REAL VAQTDA AVTOTRANSPORT VOSITALARINI ANIQLASHNING ZAMONAVIY MODELLAR TAHLILI

Babadjanov Elmurod Satimbaevich[1], Maxamatdinov Abdul-Aziz Karamatdinovich[2],

[1]Nukus davlat texnika universiteti, Kiberxavfsizlik injiniringi kafedrasi mudiri, t.f.d. (DSc), dotsent, Nukus, O'zbekiston,
[2]Toshkent davlat iqtisodiyot universiteti, Sun'iy intellekt kafedrasi tayanch doktoranti, Toshkent, O'zbekiston,
E-mail: a.makhamatdinov@tsue.uz

K E Y W O R D S

obyektlarni aniqlash, intellektual transport tizimlari, chuqur o'qitish, edge qurilma, yo'l harakati monitoringi, kompyuter ko'rishi, transformer detektorlari

A B S T R A C T

Ushbu tadqiqotda real vaqt rejimi video oqimida avtotransport vositalarini aniqlashga mo'ljallangan YOLO26n, YOLOv11n va RT-DETR arxitekturalarining qiyosiy hamda eksperimental tahlili taqdim etilgan. Tadqiqot O'zbekiston Respublikasi sharoitida yo'l monitoringi vazifalariga nisbatan zamonaviy detektorlarning qiyosiy samaradorligi yetarli darajada o'rganilmaganligi bilan bog'liq dolzarb ilmiy bo'shliqni bartaraf etishga qaratilgan. Eksperimentni o'tkazish uchun O'zbekiston yo'l harakati qoidabuzarliklarini video qayd etish kameralaridan olingan yo'l sahnalarining 1475 ta annotatsiyalangan tasvirlarini o'z ichiga olgan original Uzcardet ma'lumotlar to'plami shakllantirildi. Uchta modelni o'qitish yagona eksperimental muhitda, birlashtirilgan giperparametrlar asosida, Google Colaboratory platformasida NVIDIA Tesla T4 GPU yordamida amalga oshirildi. Aniqlik ko'rsatkichlari bo'yicha RT-DETR modeli eng yaxshi natijalarni ko'rsatdi (mAP50-95 = 0.634, F1 = 0.719), YOLO26n va YOLOv11n esa nisbatan muvozanatli ko'rsatkichlarni ta'minladi. CPU qurilmasida (Intel Core i3-1115G4) o'tkazilgan inference-eksperiment tezlik tavsiflarida prinsipial farq mavjudligini ko'rsatdi: YOLO26n va YOLOv11n mos ravishda 35.2 va 35.7 FPS ga erishgan bo'lsa, RT-DETR atigi 2.7 FPS ni ko'rsatdi, bu esa YOLO modellaridan 13 barobar pastdir. Aniqlik va tezlik mezonlari yig'indisiga ko'ra, GPU-tezlatgichsiz yo'l kuzatuvi uchun Edge-qurilmalarda joriy etish nuqtai nazaridan eng maqbul arxitektura sifatida YOLOv11n modeli e'tirof etildi.

1. KIRISH

1.1. Masalaning dolzarbligi

Avtotransport vositalarini video oqimda real vaqt rejimida aniqlash intellektual transport tizimlarining (ITT) asosiy vazifalaridan biridir [1]. Ushbu sohadagi markaziy ilmiy-texnik muammo obyektlarni lokalizatsiya qilish va klassifikatsiyalashning yuqori aniqligi bilan uzluksiz video oqimni qayta ishlashdagi minimal hisoblash kechikishi o'rtasidagi qat'iy muvozanatga erishishdan iborat.

Avtotransport vositalarini aniqlashning zamonaviy yuqori aniqlikdagi modellari yuqori hisoblash murakkabligi bilan tavsiflanadi, bu esa ularning yo'l kameralari kabi Edge qurilmalarida qo'llanilishini sezilarli darajada cheklaydi [2, 3]. Transport oqimlarini boshqarish va yo'l-transport hodisalarining oldini olish vazifalarida ma'lumotlarni qayta ishlashdagi minimal kechikish muhim ahamiyatga ega [1, 4]. Shu munosabat bilan, aniqlik va ishlash tezligi o'rtasidagi optimal nisbatni topishga qaratilgan zamonaviy, avtotransport vositalarini aniqlash modellarining qiyosiy tahliliga ehtiyoj yuzaga keladi.

Modelning real ekspluatatsiyaga yaroqliligini to'g'ri baholash uchun ikki guruh metrikalarni birgalikda tahlil qilish talab etiladi [20]. Birinchi guruh aniqlash sifatini tavsiflaydi va Precision, Recall hamda mAP ko'rsatkichlarini o'z ichiga oladi [20, 21]. Precision barcha ishga tushgan holatlar ichida to'g'ri bashoratlarning ulushini aks ettiradi va quyidagicha hisoblanadi:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

bu yerda TP (True positive) – haqiqiy musbat aniqlashlar, FP (False positive) – yolg'on musbat aniqlashlardir. Masalan, agar model 80 ta kadrda obyektlarni aniqlagan bo'lib, ulardan 8 tasi yolg'on aniqlash bo'lsa, u holda Precision =

$72/(72+8) = 0.90$ bo‘ladi. Ushbu metrika yolg‘on aniqlash narxi yuqori bo‘lgan ssenariylarda muhim ahamiyat kasb etadi.

Recall kadrdagi real obyektlarni qamrab olish to‘liqligini tavsiflaydi:

$$Recall = \frac{TP}{TP+FN} \qquad (2)$$

bu yerda FN (False Negative) – yolg‘on manfiy o‘tkazib yuborishlardir. Agar 88 ta real obyektdan model 8 tasini o‘tkazib yuborgan bo‘lsa, u holda Recall $= 80/(80+8) \approx 0.91$ bo‘ladi. Yuqori Recall buzilishlarni o‘tkazib yuborish mumkin bo‘lmagan tizimlar uchun prinsipial ahamiyatga ega.

Sifatning integral ko‘rsatkichi sifatida barcha maqsadli klasslar bo‘yicha Precision va Recall ni birlashtiruvchi mAP (mean Average Precision) xizmat qiladi [20, 21]:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \qquad (3)$$

bu yerda N – klasslar soni, AP_i – i-klass uchun Precision-Recall egri chizig‘i ostidagi maydon.

Ikkinchi guruh tizimning tezlik xususiyatlarini ifodalaydi: latency (bitta kadrni qayta ishlash kechikishi, ms) va throughput (o‘tkazuvchanlik qobiliyati, sekundiga kadrlar - FPS) [2, 9].

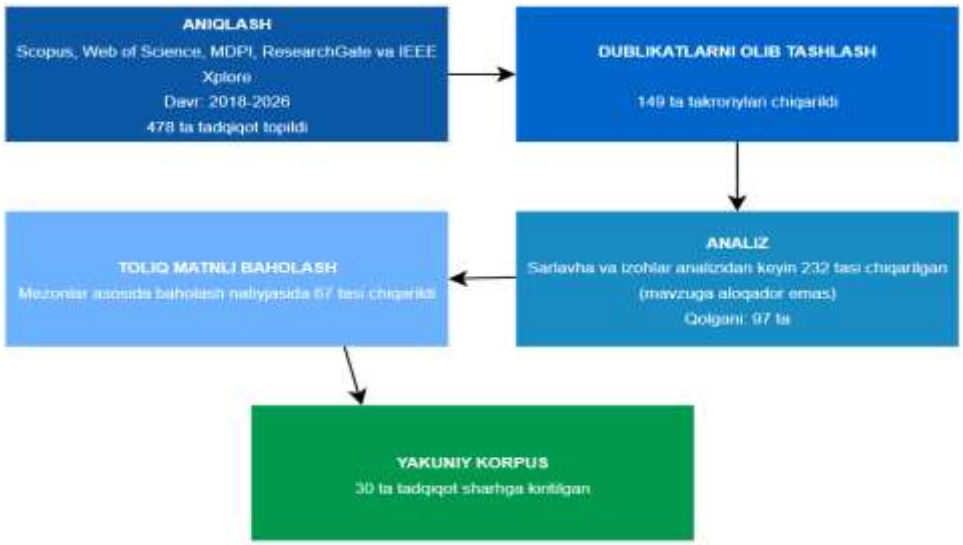
$$Latency = \frac{1}{N}\sum_{i=1}^{N} t_i \; (ms/kadr) \qquad (4)$$

$$FPS = \frac{1000}{Latency} \qquad (5)$$

Faqat ushbu ikki guruh metrikalarini birgalikda tahlil qilish modelning real vaqt rejimida foydalanishga yaroqliligini to‘g‘ri baholash imkonini beradi.

1.2. Mavjud tadqiqotlar tahlili

Avtotransport vositalarini aniqlash bo‘yicha tadqiqot JBI (Joanna Briggs Institute) [25] metodologiyasiga muvofiq scoping review formatida olib borildi, hisobot esa PRISMA-ScR talablariga muvofiq tuzildi [26]. Qidiruv Scopus, Web of Science, MDPI, ResearchGate va IEEE Xplore ma’lumotlar bazalarida quyidagi so‘rovlar bo‘yicha amalga oshirildi: “real-time object detection”, “vehicle detection”, “deep learning object detection”, “road object detection”, shuningdek, arxitekturalar nomlari bo‘yicha: “Faster R-CNN”, “SSD”, “RetinNet”, “EfficientDet”, “YOLO”, “RT-DETR”. Qidiruvning vaqt oralig‘i 2018-2026 - yillarni qamrab oldi. Deduplikatsiya, skrining hamda kiritish va chiqarish mezonlari bo‘yicha to‘liq matnli baholashdan so‘ng, yakuniy korpusga dastlab aniqlangan 478 ta nashrdan 30 tasi kiritildi. 1-rasmda nashrlarni tanlash jarayoni batafsil keltirilgan.



1-rasm. Nashrlarni tanlash sxemasi: Scopus, Web of Science, MDPI, ResearchGate, IEEE Xplore, 2018-2025

Tizimli tahlil natijalariga ko‘ra, transport vositalarini aniqlash usullari kompyuter ko‘rishida obyektlar deteksiyasining rivojlanishidagi izchil bosqichlarni aks ettiruvchi to‘rtta toifaga guruhlashtirildi. Natijada [1, 2] da klassik algoritmlar, [6, 7, 8, 9, 16] da ikki bosqichli detektorlar va bir bosqichli modellar hamda [18, 30, 19] da real vaqt rejimiga mo‘ljallangan zamonaviy arxitekturalar

belgilangan bo'lib, ularning guruhlarning qiyosiy tahlili 1-jadvalda keltirilgan.

1-jadval.

Transport vositalarini aniqlash usullari avlodlarining qiyosiy tahlili.

Usullar guruhi	**Modellar**	**Ishlash tamoyili**	**Afzalliklari**	**Cheklovlar**	**mAP (*MS COCO)**	**Real-time**
Klassik usullar	Orqa fon ayirmasi, HOG+SVM	Berilgan qoidalar bo‘yicha belgilarni qo‘lda ajratish	Oddiylik; o‘rgatuvchi ma'lumotlarni talab qilmaydi	Yorug‘lik va rakurs o‘zgarishlariga chidamliligi past	-	Past
Ikki bosqichli detektorlar	Faster R-CNN, Mask R-CNN	Dastlab nomzod sohalar ajratiladi, so‘ngra har bir soha alohida tasniflanadi	Lokalizatsiya va klassifikatsiyaning yuqori aniqligi	Yuqori hisoblash xarajatlari, 100-200 ms kechikish, Edge qurilmalari uchun yaroqsiz	37-42	Cheklangan
Bir bosqichli ilk avlod detektorlari	SSD, RetinaNet, YOLOv1-YOLOv8	Lokolizatsiya va klassifikatsiya oldindan belgilangan ramkalar asosida tarmoqning bir o‘tishida amalga oshiriladi.	Tezligi 30-80 FPS va ikki bosqichlidan tezroq	Kichik o‘lchamli obyektlarda aniqlikning pasayishi, oldindan belgilangan ramkalar parametrlariga bog‘liqlik	25-42	O‘rtacha
Zamonaviy real vaqt modellari	YOLOv11, YOLO26, RT-DETR	Tarmoqning oldindan belgilangan ramkalarsiz bitta o‘tishi; tasvirning global kontekstini hisobga olishi	Bir vaqtning o‘zida yuqori aniqlik va tezlik, kichik obyektlarni deteksiyalashning yaxshilanishi	Yo‘l monitoringi sharoitida qiyosiy samaradorlik yetarlicha o‘rganilmagan.	50+	Yuqori

*** MS COCO (Microsoft Common Objects in Context) - bu ma'lumotlar to‘plami.**

Taqdim etilgan ma'lumotlardan ko‘rinib turibdiki, so‘nggi avlod modellari va transformer arxitekturalari inference(xulosa)ning yuqori tezligini deteksiya aniqligi bilan uyg‘unlashtiradi [30, 16, 18, 19]. Shu bilan birga, ularning real sharoitlardagi yo‘l kuzatuvi vazifalariga tatbiqan qiyosiy samaradorligi yetarli darajada o‘rganilmagan bo‘lib, bu mazkur ish doirasida bartaraf etilayotgan tadqiqot bo‘shlig‘ini belgilaydi.

1.3. Tadqiqotning maqsadi va vazifalari

O‘tkazilgan tahlil asosida tadqiqotning maqsadi qilib avtotransport vositalarini real vaqt

rejimida aniqlashning zamonaviy modellarini qiyosiy va eksperimental tahlil qilish belgilandi. Maqsadga erishish uchun quyidagi vazifalar qo‘yildi:

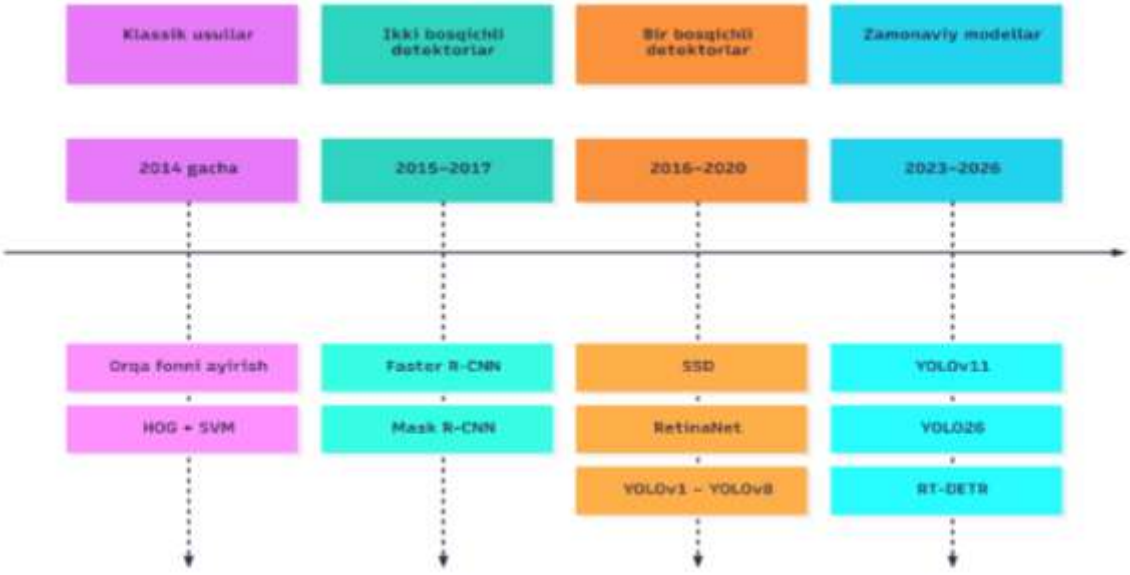
1. Avtotransport vositalarini aniqlashga oid mavjud yondashuvlarning tizimli tahlilini o‘tkazish va ularni arxitektura xususiyatlari bo‘yicha tasniflash.
2. Aniqlik va inference tezligi mezonlari asosida qiyosiy tahlil uchun modellar tanlash.
3. Maqsadli klasslarga kiruvchi transport vositalarining annotatsiyalangan tasvirlari bilan yo‘l sahnalari datasetini shakllantirish.
4. Tanlangan modellarni yagona bir xil sharoitlarda eksperimental o‘qitish.
5. O‘qitilgan modellarni aniqlik va inference tezlik metrikalari bo‘yicha qiyosiy tahlil qilish.
6. Yo‘l harakati monitoringi vazifalari uchun optimal arxitekturani tanlash bo‘yicha tavsiyalarni shakllantirish.

Ishning tuzilishi quyidagicha tashkil etilgan. 2-bo‘lim adabiyotlarning tizimli tahlili hamda obyektlarni aniqlash usullarining evolyutsiyasiga bag‘ishlangan. 3-bo‘limda qiyosiy eksperiment metodikasi, jumladan dataset tavsifi, o‘qitish konfiguratsiyasi va baholash metrikalari tizimi bayon etiladi. 4-bo‘lim eksperiment natijalarini taqdim etadi, 5-bo‘lim esa ularning muhokamasini o‘z ichiga oladi. 6-bo‘limda xulosalar shakllantiriladi va keyingi tadqiqot yo‘nalishlari belgilanadi.

2. ADABIYOTLAR TAHLILI

2.1. Obyektlarni aniqlash usullarining evolyutsiyasi

Obyektlarni aniqlash usullari klassik deskriptorlardan chuqur konvolyutsion tarmoqlargacha bo‘lgan rivojlanish yo‘lini bosib o‘tdi [23]. Burilish nuqtasi sifatida R-CNN ning paydo bo‘lishi xizmat qildi [3], u dastlab region-nomzodlar generatsiya qilinadigan, so‘ngra esa ularning klassifikatsiyasi amalga oshiriladigan ikki bosqichli yondashuvga asos soldi. Yuqori aniqligiga qaramay, ikki bosqichli usullar real vaqt tizimlari uchun zarur bo‘lgan tezlikni ta’minlay olmadi [4]. Bunga muqobil sifatida ilk avlod bir bosqichli detektorlar taklif etildi [6, 7], ular obyektlarni lokalizatsiya qilish va klassifikatsiyalashni oldindan belgilangan ramkalar asosida tarmoqdan bir marta o‘tishda bajaradi, bu esa tezlikning sezilarli oshishini ta’minladi, biroq kichik o‘lchamli obyektlarni aniqlashda aniqlikning pasayishi bilan birga kechdi [8]. Usullarning keyingi rivojlanishi zamonaviy avlod modellarining paydo bo‘lishiga olib keldi [15, 16, 18, 19], ularda oldindan belgilangan ramkalardan voz kechish va tasvirning global kontekstini hisobga olish aniqlikni oshirish hamda qayta ishlashning yuqori tezligini saqlab qolish imkonini berdi. Obyektlarni aniqlash bo’yicha modellar soni va turlari juda ko’p taklif etilgan, ushbu ishda avtotransport vositalarini qarayotganligi uchun 2-rasmda aynan avtotransport vositalarini aniqlash arxitekturalar evolyutsiyasi keltirilgan.



2-rasm. *Avtotransport vositalarini aniqlash usullarining evolyutsiyasi*

Metodlarning rivojlanishi shuni ko‘rsatadiki, tadqiqotga bo‘lgan qiziqish asta-sekin keyingi kichik bo‘limda tavsiflangan zamonaviy bir bosqichli arxitekturalarga o‘tgan.

2.2. Obyektlarni aniqlashning nazariyasi va zamonaviy modellari

Obyektlarni aniqlash o‘zaro bog‘liq ikkita kichik vazifani hal qiladi: tasvirda obyekt nusxalarini lokalizatsiya qilish va ularni klassifikatsiyalash [23]. Kiruvchi raqamli tasvir $I \in \mathrm{R}^{C\times H\times W}$ tenzori bilan ifodalanadi, bu yerda C, H, W – mos ravishda kanallar soni, balandligi va enidir. Tarmoqqa berishdan oldin piksel qiymatlari normallashtiriladi:

$$\hat{x} = \frac{x}{255} \in [0,1] \qquad (6)$$

Detektorning to‘liq hisoblash konveyeri beshta modulning ketma-ket qo‘llanishi sifatida formallashtiriladi:

$$\hat{Y} = NMS\left(\begin{matrix}Decode\\ Head\\ Neck\\ (Backbone(X))\end{matrix}\right) \qquad (7)$$

bu yerda X – kiruvchi tasvir tenzori, $\hat{Y}$ – klass yorliqlari va confidence score lari bilan bashorat qilingan bounding box larning yakuniy to‘plamidir. Endi (7) da berilgan 1`-bosqichli detektorlarning arxitekturasidagi har bir modullar batafsil tasvirlanadi.

Backbone belgi ajratish moduli kiruvchi $X \in R^{3\times 640\times 640}$ tenzorini batch normalization [10] va ReLU [11] aktivatsiya funksiyasi bilan ketma-ket konvolyutsion bloklar orqali ko‘p darajali ierarxik reprezentatsiyaga aylantiradi.

Har biri fazoviy o‘lchovni ikki baravarga kamaytiruvchi to‘rtta Cross Stage Partial (CSP) [12] bloklari bosqichi turli masshtabdagi uchta belgi xaritasini shakllantiradi:

$$F_s \in R^{128\times 80\times 80}, F_m \in R^{256\times 40\times 40}, F_l \in R^{512\times 20\times 20}$$

Uch masshtab tanlanishi yo‘l sahnasi obyektlarining o‘lchamlari to‘liq diapazonini qamrab olish zarurati bilan shartlangan: F_s kichik o‘lchamli obyektlar uchun yuqori chastotali belgilarni kodlaydi, F_m и F_l – esa katta obyektlar uchun semantik jihatdan boyitilgan belgilarni kodlaydi.

Neck belgilarni birlashtirish moduli Feature Pyramid Network (FPN) [13] va Path Aggregation Network (PAN) [14] orqali turli masshtabdagi kartalar o‘rtasida ikki tomonlama axborot almashinuvini amalga oshiradi.

FPN ning pasayuvchi yo‘li yuqori o‘lchovli kartalarga semantik kontekstni boyitadi, PAN ning ko‘tariluvchi yo‘li esa mahalliy detallarni past o‘lchovli kartalarga qaytaradi.

Chiqishda tekstura va semantik axborotni bir vaqtda intebratsiya qiluvchi uchta boyitilgan karta to‘plami $\{P_s',\ P_m',\ P_l'\}$ shakllanadi.

Detection Head. Uchta kartaning har biriga 1×1 konvolyutsion qatlam qo‘llaniladi, har bir fazoviy pozitsiya uchun A=3 ko‘rinishdagi bashoratni shakllantiruvchi $\left(t_x, t_y, t_w, t_h, t_o, \hat{p}_1, \dots, \hat{p}_{N_c}\right)$, bu yerda dastlabki to‘rtta komponent ramka geometriyasini tavsiflaydi, t_o – obyektlilik, $\hat{p}_c$ – sinflarning ehtimolliklari [6, 9]. Bir kadrga nomzodlarning umumiy soni:

$$N_{total} = A\cdot\sum_k H_k\, W_k = 3\cdot\begin{pmatrix}80^2\\ +40^2\\ +20^2\end{pmatrix} = 25\,200$$

C sinf uchun ramka ishonchliligining yakuniy bahosi quyidagicha hisoblanadi:

$$score_{i,c} = \sigma\left(t_o^i\right)\cdot\sigma\left(\hat{p}_c^i\right) \qquad (9)$$

No-maksimumlarni bostirish (NMS). $\tau_{\text{conf}} = 0.25$ chegarasi bo‘yicha filtrlashdan so‘ng, qolgan nomzodlardan IoU metrikasiga asoslangan takrorlanuvchi ramkalar olib tashlanadi [27]:

$$IoU\left(b_i, b_j\right) = \frac{|\ b_i \cap b_j\ |}{|\ b_i \cup b_j\ |} \in [0,1] \qquad (10)$$

Agar $\mathrm{IoU}(b^*, b_j) > \tau_{\text{nms}}$ bo‘lsa, b_j ramka bostiriladi, bu yerda $\tau_{\text{nms}} = 0.45$. Natijada konveyer chiqishida $\widehat{\mathrm{B}} = \{(b_x^i, b_y^i, b_w^i, b_h^i, c^i, \mathrm{score}^i)\}_{i=1}^{M}$ kompakt to‘plam hosil bo‘ladi, bu yerda $M \ll 25{,}200$ aniqlangan obyektlarning yakuniy soni.

(7) bilan tavsiflangan konveyer ilk avlod bir bosqichli detektorlarning klassik sxemasini ifodalaydi, ularda NMS majburiy yakunlovchi bosqich hisoblanadi [6, 7, 8]. Zamonaviy arxitekturalarda ushbu bosqich bartaraf etilgan: bashoratlarni obyektlar bilan bir qiymatli

moslashtirish vazifasi to‘g‘ridan-to‘g‘ri o‘qitish bosqichida hal qilinadi, bu esa konveyerni quyidagiga qisqartiradi [15, 16, 18]:

$$\hat{Y} = Decode(Head(Neck(Backbone(X))))$$

NMS ni inference konveyeridan chiqarib tashlash qayta ishlash kechikishini kamaytirish va bashorat sifatining chegara parametriga bog‘liqligini bartaraf etish imkonini beradi, bu esa Edge-qurilmalardagi real vaqt tizimlari uchun ayniqsa kritik ahamiyatga ega [16, 18]. Ushbu muammoni hal qilish yo‘lida bir-tomondan evolyutsion, ikkinchi tomondan arxitektura jihatidan prinsipial bo‘lgan ikkita parallel rivojlanish yo‘nalishi shakllandi.

YOLO seriyasi anchor-based yondashuvdan (YOLOv2–v5) YOLOv8 dan boshlab anchor-free arxitekturalarga o‘tdi , bu esa o‘qitish jarayonini soddalashtirdi va aniqlikni oshirdi [9]. Prinsipial qadam sifatida YOLOv11 [19] va YOLO26 [30] ning paydo bo‘lishi xizmat qildi, ular bir qiymatli yorliq tayinlash mexanizmi orqali NMS bog‘liqlikni bartaraf etdi, bu esa inference kechikishini kamaytirish uchun kritik ahamiyatga ega [18, 19]. Paralel ravishda obyektlarni aniqlash sohasida transformer arxitekturalari ham rivojlanishga erishdi: DETR[15] detektori birinchi bo‘lib aniqlash vazifasini NMS siz to‘g‘ridan-to‘g‘ri bashorat sifatida shakllantirdi [17], biroq yuqori hisoblash narxi uning real vaqtda qo‘llanilishini cheklardi. Ushbu cheklov RT-DETR modelida bartaraf etildi [16], u transformer yondashuvining aniqligini YOLO-detektorlar bilan taqqoslanadigan tezlik bilan birlashtirdi: NVIDIA T4 GPU da 108 FPS tezligida 53.1% AP ko‘rsatkichiga erishildi.

2.3. Muammoning qo‘yilishi va tadqiqotning asoslanishi

Avtotransport vositalarini yo‘l kameralaridan olingan video yozuvlarda aniqlash vazifasi intellektual transport tizimlarida kompyuter ko‘rishining asosiy tatbiqlaridan biri hisoblanadi [1, 23]. Shu bilan birga, mavjud qiyosiy tadqiqotlarning aksariyati asosan G‘arbiy Yevropa va Amerika yo‘llarida to‘plangan standart etalon datasetlarga (MS COCO [20], UA-DETRAC [21]) tayanadi, bu esa Markaziy Osiyoning aralash transport oqimi va o‘ziga xos sharoitlari xususiyatini aks ettirmaydi. Bundan tashqari, arxitektura yechimlari, xususan, NMS postishlovidan voz kechish va tasvirning global kontekstini hisobga olishning real yo‘l kuzatuvi sharoitlarida aniqlash aniqligiga ta’siri yetarlicha o‘rganilmagan [16, 18, 19]. Mazkur ish O‘zbekiston avtotransport yo‘li kameralaridan olingan yozuvlar asosida shakllantirilgan original dataset ustida YOLOv11, YOLO26 va RT-DETR modellarining tizimli qiyosiy eksperimenti orqali ko‘rsatilgan bo‘shliqlarni bartaraf etadi.

3. MA’LUMOTLAR TO‘PLAMI VA TAJRIBA METODOLOGIYASI

3.1. Ma’lumotlar to‘plami tavsifi

Modellarni o‘qitish va qiyosiy baholash uchun uch maqsadli klasslarga (car, truck, bus) mansub avtotransport vositalarining qo‘lda belgilangan 1475 ta yo‘l sahnasi tasvirini o‘z ichiga olgan original dataset shakllantirildi. Tasvirlar manbayi sifatida matriks o‘lchovi 5-9 Mp (2560×1920 - 4096×2160 piksel) bo‘lgan O‘zbekiston yo‘l harakati qoidabuzarliklarini videofiksatsiya qilish kameralari boldi. Suratga olish turli ekspluatatsiya sharoitlarini qamrab oladi: kunduzgi va tungi vaqt, turli ob-havo sharoitlari, bu esa datasetning real yo‘l kuzatuvi ssenariylari uchun reprezentativligini ta’minlaydi. Tasvirlarni belgilash Ultralytics Platform muhitida amalga oshirildi [9], uning tanlanishi uch omil bilan shartlangan: annotatsiya va o‘qitish vositalarining yagona muhitda integratsiyalashganligi, bu esa format konvertatsiyasini istisno qiladi va ma’lumotlarni tayyorlashdagi xato ehtimolini kamaytiradi, annotatsiya jarayonini tezlashtiruvchi SAM (Segment Anything Model) [22] modeli asosidagi AI-yordamchi belgilash qo‘llab-quvvatlanishi, shuningdek, tahlil qilinayotgan uchala arxitektura tomonidan qo‘llaniladigan YOLO formati bilan nativ moslik. Dataset 2-jadvalda keltirilgan 80/15/5 nisbatida o‘qitish, validatsiya va test to‘plamlariga bo‘lingan.

2-jadval. Ma’lumotlar to‘plamini taqsimlash

Tanlanma	Rasmlar	Annotatsiyalar
Train (80%)	1179	7432
Val (15%)	251	1014
Test (5%)	45	110
Jami	1475	8 556

3-rasmda O‘zbekiston yo‘l kameralaridan olingan izohli ma’lumotlar to‘plamlari (Uzcardet) keltirilgan. Boshqa tadqiqotchilarga qulay

bo‘lishi uchun tayyor ma’lumotlar to‘plamini olish uchun havola qoldirildi [28].



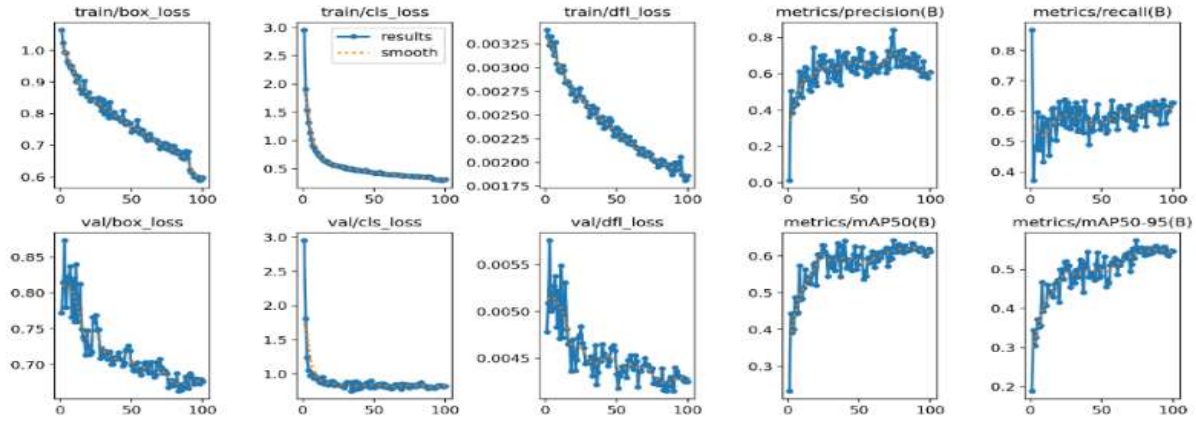
3-rasm. Uzcardetdan annotatsiyalangan ma’lumotlar to‘plamlari

3.2. Baholash metrikalari

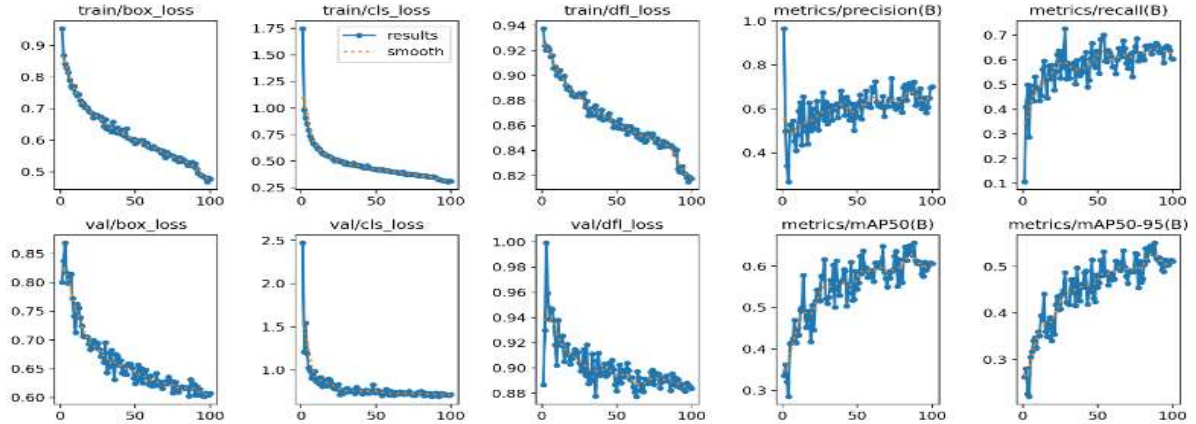
Modellarni miqdoriy solishtirish uchun 1.1-bo‘limda formallashtirilgan Precision, Recall va mAP metrikalari qo‘llanildi (1-3 - formulalar) [20, 21]. Asosiy aniqlik ko‘rsatkichi sifatida MS COCO etalon to‘plamida qabul qilingan $mAP_{(0.5:0.95)}$ ishlatildi [20]. Inference tezligi Latency (ms/kadr) kechikishi va FPS o‘tkazuvchanlik qobiliyati bo‘yicha baholandi (4-5 - formulalar), real vaqt vazifalari uchun qo‘llanilish chegarasi sifatida FPS ≥ 25 qiymati qabul qilindi [2, 4].

3.3. O‘qitish jarayoni va natijalari

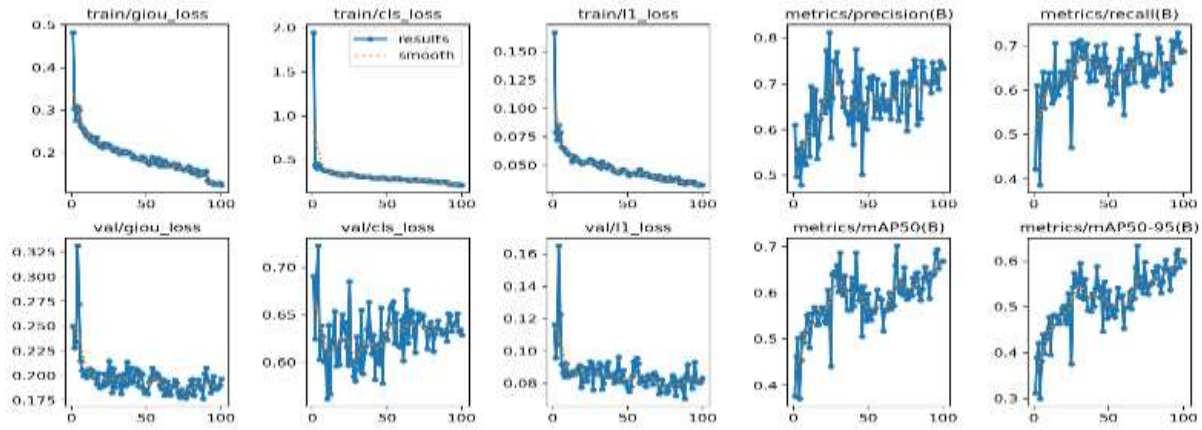
Uchala arxitekturani o‘qitish Ultralytics kutubxonasi [9] asosida yagona dasturiy muhitda NVIDIA Tesla T4 GPU bilan jihozlangan Google Colaboratory bulut platformasida amalga oshirildi. Quyidagi umumlashtirilgan giperparametrlar qo‘llandi: kirish o‘lchovi 640×640, epoxalar soni - 100, batch hajmi - 16. Barcha modellar MS COCO [20] ma’lumotlar to‘plamida oldindan o‘qitilgan boshlang‘ich og‘irliklari bilan initsializatsiya qilindi. Uchala modelning o‘qitishi 100 ta epox davomida muvaffaqiyatli yakunlandi. O‘qitish egri chiziqlari (4, 5, 6 rasmlar) qayta o‘itish belgilarisiz barqaror konvergensiyani ko‘rsatadi: o‘qitish va validatsiya yo‘qotishlari butun o‘qitish jarayoni davomida parallel ravishda kamayib bordi, bu esa o‘qitish tanlamasi hajmining yetarliligi va tanlangan giperparametrlarning to‘g‘riligidan dalolat beradi.



4-rasm. YOLOv26n modelini o‘qitish natijalari



5- rasm. YOLOv11n modelini o‘qitish natijalari



6- rasm. RT-DETR modelini o'qitish natijalari

O'qitishning umumlashtirilgan natijalari 3-jadvalda keltirilgan.

3-jadval.

O'zbekistondagi yo'l kameralaridan olingan ma'lumotlar bazasida, modellarni o'qitish natijalari.

Model	Precision	Recall	F1	mAP50	mAP_{50-95}
YOLO26n	0.608	0.628	0.618	0.613	0.547
YOLOv11n	0.701	0.603	0.648	0.606	0.511
RT-DETR	0.713	0.726	0.719	0.700	0.634

O'qitish natijalarining qiyosiy tahlili shuni ko'rsatadiki, uchala model ham O'zbekiston yo'l kameralari ixtisoslashtirilgan datasetida transport vositalarini aniqlash vazifasini muvaffaqiyatli o'zlashtirdi [9, 16, 19].

O'qitish davomida olingan mAP_{50-95} qiymatlari MS COCO [20] ma'lumotlar to'plamidagi oldindan o'qitilgan og'irliklarning tegishli bazaviy ko'rsatkichlaridan oshib ketdi, bu esa yo'l harakati monitoringi vazifalari uchun fine-tuning samaradorligini tasdiqlaydi.

Shunday qilib, tadqiq etilayotgan arxitekturalarni to'g'ri qiyosiy baholash uchun aniqlik tahlilini real joylashtirishga imkon qadar yaqin sharoitlarda tezlik tavsiflarini o'lchash bilan to'ldirish zarur.

Shu maqsadda keyingi bo'limda real yo'l kuzatuvi sharoitida CPU da inference qilishda uchala modelning har birida Latency va Throughput ni baholash bo'yicha eksperiment taqdim etiladi.

4. REAL SHAROITDA QIYOSIY EKSPERIMENTLAR

4.1. Eksperiment o'tkazish sharoitlari

O'qitilgan modellarning real yo'l kuzatuvi sharoitlarida amaliy qo'llanilishini baholash maqsadida GPU-tezlatgichsiz edge qurilmada inference-eksperiment o'tkazildi. Eksperiment quyidagi texnik xarakteristikalarga ega noutbukda o'tkazildi: Intel Core™ i3-1115G4 protsessori, 8 GB operativ xotira, Windows 10 Pro operatsion tizimi. Inference faqat markaziy protsessorda (CPU-only) amalga oshirildi. Dasturiy muhit: Python 3.15, PyTorch 2.9, Ultralytics 8.4 [9].

Video oqim 1920×1080 piksel o'lchovida va sekundiga 25 kadr chastotasida Deli S106 kamerasidan olindi [29]. Kamera yo'l qoplamasidan 8 metr balandlikda joylashtirildi, bu esa statsionar yo'l harakati qoidabuzarliklarini videofiksatsiya qilish kameralarining tipik joylashuvini taqlid qildi.





7-rasm. Tajriba o‘tkazish sharoitlari

4.2. Vizual deteksiya natijalari

Uchala o‘qitilgan model ketma-ket ravishda real yo‘l harakati sharoitlarida yozilgan bir xil video fragmentga tatbiq etildi. 8-rasmda confidence score chegarasi conf = 0.25 da har bir arxitektura uchun deteksiya natijalari bilan reprezentativ kadrlar keltirilgan.





8-rasm. YOLO26n, YOLOv11n va RT-DETR modellarining deteksiya natijalari

Deteksiya natijalarining vizual tahlili shuni ko‘rsatadiki, uchala model ham real yo‘l harakati sharoitlarida transport vositalarini muvaffaqiyatli lokalizatsiya va klassifikatsiya qiladi. Bounding box larning joylashuvi va sonidagi kuzatilayotgan farqlar har bir modelning arxitektura xususiyatlarini aks ettiradi, ularning miqdoriy baholari keyingi bo‘limda taqdim etilgan.

4.3. CPUda tezlik xususiyatlari

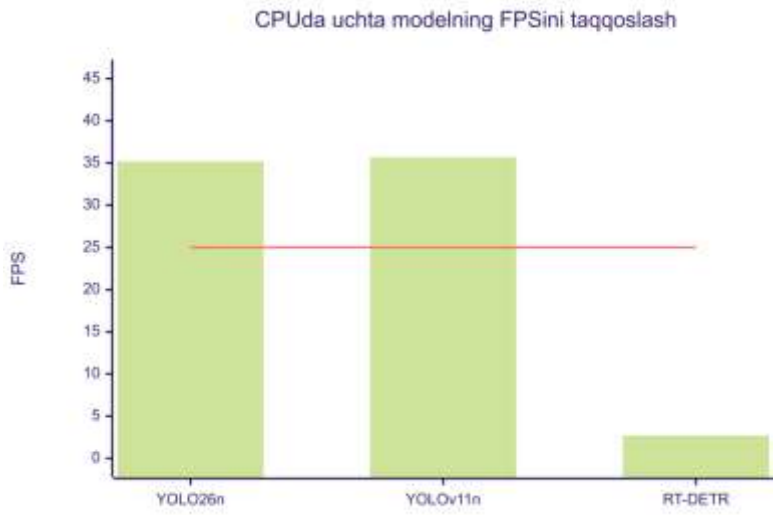
Har bir model uchun CPU da ishlashda bitta kadrni qayta ishlashning o‘rtacha Latency (ms) va Throughput (FPS) ko‘rsatkichlari o‘lchandi. Har bir o‘lchov 222 ta kadrdan iborat ketma-ketlikda amalga oshirildi. Natijalar 4-jadvalda keltirilgan.

4-jadval.

Modellarning protsessorga ta’sir qilishdagi tezlik xususiyatlari.

Model	Inference (ms)	Latency (ms)	FPS	≥25 FPS
YOLO26n	26.9	28.4	35.2	Xa
YOLOv11n	26.0	28.0	35.7	Xa
RT-DETR	369.1	372.3	2.7	Yo‘q

RT-DETR modelining bir xil video fragmentdagi inference natijasi kadrga o‘rtacha 372.3 ms Latency ko‘rsatdi, bu esa 2.7 FPS ga to‘g‘ri keladi. Olingan qiymat YOLO26n (28.4 ms) va YOLOv11n (28.0 ms) modellarining Latency ko‘rsatkichidan 13 martadan ko‘proq oshib ketadi va real vaqt rejimining chegara qiymatidan (25 FPS) sezilarli darajada past. Ushbu natija transformer detektorining arxitektura xususiyatlari bilan belgilanadi: ko‘p boshli Attention mexanizmi [17] va RT-DETR [16] ning Encoder-Decoder tuzilmasi sezilarli darajada ko‘proq hisoblash operatsiyalarini talab qiladi - ular GPU da samarali parallellashtiriladigan bo‘lsa-da, CPU da ketma-ket bajariladi va bu inference Latency ni ko‘p barobar oshiradi. 9-rasmda CPU da inference qilishda uchala modelning FPS ko‘rsatkichlari qiyosiy taqdim etilgan; gorizontal qizil chiziq real vaqt rejimining (25 FPS) chegara qiymatini bildiradi.



9-rasm. *Ustunli diagramma: uchta modelning FPS qiymatini CPU bilan taqqoslash*

Shunday qilib, real sharoitdagi tajriba natijalari test namunasida olingan aniqlik ko‘rsatkichlarini to‘ldiradi va tashqi (edge) qurilmalarda yo‘l harakati monitoringi vazifalari uchun har bir arxitekturaning amaliy qo‘llanilishining to‘liq manzarasini shakllantirishga imkon beradi. Natijalar to‘plamining batafsil izohi 5-bo‘limda keltirilgan.

5. NATIJALAR VA MUHOKAMA

5.1. Aniqlik ko‘rsatkichlarining qiyosiy tahlili

O‘zbekiston yo‘l kameralaridan yig‘ilgan datasetida uchta arxitekturani o‘qitish natijalari shuni ko‘rsatadiki, tadqiq etilayotgan barcha modellar avtotransport vositalarini aniqlash vazifasiga muvaffaqiyatli moslashdi. Erishilgan $mAP_{50\text{-}95}$ qiymatlari MS COCO [9, 16, 19, 20] etalon ma’lumotlar to‘plamidagi oldindan o‘qitilgan og‘irliklarning tegishli bazaviy ko‘rsatkichlaridan oshib ketadi, bu esa yo‘l monitoringining tor ixtisoslashtirilgan vazifalari uchun transfer learning qo‘llashning samaradorligini tasdiqlaydi [2, 4]. Aniqlik metrikalarining yig‘indisi bo‘yicha yetakchi o‘rinni RT-DETR modeli egalladi: mAP50 = 0.700, mAP50-95 = 0.634 va F1 = 0.719 qiymatlariga erishildi. Ushbu natija transformer detektorining arxitektura afzalliklari bilan belgilanadi: ko‘p boshli Attention mexanizmi [17] belgilar shakllantirishda tasvirning global kontekstini hisobga olishni ta’minlaydi, bu esa modelga qisman bir-birini qoplaydigan transport vositalari va murakkab shakldagi obyektlarni aniqroq ajratish imkonini beradi. Shuni alohida ta’kidlash lozimki, RT-DETR tadqiq etilayotgan barcha arxitekturalar orasida eng yuqori Recall qiymatini (0.726) namoyish etdi, bu esa o‘tkazib yuborilgan obyektlarning minimal sonini ko‘rsatadi – YoHQ buzarliklarini avtomatik qayd qilish tizimlari uchun muhim ahamiyat kasb etuvchi xususiyat. YOLOv11n modeli uchta arxitektura orasida Precision bo‘yicha eng yuqori qiymatni (0.701) namoyish etdi, mAP50 = 0.606 va mAP50-95 = 0.511 ko‘rsatkichlari bilan. Ushbu xarakteristika shuni ko‘rsatadiki, model soxta musbat ishlovlarning minimal sonini hosil qiladi, bu esa har bir bashorat avtomatik boshqaruv qarorini hamroh qiladigan ssenariylar uchun ayniqsa qadrlidir. YOLO26n modeli, o‘z navbatida, Precision (0.608) va Recall (0.628) ning yanada muvozanatli nisbatini ta’minladi, biroq integral metrikalar bo‘yicha – mAP50 (0.613) va mAP50-95 (0.547) – YOLOv11n dan ustun keldi. Ko‘rsatilgan tafovut o‘qitish jarayonida yorliq belgilash mexanizmidagi arxitektura farqlari bilan tushuntiriladi: YOLO26n bashoratlarni ground-truth obyektlari bilan solishtirish uchun yanada moslashuvchan strategiyani qo‘llaydi, bu esa modelning turli o‘lcham va yo‘nalishdagi obyektlarga nisbatan sezgirligini oshiradi [30].

5-jadval.

Aniqlik va tezlik metrikalari bo‘yicha umumiy natijalar.

Model	Precision	Recall	F1	mAP50	mAP50-95	Latency (ms)	FPS

YOLO26n	0.608	0.628	0.618	0.613	0.547	28.4	35.2
YOLOv11n	0.701	0.603	0.648	0.606	0.511	28.0	35.7
RT-DETR	0.713	0.726	0.719	0.700	0.634	372.3	2.7

5.2. Tezlik ko‘rsatkichlarining qiyosiy tahlili

CPU-only (Intel Core i3-1115G4) sharoitida o‘tkazilgan inference-eksperiment natijalari tadqiq etilayotgan arxitekturalarning tezlik tavsiflarida tubdan farq mavjudligini aniqladi. YOLO26n va YOLOv11n modellari deyarli bir xil Latency qiymatlarini namoyish etdi: mos ravishda 28.4 va 28.0 ms, bu esa 35.2 va 35.7 FPS Throughput ga to‘g‘ri keladi. Ikkala qiymat ham real vaqt rejimi vazifalari uchun qabul qilingan qo‘llanilish chegarasidan (≥ 25 FPS [2, 4]) taxminan 40% zaxira bilan oshib ketadi, bu esa hisoblash yukining qisqa muddatli cho‘qqilarida tizimning barqaror ishlashini ta’minlaydi. RT-DETR modeli o‘rtacha 372.3 ms Latency ko‘rsatdi, bu esa 2.7 FPS ga to‘g‘ri keladi – YOLO arxitekturalarining ko‘rsatkichlaridan 13 baravar past va real vaqt rejimining chegara qiymatidan bir necha baravar past. Alohida ta’kidlash lozimki, GPU-tezlatgich bilan joylashtirilganda RT-DETR 108 FPS [16] ga erishadi, bu esa arxitekturaning tamoyiliy cheklovini emas, balki uning apparat ixtisoslashuvi xususiyatini ko‘rsatadi.

5.3. Tadqiqotning cheklovlari

Ushbu tadqiqot natijalarini talqin qilishda hisobga olinishi zarur bo‘lgan bir qator cheklovlarga ega. Birinchidan, inference-eksperiment yagona apparat konfiguratsiyasida (Intel Core i3-1115G4, 8 GB RAM) o‘tkazildi, bu esa tezlik natijalarining boshqa edge uskunalariga (Raspberry Pi, Jetson Nano, ARM-protsessorlar) umumlashtirilishini cheklaydi. Ikkinchidan, test namunasi 45 ta tasvir va 110 ta annotatsiyani o‘z ichiga oladi; bu holat qabul qilingan dataset bo‘linish nisbati 80/15/5 va to‘plangan ma’lumotlarning cheklangan hajmi bilan belgilanadi – test namunasini kengaytirish aniqlik metrikalarining statistik ishonchliligini oshirishi mumkin edi. Uchinchidan, tadqiqot modellarning faqat nano-konfiguratsiyalarini qamrab oladi, bu esa xulosaların arxitekturalarning kattaroq variantlariga (small, medium, large) ekstrapolyatsiya qilishga imkon bermaydi.

6. XULOSA

Ushbu ishda YOLO26n, YOLOv11n va RT-DETR arxitekturalarining yo‘l kameralaridan olingan video yozuvlarda avtotransport vositalarini aniqlashga nisbatan qiyosiy tahlili amalga oshirildi. Olingan natijalarga asoslanib quyidagi to‘rtta xulosa chiqarildi:

1. Aniqlik metrikalarining yig‘indisi bo‘yicha RT-DETR modeli ikkala YOLO arxitekturasidan ustun turadi: mAP50-95 = 0.634, F1 = 0.719, Recall = 0.726. Ushbu ustunlik belgilar shakllantirishda butun tasvirning kontekstini hisobga olishga imkon beruvchi global o‘z-o‘ziga Attention mexanizmi bilan belgilanadi. Biroq aniqlikdagi ushbu ustunlik arxitekturaning CPU-qurilmalarda real vaqt rejimida joylashtirishga to‘liq yaroqsizligi ($FPS \leq 25$, chegarasida 2.7 FPS) bilan inkor etiladi.

2. YOLOv11n va YOLO26n modellari resurs cheklangan CPU-qurilmalarda real vaqt rejimida barqaror ishlashni ta’minlaydi: mos ravishda 35.7 va 35.2 FPS, chegara qiymatiga nisbatan ~40% zaxira bilan. YOLOv11n eng yuqori Precision qiymatini (0.701) va minimal inference Latency ni (28.0 ms) namoyish etadi, bu esa uni cheklangan hisoblash resurslari sharoitidagi yo‘l monitoringi sahnalari uchun maqbul qiladi.

3. Uzcardet original dataseti asosida fine-tuning uchala arxitektura uchun ham samarali bo‘ldi: erishilgan mAP_{50-95} qiymatlari MS COCO da oldindan o‘qitilgan og‘irliklarning bazaviy ko‘rsatkichlaridan oshib ketdi, bu esa mintaqaga xos yo‘l monitoringi vazifalari uchun transfer learning metodologiyasini tanlashning asosliligini tasdiqlaydi.

4. O‘zbekiston videofiksatsiya kameralari yozuvlari asosida shakllantirilgan va kunduzi hamda kechasi suratga olish sharoitlarini qamrab olgan Uzcardet original datasetining kiritilishi O‘zbekiston Respublikasi uchun belgilangan ma’lumotlarning mavjudligidagi aniqlangan bo‘shliqni to‘ldiradi va keyingi tadqiqotlar uchun etalon baza bo‘lib xizmat qilishi mumkin.

Olingan natijalarga asoslanib, GPU-tezlatgichsiz Edge-qurilmalarga joylashtirishda YOLOv11n modeli ustunlik berilishi lozim;

GPU-tezlatgich mavjud bo‘lganda va maksimal deteksiya aniqligi talablarida esa RT-DETR arxitekturasi afzalroq hisoblanadi.

Keyingi tadqiqot yo‘nalishlari. Ushbu ishni rivojlantirishning istiqbolli yo‘nalishlari quyidagilardan iborat: (1) Uzcardet datasetini kengaytirish va qo‘shimcha sinflarni kiritish (mototsikllar, maxsus texnika, odamlar, svetoforlar, yo‘l belgilari va boshqalar); (2) ixtisoslashtirilgan Edge-qurilmalarda (NVIDIA Jetson Nano, Raspberry Pi 5, Hailo-8) tezlik tavsiflarini baholash; (3) tadqiq etilayotgan arxitekturalarning o‘rta va katta konfiguratsiyalarida (YOLOv11s/m, RT-DETR-L) eksperimentni kengaytirish boladi.

7. ADABIYOTLAR RO‘YXATI

1. Janai, J., Güney, F., Behl, A., and Geiger, A., “Computer vision for autonomous vehicles: Problems, datasets and state of the art”, Found. Trends Comput. Graph. Vis., vol. 12, no. 1–3, pp. 1–308, 2020, doi: 10.1561/0600000079.
2. Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., “You only look once: Unified, real-time object detection”, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.
3. Girshick, R., Donahue, J., Darrell, T., and Malik, J., “Rich feature hierarchies for accurate object detection and semantic segmentation”, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Columbus, OH, USA, 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.
4. Ren, S., He, K., Girshick, R., and Sun, J., “Faster R-CNN: Towards real-time object detection with region proposal networks”, IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.
5. He, K., Gkioxari, G., Dollár, P., and Girshick, R., “Mask R-CNN”, in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Venice, Italy, 2017, pp. 2961–2969, doi: 10.1109/ICCV.2017.322.
6. Redmon, J. and Farhadi, A., “YOLOv3: An incremental improvement”, arXiv:1804.02767, Apr. 2018.
7. Liu, W. et al., “SSD: Single shot multibox detector”, in Proc. Eur. Conf. Comput. Vis. (ECCV), Amsterdam, Netherlands, 2016, pp. 21–37, doi: 10.1007/978-3-319-46448-0_2.
8. Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P., “Focal loss for dense object detection”, in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Venice, Italy, 2017, pp. 2980–2988, doi: 10.1109/ICCV.2017.324.
9. Jocher, G., Chaurasia, A., and Qiu, J., “Ultralytics YOLO”, Version 8.0.0, GitHub, 2023. [Online]. Available: https://github.com/ultralytics/ultralytics.
10. Ioffe, S. and Szegedy, C., “Batch normalization: Accelerating deep network training by reducing internal covariate shift”, in Proc. Int. Conf. Mach. Learn. (ICML), Lille, France, 2015, pp. 448–456, arXiv:1502.03167.
11. He, K., Zhang, X., Ren, S., and Sun, J., “Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification”, in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Santiago, Chile, 2015, pp. 1026–1034, doi: 10.1109/ICCV.2015.123.
12. Wang, C.-Y., Liao, H.-Y. M., Wu, Y.-H., Chen, P.-Y., Hsieh, J.-W., and Yeh, I.-H., “CSPNet: A new backbone that can enhance learning capability of CNN”, in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 390–391, doi: 10.1109/CVPRW50498.2020.00203.
13. Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S., “Feature pyramid networks for object detection”, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, 2017, pp. 2117–2125, doi: 10.1109/CVPR.2017.106.
14. Liu, S., Qi, L., Qin, H., Shi, J., and Jia, J., “Path aggregation network for instance segmentation”, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Salt Lake City, UT, USA, 2018, pp. 8759–8768, doi: 10.1109/CVPR.2018.00913.
15. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., and Zagoruyko, S., “End-to-end object detection with transformers”, in Proc. Eur. Conf. Comput. Vis. (ECCV), Glasgow, UK, 2020, pp. 213–229, doi: 10.1007/978-3-030-58452-8_13.
16. Zhao, Y. et al., “DETRs beat YOLOs on real-time object detection”, in Proc. IEEE/CVF

Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, 2024, pp. 16965–16974, arXiv:2304.08069.
17. Vaswani, A. et al., "Attention is all you need", in Adv. Neural Inf. Process. Syst. (NeurIPS), Long Beach, CA, USA, 2017, vol. 30, arXiv:1706.03762.
18. Wang, A. et al., "YOLOv10: Real-time end-to-end object detection", in Adv. Neural Inf. Process. Syst. (NeurIPS), 2024, vol. 37, arXiv:2405.14458.
19. Khanam, R. and Hussain, M., "YOLOv11: An overview of the key architectural enhancements", arXiv:2410.17725, Oct. 2024.
20. Lin, T.-Y. et al., "Microsoft COCO: Common objects in context" in Proc. Eur. Conf. Comput. Vis. (ECCV), Zurich, Switzerland, 2014, pp. 740–755, doi: 10.1007/978-3-319-10602-1_48.
21. Wen, L. et al., "UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking", Comput. Vis. Image Underst., vol. 193, p. 102907, Apr. 2020, doi: 10.1016/j.cviu.2020.102907.
22. Kirillov, A. et al., "Segment anything" in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Paris, France, 2023, pp. 4015–4026, arXiv:2304.02643.
23. Dalal, N. and Triggs, B., "Histograms of oriented gradients for human detection", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), San Diego, CA, USA, 2005, pp. 886–893, doi: 10.1109/CVPR.2005.177.
24. Tan, M., Pang, R., and Le, Q. V., "EfficientDet: Scalable and efficient object detection" in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, 2020, pp. 10781–10790, arXiv:1911.09070.
25. Peters, M. D. J. et al., "Chapter 11: Scoping reviews", in JBI Manual for Evidence Synthesis, JBI, Adelaide, Australia, 2020, doi: 10.46658/JBIMES-20-12.
26. Tricco, A. C. et al., "PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation" Ann. Intern. Med., vol. 169, no. 7, pp. 467–473, Oct. 2018, doi: 10.7326/M18-0850.
27. Zheng, Z. et al., "Distance-IoU loss: Faster and better learning for bounding box regression", in Proc. AAAI Conf. Artif. Intell., New York, NY, USA, 2020, vol. 34, no. 7, pp. 12993–13000, arXiv:1911.08287.
28. A.K. Makhamatdinov, "Uzcardet: Road Scene Dataset of Uzbekistan", Ultralytics Platform, 2024. [Online]. Available: https://platform.ultralytics.com/albatross-penguin/datasets/uzcardet [Accessed: Apr. 01, 2026].
29. Deli Group Co., Ltd., "Deli ES106 Security Camera User Manual", 2023. [Online]. Available: https://my.deliworld.com/uploads/file/deli-es106-security-camera-manual.pdf [Accessed: Feb. 11, 2026].
30. Ultralytics, "YOLO26 Model Documentation", 2025. [Online]. Available: https://docs.ultralytics.com/ru/models/yolo26/ [Accessed: Feb. 06, 2026].