

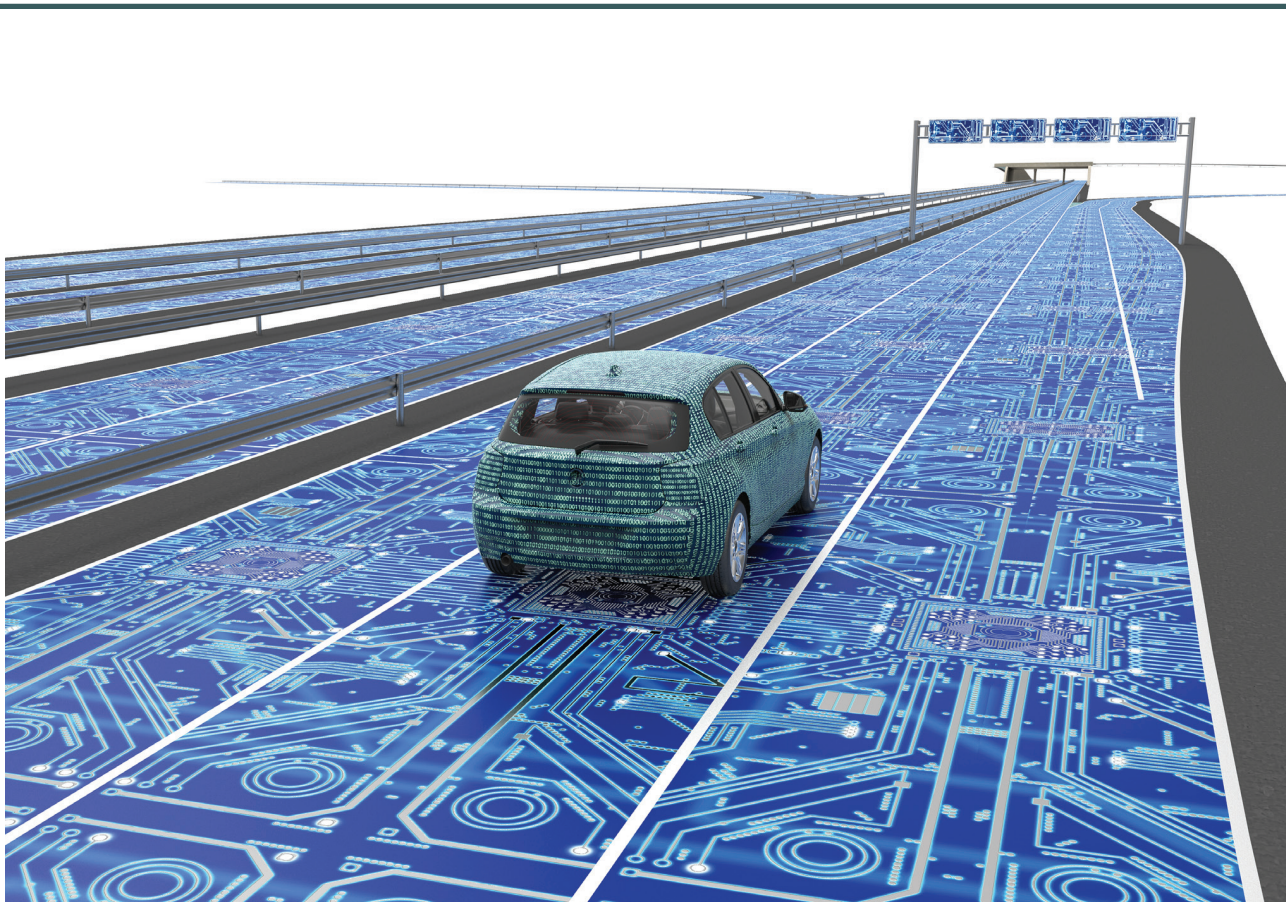


RĪGAS TEHNISKĀ
UNIVERSITĀTE

Roberts Kadiķis

EFEKTĪVAS VIDEO APSTRĀDES METODES KUSTĪGU OBJEKTU ATKLĀŠANAI UN RAKSTUROŠANAI

Promocijas darba kopsavilkums



RĪGAS TEHNISKĀ UNIVERSITĀTE

Elektronikas un telekomunikāciju fakultāte

Radioelektronikas institūts

Roberts Kadiķis

Doktora studiju programmas “Elektronika” doktorants

**EFEKTĪVAS VIDEO APSTRĀDES METODES KUSTĪGU
OBJEKTU ATKLĀŠANAI UN RAKSTUROŠANAI**

Promocijas darba kopsavilkums

Zinātniskais vadītājs

Dr. sc. comp. MODRIS GREITĀNS,
Elektronikas un datorzinātņu institūta
zinātniskais direktors, vadošais pētnieks

RTU Izdevniecība

Rīga 2018

Kadiķis R. Efektīvas video apstrādes metodes
kustīgu objektu atklāšanai un raksturošanai.
Promocijas darba kopsavilkums.
Rīga: RTU Izdevniecība, 2018. 32 lpp.

Iespiests saskaņā ar RTU promocijas padomes
“RTU P-08” 2018. gada 8. janvāra lēmumu,
protokols Nr. 42.

ISBN 978-9934-22-057-9 (print)

ISBN 978-9934-22-058-6 (epub)

PROMOCIJAS DARBS IZVIRZĪTS INŽENIERZINĀTŅU DOKTORA GRĀDA IEGŪŠANAI RĪGAS TEHNISKAJĀ UNIVERSITĀTĒ

Promocijas darbs inženierzinātņu doktora grāda iegūšanai tiek publiski aizstāvēts 2018. gada 19. aprīlī Rīgas Tehniskās universitātes Elektronikas un telekomunikāciju fakultātē, Āzenes ielā 12, 212. auditorijā.

OFICIĀLIE RECENZENTI

Profesors *Dr. habil. phys.* Andris Ozols
Rīgas Tehniskā universitāte, Latvija

Profesors *Dr. sc. ing.* Aleksandrs Grakovskis
Transporta un sakaru institūts, Latvija

Profesors *Dr. sc. ing.* Pēteris Grabusts
Rēzeknes Tehnoloģiju akadēmija, Latvija

APSTIPRINĀJUMS

Apstiprinu, ka esmu izstrādājis šo promocijas darbu, kas iesniegts izskatīšanai Rīgas Tehniskajā universitātē inženierzinātņu doktora grāda iegūšanai. Promocijas darbs zinātniskā grāda iegūšanai nav iesniegts nevienā citā universitātē.

Roberts Kadiķis (paraksts)

Datums:

Promocijas darbs ir uzrakstīts angļu valodā, tajā ir ievads, piecas nodaļas, literatūras saraksts, 37 attēli, četras tabulas, seši pielikumi, kopā 132 lappuses. Literatūras sarakstā ir 111 nosaukumu.

SATURS

SAĪSINĀJUMI	5
VISPĀRĒJAIS DARBA APRAKSTS	6
Tēmas aktualitāte	6
Darba mērķis un uzdevumi	6
Pētījumu metodika	7
Zinātniskā novitāte un galvenie rezultāti	7
Aizstāvamās tēzes.	8
Darba praktiskais lietojums un aprobācija	8
Darba struktūra un apjoms	9
1. OBJEKTU ATKLĀŠANAS METODES.	10
1.1 Transporta līdzekļu atklāšanas sensori	10
1.2 Atklāšana pēc pikseļu parametriem	10
1.3 Atklāšana pēc objektu parametriem	11
1.4 Dziļās apmācības metodes.	12
2. EFEKTĪVAS OBJEKTU ATKLĀŠANAS UN RAKSTUROŠANAS METODES	14
2.1 Eksistējošas efektīvas metodes	14
2.2 Objektu atklāšanas metode <i>IoVDL</i>	14
2.3 <i>IoVDL</i> paplašinājums objektu raksturošanai	17
2.4 Uz <i>RNN</i> balstīta virtuāla atklāšanas līnija	18
3. MARĶĒTU DATU IEGŪŠANAS PAŅĒMIENI	19
3.1 <i>RNN-VDL</i> treniņa datu formāts	19
3.2 Piedāvātās pieejas datu marķēšanai	20
3.3 Marķētu datu ģenerēšana ar 3D spēļu dzinēju	21
4. LIETOJUMI UN EKSPERIMENTĀLIE PĒTĪJUMI	22
4.1 Izstrādāto metožu efektivitāte	22
4.2 Objektu skaitīšanas precizitāte	23
4.3 Automašīnu atklāšana un klasificēšana ar paplašināto <i>IoVDL</i>	25
5. SECINĀJUMI	26
LITERATŪRA	28

SAĪSINĀJUMI

2D – 2 dimensiju

3D – 3 dimensiju

CNN – konvolūcijas neironu tīkls (*Convolutional Neural Network*)

CPU – centrālais procesors (*Central Processing Unit*)

EDI – Elektronikas un datorzinātņu institūts

FN – kļūdaina neatbilde (*False Negative*)

FP – kļūdaina atbilde (*False Positive*)

GUI – grafiskā lietotāja saskarne (*Graphical User Interface*)

IoVDL – metode “intervāli uz virtuālas atklāšanas līnijas” (*Intervals on Virtual Detection Line*)

ITS – Intelīgentas transporta sistēmas (*Intelligent Transportation Systems*)

LSTM – rekurentā neironu tīkla paveids “*Long Short-Term Memory*”

RAM – brīvpieejas atmiņa (*Random-Access Memory*)

R-CNN – metode “reģioni ar *CNN* iezīmēm”

RNN – rekurentais neironu tīkls (*Recurrent Neural Network*)

RNN-VDL – metode “uz rekurentā neironu tīkla balstīta virtuāla atklāšanas līnija”

ROI – interesējošais apgabals attēlā (*Region of Interest*)

SVM – atbalsta vektoru mašīna (*Support Vector Machine*)

TP – pareiza atbilde (*True Positive*)

YOLO – metode “*You Only Look Once*”

Pētījumu metodika

Pirmie darba uzdevumi tiek risināti ar analītiskiem pētījumiem. Analītiski tiek veikts jomas literatūras pārskats un analīze, tiek piedāvāti un izveidoti jauni kustīgu objektu atklāšanas algoritmi, kā arī tiek veidotas jaunas marķētu datu iegūšanas metodes. Piedāvātās metodes tiek īstenotas *C++* un *Python* valodās. Īstenotās programmas tiek novērtētas un salīdzinātas eksperimentālos testos ar dažādiem video. Testos tiek mērītas metožu skaitļošanas efektivitāte un atklāšanas precizitāte.

Zinātniskā novitāte un galvenie rezultāti

Darba rezultātā ir izstrādātas jaunas un efektīvas metodes kustīgu objektu atklāšanai un raksturošanai video. Lai piedāvātās metodes varētu testēt un apmācīt, ir izveidotas arī jaunas datu marķēšanas metodes.

IoVDL – intervāli uz virtuālas atklāšanas līnijas – ir skaitļošanas jaudas patēriņa ziņā efektīva metode kustīgu objektu atklāšanai. Līdzīgi jau eksistējošām efektīvām metodēm, *IoVDL* apstrādā tikai nelielu daļu no katra kadra pikseļiem, bet tā ir izmantojama plašākam uzdevumu lokam nekā alternatīvas. Automašīnu atklāšanas uzdevuma kontekstā *IoVDL* nepieprasa, lai mašīnas brauktu pa nemainīgām trajektorijām. *IoVDL* ir īstenota uz zemu izmaksu datora *Raspberry Pi Zero*. Tās precizitāte ir eksperimentāli testēta uz video, kas uzņemti dažādos laikapstākļos, iegūstot līdzīgu rezultātu esošajām efektīvajām metodēm.

Paplašinātā *IoVDL* demonstrē izvirzītās objektu atklāšanas metodes papildu priekšrocību. Izvietojojot kadrā vairākus *IoVDL* detektorus, paplašinātā *IoVDL* metode spēj izsekot objektu kustībai cauri šiem detektoriem. Metode arī raksturo atklātos objektus, nomērot to ātrumu un izmērus un nosakot objektu klasi. Salīdzinot ar eksistējošām video apstrādes metodēm objektu izsekošanai un raksturošanai, paplašinātā *IoVDL* ir efektīva no nepieciešamo skaitļošanas resursu viedokļa. Jaunā metode ir īstenota un ir testēta tās spēja klasificēt automašīnas.

RNN-VDL – uz rekurentā neironu tīkla balstīta virtuāla atklāšanas līnija – apvieno efektīvu virtuālu detektoru pieeju ar rekurentiem neironu tīkliem. Līdzīgi kā *IoVDL* gadījumā, jaunā metode apstrādā tikai vienu pikseļu līniju kadrā un atklāj objektus, kas šo līniju šķērso. Metode ir īstenota kodā un salīdzināta ar *IoVDL* automašīnu skaitīšanas uzdevumā. *RNN-VDL* balstās uz mašīnmācīšanos, kas nozīmē, ka viena un tā pati metodes implementācija var tikt apmācīta atklāt dažādus objektus. Šī metodes īpašība tiek pierādīta ar cilvēku skaitīšanas testu.

Manuāla marķēšanas metode ir uz grafiskās lietotāja saskarnes (*GUI*) balstīta metode video kadru marķēšanai. Lietotājs katram kadram nosaka atbilstošo marķējumu, tad piedāvātā metode to pārveido *RNN-VDL* saprotamā formā. Kadri un tiem atbilstošais marķējums tiek saglabāts, lai to varētu izmantot objektu atklāšanas modeļu trenēšanai.

Pusautomātiska marķēšanas metode apvieno sākotnēju automātisku video marķēšanu ar specifisku *GUI*, kas ļauj cilvēkam ātri noteikt un izlabot automātiskajā metodes daļā pieļautās kļūdas. Šī ir būtiski ātrāka marķēšanas metode par manuālo marķēšanu, bet tā ir lietojama tikai uz detektējošas līnijas balstītu modeļu trenēšanai. Promocijas darbā izveidotā pusautomātiskā metode tiek lietota, lai marķētu treniņa un testa video darbā veiktajiem testiem. Sagatavotā marķētu datu kopa pieejama [9].

Aizstāvamās tēzes

1. Izstrādātā *IoVDL* metode atklāj kustīgus objektus video ar līdzīgu precizitāti un skaitļošanas efektivitāti kā alternatīvas uz virtuāliem atklāšanas reģioniem balstītas metodes, bet tā ir mazāk jutīga pret detektēšanas reģiona šķērsojošo objektu skaitu un to kustības trajektoriju izmaiņām.
2. Vairāku *IoVDL* detektoru apvienošana (paplašinātā *IoVDL*) izseko objektu kustību un mēra to ātruma un izmēra parametrus, patērējot mazāk skaitļošanas resursu nekā tradicionālās objektu izsekošanas metodes, kas apstrādā visus kadra pikselus.
3. Virtuālas atklāšanas līnijas pieejas apvienošana ar rekurentu neironu tīklu rezultējas adaptīvā atklāšanas metodē *RNN-VDL*, ko var pārtrenēt dažādu objektu atklāšanai, nemainot atklāšanas modeļa arhitektūru un manuāli neveicot jaunu pazīmju piemeklēšanu.
4. Izstrādātā pusautomātiskā marķēšanas metode paātrina uz rekurentā neironu tīkla balstītās metodes *RNN-VDL* apmācībai nepieciešamo, vairākas stundas garo video marķēšanu vismaz desmitkārtīgi.

Darba praktiskais lietojums un aprobācija

Izstrādātās metodes *IoVDL*, paplašinātā *IoVDL* un *RNN-VDL* piedāvā jaunus paņēmienus, kā veidot efektīvus algoritmus kustīgu objektu atklāšanai video. Tiek testēta *IoVDL* spēja skaitīt automašīnas. Testi rāda, ka metode ir piemērota lietošanai *ITS*. Savukārt *RNN-VDL* metodes un piedāvāto datu marķēšanas metožu kombinācija padara uz mašīnmācīšanos balstīto pieeju izmantojamu plašākā praktisku uzdevumu lokā.

Promocijas darbs ir izstrādāts Elektronikas un datorzinātņu institūtā (EDI) saistībā ar šādu projektu izpildi:

- Eiropas Reģionālās attīstības fonda projekts “Multifunkcionāla inteliģenta transporta sistēmas punkta tehnoloģija” (MITS) Nr. 2010/0250/2DP/2.1.1.1.0/10/APIA/VIAA/086;
- Eiropas Sociālā fonda projekts “Inovātas biomedicīnisko attēlu iegūšanas un apstrādes tehnoloģijas” (InBIT) Nr. 2013/0009/1DP/1.1.1.2.0/13/APIA/VIAA/014;

- valsts pētījumu programmas “Kiberfizikālās sistēmas, ontoloģijas un biofotonika drošai un viedai pilsētai un sabiedrībai” (*SOPHIS*) projekts Nr. 4 “Tehnoloģijas drošai un uzticamaai gudrajai pilsētai”;
- Eiropas Reģionālās attīstības fonda projekts “Dziļo neironu tīklu metode auto transporta numura zīmju lokalizācijas un klasifikācijas precizitātes uzlabošanai” (DziNTA) Nr. 1.2.1.1/16/A/007;
- Eiropas Reģionālās attīstības fonda projekta “Latvijas elektrisko un optisko iekārtu ražošanas nozares kompetences centrs” Nr. 1.2.1.1/16/A/002 pētījums Nr. 11 “Pētījums par datorredzes paņēmieni attīstību industrijas procesu norises automatizācijai.”

Darba rezultāti ir aizsargāti ar Eiropas patentu [10]. Tie ir aprakstīti vairākās publikācijās [11]–[14], kā arī saistīti ar attēlu apstrādei veltītām nodaļām publikācijās [15]–[17]. Zinātniskos rezultātus autors prezentējis starptautiskās konferencēs:

- International Conference on Signal and Image Processing 2012 (ICSIP 2012), 13.–15. decembris, Koimbatore, Indija;
- The 6th International Conference on Machine Vision (ICMV 2013), 16.–17. novembris, Londona, Lielbritānija;
- Advances in Information Electronic and Electrical Engineering, 5.–7. jūlijs, 2013 Rīga, Latvija;
- Microwave and Radio Electronics Week 2015 (MAREW 2015), 21.–23. aprīlis. Pardubice, Čehija;
- 11th International Young Scientist conference. Developments in Optics and Communications, 8.–10. aprīlis, 2015, Rīga, Latvija;
- New Challenges of Economic and Business Development – 2017: Digital Economy, 18.–19. maijs, Rīga, Latvija;
- The 10th International Conference on Machine Vision (ICMV 2017), 13.–15. novembris, Vīne, Austrija.

Darba struktūra un apjoms

Promocijas darba apjoms ir 132 lpp. Tas sastāv no piecām nodaļām. Pirmā nodaļa ir objektu atklāšanas metožu pārskats un analīze. Metodes, kas neizmanto video kameru, ir salīdzinātas ar datorredzes pieeju, kuras dažādās metodes tālāk aprakstītas sīkāk. 2. nodaļa apskata esošās efektīvās video apstrādes pieejas, tad piedāvā jaunu metodi *IoVDL*. Piedāvātā metode tiek paplašināta, piešķirot tai spēju raksturot atklātos objektus. Tad tiek piedāvāta vēl viena jauna kustīgu objektu atklāšanas metode – *RNN-VDL*. Šī mašīnmācīšanās pieeja pieprasa lielu daudzumu marķētu datu, tādēļ 3. nodaļa ir veltīta datu marķēšanas pieejām. Darbā izstrādātās metodes tiek testētas automašīnu un cilvēku atklāšanas uzdevumos 4. nodaļā. Secinājumi par izstrādātajām metodēm un testu rezultātiem ir doti 5. nodaļā.

1. OBJEKTU ATKLĀŠANAS METODES

1.1. Transporta līdzekļu atklāšanas sensori

Video kamera nav vienīgais sensors, ko var lietot objektu atklāšanai. Automašīnu atklāšanai uz ceļiem ir tikuši lietoti dažādi sensori, ieskaitot invazīvus, piemēram, pneimatiskās caurules, induktīvās cilpas, magnetometri, kā arī neinvazīvus sensorus, piemēram, mikroviļņu radari, infrasarkanā lāzera radari, ultraskaņas sensori, pasīvie infrasarkanie un akustiskie sensori.

2010.–2011. gada pētījumos [18], [19] tika eksperimentāli salīdzināti dažāda tipa sensori. Video kamera izrādījās nedz precīzākais, nedz robustākais sensors. Tomēr ir vairāki apsvērumi, kādēļ šis promocijas darbs pievērsās tieši video apstrādē balstītai atklāšanai.

- Kopš minēto testu veikšanas datorredzes joma ir piedzīvojusi visstraujāko attīstību. Pirmajās *ImageNet* [20] sacensībās 2010. gadā labākās attēlu klasificēšanas metodes kļūdījās 28,2 % attēlu. 2017. gada uzvarētāji sasniedza 2,25 % kļūdu. Cilvēks *ImageNet* testā kļūdās aptuveni 5 % gadījumu [21].
- Viena kamera var nosegt plašu ceļa apgabalu un daudzas ceļa joslas, tādējādi aizstājot vairākus alternatīvus sensorus.
- Kamera ir universāls sensors. To var izmantot arī operatīvo transportlīdzekļu, gājēju, dzīvnieku un citu objektu atklāšanai.
- Tā pati kamera, kas iegūst datus atklāšanas uzdevumam, var iegūt datus arī cita veida uzdevumu veikšanai, piemēram, objektu raksturošanai, klasificēšanai un izsekošanai.

1.2. Atklāšana pēc pikseļu parametriem

2D ciparu attēls sastāv no $R \times C$ pikseļiem, kur katru pikseli apraksta kāda vērtība. Pelēktoņu attēla gadījumā šī vērtība ir pikseļa intensitāte I . Vienkāršākajos gadījumos objektus attēlā var atšķirt no fona, analizējot katra atsevišķā pikseļa vērtības. Šādas metodes nav tik robustas kā sarežģītāki paņēmieni, tomēr tās ir labs atspēriena punkts skaitļošanas resursu ziņā efektīvu metožu veidošanai.

Sliekšņošana. Ja objekti pelēktoņu attēlā ir gaišāki par fonu, tos var atklāt, atrodot attēla gaišos apgabalus. Sliekšņošanas operācija salīdzina katru pikseli ar sliekšņa vērtību T . Sliekšņošanas rezultāts ir binārs attēls, kurā pikseļi ar $I > T$ ir balti, bet pikseļi ar $I \leq T$ ir melni. Pieejai svarīgais parametrs T var tikt atrasts automātiski, piemēram, lietojot *Otsu* metodi [22].

Krāsu var izmantot kā vēl vienu objektu nošķirošu parametru. Krāsainā *RGB* attēlā katra pikseļa krāsa tiek aprakstīta ar trīs vērtībām. Attālums starp dažādu pikseļu *RGB* vērtībām ļauj izšķirt dažādas krāsas objektus un fonu. Šī ideja ir vispārināma arī uz citiem krāsu modeļiem. To var lietot arī multispektrālo un hiperspektrālo attēlu analizē, kur katram pikselim ir vairāk nekā trīs aprakstošas vērtības.

Fona atņemšana video apstrādē nozīmē šī brīža kadra salīdzināšanu ar atskaites kadru, kurā nav objektu un ir redzams tikai fons. Lai atšķirtu objektu no fona, var izmantot dažādus parametrus, piemēram, pikseļu intensitāti, krāsu [23]–[25] vai robežas un stūrus, kas ir robustāki parametri mainīgos apgaismojumos un laikstākļos [26].

Daudzām atklāšanas sistēmām ir jādarbojas mainīgos āra apstākļos, tāpēc bieži tiek lietots adaptīvs fona attēls. Pastāv nerekursīvas metodes adaptīva fona iegūšanai, kur noteikts skaits kadru tiek saglabāts buferī. Par fona pikseļa vērtību tiek pieņemta kāda statistiska vērtība (piemēram, vidējā vērtība vai mediāna) no šī paša pikseļa vērtībām buferī saglabātajos kados. Rekursīvajās metodēs fons tiek atjaunots ar tādām metodēm kā Kalmana filtrs [26], Gausa funkciju mikstūra [27], svarota fona attēla un šī brīža kadra summa [25] un citām.

1.3. Atklāšana pēc objektu parametriem

Stabilākus rezultātus parasti iegūst, atrodot un aprakstot ar objektu saistītas iezīmes.

Veidņu saskaņošana. Ja ir pieejams meklējamā objekta attēls (veidne) I_t , tad šo veidni var salīdzināt ar visiem ieejas attēla I apgabaliem, meklējot vislīdzīgāko apgabalu. Salīdzināšana bieži tiek veikta pēc slīdošā loga principa, izmantojot dažādus līdzības kritērijus.

Iezīmju punktu detektori. Objektus attēlā var atrast, meklējot objektam raksturīgus punktus vai apgabalus. Šādi iezīmju punkti var būt stūri, T-veida savienojumi vai no pārējās bildes atšķirīgi apgabali attēlā. Bieži lietotu raksturīgo punktu detektēšanas metožu piemēri:

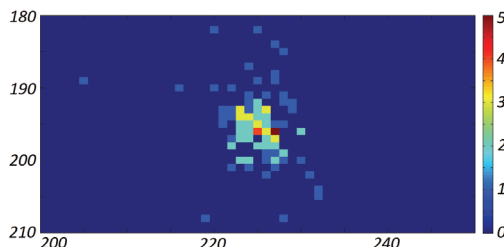
- *Harris Corner Detector* [28];
- *Hessian Matrix Detector* [29];
- *DoG – Difference of Gaussian detector* [30];
- *MSER – Maximally Stable Extremal Regions* [31];
- *FAST – Features from Accelerated Segment Test* [32], [33].

Apgabalu deskriptori. Kad ir atrasti iezīmju punkti, attēla apgabali ap šiem punktiem tiek aprakstīti ar pazīmju vektoru. Bieži lietotu deskriptoru piemēri:

- *SIFT – Scale Invariant Feature Transform* [34];
- *SURF – Speeded Up Robust Features* [35];
- *LBP – Locally Binary Patterns* [36];
- *ORB – Oriented FAST and Rotated BRIEF* [37].

Iezīmju punktu saskaņošana. Lai noteiktu, vai objekts atrodas attēlā, objekta aprakstītajiem iezīmju punktiem jāatrod atbilstoši attēla punkti. Parasti šim nolūkam tiek aprēķināti attālumi starp objekta un attēla punktu aprakstošajiem vektoriem. Šajā promocijas darbā tiek piedāvāta jauna iezīmju punktu saskaņošanas metode, kas ļauj izmantot un kombinēt dažāda veida iezīmju punktus.

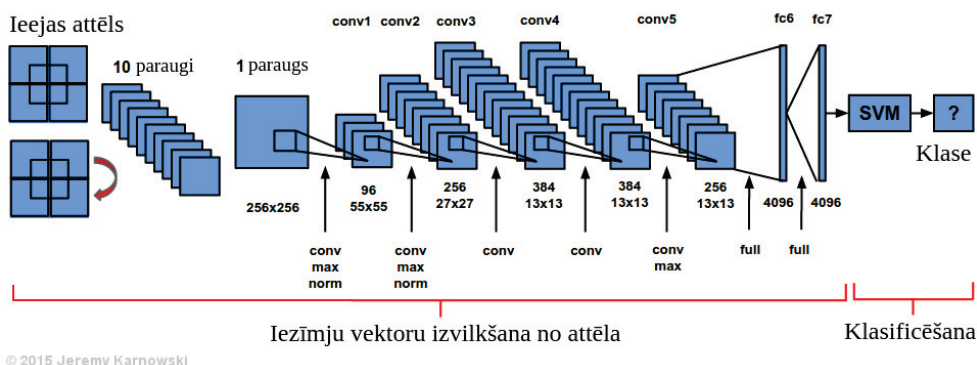
Algoritms pieņem kādu sākotnējo objekta atrašanās vietu attēlā, tad izmanto iezīmju punktus, lai atrastu vistīcāmāko pārvietojumu starp pieņemto un patieso objekta atrašanās vietu. Visi iespējamie pārvietojumi starp objekta un attēla iezīmju punktiem tiek nomērīti un izmantoti jauna attēla izveidošanai (1.1. att.). Lielākās vērtības pikseļa koordinātes ir vistīcāmākais pārvietojums starp pieņemto un patieso objekta atrašanās vietu attēlā. Piedāvāto metodi iespējams padarīt arī spējīgu atklāt objektu, pat ja tas ir rotēts un mēroģots.



1.1. att. Iespējamais pārvietojums starp pieņemto un īsto objekta atrašanās vietu attēlā.

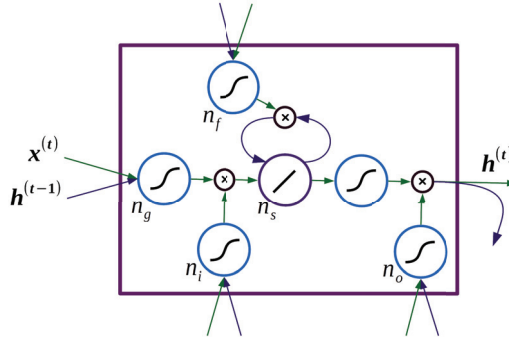
1.4. Dziļās apmācības metodes

CNN. Šī brīža labākās objektu atklāšanas metodes, piemēram *R-CNN* (reģioni ar *CNN* iezīmēm) [38], *Faster R-CNN* [39] un *YOLO (You Only Look Once)* [40], balstās uz mākslīgajiem neironu tīkliem. Redzes uzdevumos visbiežāk tiek lietoti konvolūcijas neironu tīkli (*CNN*), kas spēj iemācīties no attēliem iegūt dažādiem uzdevumiem svarīgas iezīmes. Populāra *CNN* arhitektūra *AlexNet* ir parādīta 1.2. att. Tā sastāv no pieciem konvolūcijas slāņiem, kam seko divi pilnībā savienotu neironu slāņi.



1.2. att. *AlexNet* arhitektūra [41].

RNN. Tradicionālie neironu tīkli, kā arī *CNN* nav speciāli veidoti ar domu atpazīt laikā mainīgus notikumus. Šo modeļu izeja ir atkarīga tikai no šī brīža ieejas signāla. Savukārt rekurentajos neironu tīklos (*RNN*) ir neironi ar rekurentām saitēm, kas padara *RNN* par labākajām šī brīža metodēm vairākos ar secību atpazīšanu saistītos uzdevumos [42].



1.3. att. *LSTM* tīkla atmiņas šūna.

LSTM – *Long Short-Term Memory* [43] – ir veiksmīgs *RNN* paveids. *LSTM* slēptais slānis ir atmiņas šūna, kas parādīta 1.3. att. Šūna sastāv no iekšējā stāvokļa mezgla n_s un vairākiem papildu mezgliem n_g, n_i, n_f, n_o . Visi papildu mezgli saņem šī brīža ieejas signālu $\mathbf{x}^{(t)}$ un *LSTM* šūnas izejas signālu no iepriekšējā laika soļa $\mathbf{h}^{(t-1)}$.

LSTM šūnas veiktais aprēķins katrā laika solī parādīts vienādojumos (1.1)–(1.6):

$$\mathbf{n}_g^{(t)} = \alpha(\mathbf{W}_{gx}\mathbf{x}^{(t)} + \mathbf{W}_{gh}\mathbf{h}^{(t-1)} + \mathbf{b}_g) \quad (1.1)$$

$$\mathbf{n}_i^{(t)} = \alpha(\mathbf{W}_{ix}\mathbf{x}^{(t)} + \mathbf{W}_{ih}\mathbf{h}^{(t-1)} + \mathbf{b}_i) \quad (1.2)$$

$$\mathbf{n}_f^{(t)} = \alpha(\mathbf{W}_{fx}\mathbf{x}^{(t)} + \mathbf{W}_{fh}\mathbf{h}^{(t-1)} + \mathbf{b}_f) \quad (1.3)$$

$$\mathbf{n}_o^{(t)} = \alpha(\mathbf{W}_{ox}\mathbf{x}^{(t)} + \mathbf{W}_{oh}\mathbf{h}^{(t-1)} + \mathbf{b}_o) \quad (1.4)$$

$$\mathbf{n}_s^{(t)} = \mathbf{n}_g^{(t)} \odot \mathbf{n}_i^{(t)} + \mathbf{n}_s^{(t-1)} \odot \mathbf{n}_f^{(t)} \quad (1.5)$$

$$\mathbf{h}^{(t)} = \alpha(\mathbf{n}_s^{(t)}) \odot \mathbf{n}_o^{(t)}. \quad (1.6)$$

kur \mathbf{W} ir svaru matricas un \mathbf{b} ir līdzkomponentu vektori starp dažādajiem šūnas mezgliem; $\alpha(x)$ ir kāda nelineāra aktivācijas funkcija; \odot ir Adamāra reizinājums.

RNN izmantošana video kadru plūsmas analīzei pieprasa būtiskus skaitļošanas resursus. Nākamajā nodaļā tiek piedāvāta un izstrādāta metode, kas paver iespēju efektīvai *RNN* lietošanai video apstrādē.

2. EFEKTĪVAS OBJEKTU ATKLĀŠANAS UN RAKSTUROŠANAS METODES

2.1. Eksistējošas efektīvas metodes

Viena no pieejām efektīvu algoritmu veidošanā ir apstrādāt tikai daļu kadra. Ar statistisku kameru iegūtā video var tikt definēts interešu reģions (*ROI*). Apstrādājot tikai šī reģiona pikselus, iespējams atklāt objektus, kas vai nu atrodas šajā reģionā, vai arī ar laiku tajā iekļūst. Neskatoties uz šo ierobežojumu, *ROI* balstītās metodes ir lietojamas vairākos praktiskos uzdevumos, piemēram, transporta monitorēšana uz ceļiem [44]–[46], ēkās ienākošu un izejošu cilvēku atklāšana [47], [48], ražošanas produktu monitorēšana uz konveijera lentes [49], [50].

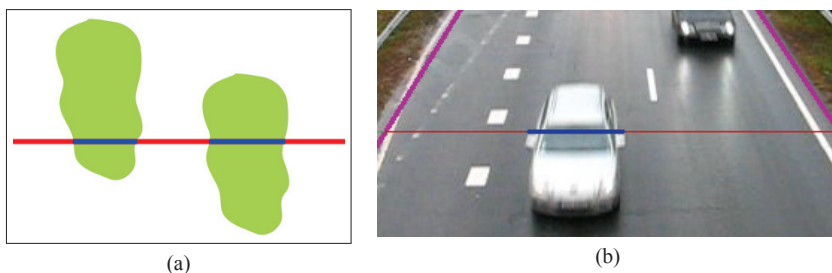
Rakstos [51]–[53] *ROI* tiek veidots no vienas vai vairākām detektēšanas līnijām kadrā. Līnijas ir novietotas perpendikulāri sagaidāmajai objektu kustībai. Literatūrā šādas līnijas sauc arī par virtuālām līnijām, detektēšanas līnijām un virtuāliem cilpu detektoriem.

Ceļu monitorēšanas gadījumā esošajām uz *ROI* balstītām metodēm (piemēram, [53]–[55]) ir precīzi jādefinē atklāšanas reģioni katrai ceļa joslai. Šādām metodēm var rasties detektēšanas kļūdas, kad garas automašīnu ēnas vai lielas automašīnas aizsedz blakus joslu detektorus, kas tādēļ kļūdaini atklāj neesošus objektus. Šādas metodes nav piemērotas arī gadījumos, kad ceļa joslu skaits vai izvietojums uz ceļa var mainīties laikā.

2.2. Objektu atklāšanas metode *IoVDL*

Šajā apakšnodaļā piedāvātā metode lietojuma ziņā ir elastīgāka par alternatīvām, jo jauna intervālu pieeja ļauj atklāt dažāda platuma objektus jebkurā detektēšanas līnijas apgabalā.

Intervāli uz virtuālas detektēšanas līnijas (*IoVDL*). Piedāvātā metode atklāj objektus, kas šķērso kadrā definētu līniju (2.1. att.) Šī līnija (pikseļu rinda) parasti tiek novietota perpendikulāri objektu kustībai. Visa turpmākā apstrāde notiek tikai uz līnijas pikseliem, padarot metodi efektīvu prasībās pret skaitļošanas jaudām.



2.1. att. Atklāšanas līnija, objekti un tiem atbilstošie intervāli kadrā.

Atklāšanas procesā uz līnijas tiek veidoti intervāli, kas atzīmē līnijas apgabalu, kuru ir nosedzis to šķērsojošs objekts. Ņemot vērā informāciju no turpmākajiem kadriem, katrs intervāls tiek atjaunināts, lai tas atbilstu mainīgajam objekta un līnijas pārklāšanās segmentam. Kad objekts pamet atklāšanas līniju, attiecīgais intervāls tiek slēgts, un tiek iegūti objekta parametri. Galvenie *IoVDL* soļi parādīti pseidokodā (1. algoritms).

1. algoritms. Objektu atklāšana uz līnijas.

```

1: procedure ATKLĀT-OBJEKTUS(līnija no krāsu kadra  $\mathbf{l}_{\text{rgb}}$ , kadra numurs  $t$ )
2:   līnija  $\mathbf{l}_g^{(t)} \leftarrow$  konvertēt  $\mathbf{l}_{\text{rgb}}$  uz pelēktoņu
3:   līnija  $\mathbf{l}_d \leftarrow$  mediānas_filtrs( $|\mathbf{l}_g^{(t)} - \mathbf{l}_g^{(t-1)}|$ )
4:   līnija  $\mathbf{l}_t \leftarrow$  sliekšņot  $\mathbf{l}_d$  ar sliekšni  $T$ 
5:   pagaidu buferis intervāliem  $\mathbb{B}_{i\_temp} \leftarrow$  ATRAST-INTERVĀLUS( $\mathbf{l}_t$ )
6:   pievienot intervālus no  $\mathbb{B}_{i\_temp}$  uz  $\mathbb{B}_i$ 
7:   for each (intervāls  $i$  un  $j$  buferī  $\mathbb{B}_i$ ) do
8:     if  $i$  un  $j$  pārklājas then
9:       apvienot  $i$  un  $j$ 
10:    end if
11:  end for
12:  līnija  $\mathbf{l}_{i\_t} \leftarrow$  INTENSITĀTES-FONS( $\mathbb{B}_i, \mathbf{l}_g^{(t)}$ )
13:  līnija  $\mathbf{l}_{e\_t} \leftarrow$  ROBEŽU-FONS( $\mathbb{B}_i, \mathbf{l}_g^{(t)}$ )
14:  līnija  $\mathbf{l}_{\text{rgb}_t} \leftarrow$  KARĀSU-FONS( $\mathbb{B}_i, \mathbf{l}_{\text{rgb}}$ )
15:  for each (intervāls  $i$  buferī  $\mathbb{B}_i$ ) do
16:    int  $Z_i \leftarrow$  balto pikseļu skaits uz  $\mathbf{l}_{i\_t}$ , kurus nosedz  $i$ 
17:    int  $Z_e \leftarrow$  balto pikseļu skaits uz  $\mathbf{l}_{e\_t}$  kurus nosedz  $i$ 
18:    int  $Z_{\text{rgb}} \leftarrow$  balto pikseļu skaits uz  $\mathbf{l}_{\text{rgb}_t}$  kurus nosedz  $i$ 
19:    if  $Z_i + Z_e + Z_{\text{rgb}} <$  konstante and  $i$  nav ticis atjaunots aizture kadru skaitu then
20:      if  $\min\text{Plat} <$   $\text{platums}(i) <$   $\max\text{Plat}$  and  $\text{ilgums}(i) >$   $\min\text{Ilgums}$  then
21:        paziņot par objekta atklāšanu
22:      end if
23:      dzēst intervālu  $i$  no  $\mathbb{B}_i$ 
24:    end if
25:  end for
26: end procedure

```

1. algoritma metode ATRAST-INTERVĀLUS(\mathbf{l}_t) atrod baltus savienotus apgabalus uz sliekšņotas līnijas un izveido šiem apgabaliem atbilstošus intervālus. Intervāli tiek saglabāti, ja tie atbilst manuāli izvēlētiem izmēra parametriem. Ceļa monitorēšanas gadījumā šie parametri nosaka, vai līdz ar automašīnām tiks atklāti arī motocikli un riteņi.

Lai intervāli tiktu uzturēti arī tad, kad uz līnijas vairs nenotiek kustība, tiek izmantota fona atņemšanas pieeja. Tiek izveidots fons vairākiem parametriem: intensitātei, robežām un krāsai. 1. algoritma 12–14 soļos attiecīgo funkciju izejas ir sliekšņotas līnijas \mathbf{l}_{i_t} , \mathbf{l}_{e_t} , $\mathbf{l}_{\text{rgb}_t}$, kuru baltie apgabali atbilst objektiem.

Kad intervāls tiek slēgts, tiek analizēti tā platuma un pastāvēšanas ilguma parametri. Tādējādi tiek ignorēti pārāk īsi un pārāk īslaicīgi intervāli. Šos parametrus var izmantot, lai atšķirtu dažādas objektu klases.

Adaptīva fona atņemšana. 2. algoritms parāda *IoVDL* fona atņemšanas procesu uz funkcijas ROBEŽU-FONS($\mathbb{B}_i, \mathbf{I}_g$) piemēra.

2. algoritms. Fona atņemšana izmantojot robežu līnijas.

```

1: function ROBEŽU-FONS(intervālu buferis  $\mathbb{B}_i$ , pelēktoņu līnija  $\mathbf{I}_g$ )
2:   līnija  $\mathbf{I}_g \leftarrow$  izplūdināt līniju  $\mathbf{I}_g$  ar Gausa filtru
3:    $\mathbf{I}_e \leftarrow$  atrast robežas uz  $\mathbf{I}_g$ 
4:   if  $\mathbb{B}_i$  ir tukšs then
5:     atjaunināt visu fona līniju  $\mathbf{I}_{e\_bg}$  pēc formulas (2.1)
6:   else
7:     atjaunināt līnijas  $\mathbf{I}_{e\_bg}$  ārpus intervāliem esošos apgabalus pēc formulas (2.1)
8:     pievienot  $\mathbf{I}_e$  buferim  $\mathbb{B}_e$ 
9:   end if
10:   $\mathbf{I}_e \leftarrow |\mathbf{I}_e - \mathbf{I}_{e\_bg}|$ 
11:  int  $T_e \leftarrow$  ATRAST-SLIEKSNI( $\mathbb{B}_e$ )
12:  līnija  $\mathbf{I}_{e\_t} \leftarrow$  sliekšņot  $\mathbf{I}_e$  ar vērtību  $T_e$ 
13:  return (sliekšņotu robežu līniju  $\mathbf{I}_{e\_t}$ )
14: end function

```

Robežas uz intensitātes līnijas tiek atrastas, aprēķinot starpības moduli starp līnijas blakus pikseliem (3. solis). Lai rekursīvi atjauninātu robežu fona līniju \mathbf{I}_{e_bg} , algoritms izmanto šādu formulu:

$$I_{bg}^{(t)} = E \cdot I_{bg}^{(t-1)} + (1 - E) \cdot I^{(t)}, \quad (2.1)$$

kur $I_{bg}^{(t)}$ – fona pikseļa intensitāte šī brīža kadrā; $I_{bg}^{(t-1)}$ – fona pikseļa intensitāte iepriekšējā kadrā; $I^{(t)}$ – pikseļa intensitāte šī brīža kadrā; E – atjaunināšanas ātrums.

Funkcija ATRAST-SLIEKSNI(\mathbb{B}_e) izmanto līnijas no bufera \mathbb{B}_e , lai izveidotu attēlu \mathbf{I}_e . Šis attēls tiek izmantots, lai ar Otsu metodi automātiski noteiktu sliekšņa vērtību T_e . \mathbf{I}_e tiek ik pa laikam atjaunots, tādēļ arī T_e pielāgojas mainīgajiem ārējiem apstākļiem.

***IoVDL* metodes uzlabojumi.** Līdz šim tika aprakstīta *IoVDL* algoritma vienkāršākā un efektīvākā forma. Promocijas darbā tiek piedāvāti papildu uzlabojumi, kas paplašina *IoVDL* iespējas, bet samazina tās efektivitāti.

- **Objektu aizklāšanas detektēšana** uzlabo objektu atklāšanas precizitāti, analizējot intervāla platuma izmaiņas tā pastāvēšanas laikā.
- **Ēnu atklāšana** palīdz, kad izteiktas objektu ēnas rada intervālus, kas ir būtiski platāki par atbilstošajiem objektiem.
- **Mašīnu atklāšana naktī** uz ceļiem ar vāju apgaismojumu ir problemātiska video analizē balstītajām pieejām. Ja tiek izmantota kamera ar lielu aizvara ātrumu (1/24500 sekundes),

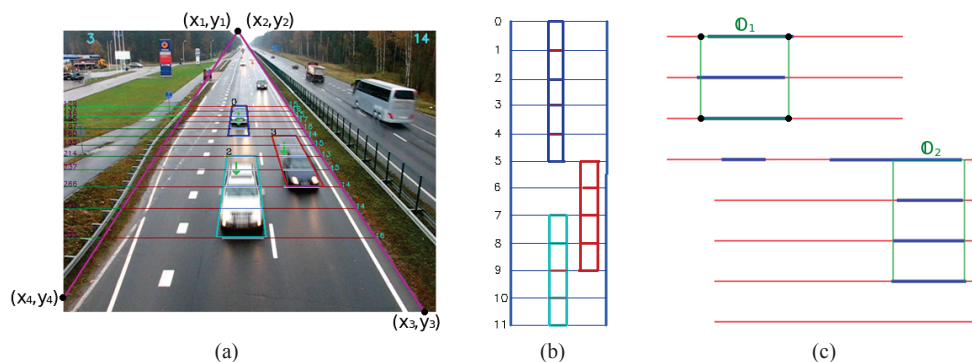
tad *IoVDL* spēj atklāt automašīnu lukturus.

- **Strauji apgaismojuma izmaiņu detektēšana** nepieciešamības gadījumā ļauj strauji atjaunot parasti lēni mainīgu fonu.
- **Palielināta jutība kustību atklāšanai** tiek panākta, apvienojot vairākas blakus līnijas viena detektora veidošanā. Šāda pieeja uzlabo fonam līdzīgu objektu atklāšanu.

2.3. *IoVDL* paplašinājums objektu raksturošanai

Intervālu apvienošana objektos. *IoVDL* metodes redzes lauks ir mazs kadra reģions. Šeit tiek piedāvāts *IoVDL* paplašinājums, kas iegūst informāciju no krietni plašāka kadra apgabala, apstrādājot dažas atklāšanas līnijas. Paplašinātā *IoVDL* spēj izsekot objektus un nomērīt to ātrumu un izmērus. Intervālu pieejas priekšrocības tiek saglabātas, un metodei ir nepieciešamas krietni mazākas skaitļošanas jaudas kā alternatīvām izsekošanas metodēm, kas apstrādā visus pikseļus objektu izsekošanas apgabalā.

Kadrā tiek izvietotas vairākas atklāšanas līnijas (2.2. a att.), un katra līnija ir perpendikulāra objektu kustībai. Pasaules koordinātēs (2.2. b att.), līnijas ir savā starpā paralēlas un izvietotas ik pēc noteikta vienāda attāluma.



2.2. att. Virtuālo atklāšanas līniju intervālu apvienošana objektos.

Dažādās atklāšanas līnijas var uzturēt intervālus, ko izveidojis viens un tas pats kustīgs objekts. Šādi intervāli tiek atrasti un apvienoti taisnstūra formas virtuālos objektos. (O_1 un O_2 2.2. c att.). Pievienojot jaunus un dzēšot esošus objekta intervālus, tiek izsekota objekta kustība cauri detektoriem.

Paplašinātā *IoVDL* metode izmanto projekcijas transformāciju, lai analizētu objektus pasaules koordinātēs. Katrā kadrā ir jātransformē tikai intervāli, kas ir aprakstāmi tikai ar divām koordinātēm. Tādējādi *IoVDL* metodes transformācijas solis ir efektīvs.

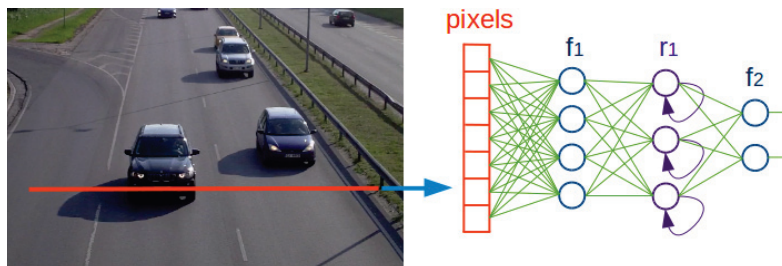
Objektu raksturošana. Visi atklātie objekti tiek analizēti pasaules koordinātēs, lai iegūtu šādus parametrus.

- **Kustības virziens** tiek noteikts pēc virtuālā objekta puses, kurā tam tiek pievienoti jauni intervāli.
- **Ātrums** tiek noteikts skaitot laiku, kas pagājis, objektam sasniedzot dažādas atklāšanas līnijas.
- **Izmērs.** Objekta garums tiek iegūts, sareizinot tā ātrumu ar tā intervālu pastāvēšanas ilgumu. Augstums un platums tiek noteikti, analizējot objekta intervālu platumus.
- **Klase** tiek noteikta no nomērītajiem izmēra parametriem. ITS lietojuma gadījumā izmērs ļauj atšķirt vieglās un smagās automašīnas.

2.4. Uz *RNN* balstīta virtuāla atklāšanas līnija

Šī apakšnodaļa piedāvā vēl vienu jaunu metodi – *RNN-VDL*. Šeit atklāšanas līnijas pieeja tiek apvienota ar rekurento neironu tīklu, kas dziļo apmācību ļauj izmantot efektīvā video apstrādē.

Arhitektūra. Virtuālais detektors ir objektu kustībai perpendikulāra līnija. Šī līnija tiek laista cauri mākslīgam neironu tīklam, kura vienkāršota arhitektūra parādīta 2.3. att.



2.3. att. Uz *RNN* balstīta virtuāla atklāšanas līnija (*RNN-VDL*).

Pirmais un pēdējais piedāvātā tīkla slānis ir pilnībā savienotu neironu slāņi. Slēptais slānis ir *LSTM* šūna, kas veido izejas signālu, ņemot vērā $S = 40$ iepriekšējos kadrus. Visas tīkla aktivācijas funkcijas ir lineāri taisngrieži $\rho(x) = \max(0, x)$. Šāda modeļa apmācības mērķis ir atrast tādas tīkla svara koeficientus, kas pie ieejas vektoru secības $\mathbf{x}^{(t-S)}, \mathbf{x}^{(t-S+1)}, \mathbf{x}^{(t-S+2)} \dots \mathbf{x}^{(t)}$ tīkla izejā dotu pareizu klasi (atrod ieejas secībai atbilstošu marķējumu).

Datu marķējuma veidi. *RNN-VDL* izdotās klases var reprezentēt dažādus notikumus uz ieejas līnijas. Šajā darbā tiek piedāvāti trīs datu marķējuma veidi. Pirmajā variantā marķējums norāda, cik objektu atrodas uz līnijas šī brīža kadrā. Otrā veida marķējums rāda, cik objektu ir pametuši līniju šī brīža kadrā. Trešā veida marķējums parāda, cik objektu ir pametuši atklāšanas līniju pēdējos $S/2$ kados.

3. MARĶĒTU DATU IEGŪŠANAS PAŅĒMIENI

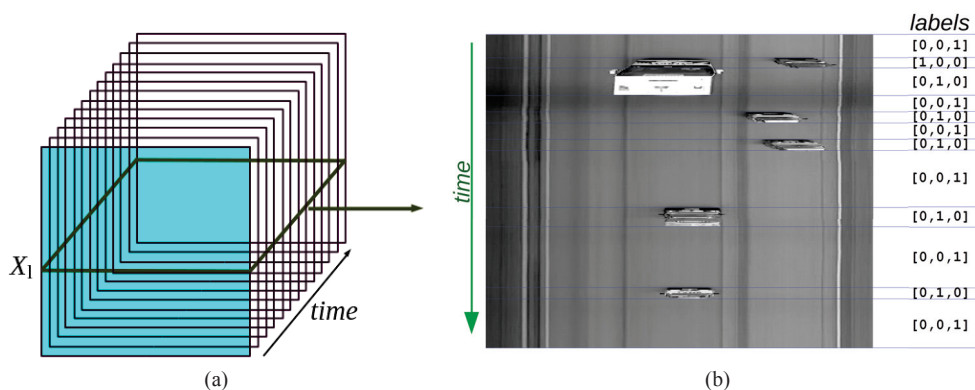
Darbā izveidotās uz mašīnmācīšanos balstītās *RNN-VDL* metodes lietojamība ir atkarīga no pieejas marķētiem treniņa datiem. Praktisku uzdevumu veikšanai, ir svarīgi spēt iegūt marķētus datus ātri un lēti. Šī nodaļa piedāvā trīs metodes vajadzīgo datu iegūšanai, sākot no manuālās marķēšanas, un beidzot ar pilnībā automatisku pieeju.

3.1. *RNN-VDL* treniņa datu formāts

RNN-VDL apmācībai nepieciešamos datus veido divas matricas – telpas-laika treniņa attēls un matrica ar bināri kodētiem marķējumiem.

Telpas-laika attēla rinda atbilst atklāšanas līnijas pikseliem konkrētā kadrā. Piemērs ir parādīts 3.1. a att., kur video ir attēlots kā taisnstūra paralēlskaldnis x, y un laika dimensijās. Treniņa attēls ir iezīmēts kā plakne, kas ir paralēla x -laika plaknei un šķeļ video paralēlskaldni cauri rindai X_1 .

Treniņa attēla piemērs un attiecīgi kodētais marķējums parādīts 3.1. b att. Piemērā apgabalus ar atšķirīgiem marķējumiem atdala tievas līnijas. Tiek izmantots pirmais marķējuma veids, tādējādi katras treniņattēla rindas marķējums norāda, cik objektu atrodas uz atklāšanas līnijas rindai atbilstošajā kadrā. *RNN* tīkla trenēšanas laikā šis attēls tiek sadalīts S kadru garās sekvencēs. Katras sekvences marķējums atbilst situācijai uz atklāšanas līnijas pēdējā secības kadra laikā.



3.1. att. Vēlamais marķēšanas rezultāts.

(a) ieejas video šķēlums, kas atbilst telpas-laika treniņa attēlam; (b) telpas-laika attēls ar līniju šķērsojušām automašīnām un tā rindām atbilstošo marķējumu.

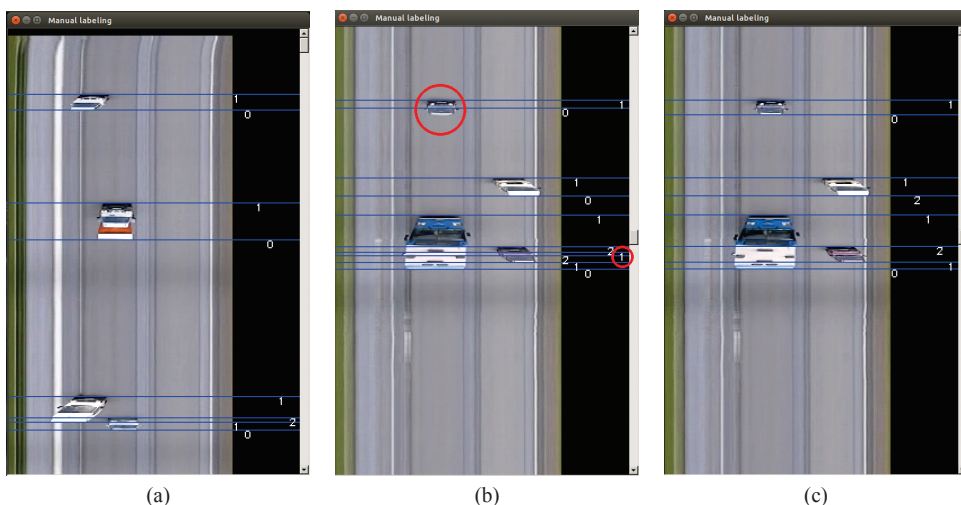
3.2. Piedāvātās pieejas datu marķēšanai

Sekojoši kadri video bieži ir līdzīgi viens otram, tādēļ tiem bieži ir vienāds marķējums. Laba saskarne ļautu lietotājam ātri noteikt un piešķirt šādiem kadriem attiecīgo marķējumu. Šajā darbā ir izstrādātas jaunas, uz grafiskām lietotāju saskarnēm balstītas metodes, kas atvieglo un paātrina datu marķēšanu. Abas metodes ir īstenotas *Python 3* valodā, izmantojot *OpenCV*, *numpy* un *PyQt4* bibliotēkas.

Manuāla marķēšanas metode izmanto specifisku grafisko saskarni, kas ļauj lietotājam ātri apskatīt visus video kadrus un manuāli piešķirt tiem marķējumu. Katrā kadrā marķētājam ir jāaplūko atklāšanas līnija un jānospiež numurs uz tastatūras. Numuram atbilstošais marķējums tiek kodēts un pievienots marķējumu matricai. Šī brīža kadra atklāšanas līnija kļūst par jaunu rindu telpas-laika treniņa attēlā. Pēc numura nospiešanas saskarne attēlo nākamo video kadru un gaida nākamo ievadu no lietotāja.

Šo marķēšanas metodi var izmantot, lai iegūtu datus arī tādām objektu atklāšanas metodēm, kas nav balstītas uz detektēšanas reģionu pieeju, bet gan apstrādā visu kadru. Šādā gadījumā lietotājam jāizskaita objekti visā kadrā, pirms norādīt kadram atbilstošu marķējumu.

Pusautomātiska marķēšanas pieeja ņem vērā *RNN-VDL* metodes specifiku, panākot vēl ātrāku datu marķēšanu. Vispirms algoritms cenšas marķēt datus automātiski ar fona atņemšanas pieeju. Rezultātā iegūtais sliekšņotais telpas-laika attēls tiek apstrādāts ar morfoloģiskajām aizvēršanas un atvēršanas operācijām, kas samazina troksni. Uz katras attēla rindas tiek izskaitīti balti nepārtraukti apgabali, un to skaits kļūst par attiecīgās rindas marķējumu (1. marķējuma veids).



3.2. att. Saskaņe datora marķēšanas kļūdu atrašanai un labošanai.

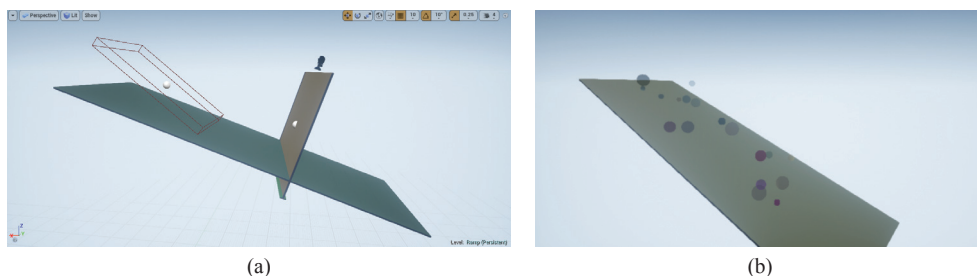
Nākamā marķēšanas fāzē sastāv no grafiskas saskarnes (3.2. att.), kas ļauj manuāli atrast un izlabot automātiskajā fāzē pieļautās kļūdas. 3.2. a att. parādīts telpas-laika attēla apgabals, kas ticis veiksmīgi nomarķēts automātiskajā fāzē. Katra zilā līnija attēlā parāda kadru, kurā marķējums ir mainījies. Cipars katras šādas līnijas labajā pusē parāda marķējumu visām zemāk esošajām rindām līdz nākamajai sadalošajai līnijai.

Cits attēla apgabals (3.2. b att.) satur divas automātiskajā fāzē pieļautas kļūdas. Izstrādātā saskarne ļauj lietotājam ātri pārskatīt visu telpas-laika attēlu un labot atrastās kļūdas ar peli un tastatūru. Telpas-laika skats ļauj cilvēkam novērtēt situāciju uz daudziem kadriem reizē. Tādējādi marķēšanas ātrums ir būtiski paātrināts, salīdzinot ar metodēm, kurās marķētājam jāapskata katrs kadrs atsevišķi.

3.3. Marķētu datu ģenerēšana ar 3D spēļu dzinēju

Jauni raksti mašīnmācīšanās jomā [56]–[59] norāda uz vēl vienu potenciālu risinājumu marķētu datu trūkumam – dziļu modeļu trenēšana ar sintētiskiem datiem. Lai novērtētu, vai šāda pieeja ir lietojama *RNN-VDL* metodei, spēļu dzinējā *Unreal Engine 4* tika izveidota simulācijas vide, kas ģenerē datus vienkāršam uzdevumam – kustīgu lodīšu atklāšanai video.

Simulācijas uzbūve parādīta 3.3. a att. Tā sastāv no slīpas bāzes plaknes, lodīšu ģeneratora, detektēšanas plaknes, apgaismojuma elementa un kameras. Kad simulācija tiek palaista (3.3. b att.), nejaušos laika brīžos un nejaušos ģeneratora apgabalos tiek ģenerētas lodītes. Tās krīt uz bāzes plakni, pa kuru tālāk ripo cauri detektēšanas plaknei. Šī plakne ļauj lodītēm brīvi rīpot cauri, bet simulācija izmēra, kuros kadros lodītes un plakne saskaras. Ar šo informāciju pietiek, lai izveidotu katram kadram atbilstošo marķējumu. Simulācijas kameras kadri tiek saglabāti līdz ar noteiktajiem marķējumiem.



3.3. att. Treniņa datu ģenerēšana ar 3D spēļu dzinēju.

Šis koncepta pierādījums parāda, ka *RNN-VDL* apmācībai nepieciešamos datus var ģenerēt arī automātiski. Tomēr reālistisku simulāciju veidošana praktiskiem uzdevumiem pieprasa specifiskas iemaņas un laiku. Datorgrafika paliek arvien reālistiskāka, tāpēc tuvā nākotnē šādai pieejai var būtiski pieaugt loma uz mašīnmācīšanos balstītajā datorredzē.

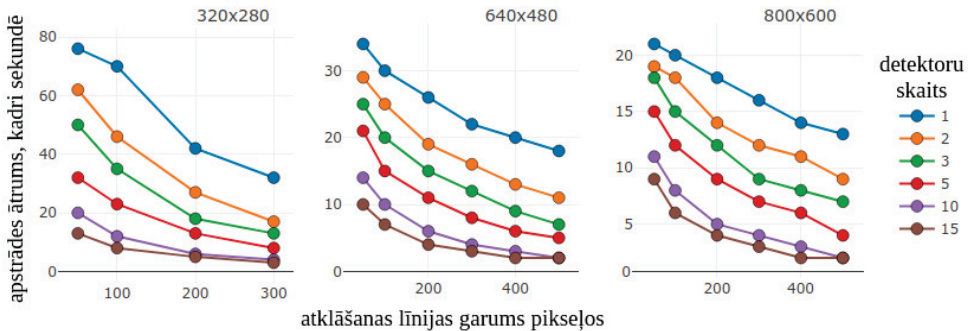
4. LIETOJUMI UN EKSPERIMENTĀLIE PĒTĪJUMI

Šīs nodaļas mērķis ir eksperimentāli pārbaudīt, vai 2. nodaļā izstrādātās metodes ir efektīvas patērēto skaitļošanas resursu ziņā, un vai tās ir pietiekami precīzas praktiskam lietojumam ITS uzdevumos.

4.1. Izstrādāto metožu efektivitāte

IoVDL ir īstenota *C++* programmēšanas valodā. Tā izmanto *OpenCV* bibliotēku video ievadei un attēlošanai, kā arī citām klasiskām apstrādes operācijām. Lai pierādītu *IoVDL* efektivitāti, metode ir īstenota uz *Raspberry Pi Zero* datora. *Zero* ir mazākais ($65 \times 30 \times 5$ mm) un lētākais (~ 5 EUR) *Raspberry* modelis, ar 1 GHz *CPU* un 512 MB *RAM*.

4.1. att. parādīts, cik kadrus sekundē *IoVDL* spēj apstrādāt uz *Raspberry Pi Zero*. Testa laikā kadri tiek iegūti reālā laikā ar *Raspberry* kameras moduli *V2*. Kadru iegūšana aizņem lielu patērētā laika daļu, kuras lielums atkarīgs no uzstādītās izšķirtspējas. Tādēļ 4.1. att. parādīti ātruma mērījumi pie trīs dažādām ieejas kadru izšķirtspējām.



4.1. att. *IoVDL* kadru apstrādes ātrums uz *Raspberry Pi Zero*.

RNN-VDL ir īstenota *Python 3* valodā, lietojot *TensorFlow* ietvaru. Galvenie metodes ātruma ietekmējošie faktori ir tīkla izmēru parametri – slāņu skaits un tīkla dziļums laikā S . Mērījumi rāda lineāru sakarību starp S un viena kadra apstrādes ilgumu F . Pie $S = 50$, $F \approx 0,004$ sekundes, pie $S = 100$, $F \approx 0,006$ s, pie $S = 150$, $F \approx 0,008$ s. Tīkla laiciskā dziļuma S izvēli nosaka konkrētais lietojuma uzdevums. Uzdevumi ar lēni kustīgiem objektiem un lielu kameras kadru uzņemšanas ātrumu var prasīt rekurentus tīklus ar dziļāku atmiņu.

Lai salīdzinātu *RNN-VDL* darbības ātrumu ar tādiem moderniem detektoriem kā *YOLO*, abas metodes tika darbinātas uz viena un tā paša procesora. *YOLO* ātrā versija *tiny-yolo* (pieejama [60]) apstrādāja vienu attēlu 0,34 sekundēs, kamēr *RNN-VDL* uz vienu kadru patērēja 0,0035 sekundes (pie $S = 40$).

RNN-VDL tika veiksmīgi īstenots un darbināts uz *Raspberry Pi 3 Model B* datora, kuram ir četru kodolu 1,2 GHz procesors un 1 GB RAM.

4.2. Objektu skaitīšanas precizitāte

Automašīnu skaitīšanas precizitāte gan ar *RNN-VDL*, gan *IoVDL* metodēm tika salīdzināta testos ar reāliem video (kadru piemēri doti 4.2. att.)



4.2. att. *IoVDL* un *RNN-VDL* testēšanai izmantotie video.

Testu rezultāti ir attēloti 4.1. tab., kur parādīta katrā testā uzrādītā precizitāte ($TP/(TP + FP)$), pārklājums ($TP/(TP + FN)$), un parametrs F_1 :

$$F_1 = 2 \cdot \frac{\text{precizitāte} \cdot \text{pārklājums}}{\text{precizitāte} + \text{pārklājums}}. \quad (4.1)$$

Tabulā parādītajam testam *Automašīnas 1* video tika iegūti no tām pašām lielceļa kamerām, no kurām tika iegūti *RNN-VDL* treniņa video (4.2. a, b att.). Saskaņā ar uzrādīto F_1 , *RNN-VDL* metode sniedz labākus rezultātus uz šāda veida video.

Izstrādājot *IoVDL*, tas tika veidots, balstoties uz video, kas pēc kvalitātes un izšķirtspējas ir līdzīgi testā *Automašīnas 2* izmantotajam video (4.2. c att.), un *IoVDL* šajā testā uzrāda labākus rezultātus.

Automašīnas 2 video savukārt būtiski atšķiras no *RNN-VDL* apmācībā izmantotajiem. Neraugoties uz to, neironu tīkls skaita automašīnas arī šāda video gadījumā. Kā daudzām dziļajā apmācībā balstītām metodēm, *RNN-VDL* objektu atklāšanas vispārināšanas spēju var uzlabot, paplašinot treniņa kopu ar video no dažādiem avotiem.

Kā tika atzīmēts 2.1. apakšnodaļā, eksistē citas efektīvas, uz detektēšanas reģioniem balstītas, bet lietojumā mazāk elastīgas objektu atklāšanas metodes. Raksti, kur šīs metodes piedāvātas, norāda uz līdzīgiem precizitātes testu rezultātiem. Rakstā [54] izstrādātajai metodei precizitāte uz dienas video variēja no 70,31 % līdz 98,39 %. Metode rakstā [53] uzrādīja 97,73 % precizitāti. Rakstā [51] piedāvātajai metodei tiek norādīta 86 %–96 % precizitāte un 89 %–96 % pārklājums. Visu metožu testēšanai ir izmantoti atšķirīgi video, tāpēc precizitātes rādījumi ir tikai indikatīvi.

Objektu atklāšanas rezultāts

Uzdevums	Automašīnas 1		Automašīnas 2		Cilvēki
	<i>RNN-VDL</i>	<i>IoVDL</i>	<i>RNN-VDL</i>	<i>IoVDL</i>	<i>RNN-VDL</i>
Patiesais skaits	4238	4238	121	121	285
Izdotais skaits	4278	3693	108	120	245
Pareiza atbilsme (<i>TP</i>)	4059	3623	103	119	209
Kļūdaina atbilsme (<i>FP</i>)	219	70	5	1	36
Kļūdaina neatbilsme (<i>FN</i>)	179	615	18	2	76
Precizitāte	0,949	0,981	0,954	0,992	0,853
Pārklājums	0,958	0,855	0,851	0,983	0,733
F_1	0,953	0,914	0,900	0,987	0,788

Papildu testi ar *IoVDL* uz dažādiem video norāda uz šādām metodes īpašībām:

- metode ir robusta pie mainīga apgaismojuma;
- metode neieskaita mākoņu un koku ēnas kā automašīnas;
- stiprs lietus būtiski neietekmē metodes precizitāti;
- kameras novietojums attiecībā pret ceļu ietekmē metodes precizitāti;
- metode veiksmīgi strādā uz ceļiem ar dažāda skaita un virziena braukšanas joslām (vai vispār bez joslām).

Cilvēku skaitīšanas tests demonstrē *RNN-VDL* metodes daudzpusību. Tā pati tīkla arhitektūra, kas tika lietota transportlīdzekļu atklāšanas uzdevumam, tika pārtrenēta cilvēku skaitīšanai (4.3. a, b att.)

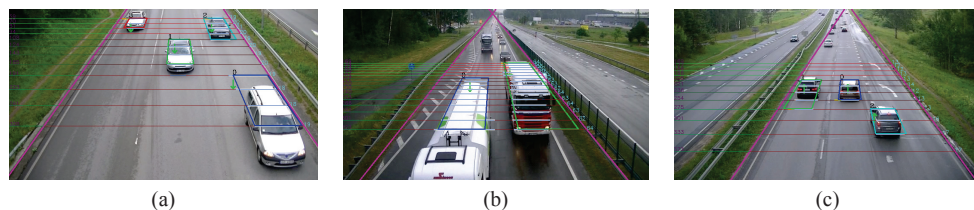
4.3. att. Cilvēku skaitīšana ar *RNN-VDL*.

Lai parādītu papildu ieguvumu no rekurentā tīkla lietošanas, cilvēku skaitīšanas uzdevums tika modificēts – ir jāskaita tika vienā virzienā ejoši cilvēki. Uzdevumam atbilstošs datu marķējums parādīts 4.3. c att. Šis piemērs ir iegūts no pusautomātiskās marķēšanas saskarnes, kas tika izstrādāta 3.2. apakšnodaļā.

Cilvēku skaitīšanas rezultāti ir parādīti 4.1. tab. kolonnā *Cilvēki*. Testa video, cilvēku plūsmas intensitāte ir līdzīga abos virzienos. Tādējādi, ja atklāšanas līnija nebūtu spējusi atšķirt cilvēku iešanas virzienus, tad rezultātos būtu daudz kļūdainas atbildes tipa kļūdu.

4.3. Automašīnu atklāšana un klasificēšana ar paplašināto *IoVDL*

Darbā izstrādātā paplašinātā *IoVDL* metode spēj noteikt objektu parametrus, kā arī klasificēt atklātos objektus. Šī klasificēšanas spēja tika testēta uz vairākiem video ar dažādiem laika apstākļiem un dažādiem kameras novietojumiem. Piemēra kadri no testa video ir doti 4.4. att.



4.4. att. Paplašinātās *IoVDL* testi.

4.2. tab. salīdzina transportlīdzekļu skaitīšanas precizitāti, kas iegūta ar vienu *IoVDL* detektoru un ar paplašināto *IoVDL*, kas testos sastāvēja no 12–16 atklāšanas līnijām. Vidējā no paplašinātās *IoVDL* metodes atklāšanas līnijām tika izmantota, lai novērtētu nepaplašinātās *IoVDL* metodes skaitīšanas precizitāti. Papildus tam tabula parāda arī skaitīšanas rezultātu, kas iegūts, aprēķinot vidējo vērtību no visu kadrā esošo līniju rādījumiem. Rezultāti parāda, ka paplašinātā *IoVDL* izdod patiesajam skaitam tuvāku lielumu nekā viens pats detektors vai vairāku detektoru vienkārša kombinācija.

4.2. tab. parāda arī paplašinātās *IoVDL* metodes klasificēšanas precizitāti. Metode pareizi klasificēja 88 % no smagajām automašīnām (*TP*), pieļaujot maz kļūdainas atbildes tipa kļūdu.

4.2. tabula

Satiksmes monitorēšana ar paplašināto *IoVDL*

Video	Patiesais skaits		Izdotais skaits			Klasificēšana	
	Visas automašīnas	Automašīnas	Parastā <i>IoVDL</i>	<i>IoVDL</i> vidējais	Paplašinātā <i>IoVDL</i>	<i>TP</i> smagās	<i>FP</i> smagās
a)	32	4	30	30	32	4	0
b)	122	8	121	124	122	8	0
c)	159	8	168	164	158	7	1
d)	43	0	45	44	44	0	0
e)	221	12	232	235	231	7	0
f)	208	5	209	209	209	5	1
g)	264	12	233	239	239	12	3
h)	117	4	118	118	116	4	1
i)	88	7	92	93	88	6	0

5. SECINĀJUMI

Promocijas darbā tika pētītas un izstrādātas video apstrādes metodes kustīgu objektu atklāšanai. Viens no tiešākajiem izstrādāto metožu lietojumiem ir automašīnu atklāšana inteligentās transporta sistēmās. Dažādo uz ceļiem izmantoto sensoru analīze 2.1. apakšnodaļā norādīja, ka metožu lietojamību praktiskiem uzdevumiem uz ceļa nosaka to spēja pielāgoties mainīgiem apstākļiem, kā arī ar metodēm saistītās izmaksas. Atbilstoši secinājumam promocijas darba izvirzītais mērķis bija uzlabot pret skaitļošanas jaudām efektīvu video apstrādi kustīgu objektu atklāšanai. Mērķa sasniegšanai tika izvirzīti seši uzdevumi.

1. Veikt literatūras analīzi par objektu atklāšanu attēlos un video. Šis uzdevums tika paveikts 1. nodaļā. No literatūras apskata tika secināts, ka šobrīd labāko objektu atklāšanu attēlos panāk ar dziļās apmācības metodēm. Rekurentie neironu tīkli tika identificēti kā potenciāli labi risinājumi video apstrādes uzdevumiem. Tomēr šodien lietotās dziļās apmācības metodes pieprasa būtiskus skaitļošanas resursus.

2. Identificēt efektīvas pieejas kustīgu objektu atklāšanai video. Tas tika paveikts 2.1. apakšnodaļā, kas parādīja, ka efektīvas metodes bieži tiek veidotas, samazinot kadrā apstrādājamo pikseļu skaitu. Tika identificēts būtisks šādu metožu trūkums – lai objektus būtu iespējams atklāt, tiem ir jānonāk virtuālajos atklāšanas reģionos. Tādēļ esošo efektīvo metožu lietojums ir ierobežots ar uzdevumiem, kuros kustīgie objekti nenovirzās no to paredzamajām kustību trajektorijām.

3. Izstrādāt efektīvas kustīgu objektu atklāšanas metodes, kam būtu uzlabotas spējas, salīdzinot ar esošām efektīvām metodēm. 2.2. apakšnodaļā tika izstrādāta metode *IoVDL*. Autotransporta atklāšanas kontekstā piedāvātā intervālu metode nepieprasa, lai automašīnas nemaina to pierastās trajektorijas. Metode ir lietojama uz ceļiem, kur braukšanas joslas nav definētas vai arī to skaits dienas laikā mainās. *IoVDL* ir spējīga noteikt vairākus objektus, kas vienlaikus šķērso virtuālo atklāšanas līniju.

Paplašinātā *IoVDL* metode spēj izsekot objektu kustībai cauri vairākiem virtuālajiem detektoriem. Metode spēj raksturot un klasificēt objektus, nomērot to ātrumu un izmērus. Metodes apraksts parādīja, ka tā joprojām ir efektīva, jo tiek apstrādāti tikai atklāšanas līniju pikseļi.

Vēl viena jauna video apstrādes pieeja kustīgu objektu atklāšanai tika izstrādāta 2.4. apakšnodaļā. *RNN-VDL* apvieno efektīvas atklāšanas līnijas ideju ar rekurento neironu tīklu. Dziļa tīkla lietošana padarīja *RNN-VDL* vēl elastīgāku un izmantojamu plašākam uzdevumu lokam nekā *IoVDL*. Tomēr šī pieeja pieprasa būtisku marķētu datu kopu neironu tīkla apmācībai.

4. Izstrādāt marķētu datu iegūšanas metodes, izveidot datu kopas darbā izstrādāto atklāšanas metožu apmācībai un testēšanai. Uzdevums tika paveikts 3. nodaļā. Tika izstrādāta grafiskā lietotāja saskarne manuālas kadru marķēšanas paātrināšanai. Tika secināts, ka šo pieeju var lietot datu marķēšanai dažādām metodēm, ne tikai uz virtuāliem detektoriem balstītām.

Tika izstrādāta vēl viena, pusautomātiska, marķēšanas metode. Tās praktiskā vērtība tika apstiprināta, ar tās palīdzību izveidojot marķētu datukopu, kas tika izmantota *RNN-VDL* metodes trenēšanai un testēšanai. Telpas-laika attēla izmantošana, lai atrastu un atzīmētu dažādus notikumus uz atklāšanas līnijas, izrādījās daudz ātrāka pieeja par katra kadra atsevišķu analīzi. Tas apstiprina promocijas darba 4. tēzi.

5. Īstenot darbā piedāvātās metodes, veikt to eksperimentālu izpēti. Visas piedāvātās atklāšanas un marķēšanas metodes tika īstenotas un testētas uz reāliem video (4. nodaļa).

Testi parādīja, ka *IoVDL* automašīnu atklāšanas precizitāte ir salīdzināma ar esošu efektīvo metožu rādījumiem. Testu rezultāti kombinācijā ar metodes aprakstu 2.2. apakšnodaļā apstiprina 1. darba tēzi.

Paplašinātā *IoVDL* izrādījās precīzāka objektu atklāšanas metode par vienkāršu *IoVDL*. Paplašinātās metodes spēja raksturot objektus tika apstiprināta ar autotransporta klasificēšanas testu. Iegūtie rezultāti un tas, ka *IoVDL* spēj izsekot objektus plašā kadra apgabalā, apstrādājot krietni mazāk pikseļus nekā klasiskas izsekošanas metodes, pierāda 2. darba tēzi.

RNN-VDL metode ir vēl plašāk lietojama par *IoVDL*. Apmācot *RNN-VDL* ar dažādiem marķētiem datiem, tā pati modeļa arhitektūra tika izmantota gan automašīnu, gan cilvēku skaitīšanai, kas apstiprina darba 3. tēzi. Vienas kadra līnijas apstrāde spēja arī atšķirt, kurā virzienā cilvēki šķērso šo līniju.

6. Izdarīt secinājumus par darbā sasniegtiem rezultātiem. Uzdevums paveiks šajā nodaļā, kur doti secinājumi par veikto literatūras analīzi, izstrādātajām atklāšanas un marķētu datu iegūšanas metodēm, kā arī par darbā veikto testu rezultātiem. Galvenais secinājums par darbā izstrādātajām kustīgu objektu atklāšanas metodēm ir tas, ka to īpašības un eksperimentālie testi norāda, ka promocijas darba mērķis ir sasniegts – ir izstrādātas uzlabotas efektīvas metodes kustīgu objektu atklāšanai video datos. Sasniegtais mērķis pierāda arī četras darba sākumā izvirzītās tēzes.

Šī darba pienesums var būt interesants gan datorredzes jomas pētniekiem, gan praktiķiem. Izstrādātās datu marķēšanas pieejas ir pienesums dziļās apmācības jomai, kur marķētu datu trūkums ir viens no šī brīža lielākajiem izaicinājumiem. Izstrādātās *IoVDL* un *RNN-VDL* metodes piedāvā jaunus veidus, kā veidot efektīvus, uz redzi balstītus, kustīgu objektu detektorus. Metodes tika galvenokārt veidotas un testētas uz ceļiem, tomēr šo metožu, īpaši *RNN-VDL*, elastība paver iespēju tās izmantot arī citos praktiskos uzdevumos.

LITERATŪRA

- [1] Saxena, A., Driemeyer, J., Ng, A. Y. Robotic grasping of novel objects using vision. No: *The International Journal of Robotics Research*, 2008, 27 (2), 157–173.
- [2] Brosnan, T., Sun, D.-W. Improving quality inspection of food products by computer vision – a review. No: *Journal of food engineering*, 2004, 61 (1), 3–16.
- [3] Maglogiannis, I., Doukas, C. N. Overview of advanced computer vision systems for skin lesions characterization. No: *IEEE transactions on information technology in biomedicine*, 2009, 13 (5), 721–733.
- [4] Kang, J., Hsu, C.-H., Wu, Q., Liu, S., Coster, A. D., Posner, B. A., Altschuler, S. J., Wu, L. F. Improving drug discovery with high-content phenotypic screens by systematic selection of reporter cell lines. No: *Nature biotechnology*, 2016, 34 (1), 70–77.
- [5] Parkhi, O. M., Vedaldi, A., Zisserman, A. Deep Face Recognition. No: *BMVC 1(3)*. 2015, p. 6.
- [6] Junior, J. C. S. J., Musse, S. R., Jung, C. R. Crowd analysis using computer vision techniques. No: *IEEE Signal Processing Magazine*, 2010, 27 (5), 66–77.
- [7] Coifman, B., Beymer, D., McLauchlan, P., Malik, J. A real-time computer vision system for vehicle tracking and traffic surveillance. No: *Transportation Research Part C: Emerging Technologies*, 1998, 6 (4), 271–288.
- [8] Li, H., Shen, C. Reading Car License Plates Using Deep Convolutional Neural Networks and LSTMs. No: *arXiv preprint arXiv:1601.05610*, 2016, p. 17.
- [9] **Kadiķis, R.** *Labeled dataset for training and testing virtual detection-line based object detectors* [tiešsaiste]. Rīga: EDI, 2017 [skafīts 2017. g. 28. novembrī]. Pieejams: <http://vault.edi.lv/index.php/s/yxGPtI2Jw2yWhLR>
- [10] Freivalds, K., **Kadiķis, R.**, Greitāns, M. System and method for video-based vehicle detection. Patents (EP 2700054 B1), Apr. 2015.
- [11] **Kadiķis, R.**, Freivalds, K. Efficient video processing method for traffic monitoring combining motion detection and background subtraction. No: *Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012)*. Springer, 2013, 131–141.
- [12] **Kadiķis, R.**, Freivalds, K. Vehicle classification in video using virtual detection lines. No: *Sixth International Conference on Machine Vision (ICMV 13)*. International Society for Optics and Photonics 2013, 90670Y–90670Y.
- [13] **Kadiķis, R.** Registration method for multispectral skin images. No: *Radioelektronika (RADIOELEKTRONIKA)*, 2015 25th International Conference. IEEE 2015, 232–235.

- [14] Nauris, D., **Kadiķis**, R., Nesenbergs, K. Vehicle type and license plate localisation and segmentation using FCN and LSTM. No: *New Challenges of Economic and Business Development 2017, Proceedings of Reports*. University of Latvia 2017, 132–140.
- [15] Tamošiūnas, M., Jakovels, D., Ļihačovs, A., Kilikevičius, A., Baltušnikas, J., **Kadiķis**, R., Šatkauskas, S. Application of fluorescence spectroscopy and multispectral imaging for non-invasive estimation of GFP transfection efficiency. No: *8th International Conference on Advanced Optical Materials and Devices*. International Society for Optics and Photonics 2014, 94210M–94210M.
- [16] Jakovels, D., Saknīte, I., Bliznuks, D., Spigulis, J., **Kadiķis**, R. Benign-A typical nevi discrimination using diffuse reflectance and fluorescence multispectral imaging system. No: *International Conference on BioPhotonics (BioPhotonics)*. IEEE 2015, 1–4.
- [17] Tamošiūnas, M., **Kadiķis**, R., Saknīte, I., Baltušnikas, J., Kilikevičius, A., Lihachev, A., Petrovska, R., Jakovels, D., Šatkauskas, S. Noninvasive optical diagnostics of enhanced green fluorescent protein expression in skeletal muscle for comparison of electroporation and sonoporation efficiencies. No: *Journal of biomedical optics*, 2016, 21 (4), 045003–045003.
- [18] Bellucci, P., Cipriani, E. Data accuracy on automatic traffic counting: the SMART project results. No: *European transport research review*, 2010, 2 (4), 175–187.
- [19] Yu, X., Sulijoadikusumo, G., Li, H., Prevedouros, P. Reliability of Automatic Traffic Monitoring with Non-Intrusive Sensors. No: *ICCTP 2011: Towards Sustainable Transportation Systems* 2011, 4157–4169.
- [20] Russakovsky, O. et al. ImageNet Large Scale Visual Recognition Challenge. No: *International Journal of Computer Vision (IJCV)*, 2015, 115 (3), 211–252.
- [21] Karpathy, A. *What I learned from competing against a ConvNet on ImageNet* [tiešsaiste]. [skatīts 2017. g. 2. augustā]. Pieejams: <http://karpathy.github.io/2014/09/02/what-i-learned-from-competing-against-a-convnet-on-imagenet/>
- [22] Otsu, N. A threshold selection method from gray-level histograms. No: *IEEE transactions on systems, man, and cybernetics*, 1979, 9 (1), 62–66.
- [23] Rodríguez, T., García, N. An adaptive, real-time, traffic monitoring system. No: *Machine Vision and Applications*, 2010, 21 (4), 555–576.
- [24] Ji, X., Wei, Z., Feng, Y. Effective vehicle detection technique for traffic surveillance systems. No: *Journal of Visual Communication and Image Representation*, 2006, 17 (3), 647–658.
- [25] Gupte, S., Masoud, O., Martin, R. F., Papanikolopoulos, N. P. Detection and classification of vehicles. No: *IEEE Transactions on intelligent transportation systems*, 2002, 3 (1), 37–47.

- [26] Cheung, S.-C. S., Kamath, C. Robust background subtraction with foreground validation for urban traffic video. No: *EURASIP Journal on Advances in Signal Processing*, 2005, 2005 (14), 2330–2340.
- [27] Stauffer, C., Grimson, W. E. L. Learning patterns of activity using real-time tracking. No: *IEEE Transactions on pattern analysis and machine intelligence*, 2000, 22 (8), 747–757.
- [28] Harris, C., Stephens, M. A combined corner and edge detector. No: *Alvey vision conference*. 15 (50), Manchester, UK 1988, 147–152.
- [29] Mikolajczyk, K., Schmid, C. An affine invariant interest point detector. No: *Computer Vision—ECCV 2002*, 2002, 128–142.
- [30] Lowe, D. G. Distinctive image features from scale-invariant keypoints. No: *International journal of computer vision*, 2004, 60 (2), 91–110.
- [31] Matas, J., Chum, O., Urban, M., Pajdla, T. Robust wide-baseline stereo from maximally stable extremal regions. No: *Image and vision computing*, 2004, 22 (10), 761–767.
- [32] Rosten, E., Drummond, T. Fusing points and lines for high performance tracking. No: *ICCV 2005. 10th IEEE International Conference on Computer Vision*. IEEE 2005, 1508–1515.
- [33] Rosten, E., Drummond, T. Machine learning for high-speed corner detection. No: *Computer Vision—ECCV 2006*, 2006, 430–443.
- [34] Lowe, D. G. Distinctive image features from scale-invariant keypoints. No: *International journal of computer vision*, 2004, 60 (2), 91–110.
- [35] Bay, H., Tuytelaars, T., Van Gool, L. Surf: Speeded up robust features. No: *Computer vision—ECCV 2006*, 2006, 404–417.
- [36] Ojala, T., Pietikäinen, M., Harwood, D. A comparative study of texture measures with classification based on featured distributions. No: *Pattern recognition*, 1996, 29 (1), 51–59.
- [37] Rublee, E., Rabaud, V., Konolige, K., Bradski, G. ORB: An efficient alternative to SIFT or SURF. No: *ICCV 2011. IEEE International Conference on Computer Vision*. IEEE 2011, 2564–2571.
- [38] Girshick, R., Donahue, J., Darrell, T., Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. No: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014, 580–587.
- [39] Ren, S., He, K., Girshick, R., Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. No: *Advances in neural information processing systems*. 2015, 91–99.

- [40] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. You only look once: Unified, real-time object detection. No: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, 779–788.
- [41] Karnowski, J. *AlexNet + SVM* [tiešsaiste]. [skatīts 2017. g. 5. decembrī]. Pieejams: <https://jeremykarnowski.files.wordpress.com/2015/07/alexnet2.png>
- [42] Lipton, Z. C., Berkowitz, J., Elkan, C. A critical review of recurrent neural networks for sequence learning. No: *arXiv preprint arXiv:1506.00019*, 2015, p. 38.
- [43] Hochreiter, S., Schmidhuber, J. Long short-term memory. No: *Neural computation*, 1997, 9 (8), 1735–1780.
- [44] Kastrinaki, V, Zervakis, M., Kalaitzakis, K. A survey of video processing techniques for traffic applications. No: *Image and vision computing*, 2003, 21 (4), 359–381.
- [45] Yue, Y. A traffic-flow parameters evaluation approach based on urban road video. No: *Int. J. Intell. Eng. Syst*, 2009, 2 (1), 33–39.
- [46] Mithun, N. C., Rashid, N. U., Rahman, S. M. Detection and classification of vehicles from video using multiple time-spatial images. No: *IEEE Transactions on Intelligent Transportation Systems*, 2012, 13 (3), 1215–1225.
- [47] Lee, G.-G., Kim, B.-s., Kim, W.-Y. Automatic estimation of pedestrian flow. No: *ICDSC 2007. First ACM/IEEE International Conference on Distributed Smart Cameras*. IEEE 2007, 291–296.
- [48] Kim, J.-W., Choi, K.-S., Choi, B.-D., Ko, S.-J. Real-time vision-based people counting system for the security door. No: *International Technical Conference on Circuits/Systems Computers and Communications*. 2002, 1416–1419.
- [49] Li, L., Han, S., Asama, H., Duan, F. An automatic parts detection system based on computer vision. No: *35th Chinese Control Conference (CCC)*. IEEE 2016, 9493–9498.
- [50] Nie, Z., Hung, M.-H., Huang, J. A Novel Algorithm of Rebar Counting on Conveyor Belt Based on Machine Vision. No: *Journal of Information Hiding and Multimedia Signal Processing*, 2016, 7 (2), 425–437.
- [51] Anan, L. Video vehicle detection algorithm based on virtual-line group. No: *APCCAS 2006. IEEE Asia Pacific Conference on Circuits and Systems*. IEEE 2006, 1148–1151.
- [52] Michalopoulos, P. G. Vehicle detection video through image processing: the autoscope system. No: *IEEE Transactions on vehicular technology*, 1991, 40 (1), 21–29.
- [53] Zhang, G., Avery, R., Wang, Y. Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras. No: *Transportation Research Record: Journal of the Transportation Research Board*, 2007, (1993), 138–147.

- [54] Lei, M., Lefloch, D., Gouton, P., Madani, K. A video-based real-time vehicle counting system using adaptive background method. No: *SITIS'08. IEEE International Conference on Signal Image Technology and Internet Based Systems*. IEEE 2008, 523–528.
- [55] Ha, D., Lee, J.-M., Kim, Y.-D. Neural-edge-based vehicle detection and traffic parameter extraction. No: *Image and vision computing*, 2004, 22 (11), 899–907.
- [56] Lerer, A., Gross, S., Fergus, R. Learning Physical Intuition of Block Towers by Example. No: *arXiv preprint arXiv:1603.01312*, 2016, p. 11.
- [57] Li, W., Azimi, S., Leonardis, A., Fritz, M. To Fall Or Not To Fall: A Visual Approach to Physical Stability Prediction. No: *arXiv preprint arXiv:1604.00066*, 2016, p. 20.
- [58] Gaidon, A., Wang, Q., Cabon, Y., Vig, E. Virtual worlds as proxy for multi-object tracking analysis. No: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, 4340–4349.
- [59] Sadeghi, F., Levine, S. CAD2RL: Real single-image flight without a single real image. No: *arXiv preprint arXiv:1611.04201*, 2016, p. 12.
- [60] *Darkflow* [tiešsaiste]. [skafīts 2017. g. 6. jūlijā]. No: <https://github.com/thtrieu/darkflow>



Roberts Kadiķis ir dzimis 1988. gadā. Absolvējis Rīgas Tehnisko universitāti, kur attiecīgi 2009. un 2011. gadā ieguvis bakalaura un maģistra grādus elektronikas studiju programmā. Kopš 2011. gada strādā kā pētnieks Elektronikas un datorzinātņu institūtā Rīgā.

Galvenās pētījumu intereses ir saistītas ar datorredzes algoritmu izstrādi, kas tiek izmantoti inteligentās transporta sistēmās, ražošanas robotos, biomedicīnas attēlu apstrādē un analīzē, kā arī citās jomās. Padziļināti interesē uz mašīnmācīšanos balstītas datorredzes metodes, kurās izmanto dziļos konvolūcijas un rekurentos neironu tīklus.