėjus šalutiniams objektas tikslumas dar labiau sumažėjo ir teko atlikti korekcijų nustatymuose.



3.2 pav. Natūralaus apšvietimo įtaka aptikimo metodui

Kaip matyti iš (3.3 pav.) vaizdo kairėje ir (3.2 pav.) nors ir pavyko rasti butelio ir plastiko ritės pozicijas vaizde, jų koordinatės apskaičiuotas gana netiksliai. Todėl vertinant objekto aptikimą jie bus atmesti, nes rastas plotas nepateks į IOU 0,5 vertės slenkstinį kriterijų. Kiti aptinkami vaizdai taip pat nebuvo pažymėti tiksliai. Vaizde (3.3 pav.) dešinėje matomi klaidingai pažymėti žali plotai, kuriuose nieko nėra. Tokie plotai mažina AP dydį ir didina MSE vertę.

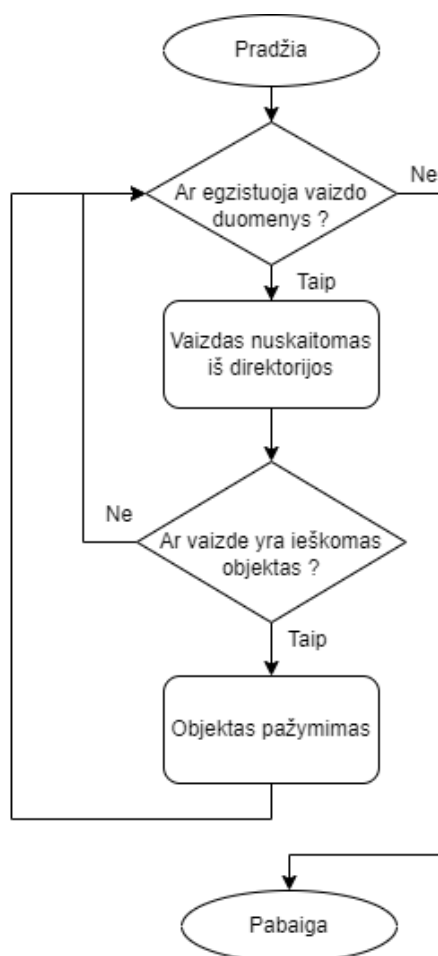


3.3 pav. Butelio paieškos rezultatai naudojantis HOG ir SVM metodu

3.2. Faster R-CNN

3.2.1. Metodo parengimas bei apmokymas

Faster R-CNN modelis yra palčiai taikomas, ir galima rasti daug pavydžių bei skirtingų struktūrų, kuriose reikia pakoreguoti tik pagrindinius parametrus. *Tensorflow* ir *Keras* pateikia išbandytą ir sukonfigūruotą Faster R-CNN metodą, kuris veikia patikimai ir yra ištestuotas su *COCO*. Turint tinklo konfigūraciją susidaroma mokymo imtis. Vaizdai atrenkami ir juose sužymimi objektai nurodant koordinatas bei objekto tipą. Visa informaciją apie rinkinį išsaugoma, prasidėjus apmokymo procedūrai remiantis šiais duomenimis bus apmokomas modelis. Pasiėkus norimą riba apmokymo procesas pertraukiamas ir modelis išsaugomas. Po kiekvienos epochos yra sugeneruojamas apmokyto modelio failas, taip išsaugant jau pasiektą progresą. Apmokymo spartą priklauso nuo turimos įrangos našumo. Naudojant grafinį procesorių apmokymas pagreitėja. Grafinis procesorius taip pat padeda sparčiau analizuoti vaizdus, nors jo startavimas užtrunka, dėl atminties atlaisvinimo. *NVIDIA* suteikia galimybę išnaudoti grafinio procesoriaus galimybes pasitelkus *CUDA* įrankį skirtą darbui su dideliais informacijos kiekiais. Derinantis prie Faster R-CNN bei *Tensorflow* programinės įrangos instaliuojama 10 programos versija. Paleidus duomenų analizatorių prasideda vaizdų svarių analizė ir priklausomai nuo epochų skaičiaus gaunama skirtinga galutinio modelio kokybė. Rekomenduojamas dydis yra 1000 epochų, tačiau jį apmokyti su naudojama įrangą užtruktų savaitę, todėl jis mažinamas iki 100. Iš pradžių mokymosi metu paklaida mažėja labai staigiai vėliau nusistovi ir mažėja labai lėtai arba praktiškai nemažėja. Mokymosi progresą galima sekti su *Tensorboard* įrankiu bei stebint paklaidos kitimą, po kiekvienos iteracijos.



3.4 pav. Faster R-CNN, YOLO, SSD metodų supaprastintas vaizdo analizės algoritmas

3.2.2. Metodo bandymai su iš anksto neparengtu modeliu

Naudojant Faster R-CNN metodą su prieš tai neapmokytu modeliu vidutinis tikslumas AP siekė 0,45, o vidutinė kvadratinė paklaida MSE 0,6. Tikslumas nebuvo didelis o, treniravimo laikas siekė 20 ir daugiau valandų. Atlikti bandymai pateikiami (3.5 pav.), kaip matyti pavyzdyje aptikti objektai žymimi gana netiksliai arba visai nerandami. Tokio modelio parengimas reikalauja našesnės įrangos, bei didesnių laiko sąnaudų, bet tai negarantuoja norimo tikslumo galutiniame rezultate. Egzistuoja jau dalinai apmokyti ir parengti modeliai, kurie geba greičiau bei efektyviau apdoroti informacija. Siekiant pagerinti metodo efektyvumą bei tikslumą buvo pasirinktas prieš tai parengtas ir išbandytas *Faster RCNN InceptionV2* modelis. Modeliui paruošti buvo naudota COCO vaizdų bazė duomenys.



3.5 pav. Faster R-CNN rezultatai su neaprengtu modeliu

3.2.3. Metodo bandymai su parengtu modeliu

Aptikimo rezultatus nulemia mokymosi laikas, vaizdų imtis ir modelio struktūra, kadangi metodo mokymosi progresą galima stebėti ir įvertinti, visiems tiriamiems objektams siekiama gauti paklaidos vertes artimas ar mažesnes už 0,1. Tinklo apmokymui naudojami vaizdai padaryti su kamera 1920x1080. Gautiems vaizdams pritaikoma augmentacija, kurios metu keičiamas apšvietimo lygis, pridedami triukšmai bei įvairūs vaizdu posūkio kampai. Pakeičiama vaizdų raiška į 640x480 taip paspartinant mokymosi procesą. Turint pilna vaizdų rinkinį pradedamas objektų žymėjimas. Kiekvienai nuotraukai sugeneruojami failai, kuriuose saugomos objekto koordinatės. Vėliau iš šių failų išrenkama tik naudinga informacija tai yra objekto koordinatės, pavadinimas bei atvaizdo pavadinimas. Remiantis šiais duomenimis sugeneruojamas tekstinis failas, kurį nusiskaito Faster R-CNN modelis, ir pradeda analizuoti bei dėti svorinius koeficientus. Mokymosi sparta priklauso nuo vaizdų raiškos bei kiekio. Baigus apmokymo procedūrą gaunamas apmokytas tinklas, kuriuo naudojantis ieškoma objektų.

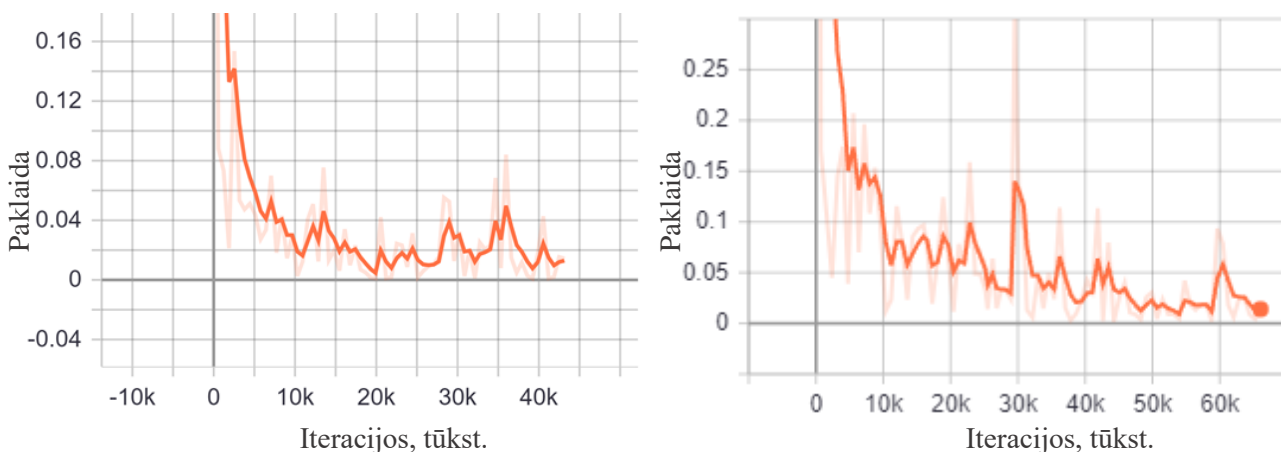


3.6 pav. Rezultatai naudojant parengta Faster R-CNN InceptionV2 modelį

Atlikus pirmus bandymus su plastiko rite matyti ženklus pagerėjimas tuose pačiuose vaizduose (3.6 pav.). Šiam rezultatui gauti užteko trumpesnio mokymosi laiko, todėl tolimesni bandymai atliekami naudojant *Faster R-CNN InceptionV2*. Apmokymui naudoti duomenys, bei gautos galutinės mokymo paklaidos pateikiamos 3.4 lentelėje. Iš lentelės matyti jog visiems objektams pamokyti reikia apytiksliai dviejų valandų. Vertei nusistovėjus mažiausiame lygyje ir siekiant modelio nepersotinti, mokymo procesas nutraukiamas. Paklaidos kitimas augant iteracijų skaičiui, kai mokoma aptikti plastiko ritę, pateikiamas (3.7 pav.) kairėje.

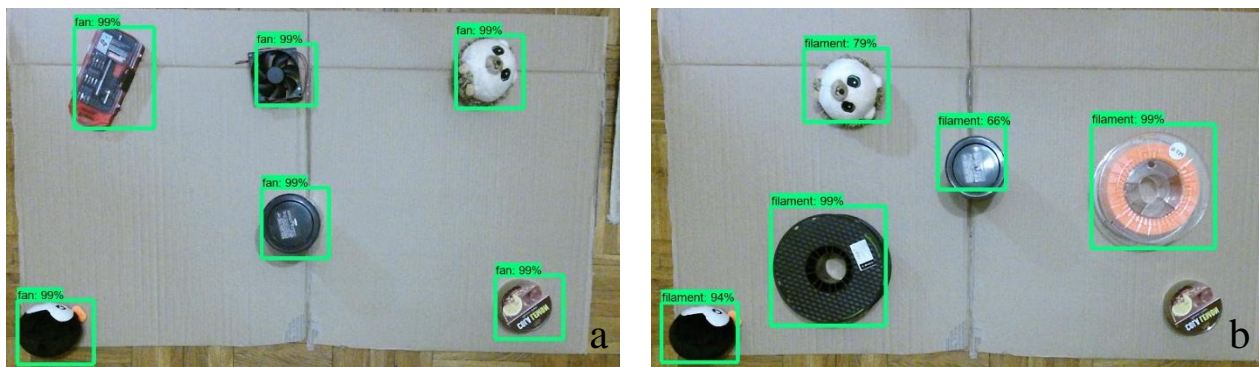
3.4 lentelė. Metodo apmokymo rezultatai

| Ieškomas objektas | Mokymosi paklaida | Mokymosi laikas | Mokymosi imtis | Iteracijų skaičius (tūkst.) |
|-------------------|-------------------|-----------------|----------------|-----------------------------|
| Plastiko ritė | 0,01436 | 2h13min14s | 600 | 43 |
| Ventiliatorius | 0.01525 | 1h57min48s | 600 | 46 |
| Butelis | 0,01050 | 2h19min54s | 600 | 54 |
| Varžtas | 0,01198 | 1h55min22s | 600 | 45 |
| Replės | 0,02878 | 2h1min49s | 600 | 45 |



3.7 pav. Modelio mokymosi regresijos pavyzdžiai, plastiko kairėje, bendras visų objektų dešinėje

Tolimesni bandymai parodė (3.8 pav.) jog modelis tampa neveiksmingas kai atsiranda šalutiniai objektai. Nors esant vienam bandomam objektui (2 Priedas 2.1 lentelė), aptikimo vidutinis tikslumas ir rastų objektų procentas buvo aukštas. Galimai modeliui per daug įsisotinus ar viršijus mokymosi ribą jis visus objektus atpažįsta kaip ieškomus.



3.8 pav. Klaidingai atpažįstami objektai

Kai naudojami tik vieni ieškomi objektai, tai sunku įvertinti modelio parengimą. Atsiradus papildomiems objektams matomas aiškus nesutapimas (3.8 pav. a) ieškant ventiliatorių, visi nuotraukoje matomi objektai buvo atpažinti kaip ventiliatoriai. Tokiu atveju imamas išsaugotas modelis, kuris turi mažiau iteracijų. Išbandžius modelius turinčius mažiau iteracijų rezultatai nepakito (3.8 pav. b). Todėl buvo priimtas sprendimas permokinti tinklą naudojant iškart visų penkių objektų vaizdus. Tokiu atveju būtų gaunama didesnė ir sudėtingesnė testavimo imtis, bei sumaišomi mokymo vaizdai. Apmokius modelį per 2h 49min 52s, gauta galutinė 0,014 vertės paklaida. Mokymo regresija matoma (3.7 pav.) dešinėje, iš jos galima pastebėti jog paklaida mažėjo lėčiau nei treniruojant tinklą vienam objektui. Atlikus bandymus rezultatai pagerėjo, kaip matyti (3.9 pav.) iš (3.10 pav.) lyginant su (3.8 pav.) atpažįstant objektus vis dar pasitaiko netikslumų bet modelis veikia patikimiau. Visi tolimesni bandymai atliekami su šiuo modeliu, gauti rezultatai matomi 3.5-3.6 lentelėse.

3.2.4. Metodo testavimas su vaizdų rinkiniais

3.5 lentelė. Testavimo rezultatai (natūrali šviesa ir tiriamieji objektai)

| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas (s) | Aptiktų/visų objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|-------------------------------|-----------------------------|--|
| Plastikas | 100 | 0,959 | 0,506 | 0,083 | 167/174 | 95,9 |
| Butelis | 100 | 0,990 | 0,682 | 0,096 | 257/257 | 100 |
| Varžtas | 100 | 0,297 | 0,748 | 0,089 | 111/195 | 56,9 |
| Ventiliatorius | 100 | 0,998 | 0,634 | 0,0917 | 204/204 | 100 |
| Replės | 100 | 0,996 | 0,558 | 0,095 | 165/170 | 97,0 |

3.6 lentelė. Testavimo rezultatai (dirbtinė šviesa ir šalutiniai objektai)

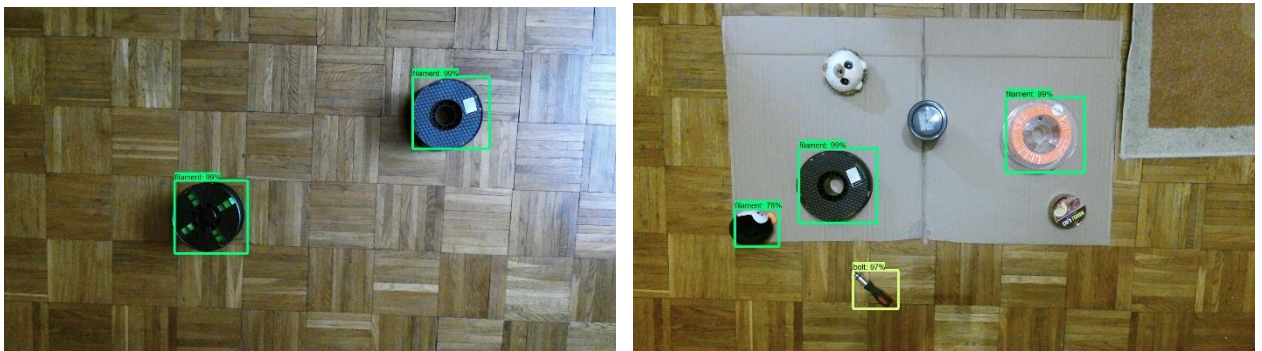
| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas | Aptiktų/visų objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|---------------------------|-----------------------------|--|
| Plastikas | 100 | 0,352 | 0,824 | 0,114 | 159/160 | 99,3 |
| Butelis | 100 | 0,263 | 0,797 | 0,112 | 137/140 | 97,8 |
| Varžtas | 100 | 0,177 | 0,892 | 0,127 | 142/214 | 66,3 |
| Ventiliatorius | 100 | 0,354 | 0,800 | 0,107 | 151/151 | 100 |
| Replės | 100 | 0,190 | 0,844 | 0,126 | 126/138 | 91,3 |

Atlikus testus su pirmu vaizdų rinkiniu, rezultatai matomi 3.5 lentelėje. Vaizdų rinkinys sudarytas esant natūraliai dienos šviesai bei nenaudojant šalutinių objektų. Iš lentelės matyt jog prasčiausiai sekėsi rasti varžtus, nes jie yra mažiausi objektai. Taip pat aptikimo metu juos painiojo su parketo lentelėmis. Didesni objektai kėlė mažiau problemų todėl vidutinis tikslumas AP buvo gana aukštas, tai yra didesnis už 0,9. Vidutinė kvadratinė paklaida MSE parodo kaip tiksliai pavyko nustatyti objekto poziciją. Kuo arčiau 0 vertės, tuo tiksliau nustatomos objekto koordinatės. Ventiliatorių grafoje matoma jos pavyko rasti visus esamus objektus, tačiau vidutinis tikslumas nėra lygus 1. Taip yra todėl, nes pasitaiko objekto aptikimų kai jo ten nėra.



3.9 pav. Faster R-CNN rezultatai kairėje natūrali šviesa, dešinėje dirbtinė ir šalutiniai objektai

Atsiradus šalutiniams objektams pakito aptikimo tikslumas, atsirado painiojamų objektų tai matyti ir (3.10 pav.). Lentelės 3.6 duomenų aptikimo rezultatai suprastėjo lyginant su 3.5 lentelės duomenimis. Sumažėjo vidutinis tikslumas, nes metodas kartais atpažindavo atsitiktinį objektą kaip ieškomą. Didelius objektus metodui aptikti sekasi geriau nei mažus, be to jiems aptikti užtenka trumpesnio laiko intervalo.



3.10 pav. Faster R-CNN rezultatai kairėje natūrali šviesa, dešinėje dirbtinė šviesa ir šalutiniai objektai

3.3. YOLOv5

3.3.1. Metodo parengimas bei apmokymas

YOLO metodas nuo savo pirmtako tobulėjo ir keitėsi. Dabar egzistuojanti naujausia YOLO v7 versija, darbe bus naudojama YOLO v5 versija. Ši versija yra stabili ir praktiškai išbandyta, be to turi realizavimo galimybes *Python* programavimo kalboje. YOLO metodui kaip ir Faster R-CNN reikalinga mokymosi imtis bei vaizdai su pažymėtais objektais. YOLOv5 versija apsimokymui bei aptikimui naudoja procesoriaus resursus. Ši versija yra skirta veikti modulinėse sistemose bei atskiruose įrenginiuose be galingo grafinio procesoriaus. Dėl to yra mažesnė ir pritaikoma daugelyje įrenginių neturinčių pakankamai vietos patalpinti didelį neuroninį tinklą. Apmokymui pasitelkiami tie patys duomenys bei objektų aprašai kaip ir Faster R-CNN. YOLO naudoja kitokia objektų pozicijos vaizduose nurodymo sistemą. Pradžioje nurodama pažymėto stačiakampio centras, po to stačiakampio ilgis bei plotis, santykiniais vienetais. Santykiai vienetai gaunami atitinkamai pagal koordinatų ašis stačiakampio ilgį pikseliais ar centrinę koordinatę dalinant iš vaizdo dydžio ta ašimi pikseliais. Analogiškai su kita ašimi statmenai šiai. Šiam metodui apmokyti sudaroma mokymo bei įsivertinimo vaizdu rinkiniai. Turimus duomenis rekomenduojama suskirstyti tokiu santykiu jog mokymo rinkinys būtų sudarytas iš 80% visų vaizdų, o įsivertinimo rinkinį sudarytų likę 20%. Pasirenkamas epochų skaičius, aprašomi mokomi objektai bei nustatomo mokomo modelio struktūra

ir pradedamas apmokymas. Kadangi šis metodas naudoja procesoriaus resursus, todėl apmokymas trunka ilgiau ir lėčiau mažėja paklaidos vertė. Tačiau galiniame rezultate aptikimo laikai bei tikslumas yra konkurencingi kitiems metodams, tuo bus galima įsitikinti tolimesniuose eksperimentų rezultatuose pateikiamuose 3.8 – 3.9 lentelėse.

3.7 lentelė. Mokymo rezultatai

| Ieškomas objektas | Apsimokymo įvertinimas (mAP) | Mokymosi laikas | Mokymosi imtis | Epochų skaičius |
|-------------------|------------------------------|-----------------|----------------|-----------------|
| Plastiko ritė | 0,812 | 1h45min32s | 600 | 5 |
| Ventiliatorius | 0,842 | 2h33min10s | 600 | 20 |
| Butelis | 0,894 | 2h39min18s | 600 | 20 |
| Varžtas | 0,791 | 2h40min20s | 600 | 20 |
| Replės | 0,856 | 2h45min10s | 600 | 20 |

Apmokius YOLO ir atlikus testus, plastiko ritė nekėlė bėdų po 5 epochų apmokymo. Norint rasti mažesnius objektus teko padidinti epochų skaičių. Atlikus bandymus su didesniu epochų skaičiumi pakeitimas pasiteisino todėl buvo likta prie didesnio epochų skaičiaus. Naudojamas epochų kiekis neleidžia modeliui įsisotinti ir randa tiek pavienius ieškomus objektus, tiek ieškomus objektus esančius šalia šalutinių objektų.

3.3.2. Metodo testavimo su vaizdų rinkiniais

3.8 lentelė. Testavimo rezultatai (natūrali šviesa ir tiriamieji objektai)

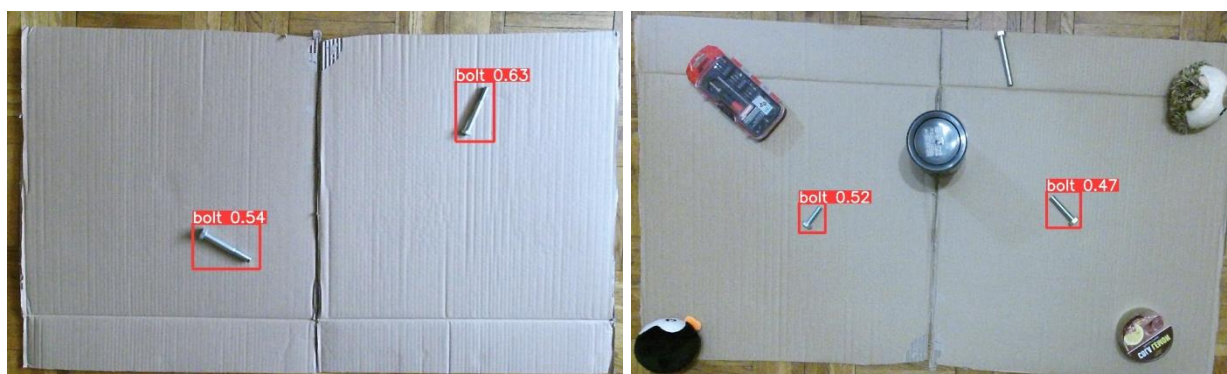
| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas | Aptiktų/visų objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|---------------------------|-----------------------------|--|
| Plastikas | 100 | 0,853 | 0,548 | 0,144 | 154/174 | 88,5 |
| Butelis | 100 | 0,665 | 0,667 | 0,1466 | 171/257 | 66,5 |
| Varžtas | 100 | 0,498 | 0,640 | 0,145 | 110/195 | 56,4 |
| Ventiliatorius | 100 | 0,881 | 0,612 | 0,153 | 180/204 | 88,2 |
| Replės | 100 | 0,829 | 0,548 | 0,1504 | 141/170 | 82,9 |

3.9 lentelė. Testavimo rezultatai (dirbtinė šviesa ir šalutiniai objektai)

| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas | Aptiktų/visų objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|---------------------------|-----------------------------|--|
| Plastikas | 100 | 0,715 | 0,670 | 0,1765 | 158/160 | 98,7 |
| Butelis | 100 | 0,107 | 0,516 | 0,1451 | 15/140 | 10,7 |
| Varžtas | 100 | 0,214 | 0,686 | 0,1456 | 46/214 | 21,4 |
| Ventiliatorius | 100 | 0,184 | 0,677 | 0,1917 | 59/151 | 39,0 |
| Replės | 100 | 0,366 | 0,473 | 0,2435 | 73/138 | 39,8 |

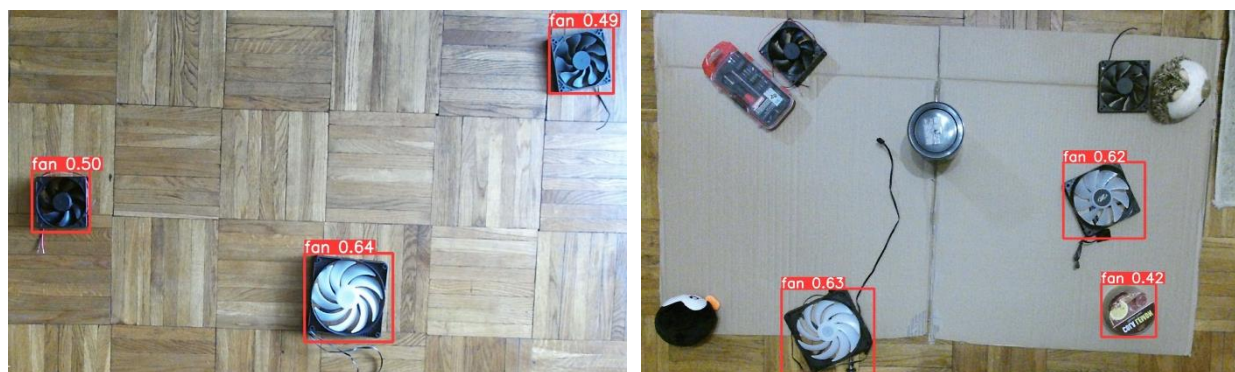
Po testų gautos 3.8 ir 3.9 lentelės su rezultatais, YOLO geriau veikė su didesniais objektais, tai rodo didesnės vidutinio tikslumo vertės AP bei mažesnės vidutinės kvadratinės paklaidos vertės MSE.

Atliekant bandymus kai naudojami šalutiniai objektai, mažiausią įtaką tai turėjo plastiko ritėms, remiantis 3.9 lentele. Viename iš pateiktų vaizdu (3.12 pav.) ieškomas objektas tai yra fenas sumaišomas su saldainių dėžute. Tai pat ne visada randami visi objektai kaip matoma (3.12 pav.) ieškant ventiliatorių ir (3.11 pav.) bandant rasti vaizde esančius varžtus.



3.11 pav. YOLO rezultatai kairėje natūrali šviesa, dešinėje dirbtinė ir šalutiniai objektai

Mažiausias rastų objektų procentas tendencingai sutampa su objektų dydžių, jam mažėjant sunkiau atlikti indentifikavimą. Panašias tendencijas išlaiko ir aptikimo laikai, didesni objektai randami sparčiausiai, mažesni savo dimensijomis lėčiau. Remiantis pateiktais vaizdais (3.11 pav.) ir (3.12 pav.) galima teigti jog metodas gana tiksliai pažymi objektus, neapimdamas didelio papildomo ploto, tai parodo ir lentelėse 3.8 ir 3.9 matomos žemos vidutinės kvadratinės paklaidos vertės.



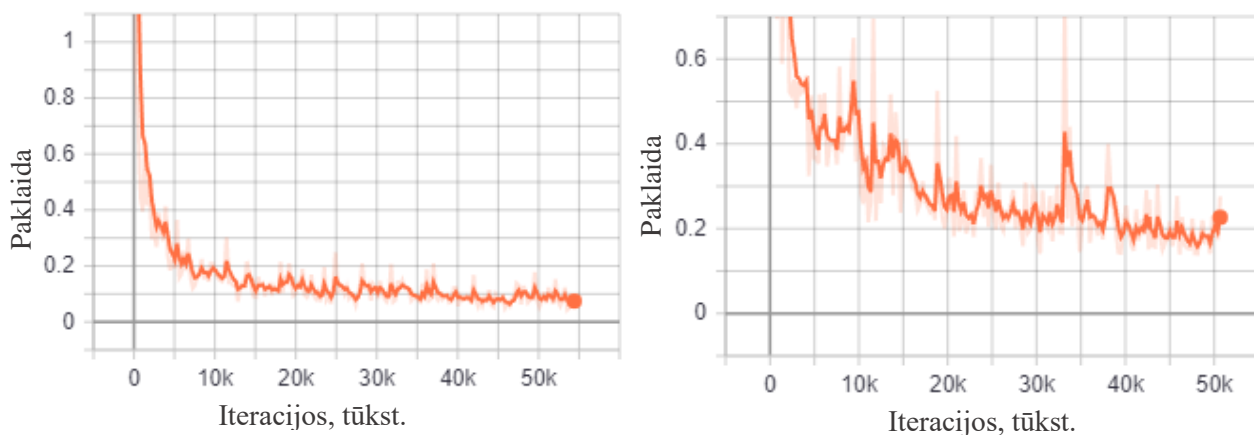
3.12 pav. YOLO rezultatai kairėje natūrali šviesa, dešinėje dirbtinė ir šalutiniai objektai

3.4. SSD

3.4.1. Metodo parengimas bei apmokymas

SSD – vieno kadro detektorius išsiskiria savo greitaveika. Šis metodas savo veikimo principu yra panašus į YOLO tačiau naudojami grafiniu procesoriumi, bei turi kitokią tinklo struktūrą. Tyrimo metu naudojamas ne standartinis SSD, o jau apmokytas ir turintis pakoreguotą tinklo struktūrą modelis – *SSD MobilnetV1*. Metodui realizuoti pasitelkiami *Tensorflow* modelių mokymo aplinka, naudota kartu su Faster R-CNN. Modeliui sukeliama reikalingi vaizdai bei failai su objektų koordinatėmis. Vaizdai padalinami į mokymo ir testavimo imtis. Sugeneravus reikalingus failus su objektų pozicijomis bei įvedus reikiamus parametrus, kurie nurodo mokymo imties charakteristikas, pradedamas modelio treniravimas. Mokymosi paklaidos kitimą grafiškai pateikiamas *Tensorboard* lange, kur jį galima sekti realiu laiku. Šio modelio mokymosi paklaida mažėja lėčiausiai. Todėl pailgėja ir mokymo laikas, nors iš pradžių paklaida per 5 tūkstančius iteracijų pasiekia, apytiksliai

0,4 vertę toliau kinta nežymiai. Visi objektai buvo mokomi apie 6 valandas, ir apie 50 tūkstančių iteracijų. Žemiau pateikiami keletas mokymosi eiga atvaizduojančių grafikų (3.13 pav.). Vaizdžiai matyti jog ne visiems objektams mokymasis vyksta vienodai todėl ir paklaidos vertė laike kinta skirtingai. Plastiko ritė yra didžiausias tiriamas ir išskirtinis forma bandinys, butelis nepastovios formos ir blizgus objektas. Sprendžiant iš 3.10 lentelės prasčiausiai sekėsi įsisavinti replių savybes, nes paklaidos vertė išliko didžiausia nors mokymas užėmė ilgiausią laiką. Mokymas buvo nutrauktas nes paklaidos vertės nustojo mažėti.



3.13 pav. Paklaidos kitimas augant mokymo iteracijoms: kairėje plastiko ritė, dešinėje butelis

3.10 lentelė. Metodo rengimo rezultatai

| Ieškomas objektas | Paklaida | Mokymosi laikas | Mokymosi imtis | Iteracijų skaičius (tūkst.) |
|-------------------|----------|-----------------|----------------|-----------------------------|
| Plastiko ritė | 0,0362 | 6h9min47s | 600 | 52 |
| Ventiliatorius | 0,0839 | 5h13min45s | 600 | 44 |
| Butelis | 0,0745 | 6h31min45s | 600 | 54 |
| Varžtas | 0,2761 | 6h7min48s | 600 | 50 |
| Replės | 0,3254 | 6h45min7s | 600 | 56 |

3.4.2. Metodo testavimas su vaizdų rinkiniais

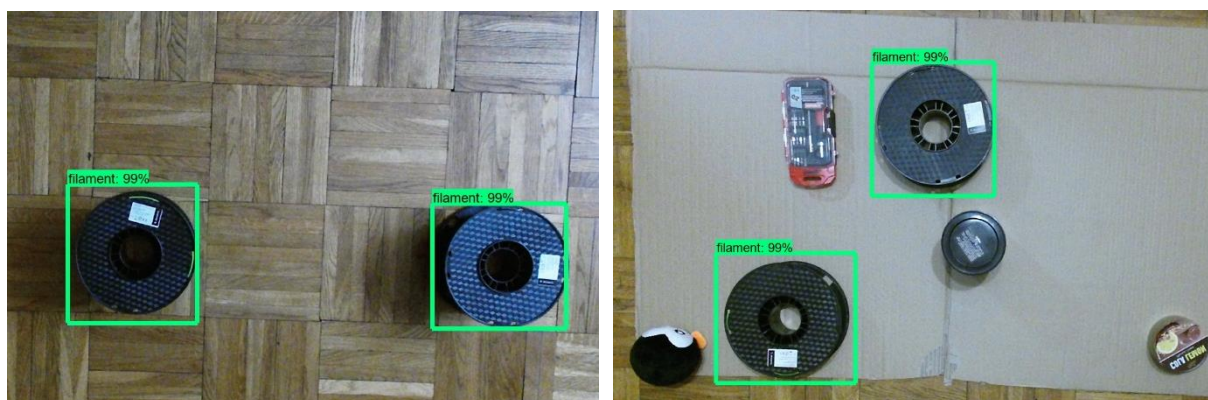
3.11 lentelė. Testavimo rezultatai (natūrali šviesa ir tiriamieji objektai)

| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas | Nerastų/rastu objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|---------------------------|------------------------------|--|
| Plastikas | 100 | 0,781 | 0,4945 | 0,0654 | 136/174 | 78,16 |
| Butelis | 100 | 0,537 | 0,690 | 0,0734 | 154/257 | 59,9 |
| Varžtas | 100 | 0,502 | 0,709 | 0,0701 | 129/195 | 66,1 |
| Ventiliatorius | 100 | 0,876 | 0,627 | 0,0855 | 179/204 | 87,7 |
| Replės | 100 | 0,300 | 0,560 | 0,6741 | 51/170 | 30,0 |

3.12 lentelė. Testavimo rezultatai (dirbtinė šviesa ir šalutiniai objektai)

| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas | Nerastų/rastu objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|---------------------------|------------------------------|--|
| Plastikas | 100 | 0,7999 | 0,470 | 0,0595 | 128/160 | 80 |
| Butelis | 100 | 0,098 | 0,760 | 0,0726 | 58/154 | 41,4 |
| Varžtas | 100 | 0,231 | 0,844 | 0,0724 | 92/214 | 23,1 |
| Ventiliatorius | 100 | 0,235 | 0,798 | 0,0753 | 108/151 | 71,5 |
| Replės | 100 | 0,059 | 0,852 | 0,0788 | 49/138 | 35,5 |

Atlikus bandymus matoma didelė aptikimo sparta bei tikslumas. Problemų kėlė tiriami vaizdai su buteliais ir varžtais tai rodo 3.11 lentelėje žemos vidutinio tikslumo vertės bei mažas aptikimo procentas apie 60. Atsiradus papildomiems objektams sumažėjo vidutinis tikslumas, bei randamų objektų procentas 3.12 lentelė. Atsiradę papildomi objektai labiausiai turėjo įtakos ventiliatorių identifikavimui. Taip galėjo atsitikti todėl jog pašaliniai objektai yra panašių dydžių ir turi panašių formų, kurios klaidina klasifikatorių. Prasčiausius rezultatus parodė replės, prisidėjus papildomiems objektams vidutinis tikslumas nukrito iki 0,059, o vidutinė kvadratinė paklaida išaugo iki 0,85 tai reiškia jog buvo daug klaidingai atpažintų objektų. Vaizde rasti plastiko rites nesutrukdė nei pašaliniai objektai nei natūrali šviesa, todėl abejuose bandymuose išlaikytos panašios vidutinio tikslumo vertės apie 0,8. Atsiradus papildomiems objektams tikslumas netgi padidėjo Taip galėjo atsitikti dėl pašalintos natūralios šviesos ir blizgesių.

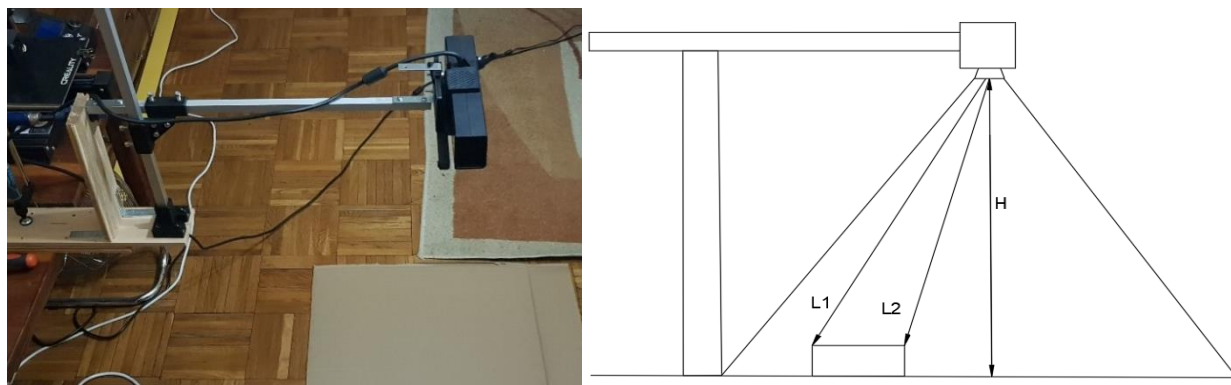


3.14 pav. Plastiko ričių aptikimo rezultatai kairėje natūrali šviesa, dešinėje dirbtinė šviesa ir šalutiniai objektai

4. Objekto padėties radimas naudojant Kinect kamerą

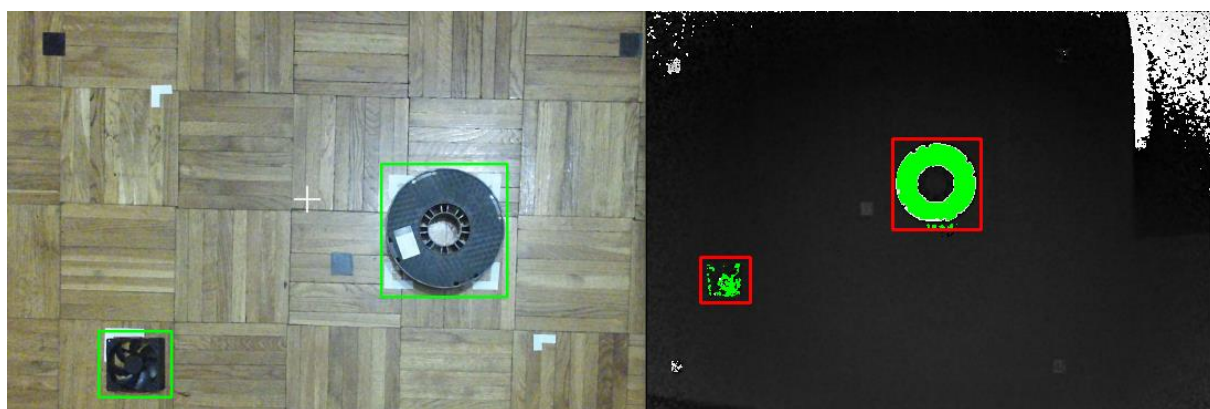
4.1. Objektų pozicijos radimas

Kamera įtvirtina $H=1$ metro (4.1 pav.) aukštyje lygiagrečiai grindimis. Objektai matomi iš viršaus, RGB bei IR kameromis, toks atstumas neiškreipia nuskaitytųjų gylio duomenų, nes patenka į kameros rekomenduojama veikimo diapazoną. Objektų pozicijoms rasti pasitelkiami du skirtingi būdai. Vienas paremtas RGB vaizdu ir proporcijomis kitas gylio kameros duomenimis. Objektų pozicijos aprašomos žinomų taškų atžvilgiu, tiksliau distancijų pokyčiu nuo žinomų taškų. Metodai išbandomi praktiškai, keičiant taškų pozicijas kadre. Gautiems rezultatams paskaičiuojamos vidutinės absoliutinės paklaidos.



4.1 pav. Bandymų sistema

Kameros programinėje įrangoje realizuoti metodai, kurie geba RGB kadro poziciją pikseliais paversti į gylio kadro poziciją pikseliais (2.2 pav.). Gavus koordinates gylio vaizde, apskaičiuojamos gaubiančio stačiakampio ribos, kuriame turi būti objektas. Gavus gylio duomenų rinkinį iš apibrėžto ploto galima apskaičiuoti vidutinę atstumo vertę ir nufiltruoti toliau už vidurkį esantį nedominantį pagrindą. Atlikus tokią operaciją su realiais vaizdais, rezultatas matomas (4.2 pav.). Vaizde kairėje matomi rasti objektai plastiko ritė ir ventiliatorius pažymėti žaliais stačiakampiais. Tuos pačius stačiakampius dešinėje atitinka raudoni stačiakampiai. Nufiltravus duomenis pagal vidutines gylio vertes, rasti objektai pažymėti žalia spalva ir matomi (4.2 pav.) dešinėje. Pasinaudojus gylio kamera išgaunama tiksli objekto poziciją ir forma kadre.



4.2 pav. Objektų pozicijos patikslinimas naudojantis gylio kameros duomenimis

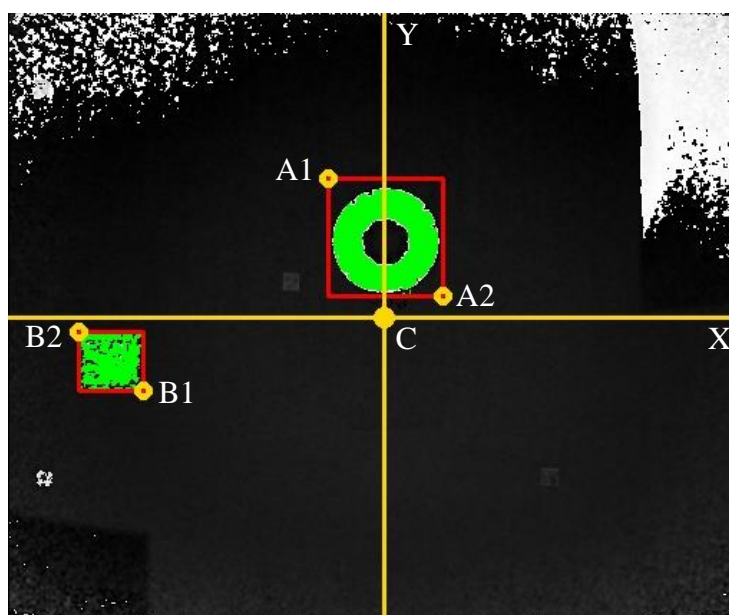
4.2. Objektų koordinatčių nustatymas naudojant gylio duomenimis

Aptikus objektą kadre, gaunamos jį gaubiančių stačiakampių kraštinės koordinatės, pikselių matricioje. Kameros atstumas iki pagrindo yra žinomas, daroma prielaida, kad kamera įtvirtina 1 metro aukštyje ir yra lygiagrečiai pagrindui. Rasti taškai RGB vaizde konvertuojami į koordinates gylio vaizde, realizuota operacija matoma (4.3 pav.). Kadangi skaičiavimams naudojami gylio kameros duomenys, nereikia papildomų markerių bei atžymų. Taip pat atsiranda galimybė pasakyti objekto koordinatės pokytį, pagrindui statmena ašimi. Objektų pozicijos bus paskaičiuotos kameros centro atžvilgiu pažymėto (4.3 pav.) C raide. Priimama, kad kameros centre kertasi dvi statmenos koordinatčių ašys X ir Y . Objektų pozicijai rasti pasitelkiama tangento funkcija (4.2). Kadro centrinėje koordinatėje 256×212 , kuri yra artimiausia lęšio centrui egzistuoja mažiausias iškraipymas, nes per ją eina lęšio optinis centras. Šioje pozicijoje gautas atstumas iki pagrindo teoriškai turėtų būti mažiausias. Skaičiavimui pasitelkiami kameros matymo kampai. Gylio kamera turi 70×60 laipsnių (žr. 1.1 lentelė) matymo lauką, kuris atitinka 512×424 vaizdo matricią. Vienam pikseliui apytikriai tenka $0,1367$ laipsnio pokytis horizontalia X ašimi, bei $0,1415$ laipsnio pokytis vertikalia Y ašimi. Turint šiuos duomenis galima gauti kampų vertes, pagal pikselių koordinates. Taško pozicija randama, atstumais lygiagrečiais X ir Y ašims, nuo taško C iki ieškomo taško pvz. A_1 . Pradžioje nuskaitomi gylio kadro duomenys taške C ir gaunamas atstumas iki pagrindo H_k . Imant pikselių pokytį X_{A1C} gaunamas Δx_p , kampas α_x apskaičiuojamas pagal (4.2). Panaudojus tangento funkciją (4.1) gaunamas realus atstumas X_{A1C} , lygiagretus X ašiai. Tokia pati operacija pakartojama poslinkiui Y ašimi rasti. Imamas pikselių atstumas, Y_{A1C} gaunamas Δy_p , kampas α_y apskaičiuojamas pagal (4.3). Naudojantis atstumu H_k (4.1) formule gaunamas realus atstumas Y_{A1C} , lygiagretus Y ašiai. Keletas gautų rezultatų pateikiama 4.1 lentelėje, taškai matomi (4.3 pav.).

$$l_{dist} = \tan(\alpha)H_k \quad (4.1)$$

$$\alpha_x = 0,1367 \Delta x_p \quad (4.2)$$

$$\alpha_y = 0,1415 \Delta y_p \quad (4.3)$$



4.3 pav. Gylio kameros matomas vaizdas su atskaitos bei ieškomais taškais

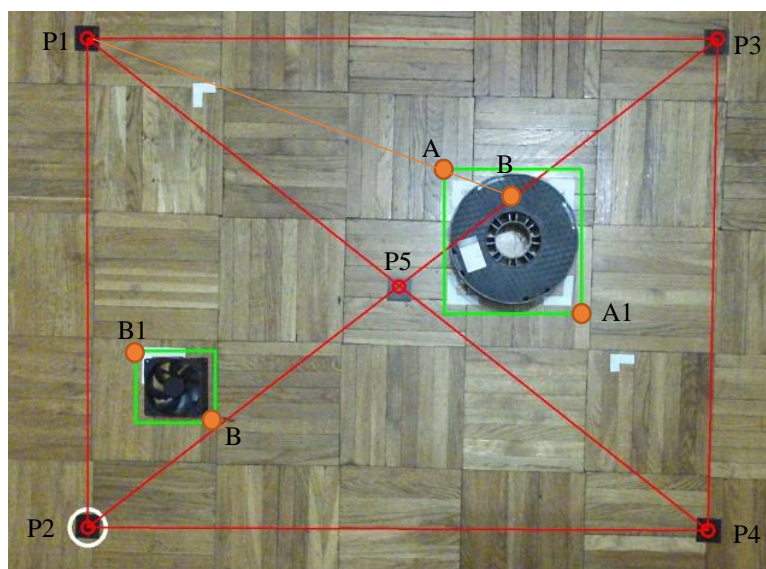
4.1 lentelė. Išmatuotų ir apskaičiuotų pozicijų vertės

| | Apskaičiuotas (mm) | Išmatuotas(mm) | | Apskaičiuotas (mm) | Išmatuotas (mm) |
|-----------|--------------------|----------------|-----------|--------------------|-----------------|
| X_{A1C} | -99 | -105 | X_{B1C} | -443 | -451 |
| Y_{A1C} | 254 | 260 | Y_{B1C} | -132 | -124 |
| X_{A2C} | 110 | 104 | X_{B2C} | -581 | -575 |
| Y_{A2C} | 39 | 44 | Y_{B2C} | -28 | -35 |

4.3. Objektų koordinatinių nustatymas naudojantis RGB vaizdu

Rastas objektų koordinatas konvertuoti į realius matmenis naudojami RGB kameros duomenys ir atskaitos taškai. Skaičiavimas paremtas vaizde matomų taškų P_1-P_5 pozicijomis, taškai naudojami kaip atskaitos sistema. Skaičiavimo būdas paremtas proporcijų metodu, kur pradžioje gauti koeficientai naudojami vėliau skaičiuojant objektų pozicijas. Toliau aprašoma skaičiavimų eiga objekto pozicijai perskaičiuoti iš vaizdo koordinatinių pikseliais į realius matmenis. Kaip atskaitos taškas matavimams naudojamas vaizde (4.4 pav.) baltu markeriu pažymėtas taškas P_2 . Atitinkamai jį kertančios linijos P_1-P_2 ir P_2-P_4 priimamos kaip statmenos ašys, jų vertė didėja tolstant nuo pažymėto taško.

Keturiems pradiniais taškams P_1, P_2, P_3, P_4 nustatomi tarpusavio atstumai ir pažymimi markeriais vaizde. Paskutinis penktas taškas randamas pasinaudojus formulėmis (4.4 – 4.7). Jis turėtų atsirasti tiesių P_{14} ir P_{23} sankirtoje. Pradžioje randami tiesių koeficientai $a_{P_1P_4}$, $a_{P_2P_3}$ ir $b_{P_1P_4}$, $b_{P_2P_3}$, su formulėmis (4.4) ir (4.5). Sankirtos taško koordinatės randamos gavus rezultatus iš (4.6) ir (4.7) formulių. Šie skaičiavimai atliekami tiek su realiomis tiek su pikselių pozicijoms randat tašką P_5 .



4.4 pav. Skaičiavimui naudojamų taškų bei ieškomų pozicijų schema

$$a = \frac{y_1 - y_0}{x_1 - x_0} \quad (4.4)$$

$$b = y_1 - ax_1 \quad (4.5)$$

$$P = \frac{b_{P_1P_4} - b_{P_2P_3}}{a_{P_2P_3} - a_{P_1P_4}} \quad (4.6)$$

$$P_{5y} = a_{P_2P_3} \cdot P_{5x} + b_{P_2P_3} \quad (4.7)$$

Sekančiame žingsnyje reikia nustatyti kuriame iš keturių trikampių yra ieškomas taškas A. Tikrinant ar taškas priklauso trikampiui reikia apskaičiuoti santykinės vertės ω_1 ir ω_2 pasinaudojus (4.8) ir (4.9) formulėmis. Jei gautos vertės tenkina sąlygas aprašytas (4.10) nelygybėse, reiškia, kad taškas priklauso tam trikampiui.

$$\omega_2 = \frac{A_y - P_{3y} - \omega_1 \cdot (P_{1y} - P_{3y})}{P_{5y} - P_{3y}} \quad (4.8)$$

$$\omega_1 = \frac{P_{3x}(P_{5y} - P_{3y}) + (A_y - P_{3y})(P_{5x} - P_{3x}) - A_x(P_{5y} - P_{3y})}{(P_{1y} - P_{3y})(P_{5x} - P_{3x}) - (P_{1x} - P_{3x})(P_{5y} - P_{3y})} \quad (4.9)$$

$$\omega_1 \geq 0 ; \omega_2 \geq 0 ; (\omega_1 + \omega_2) \leq 1 \quad (4.10)$$

Radus kuriam trikampiui priklauso taškas A, randama jam artimiausia trikampo kraštinė. Artimiausia kraštinė leidžia gauti tikslesnius rezultatus tolimesniuose skaičiavimuose. Radus artimiausią kraštinę, taškui A, nubrėžiama tiesė. Tiesė brėžiama iš taško P_1 per tašką A iki tiesės P_3P_5 kol kertasi taške B. Pasinaudojus (4.11) ir (4.12) formulėmis randami atkarpų santykiai λ_1 ir λ_2 . Toks pat tiesių padalinimo santykis egzistuoja ir realybėje. Skaičiuojant tašką A skaičiavimai vykdomi atbuline tvarka, pirmiausia randamas taškas B, jo pozicija gaunama iš (4.13) ir (4.14) formuliu.

$$\lambda_1 = \frac{X_A - X_{P_1}}{X_B - X_A} \quad (4.11)$$

$$\lambda_2 = \frac{X_B - X_{P_3}}{X_{P_5} - X_B} \quad (4.12)$$

$$X_B = \frac{X_{P_3} + X_{P_5} \cdot \lambda_2}{1 + \lambda_2} \quad (4.13)$$

$$Y_B = \frac{3 + Y_{P_5} \cdot \lambda_2}{1 + \lambda_2} \quad (4.14)$$

Radus B taško poziciją galima apskaičiuoti taško A koordinates, pasinaudojus menama tiesia. Tiesė kaip parodyta vaizde išeina iš taško B kerta tašką A ir eina link P_1 . Taškas A dalina šia tiesia λ_1 santykiniu koeficientu. Pasinaudojus formulėmis (4.15) ir (4.16) galima apskaičiuoti taško A pozicija realybėje.

$$X_A = \frac{X_{P3} + B\lambda_1}{1 + \lambda_1} \quad (4.15)$$

$$Y_A = \frac{Y_{P3} + B\lambda_1}{1 + \lambda_1} \quad (4.16)$$

Galutiniame rezultate gaunamos objekto koordinatės, konvertuotas iš pikselių vaizde į pasirinktą atskaitos sistemą. Keletas apskaičiuotų pozicijų pateikiama 4.2 lentelėje, realus taškai matomi (4.4 pav.).

4.2 lentelė. Išmatuotų ir apskaičiuotų pozicijų vertės

| | Apskaičiuotas (mm) | Išmatuotas(mm) | | Apskaičiuotas (mm) | Išmatuotas (mm) |
|------------|--------------------|----------------|------------|--------------------|-----------------|
| X_{AP2} | 584 | 590 | X_{BP2} | 201 | 206 |
| Y_{AP2} | 279 | 276 | Y_{BP2} | 162 | 168 |
| X_{AIP2} | 788 | 782 | X_{BIP2} | 85 | 89 |
| Y_{AIP2} | 495 | 502 | Y_{BIP2} | 225 | 223 |

4.4. Objektų padėties nustatymo metodų palyginimas

Buvo ištestuoti du skirtingi metodai, skirti konvertuoti turimą objekto poziciją vaizde pikseliais į realius matmenis. Pirmas aprašytas būdas remiasi kameros specifikacijomis bei gylgio kameros duomenimis ir nereikalauja jokių papildomų žymeklių ar stabilių taškų. Antras bandytas būdas naudoja markerius su žinomais atstumais tarp jų ir remiasi proporcijų skaičiavimu. Atlikus bandymus su 30 taškų, kiekvienam metodui, apskaičiuotos vidutinės paklaidos vertės – MAE pagal (4.17) formulę. Naudojant gylgio kamera paremtą metodą MAE gaunama 6,321mm, o su RGB vaizdu ir proporcijomis 7,875mm. Gylgio kameroje paklaidos didėja artėjant link kadro kampų, o su RGB kamera kadro kraštuose tolstant nuo markerių. Matavimai atlikti rankiniu būdu naudojant liniuotę ir kampainį, todėl dalį matavimo paklaidos sudaro žmogiškasis faktorius. Naudojant gylgio kamerą nereikia papildomų žymeklių ir galima išmatuoti objekto aukštį.

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x| \quad (4.17)$$

čia n – vertinamų paklaidų skaičius, x_i – prognozuojama vertė, x – reali vertė.

5. Aptikimo metodų rezultatų palyginimas

Šiame skyriuje apibendrinami ir palyginami gauti tyrimo rezultatai. Apžvelgiamos matomos tendencijos bei vienetiniai atvejai. Palyginimo tikslas nustatyti geriausiai tinkamą taikyti metodą su bandymų metu naudota įranga (2.1 – 2.2 lentelės) bei tiriamais objektais (2.3 lentelė). Žemiau pateikiamose lentelėse bei grafikuose matoma bendra rezultatų suvestinė.

5.1 lentelė. Suminė lentelė (natūrali šviesa ir tiriamieji objektai)

| Metodas | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas (s) | Aptiktų objektų, % |
|--------------|--------------------------|------------------------------------|-------------------------------|--------------------|
| HOG SVM | 0,1311 | 0,8910 | 0,6478 | 9,22 |
| Faster R-CNN | 0,8480 | 0,6256 | 0,0917 | 89,96 |
| YOLO v5 | 0,7452 | 0,6030 | 0,1478 | 76,5 |
| SSD | 0,5992 | 0,6161 | 0,0723 | 64,37 |

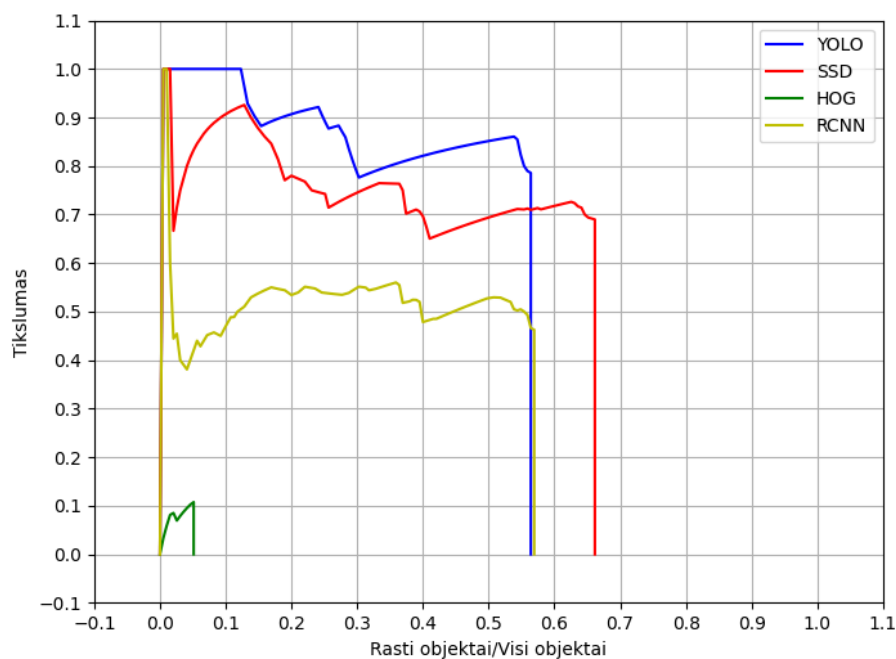
5.2 lentelė. Suminė lentelė (dirbtinė šviesa ir šalutiniai objektai)

| Metodas | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas | Aptiktų objektų, % |
|--------------|--------------------------|------------------------------------|---------------------------|--------------------|
| HOG SVM | 0,0333 | 0,9492 | 0,7492 | 4,28 |
| Faster R-CNN | 0,2672 | 0,8314 | 0,1172 | 90,94 |
| YOLO | 0,3172 | 0,6044 | 0,1805 | 41,92 |
| SSD | 0,2845 | 0,7448 | 0,0712 | 50,30 |

Lentelėse 5.1 ir 5.2 pateikiami rezultatų vidurkiai, jie gauti susumavus visų objektų rezultatus vienam metodui ir padalinus iš bandytų aptikti objektų grupių skaičiaus. Pirmiausia apžvelgiama 5.1 lentelė, kuri gauta nenaudojant šalutinių objektų ir esant natūraliam apšvietimui. Geriausius rezultatus parodė Faster R-CNN, vidutinio tikslumo vertė 0,84, o tai reiškia jog buvai mažai netinkamai klasifikuotų objektų. Aptiktų objektų procentas 89,96, o atpažinimo laikai nusileidžia tik SSD metodui, skirtumas apytiksliai 0,02 sekundės. Jam artimiausias yra YOLO, vidutinio tikslumo vertė siekė 0,74, o aptikimo procentas 76,5. Šis metodas naudojami procesoriaus resursais todėl aptikimo laikų vertės vidutiniškai nusistovi ties 0,15 sekundės. Tačiau lyginant su Faster R-CNN modeliu dydžiu ir efektyvumu tai, geras rezultatas ir taikytinas praktikoje neturint galimybės naudoti galingo grafinio procesoriaus. HOG SVM parodė prasčiausius rezultatus, vidutinis tikslumas siekia 0,13, aptikimo procentas 9 bei pats ilgiausias aptikimo laikas 0,64 sekundės. Būtent šiam metodui labiausiai kenkė saulės šviesa, bei atsiradę atspindžiai, kurie buvo klaidingai identifikuojami 1 Priedas (1.1 pav.). Nors šis metodas apsimoko sąlyginai greitai maždaug per 30 sekundžių (žr. 3.1 lentelę). Kaip geresnę alternatyvą jam galima naudoti YOLO, nes šis metodas taip pat naudoja procesoriaus resursus ir turi suspaustą modelį skirta mobilioms sistemoms.

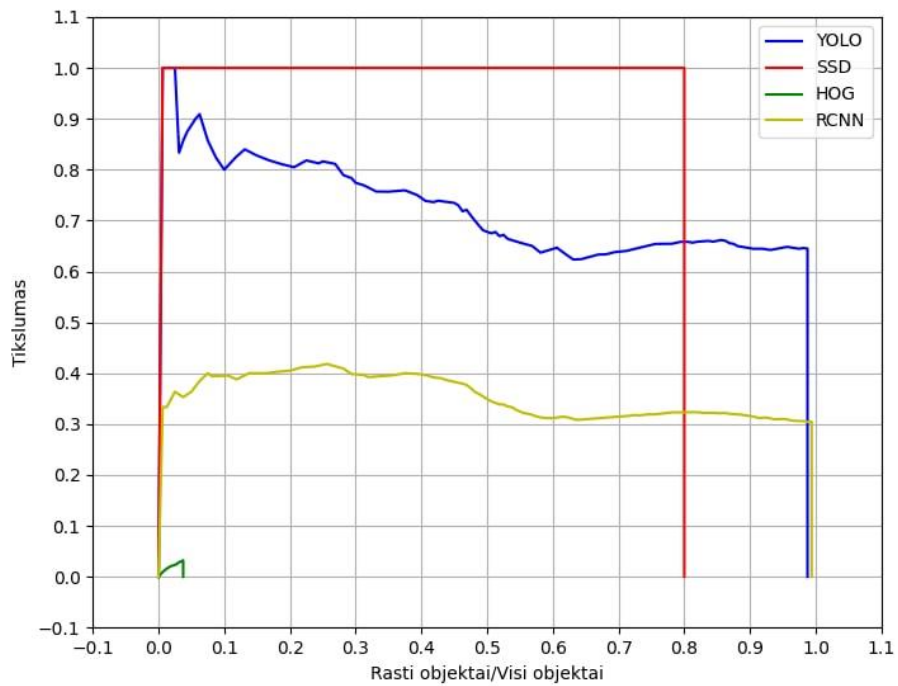
Atsiradus pašaliniais objektams aptikimo vidutinis tikslumas bei aptikimo procentas sumažėjo visuose bandytuose metoduose. Išaugo aptikimo laikai, bei vidutinės kvadratinės paklaidos. Mažiausią įtaką atsiradę pašaliniai objektai turėjo YOLO bei SSD metodams. YOLO vidutinis

tikslumas sumažėjo apie 0,43, o SSD 0,3 lyginant su bandymais be pašalinių objektų. Pagal sėkmingai aptiktų objektų vertes matyti jog geriausius rezultatus parodė Faster R-CNN 90,94 procentai, tačiau iš vidutinės kvadratinės paklaidos 0,83 bei vidutinio tikslumo vertės 0,26 galima spręsti jog buvo daug netinkamai identifikuotų objektų. Metodas nėra patikimas, nes pašalinius objektus priskyrė ieškomiems. Tai galima matyti ir 1 Priede (1.4 pav.) pateikiamose nuotraukose. Vaizdžiai tai pastebima (5.2 pav.) grafike, kur matoma jog Faster R-CNN tikslumas yra vienas mažiausių. Prasčiausiai aptinkančiu objektus metodu išliko HOG SVM, jo aptikimo vertės nesiekia nei 10 procentų, o vidutinė kvadratinė paklaida artėja į vienatę.



5.1 pav. Varžtų aptikimo bendras grafikas (natūrali šviesa ir ieškomi objektai)

Gauti grafikai atvaizduoja aptikimo tikslumo vertės kitimą, didėjant apdorotų vaizdų skaičiui. Grafike matomi, po kreivėmis esantys figūrų plotai, atitinka anksčiau pateiktose lentelėse esančias vidutinio tikslumo vertes. Jei objektai randami tiksliai ir dideliu procentu grafikas panašėja į stačiakampį, tokį grafika galima matyti 3 Priede esančiame (3.1 pav.) ir (5.2 pav.) stebint SSD metodo rezultatus. Iš visumos išsiskiria grafikas pateikiamas (5.1 pav.). Jame pateikiami rezultatai ieškant varžtų kai nėra šalutinių objektų. Tai mažiausias tiriamas objektas todėl tai turi įtakos aptikimo tikslumui. SSD ir YOLO parodo didžiausias vidutinio tikslumo vertes tai galima matyti ir iš grafiko (5.1 pav.). Dauguma metodų aptinka tik apie 60 procentų objektų, nes grafikai sustoja slinktis į dešinę ties 0.6 verte. Sekančiame pavyzdyje (5.2 pav.) pateikiami grafiniai rezultatai kai tiriama plastiko ritė. Tai vienas didesnių objektų, turintis aiškią formą, todėl matomas aptikimo rezultatų pagerėjimas. Dauguma metodų, nevertinant HOG SVM išlaiko aukštą aptikimo tikslumą, didelėje paveikslų rinkinio dalyje. Nepaisant to jog egzistuoja pašaliniai objektai, kurie apsunkina aptikimą. SSD metodas parodo aukštą aptikimo tikslumą nors neranda visų imtyje esančių objektų. Faster R-CNN ir YOLO geba rasti visus objektus tačiau pasitaiko klaidingai identifikuotų objektų todėl nukečia tikslumas.



5.2 pav. Plastiko aptikimo bendras grafikas (dirbtinė šviesa ir šalutiniai objektai)

Stabiliausiai ir patikimiausiai kintant sąlygoms veikė YOLO ir SSD, šie metodai parodė aukštus aptikimo procentus bei žemiausias vidutinės kvadratinės paklaidos vertes. YOLO reikalingas trumpesnis mokymo laikas ir mažesni apdorojimo resursai tačiau ilgiau užtrunka aptikimo operacija. Naudojant SSD galima tiksliau aptikti objektus su trumpesniu aptikimo laiku, bei apdoroti vaizdų srautus iki 14 kadrų per sekundę su turima įranga.

Rezultatai ir išvados

1. Parinkti ir realizuoti keturi skirtingi objektų paieškos metodai. Du metodai remiasi grafiniu procesoriumi Faster R-CNN ir SSD, o du įprastu HOG SVM ir YOLO.
2. Sudaryta apmokymo bei testavimo vaizdų imtis. Renkant apmokymo vaizdus panaudota augmentacija taip padidinant ir pajvairinant vaizdų skaičių. HOG SVM apmokymui naudota virš 2000 teigiamų ir neigiamų vaizdų, kurių dydis 64x64. YOLO, Faster R-CNN ir SSD naudota po 600 vaizdų, kurių matricos dydis 640x480. Visiems penkiems objektams sudarytos atskiros šimto vaizdų testavimo imtys, pasitelkus Kinect V2 kamerą. Testavimo rinkiniuose sukurtos nepastovios sąlygos keičiant apšvietimą, bandinių skaičių bei pridėdant šalutinių objektų.
3. Įgyvendinti du būdai konvertuoti pikselių koordinates į realius matmenis. Pirmas būdas naudoja gylio kameros duomenis, antras RGB vaizdą ir markerius. Gylio kameros ir RGB metodams paskaičiuotos vidutinės absoliutinės paklaidos, atitinkamai gauta 6,321mm ir 7,875mm. Matavimai atlikti su liniuote, todėl tikslumui įtakos turi žmogiškas faktorius, tačiau pastebimos lėšio poveikio tendencijos. Naudojant RGB vaizdu paremtą metodą, rezultatai prastėja tostant nuo markerių kadro kraštuose. Naudojant gylio kameros duomenis didžiausi nuokrypiai gaunami kadro kampuose. Gylio kameros metodui nereikia naudoti papildomų markerių bei galima išmatuoti objekto aukštį.
4. Metodams vertinti parenkami keturi kriterijai: vaizdo apdorojimo laikas, vidutinis tikslumas, vidutinė kvadratinė paklaida, aptiktų objektų procentas. Kai kadre nebuvo šalutinių objektų geriausius rezultatus parodė Faster R-CNN. Jis aptiko 89,96% objektų, o vidutinė aptikimo vertė 0,84. Kadruose su pašaliniais objektais YOLO ir SSD veikė geriausiai, nes aptiktų objektų procentas atitinkamai buvo 41,9% ir 50%. Vidutinis tikslumas siekė 0,31 YOLO ir 0,28 SSD metodui. Prasčiausius rodiklius viso tyrimo metu parodė HOG SVM metodas. Kai nereikia didelės greಿತaveikos su turima įranga taikytinas YOLO, jei mažas aptikimo laikas yra prioritetas tada pranašesnis yra SSD metodas.

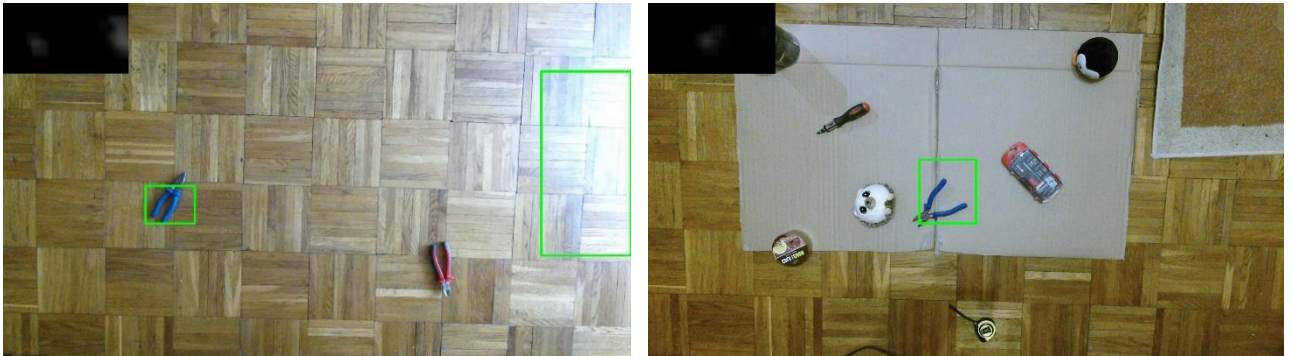
Literatūros sąrašas

1. **Rahul Kumar, Sunil Lal, Sanjesh Kumar** Object detection and recognition for a pick and place Robot [žiūrėta 2020 sausio 20 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/abstract/document/7053853>
2. **Anders Grunnet-Jepsen, Dave Tong** Depth Post-Processing for Intel® RealSense™ D400 Depth Cameras [žiūrėta: 2019 Lapkričio 20 d.] Prieiga per internetą: <https://www.mouser.ec/pdfdocs/Intel-RealSense-Depth-PostProcess.pdf>
3. **Keisuke Dohi, Yuji Yorita, Yuichiro Shibata** Pattern Compression of FAST Corner Detection for Efficient Hardware Implementation [žiūrėta 2020 Balandžio 2 d.] Prieiga per internetą: <https://ieeexplore.ieee.org/abstract/document/6044867>
4. **Deepanshu Tyagi** Introduction to FAST (Features from Accelerated Segment Test) [žiūrėta 2020 Balandžio 5 d.] Prieiga per internetą: <https://medium.com/analytics-vidhya/introduction-to-fast-features-from-accelerated-segment-test-4ed33dde6d65>
5. **Yanwei Pang, Yuan Yuan, Xuelong Li** Efficient HOG human detection [žiūrėta 2020 Balandžio 5 d.] Prieiga per internetą: https://www.sciencedirect.com/science/article/abs/pii/S0165168410003476?casa_token=OCHq1FCxSCwAAAAA:QKWOTzpSe8nbSsa2NBvBd6CwgVSSjGX081EU-WmyIyW1Zem5kvz7Pk0ils_0bH6QFnw7WL2t8w
6. **Satya Mallick** Histogram of Oriented Gradients explained using OpenCV [žiūrėta 2021 Spalio 10 d.] Prieiga per internetą: <https://learnopencv.com/histogram-of-oriented-gradients/>
7. **Ahmed Fawzy Gad** Faster R-CNN Explained for Object Detection Tasks [žiūrėta 2021 Spalio 10 d.] Prieiga per internetą: <https://blog.paperspace.com/faster-r-cnn-explained-object-detection/>
8. **Miao Kang, Xiangguang Leng, Zhao Lin** A modified faster R-CNN based on CFAR algorithm for SAR ship detection [žiūrėta 2021 Lapkričio 4 d.] Prieiga per internetą: https://ieeexplore.ieee.org/abstract/document/7958815?casa_token=rNs1fFjudbMAAAAA:nNYDEFY-zn-R1v8xRj4bg--YXFMwo3E9uY2wVSmQeGmEgEZ0Yw7ZyETwDwGQrh2Ht-knW9tKAQ
9. **Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe** Rethinking the Inception Architecture for Computer Vision [žiūrėta 2022 Balandžio 15 d.] Prieiga per internetą: <https://arxiv.org/pdf/1512.00567v3.pdf>
10. **Hmrishav Bandyopadhyay** YOLO: Real-Time Object Detection Explained [žiūrėta 2020 Balandžio 5 d.] Prieiga per internetą: <https://www.v7labs.com/blog/yolo-object-detection>
11. **Aleksa Ćorović, Velibor Ilić, Siniša Đurić** The Real-Time Detection of Traffic Participants Using YOLO Algorithm [žiūrėta 2022 Kovo 10 d.] Prieiga per internetą: https://ieeexplore.ieee.org/abstract/document/8611986?casa_token=l4MjIq4ia6UAAAAA:pspXaRbx_Xf0GIDOXOnOpH_Jq96wAscVvMGSZG1AJBwLrE956B6jCuP8S27D_PDCGpPoVTLmQg
12. **Wei Liu, Dragomir Anguelov, Dumitru Erhan** SSD: Single Shot MultiBox Detector [žiūrėta 2022 Vasario 3 d.] Prieiga per internetą: https://link.springer.com/chapter/10.1007/978-3-319-46448-0_2
13. **Andrew G. Howard, Menglong Zhu, Bo Ch** MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [žiūrėta 2022 Vasario 3 d.] Prieiga per internetą: <https://arxiv.org/abs/1704.04861>

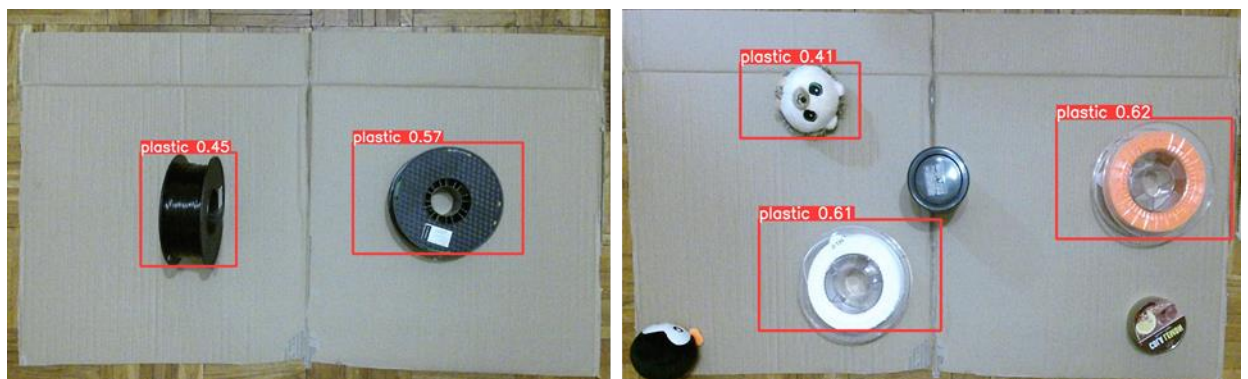
14. **Yosuke Shinya** USB: Universal-Scale Object Detection Benchmark [žiūrėta 2022 Kovo 4 d.] Prieiga per internetą: <https://arxiv.org/pdf/2103.14027.pdf>
15. **Renu Khandelwal** Evaluating performance of an object detection model [žiūrėta 2022 Kovo 5 d.] Prieiga per internetą: <https://towardsdatascience.com/evaluating-performance-of-an-object-detection-model-137a349c517b>
16. **Jonathan Hui** mAP (mean Average Precision) for Object Detection [žiūrėta 2022 Kovo 16 d.] Prieiga per internetą: <https://jonathan-hui.medium.com/map-mean-average-precision-for-object-detection-45c121a31173>
17. **Mark Everingham, S. M. Ali Eslami, Luc Van Goo** The PASCAL Visual Object Classes Challenge: A Retrospective [žiūrėta 2022 Kovo 16 d.] Prieiga per internetą: <http://host.robots.ox.ac.uk/pascal/VOC/pubs/everingham15.pdf>

Priedai

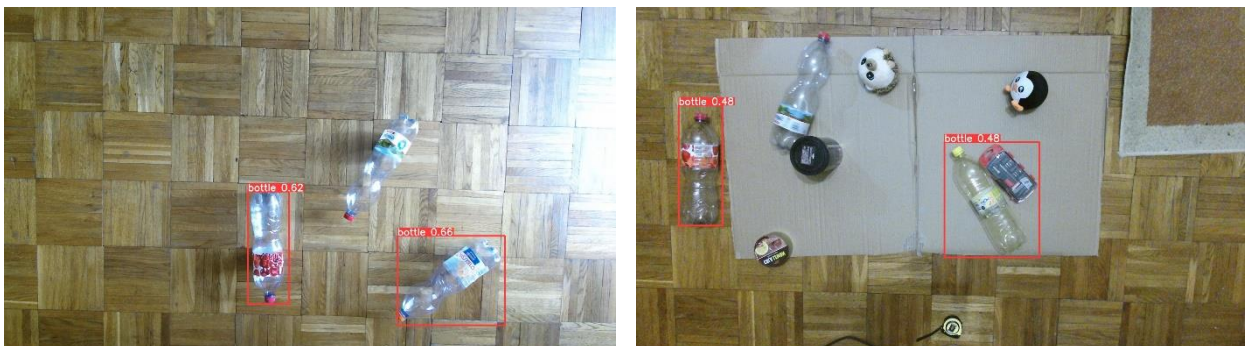
1 Tyrimo metu gauti vaizdai



1.1 pav. Replių paieškos rezultatai naudojant HOG ir SVM



1.2 pav. YOLO rezultatai: kairėje natūrali šviesa, dešinėje dirbtinė bei papildomi objektai



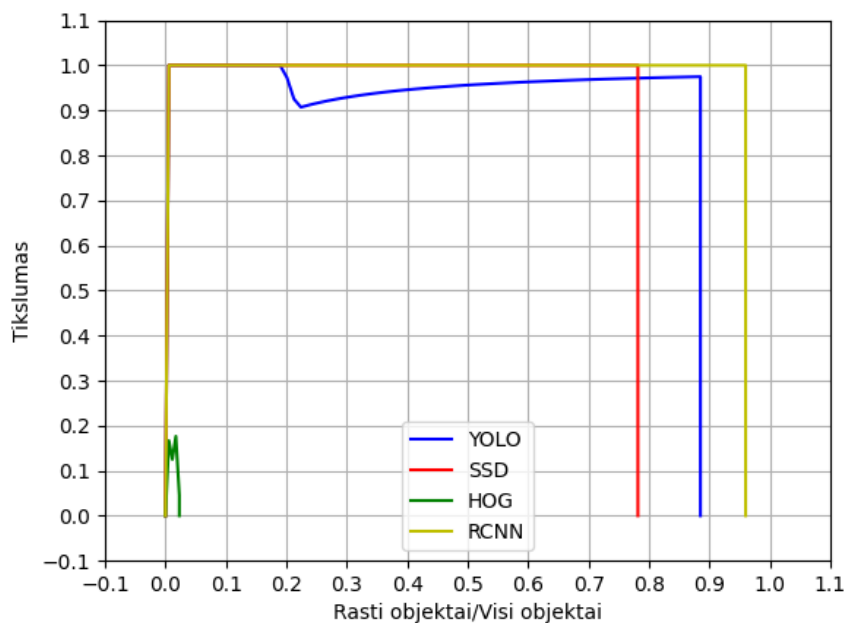
1.3 pav. YOLO rezultatai: kairėje natūrali šviesa, dešinėje dirbtinė bei papildomi objektai

2 Pirminių bandymų rezultatų lentelė

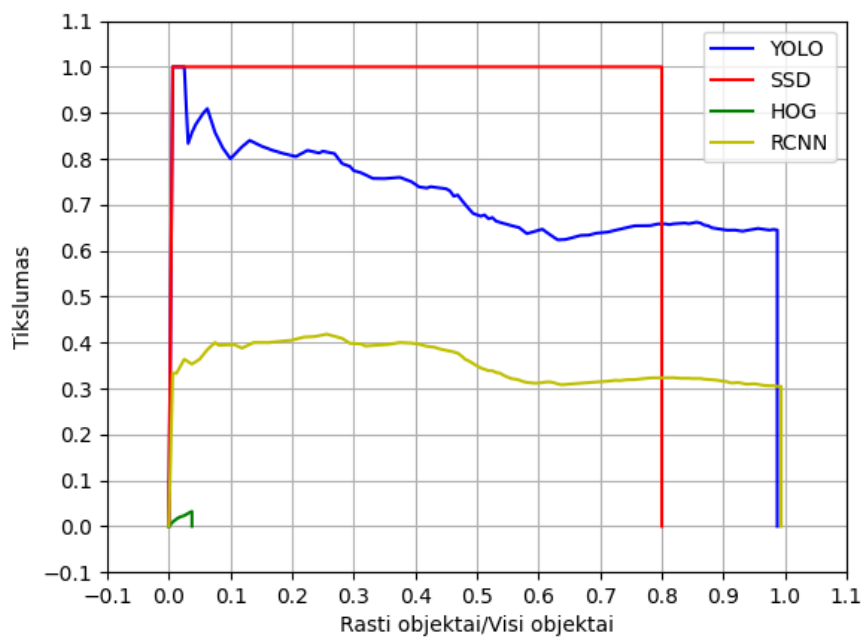
2.1 lentelė. Testavimo rezultatai (natūrali šviesa tik tiriamieji objektai)

| Vaizdų grupė | Vaizdų skaičius | Vidutinis tikslumas (AP) | Vidutinė kvadratinė paklaida (MSE) | Aptikimo vidutinis laikas (s) | Aptiktų/visų objektų kiekis | Teisingai aptiktų objektų procentas, % |
|----------------|-----------------|--------------------------|------------------------------------|-------------------------------|-----------------------------|--|
| Plastikas | 100 | 0,982 | 0,504 | 0,19 | 171/174 | 98,2 |
| Butelis | 100 | 0,999 | 0,682 | 0,18 | 257/257 | 100 |
| Varžtas | 100 | 0,325 | 0,759 | 0,22 | 129/195 | 66,1 |
| Ventiliatorius | 100 | 0,998 | 0,631 | 0,21 | 204/204 | 100 |
| Replės | 100 | 0,991 | 0,558 | 0,20 | 169/170 | 99,4 |

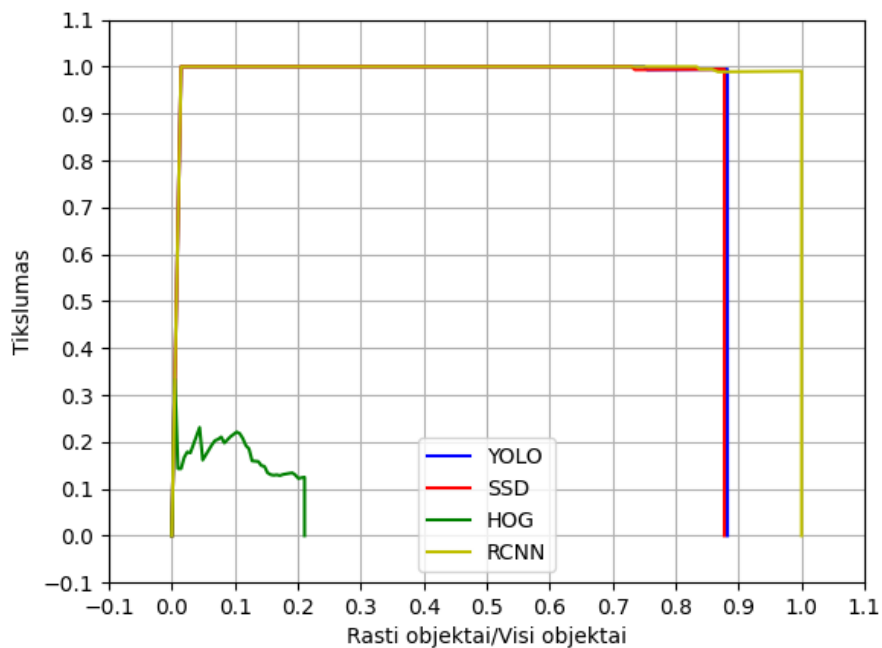
3 Suminiai grafikai



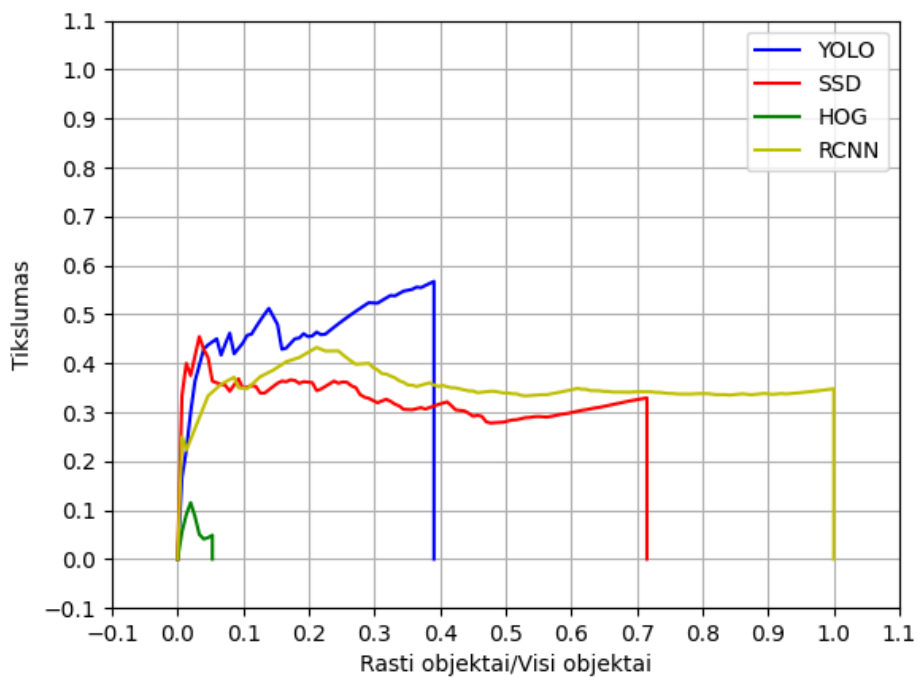
3.1 pav. Plastiko ritės aptikimo skirtingais metodais rezultatų regresija (dienos šviesa)



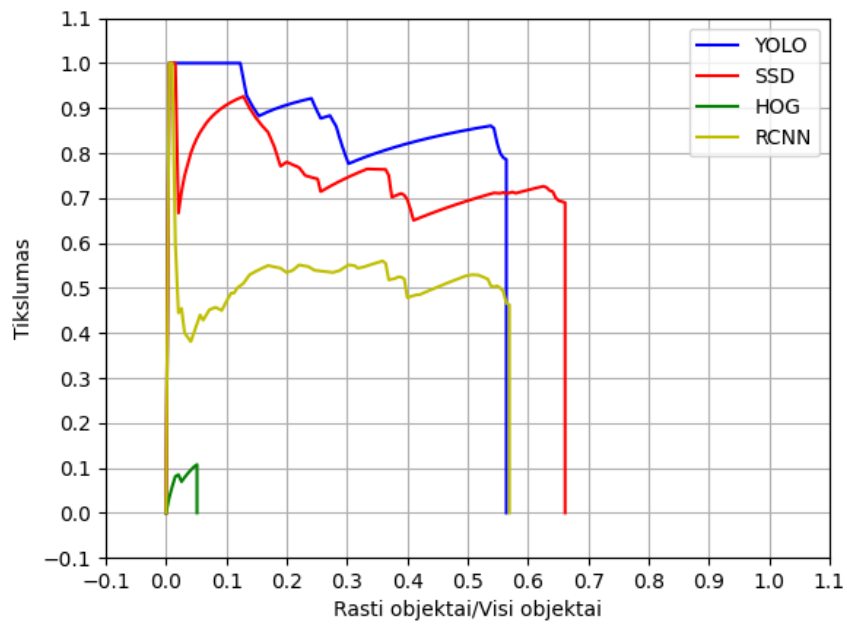
3.2 pav. Plastiko ritės aptikimo skirtingais metodais rezultatų regresija (šalutiniai objektai ir dirbtinė šviesa)



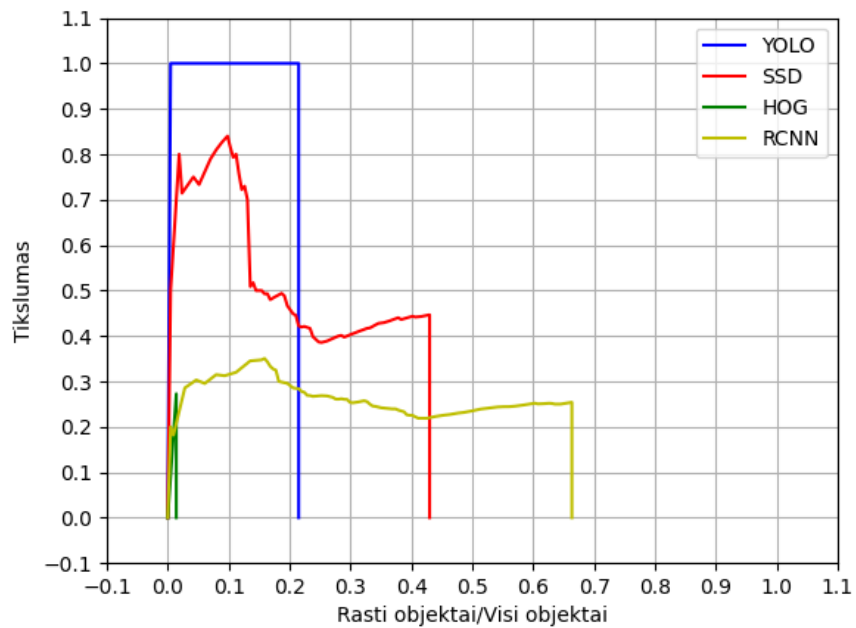
3.3 pav. Ventiliatoriaus aptikimo skirtingais metodais rezultatų regresija (dienos šviesa)



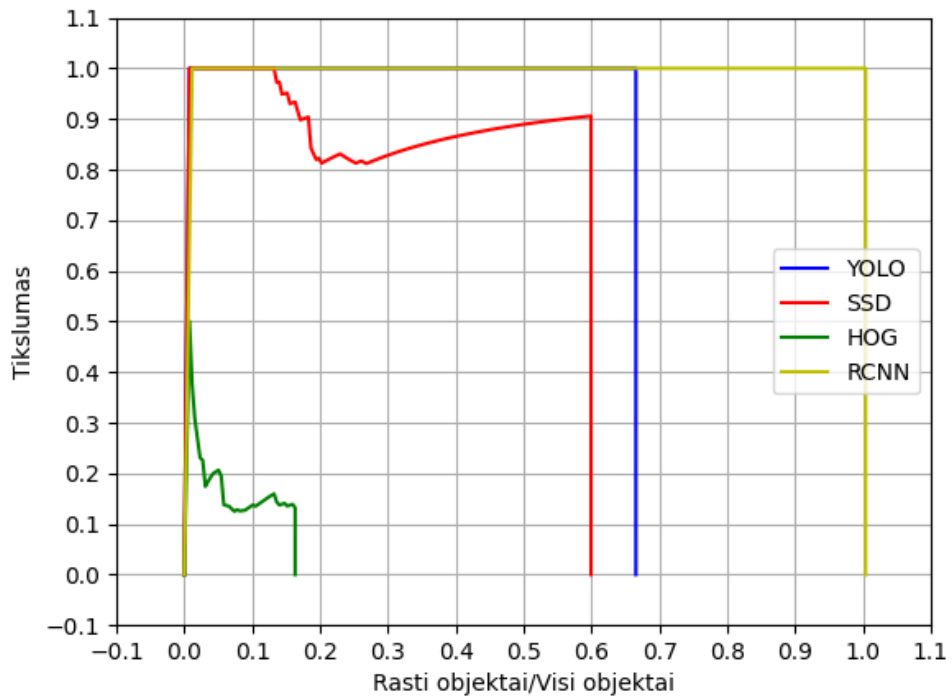
3.4 pav. Ventiliatoriaus aptikimo skirtingais metodais rezultatų regresija (šalutiniai objektai ir dirbtinė šviesa)



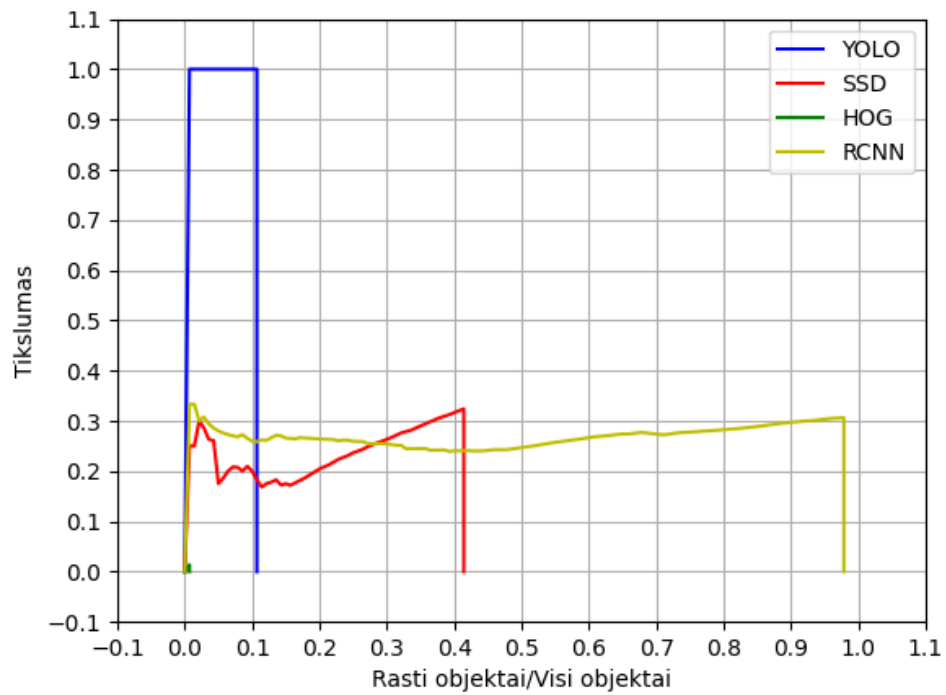
3.5 pav. Varžto aptikimo skirtingais metodais rezultatų regresija (dienos šviesa)



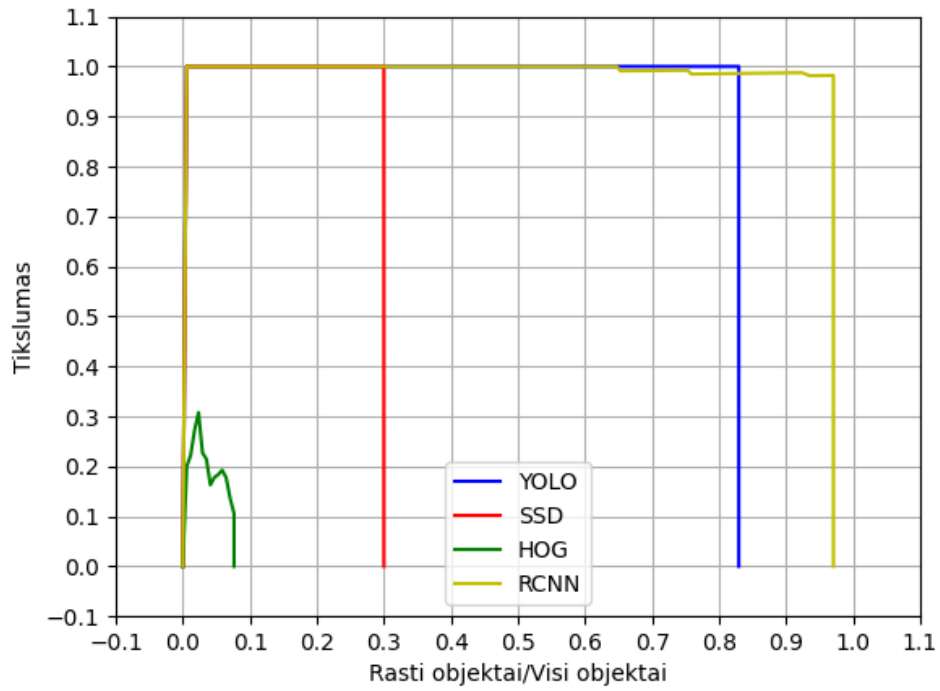
3.6 pav. Varžto aptikimo skirtingais metodais rezultatų regresija (šalutiniai objektai ir dirbtinė šviesa)



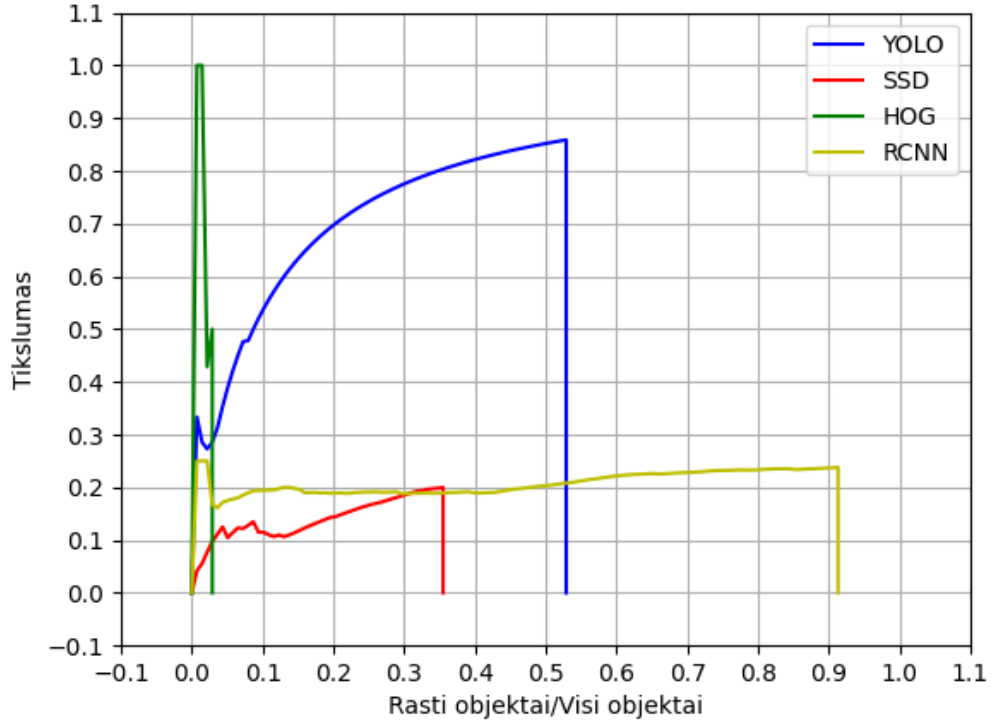
3.7 pav. Butelio aptikimo skirtingais metodais rezultatų regresija (dienos šviesa)



3.8 pav. Butelio aptikimo skirtingais metodais rezultatų regresija (šalutiniai objektai ir dirbtinė šviesa)



3.8 pav. Replių aptikimo skirtingais metodais rezultatų regresija (dienos šviesa)



3.10 pav. Replių aptikimo skirtingais metodais rezultatų regresija (šalutiniai objektai ir dirbtinė šviesa)