

COMPUTER ENGINEERING AND INFORMATION SYSTEMS / ЕСЕПТЕУ ТЕХНИКАСЫ ЖӘНЕ АҚПАРАТТЫҚ ЖҮЙЕЛЕР / ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ИНФОРМАЦИОННЫЕ СИСТЕМЫ

COMPUTER ENGINEERING AND INFORMATION SYSTEMS / ЕСЕПТЕУ ТЕХНИКАСЫ ЖӘНЕ АҚПАРАТТЫҚ ЖҮЙЕЛЕР / ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Industrial Transport of Kazakhstan

ISSN 1814-5787 (print)

ISSN 3006-0273 (online)

Vol. 23. Is. 3. Number 91 (2026). Pp. 54–64

Journal homepage: <https://prom.mtgu.edu.kz>

<https://doi.org/10.58420.ptk.2026.91.03.005>

UDC / УДК / ЭОЖ 28.23.15

OBSTACLE DETECTION ON RAILWAY TRACKS FROM UAV IMAGERY: A LEAKAGE-AWARE EVALUATION

*Y. Bagdollauly**

Satbayev University, Almaty, Kazakhstan.

E-mail: y.bagdollauly@stud.satbayev.university

Yessen Bagdollauly — PhD student, Satbayev University, 22 Kanysh Satpayev St., Almaty, Kazakhstan

E-mail: y.bagdollauly@stud.satbayev.university, <https://orcid.org/0000-0003-1583-0034>.

© Y. Bagdollauly

Abstract. On industrial sidings and station tracks, foreign objects and people create a collision risk during shunting, and automatic analysis of images taken by an unmanned aerial vehicle (UAV) could speed up track inspection. Such detectors are evaluated on the split published with an open dataset, but whether the test images are independent of the training images is rarely checked. This study measures data leakage in the published split of the open UAV-RSOD dataset (2,002 images, six obstacle classes) and its effect on the estimated detector quality. Near-duplicates were searched with SIFT keypoints and RANSAC geometric verification; a blocked split that reduces near-duplicates was built from the near-duplicate graph with a buffer zone; a YOLOv8n detector was fine-tuned with identical settings on a random and a blocked split of equal size; COCO-pretrained YOLOX detectors were tested on the person class without fine-tuning. Of the published test images, 390 of 391 (99.7 %) have a RANSAC-verified near-duplicate (augmented copy or adjacent frame) in the training set. Compared with a random split of equal size, the blocked split lowered the share of test images with a same-class near-duplicate from 99.8 % to 64.8 % but could not remove it, because all images come from one track section. YOLOv8n reached mAP50 = 0.806 on the random split and 0.719 on the blocked split (one run); the difference varied across classes. The best zero-shot AP50 for people seen from above was 0.455. All images come from one track section with staged obstacles, and false-alarm rates could not be estimated because the dataset has no obstacle-free images. The contribution is a quantified leakage estimate for UAV track-inspection data and a procedure to reduce it; in practice, detectors should be accepted on track sections not used for training.

Keywords: computer vision, obstacle detection, railway track, unmanned aerial vehicle, data leakage, object detection

For citation: Bagdollaully Y. (2026). Obstacle Detection on Railway Tracks from UAV Imagery: A Leakage-Aware Evaluation // Industrial Transport of Kazakhstan. Vol. 23. No. 91. Pp. 54–64. <https://doi.org/10.58420.ptk.2026.91.03.005> (In Kaz.).

Conflict of interest: The author declares no conflict of interest.

ҰШҚЫШСЫЗ ҰШУ АППАРАТЫ ТҮСІРГЕН КЕСКІНДЕР БОЙЫНША ТЕМІРЖОЛ ЖОЛЫНДАҒЫ КЕДЕРГІЛЕРДІ АНЫҚТАУ: ДЕРЕКТЕРДІҢ АҒЫП КЕТУІН ЕСКЕРЕТІН БАҒАЛАУ

Е. Бағдоллаұлы*

Satbayev University (Сәтбаев университеті), Алматы, Қазақстан.

E-mail: y.bagdollauly@stud.satbayev.university

Есен Бағдоллаұлы — докторант, Satbayev University, Қаныш Сәтбаев көшесі, 22, Алматы, Қазақстан

E-mail: y.bagdollauly@stud.satbayev.university, <https://orcid.org/0000-0003-1583-0034>

© Бағдоллаұлы Е.

Аннотация. Өнеркәсіптік кірме жолдар мен станция жолдарында бөгде заттар мен адамдар маневрлік жұмыс кезінде соқтығысу қаупін туғызады, ал ұшқышсыз ұшу аппаратымен (ҰҰА) түсірілген кескіндерді автоматты талдау жолды тексеруді жеделдетуі мүмкін. Мұндай детекторлардың сапасы ашық деректер жиындарында жарияланған бөлу бойынша бағаланады, бірақ тест кескіндерінің оқыту кескіндерінен тәуелсіздігі әдетте тексерілмейді. Зерттеудің мақсаты – ашық UAV-RSOD жиынында (2002 кескін, алты кедергі класы) жарияланған бөлудегі деректердің ағып кетуін өлшеу және оның детектор сапасының бағасына әсерін анықтау. Жақын көшірмелер SIFT кілттік нүктелері мен RANSAC арқылы геометриялық тексеру әдісімен ізделді; жақын көшірмелерді азайтатын блоктық бөлу көшірмелер графы бойынша буфермен құрылды; YOLOv8n детекторы бірдей көлемдегі кездейсоқ және блоктық бөлулерде бірдей параметрлермен қосымша оқытылды; COCO-да оқытылған YOLOX детекторлары адам класы бойынша қосымша оқытусыз тексерілді. Жарияланған тест кескіндерінің 390/391-інің (99,7 %) оқыту жиынында RANSAC арқылы тексерілген жақын көшірмесі (бір сахнаның түрлендірілген нұсқасы немесе көршілес кадр) бар. Бірдей көлемдегі кездейсоқ бөлумен салыстырғанда блоктық бөлуде бір кластағы жақын көшірмесі бар тест кескіндерінің үлесі 99,8 %-дан 64,8 %-ға дейін азайды, бірақ толық жойылмады, себебі барлық кескіндер бір учаскеден алынған. YOLOv8n детекторының mAP50 мәні кездейсоқ бөлуде 0,806, ал блоктық бөлуде 0,719 болды (бір рет оқыту); айырма кластар бойынша біркелкі емес. Жоғарыдан түсірілген адамдар үшін COCO детекторларының ең жақсы AP50 мәні – 0,455. Нәтижелер бір учаскеде, әдейі қойылған кедергілермен алынды және кедергісіз кескіндер болмағандықтан жалған дабылдар жиілігі бағаланбады. Жұмыстың жаңалығы – ҰҰА-мен жолды тексеру деректеріндегі ағып кетуді сандық бағалау және оны азайту тәсілі; практикалық құндылығы – детекторларды қабылдау кезінде оқытуға кірмеген учаскелерде бағалау талабы.

Түйін сөздер: компьютерлік көру, кедергілерді анықтау, теміржол жолы, ұшқышсыз ұшу аппараты, деректердің ағып кетуі, объектілерді детекциялау

Дәйексөз үшін: Бағдоллаұлы Е. (2026). Ұшқышсыз ұшу аппараты түсірген кескіндер бойынша теміржол жолындағы кедергілерді анықтау: деректердің ағып кетуін ескеретін бағалау // Қазақстан өндіріс көлігі. Том. 23. № 91. 54–64 бет. <https://doi.org/10.58420.ptk.2026.91.03.005> (Қаз. тіл.).

Мүдделер қақтығысы: Автор мүдделер қақтығысының жоқтығын мәлімдейді.

ОБНАРУЖЕНИЕ ПРЕПЯТСТВИЙ НА ЖЕЛЕЗНОДОРОЖНОМ ПУТИ ПО СНИМКАМ С БЕСПИЛОТНОГО ЛЕТАТЕЛЬНОГО АППАРАТА: ОЦЕНКА С УЧЁТОМ УТЕЧКИ ДАННЫХ

Е. Бағдоллаұлы*

Satbayev University, Алматы, Казахстан.

E-mail: y.bagdollauly@stud.satbayev.university

Есен Бағдоллаұлы — докторант, Satbayev University, ул. Каныша Сатпаева, 22, г. Алматы, Казахстан.

E-mail: y.bagdollauly@stud.satbayev.university, <https://orcid.org/0000-0003-1583-0034>.

© Бағдоллаұлы Е.

Аннотация. На подъездных и станционных путях промышленных предприятий посторонние предметы и люди создают риск столкновения при маневровой работе, а автоматический анализ снимков с беспилотного летательного аппарата (БПЛА) может ускорить осмотр пути. Качество таких детекторов оценивают по разбиению, опубликованному вместе с открытым набором данных, но независимость тестовых снимков от обучающих обычно не проверяют. Цель исследования – измерить утечку данных в опубликованном разбиении открытого набора UAV-RSOD (2002 снимка, шесть классов препятствий) и определить её влияние на оценку качества детектора. Близкие дубликаты искали по особым точкам SIFT с геометрической проверкой RANSAC; блочное разбиение, уменьшающее число близких дубликатов, построено по графу почти-дубликатов с буферной зоной; детектор YOLOv8n дообучен с одинаковыми параметрами на случайном и блочном разбиениях одинакового размера; детекторы YOLOX, обученные на COCO, проверены для класса «человек» без дообучения. 390 из 391 (99,7 %) опубликованных тестовых снимков имеют в обучающей выборке близкие дубликаты, подтвержденные RANSAC (преобразованные копии или соседние кадры). По сравнению со случайным разбиением того же размера в блочном доля тестовых снимков с близким дубликатом того же класса снизилась с 99,8 % до 64,8 %, но не исчезла, так как все снимки сделаны на одном участке. mAP50 детектора YOLOv8n составила 0,806 на случайном разбиении и 0,719 на блочном (один запуск); различие неоднородно по классам. Лучшее значение AP50 детекторов COCO для людей, снятых сверху, – 0,455. Результаты получены на одном участке пути с расставленными препятствиями; частоту ложных тревог оценить нельзя, так как в наборе нет снимков без препятствий. Новизна – количественная оценка утечки в данных осмотра пути с БПЛА и способ её снижения; практическая значимость – требование оценивать детекторы на участках, не входивших в обучение.

Ключевые слова: компьютерное зрение, обнаружение препятствий, железнодорожный путь, беспилотный летательный аппарат, утечка данных, детекция объектов

Для цитирования: Бағдоллаұлы Е. (2026). Обнаружение препятствий на железнодорожном пути по снимкам с беспилотного летательного аппарата: оценка с учётом утечки данных // Промышленный транспорт Казахстана. Т. 23. No. 91. Стр. 54–64. <https://doi.org/10.58420.ptk.2026.91.03.005> (На каз.).

Конфликт интересов: Автор заявляет об отсутствии конфликта интересов.

Кіріспе.

Өнеркәсіптік кәсіпорындардың кірме жолдары, станциялардың қабылдау-жөнелту жолдары және карьерлердегі теміржол тармақтары көбіне күзетілмейді, ал маневрлік жұмыс кезінде жолдағы бөгде зат немесе адам жылжымалы құрамның соқтығысуына, рельстен шығуына және жарақатқа әкелуі мүмкін. Жолды машинистің көзімен немесе жаяу айналып тексеру уақытты көп алады; ұшқышсыз ұшу аппараты (ҰҰА) жолды жоғарыдан түсіріп, кескіндерді автоматты талдауға беруге мүмкіндік береді. Мұндай жүйенің сенімділігі кедергіні анықтайтын модельдің жаңа, бұрын көрмеген жол учаскесіндегі сапасына тәуелді.

Кедергілерді анықтау модельдерінің сапасы әдетте ашық деректер жиындарында жарияланған бөлу (оқыту және тест жиындары) бойынша бағаланады. Егер тест кескіндері оқыту кескіндерінің бұрылған, айналы көрінісі немесе жарықтығы өзгертілген көшірмелері болса, модель жаңа сахнаны емес, бұрын көрген сахнаны таниды, ал өлшенген дәлдік пайдалану кезіндегі дәлдіктен жоғары болып шығуы мүмкін. Бұл құбылыс деректердің ағып кетуі (data leakage) деп аталады және машиналық оқытуға негізделген зерттеулердің қайталанбауының негізгі себептерінің бірі ретінде сипатталған (Karoor, Narayanan, 2023).

Көру арқылы кедергілерді анықтау бойынша зерттеулердің басым бөлігі локомотивтегі борттық камераларға арналған, ал олардың негізгі шектеуі ретінде нақты пайдалану жағдайындағы белгіленген деректердің жетіспеушілігі аталады (Ristić-Durrant және т.б., 2021). ҰҰА кескіндері жолды жоғарыдан, бір қалыпты масштабта көрсетеді, сондықтан маневрлік жұмыс басталар алдында учаскені тексеру міндетіне ыңғайлы. Бұл бағытта жарияланған ашық деректер жиындары аз, ал олардың бірі – UAV-RSOD – жоғары дәлдік нәтижелерімен бірге жарияланған (Rampriya және т.б., 2024). Мұндай нәтижелер жүйені енгізу туралы шешім қабылдау кезінде бағдар ретінде пайдаланылуы мүмкін, сондықтан олардың қандай жағдайда алынғанын тексеру практикалық маңызға ие.

Тәжірибеде деректер жиынын бөлудің әділдігі сирек тексеріледі. Кескіндер жиынында оқыту мен тест арасындағы жақын көшірмелер модельдердің рейтингін бұрмалайтыны белгілі (Barz, Denzler, 2020), ал бейнеден алынған кадрлар мен аугментацияланған кескіндерде мұндай көшірмелердің пайда болу ықтималдығы жоғары. Егер детектор кейін басқа кірме жолда немесе басқа маусымда пайдаланылса, тапсырыс берушіге оқытуда көрмеген учаскедегі сапа маңызды; ал оқыту кескіндерінің көшірмелерінде өлшенген сапа бұл сұраққа жауап бермейді. Бұл жұмыста бастапқыда борттық камерамен түсірілген RailGoerl24 жиыны жоспарланған, бірақ жұмыс кезінде ол жүктеп алуға қолжетімсіз болды, сондықтан талдау ҰҰА кескіндерінің ашық жиынында жүргізілді.

Зерттеу нысаны – теміржол жолын ҰҰА-мен түсіру арқылы кедергілерді анықтау. Зерттеу пәні – кедергілерді анықтау модельдерін бағалау әдістемесі, атап айтқанда оқыту мен тест жиындарын бөлу тәсілі. Мақсат – ашық UAV-RSOD деректер жиынында жарияланған бөлудегі ағып кетудің ауқымын өлшеу, жақын көшірмелерді азайтатын блоктық бөлуді құру және екі бөлудегі детектор сапасын салыстыру, сондай-ақ СОСО жиынында оқытылған дайын адам детекторларын қосымша оқытусыз тексеру.

Мақсатқа жету үшін келесі міндеттер қойылды: (1) жарияланған тест жиынының оқыту жиынымен геометриялық тексерілген жақын көшірмелерін анықтау және олардың үлесін бағалау; (2) барлық кескіндердің жақын көшірмелер графын құру және оның негізінде буфері бар блоктық бөлуді жасау; (3) бөлудің сапасын арнайы салыстырумен тексеру; (4) бірдей көлемдегі кездейсоқ және блоктық бөлулерде бір детекторды бірдей параметрлермен оқытып, сапасын салыстыру; (5) СОСО-да оқытылған детекторлардың жоғарыдан түсірілген адамдарды анықтау қабілетін бағалау. Мақала келесі түрде құрылған: әдебиетке шолу бөлімінде олқылық пен зерттеу сұрақтары тұжырымдалады, «Материалдар мен әдістер» бөлімінде деректер, бөлу тәсілдері және бағалау хаттамасы сипатталады, «Нәтижелер» бөлімінде үш зерттеу сұрағына жауап беріледі, ал «Талқылау» мен

«Қорытынды» бөлімдерінде нәтижелердің практикалық салдары мен шектеулері қарастырылады.

Теміржолдағы кедергілерді көру арқылы анықтау бойынша шолуда борттық камералар мен қашықтықты бағалау әдістері жүйеленген және нақты пайдалану жағдайындағы деректердің жетіспеушілігі негізгі шектеу ретінде аталған (Ristić-Durrant және т.б., 2021). Кейін теміржол сахналарының ашық деректер жиындары пайда болды: RailSem19 – жол сахналарын семантикалық сегменттеуге арналған жиын (Zendel және т.б., 2019: 1221–1229), OSDaR23 – локомотивтегі көп сенсорлы жиын (Tagiew және т.б., 2023: 270–276), RailGoerl24 – сынақ полигонында адамдар белгіленген борттық бейнежиын (Tagiew және т.б., 2025). Бұл жиындар борттық камера көрінісіне арналған; жолды жоғарыдан түсіру үшін ашық жиын ретінде UAV-RSOD жарияланды (Rampriya және т.б., 2024).

ҰҰА кескіндеріндегі нысандарды анықтау жалпы есеп ретінде VisDrone жарыстарында зерттелген, онда нысандардың кішілігі мен жоғарыдан қарау бұрышы негізгі қиындық ретінде көрсетілген (Du және т.б., 2019: 213–226); кішкене нысандар үшін кескінді бөліктерге бөліп талдау ұсынылған (Akyon және т.б., 2022: 966–970). Бір сатылы детекторлар (Redmon және т.б., 2016: 779–788) және олардың кейінгі нұсқалары, мысалы YOLOX (Ge және т.б., 2021) мен YOLOv8 (Jocher және т.б., 2023), шектеулі есептеу ресурсында жұмыс істеуге ыңғайлы. Бұл модельдердің көпшілігі COCO жиынында алдын ала оқытылған (Lin және т.б., 2014: 740–755), ал COCO кескіндері негізінен күнделікті сахналардың суреттері, ҰҰА-дан тігінен түсірілген кескіндер емес.

Бағалау әдістемесіне келсек, детекторлар PASCAL VOC (Everingham және т.б., 2010: 303–338) мен COCO метрикалары бойынша салыстырылады; метрикаларды есептеудегі айырмашылықтар нәтижелерге әсер ететіні көрсетілген (Padilla және т.б., 2021). Оқыту мен тест арасындағы жақын көшірмелер CIFAR жиынында анықталып, олардың модельдер рейтингін бұрмалайтыны көрсетілді (Barz, Denzler, 2020). Қазақстанда теміржолдағы төтенше жағдайларға компьютерлік қолдау көрсету мәселесі осы журналда да қарастырылған (Shalabayeva, 2025: 90–104), бірақ кескіндер бойынша кедергілерді анықтау модельдерін бағалау әдістемесі онда зерттелмеген.

Әдебиеттен мынадай олқылық шығады: UAV-RSOD жиынында жарияланған нәтижелер (Rampriya және т.б., 2024) оқыту мен тест жиындарының тәуелсіздігі тексерілмеген бөлуге негізделген, ал сипаттамаға сәйкес кескіндер аугментациядан кейін бөлінген. Сондықтан келесі сұрақтар қойылды: RQ1 – жарияланған тест жиынының қандай үлесінің оқыту жиынында жақын көшірмесі бар? RQ2 – жақын көшірмелерді азайтатын блоктық бөлуде детектордың сапасы бірдей көлемдегі кездейсоқ бөлумен салыстырғанда қалай өзгереді? RQ3 – COCO-да оқытылған адам детекторлары жоғарыдан түсірілген адамды қосымша оқытусыз анықтай ала ма? Екі болжам тексерілді. H1: кездейсоқ бөлуде орташа дәлдік (mAP50) блоктық бөлумен салыстырғанда жоғары болады. H2: COCO-да оқытылған детекторлардың жоғарыдан түсірілген адамдарға арналған AP50 мәні 0,5-тен төмен.

Материалдар мен әдістер.

Деректер. UAV-RSOD жиынының кедергілерді анықтауға арналған бөлігі (V2) пайдаланылды (Rampriya және т.б., 2024; Rampriya, 2024; CC BY 4.0). Авторлардың сипаттамасы бойынша кескіндер DJI Phantom 4 PRO ҰҰА-мен Тамилнад штатындағы (Үндістан) пайдаланылмайтын жолда 10,5 м биіктіктен түсірілген бейнеден алынған; кедергілер жолға әдейі қойылған. Архивте 2002 кескін (1920×1080 немесе 1080×1920) және PASCAL VOC пішіміндегі белгілер бар; кескіндердің біреуінен басқасында бір ғана нысан белгіленген. Кластары: бұтақ, темір шыбық, бөшке, тас, канистр, адам. Сипаттама бойынша 315 бастапқы кескін бұру, айналы бейнелеу, масштабтау, қисайту (shear) және жарықтықты өзгерту арқылы көбейтілген және 80/20 қатынасында бөлінген (мақалада 1602/400, архивте

1611/391 кескін). Жүктелген архивтің MD5 бақылау сомасы Zenodo жазбасындағы мәнмен сәйкес келді (тексеру журналы кодпен бірге сақталған).

Жақын көшірмелерді анықтау (RQ1). Әр кескіннен ұзын жағы 960 пиксельге кішірейтілген сұр нұсқада SIFT кілттік нүктелері (keypoints) алынды, бір кескінге 800-ге дейін (Lowe, 2004: 91–110). Барлық оқыту кескіндерінің дескрипторлары бір FLANN индексіне жинақталды; тест кескінінің әр дескрипторы Лоу қатынасы сынағынан (0,75) өткен жағдайда өзі сәйкес келген оқыту кескініне дауыс берді. Ең көп дауыс алған бес кескін RANSAC әдісімен гомографияны бағалау арқылы (Fischler, Bolles, 1981: 381–395; проекция қатесінің шегі 5 пиксель) геометриялық тексеруден өткізілді; айналы аугментацияны ескеру үшін тест кескіні айналы түрде де тексерілді. Геометриялық тексеруден өткен сәйкестіктер (inliers) саны 40-тан кем болмаса, жұп жақын көшірме деп саналды; 20 және 100 шектері сезімталдықты тексеру үшін қолданылды. Жұптағы кескіндердің кластары бірдей болса, жақын көшірме бір кластағы деп саналды: бұл бір жерге қойылған бір кедергінің сахнасы болуы ықтимал, ал әртүрлі кластағы жақын көшірме жолдың сол жерін басқа кедергімен көрсетеді. Тексерілген кандидаттар саны шектеулі (5) болғандықтан, бұл үлестер төменгі баға болып табылады. Тексеру үш бөлудің әрқайсысында бөлек жүргізілді.

Блоктық бөлу (RQ2). Барлық 2002 кескін бір графқа біріктірілді: түйін – кескін, қабырға – тексерілген сәйкестіктер саны. Графты құру үшін әр кескін (түпнұсқасы және айналы нұсқасы) барлық басқа кескіндердің дескрипторлары бойынша сұралды, өз дескрипторлары алынып тасталды, кемінде 4 дауыс алған ең көп 8 кандидат геометриялық тұрғыда тексерілді. Байланысты компоненттер (connected components) бойынша бөлу мүмкін болмады, себебі компоненттердің біреуі кескіндердің басым бөлігін қамтиды (Нәтижелер бөлімін қараңыз). Сондықтан блоктық бөлу қолданылды: әр класс ішінде кескіндер графтың салмақтары бойынша спектрлік кластерлеумен бес топқа бөлінді; класс кескіндерінің 10–35%-ын қамтитын топтардың ішінен өз класындағы буфер көлемінің топ көлеміне қатынасы ең аз топ тест жиынына берілді. Графта тест кескінімен кемінде 40 сәйкестігі бар кез келген кескін (буфер) оқытудан алынып тасталды. Салыстыру үшін дәл осындай көлемдегі (1000/458) кездейсоқ бөлу құрылды. Екі бөлудің көлемі бірдей, бірақ тест және оқыту жиындарының кластық құрамы мен сахналары әртүрлі (Кесте 2), сондықтан олардың айырмасы тек ағып кетумен түсіндірілмейді.

Детектор және оқыту. COCO-да алдын ала оқытылған YOLOv8n моделі (Jocher және т.б., 2023; Ultralytics 8.4.175 кітапханасы) екі бөлуде де бірдей параметрлермен қосымша оқытылды: кескін өлшемі 416 пиксель, 15 дәуір, бума (batch) өлшемі 16, AdamW оңтайландырғышы (кітапхананың автоматты таңдауы), кітапхананың әдепкі аугментациялары, кездейсоқ сандар генераторының бастапқы мәні 0. Есептеу ресурсын үнемдеу үшін кескіндер алдын ала ұзын жағы 640 пиксельге кішірейтілді. Бөлек валидация жиыны болмағандықтан, нүктелерді тест жиыны бойынша таңдамау үшін бағалауға соңғы (15-) дәуірдің салмақтары алынды. Сапа тест жиынында mAP50 және mAP50:95 метрикаларымен, сондай-ақ әр класс бойынша AP50 мәнімен бағаланды. Есептеу графикалық процессорсыз, 2 виртуалды ядролы (2 vCPU) Intel Xeon 2,80 ГГц процессорында орындалды.

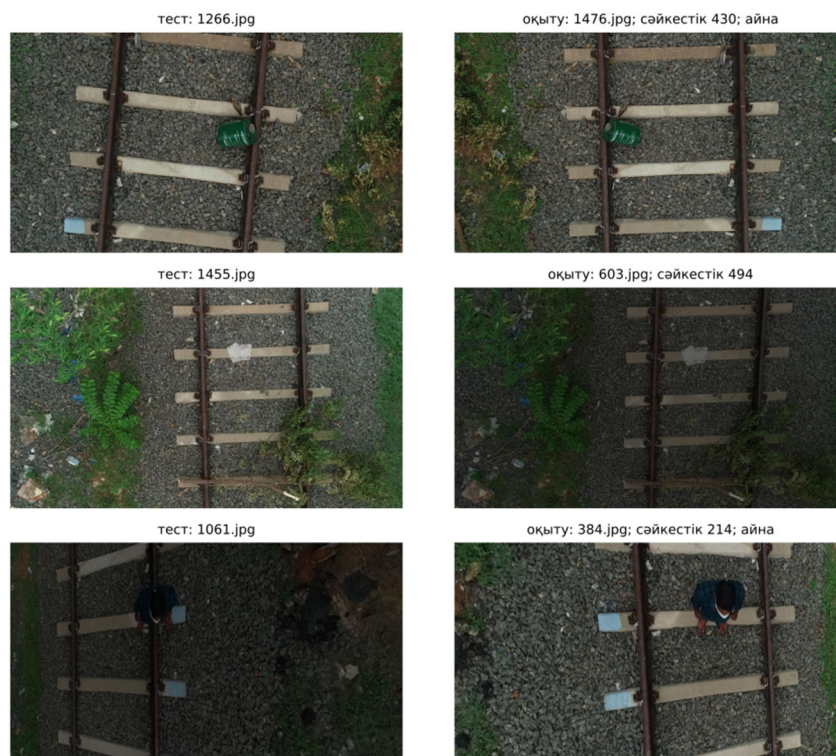
COCO детекторларын қосымша оқытусыз тексеру (RQ3). YOLOX тобының ресми ONNX модельдері (nano, tiny, s, m) (Ge және т.б., 2021) барлық кескіндерде COCO «person» класы бойынша іске қосылды. Модельдер бұл жиында оқытылмағандықтан, бұл бағалауда ағып кету болмайды, бірақ адам кескіндері аз ғана бастапқы кадрлардың аугментацияланған нұсқалары, сондықтан олар өзара тәуелсіз емес. AP50 мәні PASCAL VOC әдістемесі бойынша барлық нүктелік (all-point) интерполяция арқылы, IoU $\geq 0,5$ шартымен есептелді (Everingham және т.б., 2010: 303–338); қосымша 0,3 сенімділік шегіндегі толықтық пен жалған анықтаулар саны келтірілді.

Нәтижелер.**Кесте 1 — Бөлулердегі тест кескіндерінің оқыту кескіндерімен жақын көшірмелері**

Көрсеткіш	Жарияланған бөлу	Кездейсоқ бөлу	Блоктық бөлу
Оқыту / тест кескіндері	1611 / 391	1000 / 458	1000 / 458 (буфер 544)
Оқытуда кез келген кластағы жақын көшірмесі бар тест кескіндері, ≥ 40 сәйкестік	390/391 (99,7%)	458/458 (100,0%)	412/458 (90,0%)
Оқытуда бір кластағы жақын көшірмесі бар тест кескіндері, ≥ 40 сәйкестік	386/391 (98,7%)	457/458 (99,8%)	297/458 (64,8%)
Сол үлес, ≥ 100 сәйкестік	381/391 (97,4%)	455/458 (99,3%)	241/458 (52,6%)
Бір кластағы ең жақын оқыту кескінімен сәйкестіктер медианасы	456	447	104

Дереккөз: автордың есептеулері. Әр бөлуде тест кескіндері тек сол бөлудің оқыту жиынымен салыстырылды. Сәйкестік – RANSAC тексеруінен өткен SIFT нүктелерінің жұбы. Блоктық бөлуді құрған біріктірілген граф (кандидаттар саны шектеулі) жарияланған бөлуде 95,9 % үлесті табады, яғни бұл кестедегі арнайы салыстырудан біршама аз.

RQ1 (ашық деректерді талдау). Жарияланған тест жиынының 391 кескінінің 390-інің (99,7 %) оқыту жиынында RANSAC арқылы тексерілген жақын көшірмесі бар; сәйкестіктердің медианасы – 458, ал 42,7 % жағдайда ең жақын жұп айналы бейнеде табылды. 100 сәйкестік шегінде де үлес 98,0 % болды (Кесте 1). Сурет 1-дегі мысалдарда тест кескіні оқыту кескінінің айналы бейнеленген, жарықтығы немесе масштабы өзгертілген нұсқасы; басқа жұптардың бір бөлігі бейнедегі көршілес кадрлар болуы мүмкін, бұл бөлек тексерілмеді. Демек, жарияланған тест жиыны модельдің жаңа жол учаскесіндегі сапасын емес, негізінен бұрын көрген сахнаны тануын өлшейді.



Сур. 1. Жарияланған тест кескіндері (сол жақта) және олардың оқыту жиынындағы жақын көшірмелері (оң жақта)

Дереккөз: UAV-RSOD кескіндері (Rampriya және т.б., 2024; CC BY 4.0). Жұптар алгоритм бойынша таңдалды: айналы және айналы емес жұптардың медианасы және сәйкестік санының 10-процентілі.

Біріктірілген графта 40 сәйкестік шегінде 10 компонент шықты, олардың ең үлкені 1949 кескінді қамтиды; шекті 200-ге дейін көтергенде де ең үлкен компонентте 1867 кескін қалды. Бұл жиынның бір жол учаскесінің бір-бірімен жалғасқан сахналарынан тұратынымен сәйкес келеді. Блоктық бөлуде тест жиынына 458 кескін берілді, буфер ретінде 544 кескін алынып тасталды, оқытуда 1000 кескін қалды. Графтың өзінде тест пен оқыту арасында 40 сәйкестіктен асатын жұп қалмады, бірақ арнайы салыстыру жақын көшірмелердің толық жойылмағанын көрсетті (Кесте 1): бір кластағы жақын көшірмесі бар тест кескіндерінің үлесі бірдей көлемдегі кездейсоқ бөлудегі 99,8 %-дан блоктық бөлуде 64,8 %-ға дейін төмендеді, ал бір кластағы сәйкестіктердің медианасы 447-ден 104-ке дейін азайды. Графты құру кезінде әр кескін үшін тек 8 кандидат тексерілгендіктен, оның бір бөлігі буферден тыс қалды; ал барлық кескіндер бір учаскеден алынғандықтан, жолдың сол жерінің басқа кедергімен түсірілген кескіндері (әртүрлі кластағы жақын көшірмелер) кез келген бөлуде қалады. Сондықтан блоктық бөлу ағып кетуді азайтады, бірақ жоймайды; адам класында тест кескіндері оқыту жиынынан толық бөлек болды.

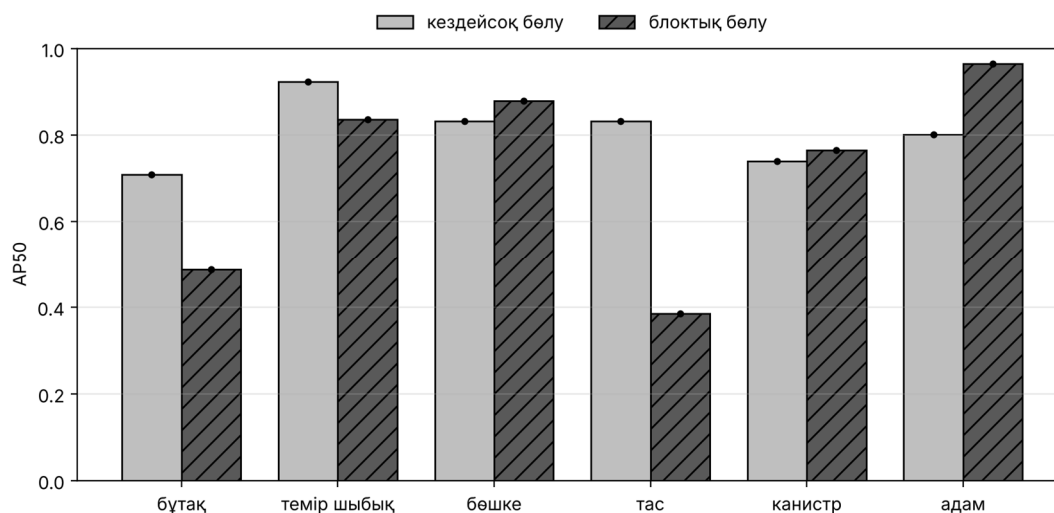
Кесте 2 — YOLOv8n детекторының кездейсоқ және блоктық бөлулердегі сапасы

Класс	Тест: кездейсоқ / блоктық	AP50, кездейсоқ бөлу	AP50, блоктық бөлу	Айырма
бұтақ	115 / 169	0,708	0,487	0,221
темір шыбық	83 / 98	0,924	0,836	0,088
бөшке	64 / 81	0,832	0,878	-0,046
тас	47 / 24	0,832	0,385	0,447
канистр	87 / 42	0,739	0,764	-0,026
адам	62 / 44	0,801	0,965	-0,164
Барлық класс, mAP50	458 / 458	0,806	0,719	0,087
Барлық класс, mAP50:95		0,524	0,468	0,056

Дереккөз: автордың есептеулері; бір рет оқыту (бастапқы мән 0), соңғы дәуірдің салмақтары. Екі бөлудің оқыту және тест жиындарының көлемі бірдей (1000/458), кластық құрамы әртүрлі.

RQ2 (есептеу эксперименті). Кездейсоқ бөлуде YOLOv8n mAP50 = 0,806, ал блоктық бөлуде mAP50 = 0,719 нәтижесін көрсетті (Кесте 2); айырма 0,087, mAP50:95 бойынша 0,056. Әсер кластар бойынша біркелкі емес (Сурет 2): AP50 мәні «бұтақ» (0,708 → 0,487), «темір шыбық» (0,924 → 0,836), «тас» (0,832 → 0,385) кластарында төмендеді, ал «бөшке» (0,832 → 0,878), «канистр» (0,739 → 0,764), «адам» (0,801 → 0,965) кластарында блоктық бөлуде жоғары болды. Айырманың негізгі бөлігі «тас» класына тиесілі (блоктық тест жиынында 24 кескін): бұл класты есептемегенде орташа айырма 0,014. Кейбір кластарда блоктық бөлудегі мәннің жоғары болуы тест және оқыту жиындарының құрамындағы айырмашылықпен байланысты болуы мүмкін; бұл тексерілмеді. Бір рет оқытуда нәтиже H1-ге сәйкес келді, бірақ оқытулар арасындағы ауытқу бағаланбағандықтан H1 түпкілікті тексерілген жоқ.

RQ3 (есептеу эксперименті, қосымша оқытусыз). COCO детекторларының ең жақсысы – YOLOX-m – адам класы бойынша AP50 = 0,455 нәтижесін берді, 0,3 сенімділік шегіндегі толықтық – 0,244 (Кесте 3). Төрт модельдің бәрінде AP50 < 0,5, яғни нәтиже H2-ге сәйкес келді: бұл деректерде мұндай детекторлардың қосымша оқытусыз сапасы төмен. Мүмкін себебі – жоғарыдан түсірілген адамның көрінісі (бас пен иық) COCO-дағы адам кескіндерінен өзгеше; бұл бөлек тексерілмеді.



Сур. 2. Кластар бойынша AP50: кездейсоқ және блоктық бөлу

Дереккөз: автордың есептеулері; бір рет оқыту (бастапқы мән 0).

Кесте 3 — СОСО-да оқытылған YOLOX детекторларының жоғарыдан түсірілген адамдарды қосымша оқытусыз анықтауы

YOLOX моделі	AP50 (адам)	Толықтық, сенімділік $\geq 0,3$	Жалған анықтаулар, сенімділік $\geq 0,3$	Жалған анықтауы бар адамсыз кескіндер (1768-ден)
nano	0,352	0,227	27	5
tiny	0,328	0,094	11	1
s	0,440	0,227	17	2
m	0,455	0,244	23	6

Дереккөз: автордың есептеулері; барлық 2002 кескін, адам белгілері – 234.

.Талқылау.

Нәтижелер UAV-RSOD бойынша жарияланған жоғары дәлдікті (Rampriya және т.б., 2024) жаңа жол учаскесіндегі сапаның бағасы ретінде қабылдауға болмайтынын көрсетеді: жарияланған тест кескіндерінің 99,7 %-ының оқыту жиынында жақын көшірмесі бар. Бұл CIFAR жиынында сипатталған жағдайға ұқсас (Barz, Denzler, 2020), бірақ мұнда жақын көшірмелер үлесі әлдеқайда жоғары; мүмкін себебі – аугментацияның бөлуге дейін жасалуы. Ағып кетудің тағы бір көзі – кескіндердің бір бейнеден және бір учаскеден алынуы: аугментациясыз кездейсоқ бөлу де көршілес кадрларды екі жиынға таратуы мүмкін.

Практикалық салдары: өнеркәсіптік кірме жолдарға ҰҰА-мен тексеру жүйесін енгізуде детектор сапасын сол учаскелер оқытуға кірмеген деректерде, яғни учаскелер, түсіру күндері немесе маршруттар бойынша бөлінген жиында бағалау қажет. Ұсынылған тәсілді – SIFT пен RANSAC арқылы тексерілген жақын көшірмелер графы, буфер және бөлінген жиындарды арнайы қайта тексеру – басқа кескіндер жиындарына да қолдануға болады, ол тек ашық кітапханаларды қажет етеді; бір учаскеден алынған деректерде ол ағып кетуді тек азайтады. Блоктық бөлуде алынған mAP50 = 0,719 мәні жарияланған бөлудегі нәтижелерге қарағанда шынайырақ бағдар болуы мүмкін, бірақ ол ағып кетуі азайған, бірақ жойылмаған бөлуде, бір учаскеде, бір рет оқыту нәтижесінде және кластық құрамы әртүрлі тест жиындарында алынды.

Зерттеудің шектеулері. Біріншіден, авторлардың сипаттамасы бойынша барлық кескіндер бір учаскеде, бір түсірілімде (15.09.2020), әдейі қойылған кедергілермен түсірілген; блоктық бөлу де сол учаскенің ішінде, сондықтан басқа жолдарға, маусымдарға және Қазақстан жағдайына (қар, шаң, бетон шпалдар) көшіру тексерілмеді. Екіншіден,

жиында кедергісіз кескіндер жоқ, сондықтан жүйе деңгейіндегі жалған дабылдар жиілігін бағалау мүмкін болмады. Үшіншіден, есептеу ресурсына байланысты детектор 416 пиксельде, 15 дәуір және бір бастапқы мәнмен оқытылды; оқыту қисықтары 15-дәуірде әлі өсіп келе жатқан, сондықтан ұзағырақ оқыту абсолютті мәндерді өзгертуі мүмкін. Төртіншіден, блоктық бөлу жақын көшірмелерді азайтты, бірақ жоймады (Кесте 1); буферді арнайы салыстыру бойынша кеңейту оқыту жиынын едәуір қысқартады және келесі кезеңнің міндеті болып қалады. Бесіншіден, бастапқы жоспарланған RailGoerl24 жиыны жұмыс кезінде жүктеп алуға қолжетімсіз болды (сервер HTTP 503 қатесін қайтарды).

Қорытынды.

Зерттеу мақсатына қол жеткізілді: ашық UAV-RSOD жиынындағы жарияланған бөлудің ағып кетуі өлшенді, жақын көшірмелерді азайтатын блоктық бөлу құрылды және детектор сапасы бірдей көлемдегі екі бөлуде салыстырылды; сонымен қатар COCO-да оқытылған дайын адам детекторлары қосымша оқытусыз тексерілді. Барлық есептеулер ашық деректерде және ашық бағдарламалық кітапханалармен орындалды, ал код пен бөлулер тізімдері нәтижелерді қайталауға мүмкіндік береді.

Алынған нәтижелер. Біріншіден (ашық деректерді талдау), жарияланған тест кескіндерінің 99,7 %-ының оқыту жиынында RANSAC арқылы тексерілген жақын көшірмесі бар, ал бір кластағы жақын көшірмесі барлардың үлесі 98,7 %; демек жарияланған дәлдік көрсеткіштері негізінен бұрын көрген сахналарды тануды сипаттайды. Екіншіден, бүкіл жиын бір учаскенің бір-бірімен жалғасқан сахналарынан тұрады: жақын көшірмелер графында бір компонент 1949 кескінді қамтиды, сондықтан сахналар бойынша толық тәуелсіз бөлу бұл деректерде мүмкін болмады. Блоктық бөлу бір кластағы жақын көшірмелер үлесін бірдей көлемдегі кездейсоқ бөлумен салыстырғанда 99,8 %-дан 64,8 %-ға дейін азайтты, бірақ толық жоймады. Үшіншіден (есептеу эксперименті, бір рет оқыту), YOLOv8n детекторының mAP50 мәні кездейсоқ бөлуде 0,806, ал блоктық бөлуде 0,719 болды; айырманың негізгі бөлігі «тас» класына тиесілі, ал кейбір кластарда блоктық бөлудегі мән жоғары болды, сондықтан H1 болжамы түпкілікті тексерілген жоқ. Төртіншіден (есептеу эксперименті, қосымша оқытусыз), COCO-да оқытылған YOLOX детекторлары жоғарыдан түсірілген адамдар үшін AP50 $\leq 0,455$ нәтижесін берді, ал 0,3 сенімділік шегіндегі толықтық 0,244 болды; бұл H2 болжамына сәйкес келеді.

Әдістемелік қорытынды. Бөлудің сапасын оны құрған құралмен ғана емес, тәуелсіз арнайы салыстырумен тексеру қажет: блоктық бөлуді құрған графта тест пен оқыту арасында 40 сәйкестіктен асатын жұп қалмаса да, арнайы салыстыру кез келген кластағы жақын көшірмелер үлесін 412/458 (90,0 %) деп көрсетті. Сондай-ақ бір кластағы және әртүрлі кластағы жақын көшірмелерді бөлек есептеу маңызды: біріншісі бір кедергінің сахнасының қайталануын, екіншісі жолдың сол жерінің басқа кедергімен түсірілуін көрсетеді, ал олардың модель сапасына әсері әртүрлі болуы мүмкін.

Ғылыми жаңалығы – ҰҰА-мен жолды тексеру деректеріндегі ағып кетудің сандық бағасы (кез келген кластағы және бір кластағы жақын көшірмелер бойынша бөлек) және оны азайтудың қайталанатын тәсілі: SIFT пен RANSAC арқылы тексерілген жақын көшірмелер графы, класс ішіндегі спектрлік кластерлеу, буфер және бөлінген жиындарды арнайы қайта тексеру. Бұл тәсіл бір учаскеден алынған деректерде ағып кетуді толық жоя алмайтыны да көрсетілді, яғни деректер жиынын жоспарлау кезеңінде түсіру учаскелерін алдын ала бөлу қажет.

Практикалық маңызы. Өнеркәсіптік кірме жолдар мен станция жолдарын ҰҰА-мен тексеру жүйесін сатып алу немесе әзірлеу кезінде тапсырыс беруші детектордың сапасын оқытуға кірмеген учаскелерде, басқа түсіру күндерінде және басқа маусымда алынған деректерде көрсетуді талап етуі қажет, ал әзірлеуші деректер жиынын бөлуге дейін жақын көшірмелерге тексеруі керек. Дайын COCO детекторларын жоғарыдан түсірілген адамдарды анықтау үшін қосымша оқытусыз пайдалану бұл деректерде төмен сапа берді, сондықтан адамдарды анықтау міндеті үшін ҰҰА көрінісіндегі белгіленген деректер қажет.

Шектеулер мен болашақ жұмыс. Нәтижелер бір учаскеде, бір түсірілімде, әдейі қойылған кедергілермен алынған; жиында кедергісіз кескіндер жоқ, сондықтан жалған дабылдар жиілігі бағаланбады; детектор шектеулі есептеу ресурсында оқытылды. Келесі кезеңде бірнеше учаскеден және әртүрлі маусымдардан, оның ішінде Қазақстанның қысқы жағдайынан алынған деректер жинау, кедергісіз кескіндерді қосу, оқытуды бірнеше бастапқы мәнмен қайталау және RailGoerl24 жиыны қолжетімді болғанда нәтижені борттық камера деректерінде тексеру жоспарланады.

REFERENCES

- Akyon F.C., Onur Altinuc S., Temizel A. (2022). Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection // 2022 IEEE International Conference on Image Processing (ICIP). Pp. 966–970. <https://doi.org/10.1109/ICIP46576.2022.9897990>
- Barz B., Denzler J. (2020). Do We Train on Test Data? Purging CIFAR of Near-Duplicates // Journal of Imaging. Vol. 6. No. 6. — Art. 41. <https://doi.org/10.3390/jimaging6060041>
- Du D., Zhu P., Wen L., Bian X., Lin H., Hu Q. et al. (2019). VisDrone-DET2019: The Vision Meets Drone Object Detection in Image Challenge Results // 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW).— Pp. 213–226. <https://doi.org/10.1109/ICCVW.2019.00030>
- Everingham M., Van Gool L., Williams C.K.I., Winn J., Zisserman A. (2010). The Pascal Visual Object Classes (VOC) Challenge // International Journal of Computer Vision. Vol. 88, No. 2. — Pp. 303–338. <https://doi.org/10.1007/s11263-009-0275-4>
- Fischler M.A., Bolles R.C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography // Communications of the ACM. Vol. 24. No. 6. — Pp. 381–395. <https://doi.org/10.1145/358669.358692>
- Ge Z., Liu S., Wang F., Li Z., Sun J. (2021). YOLOX: Exceeding YOLO Series in 2021 [Preprint] // arXiv. arXiv:2107.08430. <https://doi.org/10.48550/arXiv.2107.08430>
- Jocher G., Chaurasia A., Qiu J. (2023). Ultralytics YOLOv8 [Software]. — Version 8.0.0. (used: ultralytics 8.4.175). — URL: <https://github.com/ultralytics/ultralytics> (accessed: 10.10.2026).
- Kapoor S., Narayanan A. (2023) Leakage and the reproducibility crisis in machine-learning-based science // Patterns. Vol. 4. No. 9. — Art. 100804. <https://doi.org/10.1016/j.patter.2023.100804>
- Lin T.-Y., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollár P., Zitnick C.L. (2014). Microsoft COCO: Common Objects in Context // Computer Vision — ECCV 2014. Lecture Notes in Computer Science. Vol. 8693. — Pp. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48
- Lowe D.G. (2004). Distinctive Image Features from Scale-Invariant Keypoints // International Journal of Computer Vision. Vol. 60. No. 2. — Pp. 91–110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- Padilla R., Passos W.L., Dias T.L.B., Netto S.L., da Silva E.A.B. (2021). A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit // Electronics. Vol. 10, No. 3. — Art. 279. <https://doi.org/10.3390/electronics10030279>
- Rampriya R.S., Al-Shehari T., Nathan S., Jenefa A., Suganya R., Shunmuga Perumal P., Alfakih T., Als Salman H. (2024). An unmanned aerial vehicle captured dataset for railroad segmentation and obstacle detection // Scientific Data. Vol. 11. — Art. 1315. <https://doi.org/10.1038/s41597-024-03952-3>
- Rampriya R.S. (2024). UAV Railroad Images for Segmentation and Obstacle Detection [Dataset]. — Version 1. — Zenodo, <https://doi.org/10.5281/zenodo.12606374>
- Redmon J., Divvala S., Girshick R., Farhadi A. (2016). You Only Look Once: Unified, Real-Time Object Detection // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Pp. 779–788. <https://doi.org/10.1109/CVPR.2016.91>
- Ristić-Durrant D., Franke M., Michels K. (2021). A Review of Vision-Based On-Board Obstacle Detection and Distance Estimation in Railways // Sensors. Vol. 21. No. 10. — Art. 3452. <https://doi.org/10.3390/s21103452>
- Shalabayeva M. (2025). Computer support for responding to railway emergencies // Industrial Transport of Kazakhstan. Vol. 22, No. 2(86). — Pp. 90–104. <https://doi.org/10.58420/ptk/2025.86.02.008>
- Tagiew R., Klasek P., Tilly R., Köppel M., Denzler P., Neumaier P., Klockau T., Boekhoff M., Schwalbe K. (2023). OSDaR23: Open Sensor Data for Rail 2023 // 2023 8th International Conference on Robotics and Automation Engineering (ICRAE). Pp. 270–276. <https://doi.org/10.1109/ICRAE59816.2023.10458449>
- Tagiew R., Wunderlich I., Sastuba M., Seitz S. (2025). RailGoerl24: Görlitz Rail Test Center CV Dataset 2024 [Preprint] // arXiv. arXiv:2504.00204. <https://doi.org/10.48550/arXiv.2504.00204>
- Zendel O., Murschitz M., Zeilinger M., Steininger D., Abbasi S., Beleznaï C. (2019). RailSem19: A Dataset for Semantic Rail Scene Understanding // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). — Pp. 1221–1229. <https://doi.org/10.1109/CVPRW.2019.00161>