

TARTU ÜLIKOOL
Arvutiteaduste instituut
Informaatika õppekava

Ilmar Möls

Fiery semantilise segmentatsiooni mudeli efektiivsuse hindamine

Bakalaureusetöö (9 EAP)

Juhendaja: Markus Kängsepp

Tartu 2024

Lühendite sõnastik

CC BY-NC-SA 4.0 (*Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Public License*) - Creative Commonsi 4 versiooni litsents, mille korral tohib tööd kasutada mitteäriline eesmärkidel juhul kui viidatakse autorile. Seda tööd võib edasi jagada samadel tingimustel.

SSD (*Solid State Drive*) - pooljuhtketas

NAND (*NOT AND*) - NING-EI on loogikatehe, kus kui talle sisse antud sisenditest kõik ei ole õiged annab ta tõese väärtuse.

NOR (*NOT OR*) - EI-VÕI on loogikavärav, mis annab tõese tulemuse vaid juhul, kui ükski sisenditest pole tõene.

LiDAR (*Light Detection And Ranging*) - LiDAR on laserskaneerimisseade, mis kasutab laserkiirt objektide tuvastamiseks ja nende kauguste mõõtmiseks.

RADAR (*Radio Detection And Ranging*) - RADAR on õhuruumi raadioseireseadeldis, mis kasutab raadiolaineid objektide tuvastamiseks ja nende kauguste mõõtmiseks.

CNN (*Convolutional Neural Network*) - Konvolutsiooniline närvivõrk (CNN) on mitmekihiliste sügavõppe algoritmi grupp, mis on kasutusel pilditöötlustes ja pildilt objektituvastuses. CNN koosneb tavaliselt mitmest kihist, sealhulgas konvolutsiooni kihtidest, aktiveerimise kihtidest, *pooling layers* ehk koonduvuskihtidest ja täielikult ühendatud kihtidest.

SEN (*Squeeze-and-Excitation Networks*) - Pigistamine ja väljundamine on närvivõrkude õppearhitektuur, mis suurendab võrgu jõudlust ja tõhusust, võttes arvesse kanalitevahelist seost ja nende tähtsust.

Sisukord

Sissejuhatus	6
1. Fiery mudel	7
1.1 Ümbruse info sisselugemine	7
1.2 Pildi töötlemine	7
1.3 Fiery mudeli kasulikkus	8
1.4 EfficientNet	9
1.5 NuScene andmestik	10
1.6 Semantiline segmentatsioon	11
2. Isesõitvate autode teaduslik taust	12
2.1 Autonoomsete sõidukite kontseptsioon ja katsumused	12
2.2 Isesõitvate autode ohutus	12
2.3 Aukoha määramise tehnikad	13
2.3.1 Kaamerad	14
2.3.1.1 Monokulaarne kaamera	14
2.3.1.2 Stereoskoopiline kaamera	15
2.3.1.3 Panoraamkaamera	15
2.3.1.4 Termokaamerad	15
2.3.2 LiDARid	16
2.3.3 Radarid	16
2.3.4 Globaalne Positsioneerimissüsteem (GPS)	17
2.4 Andmete salvestamine	18
3. Fiery mudelilt andmete kogumine	19
4. Fiery mudeli analüüs	21
4.1 Eeldused Fiery mudeli analüüsimisel	21
4.2 Tulemuste analüüs	23
4.3 Võimalikud Edasiarendused	30
Kokkuvõte	31
Viited	32
Lisad	35

Fiery Semantilise Segmentatsiooni Mudeli Efektiivsuse Hindamine

Lühikokkuvõte:

See uurimistöö keskendub Fiery mudeli jõudluse analüüsimisele inimeste tuvastamisele isesõitva auto ees oleva kaamera pildist kasutades Fiery mudelit. Isejuhtivate autode "nägemissüsteemi" (inglise keeles *perception system*) või tajumissüsteemi täpsus on kriitiline, kuna see aitab tagada nii liiklusohutuse selliste sõidukitega liikluses kui ka sõidukite juhtimissüsteemi tõhusa toimimise.

Töös uuritakse, mis on Fiery mudel ja mis on selle mudeli võimekus inimeste tuvastamiseks erinevate isikute või pildi mõjutavate tingimuste korral. Lisaks pakub töö ülevaadet isesõitvate autodega seotud teaduslikust taustast, sealhulgas autonoomsete sõidukitega kaasakäivatest ohutusest ja nende poolt kasutatavate asukoha määramise tehnikatest. Kokkuvõttes võiks öelda, et Fiery mudelil paistab olevat tuvastusprobleeme piltide äärtes ja teel lamavate isikutega.

Võtmesõnad:

Fiery mudel, isikute tuvastamine, süvaõpe, arvutinägemine, pilditöötlus, masinõpe, algoritm, isesõitev auto

CERCS:

P170 Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine (automaatjuhtimisteooria)

The analysis of the effectiveness of Fiery semantic segmentation model

Abstract:

This Bachelor's thesis focuses on analysing the performance of the Fiery model in detecting people from frontal camera images of a self-driving car using the Fiery model. The accuracy of the perception system of self-driving cars is critical, as it helps to ensure both road safety in traffic with such vehicles and the efficient operation of the vehicle's control system.

The thesis covers what the Fiery model is and what is the ability of this model to identify people in the case of different persons or conditions affecting the image. In addition, the paper provides an overview of the scientific background related to self-driving cars, including the safety issues associated with autonomous vehicles and the positioning techniques used. In conclusion, the Fiery model seems to have detection problems with pedestrians on the edges of images and pedestrians lying on the road.

Keywords:

Fiery Model, Person Recognition, Deep Learning, Computer Vision, Image Processing, Machine Learning, Algorithm, Self-Driving Car

CERCS:

P170 Computer science, numerical analysis, systems, control

Sissejuhatus

Lõputöö teema valimine

Fiery mudeli analüüsimine valiti lõputöö teemaks kuna, juhendaja Markus Kängsepp oli selle mudeliga varasemalt tegelenud ning märganud väiksemat tuvastusprotsenti istuvate ja pikali olevatele inimestega piltidel. See tekitas küsimuse, kas see oli vaid erandlik juhtum või viitab mudeli üldistele probleemidele. Huvi teema vastu tekkis seega soovist süvitsi uurida, millised tegurid võivad mõjutada Fiery mudeli töökindlust erinevates valgusoludes ning kas ja kuidas saaks mudelit paremaks kohandada erinevate tingimustega toimetulekuks.

Bakalaureusetöö olulisus

Suuremate tuvastusprobleemide korral võivad isesõitvad autod mitte tajuda teede ääres olevaid takistusi või piire, mis suurendab autos viibijate surmaga lõppevate õnnetuste riski. Isegi selliste isesõitvate autodega, millel on väiksemad tajumisprobleemid võivad põhjustada inimohvritega õnnetused, kuna kui süsteem ei tuvasta väiksemaid objekte, nagu jalakäijad või rattureid, siis võib auto neile otsa sõita.

Bakalaureusetöö eesmärk

Uurimistöö eesmärk on saavutada arusaam Fiery mudeli tõhususest inimeste tuvastamisel kolmemõõtmelises stseenis kahemõõtmelise pildi alusel. Selle tulemusel saaks paremini hinnata Fiery mudeli tugevusi ja kitsaskohti ning avada uusi võimalusi selle mudeli täiendamise ja kohandamise jaoks, võttes arvesse selles töös kogutud infot kasutusstenaariumite kohta. Saadud arusaamine toeatab mitte ainult Fiery mudeli teoreetilise aluse täiustamise eesmärki, vaid pakub ka praktilisi suuniseid mudeli tõhusaks rakendamiseks mitmesugustes autonoomsetes sõitmissüsteemides.

Bakalaureusetöö struktuur

Antud bakalaureusetöö esimene sisupeatükk annab ülevaate, mis on Fiery mudel, mis närvivõrke ja treeningandmeid ta kasutab ja kuidas ta töötab. Teises peatükis on antud konteksti isesõitvate autode poolt kasutatavate tehnoloogiate kohta, nagu kaamerad ja teised objektide tuvastussüsteemid. Kolmandas peatükis saab ülevaate selle kohta, kuidas

bakalaureusetöö praktiline osa läbi viidi ja neljandas tehakse saadud tulemuste põhjal analüüs.

1. Fiery mudel

Fiery mudel on isesõitvatele autodele loodud autode tuleviku asukoha ennustamismudel, mis selle uurimistöö jaoks ümbertöötatud vaid isikute tuvastamiseks. Selleks, et mudel saaks ennustada tulevikku, antakse talle sisse kuuest autol olevast kaamerast kaadrid - kaks minevikukaadrit ja üks oleviku kaader. Mudel tuvastab nendest kaadritest isikud või autod ning seejärel töötleb neid, et luua 2D semantiline segmentatsiooni faili, mis kujutab olukorda linnusilmavaates. [1]

1.1 Ümbruse info sisselugemine

Fiery mudelis kasutatakse kuut monokulaarset kaamerat, et teha pildid 360 kraadi ümbritsevast keskkonnast. Kuue kaameraga saab auto ümbrusest peaaegu täiskatte. Kuna autod on enamasti pikemad kui laiemad, siis paigaldatakse 2 kaamerat mõlemale auto küljele. [1]

Monokulaarsed kaamerad kasutavad ühte objektiivi pildi jäädvustamiseks, mis erinevalt stereokaamerateist ei võimalda otsest sügavustajamist. Seevastu aitab nende kaamerate vaatealade osaline kattuvus 3D pildi loomist, kuna sama objekti erinevad vaatenurgad on saadaval erinevatelt kaameratelt. [1]

Kuna Fiery mudel ei kasuta objektide kauguse hindamiseks LiDAR'eid ega muid sarnaseid sensoreid, mis on tavalised paljudes isejuhtivates sõidukites, vaid kasutab ainult kaamerate abil saadud ja salvestatud teavet. See lähenemine võimaldab vähendada süsteemi maksumust ja keerukust, samal ajal säilitades võime luua täpne ja dünaamiline ümbritseva keskkonna mudel. [1]

1.2 Pildi töötlemine

Pildi tötluse esimeses etapis tuvastab Fiery mudel kaadril olevad autod või isikud segmenteerides pildid eraldi seisvateks pikslite kogumiteks. Selleks kasutades süvaõppe konvolutsioonilist võrku nimega EfficientNeti, mis on treenitud suurel hulgal märgendatud andmetel. Selle tulemusena saame eraldi seisvad objektid ning siis leitakse tõenäosus, et antud kogumik pildil on auto või isik. [1]

Objektide tuvastamise 2D piltidel teeb Fiery mudeli jaoks keeruliseks piltide tegemiseks kasutatavad monokulaarsed kaameraid, kuna niimoodi piltide salvestamisel ei ole võimalik piltidelt kauguse infot välja lugeda. Seda protsessi korratakse iga kaamera jaoks eraldi. [1]

Kaamerate sisendite 2D-st 3D-sse tõstmiseks kasutab Fiery mudel EfficientNet'i konvolutsioonilise kodeerija poolt loodud tunnuste komplekti, mille abil arvutab Fiery iga objekti kõige tõenäolisemaid sügavused ja seejärel tõstab need kolmemõõtmelisse punktipilve. [1]

Kasutatakse infot kaamera poolt tehtud piltide paiknemisest võrreldes kaamerate asukoht sõiduki massikeskme suhtes, tõstab Fiery mudel kaadrid ühisesse 3D-sse raamistikku võttes arvesse tehtud piltide aega ja asukohta. Selle tulemusena loob mudel igast kaamerast kogutud andmete põhjal ühtse 3D-kujutise, mis esindab stseeni mitmest vaatenurgast. [1]

Seejärel puhastab ja täpsustab Fiery mudel 3D-pilti erinevate filtrite ja algoritme abiga, eemaldades müra ja objekte, mis ei anna kaadril olevate objektide kohta mingit infot. Kui müra, näiteks vihmapiisad, on eemaldatud, siis mudel täidab tühjad alad ja moodustab ruumilise 3D maatriksi. [1]

Pärast 3D kujutise loomist projekteeritakse need ülalt alla vaate pildile. Pärast 3D kujutise loomist muudetakse need ülalt alla vaate segmentatsiooniks. Selleks jaotatakse ruum sõiduki ümber kastidesse mille mõõtmed on $0,50\text{ m} \times 0,50\text{ m}$, kattes niimoodi kokku ala ulatuses $100\text{ m} \times 100\text{ m}$. 3D omadused summeeritakse vertikaalses suunas, et moodustada ülalt alla vaate omaduste kaardid. See protsess võimaldab luua ülevaatliku semantilise segmentatsioonistseenist ülalt alla vaates, mis võimaldab analüüsida sõidukite ja muude objektide käitumist nende liikumisel erinevates keskkondades. Selline ühine pilt, mis võimaldab mudelil ennustada sõidukite või jalakäijate käitumist ja manöövreid erinevates liiklusolukordades. [1]

1.3 Fiery mudeli kasulikkus

Fiery mudel aitab lahendada probleeme, mis tekivad isesõitvatel autodel liiklemisel dünaamilises liikluses. Traditsioonilised autonoomsed sõidukite juhtimissüsteemid lähtuvad eeldusest, et dünaamilised takistused jätkavad oma praegust käitumist tulevikus. Fiery aga õpib ennustama tuleviku liikumist ilma eelduseta, et see jätkub lineaarselt, võimaldades süsteemil paremini arvestada võimalike tulevikustsenaariumitega. [1]

1.4 EfficientNet

EfficientNet on konvolutsioonilisel (CNN) närvivõrgul põhinev arhitektuur, mis töötab "ühendatud skaleerimise" (inglise keeles *compound scaling*) alusel. Ühendatud skaleerimise idee on suurendada korraga pildi, sügavust, laiust ja lahutust. [2]

EfficientNeti treenimine algab mõistliku suurusega närvivõrguga. Kui see närvivõrk näitab, et ta töötab antud ülesandega hästi, siis võetakse parameetrina kasutusele liitkoefitsient (*Partition coefficient*), mis määrab, kui palju närvivõrgu milliseid mõõtmeid skaleerida nii, et mudeli laius, sügavus ja eraldusvõime suurenevad ühtlaselt. [2]

Fiery poolt kasutatav EfficientNet on treenitud suure hulga märgendatud piltide abil, kus iga pilt on seotud vastava objektiklassiga, mida süsteem peaks tuvastama. Treenimisprotsessis toimub närvivõrkudele standartne kaalude õppimine, kus süsteem püüab optimeerida võrgu parameetreid nii, et see saaks anda võimalikult täpseid ennustusi uute piltide kohta, mida tema treeningandmetes ei ole. [1,2]

Oma arhitektuuri baasina kasutab EfficientNet Mobile Inverted Bottleneck (MBConv) kihte, mis koosnevad sügavast kihist eraldiseisvatest konvolutsioonidest ja pööratud jääkklassidest. Lisaks kasutab arhitektuur *Squeeze-and-Excitation* (SE) optimeerimist, et edasi parandada mudeli jõudlust. [1,2]

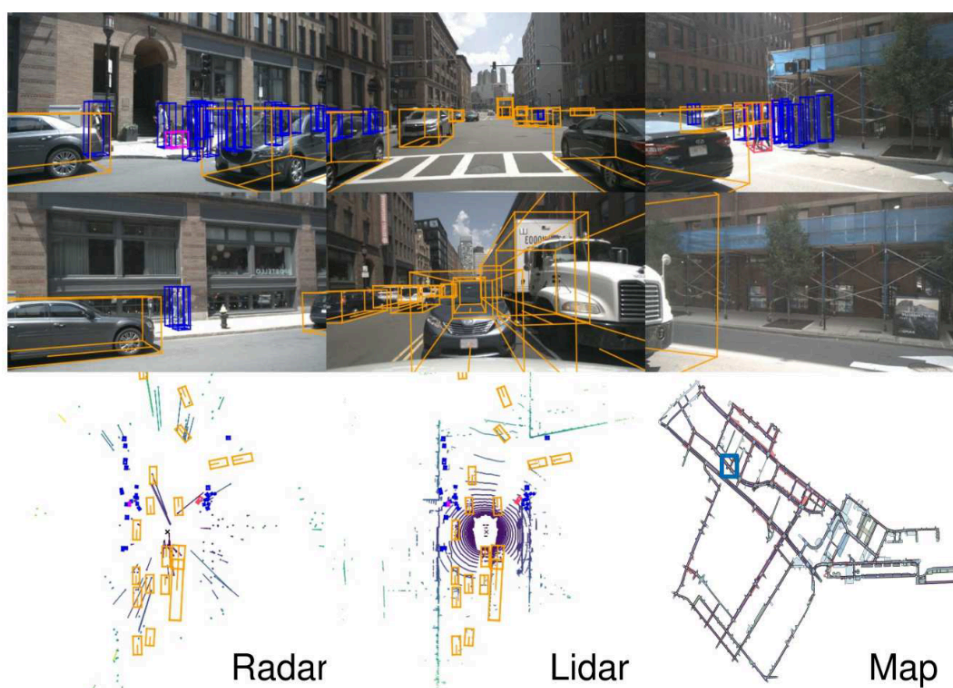
Squeeze-and-Excitation Networks (SENet) on sügav õppearhitektuur, mis suurendab võrgu võimsust ja tõhusust. *Squeeze* etapis vähendab see süsteem kanalite ruumala nii, et säilitatakse info olulised omadused. *Excitation* etapis, õpib võrk iga kanali omadust, lastes närvivõrgul keskenduda olulisematele omadustele ja ignoreerida kasutumaid. [3]

FIERY mudelis kasutati EfficientNeti arhitektuuri osana sügavat konvolutsioonilist võrku, et analüüsida ja töödelda autodele paigaldatud kaamerate poolt salvestatud kaadreid. EfficientNet võimaldas mudelil täpselt tuvastada olulisi tunnuseid ja mustreid kaamerate piltidelt, mis omakorda aitas kaasa tuleviku asukohtade ennustamisele. Tänu EfficientNeti võimele tasakaalustada täpsust ja arvutuslikku efektiivsust oli FIERY mudel võimeline saavutama kõrget jõudlust ennustuste tegemisel reaajas videote analüüsil. [1]

1.5 NuScene andmestik

NuScene on laiendatud andmekogum, mis on spetsiaalselt loodud autonoomsete sõidukite uurimiseks ja arendamiseks. Need andmed hõlmavad mitmesuguseid stseene ja olukordi, mis on salvestatud reaalistest sõidukitest Bostonist ja Singapurist, kus on äärmiselt tihedad sõidutingimused. NuScene'i andmestik sisaldab mitmesuguseid sensorandmeid, sealhulgas kaamerate, LiDARi, radari ja GPS-seadmete salvestisi, mis võimaldavad süvaõppe algoritme treenida. [4, 5]

NuScene on ulatusliku sensoriteabega kogumik erinevates liiklusolukordades, hõlmates teekatte seisundit, liiklusmärke, liiklusvoolu ja jalakäijate käitumist. Tänu sellele andmekogu mitmekesisusele saavad NuScene'i kasutavad uurijad ja arendajad treenidaja testida autonoomsete sõidukite algoritme erinevates ja keerukates keskkondades, kus on mitmeid väljakutseid. [4, 5]



Joonis 1: Näidis informatsioon, mis on igas NuScene'i ploki kättesaadav. [5]

NuScene andmekogumi kasutamine on lubatud mitteärilistel eesmärkidel, tingimusel, et järgitakse litsentsis CC BY-NC-SA 4.0 kirjas olevaid reegleid, mis lubab andmekogu kasutada tasuta mitteärilistel eesmärkidel. Selles andmekogumis olevate kolmandate osapoolte andmete puhul peab nende kasutamisel täitma igale juhule konkreetseid

eritingimusi. Lisaks tuleb järgida Singapuri ja Ameerika Ühendriikide privaatsuseeskirju, mille rikkumine võib lõppeda kohtuasjaga. [4, 5]

FIERY mudelis kasutati Nuscenide andmekogumit treeninguks ja testimiseks, et analüüsida autonoomsete sõidukite käitumist reaalse maailma stsenaariumides. Kuigi NuScenid sisaldavad lidarite ja radari andmeid, ei kasuta Fiery mudel neid konkreetseid andmeid oma töös. NuScenide mitmekülgsete andmete abiga on Fiery mudelil kontekst erinevate keerulisemate liiklusolukordade jaoks ja on selle abil võimaline saavutada kõrget täpsust ennustuste tegemisel autonoomsete sõidukite käitumise kohta reaalajas. [4, 5]

1.6 Semantiline segmentatsioon

Semantiline segmentatsioon on arvutinägemise tehnika, mis jaotab pilt eraldi segmentideks, ehk osadeks, kus iga osa vastab konkreetsele objektile või objektiosale. Seda viiakse läbi treenitud närvivõrkude abil, millele antakse treenimiseks ette pildid, kus on käsitsi igale pikslile määratud nende klass. [6]

Erinevalt lihtsast kujutiste klassifitseerimist, mis aitab masinal mõista, millist teavet kujutis sisaldab, võimaldab semantiline segmenteerimine masinal tuvastada erinevat tüüpi visuaalset teavet objektide täpsest asukohast, näiteks, kus üks asi algab ja lõpeb. Fail, mis sisaldab teavet semantilise segmentatsiooni tulemustest kutsutakse segmentatsioonifailiks. [6]

2. Isesõitvate autode teaduslik taust

2.1 Autonoomsete sõidukite kontseptsioon ja katsumused

Autonoomsete sõidukite idee seisneb sõidukites, mis suudavad iseseisvalt, või vähemalt ilma inimjuhi otsese sekkumiseta, navigeerida ja liikuda. See tähendab, et autonoomsed sõidukid peaksid suutma iseseisvalt teedel navigeerida, tuvastada ümbruskonna objekte, teha otsuseid olukordade ohutuse kohta ja reageerida kõikvõimalikele liiklusolukordadele.

See kontseptsioon on atraktiivne mitmel põhjusel. Esiteks võiksid autonoomsed sõidukid suurendada liiklusohutust, kuna sellised Tehisintellektil põhinevad masinad oleksid võimelised kiiremini reageerima ohtlikele olukordadele ja vähendama inimvigade tõenäosust. Lisaks sellele võiks nende suurem kasutuselevõtt parandada liiklusvooge, vähendades ummikuid ja optimeerides sõiduteede kasutamist. Autonoomsete sõidukite laialdane kasutuselevõtt võib samuti aidata kaasa keskkonnasäästlikumale liiklusele, vähendades heitgaase ja kütusekulu. [8]

Autonoomsete sõidukite rakendamisega kaasnevad paljud raskused ja väljakutsed. Üks suuremaid raskuseid on nende integreerimine traditsioonilise liiklusega, kus toimub suhtlus teiste liikuvate objektidega. Kuna enamik teedel olevad autosid ei ole isesõitvad autod ja iga autojuht peab ka arvestama autoteedele sattunud jalakäijatega, ei saa nende käitumise mõistmiseks neilt arvuti teel infot küsida või siis nende poolt kasutatavate algoritme teades nende tuleviku asukohta välja arvutada.

See nõuab mitte ainult tehnilisi lahendusi, vaid ka kohanemist ühiskondlike ootustega autonoomsete sõidukite rolli osas liiklusmaastikul. Lisaks tuleb lahendada küsimusi seoses küberturvalisusega, kuna autonoomsed sõidukid sõltuvad tugevalt elektroonikast ja andmetöötlusest, suurendades seeläbi potentsiaalseid riske ja väljakutseid, nagu näiteks küberkuritegevus ja andmepüük. [3]

2.2 Isesõitvate autode ohutus

Kuigi autonoomsed sõidukid võivad potentsiaalselt vähendada liiklusõnnetuste arvu ja parandada liiklusohutust, on isesõitvate autode kasutuselevõtt kaasa toonud mitmeid ohutusega seotud väljakutseid. Selle masina süsteem peab tuvastama ümbritsevaid objekte ja

ennustada nende käitumist. See muutub eriti keeruliseks rasketes ja muutlikes liiklusoludes. Isejuhtivate autode süsteemides võivad tekkida tehnilised rikked, nagu sensorite ebaõige kalibreerimine või tarkvaravead, mis omakorda võivad põhjustada jalakäijate või teiste autode ignoreerimise sõiduki juhtimissüsteemi poolt, mis oleks suureks ohuks kõikidele teistele liikluses osijatele. [9, 10]

Isegi siis kui saaks isesõitva auto ohutult liiklema panna peab arvestama ka seda, kuidas teised autojuhid arvestavad sellega, kuidas isesõitev auto liikleb. Nimelt, kui isesõitvad autod käituvad teedel neile loogiliselt, kuid inimestele ettenägematul või tavapärasest erineval viisil, võivad teised juhid olla segaduses või teha ettearvamatuid otsuseid. Näiteks kui isesõitev auto on välja arvutanud, et ta jõuaks üle risttee enne seda kui teine auto talle sisse sõidab, siis ei tähenda see seda, et vastassuuna vööndist tulevad autod selle isesõitva auto otsust oskaksid ette näha. [11]

Peale selle on isesõitvate autode puhul sageli ka eetilisi dilemmasid hädaolukordades, kus süsteem peab otsustama, kuidas reageerida ohuolukorrale, kus sõiduk peab valima kahe halva variandi vahel. Näiteks, kui sõidukil on liiga suur kiirus sees, et pidurdada enne jalakäijateni jõudmist, peab isesõitev auto otsustama, kas jalakäijatest üle sõita või auto peatada mingile betoon objektile sissesõites [12]. Sellistele dilemmadele lisanduvad ka majanduslikud kaalutlused, kuna üldiselt ostetakse vähem autosid, mis on programmeeritud eetilistel põhjustel oma juhi tapma.

2.3 Asukoha määramise tehnikad

Asukoha määramiseks isesõitvate autode jaoks on kasutatud mitmeid erinevaid tehnikaid, mis võimaldavad autodel täpselt tuvastada oma asukohta ja ümbritsevat keskkonda. Üks esimesi katseid autole tema sõiduinfo väliselt kätte anda on pärit 1920-ndatest aastatest, kui Houdina Radio Control demonstreeris raadiojuhitavat "American Wonder" sõidukit New Yorgi tänavatel. See sõiduk oli varustatud antenniga ning sai endale sõiduks vajaliku info raadiosignaalide abil inimestelt, kes sõitsid temaga kaasa teises autos. Selline meetod, kuigi algeeline, andis esimese pilgu sellele, kuidas isesõitvad autod võiksid tulevikus liikuda ja navigeerida. [13]

Esimene katse ilma inimese silmade kasutamiseta auto juhtimiseks tuli 1961. aastal Stanfordi Ülikoolis välja töötatud väikese käruga, mis navigeeris Kuul. Sellel projektil polnud võimalik kasutada kärul olnud kaamerat, mille abil saata maapeal olevatele roboti kontrollijatele infot, kuna sõnumi saatmisel kuule tekkis pikk viivitus, mistõttu roboti hetkene asukoht ei olnud sama, kui see info, mida insenerid maa peal nägid. Seetõttu pidid Stanfordi ülikooli teadlased välja töötama arvutinägemisalgoritmi, kus kärul oli võime mõõta ja hinnata ümbritsevat keskkonda ilma inimese sekkumiseta. [14]

Need varased eksperimendid andsid aluse edasisteks arenguteks asukoha määramise tehnikates. Autonoomsete sõidukite jaoks on hetkel asukoha määramiseks sobivate lahenduste hulgas kaamerad, LiDARid ja radarid.

2.3.1 Kaamerad

Kaamerad on inimsilmale kõige lähedasem sensorid, mida isesõitvad autod kasutavad, lubades neil "näha" ja "tajuda" ümbritsevat keskkonda sama visuaalselt. Kaamerad pakuvad mitmeid eeliseid võrreldes LiDARite ja radaritega, kuna nende poolt loodud pildid on kõige sarnasemad sellele, mida inimsilm näeb, kuna on värvilise kujutise esitamise võime ja tal on suhteliselt odavad hanke kulud. Lisaks on kaamerate kasutamine tuttav paljudele inimestele ja neid on lihtne paigaldada autodele, mistõttu selle lahenduse kasutajaid on isesõitvate autode töötamisel palju. [15]

Kuna isesõitvad autod kasutavad peale kaamerate ka teisi "silmi", võib oletada, et kaamerate kasutamisel on ka puuduseid. Üks suurimaid väljakutseid on halb võimekus tingimustes, kus on nõrk valgustus või tugev kontrastsus, nagu näiteks hämaras või vastuvalguses sõites. Nimelt võivad isesõitvate autode närvivõrgud anda teistsuguseid tulemusi muutunud ilmastikuoludes, ning hoobiski lõpetada vihma või lume puhul piisava täpsusega töötamise, kui isesõitev auto oli algul treenitud päeval vihmatal suveilmal. Samuti võivad kaamerad olla vähem usaldusväärsed ruumiliste objektide korrektselt tuvastamisel või juhtudel, kus on vaja täpset kauguse mõõtmist. [16]

2.3.1.1 Monokulaarne kaamera

Monokulaarse, ehk üksikobjektiiviga kaamera pilt on kujutis, mis pakub ühte vaatepunkti ümbritsevast keskkonnast. Selliseid pilte tegevatel kaameratel on kõrge lahutusvõime, mis

tähendab, et pildil on suur detailide hulk ja need detailid on hea selgusega, kuid sellistel kaameratel puudub sügavuse tajumise võime. Tehniliselt on võimalik pildidel olevate objektide kauguse kaamerast välja arvutada, kui teada nende objekti tegelikke mõõtmeid, kuid ebastandardsete objektide puhul võib sellega tekkida probleeme. [10, 17, 18]

Üksikobjektiiviga kaamera pildid võivad olla, kas must-valged või värvilised ning neid kasutatakse sageli 2D kujutiste või videote salvestamiseks. Monokulaarsete kaamerate kasutuse eelis teiste kaamerate ees on nende odavus ja lihtne paigaldamine. [10, 17, 18]

2.3.1.2 Stereoskoopiline kaamera

Stereoskoopiline pilt on 3D pilt, mida tehakse enamasti stereokaamera abil. Need kaamerad on varustatud kahe või rohkem objektiiviga ning iga objektiivi jaoks on eraldi pildisensor või filmiraam mille kasutamisel tekib sama efekt kui inimsilmade vaatenurkade kattumisel, mis lubab objektide kaugust hinnata. Seega kasutatakse selliseid kaameraid ala kohta ruumilise teabe kogumiseks, näiteks autonoomsete sõidukite süsteemides, kus on oluline täpne kauguse mõõtmine objektideni. [10, 19]

2.3.1.3 Panoraamkaamera

Enamik Panoraampilte tehakse spetsiaalsete turvakaameratena kasutatavate panoraamkaamerate abiga, mis jäädvustavad oma ümbrusest 360-kraadise ühese pildi. Kuigi monokulaarsed kaamerad võivad luua panoraampilte, ei ole need nii efektiivsed kui spetsiaalselt selleks kavandatud panoraamkaamerad. Panoraampiltideks kutsutakse peale 360-kraadised piltide ka horisontaalselt laiasid pilte. [18, 20]

2.3.1.4 Termokaamerad

Termokaamerad, tuntud ka kui Infrapunakaamerad või soojuskaamerad, on andurid, mis tuvastavad infrapunakiirgust, mida inimsilmad ei suuda näha. Infrapuna pilte ei ole monokulaarsed kaamerad võimelised tavaliselt salvestama, kuid neid monokulaarseid kaameraid, mis seda suudavad teha, kutsutakse infrapunakaamerateks. [21]

Infrapuna pildid võivad näidata objektide vahelisi temperatuurierinevusi, mistõttu võivad infrapunakaamerad tuvastada metaani, ammoniaaki, vääveldioksiidi ja muude gaaside lekkeid, kuna need gaasid eraldavad iseloomulikku infrapunakiirgust, mida saab tuvastada ja

mõõta. Peale selle on infrapunakaamerate eeliseks teiste kaamerate ees veel see, et neid saab ka kasutada öiste vaatluste ja halva nähtavuse tingimuste korral. [21]

2.3.2 LiDARid

Light Detection And Ranging, ehk LiDARid on sensorid, mis kasutavad laserkiirt, et mõõta pidevalt kaugusi ümbritsevatest objektidest, luues seeläbi täpse kolmemõõtmelise punkt pilve. LiDARid kasutavad laserkiiri, mis saadetakse objektidele ja ning mille tagasi peegeldunud kiirtelt saadakse info, mille abil mõõdetakse objektide kaugust ja arvutatakse välja nende kuju. [22]

LiDARid erinevad kaamerate ja radaritest, kuna need kasutavad laserkiirt, mitte valgust või raadiolaineid, objektide tuvastamiseks. Nad pakuvad kõrget ruumilist eraldusvõimet ja täpsust ning võimaldavad objektide tuvastamist ja mõõtmist erinevates tingimustes, sealhulgas madala valgustuse ja halbade ilmastikutingimuste korral. Lisaks saavad LiDARid luua kolmemõõtmelisi kaarte ümbritsevast keskkonnast, mis võimaldab autodel täpsemalt tajuda ja reageerida teedel olevatele objektidele. [23]

Isesõitvad autod kasutavad LiDAREid sageli objektide tuvastamiseks ja jälgimiseks nende ümbruses. LiDARid võimaldavad autodel täpsemalt mõõta objektide kaugust ja asukohta ning aidata neil navigeerida keerulistes liiklusolukordades ja teha turvalisi otsuseid teel. Need sensorid on olulised osad autonoomsete sõidukite süsteemides, kuna need pakuvad täiendavat teavet ümbritseva keskkonna kohta, mis aitab autodel ennetada kokkupõrkeid ja tagada sõiduohutus. [23]

LiDARite nõrgad küljed on nende kasutuselevõtmise kõrged kulud võrreldes teiste sensoritega. Peale selle on neil piiratud võimekus objektide tuvastamiseks kõrgemate takistuste või objektide puhul, mis ei peegelda hästi laserkiiri. Lisaks võivad LiDARid olla tundlikud ilmastiku- ja keskkonnatingimuste suhtes ning nende täpsus võib halveneda näiteks peegeldavate klaasidega või tiheda uduga tingimustes. [23]

2.3.3 Radarid

Radarid on sensorid, mis kasutavad raadiolainete saatmist ja vastuvõtmist, et tuvastada objekte nende ümbruses ning mõõta nende kaugust ja kiirust. Need sensorid erinevad kaamerate ja LiDARitest, kuna need sensorid kasutavad erinevat tüüpi signaale objektide tuvastamiseks ning suudavad töötada erinevates ilmastiku- ja valgustingimustes. Radarid on

tuntud oma võime poolest tuvastada objekte suuremate kauguste ja kiiruste peal võrreldes teiste sensoritega ning nende kasutamine on levinud autonoomsetes sõidukites objektide tuvastamiseks ja jälgimiseks. Kuna radareid ei häiri keerulisemad ilmastikutingimused on nad kasutusel ka lennunduses, kus objektide meetripealset täpset asukohta ei ole vaja teada. [24, 25]

Üks Radarite nõrgemaid külgi võrreldes LiDARitega oleks väiksem kauguse mõõtmise täpsus. Peale selle on radard vähem tundlikumad ruumiliste objektide korrektsel tuvastamisel ja keerulistes keskkonnatingimustes, näiteks tiheda liikluse või segavate signaalide korral, kus võib tekkida elektromagnetiline üleküllus. Lisaks võivad radarid olla alati signaali hajutamisele ja moonutamisele teatud tingimustes, mis võib mõjutada nende täpsust ja usaldusväärsust. [19]

Kuna radarid tuvastavad liikuvaid objekte paremini kui teised sensorid kasutavad iseseisvalt autod radareid sageli objektide jälgimiseks erinevates liiklusolukordades. See võimaldab teha turvalisi manöövreid ja reageerida kiiresti autoteedel muutuvatele oludele isegi kehvemate ilmade puhul. Kuna ükski sensor ei ole päris täiuslik kasutatakse radareid tavaliselt koos teiste sensoritega, nagu kaamerad ja LiDARid, et anda autodele täpsem ülevaade ümbritsevast keskkonnast ja aidata neil navigeerida erinevates tingimustes. [24, 25]

2.3.4 Globaalne Positsioneerimissüsteem (GPS)

Globaalne Positsioneerimissüsteem (GPS) on satelliitnavigatsioonisüsteem, mis võimaldab kasutajatel täpselt määrata oma asukohta. Nagu ka satelliitnavigatsioonisüsteemi nimi ütleb kasutab GPS satelliitidelt Maale saadud signaale. Maal asuvad GPSi kasutajad ehk vastuvõtjad mõõdavad nende signaalide ajalisi erinevusi, et kindlaks teha oma asukoht. See süsteem on saavutanud laialdase kasutuse kogu maailmas erinevates valdkondades, sealhulgas navigatsiooni-, transpordi-, militaar- ja tsiviilsektoris. [26]

GPS-i tööpõhimõte põhineb mitme satelliidi signaalide vastuvõtmisel ja nende signaalide põhjal positsiooni määramisel. Iga GPS-satelliit saadab pidevalt signaale, mis sisaldavad satelliidi täpset asukohta ja aega, millal signaal saadeti. Maa peal asuv GPS-vastuvõtja kogub signaale mitmelt satelliidilt ning kasutab neid signaale kolmemõõtmelise positsiooni kindlakstegemiseks. [26]

GPS erineb kaamerateest ja LiDARitest, kuna see kasutab satelliitide signaale, mitte valgust või laserkiiri, asukoha määramiseks. GPS-i eelised hõlmavad suurt kättesaadavust kogu maailmas, kõrget positsioneerimistäpsust ning võimet töötada igasugustes ilmastiku- ja valgustingimustes. GPS-i kasutatakse laialdaselt autonoomsete sõidukite asukoha määramisel, marsruudi planeerimisel ja navigeerimisel, võimaldades neil teha turvalisi ja täpseid manöövreid teel. [26]

Need samad GPS-i eelised on suuresti ka tema nõrkused, kuna signaali võib takistada hoone, mägi või muu takistus, mis võib mõjutada vastuvõetud signaali kvaliteeti ja täpsust. Peale selle saab teda häirida atmosfääri tingimuste muutumine, näiteks tiheda pilvekihi tekkimine signaali soovija kohal, mis saab häirida GPS-signaali stabiilsust ja täpsust. Seetõttu on oluline kaaluda GPS-i kasutamist osana autonoomsete sõidukite navigatsioonisüsteemist, kuid arvestada ka selle võimalike piirangutega ning kaaluda täiendavaid positsioneerimisvõimalusi. [26]

2.4 Andmete salvestamine

Selleks, et tagada andmete tõhus kogumine ja salvestamine, kasutatakse isesõitvates autodes enamasti selliseid seadmeid nagu pooljuhtdraivid (SSD), NAND välgmälu ja NOR välgmälu, mis pakuvad kiiret andmeedastust ja suurt salvestusmahtu. Need seadmed on valitud nende vastupidavuse, kiiruse ja tõhususe poolest. Probleemid võivad tekkida andmete kaotsimineku või korrumppeerumise korral ning seetõttu tuleb tagada, et andmete salvestamise süsteemid oleksid usaldusväärsed ja hästi varundatud. [6]

Kuna autonoomsed autod genereerivad tohutul hulgal andmeid, sealhulgas kujutisi, videoid ja andurite lugemisi, mis võivad kiiresti täita traditsioonilisi salvestusseadmeid, siis peavad salvestusseadmed olema piisavalt suured, et neid läbi töödelda. Samuti peavad need olema piisavalt vastupidavad, et taluda sõiduki liikumisest tingitud vibratsiooni ja muutuvaid ilmastikutingimusi. [6]

3. Fiery mudelilt andmete kogumine

Bakalauruse tööd alustati Fiery mudeli süva analüüsiga, vaadates, mis andmeid seesisse võtab ja mis asju see mudel väljundiks genereerib. Seejärel lõin pythoni `.py`¹ faili millega saab jooksutada eeltreenitud Fiery mudelit uute plokkide (inglise keeles *batch*) loomiseks valitud algandmete põhjal. Seda pythoni faili jooksutades genereerib Fiery mudel sellele sisse antud piltide põhjal jalakäijate ja teel ning selle ümbruses olevate isikute tuvastustus tõenäosusega segmentatsiooni faili (`.p`), mille järgi saab hinnata kui hästi mudel inimesi või autosid tuvastab.

Seejärel kirjutasin Jupyter Notebookis `.ipynb`² faili, kuhu kirjutasin koodi Fiery genereeritud segmentatsiooni hindamiseks ja plokilt piltide välja võtmiseks. Selleks, et mõõta mudeli täpsust oleks mudelile vaja sisse anda uusi pilte, mille peal seda ei ole veel treenitud, algpildid millega ta on treenitud on võetud NuScene'i valideerimisandmestikust.

Täiesti uute piltide plokile lisamine osutus liiga keeruliseks, kuna see eeldaks piltide omavahelise asetuse välja arvutamist ja Fiery mudeli kaamerate ja piltide andmete muutmine koos piltide salvestamisega õigetes kogumisse. Kuna igale plokile oleks vaja 2 mineviku, ühte oleviku ja 4 tuleviku pilti, mida ei saanud lihtsalt asjade ülevaatamisega hinnata, sest algandmetes isesõitvad autod liiguvad erisuundades ja seetõttu peaks ka seda arvesse võtma. Seega sai plaaniks Fiery mudelit jooksutada originaalsetel piltidel, kus jalakäijaid veel ei ole ja kus auto on seismas ühes kohas, mis hõlbustas protsessi ja võimaldas saada selget ülevaadet mudeli toimimisest.

Mudelile sama hea võimaluse jalakäiate tuvastamiseks andmiseks nagu ta autoteel töötaks otsustasime lisada sinna sama kvaliteediga pildelt võetud jalakäiaid, kes olid teistel NuScene'i piltidel. Selliseid pilte lõikasin välja Gimp³iga ja salvestasin `.png` failiformaati, kus fail ei täida tühjasid alasid default pikslitega vaid jätab need tühjaks. Seetõttu on `.png` formaadilt võimalik teise pildi peale kleepida *for*-tsükli kasutades vaid algsest pildist välja lõigatud osa. Kuna see bakalaureuse töö keskendub Fiery mudeli jalakäijate tuvastamise võime analüüsimisele, siis teiste autode tuvastamisega see töö ei tegele.

¹ <https://www.python.org/>

² <https://jupyter.org/>

³ <https://www.gimp.org/>

Viimaks salvestasin pildi plokki tagasi ja jooksutasin *.py* failiga Fiery mudelit, mille tulemuse sai *.ipynb* failis hinnata. Selleks, et teha isiku tuvastamine pildilt võimalikult heaks, tuli kõik teised pildid plokis asendada oleviku piltidega, jättes plokile mulje, et auto seisab, tehes mudeli oletused staatiliseks ja stabiilsemaks. Siit edasi üritasin teha võimalikult suurt *for*-tsüklit, et suurendada ühe korraga *.py* faili jooksutamisega saadavad *.ipynb* failis uuritavad tulemused, üritades samal ajal hoidudes programmikoodi liigselt segaseks muutumisest.

4. Fiery mudeli analüüs

4.1 Eeldused Fiery mudeli analüüsimisel

Seda bakalaaurusetööd alustati, kuna Markus Kängseppal oli kahtlus, et Fiery mudel tuvastab maas lebavat isikut kehvemini kui seisvat isikut. Selleks lisati pildile isikuid, keda oli kallutatud 90° või -90° selleks, et jätta muljet maas vedelevast isikust. Selle oletuse kinnitamiseks sai tõstatatud vastupidine eeldus, et Fiery mudel ei tuvasta maas vedelevaid isikuid seisvatest inimestest kehvemini kui 50%.

Samuti oli bakalaureusetöö algseks eelduseks, et Fiery mudeli tuvastamisvõime langeb pimedate piltide korral, kus enamikul pildil puudub selge valgustus, kuid seal on piisavalt valgust, et selle mudeli hindamisvõimet saaks analüüsida (Joonis 12), nii palju, et mudel muutub liiga kehvaks. Selle oletuse kinnitamiseks oli vajalik uurida vastupidist eeldust, et pimedatel aegadel tehtud pildid, kus on jalakäija, ei näita kehvemat tuvastusprotsenti kui 80% võrreldes vihmatutel päevadel tehtud piltidega, kus jalakäija on tuvastatud.

Kui Fiery mudeli äratundmiseksust võiks erineda pimedatel oludel, siis on vaja ka uurida, kuidas selle tuvastussüsteemi mõjutab ka valguserohkne, kuid vihmane ilmaolu, sest isegi siis kui Fiery ei suudaks pimedas inimesi tuvastada, siis ei tähenda see seda, et probleemiks võib olla pildi hägusus. Selle kohta info saamiseks oli uurimise käigus vaja testida mudeli jõudlust erinevates ilmastikuoludes, et hinnata selle reaalses maailmas kasutatavust erinevates olukordades, alates päikesepaistelisest ilmast kuni vihma, lume või udu tingimusteni.

Fiery mudeli testimise alguses ei olnud sinna lisatavad isikud alati kõige parema kvaliteediga piltidest, kuid uute isikutega piltide loomisel tekkis küsimus, kas mudel tuvastab taustaga võrdse teravusega inimesi paremini kui pildile lisatud inimesi, kelle tunnused olid, kas selgemad või kehvemad kui taustal. Selle all mõtleme seda, et nende jalakäijate tuvastusprotsent oleks väiksem juhtudel, kui lisatud inimese pilt ei sobi väga hästi tema taustaga kokku. Seda hüpoteesi saaks tõstatada, kui ümber lükata eelduse, et süsteem tuvastas kehvema kvaliteediga pilte, mida siin klassifitseeritakse kui isikut piltidel 1, 5 ja 6, mitte rohkem kui 50% kehvemini kui isikut piltidel 2, 3 ja 4 nende kõige paremini tuvastatud kaldes ja ka paremini kui isik, kelle kvaliteet on tunduvalt suurem kui teistel piltidel, milleks tegin pildi 7.



Joonis 2: Vasakult paremale on esimene pilt failist *person_1.png*, teine isik failist *person_2.png*, kolmas isik failist *person_3.png*, neljas isik failist *person_4.png*, viies isik failist *person_5.png* ja kuues isik failist *person_6.png*. Seitsmes ning viimane pilt lugedes vasakult on Bakalaureusetöö autori enesepilt, mida mujal kutsume pildiks seitsmendaks isikuks, failist *person_7.png*.

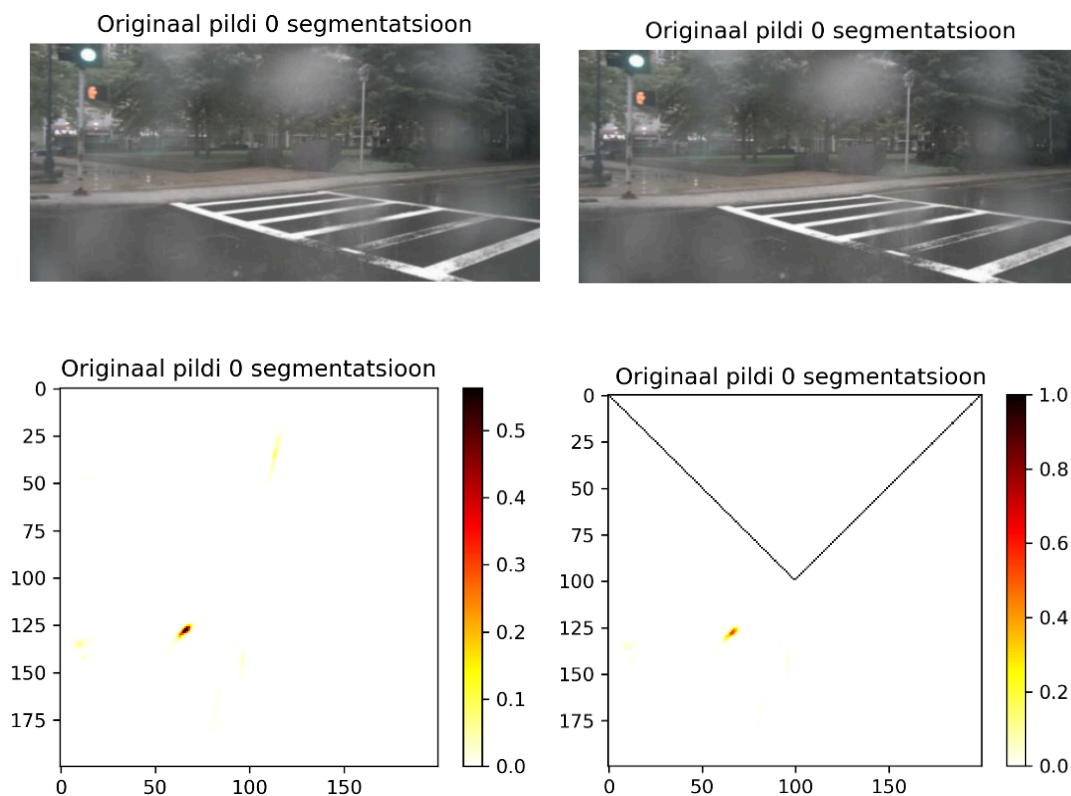
Lisaks oli eeldus, et mida paremini isik pildile sobib, seda paremini Fiery mudel ta ära tuvastab. See eeldus tugines arusaamale, et kui isik on pildil loomulikus keskkonnas ja positsioonis, on tema tuvastamise jaoks tal mingi sidepunkt võtta oma treeningandmetest. Selle hindamine on subjektiivsem ja selle täpse tõestamiseks puuduvad selged kvantitatiivsed mõõdikud. Kuigi see uurimustöö ei avalda siin kontekstis kindlaid numbreid, oleks siiski oluline arutada ja väljendada, kuidas eeldus paika pidas.

Inimesed kalduvad loomulikult tähelepanu pöörama objektidele, mis asuvad pildi keskosas või on esiletõstetud. Fiery mudeli analoogselt toimumise juhul, võiks tema tuvastamisvõime olla inimese omale sarnane ja optimaalne juhul, kui objektid asuvad pildi keskmes. Seega oli vajalik uurida, kuidas mudel reageerib erinevate objektide paigutamisele pildidel ning millist mõju see avaldab tuvastustäpsusele.

Lisaks eelmainitud eeldustele oleks hea ka teada, kas Fiery mudel suudab paremini tuvastada pilte, kus kujutatakse üksteisele lähedal kahte inimest. See hüpotees tugineb arusaamale, et mudel võimekus tuvastada inimeselaadset kujundit pildi peal suureneb, kui nende lähedal on inimeste sarnaseid pikslikogumikke. Selle hüpoteesi testimiseks vaatame pilti *Person_3.png* ja teiste piltide tulemuste erinevusi, keskendudes teistele paremate kvaliteetidega piltide tuvastustulemustele.

4.2 Tulemuste analüüs

Tulemuste analüüsimiseks käsitlesin segmentatsioonifaile, mille kõrval hoian selle kohta käivat vajalikku informatsiooni. Ehkki muudatus toimus vaid ühel pildil kuuest, annab analüüsitav segmentatsioonifail kõigi kuue kaamera pildist kokku pandud inimeste ja autode tuvastatud ala kaardi. Analüüsi visualiseerimiseks on siin pildid alaga kus muudatus on tehtud.



Joonis 3: üleval on kaks pilti enne isiku lisamist ja nende all on pildid neile kuuluvatest segmentatsioonifaildest. All paremal on segmentatsioonifail peale nende piltide poolt mõjutatava ala ülejäänud failist välja piiritlemist ja all vasakul on pilt enne selle tegemist. Segmentatsioonifailile tulemuste kuvamiseks piiri lisades fikseerima skaala, kuna on tähtsam teada kas Fiery mudel tuvastab objekti vähemalt 70 % tõenäosusega ja kuigi 0.1 % tõenäosusega tulemuse visualiseerimine näitab, et mudel siiski tuvastab midagi, ei aita see tulemuse visualiseerimisele soovitud viisil kaasa.

Üleval olev pilt on see, mida me modifitseerima hakkame, ja paremal on V märgi sees ala, mida selle muutmisel Fiery tuvastusmeetod hakkab muutma.

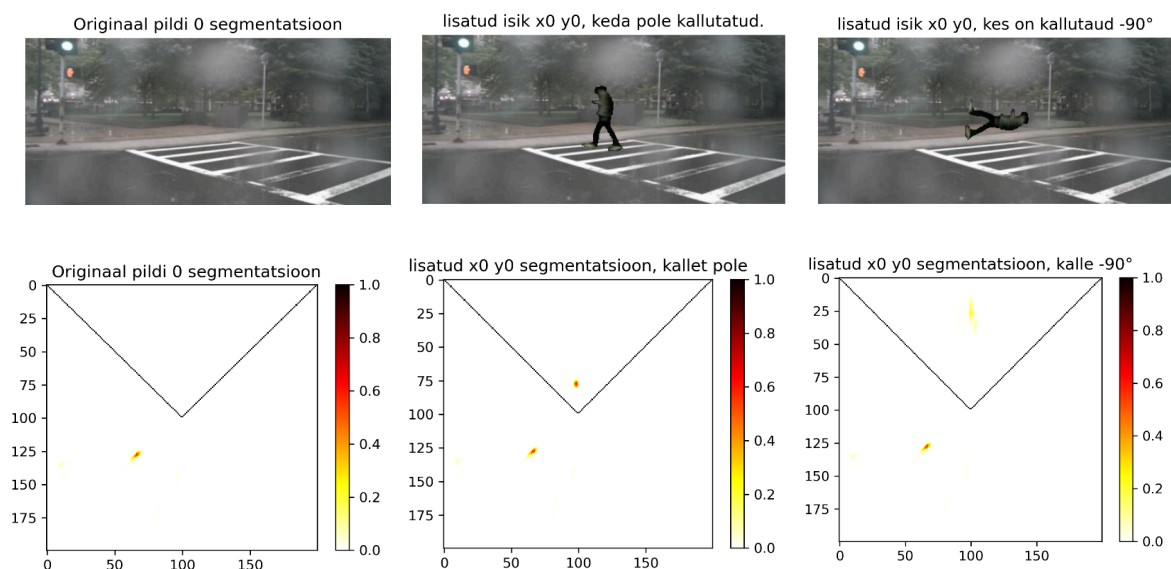
Tuvastustõenäosuste summa on siin töös käsitletud kui kõikide käsitletud piltide tõenäosuste summa. Tuvastusprotsendi arvutamise valem on kõikide käsitletud piltide tõenäosuste summa (igal pildil võis Fiery mudel tuvastada isikuid väärtusega ühest nullini) jagatud kõikide käsitletud piltide arvuga.

Analüüsidest Fiery mudeli efektiivsust kaldasendis isikute tuvastamisel, tuli välja, et mudeli tuvastustõenäosus varieerub rohkem kui 50% pöörates pildile lisatavat isikut vähemalt 45°. Pöörates pildile lisatavat isikut 90° oli tuvastustõenäosus rohkem kui 90% väiksem kui kallutamata isiku puhul, andes kõige kehvemad tulemused -90° ja 90° kraadise kalde puhul.

Tabel 1: Fiery isikutuvastus vastavalt isiku kaldele. Iga kaldel oli 126 pilti analüüsimiseks.

Kalle	Tuvastustõenäosuste summa	Tuvastamise tõenäosus
-90°	1.9	1.51 %
-45°	17.6	13.97 %
0°	98	77.78 %
45°	29.9	23.73 %
90°	9.2	7.30 %

Tabel 1 esitab Fiery mudeli tuvastustõenäosust erinevates kalletes. Siit on välja jäetud pildi äärtel asuvaid isikuid, defineeritud kui isikud, kes lisati pildi keskkohast vertikaalselt kaugemale kui 200 pikslit, kuna mudeli tuvastus nendes piirkondades oli puudulik.



Joonis 4: vasakul on üleval pilt ja all tema segmentatsioon enne isiku lisamist, keskel üleval on isik lisatud pildi keskele ja all on pilt tema segmentatsioonifailist, kus on näha, et Fiery mudeli ennustatud jalakäija paiknemise tõenäosus on pildile lisatud alas 1, ehk seal on mudeli meelest kindlasti üks isik. Vasakul üleval on isik lisatud küllili pildile ja all on tema segmentatsioon, kus on määratud isiku paiknemise tõenäosusega 0.2.

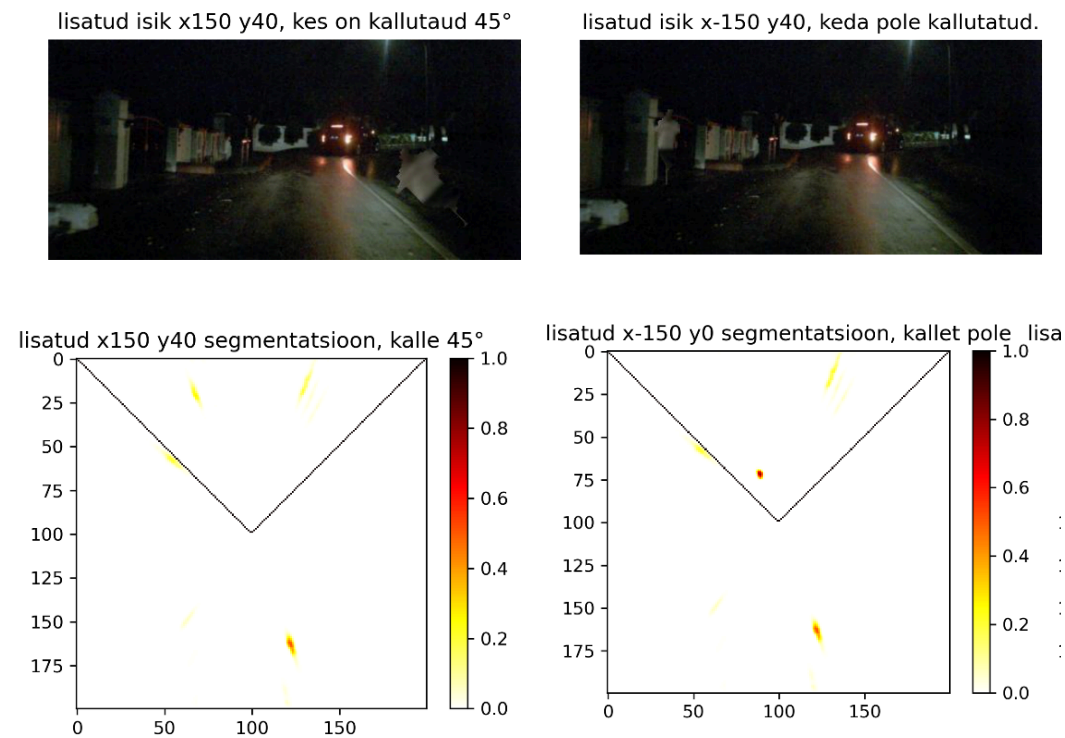
Uurides Fiery mudeli tuvastamisvõimet pimedate tingimustega olukordades, kus oli siiski piisav valgustus, et mudeli hindamist oleks võimalik läbi viia, tuli siin uurimuses välja, ja viimasel päeval, Fiery mudeli tuvastustõenäosus ei muutunud rohkem kui 80% nende ilmaolude muutumisel (Tabel 2). Plokk (inglise keeles *batch*) 13 korral (Joonis 9), mis oli vihmatal ja pilvetul päevase ajal tehtud piltide põhjal loodud plokk, ei olnud tulemused nii erinevad kui plokk 34 (Joonis 10), mis oli vihmatal ja päeval ajal loodud piltide põhjal kokku pandud plokk või plokk 68 (Joonis 11), mis oli öisel ajal tehtud piltide korral loodud plokk, nii et algset eeldust ei saanud selle töö puhul vastu võtta.

Tabel 2: Fiery isikutuvastus vastavalt ilmale ja päevaajale.

plokid	Tuvastustõenäosuste summa	Tuvastamise tõenäosus
plokk 13	60.4	28.76 %
plokk 34	48.5	23.1 %
plokk 68	47.7	22.71 %

Tuvastustõenäosus on Fiery poolt korrektseks määratud isiku paiknemise tõenäosus jagatud kõikide piltide arvuga, kuhu lisati isik.

Joonisel 5 on näha mudeli tendentsi, et isegi siis kui isikud, mis lisati pildile nii, et nad nägid seal inimsilmale loomulikud ja sobivad, ei tahtnud mudel neid väga tuvastada. Peaaegu kõik sellised juhtumid olid kalde all lisatud inimeste korral.



Joonis 5: Vasakul üleval pildil on öisest ajast pilt, kus paremal nurgas kukub üks isik autoteele. Vasakul all on segmentatsioon, kus süsteem ei tuvasta tema asukohta ära. Paremal üleval on pilt isikuga seina ääres, ning all on segmentatsioon, kus ta on ära tuvastatud.

Hinnates Fiery mudeli võimet tuvastada erineva kvaliteediga isikuid pildil, saime tulemused, et mudel tuvastas pilte 2, 3 ja 4, mille kvaliteet oli parem kui 1, 5, 6 ja kehvem kui 7 pildi oma üldiselt paremini, kuid erinevused tuvastustõenäosuses ei ole piisavalt suured, et võtta algne eeldus vastu. Välja peab tooma ka pildi 6, mille korral on isiku tuvastamise tõenäosus kõikides asendites kokku suurem kui ühegi teise pildi puhul.

Kõikidest teistest parema kvaliteediga pildi 7 korral tuli tuvastusprotsent kõikide teiste isikutega võrreldes kõige madalam, andes kõige madalama tulemuse kõikide kokku loetud

kallete korral tehtud testides, tuli tulemuseks tuvastusprotsent 20.4%. Püstiseisu asendis oli olukord parem, andes paremaid tulemusi kui enamik kehva kvaliteediga pilte, kuid kehvema kui enamik parema kvaliteediga pilte. Siit saab tagantjäreli järeldada, et kui uurimuse alguses jagati pildid parema ja halvema kvaliteediga piltide gruppi, siis võisid mõned neist olla selle mudeli kontekstis olla ekslikud.

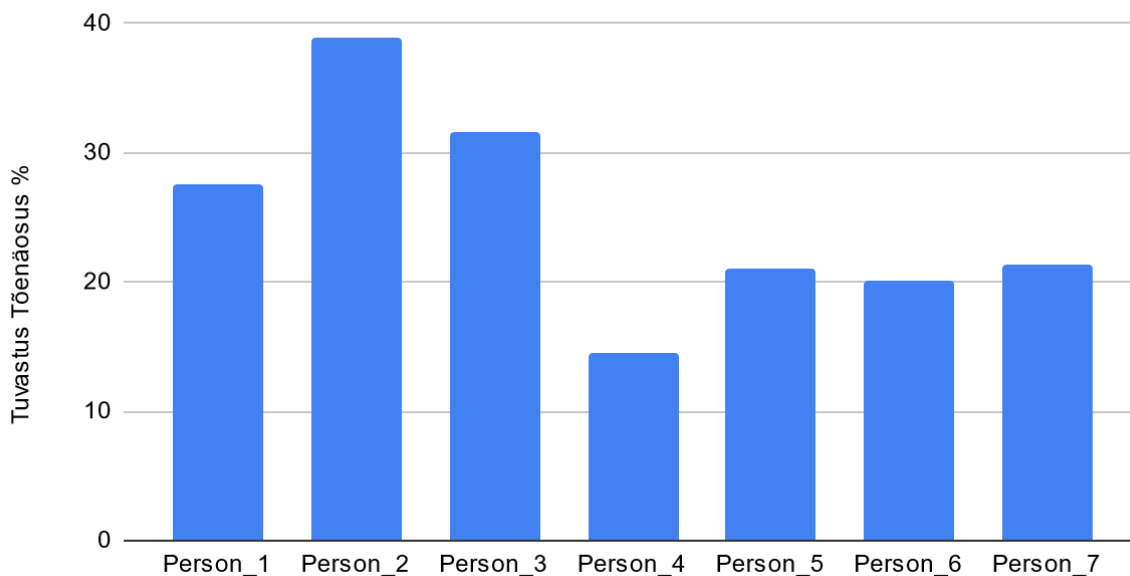
Tabel 3: Fiery isikutuvastus vastavalt pildile lisatud isikule, juhul kui ta kalle 0

Lisatud isik	Tuvastustõenäosuste summa	Tuvastamise tõenäosus
<i>Person_1.png</i>	18	100 %
<i>Person_2.png</i>	15.2	84.44 %
<i>Person_3.png</i>	17.2	95.56 %
<i>Person_4.png</i>	8.6	47.78 %
<i>Person_5.png</i>	12.2	67.78 %
<i>Person_6.png</i>	12.2	67.78 %
<i>Person_7.png</i>	14.6	81.11 %

Tabel 5: Fiery isikutuvastus vastavalt pildile lisatud isikule eemaldades ekstreemumkohad x200, y0, x200, y40, x-200, y0 ja x-200, y40 .

	Tuvastustõenäosuste summa	Tuvastamise tõenäosus
<i>Person_1.png</i>	24.8	27.56 %
<i>Person_2.png</i>	35	38.89 %
<i>Person_3.png</i>	28.5	31.67 %
<i>Person_4.png</i>	13.1	14.56 %
<i>Person_5.png</i>	18.9	21 %
<i>Person_6.png</i>	18.1	20.11 %
<i>Person_7.png</i>	19.2	21.33 %

Fiery poolt erinevate isikute tuvastus tõenäosus ilma ekstreemumiteta



Joonis 6: Fiery isikutuvastus vastavalt pildile lisatud isikule eemaldades ekstreemumkohad x200, y0, x200, y40, x-200, y0 ja x-200, y40 .

Eeldus, et Fiery mudel suudab paremini tuvastada pilte, kus kujutatakse üksteisele lähedal kahte inimest, võib selle uurimuse puhul pidada tõseks, kuna hoolimata sellest, et Person_3.png tuvastustõenäosus ei olnud kõikide piltide pealt kokku arvatud tuvastustõenäosus suurem kui teistel piltidel, püsis ta edukalt esikolmikus. Seega võib selle uurimuse puhul öelda, et Fiery mudel tuvastab inimesi suuremates kogustes paremini kui üksikjalakäiaid.

Eeldus, et mudel tuvastab jalakäijaid paremini siis, kui nad on lisatud pildi keskosale lähemale sai igati kinnitust. Kuigi on võimalik, et pildi suurendus ta pildi alla ossa tõstmisel oli natuke paigast ära, mistõttu annab mudel allosas kehvemaid tulemusi, ei olnud see siiski märgatav, kuna ülemisel paigutusel, kus objektid asusid keskosaga samal y-teljel, oli jalakäijate tuvastamise tõenäosus 17.9 % ja keskosa y-teljest 40 pikslit all, koos suurendusega olid isiku pildil tuvastus 12.83 %. On oluline mainida ära Fiery mudeli treenimisega loodud eripära. Kuigi mudel ei kasuta LiDAR-i ega stereoskoopilisi kaameraid, kasutades vaid monokulaarseid kaameraid, suudab ta siiski hinnata objektide kaugust kaamerast, kuna teda

on treenitud ootama inimesi teatud kõrguse ja laiusega kindlatel pildi kõrgustel. Selleks, et jalakäijate proportsioonid pildil oleksid nende asukoha jaoks mudelile realistlikud, pidime lähedal asuvad jalakäijad pildil suuremaks muutma ja kaugemal asuvad jalakäijad väiksemaks.

Mudel ei töötanud optimaalselt juhul kui isik lisati pildi äärtesse. Ekstreemsete lokatsioonide korral, mis on pildi vasakul ja paremal äärtel, mille asukohad olid x_{200} ja y_0 , x_{200} ja y_{40} ja x_{-200} , y_0 ning x_{-200} ja y_{40} , ei tuvastanud Fiery mudel peaaegu ühtegi sinna lisatud isikut, isegi siis kui sinna lisatud isikud olid püstiasendis.

Siin töös kasutatav asukohtade märkimis süsteemi näitab isiku keskkoha kaugust pildi keskkohast. Pilt mille keskkoha asukoht on märgitud kui x_0 ja y_0 on pandud pildi keskele, ning kui pildi asukoht on märgitud kui x_{200} ja y_{40} siis on pildile lisatud isiku keskkoha asukoht 200 pikslit paremale ja 40 pikslit allapoole. Pildi keskkohast allapoole lisatud isikute suurust pidi suurendama, et Fiery mudel saaks nad ära tunda.

Põhjusteks võis olla äärtele lisatud isikute korral mudelil soov saada kinnitust teistelt kaameratelt, kuna liikuva auto korral ei ole nad väga ohtlikes kohtades eestvaates ja juhul kui info nende kohta on juba eestvaates, siis üldiselt kõrvalevaate kaamerad tuvastasid nad ära kui auto edasi sõidaks.

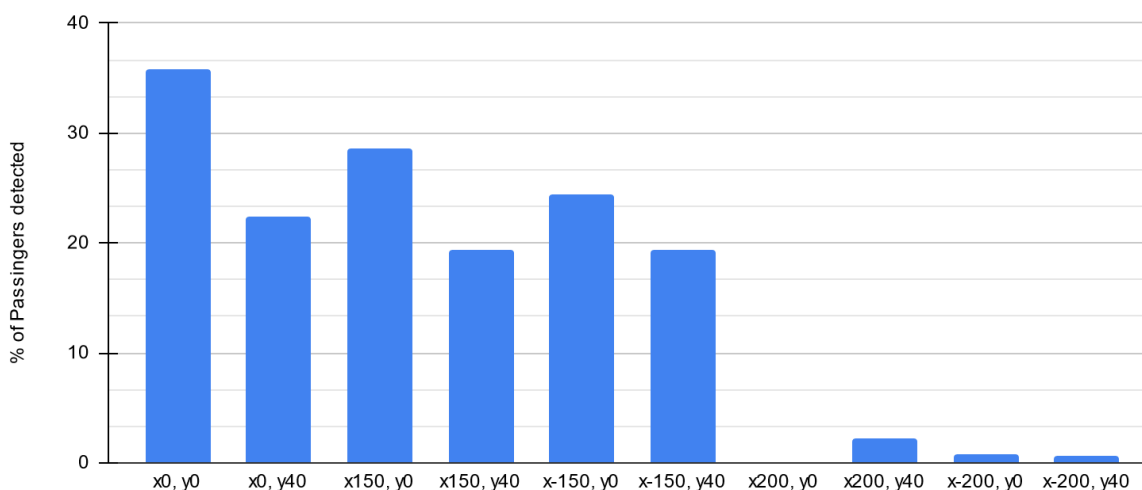
Kuna paigaldades isiku pildil ekstreemsetesse äärtesse tulid tulemused nii ekstreemselt kehvad, et jäi mulje et mudelil ei ole nende äärte tuvastamine tehtud esikaameraga, kuigi teised kaamerad seda ala ka näha ei saa (Joonis 8), ei kasutanud positsioone x_{200} , y_0 , x_{200} , y_{40} x_{-200} , y_0 ja x_{-200} , y_{40} ülejäänute eelduste, nagu pildi kvaliteedi (Tabel 4), analüüsis.

Tabel 5: Fiery isikutuvastus vastavalt pildile lisatud isiku asukohale.

Kus	Tuvastustõenäosuste summa	Tuvastamise tõenäosus
x_0, y_0	37.6	35.81 %
x_0, y_{40}	23.6	22.48 %
x_{150}, y_0	30	28.57 %
x_{150}, y_{40}	20.4	19.43 %

x-150, y0	25.6	24.38 %
x-150, y40	19.4	18.48 %
x200, y0	0	0 %
x200, y40	2.4	2.29 %
x-200, y0	0.8	0.76 %
x-200, y40	0.6	0.57 %

The Chance of Fiery detecting the pedestrian by Location



Joonis 7: Erinevates asukohtades saadud tuvastustõenäosus.

4.3 Võimalikud Edasiarendused

Kuigi see uurimistöö andis algse ülevaate Fiery mudeli jõudluse mõistmisel, ei tohiks see kindlasti olla viimane sarnane uurimus. Selle bakalaaurusetöös saadi hinnang Fiery mudeli võimekusele tuvastada objekte erinevates tingimustes ja keskkondades ning arusaam, kuidas isiku kalle piltides mudelit mõjutas.

Töö käigus tekkis küsimus, kas objekti paigutamise kõrgus mõjutab Fiery mudeli efektiivsust. Kahjuks ei olnud piisavalt aega, et testida kõikide võimalike piltide suurus selles valdkonnas. Seega oleks mõistlik jätkata uurimist eesmärgiga aru saada, mis on objektide paigutamise mõju Fiery mudeli tuvastamise täpsusele.

See võiks hõlmata ka siin töös probleemiks tulnud Fiery mudeli raskustega tuvastada objekte oma pildi äärtel. Samuti jäi siin töös vastamata küsimus, kuidas jagab mudel infot kaamerate vahel, mida võiks alustada testides kui hästi Fiery mudel tuvastab külgkaamerate piltidele lisatud isikuid, kuna sealsete kaamerate vaateala kattub kergelt.



Joonis 8: kaamerate piltide tuvastusalad [1].

Kuna see töö oli tehtud seisva autoga, on oluline kaaluda, kuidas see mudel töötab liikuva auto korral. Kuna enamik pilte, mille peal mudelit on treenitud, olid tehtud autodel, mis olid tol hetkel liikumas, võiks eeldada, et isikute tuvastamine muutuks sellel mudelil täpsemaks liikuvate autode korral.

Kokkuvõte

Käesoleva töö eesmärgiks oli Fiery mudeli headuse hindamine isikute tuvastamises erinevates keskkondades ja olukordades. Fiery mudeli jõudluse analüüsimiseks loodud kood lubas Fiery mudelit jooksutada korraga mitme plokki peal ja selle hindamiseks kirjutatud *.ipynb* faili lubas neid plokkide korraga analüüsida. Tulemused näitasid, et Fiery mudeli tuvastustäpsus sõltub suures osas objektide kallutusnurgast, kuid mitte isiku asukohast pildil, jättes siin välja pildi ääred. Peale selle võib mudelil olla raskusi objektide tuvastamisel nende pildikvaliteedi tõttu, eriti juhtudel kui isikute tuvastuskvaliteet oli võrreldes ülejäänud pildiga liiga kõrge. Lisaks näitas töö, et Fiery mudel ei ole oluliselt kehvema jõudlusega vihmasel või öisel ajal võrreldes päevase ajaga. Edasised Fiery mudeli uuringud võiksid vaadata mudeli efektiivsust dünaamilisemates olukordades, näiteks liikuva auto korral. See võiks anda täiendavat arusaama selle mudeli headusest.

Viited

- [1] Anthony Hu1, Zak Murez, Nikhil Mohan, Sofia Dudas, Jeffrey Hawke, Vijay Badrinarayanan, Roberto Cipolla, Alex Kendall, Wayve, UK. University of Cambridge, UK. "FIERY: Future Instance Prediction in Bird's-Eye View from Surround Monocular Cameras.", pages 1-14, 2023. Kättesaadav: <https://arxiv.org/pdf/2104.10490v3.pdf> [Külastatud: 17 aprill 2024]
- [2] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ArXiv. Kättesaadav: <https://arxiv.org/pdf/1905.11946> [Külastatud: 3 mai 2024]
- [3] Ramadoss, V. (2018). Squeeze-and-Excitation Networks. Medium. Kättesaadav: <https://medium.com/@Vinoth-Ramadoss/squeeze-and-excitation-networks-84e3db0e04e2> [Külastatud: 15 mai 2024]
- [4] NuScenes ametlik veebileht, Kättesaadav: <https://www.nuscenes.org/> [Külastatud: 15 mai 2024].
- [5] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, Oscar Beijbom, "nuScenes: A multimodal dataset for autonomous driving", nuTonomy: an APTIV company. Kättesaadav: https://openaccess.thecvf.com/content_CVPR_2020/papers/Caesar_nuScenes_A_Multimodal_Dataset_for_Autonomous_Driving_CVPR_2020_paper.pdf [Külastatud: 15 mai 2024]
- [6] "What is semantic segmentation?." IBM ametlik veebileht, Kättesaadav: <https://www.ibm.com/topics/semantic-segmentation> [Külastatud: 15 mai 2024].
- [7] "Segmentation File." Spraak. Kättesaadav: https://www.spraak.org/documentation/spr_dt_seg.html [Külastatud: 15 mai 2024].
- [8] Zhang, L. "Human Ridehail Crash Rate Benchmark." Cruise LLC. Kättesaadav: <https://www.getcruise.com/news/blog/2023/human-ridehail-crash-rate-benchmark/?ref=warpnews.org> [Külastatud: 15 mai 2024].
- [9] Manikandasriram Srinivasan Ramanagopal, Cyrus Anderson, Ram Vasudevan, Matthew Johnson-Roberson. "Failing to Learn: Autonomously Identifying Perception Failures for Self-Driving Cars." IEEE Xplore. Kättesaadav: <https://ieeexplore.ieee.org/abstract/document/8412512> [Külastatud: 7 mai 2024].
- [10] De Jong Yeong, Gustavo Velasco-Hernandez, John Barry, Joseph Walsh. "Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review." National Center for

- Biotechnology Information, U.S. National Library of Medicine, 2021. Kättesaadav: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8003231/> [Külastatud: 12 mai 2024].
- [11] Surden, Harry, Mary-Anne Williams. "Technological Opacity, Predictability, and Self-Driving Cars." *Cardozo Law Review*, vol. 38, no. 1, 2016, lk 125. HeinOnline. Kättesaadav: <https://heinonline.org/HOL/Page?handle=hein.journals/cdozo38&id=131&collection=journals&index=> [Külastatud: 10 mai 2024].
- [12] Mayer, Maike M., Raoul Bell, Axel Buchner. "Self-protective and self-sacrificing preferences of pedestrians and passengers in moral dilemmas involving autonomous vehicles." National Center for Biotechnology Information, U.S. National Library of Medicine, 2021. Kättesaadav: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8700044/> [Külastatud: 12 mai 2024].
- [13] Time Magazine. Science: Radio Auto. Kättesaadav: "https://content.time.com/time/subscriber/article/0,33009,720720,00.html". [Külastatud: 9 aprill 2024]
- [14] Stanfordi Ülikooli ametlik veebileht, Les Earnest, "Stanford Cart" , Kättesaadav: <https://web.stanford.edu/~learnest/sail/oldcart.html> [Külastatud: 13 aprill 2024]
- [15] J. Stähler, C. Markgraf, M. Pechinger, and D. W. Gao, "High-Performance Perception: A camera-based approach for smart autonomous electric vehicles in smart cities," in *IEEE Electrification Magazine*, vol. 11, no. 2, pp. 44-51, June 2023. Kättesaadav: <https://ieeexplore.ieee.org/abstract/document/10143714> [Külastatud: 12 aprill 2024]
- [16] J. Vargas, S. Alsweiss, O. Toker, R. Razdan, and J. Santos, "An Overview of Autonomous Vehicles Sensors and Their Vulnerability to Weather Conditions," *Sensors*, vol. 21, no. 16, article no. 5397, Aug. 2021. Kättesaadav: <https://www.mdpi.com/1424-8220/21/16/5397> [Külastatud: 10 aprill 2024].
- [17] Cohen, N. "Are Two Cameras Better Than One?" *Visionary AI Blog*. Kättesaadav: <https://www.visionary.ai/blog/are-two-cameras-better-than-one> [Külastatud: 15 mai 2024].
- [18] "Panoramic Cameras." *Photonics.com*. Kättesaadav: https://www.photonics.com/EDU/panoramic_camera/d5933 [Külastatud: 15 mai 2024].
- [19] Yurtsever, E., Lambert, J., Carballo, A., Takeda, K. "A Survey of Autonomous Driving: Common Practices and Emerging Technologies." *IEEE Xplore*. Kättesaadav: <https://ieeexplore.ieee.org/abstract/document/9046805> [Külastatud: 15 mai 2024].

- [20] Fu, M., Wang, X., Ma, H., Yang, Y., & Wang, M. (2014). Multi-lanes detection based on panoramic camera. *IEEE*. Kättesaadav: <https://ieeexplore.ieee.org/document/6870997> [Külastatud: 10 aprill 2024].
- [21] Gade, R., & Moeslund, T. B. "Thermal cameras and applications: a survey." *Machine Vision and Applications*, 24(3), 537-541. Kättesaadav: <https://link.springer.com/article/10.1007/s00138-013-0570-5> [Külastatud: 15 mai 2024].
- [22] "What is lidar?," National Ocean Service, U.S. National Oceanic and Atmospheric Administration. Kättesaadav: <https://oceanservice.noaa.gov/facts/lidar.html> [Külastatud: 21 aprill 2024].
- [23] You Li, Javier Ibanez-Guzman. "Lidar for Autonomous Driving: The Principles, Challenges, and Trends for Automotive Lidar and Perception Systems." *IEEE Signal Processing Magazine*, vol. 37, no. 4, July 2020. Kättesaadav: <https://www.sciencedirect.com/science/article/pii/S2590198220301354> [Külastatud: 15 mai 2024].
- [24] Taohua Zhou, Mengmeng Yang, Kun Jiang, Henry Wong, Diange Yang. "MMW Radar-Based Technologies in Autonomous Driving: A Review." *MDPI, Sensors*, 2020. Kättesaadav: <https://www.mdpi.com/1424-8220/20/24/7283> [Külastatud: 15 mai 2024].
- [25] Igal Bilik, Oren Longman, Shahar Villeval, Joseph Tabrikian. "The Rise of Radar for Autonomous Vehicles: Signal Processing Solutions and Future Research Directions." *IEEE Xplore*, 2019. Kättesaadav: https://ieeexplore.ieee.org/abstract/document/8828025?casa_token=1intOLPrSwEAAAAA:r_mMeSEUruX65iXXm4q70kvd9MKTMBYRDlThVw2xmOP0OAQVgsLE19DSSpJqUliugZzaQeCR2XSeq [Külastatud: 15 mai 2024].
- [26] "What is GPS: Getting Started." *GPS Central*. Kättesaadav: <https://www.gpscentral.ca/what-is-gps-getting-started/> [Külastatud: 15 mai 2024].

Lisad



Joonis 9: plokis 13 muudetav pilt



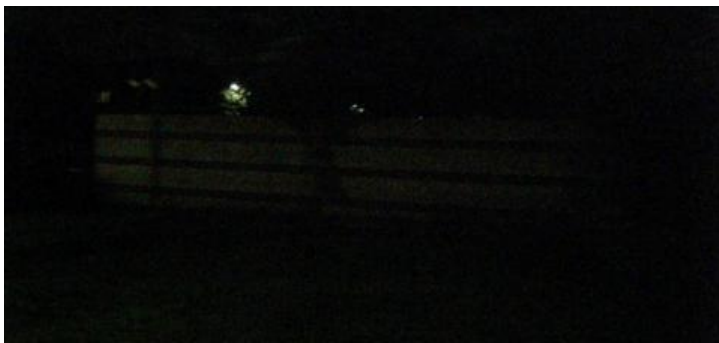
Joonis 10: plokis 34 muudetav pilt



Joonis 11: plokis 68 muudetav pilt

Tabel 4: Fiery isikutuvastus vastavalt pildile lisatud isikule kõikide kallete korral.

	Tuvastustõenäosuste summa	Tuvastamise tõenäosus
<i>Person_1.png</i>	24.8	24.31 %
<i>Person_2.png</i>	36	35.29 %
<i>Person_3.png</i>	29.5	28.92 %
<i>Person_4.png</i>	13.5	13.24 %
<i>Person_5.png</i>	18.9	18.53 %
<i>Person_6.png</i>	69.9	68.53 %
<i>Person_7.png</i>	20.4	20 %



Joonis 12: pilt, mida peeti uurimuse hindamiseks liiga pimedaks

Litsents

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, **Ilmar Möls**,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose
Fiery semantilise segmentatsiooni mudeli efektiivsuse hindamine,
mille juhendaja on **Markus Kängsepp**,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Ilmar Möls

15.05.2024