

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Ivar Vipper, 183111IAPM

Laoliikumiste tuvastamine ja jälgimine turvakaamerate abil

Magistritöö

Juhendaja: Juhan-Peep Ernits, PhD

Tallinn 2020

Autorideklaratsioon

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Ivar Vipper

05.01.2021

Annotatsioon

Töö kirjutamise ajal on laondus ning laologistika jätkuvalt suur probleem keskmiste ning väiksemate ettevõtete jaoks, kellel ei ole võimalik investeerida suuri summasid probleemi lahendamiseks või leevendamiseks. Seetõttu uuritakse järgnevas töös lahendust probleemile, kasutades turvakaamerate infrastruktuuri, mis on tõenäoliselt erinevates ettevõtetes juba olemas.

Uurimuse raames viiakse läbi eksperiment, mille tulemusel valmib uudne poolautomaatne laoliikumiste jälgimise süsteem, kus konkreetsete toodete asemel jälgitakse laotransporditööriid liikumisi ning hõivatuse staatuseid, jättes konkreetse toote tuvastuse eraldi ülesandeks.

Selline süsteem koosneb kolmest peamisest etapist. Esiteks tuvastatakse videokaadritel asuvad käsikahveltõstukid või tõstukid ning nende hõivatuse staatused. Seejärel teisendatakse nende asukohad kaameratevahelisse ühtsesse koordinaatsüsteemi ning viimaseks jälgitakse leitud objekti ning sellel asuvat toodet ajas ning maapinnal.

Järgnevalt analüüsitakse töös väljapakutud süsteemi tõhusust kolme stsenaariumi põhjal, kus hinnatakse valminud laojälgimissüsteemi prototüübi headust. Uuritakse nii süsteemi toote hoidmise võimet, toote üles korjamise ning maha panemise täpsust meetrites ning leitakse probleeme süsteemi töös.

Sooviga jõuda lähemale keskmiste ning väikeste ettevõtete laoprobleemide lahenduse leidmisele, hinnatakse töös süsteemi tugevaid ning nõrkasid osasid põhinedes stsenaariumite analüüsi tulemustele. Samuti koostatakse põhjalik kvalitatiivne süsteemi kuluanalüüs ning arutletakse lahenduse piirangute ning edasiste lubavamate arengusuundade üle, mis prototüübi töökindlust parandaks. Viimaseks konsulteeritakse laondusala eksperdiga ning arutletakse valminud süsteemi kasutatavuse üle ühe tööstusettevõtte laos.

Lõputöö koos lisadega on kirjutatud eesti keeles ning sisaldab teksti 70 leheküljel, 7 peatükki, 27 joonist.

Abstract

Inventory movement tracking using video surveillance system

At the time of writing this thesis, the inventory tracking is still a big problem especially for small to medium sized companies, who do not have the means to invest large sums to solve or alleviate the problem. Therefore the following thesis aims to find a possible solution to the inventory tracking problem by proposing a novel solution using video surveillance infrastructure, which highly likely already exists in most companies.

An experiment is conducted in order to create a prototype inventory tracking system, where instead of tracking specific stock items, the system tracks pallet jacks and forklifts and their occupied statuses. Identifying specific stock items is left as a separate task, where some possibilities to achieve that goal are explored.

The proposed system is then analysed based on three scenarios, where its strengths, weaknesses and overall effectiveness is judged. Also based on the scenarios, an extensive cost analysis is constructed, as well as the numerous limitations and further advancements of the proposed system are explored. Last, in order to assess the prototype solution's viability in manufacturing environment, a discussion with an industry expert is held.

The thesis is written in Estonian and contains 70 pages of text, 7 chapters, 27 figures.

Lühendite ja mõistete sõnastik

RFID	<i>Radio-frequency identification</i> , raadiolainetel põhinev tuvastus
RTSP	<i>Real Time Streaming Protocol</i> , videovoogesitus protokoll
CAD	<i>Computer-aided design</i> , tarkvarapõhine disain
AP	<i>Average precision</i> , keskmine täpsus
IoU	<i>Intersection over Union</i> , objektituvastuse headuse mõõtühik
MSRP	<i>Manufacturer's Suggested Retail Price</i> , tootja soovituslik jaehind
QR	<i>Quick Response code</i> , triipkoodiliik

Sisukord

1	Sissejuhatus	11
1.1	Probleemi kirjeldus	12
1.2	Eesmärk	12
1.3	Töö struktuur	13
2	Eksperiment	15
2.1	Turvakaamerate infrastruktuur	15
2.2	Kaamerate kalibratsioon	17
2.3	Lähteandmed	18
2.3.1	Huvialused klassid	19
2.3.2	Andmete jagunemine	20
2.4	Objektituvastus	21
2.4.1	Võrreldavad mudelid	22
2.4.2	Katsetused	22
2.4.3	Mudelisisene võrdlus	23
2.4.4	Kaalutud keskmine hinne	23
2.4.5	Tulemused	25
2.5	Objektijälgimine	27
2.5.1	Validatsioon	28

3 Lahendus	29
3.1 Radade ühendamise ning lahkulöömine	29
3.2 Objekti staatuse leidmine	31
4 Süsteemi analüüs	32
4.1 Stsenaarium 1	34
4.2 Stsenaarium 2	37
4.3 Stsenaarium 3	41
4.4 Üldine hinnang	42
4.5 Süsteemi kulu	43
5 Piirangud ning täiendused	46
5.1 Tuvastatavate transportöörde iseärasused	46
5.2 Koordinaatide teisendus	47
5.3 Objektituvastus	49
5.4 Objektijälgimine	51
6 Eksperdi hinnang	53
7 Kokkuvõte	55
Kasutatud kirjandus	57
Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks	63

Lisad	64
7.1 Intersection over Union headuse näitaja	64
7.2 Average Precision ehk keskmine täpsus	65
7.3 Mudelite treenimise parameetrid	65
7.3.1 Faster-RCNN	66
7.3.2 YOLOv5	67
7.3.3 EfficientDet	67
7.3.4 YOLOv4	68
7.4 Kaalutud keskmine objektituvastus hinne	69

Jooniste loetelu

1	Kaamerate positsioonid ning kattuvused	16
2	Kaamera kalibratsiooni näidis	18
3	Näidised sildistatud klassidest	19
4	Treening- ning lõppvalidatsiooni andmete erinevuste näidis	20
5	Süsteemi lõpplahenduseni jõudmise protsess	21
6	Klasside esindatus	21
7	Andmete suuruse jaotus	24
8	Objektituvastus täpsused	26
9	Mudelite paigutamise võime	27
10	Objekti raja muutuse näidis	28
11	Analüüsi alusjoonis asukohas tagahoov	33
12	Stsenaarium 1: tõstuk laadib toodet veokilt maha	35
13	Vale toote mahapaneku tuvastus	35
14	Tõstuk laadib maha toodet	35
15	Tõstuki liikumine kaamerate vahel	36
16	Objekti rada lüüakse lahku	37
17	Käsikahveltõstuk liigub kaamerate vahel	38
18	Käsikahveltõstuk liigub tõstuki tagant läbi	38

19	Käsikahveltõstuki liikumise võrdlus kaamerasuhtes peidetuse korral	39
20	Tõstuk liigutab toodet kaks korda	40
21	Kahe tõstuki kattuvuse näidis	41
22	Näidised erinevatest laotransportööridest	46
23	Kaameranurkade näidis	50
24	Sama objekt kahe kaamera vaates	51
25	Laotransportööri asukoha ennustus	52
26	IoU võrdlus	64
27	Keskmise suurusega objekti näidis	70

1 Sissejuhatus

Vaatamata tõsiasjale, et suurtes ettevõtetes, nagu näiteks Amazon, lahendatakse laohalduse automatiseerimise probleeme nii jälgimissüsteemide kui ka robottransportöride abil [1, 2, 3], on väikeste ning keskmise suurusega ettevõtete jaoks laologistika haldamine siiani oluline probleem. Traditsiooniliselt töötab laohaldus tänu inimeste käsitööle, nagu näiteks toodete positsioonide laohaldussüsteemi sisestamine peale igat lao liikumist. Seda tegevust on küll lihtsustatud kasutades triipkoodide skaneerimist ja raadiosagedustuvastust, kuid sellegipoolest sõltub traditsiooniline laohaldus täna siiski veel suures osas inimestest.

Vähemal määral on kasutusel laohaldusele ja siseruumides positsioneerimisele suunatud tehnoloogiaid [4, 5], mis aitavad automatiseerida lao jälgitavust. Praktikas aga on sellised erilahendusi vajavad süsteemid tihti liiga kallid võrreldes traditsiooniliste lahendustega, mistõttu neid tihti kasutusele ei võeta.

Teisest küljest aga on viimastel aastatel kiiresti arenenud tehisnägemismeetodid kasutades ainult kaameraid. Aktiivselt tegeletakse mitme objekti jälgitavuse etalonide loomisega [6] ning universaalsete objektituvastusmeetodite arendamisega [7]. Veel annab motivatsiooni fakt, et Amazon Go (pood, kus ei ole kassasid) kasutab videokaameraid, et jälgida, kes mida ostab [8]. Tõepoolest, tööstuses investeeritakse järjest enam automaatsesse laohaldusesse kasutades kaameraid [9], kuid samas valitseb suur teadmatus kaamerate kasutatavuse kohta konkreetses kontekstis [10].

Käesolevas töös uuritakse, kui kasutatav on olemasolevatel turvakaameratel põhinev laoliikumiste jälgimise süsteem. Selleks viiakse läbi eksperiment erinevate tehisnägemistehnoloogiate hindamiseks laonduse kontekstis ning pakutakse välja laoliikumiste jälgimise prototüüplahendus. Väljatöötatud lahendusele annab hinnangu valdkonna ekspert, kelle arvamus on ka lõpphinnanguks selle kasutatavuse kohta tööstuslikes tingimustes.

1.1 Probleemi kirjeldus

Laoliikumiste jälgitavus on siiamaani suur probleem, mille tõttu ettevõtted kaotavad igal aastal raha [11]. Üheks otseseks kuluks on laokaupade leidmise ja ladustamise ajakulu. Aega, mis kulub töölistel laohalduseks, on võimalik kokku hoida, kui on teada, kus on otsitud toode või kuhu peaks uue toote ladustama.

Olemasolevad lahendused on enamasti väikeste ning keskmiste suurustega ettevõtete jaoks ebapiisavad. Tõenäoliselt kõige populaarsem lahendus – triipkoodid koos skänneritega, eeldab üldiselt otsest inimese sekkumist nii laoartikli liikumaasumise alg- kui ka lõppasukohas. Lisaks sellele peavad igal tootel olema nähtavad märgised [12].

Raadiosagedustuvastuse põhised lahendused vajavad samuti igal jälgitaval objektil asuvat märgist. Samuti vajavad eelnimetatud lahendused suuri infrastruktuurilisi muudatusi, mis võivad väiksemate ettevõtete jaoks muuta laohalduse automatiseerimise liiga kulukaks [13].

Objektituvastus- ning jälgimismeetodid on viimaste aastatega kiiresti arenenud [14, 15]. Samas on kritiseeritud eelnimetatud alade teadustööde fookust, kuna sageli keskendutakse täpsema objekti tuvastus- või jälgimismeetodi loomisele selmet uurida olemasolevate lahenduste kasutatavust reaalseste probleemide lahendamiseks [16].

Seega püstitan hüpoteesi, et kasutades olemasolevaid masinõppel põhinevaid tehisenägemise tehnoloogiaid on turvakaamerate põhise süsteemi abil võimalik laoliikumiste jälgimine lahendada väiksemate kuludega kui konkureerivad lahendused ning seeläbi teha laohaldussüsteem kättesaadavamaks ka väiksematele ettevõtetele.

1.2 Eesmärk

Käesoleva töö peamine eesmärk on pakkuda välja prototüüp süsteemist, mis võimaldab poolautomaatselt jälgida laoliikumisi erinevate ruumide vahel. Süsteemi arhitektuuris kasutatakse olemasolevat turvakaamerate infrastruktuuri, mis võimaldab töö autori poolt välja pakutud lahendust kasutada igas laos, kus on olemas turvakaamerad.

Täpsemalt on töö eesmärk leida vastused järgnevatele alamküsimustele:

- Kui täpselt on võimalik laotranspordivahendeid tuvastada kasutades turvakaameraid, et nende liikumisi jälgida? Täpsus on defineeritud peatükis 2.4.3.
- Kui kasutatav on turvakaameratepõhine laojälgimissüsteem? Täpsus on defineeritud peatükis 2.5.1.
- Millised on suuremad piirangud turvakaameratepõhise laojälgimissüsteemi juures?

Konkreetsel laotranspordivahendite tuvastamine ei ole töö ülesanne, selle asemel tuvastatakse erinevate laotranspordivahendite liikumisi ning kasutaja otsustada jääb, kuidas konkreetne laotranspordivahend siduda laotranspordivahendiga. Ainult laotranspordivahendite tuvastamise põhine lähenemine võimaldab ehitada süsteemi, mida saab kasutada võimalikult paljudes ladudes: reeglina on tuvastatavad tava- ning käsikahveltõstukid sarnase väljanägemisega.

Samuti ei ole konkreetse kauba täpse asukoha leidmine käesoleva magistritöö ülesanne. Hüpotees loetakse töö autori poolt kinnitatuks, kui laoliikumisi suudetakse jälgida mitme kaamera vahel ning toode jääb tegelikult asukohast mõistlikusse raadiusesse.

Selline valideerimiskriteerium valiti, kuna arvestades jälgitavate huvialuste objektide suurus¹, on töö autori hinnangul võimalik laotöötajal otsitav alus leida visuaalse vaatluse teel eeldusel, et otsinguid alustatakse toote tegelikult asukohast mõistlikult lähedalt.

1.3 Töö struktuur

Töö eksperimendi peatükk 2 selgitab üksikasjalikult, kuidas igale uurimisküsimusele vastus leitakse ning mille põhjal leitud lahendusi valideeritakse. Sellele järgneb välja töötatud lahenduse kirjeldamine peatükis 3, kust on võimalik leida, mis algoritme lahendus kasutab.

Lahendusele järgnevas peatükis on koostatud süsteemi analüüs, kus hinnang prototüübi tööle antakse analüüsides kolme erinevat laoliikumise juhtumit. Peatükk 4 lõppeb detailse kvalitatiivse süsteemi kuluanalüüsiga, kus antakse hinnang muuhulgas prototüüplahenduse ülesseadmiskulule ning halduskulule.

Peatükis 5 on kirjeldatud piiranguid, mida peaks kindlasti arvesse võtma enne süsteemi

¹EUR aluse mõõtmed on 1200mm x 800mm

reaalset kasutust. Samuti analüüsitakse käesoleva töö peatükis 5 süsteemi edasiarendusi ning antakse hinnang, millised väljatöötatud süsteemi edasiarendused on prioriteetsemad. Viimaseks antakse peatükis 6 hinnang väljatöötatud prototüüplahendusele laondusala eksperdi poolt.

2 Eksperiment

Eksperimendid uurimisküsimustele lahenduse leidmiseks viidi läbi tööstusettevõtte laoruumides ning analüüsiti reaalseid laoliikumisi. Reaalandmetega eksperimendil on võrreldes simulatsiooniga järgmised eelised:

- Kasutatavate turvakaamerate asukohad on eelnevalt fikseeritud. Seega eksperimendi tulemusena saab anda hinnangu süsteemi toimivusele infrastruktuuriliste muudatusteta.
- Lahendust saab analüüsida kasutades reaalist laoliikumiste sagedust. Üldjuhul on erinevates ladudes liikumiste tihedus sarnane eriti maksimaalsete samaaegsete liikumiste suhtes: liiga paljude laotranspordööride väiksel alal liikumise korral võib tekkida liiklusprobleeme.

Siinkohal on oluline mainida eksperimendis kasutatud riistvara. Mudeleid trenniti Nvidia GTX1060 6Gb graafikakiirendiga, mistõttu ei olnud võimalik kõiki objektituvastuse ajakohaseid mudeleid võrrelda.

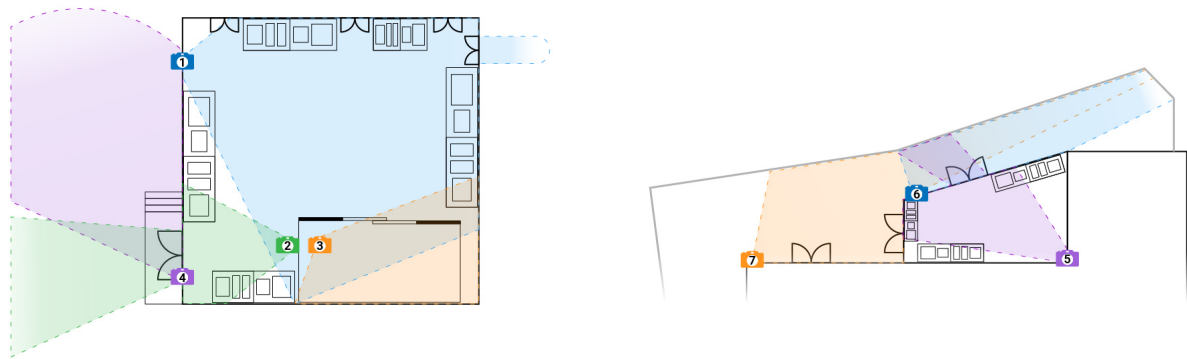
2.1 Turvakaamerate infrastruktuur

Eksperimendis kasutatavad andmed on kogutud 7-st võrgukaamerast kahes erinevas asukohas. Kõik kaamerad on esialgselt üles seatud turvakaameratena, millele on võimalik läbi kohaliku võrgu ligi pääseda kasutades RSTP protokoll². Selguse mõttes on igale kaamerale antud unikaalne tuvastusnumber ning nimetus. Kaamerate positsioonid ning kattuvused on kujutatud joonisel 1.

Kaamera 1: Montaaž

Kuvab pildi peamisest hoone sisesest laoalast. Väheine kattuvus on kaameraga 2 ning täielik kattuvus on kaameraga 3 juhul, kui ruumi Robot 3 uks on avatud. Kaamera 1 vaatevälja jäävad

²Real Time Streaming Protocol [17]



(a) Koosteala ja esihoov

(b) Tagahoov

Joonis 1: Kaamerate positsioonid ning kattuvused.

tootmishoone pooled uksed ning ruumi sisenevad ning sealt väljuvad tooteid on võimalik Kaamera 1 abil tuvastada.

Kaamera 2: Uks nr. 5

Kuvatakse pilt välisuksest ning kui see on avatud, siis on kattuvus kaameraga 4. Juhul kui välisuks on täielikult avatud, takistab vähesel määral kaamera 2 vaatevälja estakaad. Eksperimendi mõttes ei ole selles negatiivset mõju, kuna estakaadi all reeglina ei liigu käsikahveltõstukeid, mille nägemist estakaad kaamera 2 kaadris takistaks.

Kaamera 3: Robot 3

Katab siseruumi, kus asub automaattööpink. Kaamera näeb lähemalt robotis valmivaid detaile ning nende ruumist välja viimist, väljaarvatud väikeses pimenurgas, mida katab kaamera 1.

Kaamera 4: Tõkkepuu

Esihoovis asuv kaamera näeb tõstukeid, mis laadivad kaupa väikestele kaubikutele. Kaamera 4 katab täielikult estakaadi, mis takistab kaamera 2 vaadet.

Kaamera 5: Uks nr. 1 ja 2

Sellest kaamerast on võimalik näha tootmishoone tagumist ladu, mis peamiselt koosneb lühiajalisest laost ning toormaterjali laost. Kui uks 1 on lahti, siis kaamera 5 ja kaamera 6 vaateväli kattub.

Kaamera 6: Tagahoov uks nr. 1

Katab hoone külgmise ala. Üldjuhul tuuakse siin käsikahveltõstukitega alused välja kaamera 5

alast ning edasi liigutakse tõstukiga kaamera 7 vaatevälja. Kõik uks 1 kaudu hoonest lahkuvad laoliikumised on sellest kaamerast tuvastatavad.

Kaamera 7: Tagumine hoov

Kaamera asub tootmishoone peamisel kaubalaadimisalal. Vähene kattuvus on kaameraga 6.

Kõik kaamerad ei ole samade parameetritega. Näiteks on RTSP videovoos kaamerati erinev resolutsioon, kaadrisagedus, video kvaliteet (bitrate) ning ka pildiedastuse viide võib kohati erinev olla. Sisse on toodud eeldus, et kaamerad on statsionaarsed ehk nende vaateväli on alati sama.

2.2 Kaamerate kalibratsioon

Kuna eksperimendi eesmärk on luua laokaupade liikumiste jälgija, on vaja teisendada kaameratest nähtavate objektide asukohad ühtsesse koordinaatsüsteemi. Seeläbi on võimalik objekti asukoht kindlaks määrata olenemata kaamerast. Seega on vaja kaamerad kalibreerida ehk leida iga kaamera jaoks kalibreerimismatriks ning moonutuskoefitsiendid [18].

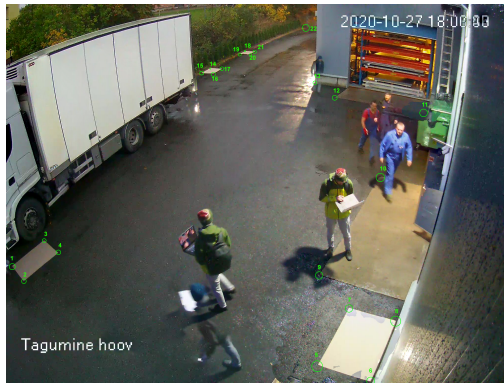
Eksperimendis viidi see läbi järgnevalt. Esiteks kalibreeriti kaamerad kasutades 6x9 malelaua kalibratsiooni, mille tulemusel leiti kaamera kalibreerimismatriks ning moonutuskoefitsiendid.³ Kalibreerimise tulemusel jäeti alles kõik originaalkaadris asuvad pikslid. Kalibreerimise tulemust on võimalik näha joonisel 2b.

Järgnevalt positsioneeriti iga kalibreeritava kaamera kaadrisse tähised. Näidis tähistest on toodud joonisel 2a. Nende tähiste eesmärk oli teisendada igas kalibreeritavas kaameras nähtav piksel ühtsesse koordinaatsüsteemi, mis ühtiks reaalmaailmaga põranda / maapinna tasandil.

$$\begin{bmatrix} X_n \\ Y_n \\ n \end{bmatrix} = H^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (1)$$

Selleks arvutati maailma ning kaamera koordinaatide vastavuste põhjal homograafia

³Kalibratsiooni meetod nähtaval: https://docs.opencv.org/master/dc/dbb/tutorial_py_calibration.html



(a) Originaal kaader



(b) Ennustatud markerid

Joonis 2: Kalibratsiooni näidis. Joonis *a* kujutab originaalkaardit ning joonis *b* kujutab sirgendatud kaadrit ning kalibreerimise punktide viga. Punane joon algab markeri mõõtepunktist (tähistatud rohelse ringiga) ning näitab koordinaatsüsteemide teisendusvea ulatust.

maatriks H ning läbi selle arvutati maailma koordinaadid igale vastavale punktile [19] kasutades valemit 1.

Siinkohal on oluline märkida, et kuna tulemus oli vajalik saada teoreetilisel tasapinnal, siis kõrgust kalkulatsioonides arvesse ei võetud. Reaalsuses aga maapind ei ole täiesti tasane, seega tuli siit sisse mõningane arvutuslik viga.

Vea tõi sisse ka mõõtemetod. Iga tähise koordinaat mõõdeti kasutades laser distantismõõdikut ning kaugus mõõdeti eksperimendi asukohas teadaolevatesse punktidesse, kus teadaolevad punktid saadi eksperimendi asukoha CAD⁴ jooniselt. Et mõõtmisviga vähendada korraldi kauguse mõõtmist igasse punkti vähemalt 5 korda ning võeti tulemus, mis kordus enim.

2.3 Lähteandmed

Eksperimendiks valmistatud andmekogu koosneb $28 + 9$ käsitsi sildistatud videost, milles on kokku $96105 + 27911$ kaadrit. Esimest 28 videot 96105 kaadriga kasutatakse objektivastusmodelite treenimiseks, testimiseks ja validatsiooniks. Viimast üheksat videot 27911 kaadriga kasutatakse lõppmodeli hinnanguks. Need videod on kogutud ajavahemikul juuni 2020 kuni november 2020 ning hõlmavad endas nii sise kui välistingimustes kui ka

⁴Computer-aided design

erineva objektitihedusega kaadreid.

2.3.1 Huvialused klassid

Sildistused on loodud kahe peamise klassi kohta, mis jagunevad omakorda kaheks alamklassiks. See jagunemine on illustreeritud järgnevalt joonisel 3.



(a) Tõstuk hõivamata



(b) Tõstuk hõivatud



(c) Käsikahveltõstuk hõivamata



(d) Käsikahveltõstuk hõivatud

Joonis 3: Näidised sildistatud klassidest.

Lisaks klassidele on sildistatud ka unikaalsed objektid ehk teisisõnu rajad. See tähendab, et kui näiteks konkreetne tõstuk liigub kaamerate vahel, siis objekti raja number jääb samaks. Tänu sellele on võimalik hinnata objekti jälgimise täpsusi mitme kaamera vahel.

Oluline on veel mainida, et raja number võib jääda samaks isegi siis, kui objekti klass vahetub. Näiteks kui tõstuk võtab toote peale, muutub tema klass, kuid mitte raja number. See tähendab, et raja number on seotud konkreetse klassi asemel peaklassiga (käsikahveltõstuk või tõstuk).

2.3.2 Andmete jagunemine

Lähteandmed on jagatud nelja osasse: treening-, test- ning validatsiooniandmed ja viimaseks lõppmudeli hindamise andmed. Sellise andmete jagamise eesmärk on eraldada eksperimentis kasutatavate meetodite ning tehnoloogiate sõltuvus lähteandmetest: arendatav süsteem peab saama hakkama ka ennenägemata videokaadritega. Tõepoolest, objektituvastusmodelite valideerimisandmed on kogutud erinevatest kaameratest või erinevates keskkonnatingimustes kui treenimis ning testandmed.

Oluline on ka välja tuua, et lõppsüsteemi validatsioonis siiski uute kaameranurkadega andmeid ei kasutata kaamerate kalibreerimise võimatuse tõttu. Et vähendada objektituvastusmeetodite kallutatust lõpptulemusel kasutatavate kaadrite suhtes, on lõppvalidatsioonis kasutatud andmeid, mis erinevad treeningandmetest ilmastikutingimuste poolest. Näidis on toodud treening ning lõppvalidatsiooni andmetest joonisel 4.



(a) Treeningandmed



(b) Lõppvalidatsiooniandmed 1

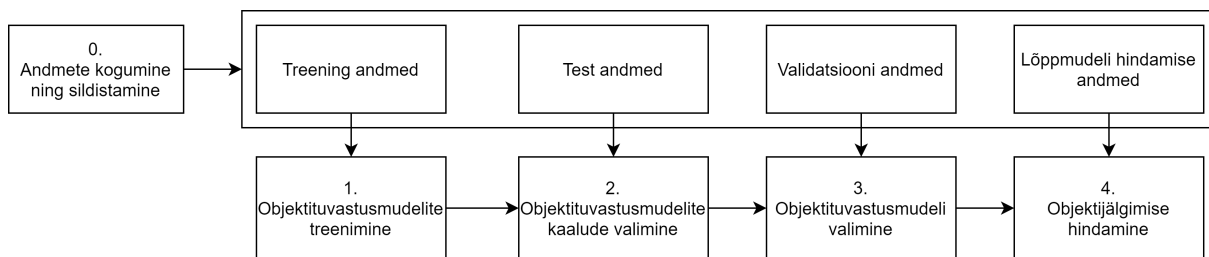


(c) Lõppvalidatsiooniandmed 2

Joonis 4: Erinevused treening- ning lõppvalidatsiooni andmete vahel (ilmastikutingimused ning valgustingimused).

Eksperimentis kasutatavate treening- ning testandmete alamjaotuste erinevused käeoleva töö raames on järgnevad. Esiteks kasutatakse testandmeid ainult objektituvastusmodelite vahetulemuste hindamiseks meetodipõhiselt. Näiteks objektituvastuseksperimentides saab küsimusele "kui hea on mudel praegu" vastuse hinnates mudelit testandmekogul. Selline küsimus loob aga kallutatud hinnangu testandmete suhtes, seega mudelile lõpphinnangu andmiseks kasutatakse validatsiooniandmeid. Täpsemalt on erinevate alamjaotatud andmete kasutuskohad illustreeritud joonisel 5.

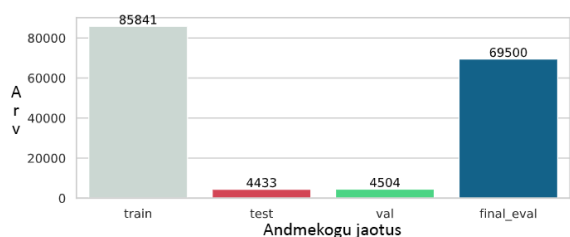
Kurioosse asjaoluna võib mainida, ei töö autori test- ning validatsiooni andmekogude nimed võivad erineda töö lugeja eelnevatest teadmistest. Kui töö lugeja on tutvunud Russell ja Norvig raamatuga "Artificial Intelligence: A Modern Approach" [20], siis selle kirjutise



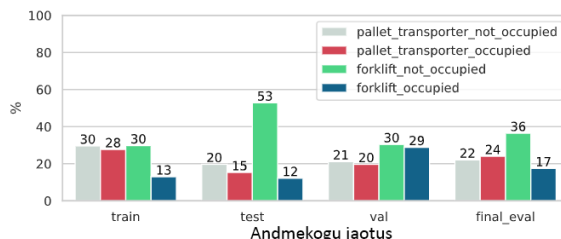
Joonis 5: Süsteemi lõplahenduseni jõudmise protsess ning lähteandmete kasutusala etapi põhiselt.

autorid defineerivad test- ja validatsiooni andmejaotused vastupidiselt. Teisest küljest Kuhn ja Johnson oma raamatus sellist andmejaotuse nimetuse vahet ei tee [21].

Klasside esindatus objektivastusandmekogus on illustreeritud joonisel 6. Kõige olulisemana tuleks tähelepanu pöörata alamjoonisele 6b, kus võib näha, et validatsiooniandmekogu on võrdlemisi ühtlaselt jagunenud klasside vahel. Seega objektivastuseksperimenti tulemused eelnimetatud alamjaotusel annavad hea ülevaate iga klassi kohta samaväärselt.



(a) Objektide arv alamandmekogu kaupa, kus üheks objektiks loetakse ühes kaadris asuvat piirkasti.



(b) Klasside jagunemised alamandmekogude kaupa protsendilisel kogul alamandmekogu suurusest.

Joonis 6: Objektide jagunemine ning klasside esindatus tükeldatud andmekogu kaupa.

2.4 Objektivastus

Laotranspordivahendite ning nende staatuste tuvastamiseks videokaadritest kasutati eksperimentis ainult närvivõrkudel baseeruvaid meetodeid. See otsus langetati põhjendusega, et traditsioonilised objektivastusmeetodid kasutavad üldiselt käsitsi valitud objektiomadusi, sealjuures olles tihedalt sõltuvad treeningandmetest [14].

2.4.1 Võrreldavad mudelid

Eksperimendis võrreldi meetodeid, mis on üldlevinud ning hea täpsusklassiga: YOLOv4 [22], YOLOv5 [23], Faster-RCNN [24], samuti EfficientDet [25]. Täpsemat mudelite võrdlust ning põhjust, miks näiteks EfficientDet mudel edasisest võrdlusest välja jäi, võib näha lisas 7.3.3.

Välja jäeti mudelid, mida ei olnud võimalik mõistliku kiirusega saadaval oleva videokiirendiga kasutada (6GB videokiirendi): videokiirendiga mudelite treenimine võib olla kordades kiirem CPU'1 treenimisest [26, 27]. Eksperimendis kasutatava andmekogu suurust arvesse võttes tähendas ühe mudeli GPU'1 treenimine ajakulu üks päev, võrreldes CPU'1 treenimisega, kus sama mudeli treenimine oleks võtnud ligemale üks nädal.

Lisaks peab mainima, et YOLOv4 ning YOLOv5 mudelid olid eeltreenitud MSCOCO andmekogul [28] ning Faster-RCNN mudel oli eeltreenitud ImageNet andmekogul [29]. Ülekandeõpet kasutati ainult viimase närvivõrgu kihi treenimiseks intuitsiooniga, et eelnimetatud andmekogudes olevad klassid (autod, veokid, rongid jms) on võrdlemisi sarnased eksperimendi huvialuste klassidega ning seetõttu on võimalik mudeli treenimist kiirendada.

Siinkohal on oluline ka mainida, et mudelite treenimise validatsioon viidi läbi peale iga andmekogu pildi nägemist. See tähendab, et eksperimendis osalenud mudelid võisid saavutada täpseima tulemuse ajal, kus validatsiooni ei teostatud. Selle põhjuseks oli jälle ajakulu, kus üks validatsiooni etapp võttis olenevalt mudelist kuni 20 minutit.

2.4.2 Katsetused

Esmased katsed viidi läbi kasutades andmekogu osa, millest olid välja võetud kaadrid tuvastatavatest objektidest, mis olid üleni nähtavad. Selle katse eesmärk oli saada esialgne kindlus, et eksperimendi hüpotees võib leida kinnitust: kui objektituvastusmeetodid ei suuda õppida võrdlemisi lihtsat uurimisdomeeni andmekogu, siis ei ole ka lootust osaliselt kaetud objekte tuvastada ning laoliikumiste jälgimis süsteemi selle põhjal valmistada.

Kui esmased katsed olid õnnestunud, võrreldi järgnevalt mudeleid kõikidel andmekogu varem nägemata validatsiooni andmetel. Seejärel valiti välja parim mudel, mis ühtlasi valiti

objektijälgimise algoritmi sisendiks.

Kõik eksperimendis läbi proovitud mudelid, nende treenimise parameetrid ning ajad on toodud välja lisas 7.3.

2.4.3 Mudelisisene võrdlus

Esmased mudelisesed validatsioonid tehti testandmekogul. Seal võrreldi mudelite täpsust, kasutades erinevaid objektituvastusmeetodite treenimisel saadud kaale ning videokaadri resolutsioone meetodi põhiselt. Parim mudel valiti vastavalt AP50⁵ (AP) headuse kriteeriumile.

Esiteks ning kõige olulisemalt võimaldas AP50 headuse kriteeriumi kasutamine võrrelda mudeleid ühtsemalt: kõik kasutatavad mudelid võimaldasid kasutada tööriistasid, mis automaatselt arvutasid välja eelnimetatud AP50 väärtuse. See võimaldas mudeleid eksperimendi algfaasis iteratiivselt testida, samas andes piisavalt hea intuitsiooni lõppmudeli headusest.

Teiseks võimaldas treenitud mudeli AP50 väärtus võrrelda tulemusi baasmudeli andmestiku tulemustega. Töö autori hinnangul on see hea indikatsioon treenitud mudeli üldisest headusest, kui saavutatud AP50 on võrdlemisi sarnane ülekantud mudeli headusega eeldusel, et andmete 'raskus' on subjektiivsel hinnangul sarnane (sarnased keskkonnatingimused, sarnased objektide suurused, sarnased objektide peidetuse tasemed jne).

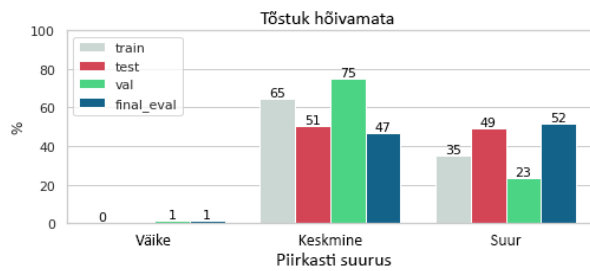
Viimaseks põhjuseks, miks kasutati AP50 hinnet näiteks AP50-AP95 sammuga 0,05 asemel oli põhjus, et pildilt otsitavate objektide tuvastuse lokaalsus (leitud piirkasti kattuvus tegeliku piirkastiga) ei mõjuta soovitud lõpptulemust märkimisväärselt.

2.4.4 Kaalutud keskmine hinne

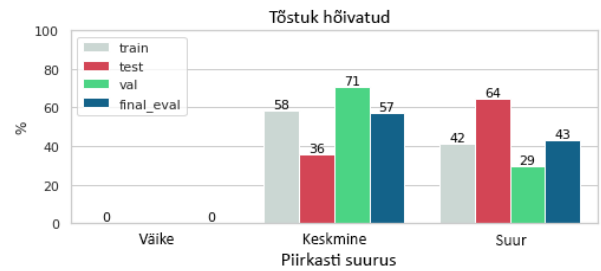
Objektituvastusmeetodite lõppvõrdlus viidi läbi kasutades kaalutud aritmeetilist keskmist AP50 hinnet IoU⁶ 50% juures, kus töö kontekstis huvitavamad klassid ning tuvastatud objekti

⁵Lisa 7.2 - Average Precision ehk keskmine täpsus

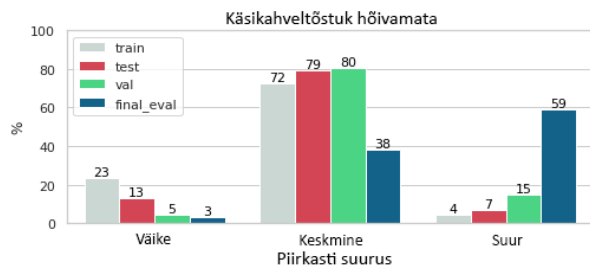
⁶Lisa 7.1 - Intersection over Union headuse näitaja



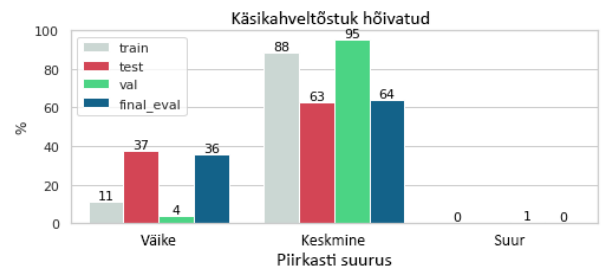
(a) Tõstuk hõivamata



(b) Tõstuk hõivatud



(c) Käsikahveltõstuk hõivamata



(d) Käsikahveltõstuk hõivatud

Joonis 7: Andmete jaotus objekti piirkasti suurust arvestades. Kõik väärtused on toodud välja protsentides kogu andmekogu alamjaotusest.

piirkastid said suurema kaalu.

Täpsemalt jagati objekti piirkastid kolme kategooriasse: väike, keskmine ning suur. Jaotust pikslites on illustreeritud tabelis 1 ning näidised, kuidas jaotus leiti on toodud pildil 27.

Siinkohal on oluline välja tuua kuidas objekti piirkastide suurused leiti. Esiteks otsiti välja väike osa kaadritest, kus huvialused objektid on näha vähemalt kahes turvakaameras korraga iga kaamera kohta. Seejärel märgiti üles nende mõõtmed pikslites (ning ühtlasi ka pindala).

Eesmärgiga luua ühtne suuruste skaala kõikidele laobjektidele tehti otsus keskmise suurusega pindalaks valida ainult käsikahveltõstuki tüübi kaadrid: reeglina on igas kaameras samas asukohas käsikahveltõstuk füüsiliselt väiksem kui tavatõstuk. Seetõttu on klassi jaotus objekti piirkasti suuruse suhtes kallutatud. See kallutus on illustreeritud joonisel 7.

Lisaks eelnevale olid töö kontekstis huvitavamad klassid, kus tooteid liigutati laotranspordivahenditega: eesmärk oli tuvastada transpordivahenditel asuvate toodete liikumisi ning mitte laotranspordivahendite endi liikumisi. Seega suurem kaal antakse klassidele

hõivatud käsikahveltõstuk ning hõivatud tõstuk.

$$\bar{x} = \sum_{i=1}^n w_i x_i \quad (2)$$

Kogu kaalutud keskmise hinde arvutus käis seega kaaludega, mis on illustreeritud tabelis 2 ning kasutades kaalutud aritmeetilise keskmise valemit 2, kus w tähistab kaalu ja x kaalu klassile ja objekti piirkasti suurusele vastavat AP50 väärtust.

Siinkohal on oluline märkida, et kõiki klassi ning piirkasti suuruse kombinatsioone ei võetud arvesse. Täpsemalt ei ole kalkulatsioonis arvestatud klasse tõstuk suurusega väike ning käsikahveltõstuk suurustega väike ning suur: eelnimetatud klassides ei leidunud piisavalt andmeid, et töö autori hinnangul neid klasse arvestada lõpphinde arvutamisel. Joonis 7 illustreerib seda hinnangut.

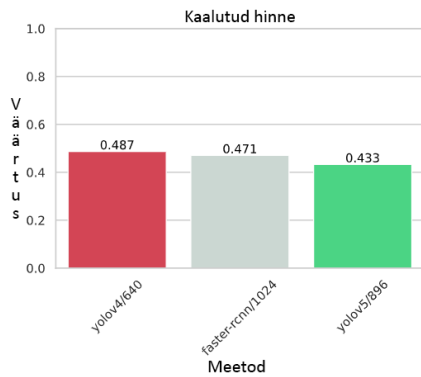
2.4.5 Tulemused

Kasutades defineeritud kaaludega (tabel 2) keskmist hinnet, saavutati parim tulemus YOLOv4 mudeliga. Kõiki mudeleid võrreldi 25% kindluse juures kasutades IoU läve 50% juures. Tulemused on illustreeritud joonisel 8.

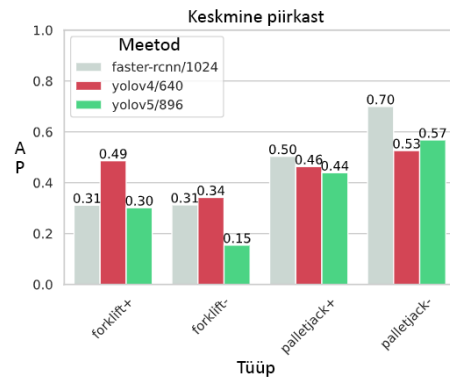
Uurides tulemusi joonistel 8b, 8c ja 8d võib näha, et YOLOv4 mudel on arvestades tõstuki klassi parim. Käsikahveltõstuki klassis on YOLOv4 mudel paremuselt teine hõivatud staatuses ning halvim hõivamata staatuses. Meeldetuletuseks kaalutud hindede hõivamata staatuses on antud väiksem kaal, seega loetakse YOLOv4 mudel parimaks.

Samuti kui võrrelda ainult keskmise suurusega klasside tulemusi joonisel 8b võib näha, et YOLOv4 mudel on parim tõstuki klassi tuvastaja. See on oluline, kuna keskmise suurusega tõstuki klass on kõige väiksema keskmise AP50 väärtusega: parim tulemus nõrgima klassi juures tähendab, et antud objekti tuvastusmeetod on hästi tasakaalustatud ning sobilik iga uurimisaluse klassi tuvastamiseks.

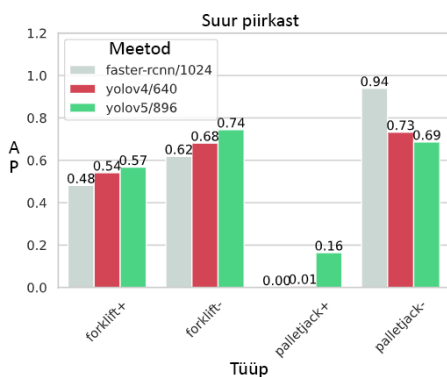
Võrdluseks võib tuua faster-rcnn mudeli, mis saab kõige paremini hakkama käsikahveltõstuki tuvastamisega, kuid hõivatud staatuses tõstuki tuvastamises on eelnimetatud mudel kokkuvõttes halvim.



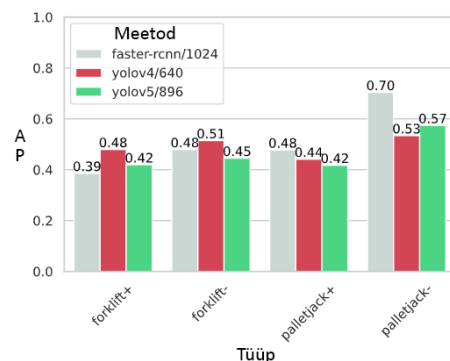
(a) Tõstuk hõivamata



(b) AP50, keskmine objekti piirkast.



(c) AP50, suur objekti piirkast.



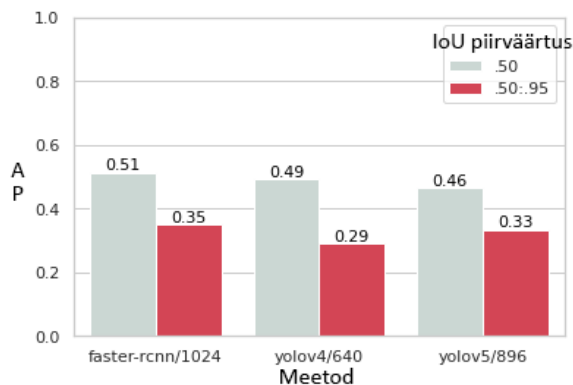
(d) Klassipõhine AP50

Joonis 8: Objektituvastuse täpsused.

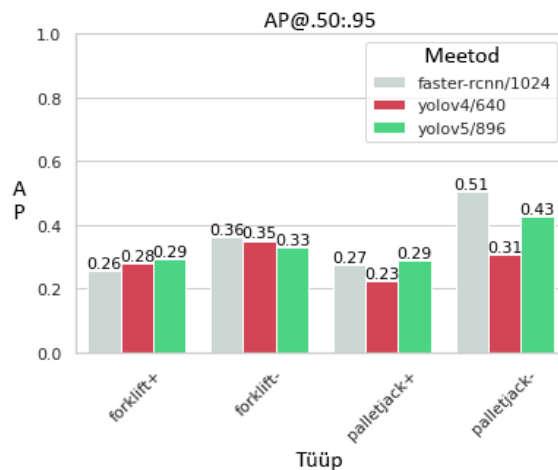
Teisest küljest peab välja tooma ka aspektid, kus YOLOv4 mudel võib jääda nõrgaks.

Esiteks ei ole YOLOv4 mudeli lokaliseerimise ehk paigustamise täpsus väga hea. See on illustreeritud joonisel 9, kust võib näha, et IoU .50:.95 juures on YOLOv4 kõige halvem. See ei ole aga eksperimendi tulemuste suhtes suure kaaluga, kuna nagu eelnevalt mainitud, paigustamise täpsus ei ole otsustava kaaluga. Küll aga tähendab halb mudeli paigustamine täpsete laosukohtade leidmisel suuremat viga.

Samuti ei ole valitud mudeli üldine AP50 tulemus kõige parem. See tähendab, et kõikides objekti klassides ei ole YOLOv4 mudel kõige parem. Näiteks oleks see oluline siis, kui eksperimendi asukohas oleks rohkem käsikahveltõstukeid ning vähem tavalisi tõstukeid. Kuna aga eelnimetatud klassid on uurimuses võrdse kaaluga, siis eksperimendi tulemusi see faktor ei mõjuta.



(a) AP kokku



(b) Klassipõhine AP

Joonis 9: Mudelite paigustamise võime vastavalt AP headuse kriteeriumile.

2.5 Objektijälgimine

Objektijälgimiseks kasutatakse modifitseeritud versiooni FastMOTist [30], mis baseerub DeepSORT [31] algoritmil. Lühidalt töötab eelnimetatud objektijälgija järgnevalt.

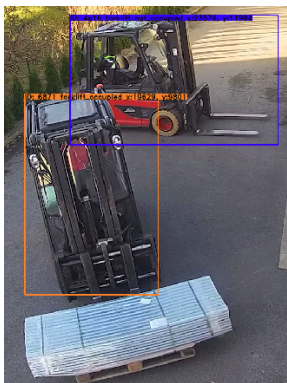
Esiteks tuvastatakse kaadris asuvate huvialuste objektide piirkastid kasutades objektituvastusmeetodit. Kasutatavale objektituvastusmeetodile siinkohal muid nõudmisi ei ole, seega sobib see käesoleva süsteemi jälgimismeetodiks hästi.

Samuti leitakse iga tuvastatud objekti kohta tema tunnusvektor kasutades närvivõrkudel põhinevat objekti taastuvastusmeetodit. Käesolevas töös kasutati sellise tunnusvektori mudeli treenimiseks FastReID tarkvara [32].

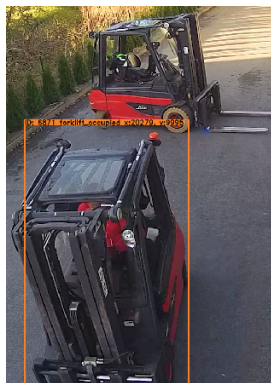
Objektid samastatakse, kui need on kahes järgnevas kaadris piisavalt sarnased, kusjuures sarnasust hinnatakse eelmise kaadri objekti tunnusvektori ning Kalmani filtrist tuleneva liikumise ennustuse põhjal. Rada määratakse kadunuks, kui etteantud aja vältel objekti ei suudeta taastuvastada.

2.5.1 Validatsioon

Käesolevas töös soovime peamiselt hinnata, kui kasutatav on kaamerate põhine laoliikumiste jälgimise süsteem. Sellest tulenevalt on peamine objekti jälgimise kriteerium toote tegelikust algasukohast tema tegelikku lõppasukoha jõudmise tuvastuse täpsus. See täpsus saadakse võrreldes käsitsi sildistatud andmetest tulenevat toodete asukoha muutuse täpsust eksperimendi tulemusena valminud süsteemi täpsusega. Valetuvastusteks loetakse siin kõiki toodete paigutusi, mis asuvad kaugemal kui 5 meetrit käsitsi sildistatud andmetest tulenevast toote asukohast.



(a) Kaader 1



(b) Kaader 2



(c) Kaader 3

Joonis 10: Näidis sama objekti raja muutusest. Kaamerale lähima tõstuki unikaalne tuvastus (oranz) omistatakse valele objektile.

Täpsemalt hinnatakse objektide jälgimise puhul järgmiseid omadusi:

- Radade vahetuvus, kus füüsiliselt üks objekt kaotab oma identiteedi (joonis 10).
- Kaamerate vaheline radade samastamine, kus objekti jälgimis algoritm peab suutma sama objekti kokku viia erinevate kaamerate vaatenurkadest.
- Erinevates kaamerates nähtava raja lahkulöömine juhul, kui eelnevalt viidi rada valesi kokku.

3 Lahendus

Järgnevalt on täpsemalt kirjeldatud lõpplahendust, mis valmis uurimustöö käigus ja mille toimimist kontrolliti empiiriliste katsetega.

Süsteemi põhiarhitektuur baseerub FastMOT objektijälgimise tarkvaral, kus kasutatakse käesoleva magistr töö raames treenitud laotransporditööde mudeleid objektituvastuseks ning taastuvastuseks. FastMOT üksi suudab objekte jälgida ainult ühe kaamera vaate põhjal, seega on lõpplahenduses DeepSORT algoritmi sisse viidud teatud muudatused.

Siinkohal peab mainima, et kuna algandmete videod on sünkroniseeritud vastavalt kõige suuremale kaamera kaadrisagedusele 25 kaadrit sekundis, siis kahes kaameras on mitmeid kaadreid, kus kahe järjestikuse kaadri tulemusel liikumist ei toimu. Sellistel juhtudel kasutatakse optilist voogu.⁷

3.1 Radade ühendamine ning lahkulöömine

Kõige suurem muudatus, mis süsteemi sisse viidi oli mitme videovoog töötlemine ning radade ühendamine. Selleks kasutatakse algoritmi 1, mis käivitub koos prototüüplahenduse programmiga. Algoritmi sisendiks on video objektide andmestruktuur, kus igas video objektis sisaldub videovoog, unikaalne FastMOT jälgija ning videovoole vastava kaamera kalibratsiooni parameetrid. Algoritmi 1 töö lõppeb, kui kõik videod on lõppenud. Oluliseim samm toimub algoritmis 1 real 10, kus igale video objektile vastavast FastMOT jälgijast tulenevad sama objekti rajad sünkroniseeritakse.

Rajad erinevate kaamerate vahel sünkroniseeritakse kasutades iga tuvastatud objekti vahelist kaugust. Täpsemalt, esmalt arvutatakse igas kaameras iga nähtaval oleva objekti vahel eukleidiline kaugus, kus objekti koordinaat leitakse tema piirkasti alumise keskmise punkti teisendamisel kaamerate vahel ühtsesse koordinaatsüsteemi.

⁷Tehnika objekti liikumise ennustamiseks, kasutades pikslite heledusi ning kahte järjestikust kaadrit.

Algorithm 1: FastMOT mitme kaameraga

Input: Video objektid $V = \{v_1, v_2, \dots, v_n\}$

```
1 while Kõik videod ei ole lõppenud do
2   for  $v$  in  $V$  do
3     if Eksisteerib järgmine kaader videos  $v$  then
4        $k =$  järgmine kaader videos  $v$ 
5        $t =$  FastMOT jälgija videos  $v$ 
6       Objektijälgija  $t$  uuenduse samm kaadril  $k$ 
7   if Ühtegi uut kaadrit ei leidu then
8     Kõik videod on lõppenud
9   else
10    Sünkroniseeritakse rajad videotest  $V$ 
```

Seejärel proovitakse lahku lüüa radasid, mis on valesti kokku pandud. Selleks kasutatakse lävendit 6,25 meetrit (5 m + 25%) ning kui sama identiteediga raja distantis erinevate kaamerate vahel ületab seda lävendit, siis antakse radadele uued identiteedid.

Järgnevalt viiakse kokku rajad kaamerate vahel, kasutades sama kauguste maatriksit, kuid seekord lävendiga 5 meetrit. Rada viiakse kokku kõige lähemal asuva rajaga, mis asub erinevas kaameras kui esialgne objekt, kusjuures alles jäetakse vanema raja identiteet. Kui kaks või enam rada on ühendatud, interpoleeritakse tema maailma koordinaat kõikide ühendatud radade alusel ning ühendatud raja maailma koordinaadiks igas kaameras saab tema aritmeetiline keskmine maailma koordinaat.

Lävendit 5 meetrit otsustati kasutada, põhinedes kaamera kalibratsiooni täpsuse katsete tulemustele, kus maksimaalne kalibratsiooni viga oli 5,2 meetrit. Maksimaalne kalibratsiooni viga valiti lävendiks, kuna kaamerate vaateväljas asuv füüsiliselt sama objekt võib olla nähtav erinevatest külgedest (joonis 16).

Näiteks kui tuvastatakse tõstuki tagumine külg ühes kaameras ning teises kaameras tõstuki käpad, siis arvestades tõstuki pikkust (3-3,5 meetrit) ning kaamerate kalibratsiooni keskmist viga 1,0 meetrit, on 5 m töö autori hinnangul mõistlik radade kokkuviiamise lävend.

3.2 Objekti staatuse leidmine

Objekti staatus arvutatakse vastavalt kõikide kaamerate staatuse hinnangule. Staatuse all mõeldakse siin laotranspordööri hõivatud või hõivamata staatust.

Hõivatud staatusesse määratakse objekt, kui vähemalt üks kaamera hindab objekti staatuse olevat hõivatud olekus. Selline lähenemine valiti, kuna hõivatud oleku tuvastamine on töö kontekstis kõige olulisem (et jälgida toote liikumist) ning hõivatud oleku tuvastustäpsus oli madalam kui hõivamata oleku täpsus.

Hõivatud olekus hoitakse objekti 2 sekundit enne, kui staatus muudetakse hõivamata olekusse, misjärel ennustatakse toote mahapanek. Teisisõnu, kui kahe sekundi jooksul ei ole objektile vähemalt ühest kaamerast ennustatud hõivatud olekut, siis muudetakse tema staatus hõivamata olekusse.

4 Süsteemi analüüs

Süsteemi headuse analüüs viiakse läbi visuaalse hindamise abil. See on võimalik, kuna valideeritavaid andmeid on võrdlemisi vähe. Visuaalne kontroll antud süsteemi valideerimiseks on mõistlik, kuna leitud lahenduse headust ei ole antud töö kontekstis vaja võrrelda erinevate lahendusmeetoditega.

Süsteemi edasiarendusel oleks aga mõistlik välja töötada numbrilised headuse näitajad, mis võimaldaks efektiivsemalt süsteeme omavahel võrrelda.

Täpsemalt analüüsitakse käesolevas peatükis järgmisi süsteemi headuse näitajaid: toodete kaotuste arv (objekti tuvastuse headus), jälgitavate radade kaotused (jälgimisalgoritm ei suutnud füüsiliselt sama objekti jälgida), kaamerate vaheline sama objekti tuvastus, kui kaugele toode jõudis enne, kui see kadus. Toodete kaotusi analüüsitakse vastavalt transportööride tüübile ning nende hõivatus staatusele.

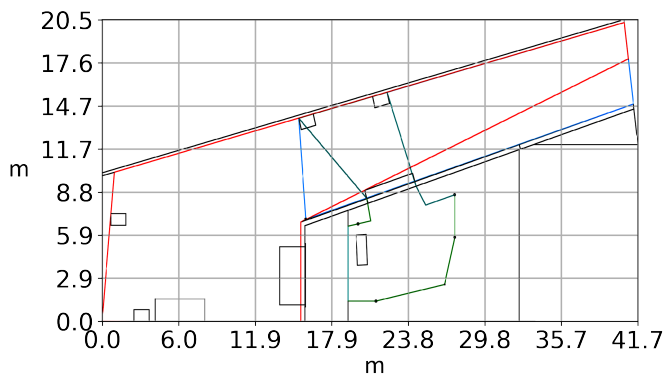
Järgnevalt analüüsitakse süsteemi headust kolme erineva stsenaariumi põhjal, kus igas olukorras osaleb kolm kaamerat. Kõik olukorrad leiavad aset tagahoovi kaamerate vahel, mille geomeetria ning visuaalne kattuvus on kujutatud joonisel 1.

- Stsenaarium 1: Tõstuk laadib tagahoovis alustel tooteid maha veokilt ning viib need siseruumi.
- Stsenaarium 2: Käsikahveltõstuk viib tooteid siseruumist tagahoovi ning samal ajal viib tõstuk tooteid siseruumist tagahoovi.
- Stsenaarium 3: Kaks tõstukit viivad tooteid siseruumist tagahoovi.

Jooniste legend Järgnevas stsenaariumite analüüsis on kõik radade joonised koostatud ühtsel põhimõttel ning sellest tulenevalt on legend siinkohal eraldi välja toodud.

Jooniste põhjaks on võetud analüüsitava ala joonis 11, kus mõõtmised on saadud otse hoone joonistest. Seega on arvestatud, et joonisel 11 mõõtmisviga puudub. Sellest tulenevalt

on radadel, mis liiguvad mitme kaamera vahel, lüngad sees. Need lüngad on näha nii süsteemi ennustatud radade joonisel kui ka käsitsi sildistatud andmetest tulenevatel joonistel.



Joonis 11: Analüüsi alusjoonis asukohas tagahoov. Kaamerate paigutused on nähtavad joonisel 1b.

Kui ei ole kirjeldatud teisiti, siis kõikidel radade joonistel on parempoolne kujutis parim tulemus, mida on võimalik saavutada täiusliku objektituvastusalgoritmi korral. Selle alamjoonise sildistused on saavutatud käsitsi laotransportööride annoteerimise teel. Vasakpoolsetel radade joonistel on kujutatud töö tulemusena valminud süsteemi ennustusvõimet ning töö lugejal on võimalik seda alamjoonist võrrelda parima võimaliku tulemusega paremal.

Samadel joonistel on kujutatud ka sündmuseid erinevate värvide ning ringidega. Joone värvitooni vahetus näitab unikaalsete laotransportööride tuvastusi ning valetuvastusi, kus värvi küllastus tähistab transportööri staatust: erk värv märgib hõivatud staatuses transportööri ning tuhm värv hõivamata staatuses laotransportööri.

Ringe võib eelnimetatud joonistel näha suuri ning väikseid, täidetud ning seest tühjasid. Suured ringid tähistavad kaameras nähtavaid toote tegelikke peale- ning mahalaadimisi. Need andmed on käsitsi sildistatud ning toodud võrdluseks süsteemis kasutatava algoritmi leitavate andmetega.

Väikesed ringid näitavad tootetuvastusalgoritmi poolt tuvastatud toodete peale- ning mahalaadimisi. Täidetud suured ning väikesed ringid märgivad ennustatud toodete pealelaadimisi ning tühjad ringid tähistavad toodete mahalaadimiste ennustusi. Joonistel võib näha ka olukordi, kus parempoolsetel joonistel algoritmi poolt tuvastatud toodete paigutuste asukoha lähedal ei asu suurt ringi ehk tegelikku toote paigutust. See tähendab, et tegelik toote peale- või mahalaadimine laotransportöörile toimus uuritava ala kaamerate vaatevälja väliselt.

Radade joonistel on võimalik hinnata nelja erinevat viga:

- Kaamerate kalibratsioonist tulenev viga - suured lüngad radades, mis tulevad eriti esile mitme kaamera kattuvuse piiridel.
- Kaameratevahelise sama laotransportööri tuvastusviga - raja värvi vahetus, mis tuleb samuti esile kaamera kattuvuse piiridel. Kui raja värv vahetub tagasi algseks raja värviks, siis seda veaks ei loeta. Näiteks vasakpoolisel joonisel 19b roheline rada taastuvastatakse tumepunasena.
- Laotransportööri identiteedi kaotusviga - raja värvi vahetus.
- Toodete pealevõtmise ning maha paigutamise viga - valetuvastused on kõik tühjad ning täidetud väiksed ringid, mis ei ole jõudnud käsitsi sildistatud andmetest tulenevate toodete paigutamise asukohast viie meetri raadiusesse.

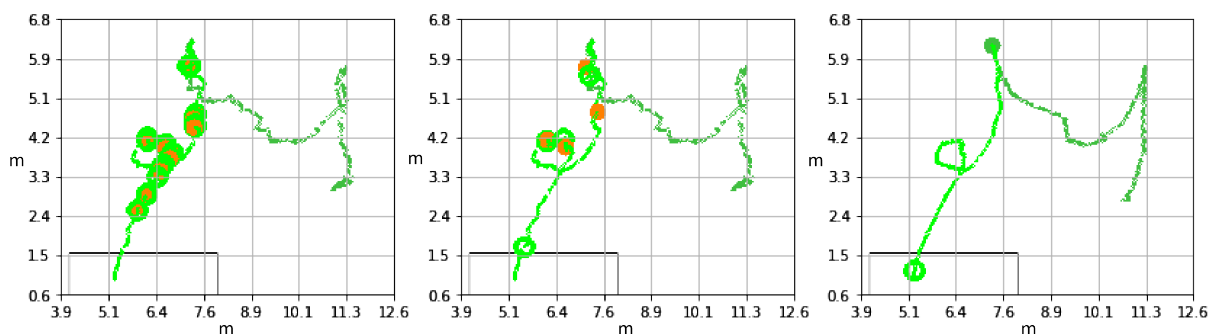
4.1 Stsenaarium 1

Esimene olukord hõlmab endas järgnevaid sündmuseid: tõstuk laeb alustel tooteid veokilt maha 4 korda, millest 2 toodet liigub kaamerate vaateväljast välja ning 2 toodet liigub kolme kaamera vahel. Käsikahveltõstuk ilmub kaadrisse kandes toodet ning järgnevalt võtab peale tõstuki poolt maha pandud toote.

Stsenaariumis 1 on võimalik peamiselt hinnata toodete kaduvust, ehk kui tihti toode jõudis tegelikku lõpp-punkti ning mis olid lõpp-punkti jõudmise probleemid.

Joonisel 12 on kujutatud järgnev olukord: hõivamata tõstuk liigub veokini (tuhmroheline rada), mis järel võetakse peale toode (tuhmroheline täpp parempoolsemal joonisel) ning viiakse toode siseruumi (erkroheline rada). Oranž täpp tähistab toote pealevõtmist olukorras, kui tegelikult transportöör juba on hõivatud staatuses ning roheline seest tühi ring tähistab toote kaotamist.

Kõige olulisem toodete kaduvuse analüüsi kohalt on roheline raja vastavus käsitsi sildistatud rajaga. Selle kaudu on võimalik hinnata, kas tõstuk leidis toote ning kui kaua toode püsis tõstukil enne, kui tehti valetuvastus ning toode kaotati.

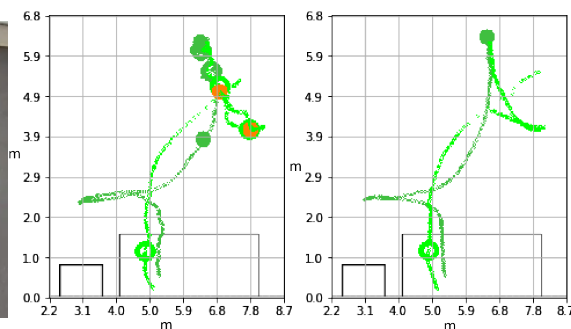


Joonis 12: Tõstuk laadib toodet veokilt maha. Vasakul on kujutatud tõstuki staatuste muutumised, keskel on kujutatud filtreeritud tõstuki staatuste muutused ning parempoolsel joonisel on illustreeritud käsitsi sildistatud staatuste muutumised.

Joonisel 12 toodud olukorra põhjal võib näha, et tühja tõstuki jälgimine töötab hästi ning toodud situatsioonis toodete üles korjamisi, mida tegelikult ei esinenud, ei tuvastata. Samas kui toode on peale võetud, kaotatakse see üsna kiiresti peale tõstuki ringi keeramist.



Joonis 13: Vale toote mahapaneku tuvastus.



Joonis 14: Tõstuk laadib maha järgmist toodet.

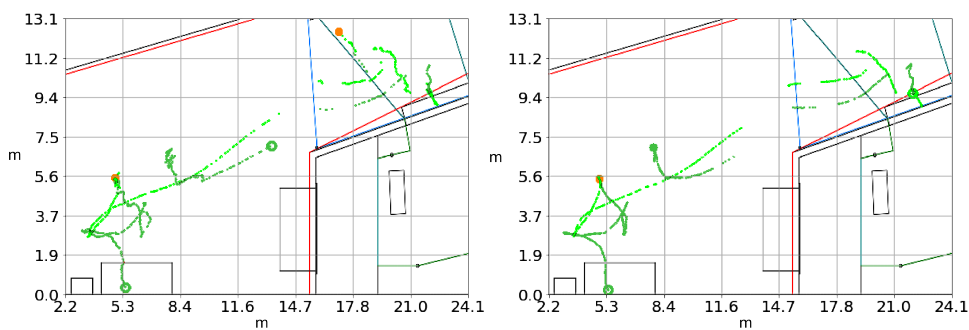
Joonisel 13 on kujutatud tõstuk, kus objektituvastus ei suutnud tõstukil olevat toodet tuvastada ning toode läks kaduma. Tõstuk seisis asendis, kus selle käpad olid nähtavad ligemale 20 sekundit. Võib eeldada, et just käppade nähtavus kaadris viis objektituvastaja valele teele, mistõttu toode kaotati.

Võrdluseks võib tuua olukorra, mis on kujutatud joonisel 14. Esiteks võib siin näha, et toote leidmine on sellel korral palju ebakindlam, kui oli seda olukorras, mis on illustreeritud joonisel 12. Seda illustreerib tumeroheliste täidetud ringide rohkus. Kui aga vaadelda toote hoidmise sooritust, siis selles situatsioonis kaotatakse toode vaid kahel korral, kus esimesel korral tõstuk on füüsiliselt liikumas.

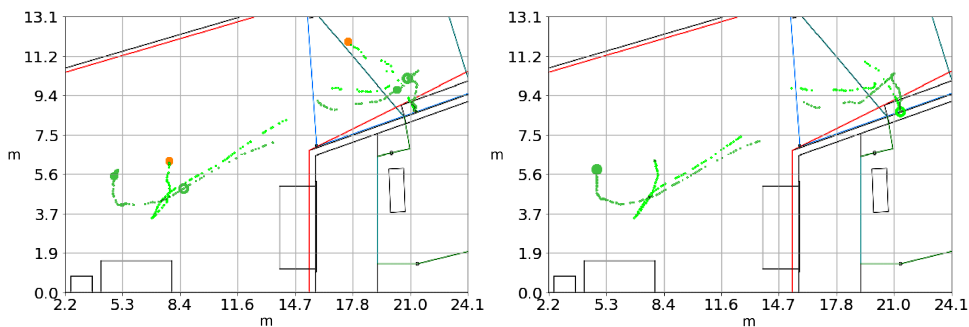
Teisest küljest võib jooniselt näha, et peale toote kaotust taastuvastatakse toode väga

väikese vahega (oranž täpp): esimene kord kulub selleks kolm kaadrit peale kaotust, teine kord tuvastatakse toode üks kaader peale toote kaotust. Selliseid valetuvastusi on võimalik vältida täiustades toote hoidmise algoritmi nii, et see võtaks arvesse laotransportööri liikumist ühtses maailma koordinaatsüsteemis: reeglina kiirema liikumise pealt tooteid üles ei korjata ega ka panda maha.

Joonisel 15 on illustreeritud 2 sarnast sündmust, kus tõstuk liigub läbi 3 kaamera.



(a) Alasündmus 1.



(b) Alasündmus 2.

Joonis 15: Tõstuki liikumine kaamerate vahel.

Esiteks võib illustratsiooni 15 alamjoonistel 15a ning 15b näha, et isegi käsitsi sildistatud rada ei ole ühtlane.

Kaamerate vaheline radade ühendamine toimus käesolevas stsenaariumis hästi. Kokku toimus 8 radade ühendamist, millest 6 toimus 0 kaadrise viivitusega, 1 toimus 19 kaadrise viivitusega ning 1 30 kaadrise viivitusega. Samas peab mainima, et kuna käesolevas olukorras osales ainult 1 kaamerate vahel liikuv tõstuk, siis sellest radade ühendamise analüüsist ei ole mõistlik täpsemaid järeldusi teha.

Stsenaariumis 1 toimus ainult 1 raja vahetus. Sellele on ka lihtne leida põhjus: kuna raja lahkulöömise algoritm võtab arvesse ainult objekti positsiooni ühtses koordinaatsüsteemis, siis

kaamerate kalibratsiooni täpsusest ning kaameranurkadest olenevad vead mõjutavad tulemust palju. Olukord, kus objekti jälgija löi raja valesti lahku, on kujutatud joonisel 16.

Kalibratsiooni ning kaamera vaatenurgast tulenevat viga võib näha ka joonise 15 alamjoonistel 15a ning 15b. Parempoolsetel joonistel on näha, et isegi käsitsi sildistatud rada ei ole eelnimetatud asjaoludest sõltuvalt ühtlane.



(a) Kaamera 7: Tagumine hoov



(b) Kaamera 6: Tagahoov Uks nr.1

Joonis 16: Objekti rada lüüakse lahku.

Joonisel 16 tuvastab kaamera 6 tõstuki tagaosas ning kaamera 7 tuvastab tõstuki esiosa. Samuti on kaamerate kattuvus alas, kus kaamera 7 kalibreerimisviga on 2,07 meetrit, mis on võrreldes tõstuki mõõtmetega võrdlemisi suur. Eelnevast tulenevalt viiakse rajad antud situatsioonis lahku.

4.2 Stsenaarium 2

Olukorras 2 toimuvad järgnevad sündmused: käsikahveltõstuk ilmub kaadrisse koos tootega kaameras 5 ning liigutab selle tagahoovi kaamera 7 vaatevälja. See toimub kokku neli korda, kusjuures kolm esimest korda veab käsikahveltõstuk suuri kaste, mis peidab turvakaamerate vaatenurgast laotranspordööri ning neljas kord veab käsikahveltõstuk madalat alust.

Samuti osaleb olukorras kaks tõstuk, mis sõidab kaamera 7 vaateväljast tühjana kaamera 5 vaatevälja, võtab sealt peale toote ning viib selle tagasi kaamera 7 vaatevälja. Samal ajal liigub tõstuki kõrvalt mööda käsikahveltõstuk. Juhtum lõppeb sama tõstuki uue toote peale võtmisega kaamerast 5, misjärel tõstuk lahku ruumist koos tootega.

Selles stsenaariumis hinnatakse peamiselt radade taastuvastusi ning toodete kaduvust. Olukorrad, kus taastuvastused oluliseks muutuvad, on siis, kui käsikahveltõstuk liigub

kaamerate vahel suure tootega (joonis 17) ning kui käsikahveltõstuk liigub tõstuki tagant läbi nii, et ühes kaameras on eelmainitud laotransportöör nähtav, kuid teises mitte (joonis 18).



Joonis 17: Käsikahveltõstuk liigub kaamerate vahel, kus ühes kaameras on see nähtav ning teises toote taga.



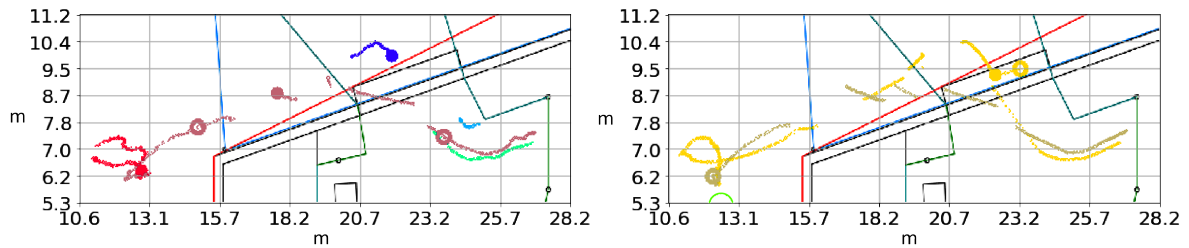
Joonis 18: Käsikahveltõstuk liigub tõstuki tagant läbi.

Olukorras 2 käsikahveltõstuki poolt liigutatud tooted jõuavad lõpp-punkti lähedusse ühel korral neljast. Samas on selgelt näha juhud, kus ning miks süsteem ei suutnud jälgida toote kaameratevahelist liikumist.

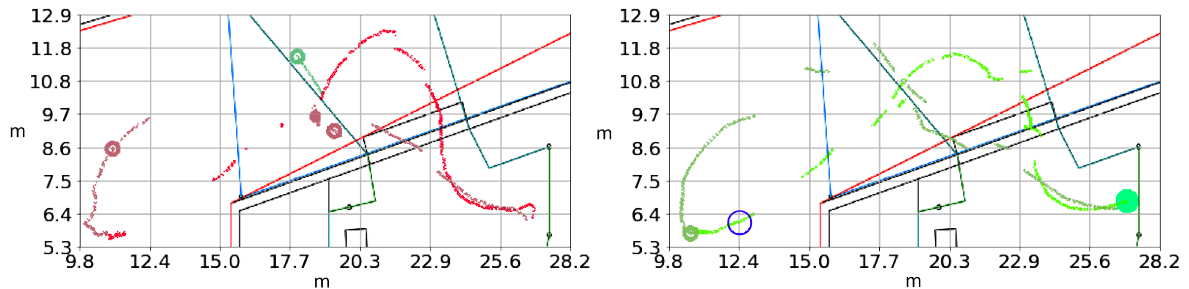
Täpsemalt, kõik kolm toote kaduvust on sellel juhul sarnased ning põhjus on näha joonisel 17. Nimelt katab toode käsikahveltõstuki kaamera 5 vaateväljas piisavalt kauaks, et nii kaamera 6 kui ka 7 vahel ei kattu piisavalt kaadreid, et radasid ühendada. See on illustreeritud joonisel 19a. Samalt jooniselt on võimalik ka näha, et isegi ideaalse objektituvastuse täpsuse korral valitud algoritm ei tööta: käsitsi sildistatud toode läheb samas kohas kaduma.

Joonised 19 ning 20 vajavad täpsemat selgitamist. Esiteks on vasakpoolsetel joonistel 19a, 19b, 20a ja 20b välja toodud identiteedi vahetused erinevate värvitoonidega (vasakpoolsetel joonisel 19a roheline, sinine ning punane). Värvide erksus märgib transportööri hõivatud staatust: erk värv tähendab käsitsi sildistatud hõivatud staatust ning tuhm värv tähendab sama identiteedi mittehõivatud staatust.

Sarnaselt stsenaariumis 1 näidatud joonistele 15 ja 14 märgib ka siin täidetud väike täpp ennustatud toote pealevõtmise asukohta, suur täpp tegeliku toote pealevõtmise asukohta, väike



(a) Vähemalt kahes kaameras ei tuvastata käsikahveltõstuki.



(b) Käsikahveltõstuk nähtava ühes kaameras ning peidetud teises.

Joonis 19: Käsikahveltõstuki liikumise võrdlus kaamerasuhtes peidetuse korral.

ring ennustatud toote mahapaneku asukohta ning suur ring tegelikku toote mahapaneku asukohta.

Vasakul on tuvastatud rada, identiteedi vahetused ning toodete paigutamised ning peale võtmised, parempoolsel joonisel on sama informatsioon käsitsi sildistatud. Erksa värviga rada tähistab laotransportööri liikumist hõivatud staatuses ning tuhm, sama tooni rada tähistab hõivamata staatust.

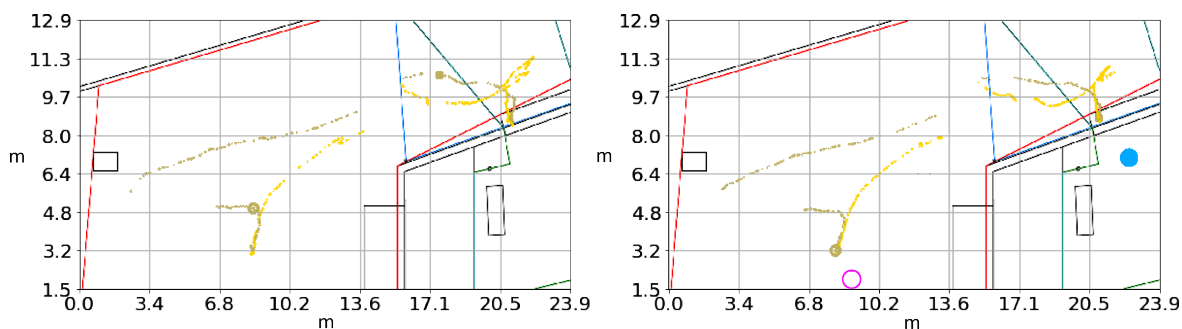
Seega joonisel 19a on illustreeritud laotransportööri identiteedi vahetused. Kui üldjuhul radade kokkupanemise algoritm töötab hästi, siis olukorras, kus kaks kaamerat ei näe sama objekti korraga, ei suudeta rada ka kokku panna. Siin võiks näiteks laotransportööri trajektoori ennustuse põhjal radade kokkupanemise algoritmi täiustada.

Neljas käsikahveltõstuki toote liigutamise olukord (joonised 18 ning 19b) on sarnane esimese kolmega, kus ainuke vahe on selles, et kantava aluse väiksuse tõttu on käsikahveltõstuk kaameras 5 nähtav kuni kaamera 6 ning järgnevalt kaamera 7 vaatevälja ilmumiseni. Sellisel juhul valetuvastatud toote mahapanekuid ei ole ning toode jõuab tegelikust lõppasukohast 3 meetri kaugusele.

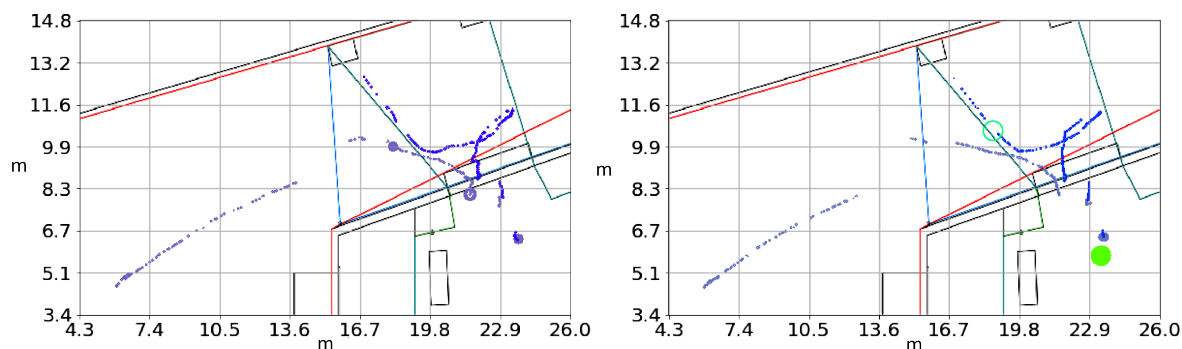
Samuti on joonisel 19b näha radade ühendamise algoritmi järjekordne puudjääk. Isegi kui

rajad ühendatakse kaamerate vahel, siis seda tehakse ainult edasiulatuvalt. Selle tagajärg on toote paigutuse valetuvastused.

Näiteks võib eelnimetatud joonisel näha väikest rohelist ringi ning tuhmrohelist rada, mis märgib huvialuse objekti tuvastust mitte hõivatud staatuses, ning selle liikumist. Samas see rada ühendatakse järgnevalt punase raja identiteediga, muutes toote asetustuvastuse üleliigseks valetuvastuseks.



(a) Alasündmus 1.



(b) Alasündmus 2.

Joonis 20: Tõstuk liigutab toodet kaks korda.

Vaadates tõstuki liikumist joonisel 20 võib näha, et tõstukil asuv toode jõuab mõlemal juhul tegeliku asukoha lähedale. Siin esines valetuvastatud toote üles korjamine juhul, kui tegelikult tõstukil toodet ei asunud, kaks korda: mõlemal korral samas kohas ning põhjusel, et kaameras 7 oli näha vaid tõstuki tagumine ots. Kui tõstuk ilmus kaamera 6 vaatevälja, tuvastati tõstukil jälle mittehõivatud staatus korrektselt.

4.3 Stsenaarium 3

Kolmandas, ning ühtlasi ka viimases videovoos viivad tõstukid tooteid siseruumist veoautole, kusjuures veoauto jääb kaadrist välja. Selline tegevus toimub 5 korda, millest kolmel korral seisab üks tõstuk paigal.

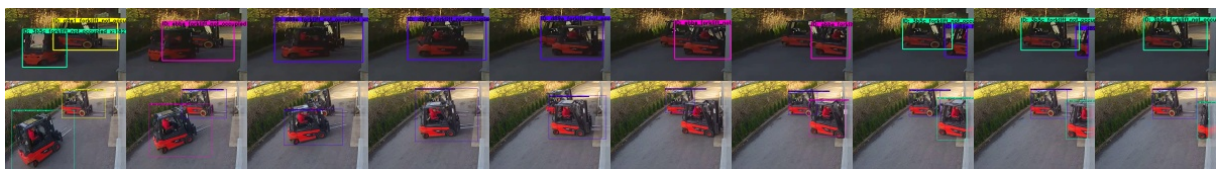
Samuti esineb kolmandas stsenaariumis 5 korda olukord, kus sama toode pannakse maha ning järgnevalt korjatakse üles 90 kraadi pööratult. Eelnimetatud olukord toimub 4 korda samas asukohas ning kõikide kaamerate vaate suhtes samas asukohas.

Täpsemalt pannakse toode maha üks kord käsikahveltõstuki poolt ning korjatakse üles tõstuki poolt, kolm korda pannakse toode maha, misjärel korjatakse see sama tõstuki poolt üles. Viimane toote keeramine toimub kaameras 7 ning kaadri nurgas, kus on kaamera kalibratsiooni viga suurim.

Stsenaariumi 3 teeb huvitavaks siinkohal asjaolu, et korraga on nähtaval kaks identset tõstukit ning ühel juhul ka käsikahveltõstuk. Analüüsides sellist olukorda, saab anda hinnangu objektijälgija tööle segavate objektide korral, milleks antud juhul on seisev tõstuk.

Seega, kolmanda stsenaariumi põhjal hinnatakse peamiselt huvialuste objektide identiteedi vahetusi. Samuti antakse hinnang süsteemi headusele olukorras, kus tõstuk paneb toote maha ning järgnevalt korjab selle sama toote üles.

Kokku oli käesolevas stsenaariumis kokku 42 identiteedi vahetust, kus objekti rada kadus või objekt millele rada omistati, vahetus. Tõstukite puhul kaotati rada objektituvastuse parameetrite tõttu, väikses alas suudeti tuvastada ainult üks objekt. Üks selline olukord on illustreeritud joonisel 21.



Joonis 21: Kahe tõstuki kattuvuse näidis.

Käsikahveltõstuki puhul toimus objekti identiteedi vahetus samal põhjusel, mis stsenaariumis 2 suurte toodete liigutamisel. Tööline lükkas toodet enda ees ehk

kaamerasilmale oli käsikahveltõstuk piisavas ulatuses peidetud ning kahe kaamera vahel ei suudetud vaadeldava objekti identiteeti kokku viia.

Kaduma läks tooteid kuuel korral, millest kõik juhtumid olid põhjustatud raja identiteedi muutusest. Samas peab mainima, et kui toode maha pandi ning ümber keerati tõstuki suhtes, siis ühel korral ei suudetud tuvastada sellist tegevust (ümber keeramist).

Analüüsidest ainult objektide staatuseid ning mitte arvesse võttes objektide identiteedi vahetusi, olid kolmanda stsenaariumi tulemused kõige paremad. Täpsemalt siis toode kaotati tuvastustäpsuse tõttu vaid ühel korral, kui käsikahveltõstuk oli kaamera suhtes peidetud kauem, kui süsteem rada elus hoidis.

Samuti tuvastati toodete pealevõtmise asukohad siin kõige täpsemalt. Kui mitte arvestada objektide identiteedi vahetustest tulenenud toodete pealevõtmiste ning mahapanekute ennustusi, võeti toode peale keskmise täpsusega 1,5 meetrit, mis on eeldatavat süsteemi täpsust arvesse võttes väga hea tulemus.

Võrdluseks võib tuua tõstuki käppade pikkuse 1,2 meetrit ning käsikahveltõstuki käppade pikkuse 1,15 meetrit. Meeldetuletuseks lugejale, objektituvastusmeetodit treeniti nii, et kui tõstuki ning käsikahveltõstuki käpad on piisavalt palju toote poolt peidetud, määratakse laotransportöör hõivatud staatusesse.

4.4 Üldine hinnang

Esiteks tuleb mainida, et koostatud analüüs ei ole täielik, kuna analüüsida oli võimalik ainult piiratud süsteemi osa. Näiteks kaamerate kattuvuse erinevuse, kaamerate kalibratsiooni erinevuse ning tuvastavate objektide arvu erinevuse korral on süsteemi headus tõenäoliselt erinev ning vajab täiendavat uurimist. Selliste parameetrite muutmine antud töö raames ei olnud võimalik.

Sellel põhjal on töö autori hinnangul koostatud analüüsist saadud järelduste põhjal võimalik hinnata väljatöötatud lahenduse erinevaid tugevaid ning nõrkasid kohtasid ning anda adekvaatne hinnang süsteemi edasiarenduse lubavamateks suundadeks. Sellisele järeldusele jõuti peale erinevate stsenaariumite analüüsi: vead süsteemis tekkisid väga sarnastel põhjustel

ning samuti nende vigade elimineerimine hõlmab suures osas samu samme.

Veel lisaks tuleb mainida, et erinevates valgus- ning ilmastikutingimustes süsteemi headust ei hinnatud. Selle põhjused olid järgnevad. Esiteks ei olnud töö kirjutamise ajal võimalik koguda piisavalt erinevates valgustingimustes andmeid ning teine põhjus oli esimesega otseselt seotud: väga halbades ilma- ning valgustingimustes üldjuhul laoliikumisi ei esine (pimedad ruumid, valgustamata õuealad, vihmatormid, lumetormid jms).

Hüpotees, et laoliikumisi on võimalik jälgida turvakaamerate abil loetakse töö autori poolt kinnitatuks. Seda kinnitavad ka katsete tulemused: tõstukite ning käsikahveltõstukite hõivatud ning hõivamata staatuses tuvastus töötas üldiselt hästi.

Kuigi objektijälgimise katsetustes võib näha süsteemis suuremaid piiranguid, võib siiski ka objektijälgimise edukaks lugeda. Näiteks on võimalik paranda valepositiivsete toodete mahalaadimise algoritmi, võttes arvesse fakti, et kui laotransportöör liigub kiiremini kui etteantud lävend, siis suure tõenäosusega laotransportööri hõivatuse staatus muutuda ei saa.

Välja võib siinkohal tuua näidiskasutusasukoha, kus arendatud süsteem adekvaatselt toimib - väheste samaliigiliste laotransportööridega ruumid, kus kaamerad näevad ühte objekti kahe erineva nurga alt. Selline olukord väldib süsteemi peamisi nõrkuseid: mitme samaliigilise laotransportööri tuvastamine väiksel alal ning samuti minimeerib kahe kaamera kattuvus transportööril asuva toote kaduvust kaamera vaatenurga tõttu.

4.5 Süsteemi kulu

Järgnevalt on välja toodud võimalikud kuluallikad, mis lahenduse kasutuselevõttu mujal kui eksperimendi asukohas mõjutada võivad.

Esiteks peab mainima, et väljapakutud tehisenägemise süsteemi kogukulu oleneb suures osas kasutusasukohas juba olemasolevast turvakaamerate infrastruktuurist. Kui kasutusasukohas on turvakaamerate infrastruktuur paigas, siis on võimalik koheselt väljapakutud süsteemi headust testida andmetöötluseks pilveteenuseid kasutades. Võrdluseks võib tuua raadiolainetel põhinevad laojälgimise süsteemid, mis vajavad suuremahulisi algseadistuse investeeringuid [12].

Kui aga turvakaamerate infrastruktuur soovitavas laos asukohas puudub, siis on ka väljapakutud süsteemi algseadistuskulu võrdlemisi suur, ulatudes tuhandettesse eurodesse [33]. See investeering võib siiski ennast kordades ära tasuda juba turvakaalutlustel. Võimalus kasutada laoliikumiste jälgimise süsteemi oleks sellisel juhul vaid üheks sellise investeeringu motiveerivaks faktoriks.

Süsteemi kulu mõjutab ka kaamerate arv. Rohkematel kaameratel võib olla mitmeid põhjuseid: näiteks hoone arhitektuur tekitab palju pimenurkasid, kaamera vaateväli on piiratud kõrgete riiulite tõttu või kasutatakse ülilainurk kaameraid, mille nägemiskaugus on piiratud. Samuti on kaamerate rohkus otseselt seotud väljapakutud süsteemi täpsusega, kus rohkemate kaamerate korral on valetuvastusi ning toodete kaotuseid vähem.

Veel mõjutab süsteemi kulu soov andmetöötlus läbi viia kohalikus võrgus või reaalajas. Peamine põhjus, miks andmetöötlus peaks kohalikul riistvaral toimima on jällegi seotud turvakaalutlustega: turvakaamerates nähtav pilt võib tihti olla tundliku sisuga, mida ei soovita kohalikust võrgust väljapoole saata. Reaalajas toodete liikumise jälgimise põhjuseks võivad olla näiteks olukorrad, kus ettevõtte kliendid saavad toodetele lattu ise järele tulla.

Ka siin oleneb süsteemi kulu kaamerate arvust ning kasutatavast riistvarast. Võrdluseks võib tuua, et eksperimendis kasutatava riistvaraga eksperimendi asukohas ühe 1920x1080 kaadri töötlus võttis keskmiselt aega 66,5ms, millest objekti tuvastus ning taastuvastus võttis 64,8ms ehk 97,4%. Näiteks YOLOv4 autor on defineerinud reaalajaliseks objektituvastuseks 30 kaadrit sekundis ehk 33,3ms [22].

Süsteemi jõudlust kiiruse mõistes on võimalik vähese vaevaga märkimisväärselt parandada investeerides kallimasse riistvarasse. Sealjuures tuleb tuua välja, et parimad töö kirjutamise ajal saadaolevad tavatarbijale mõeldud videokiirendid⁸ on oma hinna poolest konkurentsivõimelised ühe passiivse raadiosidelaine ehk RFID lugeja lisamisega RFID põhisesse laoliikumiste jälgimissüsteemi [13].

Haldamiskulu ning ajakulu turvakaamerate põhisel süsteemil on võrreldes RFID süsteemiga väike. Selleks on kaks põhjust. Nii väljapakutud lahenduse kui ka turvakaamerate süsteemi hooldust saab viia läbi samaaegselt: olenevalt kasutuskeskkonnast võib aegajalt

⁸GeForce RTX 3090 soovituslik jaehind \$ 1499 - <https://www.nvidia.com/en-us/geforce/graphics-cards/30-series/rtx-3090/>

osutada vajalikuks kaameraid puhastada tolmust või muust mustusest. Pigem tuleb sellist hooldamist ette väga harva ning enamus ajast on süsteem võimeline töötama ilma täiendavat tähelepanu nõudmata. RFID põhises süsteemis on aga vaja igale jälgitavale objektile lisada silt, või olemasolevat silti ümber häälestada, mida on üldjuhul automatiseerida keeruline või ressursimahukas.

Et aga tuvastada konkreetsete toodete või tellimuste liikumisi on ka kaamerapõhise laojälgimissüsteemi kulu suurem. Selleks on lahendused järgnevad. Tootmisetevõtetes võib ühendada automatiseeritud tööpingid väljapakutud süsteemiga, misjärel süsteem saab automaatselt konkreetse toote teabe. Kõige lihtsam lahendus selleks on luua kaamera vaatevälja ala, kuhu automatiseeritud tööpink tooteid paneb ning kui toote liikumine tuvastatakse, tehakse päring ala kohta, kust toode üles korjati.

Palju kulukam on kasutada manuaalset toote tuvastust. Kuigi võrreldes raadiosidelainetel põhinevate süsteemidega on sellisel juhul võimalik kasutada näiteks nutitelefoni spetsiifilise riistvara asemel, peab inimene selleks siiski aega kulutama.

5 Piirangud ning täiendused

Käesolevas töös on kirjeldatud laoliikumiste jälgimise prototüüplahendust. Magistritöö mahu piirangute tõttu on väljapakutud lahenduses mitmeid puudujääke, mida ei olnud töö keerukuse tõttu võimalik läbi töötada. Seetõttu on järgnevalt välja toodud olulisemad piirangud ning võimalikud lahendused, kuidas nende mõju vähendada.

5.1 Tuvastatavate transportöride iseärasused

Üheks olulisemaks piiranguks tulemuste analüüsil on andmete piiratus. Selletõttu ei ole võimalik prototüüplahenduses tuvastada elektrifitseeritud käsikahveltõstukeid. Kaamerate asukohtadest tulenevalt ei olnud võimalik koguda piisavalt videomaterjali, et anda adekvaatne hinnang eelnimetatud laotransportööri liikumise tuvastustäpsusele.

Samuti ei uuritud käsitsi toodete liigutamise jälgimisi. Eksperimendi asukohas liiguvad tooted üldjuhul partiidena, mida peab liigutama alustel kasutades laotransportööre. Seega ei loetud eelnimetatud laoliikumisi magistritöö ülesande sisse.



Joonis 22: Näidised erinevatest laotransportöridest.⁹

Täpsemalt saaks suuremat variatsiooni laotransportöridest tuvastada treenides objektituvastusmudelit rohkematel andmetel, fookuseerides andmete lisamisele erinevatest

⁹Allikas: <https://info.japanesecartrade.com/content-item/916-japan-forklift-machinery>

kaameranurkadest ning erineva värvi ning kujuga laotransportööridest. Kui käsikahveltõstukid erinevad üldjuhul ainult värvilt, siis tõstukid võivad erineda ka välimuselt. Näidised erinevatest laotransportööridest on välja toodud joonisel 22.

Viimaseks peab mainima, et uues asukohas süsteemi ülesseadmisel on võimalik objektituvastuse mudelit täiendavalt treenida, seejuures parandades süsteemi täpsust. See on võimalik, kuna üldjuhul on ladudes kasutatavad turvakaamerad fikseeritud vaatenurgaga, mis tähendab, et ühes kaamera vaateväljas on pimenurgad alati samades asukohtades. Objektituvastusmudelit täiendavalt treenides on võimalik sellistes kaadri asukohtades asuvaid laotransportööre täpsemini tuvastada. Samuti aitab mudeli täiendav treenimine kasutusasukoha kaadritel arvestada kohalike laotransportööride eripäradega, nagu eri värvi tõstukid ning ka näiteks kulunud värviga tõstukid. Selline mudeli täiustus aga jällegi tõstab süsteemi ülesseadmiskulu ning seetõttu võib olla olenevalt kasutusasukohast ning täpsusnõuetest ebapraktiline.

5.2 Koordinaatide teisendus

Kaadri uv koordinaatide teisenduse täpsust maailma koordinaatidesse XY mõjutab enim reaalmaailma koordinaatide saadavuse täpsus, nende kaugus kaamerast, algkaadri resolutsioon ning üldjuhul ka kaamera kalibratsioon. Kalibratsioon ei mängi rolli vaid juhul, kui on saadaval täpne teisendus iga videokaadri piksli ning reaalmaailma koordinaatide vahel.

Objekti kaugus reaalkoordinaatide mõõtühikus ning algkaadri resolutsiooni suurus on omavahel seotud. See tähendab, et kui huvialune punkt on kaamera vaateväljas kaugemal ehk objekt on projekteeritud kaadrisse väiksemalt, siis saab viga vähendada kasutades suurema videoresolutsiooniga kaamerat. Suurem kaamera lahutusvõime annab suurema võimaliku pikslite vahemiku ning see on oluline, kuna pikslite koordinaate saab täpselt mõõta ainult täisarvuliselt.

Kaamera kalibratsioonist tulenevat viga võib vähendada kasutades väiksemate moonutustega objektiiv. Nagu võib näha võrreldes jooniseid 2a ja 2b on moonutustest tulenev viga suurim kaadri servades, kus objektiivist tulenev moonutus on suurim.

Samas on suur kaadri moonutus ka oodatud, võttes arvesse eksperimendi eesmärki

kasutada turvakaameraid, kuna turvakaamerad on enamasti laia vaatenurgaga ning seega ka suure moonutusega. Näiteks võib tuua Godbolt kompileeritud nimekirja Amazoni parimate ülevaadetega turvakaameratest, kus 7 kaamerat 10st on üle 100 kraadise vaatenurgaga [34].

Teisalt võib kaamera kalibreerimine osutada uues süsteemi kasutuskohas üldse ebapraktiliseks või võimatuks. Näiteks võib tuua siin turvakaamerad, mis on paigutatud raskelt ligipääsetavatesse asukohtadesse (kõrgel lagede all, tööstuses tööpinkide läheduses mida ei ole mõistlik plaaniväliselt sulgeda jms).

Süsteemi algseadistus ehk kalibratsioon ongi töös arendatud süsteemi üheks peamiseks kasutuselevõtu takistuseks. Olles küll lihtsam ning rahaliselt odavam kui raadiosidelainetel põhinevad süsteemid (eeldusel, et kasutuselevõtu asukohas on olemas turvakaamerad), jääb siiski väikeste ettevõtete kartus muutustele.

Näiteks võib tuua uuringu, kus küsitledes väikseid ehitusettevõtteid leiti, et peamiseks takistuseks uute tehnoloogiate kasutuselevõtul sellistes firmades on esiteks uute lahenduste usaldusväärsus ning kasutuselevõtu kiirus ja silmnähtav kasu [35].

Siin võivad appi tulla tehisnärvivõrkudel baseeruvad mudelid, mille põhjal on teoreetiliselt võimalik kaamerasisene kalibratsioon kaadris asuvate moonutuste väljaviimiseks. Näiteks DeepCalib meetodit [36, 37] kasutades peaks kaamerasisene kalibratsioon toimuma vaid kaamera kaadrit meetodile näidates. Oluline on siinkohal veel mainida, et eelnimetatud tehisnärvivõrgu põhise kalibreerimise meetodi kasutatavust käesoleva töö raames valminud süsteemiga ühendamiseks peaks täiendavalt uurima: DeepCalib kasutab moonutuste modelleerimiseks ühendatud sfäärilist mudelit [38, 39], aga töös on kasutusel traditsiooniliselt levinud Brown-Conrady polünoomiaalsel mudelil [40] põhinev lahendus.

Samuti on võimalik teostada kaamerate väline kalibratsioon, kus kaamerate vaateväljad on vaja projekteerida ühtsesse koordinaatsüsteemi, ilma kaameratele füüsiliselt ligipääsemiseta. Kui süsteemi uues asukohas või lisa kaamerate kasutuselevõtul on võimalik kaamera kaadris näha objekte, mis asuvad ka hoone joonistel, siis läbi selle on samuti võimalik kaameravälist kalibratsiooni teostada. Seda protsessi on võimalik teostada ka ainult kaamera pilti kasutades, nagu näitasid Wu et al. allikas [41].

5.3 Objektivastus

Väljatöötatud süsteemi üheks peamiseks piiranguks sai suutmatus tuvastada mitut laotransportööri väikesel alal. Sellest tingituna tekkis algoritmis mitu objekti identiteedi vahetust ning jälgimise all olev toode läks kaduma. Probleemiks on siin objektivastusmeetodi parameetrid, kus väikesel alal lubatakse esineda ühes kaameras ainult ühel laotransportööril (joonis 21). Selle vea parandamiseks on vaja uurida eelnimetatud parameetrite muutmisi ning selle mõju terve süsteemi tööle: lubades rohkem objektivastusi väikesel alal, võib esineda ka rohkem valetuvastatud laotransportööre.

Töös välja pakutud süsteemi üldine objektivastustäpsus sõltub objektivastuse baasmeetodist. Kõik uuritud (ning tänapäeval ka parimad ehk täpseimad [14]) meetodid põhinevad sügavnärvivõrkudel ning seega täpsuse parandamine on seotud otseselt sügavnärvivõrgu põhitõega: rohkem asjakohaseid andmeid annab üldjuhul täpsema mudeli.

Samuti on võimalik väikese vaevaga objektivastuse täpsust parandada tingimusel, et süsteemi kasutuselevõtu kohas on ainult käsikahveltõstukid. Vaadates joonist 8d võib näha, et isegi enne kaalutud hinde arvutamist on Faster-RCNN mudel ainult käsikahveltõstuki kategooriat arvestades parim, saavutades parima AP tulemuse nii hõivatud kui hõivamata staatustes. Seega vahetades YOLOv4 objektivastuse Faster-RCNN mudeli vastu on võimalik eelnimetatud nõuete korral süsteemi täpsust parandada.

Üheks põhjuseks, miks käsikahveltõstukite hõivatud ning hõivamata staatust suudeti tõstuki kategooriast paremini hinnata, oli töö autori hinnangul asjaolu, et transporditava aluse või toote suurusega võrreldes on käsikahveltõstuk tavatõstukist palju väiksem. Kuna algandmeid sildistati nii, et hõivatud staatustes on objekt siis, kui laotransportööri käppasid pole näha, siis olukordades kus tõstuk on tagumise otsaga kaamera poole, ei ole ka inimesel tihti võimalik hinnata tõstuki hõivatusstaatust. Näidis on sellisest olukorrast toodud joonisel 23, kus eriti raske on hõivatud staatust hinnata joonisel 23a, vähesel määral on hindamine keeruline joonisel 23b ning alles joonisel 23c on selgelt nähtav, et tõstukil hetkel tooteid pole.

Kuna süsteemi täpsusele ei olnud võimalik hinnangut anda elektrilise käsikahveltõstuki tuvastus- ning jälgimistäpsuse kohta, siis nende leidmiseks on vaja teha edasisi katseid. Kui manuaalne andmete sildistamine ei ole võimalik, oleks üheks uurimissuunaks ka andmete



(a)

(b)

(c)

Joonis 23: Erinevatest kaamera nurkadest tõstuki hõivatus staatus nähtav selgelt ainult alajoonisel 23c genereerimine simulatsiooni teel. Näiteks Nowruzi et al. näitasid, et sünteetiliste andmete kombineerimine reaalandmetega võib olla väga paljulubav viis andmeid koguda [42].

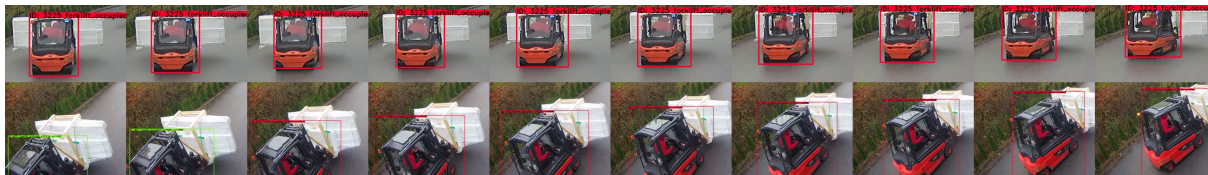
Samuti oleks vaja veel täiendavalt uurida erinevate objektituvastusmeetodite headust. Esiteks ei võimaldanud eksperimendis kasutatav riistvara kasutada töö kirjutamise ajal täpsemaid objektituvastusmeetodeid, nagu EfficientDet-D7 [25]. Seetõttu on võimalus, et olemasolevatel andmetel suuremaid objektituvastusmudeleid trenides suudetakse saavutada parem objektituvastustulemus, kui töös valminud mudelitega. Motivatsiooni annab selleks fakt, et käsikahveltõstuki kategooriat suudetakse kõige täpsemalt tuvastada suurima mudeliga.

Uurida oleks vaja ka kolmedimensioonilise objektituvastusmeetodi mõju süsteemi täpsusele, kasutades näiteks lahendust, mille on välja pakkunud Fang et al. artiklis [43]. Selline lähenemine peaks täpsust parandama, kuna laotranspordööri oleks sellisel juhul võimalik kaameratevahelises ühtses koordinaatsüsteemis positsioneerida täpsemalt. Näiteks peaks 3D objektituvastus vähendama vigasid, mida kirjeldati peatükis 4.1 ning illustreeriti joonisel 16.

Viimaseks oleks vaja leida lahendus toodete unikaalseks tuvastamiseks, et nende liikumisi täisautomaatselt jälgida. Näiteks üheks lihtsamaks lahenduseks võib siin tuua manuaalse süsteemi integratsiooni, kus laotöölaine peab valima terminalist toote või tooted, mida ta järgnevalt liigutama hakkab. Uurima peaks veel triipkoodide või QR koodide põhise tootetuvastust ning selle integratsiooni käesoleva süsteemiga, kuna eelnimetatud koodid on üks ühiseid omadusi erinevates ladudes ning nende olemasolu võiks ära kasutada.

5.4 Objektijälgimine

Objektijälgimise juures on prototüüplahenduses suurimaks probleemiks toodete kadumine liiga vara, ehk toode ei jõua tegeliku asukoha lähedusse. Selle põhjus on liialt lihtsustatud objekti staatuse tuvastusalgoritm, mis ei võta arvesse näiteks ennustuskindlust ning ka näiteks tõsiasja, et reeglina kiirema liikumise pealt laotransportöör ei võta peale tooteid ega pane neid maha.



Joonis 24: Sama objekt kahe kaamera vaates. Ülemises jadas objekti liikumise suund piirkasti alumise keskmise punkti suhtes liigub, alumises jadas sama punkt seisab praktiliselt paigal.

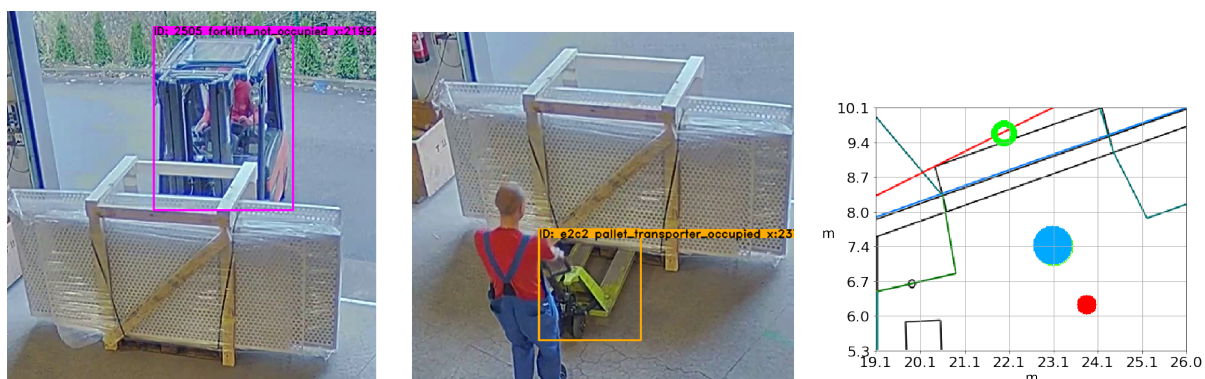
Samuti on kaamerate vaheliste radade kokkupanemise algoritm keerulisemates ladudes kasutuseks liialt primitiivne: arvesse võetakse ainult objekti asukohta ühtses koordinaatsüsteemis ning mitte näiteks tema liikumissuunda ega kiirust. See otsus tehti objektituvastuse algoritmi tulemustele põhinedes, kuna kaamerate pimenurgad tegid eksperimendi asukohas objekti liikumise suuna hindamise kaamerate vahel väga ebakindlaks (joonis 24).

Lisaks eelnevale on süsteemi universaalsust laonduse kontekstis silmas pidades piiranguks ka jälgija suutmatus kõrgustega arvestada. Teisisõnu, kasutades väljatöötatud prototüüplahendust, ei ole võimalik tuvastada toodete paigutamisi maapinnast kõrgemal asuvatele riiulitele, mis võib erinevates ladudes süsteemi kasutuselevõtul saada takistavaks faktoriks.

Teisalt ei ole ülitäpne toote paigustamine ka käesolevas töös arendatud süsteemi eesmärk. Süsteem arendati eesmärgiga demonstreerida laoliikumiste jälgimise võimalikkust, tuvastades laotoodete liikumisi tasapinnal ehk maapinnal ning konkreetne toote asukoha täpne määramine ei ole süsteemi ülesanne.

Samuti peab märkima, et paigustamise viga tuleb sisse juba käsitsi sildistatud andmetest, nagu on illustreeritud joonisel 25. Sellest joonisest võib näha, et piirkastide alumise keskmise punkti järgi teisendatud toote ehk aluse maailma koordinaat konkreetset juhul erineb tõstuki ja

käsikahveltõstuki vahel 4 meetrit.



Joonis 25: Sama objekti erinevate asukohtade ennustused vastavalt objekti tüübile ning kaamera nurgale. Parempoolseimal joonisel roheline täpp tähistab toote mahapanekut tõstuki poolt, punane täpp sama toote pealevõtmist käsikahveltõstuki poolt ning sinine tegelikku toote asukohta.

Sellegipoolest võib määratleda süsteemi edasiarendust võtmaks arvesse kõrgust ühe olulisema arengusuunana, kuna sellise funktsionaalsuse lisamine võimaldaks süsteemi kasutada enamikes ladudes: võimalikuks muutuks riuli täpsusega toote paigustamine.

Üheks lubavamaks uurimissuunaks on töö autori hinnangul ka erinevate olemasolevate parimate objektijälgimismeetodite ületoodavuse analüüs laonduse valdkonda. Näiteks on eriti lootustandvad meetodid objekti tuvastust ning jälgimist kombineerivad meetodid nagu JDE [44] ning FairMOT [45].

Süsteemi praktikas kasutatavaks tegemiseks oleks vajalik uurida veel meetodeid, kuidas tuvastada käsitsi toodete ümberpaigutamisi kasutades turvakaameraid. See probleem on eriti aktuaalne näiteks jaekaubanduses [46]. Probleemi on korduvalt analüüsitud RFID tehnoloogiat kasutades [47, 48], millega käesolevas töös arendatud süsteem konkureerib, seega vajaks turvakaamerapõhine lähenemine samuti täpsemat analüüsi.

6 Eksperdi hinnang

Käesolevas magistritöös valminud laojälgimissüsteemile anti hinnang eksperimendiasukoha ettevõtte logistikaalajuhi poolt. Taustaks võib öelda, et asutuses töötab 50 inimest ning nende valduses on kaks hoonet, mis asuvad teineteisest 3 kilomeetri kaugusel.

Tootmisettevõtte peamiseks tegevusalaks on lehtmetsa töötlus, kus üldjuhul iga valmistatav toode läbib vähemalt metalli lõikamise, painutamise ning värvimise protsesse. Ideaalis peaks iga sellise operatsiooni juures olema toodete ladustamise ala, kus hoiustatakse nii toormaterjali operatsiooniks kui ka operatsioonis valminud pooltoodet. Samuti on iga toote või pooltoote juures triipkoodi kleebis, mis seda toodet unikaalselt tuvastab.

Laonduse kontekstis on ettevõttes logistikajuhi hinnangul kaks suuremat probleemi. Esiteks, tihti operatsioone ei ole võimalik või otstarbekas täide viia vahetult teineteise järel. Näiteks suuremahulised tellimused viiakse sageli täide osadena, kus metalli lõikus toimub terve tellimuse jaoks ühes osas ning järgnevad operatsioonid on jagatud vastavalt tööpinkide täituvusele. Seetõttu on ettevõttesse loodud ka mitu lühiajalist vaheladu, kus toodete liikuvus on suur ning segadust tekitab.

Teine suurem probleem on ettevõtte jaoks toodete hoonete vahelise liikumise jälgimine. Suurema töömahu korral toimub ettevõtte teises hoonetes toodete koostamise operatsioon, mille käigus mitmest pooltootest valmib üks lõpp-produkt. Need pooltooted asetsevad sageli ettevõtte peahoone erinevates vaheladudes, kus nende leidmine võtab tihti liigselt aega. Samuti on tulnud ette olukordi, kus pooltoode on juba teise hoonesse liikunud, kuid sellekohast informatsiooni pole tulnud. Sellistel juhtudel läheb tavaliselt eriti kaua enne, kui pooltoode leitakse.

Logistikajuht on eelnimetatud probleemide lahendamiseks uurinud erinevaid võimalusi, millest paljulubavamad süsteemid on kõik põhinenud raadiosagedustuvastusel. Suurimaks takistuseks nende kasutuselevõtul sai tõsiasi, et selliste süsteemide nii ülesseadmiskulud kui ka halduskulud on liiga suured ning pigem sobivad olukordadesse, kus olemasolev toodete unikaalse tuvastamise võimalus puudub. Sellisel juhul oleks võimalik kleebisepõhine

tootetuvastus asendada raadiosidelainel põhineva tootetuvastusega, ning ühtlasi ka läbi majakate jälgida nende toodete liikuvusi.

Käesolevas magistritöös valminud laojälgimissüsteemi prototüüplahenduses samas nähakse potentsiaalset lahendust ettevõtte laohalduslikele probleemidele. Seda eelkõige lähtudes tõsiasjast, et süsteemi demonstratsioonis võib näha olukordi, kus juba praegu süsteem suudab kauba liikumisi mitte kaotada. Samuti annab motivatsiooni väljapakutud süsteemi madal kulu. Lahenduse nii ülesseadmiskulud kui ka halduskulud on väikesed, samuti saab vajadusel turvakaameraid süsteemi olemasoleva turvakaamerate infrastruktuuritõttu juurde lisada vähese vaevaga.

Suuremaks probleemiks on väljapakutud süsteemi juures ettevõtte logistikajuhi sõnul süsteemi suutmatust konkreetset kaupa tuvastada. Seetõttu ei ole hetkel võimalik demonstreeritud süsteemi kasutusse võtta, kuna probleem jääks lahendamata: konkreetsed tooted ning pooltooted läheks ikka kaduma. Kui aga suudetakse toodete unikaalne identifitseerimine lahendada, siis oleks võimalik süsteemi reaalselt testida ettevõtte probleemide lahendamisel. Seega unikaalsete toodete tuvastamine on järgmine suur väljakutse süsteemi edasiarendusel, mis vajab kindlasti läbi töötamist.

7 Kokkuvõte

Käesolevas töös pakuti välja poolautomaatne laoliikumiste jälgimissüsteem, mille eesmärk on pakkuda konkurentsi RFID tehnoloogial põhinevatele süsteemidele. Üheks peamiseks töö eesmärgiks oli tõestada, et laohaldus on võimalik lahendada kasutades olemasolevat turvakaamerate süsteemi, kuna võrreldes juba eksisteerivate laohaldussüsteemidega on turvakaamerate põhine süsteem kättesaadavam suuremale hulgale ettevõtetele väiksema investeeringu mahu tõttu.

Selleks viidi läbi eksperiment, mille tulemusena töötati välja tehisinägemise tehnoloogiatel põhinev mitme kaamera vaheline mitme laotranspordööri jälgimissüsteemi prototüüplahendus, mis suudab jälgida toodete ümberpaigutusi juhul, kui seda tehakse tõstukiga või käsikahveltõstukiga. Leitud lahendus arendati eraldiseisvate osadena, jättes süsteemile mitu edasist arengusuunda ning täiustusvõimalust.

Demonstreerimaks, et väljapakutud lahendus täidab oma eesmärgi, viidi läbi juhtumi analüüs, kus hinnati süsteemi täpsust ning selle tugevaid ning nõrkasid külgi. Analüüsi tulemusena jõuti järeldusele, et kuigi lihtsamatel juhtudel arendatud süsteem täidab oma eesmärgi, on keerulisematel juhtudel süsteemi funktsionaalsuses veel olulisi puudujääke. Näiteks üheks selliseks puuduseks on samaliigiliste laotranspordööri jälgimine, juhul kui eelnimetatud objektid asuvad kaamera vaates üksteise taga.

Jõuti ka järeldusele, et paljud süsteemi puudused on lahendatavad, suurendades kaamerate arvu, täiustades kaamerate kalibratsiooni kvaliteeti ning modifitseerides objektituvastuse ning jälgimise meetodeid või nende parameetreid. Näiteks toodi võimalus eksperimenteerida veel nii objektituvastus- ning jälgimist kombineerivate meetoditega ning ka objektijälgimisega kolmedimensioonilise objektituvastuse põhisedelt.

Samuti esitati töös edasisteks arengusuundadeks ettepanekuid, kuidas süsteemi veelgi täiendada. Olulisemateks sellisteks suundadeks leiti olevat objektituvastusmudeli treenimine seninägemata laotranspordööridel ning ka toodete asukoha tuvastusel kõrguse arvestamine.

Magistritöös koostati veel ka kvalitatiivne süsteemi kuluanalüüs, võttes võrdluseks RFID tehnoloogiatel põhinevad laojälgimise süsteemide kulu ning viimaseks anti hinnang väljatöötatud süsteemi kasutatavusele laondusala eksperdi poolt. Kuigi loodud lahendus ei paku veel terviklikku funktsionaalsust süsteemi juurutamiseks praktikas, on tehtud oluline samm töös näidatud tehnoloogial põhineva lahenduse loomise suunas.

Kasutatud kirjandus

- [1] Dilip Kumar, Elie Micah Kornfield, Alexander Clark Prater, Sridhar Boyapati, Xiaofeng Ren, and Chang Yuan. Detecting item interaction and movement, 2015. URL <https://www.lens.org/lens/patent/038-154-039-982-162>. Patent # US20150019391A1.
- [2] Kalyan Atishkumar and Dunn Steven Gregory. Automated inventory management system, 2015. URL <https://www.lens.org/lens/patent/057-175-991-956-890>. Patent # US009216857B1.
- [3] Wellman Parris S. Entity monitoring for Kiva robotic floors, 2017. URL <https://www.lens.org/lens/patent/101-945-092-203-133>. Patent # US009607285B1.
- [4] Luca Mainetti, Luigi Patrono, and Ilaria Sergi. A survey on indoor positioning systems. In *2014 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 111–120. IEEE, 2014.
- [5] Faheem Zafari, Athanasios Gkeliias, and Kin K. Leung. A survey of indoor localization systems and technologies. *CoRR*, abs/1709.01015, 2017. URL <http://arxiv.org/abs/1709.01015>.
- [6] Paul Voigtlaender, Michael Krause, Aljosa Osep, Jonathon Luiten, Berin Balachandar Gnana Sekar, Andreas Geiger, and Bastian Leibe. MOTs: multi-object tracking and segmentation. *CoRR*, abs/1902.03604, 2019. URL <http://arxiv.org/abs/1902.03604>.
- [7] Hang Xu, ChenHan Jiang, Xiaodan Liang, Liang Lin, and Zhenguo Li. Reasoning-RCNN: Unifying adaptive global reasoning into large-scale object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6419–6428, 2019.
- [8] Alex Polacco and Kayla Backes. The Amazon Go concept: Implications, applications, and sustainability. *Journal of Business & Management*, 24(1), 2018.

- [9] Paul Sawers. Radar raises \$16 million to automate inventory management with RFID and computer vision, 03 2019. URL <https://venturebeat.com/2019/03/28/radar-raises-16-million-to-automate-retailers-inventory-management-with-rfid-and-computer-vision/>. (klastatud 19.11.2019).
- [10] Emma Cosgrove. Automated inventory management is becoming a 'key battleground', 08 2018. URL <https://www.supplychaindive.com/news/automated-inventory-management-RFID/530071/>. (klastatud 19.11.2019).
- [11] Max Muller. *Essentials of inventory management*. HarperCollins Leadership, 2019.
- [12] Shelly Stazzone. Using RFID for inventory management: Pros and cons, 06 2019. URL <https://www.camcode.com/asset-tags/using-rfid-for-inventory-management-pros-and-cons/#>. (klastatud 31.10.2020).
- [13] Brian Ray. A breakdown of 7 RFID costs, from hardware to implementation, 02 2018. URL <https://www.airfinder.com/blog/rfid-cost>. (klastatud 31.10.2020).
- [14] Z Zou, Z Shi, Y Guo, and J Ye. Object detection in 20 years: A survey. arxiv 2019. *arXiv preprint arXiv:1905.05055*, 2019.
- [15] Patrick Dendorfer, Hamid Rezatofighi, Anton Milan, Javen Shi, Daniel Cremers, Ian Reid, Stefan Roth, Konrad Schindler, and Laura Leal-Taix. Mot20: A benchmark for multi object tracking in crowded scenes. *arXiv preprint arXiv:2003.09003*, 2020.
- [16] Hannah Kerner. Too many ai researchers think real-world problems are not relevant, 08 2020. URL <https://www.technologyreview.com/2020/08/18/1007196/ai-research-machine-learning-applications-problems-opinion/>. (klastatud 31.10.2020).
- [17] Henning Schulzrinne, Anup Rao, and Robert Lanphier. Real time streaming protocol (rtsp). RFC 2326, RFC Editor, April 1998. URL <http://www.rfc-editor.org/rfc/rfc2326.txt>. <http://www.rfc-editor.org/rfc/rfc2326.txt>.
- [18] Richard Hartley and Andrew Zisserman. *Computation of the Camera Matrix P*, page 178–194. Cambridge University Press, 2 edition, 2004. doi: 10.1017/CBO9780511811685.011.

- [19] Richard Szeliski. *Computer Vision: Algorithms and Applications*, chapter 2.1.5, pages 46–58. Springer-Verlag, Berlin, Heidelberg, 1st edition, 2010. ISBN 1848829345.
- [20] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*, page 709. Pearson, 3rd edition, 2009. ISBN 9789332543515.
- [21] Max Kuhn and Kjell Johnson. *Applied predictive modeling*, page 67. Springer, 1st edition, 2013. ISBN 9781461468486.
- [22] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection, 2020.
- [23] Glenn Jocher, Alex Stoken, and Jirka Borovec. ultralytics/yolov5: v3.1 - Bug Fixes and Performance Improvements, oct 2020. URL <https://doi.org/10.5281/zenodo.4154370>.
- [24] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks, 2016.
- [25] Mingxing Tan, Ruoming Pang, and Quoc V. Le. Efficientdet: Scalable and efficient object detection, 2020.
- [26] Yanzhao Wu, Ling Liu, Calton Pu, Wenqi Cao, Semih Sahin, Wenqi Wei, and Qi Zhang. A comparative measurement study of deep learning as a service framework. *IEEE Transactions on Services Computing*, page 1–1, 2019. ISSN 2372-0204. doi: 10.1109/tsc.2019.2928551. URL <http://dx.doi.org/10.1109/TSC.2019.2928551>.
- [27] Yu Emma Wang, Gu-Yeon Wei, and David Brooks. Benchmarking tpu, gpu, and cpu platforms for deep learning, 2019.
- [28] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft coco: Common objects in context. *Lecture Notes in Computer Science*, page 740–755, 2014. ISSN 1611-3349. doi: 10.1007/978-3-319-10602-1_48. URL http://dx.doi.org/10.1007/978-3-319-10602-1_48.
- [29] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg,

and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015. doi: 10.1007/s11263-015-0816-y.

- [30] GeekAlexis. Fastmot. <https://github.com/GeekAlexis/FastMOT>, 2020.
- [31] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and realtime tracking with a deep association metric. *2017 IEEE International Conference on Image Processing (ICIP)*, Sep 2017. doi: 10.1109/icip.2017.8296962. URL <http://dx.doi.org/10.1109/ICIP.2017.8296962>.
- [32] Lingxiao He, Xingyu Liao, Wu Liu, Xinchun Liu, Peng Cheng, and Tao Mei. Fastreid: A pytorch toolbox for general instance re-identification. *arXiv preprint arXiv:2006.02631*, 2020.
- [33] Kait Hobson. How much does a business security system cost?, 02 2020. URL <https://www.getkisi.com/blog/how-much-does-a-business-security-system-cost>. (klastatud 15.12.2020).
- [34] Alston Godbolt. What is the average viewing angle of security cameras?, 2020. URL <https://thesecuritycameraguy.com/what-is-the-average-viewing-angle-of-security-cameras>. (klastatud 03.11.2020).
- [35] Martin Sexton, Peter Barrett And, and Ghassan Aouad. Motivating small construction companies to adopt new technology. *Building Research & Information*, 34(1):11–22, 2006.
- [36] Oleksandr Bogdan, Viktor Eckstein, Francois Rameau, and Jean-Charles Bazin. Deepcalib: a deep learning approach for automatic intrinsic calibration of wide field-of-view cameras. In *Proceedings of the 15th ACM SIGGRAPH European Conference on Visual Media Production*, 2018.
- [37] Jianxiong Xiao, Krista A Ehinger, Aude Oliva, and Antonio Torralba. Recognizing scene viewpoint using panoramic place representation. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

- [38] João Pedro Barreto. A unifying geometric representation for central projection systems. *Comput. Vis. Image Underst.*, 103(3):208–217, 2006. doi: 10.1016/j.cviu.2006.06.003. URL <https://doi.org/10.1016/j.cviu.2006.06.003>.
- [39] Christopher Mei and Patrick Rives. Single view point omnidirectional camera calibration from planar grids. In *2007 IEEE International Conference on Robotics and Automation, ICRA 2007, 10-14 April 2007, Roma, Italy*, pages 3945–3950. IEEE, 2007. doi: 10.1109/ROBOT.2007.364084. URL <https://doi.org/10.1109/ROBOT.2007.364084>.
- [40] C Brown Duane. Close-range camera calibration. *Photogramm. Eng.*, 37(8):855–866, 1971.
- [41] J. Wu, L. Ma, and X. Hu. Predicting world coordinates of pixels in rgb images using convolutional neural network for camera relocalization. In *2016 Seventh International Conference on Intelligent Control and Information Processing (ICICIP)*, pages 161–166, 2016. doi: 10.1109/ICICIP.2016.7885894.
- [42] Farzan Erlik Nowruzi, Prince Kapoor, Dhanvin Kolhatkar, Fahed Al Hassanat, Robert Laganieri, and Julien Rebut. How much real data do we actually need: Analyzing object detection performance using synthetic and real data, 2019.
- [43] Jiaojiao Fang, Lingtao Zhou, and Guizhong Liu. 3d bounding box estimation for autonomous vehicles by cascaded geometric constraints and depurated 2d detections using 3d results, 2019.
- [44] Zhongdao Wang, Liang Zheng, Yixuan Liu, and Shengjin Wang. Towards real-time multi-object tracking, 2019.
- [45] Yifu Zhang, Chunyu Wang, Xinggang Wang, Wenjun Zeng, and Wenyu Liu. Fairmot: On the fairness of detection and re-identification in multiple object tracking, 2020.
- [46] Joshua Bamfield. Shrinkage, shoplifting and the cost of retail crime in europe: a cross-sectional analysis of major retailers in 16 european countries. *International journal of retail & distribution management*, 2004.
- [47] AG De Kok, KH Van Donselaar, and Tom van Woensel. A break-even analysis of rfid technology for inventory sensitive to shrinkage. *International Journal of Production Economics*, 112(2):521–531, 2008.

- [48] Ti-Jun Fan, Xiang-Yun Chang, Chun-Hua Gu, Jian-Jun Yi, and Sheng Deng. Benefits of rfid technology for reducing inventory shrinkage. *International Journal of Production Economics*, 147:659–665, 2014.
- [49] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. <https://github.com/facebookresearch/detectron2>, 2019. (küstatud 01.11.2020).
- [50] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. URL <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- [51] zylo117. Yet-another-efficientdet-pytorch. <https://github.com/zylo117/Yet-Another-EfficientDet-Pytorch>, 2020. (küstatud 01.11.2020).
- [52] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Gregory S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian J. Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Józefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Gordon Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul A. Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda B. Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *CoRR*, abs/1603.04467, 2016. URL <http://arxiv.org/abs/1603.04467>.
- [53] Joseph Redmon. Darknet: Open source neural networks in c. <http://pjreddie.com/darknet/>, 2013–2016.

Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹⁰

Mina, Ivar Vipper

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose "Laoliikumiste tuvastamine ja jälgimine turvakaamerate abil" , mille juhendaja on Juhan-Peep Ernits
 - 1.1. reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2. üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

05.01.2021

¹⁰Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

Lisad

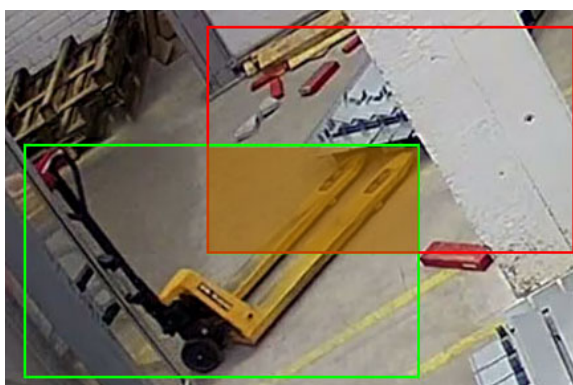
Järgnevalt on välja toodud enimlevinud terminite lühikirjeldused ning samuti objektivastus mudelite lühikirjeldused.

7.1 Intersection over Union headuse näitaja

Intersection over Union, samuti tuntud kui *Jaccard index*, on termin, mida kasutatakse ennustatud ala paigutamise täpsuse määramiseks võrrelduna algtõega. Parim tulemus on 100%, mis tähendab, et ennustus kattub täielikult algtõega ning halvim on 0%, mis tähendab, et ennustus ei saanud algtõele ligilähedale.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

Matemaatiliselt käib arvutus valemi 3 järgi ning IoU illustratsioon on toodud joonisel 26.



(a) Ennustatud ala kattub vähesel määral algtõega ehk $IoU < 50\%$.



(b) Ennustatud ala kattub hästi algtõega ehk $IoU > 50\%$.

Joonis 26: IoU võrdlus. Roheline ala märgib algtõde ning punane ala märgib ennustatud ala.

7.2 Average Precision ehk keskmine täpsus

Intersection over Union annab meile objektituvastuses ainult objekti paigustamise headuse kohta teadmisi jättes välja informatsiooni konkreetse klassi, mida soovitakse leida. Täpsemalt siis on generaliseeritud objektituvastuse headuse hinnanguks vaja teadmisi nii paigustamise headuse ning klassi sildistamise headuse kohta kombineeritult.

Siin on võetud kasutusele *AP* ehk *Average Precision* headuse näitaja, mis kombineeribki eelnevalt kirjutatud objektituvastuse omadusi. Täpsemalt kombineeritakse iga ennustatava klassi kohta näitajad *precision* ehk täpsus (valem 4) ning *recall* ehk saagis (valem 5) ning piiritletakse seda IoU väärtusega.

$$precision = \frac{TP}{TP + FP} \quad (4)$$

$$recall = \frac{TP}{TP + FN} \quad (5)$$

Selline IoU piiritus tähendab, et näiteks 50% lävise korral määratletakse kõik klassid, mis paigustati IoU väärtusega $\geq 50\%$ ning klassifitseeriti õigesti kategooriasse tõeseks (*TP*). Kui objektid klassifitseeritakse küll õigesti, kuid nende paigustamine on halb (antud näites $< 50\%$), siis määratletakse sellised juhud valedeks (*FP*). Kui aga objektid klassifitseeritakse valesti, kuigi tegelikult on objektid leitavad, siis määratletakse need juhud samuti valedeks (*FN*).

7.3 Mudelite treenimise parameetrid

Kõik järgnevalt nimetatud mudelid on treenitud terve andmekogu treeningandmetel, kus toimub liikumist ning kasutades MSCOCO eeltreenitud kaalusid, väljaarvatud eraldi välja toodud juhud. Treenimisandmekogus on kokku 42240 kaadrit, testandmekogus on 3304 kaadrit ning validatsiooni andmekogus on 3510 kaadrit.

7.3.1 Faster-RCNN

Faster-RCNN [24] mudelit treeniti kasutades Detectron2 tarkvararaamistiku [49]. Kõikide konfiguratsioonide kaalud baseerusid ImageNet [29] kaaludel. Treenimisparameetreid ei muudetud baaskonfiguratsiooni failis, seega mudeli treenimisparameetritega on võimalik tutvuda Detectron2 konfiguratsiooni failides.¹¹

Tulemused olid järgnevad:

faster_rcnn_R_50_C4_1x

Andmekogu: Ainult täielikult nähtavad objektid.

Ülekantud kaalud: ImageNet [29].

Treenitud kaadreid: 6000.

Treenitud aeg: 1,7h.

Kasutatud treenimis täiustused: Horisontaalne pildi peegeldus.

Treenimise lõppemise põhjus: Mudel ei suutnud õppida peale ~5500 kaadri nägemist.

Kommentaar: Pildi suurus 640px lühem külg.

faster_rcnn_R_101_FPN_3x

Andmekogu: Kõik.

Treenitud kaadreid: 100000.

Treenitud aeg: 17h.

Kasutatud treenimis täiustused: Horisontaalne pildi peegeldus.

Treenimise lõppemise põhjus: Ajalimiit täis.

Kommentaar: Mudeli treenimine ebastabiilne, samas võimalik mudelid veel treenida väiksema õppimise parameetri korral. Pildi suurus 640px lühem külg.

faster_rcnn_R_50_C4_1x

Andmekogu: Kõik.

Treenitud kaadreid: 10000.

Treenitud aeg: 3h.

Kasutatud treenimis täiustused: Horisontaalne pildi peegeldus.

¹¹Saadaval: <https://github.com/facebookresearch/detectron2/tree/master/configs/COCO-Detection>

Treenimise lõppemise põhjus: Ajalimiit täis.

Kommentaar: Mudel lõpetas kiirema õppimise peale 3500 pildi nägemist, samas õppis aeglaselt edasi. Mudeli hindamine tehti testandmekogul iga 500 pildi järel, seega kokku oli treenimisaeg 20h, millest mudeli hindamisele kulus 17h. Pildi suurus 640px lühem külg.

7.3.2 YOLOv5

YOLOv5 [23] mudelit treeniti kasutades PyTorch tarkvararaamistiku [50] ning YOLOv5m baas mudelit.

faster_rcnn_R_50_C4_1x

Andmekogu: Kõik.

Treenitud kaadreid: 422000.

Treenitud aeg: 30h.

Kasutatud treenimis täiustused: Horisontaalne pildi peegeldus.

Treenimise lõppemise põhjus: Ajalimiit täis.

Kommentaar: Pildi suurus 640x640.

7.3.3 EfficientDet

EfficientDet alammudeleid d1 ja d3 treeniti Yet-Another-EfficientDet-Pytorch tarkvaraga [51] ning d0 alammudelit treeniti TensorFlow tarkvara raamistikuga [52].

Mudel ei osutunud valituks, kuna saavutatud headus ei olnud piisavalt hea võrreldes erinevate proovitud mudelitega. Üheks suuremaks põhjuseks siin oli eksperimendis kasutatava riistvara jõudlus: EfficientDet mudel soovib treenimiseks palju videomälu.

Näiteks EfficientDet-d2 ning -d3 mudelid, mis on võrreldavad oma täpsuses MSCOCO andmestikul YOLOv4 mudeliga [22], suudeti NVIDIA GTX1060 6GB graafikakiirendil treenida bloki suurusega 1. Selline treeningkonfiguratsioon aga oli väga ebastabiilne ning ei suutnud paraneda. Võrdluseks YOLOv4 treeniti sama graafikakiirendiga bloki suurusega 64.

EfficientDet-d1

Andmekogu: Kõik.

Treenitud kaadreid: 89000.

Treenitud aeg: NA.

Kasutatud treenimis täiustused: NA.

Treenimise lõppemise põhjus: Mudel ei parane.

EfficientDet-d3

Andmekogu: Kõik.

Treenitud kaadreid: 211000.

Treenitud aeg: NA.

Kasutatud treenimis täiustused: NA.

Treenimise lõppemise põhjus: Ajalimiit täis ning tulemused halvad.

Kommentaar: Esimesed 80000 sammu treenitud õpisammuga $1e-4$, järgnevalt sammuga $1e-5$.

Pildi suurus 512x512.

EfficientDet-d0

Andmekogu: Kõik.

Treenitud kaadreid: 220000.

Treenitud aeg: 40h.

Kasutatud treenimis täiustused: NA.

Treenimise lõppemise põhjus: Mudel ei parane.

Kommentaar: Mudel õpib liigselt treening andmeid, test andmetel tulemus ei parane. Pildi suurus 512x512.

7.3.4 YOLOv4

YOLOv4 mudelit [22] treeniti kasutades ametliku Darknet tarkvararaamistiku [53].

YOLOv4

Andmekogu: Kõik.

Treenitud kaadreid: 550000.

Treenitud aeg: 26h.

Kasutatud treenimis täiustused: NA.

Treenimise lõppemise põhjus: Ajalimiit täis.

Kommentaar: Pildi suurus 416x416.

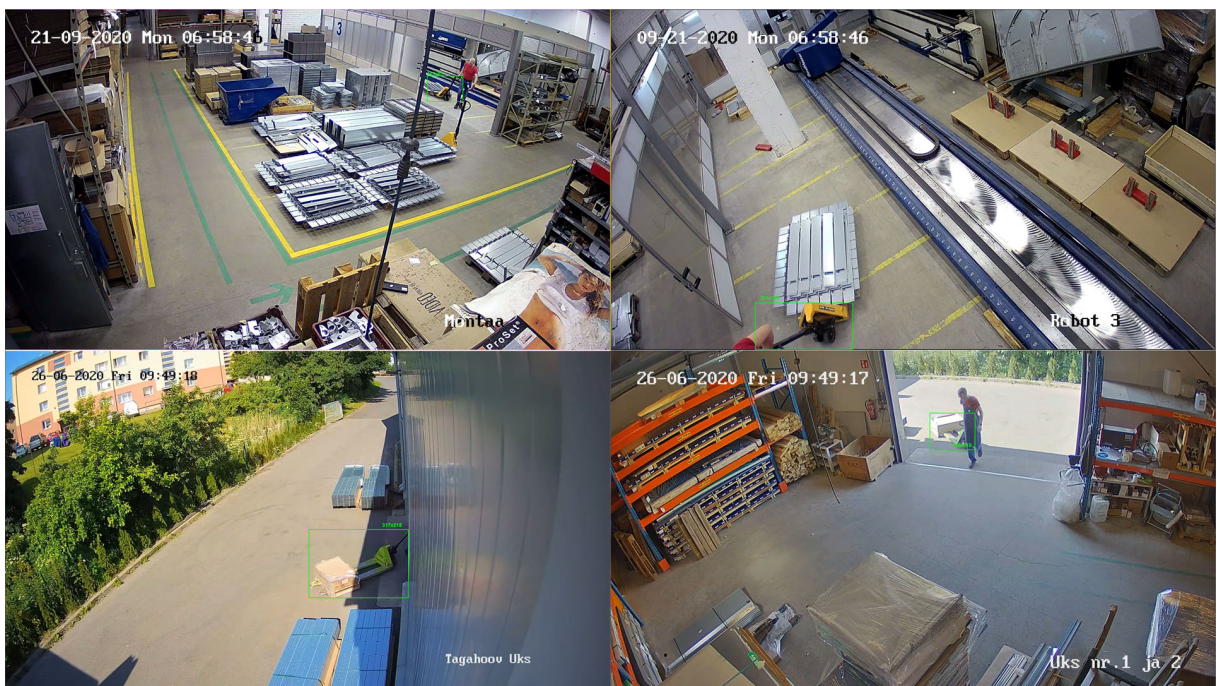
7.4 Kaalutud keskmine objektituvastus hinne

Suurus	Pindala pikslites
Väike	0-5313
Keskmine	5314-89958
Suur	59959-1e10

Tabel 1: Objektituvastus meetodi tulemusel leitud objekti piirkasti suurused pikslites.

Klass	Suurus	Kaal
Tõstuk hõivamata	keskmine	0,085
Tõstuk hõivatud	keskmine	0,220
Tõstuk hõivamata	suur	0,040
Tõstuk hõivatud	suur	0,155
Käsikahveltõstuk hõivamata	keskmine	0,125
Käsikahveltõstuk hõivatud	keskmine	0,375

Tabel 2: Kaalutud keskmise hinde kaalud objekti klassi ja suuruse kaupa.



Joonis 27: Keskmise suurusega objekti piirkasti suuruse arvutuse näidispildid. Keskmise suurusega klassi tüüp jääb väikseima ning suurima leitud käsikahveltõstuki piirkasti pindala vahele.