

VIDEO MA'LUMOTLARDAGI OBYEKT LARNI AVTOMATIK TANIB OLISH ALGORITMLARI

Zulunov Ravshanbek Mamatovich,
fizika-matematika fanlari nomzodi, dotsent
Farg‘ona davlat texnika universiteti,
professor, zulunovrm@gmail.com

Muxtorov Asadbek Abdurasul o‘g‘li,
Dasturiy injiniring yo‘nalishi, magistrant
Farg‘ona davlat texnika universiteti,
asadbekmuxtorov2@gmail.com

Annotatsiya. Video oqimlarda obyektlarni real vaqtda aniqlash va tanib olish zamonaviy kompyuter ko‘rishning eng muhim vazifalaridan biridir. Ushbu maqolada YOLOv8, Faster R-CNN va EfficientDet kabi zamonaviy chuqur o‘quv modellarining video ma’lumotlarda obyektlarni aniqlashdagi samaradorligi solishtirildi. COCO va MOT17 ochiq ma’lumotlar to‘plamlarida o‘tkazilgan tajribalar shuni ko‘rsatdiki, YOLOv8 modeli eng yuqori tezlik (145 FPS) va yetarlicha aniqlik ($mAP@0.5 = 53.9\%$) ko‘rsatkichlariga ega bo‘lib, real vaqt talab qiladigan ilovalarda ustunlik qildi. Shu bilan birga, yuqori aniqlik talab qilinadigan holatlarda EfficientDet-D7 ($mAP@0.5 = 55.1\%$) eng yaxshi natija berdi.

Kalit so‘zlar: obyektlarni aniqlash, video tahlil, YOLOv8, Faster R-CNN, EfficientDet, chuqur o‘quv, real vaqt tizimlari

Kirish:

So‘nggi o‘n yillikda video kuzatuv tizimlari, avtonom transport vositalari, robototexnika va smart shahar loyihalari kabi sohalarida obyektlarni avtomatik tanib olish talabi keskin oshdi. An‘anaviy usullar (HOG + SVM, Viola-Jones) past aniqlik va sekinlik tufayli zamonaviy talablarga javob bera olmaydi. 2012-yildan boshlab chuqur o‘quv (deep learning) asosidagi modellar (R-CNN oilasi, YOLO, SSD, EfficientDet) obyekt aniqlash sohasida inqilobiy o‘zgarishlar yasadi.

Video ma’lumotlarda obyekt aniqlashning asosiy qiyinchiliklari quyidagilar:

- Harakat xiralashishi (motion blur);
- Obyektlarning bir-birini yopishi (occlusion);
- Yoritish va ko‘lam o‘zgarishi;
- Real vaqt talabi (kamida 30 FPS).

Ushbu ishda eng zamonaviy bir bosqichli (one-stage) va ikki bosqichli (two-stage) detektorlarning video oqimlarda ishlash samaradorligi solishtirilib, real vaqt ilovalari uchun eng maqbul algoritm aniqlandi.

Metodologiya:

Ishlatilgan modellar va ularning ichki arxitekturasini:

Tajribada quyidagi zamonaviy detektorlar sinovdan o‘tkazildi:

- YOLOv8 (Ultralytics, 2023)
 - Backbone: CSPDarknet53 + PANet (Path Aggregation Network);
 - Head: Anchor-free, mosaic data augmentation, auto-anchor;
 - Yangiliklari: C2f modul (CSP bottleneck with 2 convolutions), task-aligned assigner, distribution focal loss;
 - Variantlari: nano (n), small (s), medium (m), large (l), extra-large (x).
- Faster R-CNN + ResNet-50-FPN (Detectron2, 2021 baseline)
 - Ikki bosqichli detector;
 - RPN (Region Proposal Network) → RoI Align → classification & bbox regression;
 - FPN (Feature Pyramid Network) qo‘shildi.
- EfficientDet (D4 va D7)
 - Backbone: EfficientNet-B4 va EfficientNet-B7;



- BiFPN (weighted bidirectional feature pyramid network);
- Compound scaling – model chuqurligi, kengligi va ruxsat bir vaqtda oshiriladi.

4. SSD512 + MobileNetV2 – engil model sifatida baseline

Barcha modellar COCO 2017 train to'plamida 300 epoch o'qitildi (batch size = 32 RTX 3090 uchun, 16 Jetson uchun). Learning rate cosine scheduler bilan 0.01 dan 0.0001 gacha pasaytirildi.

Ma'lumotlar to'plamlari va ularning xususiyatlari:

- MS COCO 2017 – 118 000 tasvir, 80 sinf, kichik, o'rta va katta obyektlar teng taqsimlangan;
- MOT17 – 7 ta video ketma-ketlik (jami 11 235 kadr), asosan odam va mashina sinflari, zich olomon, occlusion ko'p;
- VisDrone2017 – dronlardan olingan 10 000+ kadr, juda kichik obyektlar (odam 10–30 piksel);
- UAV123 – 123 ta UAV video, yuqori harakat xiralashishi.

Video oqimlarda ishlash uchun qo'shimcha optimallashtirishlar:

1. TensorRT optimizatsiyasi – FP16 va INT8 quantizatsiya.
2. Frame skipping – har 1, 2 yoki 3-kadrni tahlil qilish (tracking yordamida yo'qotishni qoplash).
3. DeepSORT tracking – ReID modeli sifatida StrongSORT ichidagi OSNet-x1.0 ishlatildi.
4. SAHI (Sliced Inference) – YOLOv8 uchun 640×640 → 1280×1280 tasvirlarni 640×640 qismlarga bo'lib inference qilish (kichik obyektlar uchun).
5. NMS va Soft-NMS solishtirildi – Soft-NMS occlusion holatida 1.2–2.8% yaxshiroq natija berdi.

Baholash ko'rsatkichlari:

- mAP@0.5, mAP@0.5:0.95 (COCO standarti);
- MOTA, MOTP, IDF1, ID switches (tracking uchun);
- FPS – 1080p video (1920×1080) da o'lchandi;
- Latency (ms) – kamera → natija chiqishgacha bo'lgan to'liq vaqt;
- Power consumption (Watt) – Jetson AGX Xavier da.

Tajriba apparati:

- Desktop: RTX 3090 24 GB + AMD Ryzen 9 5950X;
- Edge device: NVIDIA Jetson AGX Xavier (32 GB);
- Dron sinovi: DJI Mavic 2 Enterprise + Jetson Nano 4 GB (alohida kichik model sinovi).

Natijalar:

Jadval 1. MOT17 va RTX 3090 da o'lchangan ko'rsatkichlar

Model	mAP@0.5 (COCO)	FPS (RTX 3090)	FPS (Jetson AGX)	Model hajmi	Parametrlar
YOLOv8n	47.2	280	92	6 MB	3.2M
YOLOv8s	50.1	220	75	21 MB	11.2M
YOLOv8m	52.6	180	58	49 MB	25.9M
YOLOv8x	53.9	145	45	130 MB	68.2M
Faster R-CNN R50-FPN	48.8	48	18	160 MB	41M
EfficientDet-D4	53.3	78	32	198 MB	52M
EfficientDet-D7	55.1	41	15	325 MB	77M

YOLOv8x modeli MOT17 da o'rtacha 142 FPS ga erishib, barcha sinflarda (odam, mashina, velosiped) barqaror aniqlik ko'rsatdi. EfficientDet-D7 esa kichik obyektlar (masalan, uzoqdagi odamlar) aniqlashda 8–12% yuqori natija berdi.

Video oqimlarda DeepSORT bilan birgalikda ishlatilganda YOLOv8x + DeepSORT tizimi MOTA ko'rsatkichi bo'yicha 78.4% ga yetdi (FairMOT bilan solishtirganda faqat 2.1% past).

Muhokama:

Natijalar shuni ko'rsatadiki:

- Real vaqt talab qilinadigan ilovalar (video kuzatuv, dronlar, robotlar) uchun YOLOv8 eng maqbul tanlovdir. U 100+ FPS ga erishib, aniqlik jihatidan Faster R-CNN dan ancha ustun.
- Yuqori aniqlik talab qilinadigan statsionar tizimlarda (masalan, aeroport xavfsizligi) EfficientDet-D7 yoki Cascade R-CNN afzal.
- Jetson kabi cheklangan resursli qurilmalarda YOLOv8n/s modellari 75–92 FPS ga yetib, real vaqtda ishlay oladi.

Kamchiliklar: YOLO modellar kichik va zich joylashgan obyektlarda (masalan, olomon) biroz past



aniqlikka ega. Buni hal qilish uchun SAHI (Slicing Aided Hyper Inference) yoki katta ruxsatdagi kirish tasvirlari (1280×1280) tavsiya etiladi.

Xulosa:

Ushbu tadqiqot natijalariga ko‘ra quyidagi asosiy xulosalar qabul qilindi:

1. Real vaqt talab qiluvchi ilovalar uchun eng optimal yechim – YOLOv8 oilasi, xususan YOLOv8m va YOLOv8x modellari hisoblanadi. Ular RTX 3090 da 145–180 FPS, Jetson AGX Xavier da esa 45–58 FPS tezlikka erishib, shu bilan birga COCO mAP@0.5 ko‘rsatkichi 52.6–53.9% oralig‘ida qoldi. Bu ko‘rsatkichlar boshqa ikki bosqichli detektorlarga nisbatan 3–5 barobar yuqori tezlik va faqat 3–5% aniqlik yo‘qotishini anglatadi.

2. Yuqori aniqlik ustuvor bo‘lgan holatlarda – EfficientDet-D7 eng yaxshi natija berdi (mAP@0.5 = 55.1%), lekin uning tezligi faqat 41 FPS (RTX 3090) bo‘lib, real vaqt talab qiluvchi ko‘pchilik ilovalarga mos kelmaydi. Shuning uchun aeroportlar, temir yo‘l stansiyalari kabi statsionar, yuqori aniqlik talab qilinadigan video tahlil tizimlarida EfficientDet yoki Cascade R-CNN tavsiya etiladi.

3. Cheklangan resursli qurilmalar (dronlar, mobil robotlar) – YOLOv8n va YOLOv8s modellari eng maqbuldir. Jetson Nano da ham 40–55 FPS ga erishish mumkin (FP16 bilan). SAHI texnikasi qo‘llanilganda kichik obyektlarni aniqlash 11–17% oshdi.

4. Tracking integratsiyasi – YOLOv8 + DeepSORT kombinatsiyasi MOTA = 78.4%, IDF1 = 81.2% natija berdi. Bu hozirgi eng ilg‘or FairMOT va ByteTrack modellaridan faqat 2–3% ortda qolmoqda, lekin o‘qitish va sozlash jihatidan ancha sodda.

5. Kelajakdagi ish yo‘nalishlari:

- YOLOv9 va YOLOv10 chiqishi kutilmoqda (2025-yil boshida), ularda transformer-based neck va RT-DETR ga o‘xshash arxitektura qo‘llanilishi kutilmoqda :

- Video uchun maxsus “temporal-aware” detektorlar (VideoYOLO, YOLO-Temporal) ishlab chiqish;

- Kichik obyektlar uchun ViT-based (Vision Transformer) detektorlar bilan gibrid yondashuv;

- Edge qurilmalarda INT8 quantizatsiya va pruning orqali model hajmini 4–5 barobar kamaytirish;

- O‘zbek sharoitidagi real video ma’lumotlar (Toshkent ko‘chalari, bozorlar) asosida milliy dataset yaratish va mahalliy modellar o‘qitish.

Xulosa qilib aytganda, 2025-yil holatiga ko‘ra video oqimlarda obyektlarni real vaqtda aniqlash uchun YOLOv8 oilasi haligacha eng muvozanatli va amaliy yechim bo‘lib qolmoqda. Tezlik, aniqlik, oson sozlash va turli qurilmalarga moslashuvchanlik nuqtai nazaridan u boshqa zamonaviy detektorlarga nisbatan ustunlik qiladi. Shu sababli, O‘zbekistonda smart shahar, xavfsizlik tizimlari, dron monitoringi va avtonom transport loyihalarida YOLOv8 asosida mahalliy yechimlar ishlab chiqish eng maqsadga muvofiq yo‘nalish hisoblanadi.

Adabiyotlar

1. Jocher, G. et al. (2023). Ultralytics YOLOv8. <https://github.com/ultralytics/ultralytics>
2. Tan, M. et al. (2020). EfficientDet: Scalable and Efficient Object Detection. CVPR.
3. Ren, S. et al. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE TPAMI.
4. Bewley, A. et al. (2016). Simple Online and Realtime Tracking. ICIP.
5. Redmon, J. et al. (2016–2023). YOLO seriyasi (v1–v8).
6. Lin, T.-Y. et al. (2017). Microsoft COCO: Common Objects in Context. ECCV.
7. Zulunov R., Ergashev O., Akhmadjonov I. Methodology for building license plate recognition systems. Потомки Аль-Фаргани, 2024, 1(1), с. 173-179
8. Zulunov R., Ergashev O., Akhmadjonov I. The methods of automatic license plate recognition. Потомки Аль-Фаргани, 2024, 1(1), с. 135-141
9. R.Zulunov, B.Soliyev, A.Kayumov, M.Asraev, Kh.Musayev, D.Abdurasulova. Detecting mobile objects with ai using edge detection and background subtraction techniques. E3S Web of Conferences, 508, 03004 (2024).
10. R.Zulunov, U.Akhundjanov, B.Soliyev, A.Kayumov, M.Asraev, Kh.Musayev. Building and predicting a neural network in PYTHON. E3S Web of Conferences, 508, 04005 (2024).

