

TASVIRDAN OB'YEKTNİ ANIQLASH USULLARI

Eshankulov Hamza Ilhomovich¹, Fayziyev Muhriddin Bahriddin o'g'li²

¹Buxoro davlat universiteti, Axborot texnologiyalari fakulteti dekani, PhD, dots., x.i.eshonkulov@buxdu.uz

²Buxoro davlat universiteti tayanch doktoranti, m.b.fayziyev@buxdu.uz

KEYWORDS

sun'iy intellekt, kompyuter ko'rish, ob'yektni aniqlash, mashinali o'qitish, chuqur o'qitish, algoritm, CNN, YOLO

ABSTRACT

Ushbu maqolada tasvirdan ob'yektni aniqlash masalalari va algoritmlari, xususan, R-CNN, Faster R-CNN va YOLO algoritmlari, algoritmlarning tasvirdan ob'yektni aniqlashdagi samaradorligi tahlil qilingan. Algoritmlarni taqqoslashda YOLO algoritmining YOLOv3 timsoli tanlab olindi, chunki, bu algoritm YOLO ning YOLOv1 va YOLOv2 timsollariga nisbatan sodda va tezroq ishlaydi. Tanlab olingan tasvirdan ob'yektni aniqlash modelini o'rgatish (o'qitish) uchun COCO ma'lumotlar to'plamidan foydalanildi hamda modelni o'qitish va natijalarni olish Colab Notebookda (<https://colab.research.google.com/>) amalga oshirildi.

I. KIRISH

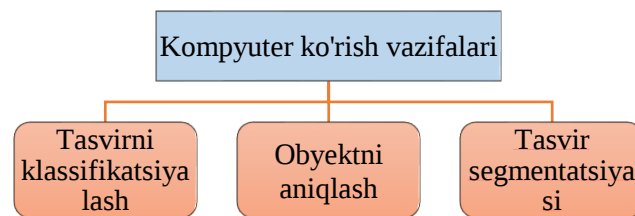
Bugungi kunda jahonda tasvirlarga dastlabki ishlov berish, tasvirlardan muhim ob'ektlarni aniqlash, ajratib olish va sinflashtirish algoritmlarini ishlab chiqishga alohida ahamiyat berilmoqda[1]. Bunga asosiy sabablar sifatida, hozirda axborot texnologiyalarining rivojlanib borayotganligi, videokuzatuv tizimlarini ko'payib borishi, ular bilan o'zaro aloqalarning murakkablashishi kabilarni ko'rsatish mumkin[2]. Axborot texnologiyalarini qo'llash ixtiyoriy sohani rivojlantiradi hamda videotasvirlardan ob'ektlarni tanib olish orqali jarayonlarni kuzatish va samarali boshqarish imkoniyatini beradi. Sun'iy intellekt texnologiyalari yordamida bunday masalalarni yechish bilan ham iqtisodiy, ham ijtimoiy samaradorlikni oshirish bugungi kunning dolzarb masalalaridan biri hisoblanadi.

Tasvirdan ob'ektlarni aniqlash, tanib olish, tasvirlarni qayta ishlash masalalari sun'iy intellekt texnologiyalarining kompyuter ko'rish sohasida o'rganiladi.

Kompyuter ko'rish - bu sun'iy intellekt (SI) sohasi bo'lib, u kompyuterlarni tasvir va video ma'lumotlardan ma'lumot olish hamda talqin qilishga o'rgatadi, shuningdek, u odamlar ko'rish va idrok etishi mumkin bo'lgan barcha narsalar bilan shug'ullanadi. Kompyuter ko'rish

kompyuterlar va tizimlarga raqamli fotosuratlar, raqamli videolar hamda boshqa vizual kontentlardan to'liq mazmunli ma'lumotlarni olish imkonini beradi. Kompyuter ko'rish ob'ektlar, yuzlar, hissiyotlar, ranglar yoki matnlarni aniqlash kabi vizual ma'lumotlarni qayta ishlash va tahlil qilish uchun algoritm va modellardan foydalanishni o'z ichiga oladi[1].

Kompyuter ko'rish sohasi sun'iy intellektning tasvirni klassifikatsiyalash, tasvirdan ob'yektni aniqlash, yangi tasvir generatsiya qilish, anomaliyalarni aniqlash, o'xshash tasvirlarni qidirish masalalari bilan shug'ullanadi[2] va bu masalalarni 1-rasmda keltirilgan kompyuter ko'rish vazifalarini amalga oshirish orqali yechadi.



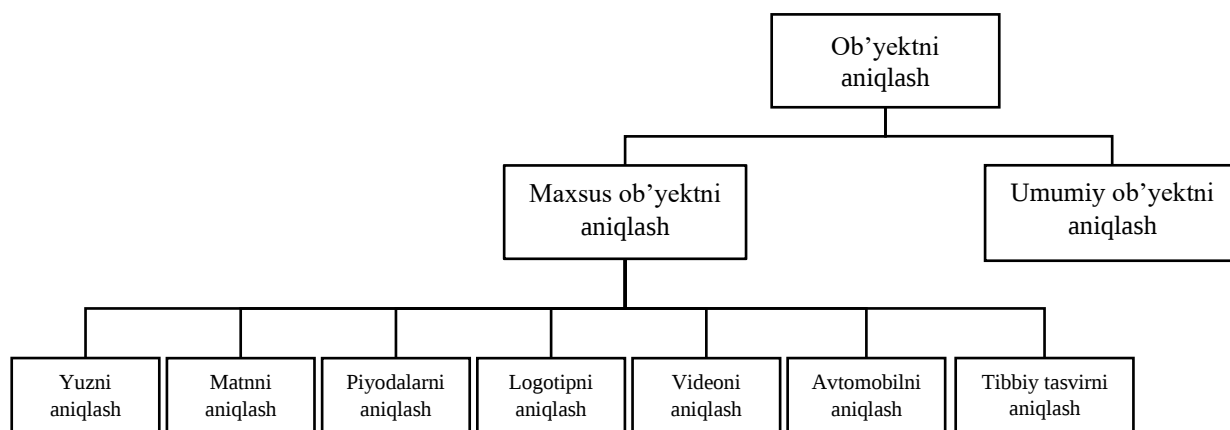
1-rasm. Kompyuter ko'rish vazifalari.

Yuqoridagi 1-rasmda keltirilgan kompyuter ko'rishning vazifalarini quyidagicha tavsiflashimiz mumkin[3].

- **tasvirni klassifikatsiyalash** - berilgan tasvirni mumkin bo'lgan sinflar to'plamiga ko'ra tasniflash vazifasi.
- **ob'yektni aniqlash** - berilgan tasvirdagi bir nechta ob'yektlarni lokalizatsiya qilish va tasniflash vazifasi.
- **tasvir segmentatsiyasi** - berilgan tasvirni bir nechta piksellar to'plamiga bo'lish yoki bir xil tegli piksellar umumiy vizual xususiyatlarga

ega bo'lishi uchun har bir pikselga teglar belgilash vazifasi.

Ob'yektni aniqlash kompyuter ko'rishning muhim sohasi bo'lib, ilmiy tadqiqotlar va amaliy sanoat ishlab chiqarishida yuzni aniqlash, matnni aniqlash, piyodalarni aniqlash, logotipni aniqlash, videoni aniqlash, transport vositasini aniqlash va tibbiy tasvirni aniqlash kabi muhim ilovalarga ega[4] (2-rasm).



2-rasm. Ob'yektni aniqlash masalalari

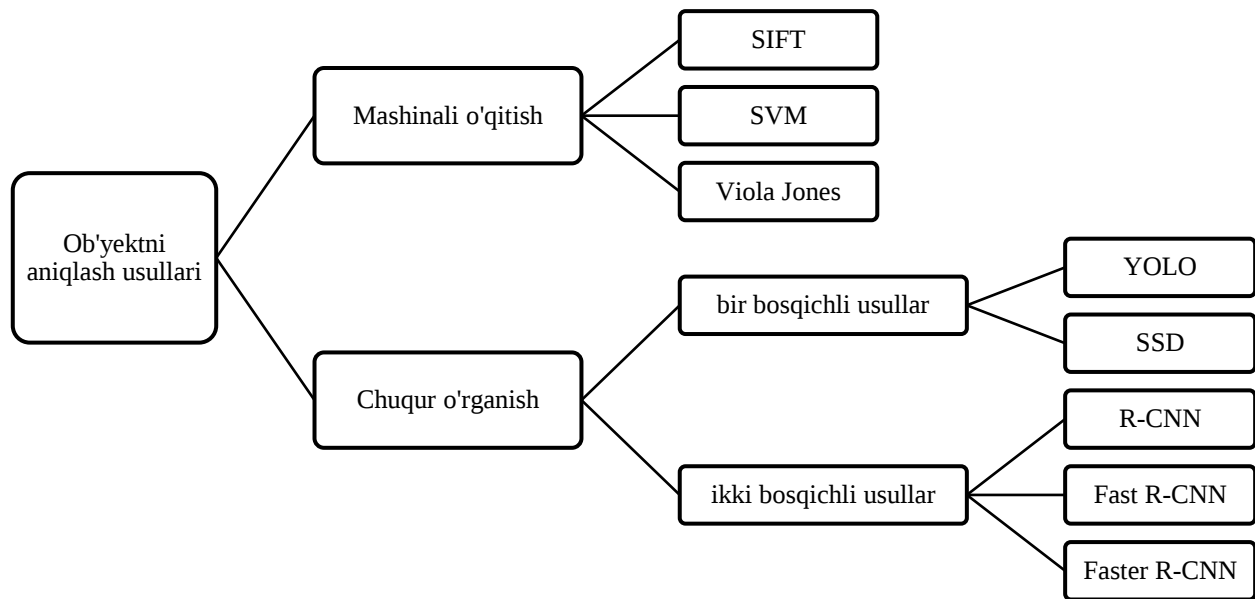
Kompyuter ko'rishning ob'yektni aniqlash vazifasi mashinali o'qitish (Machine Learning) va chuqur o'rganish (Deep Learning) yondashuvlari asosida amalga oshiriladi[5].

Mashinali o'qitish - bu sun'iy intellektning qo'llanilishi bo'lib, unda kompyuter berilgan ma'lumotlardan o'rganadi va ular asosida qarorlar qabul qiladi yoki bashorat qiladi. Odatda ma'lumotlarni tahlil qilish uchun algoritmlardan foydalanish, keyin esa berilgan ma'lumotlar asosida bashorat qilishni va narsalarni aniqlashni o'rganish jarayonidir[6].

Chuqur o'rganish (Deep Learning) - bu mashinali o'qitish algoritmlari sinfidir. Deep

Learning shuningdek, **deep structured learning** deb ham ataladi. Chuqur o'rganish modeli tasvirdan ob'yektni aniqlash uchun unga taqdim etilgan ma'lumotlardan yuqori darajadagi xususiyatlarni olish uchun ko'p qatlamli yondashuvdan foydalanadi. Chuqur o'rganish modeli tasniflash uchun qo'lda taqdim etilishi kerak bo'lgan hech qanday xususiyatni talab qilmaydi. Buning o'rniga, u o'z ma'lumotlarini mavhum tasvirga aylantirishga harakat qiladi. Ko'pgina chuqur o'rganish usullari natijalarga erishish uchun neyron tarmoqlarni amalga oshiradi[7].

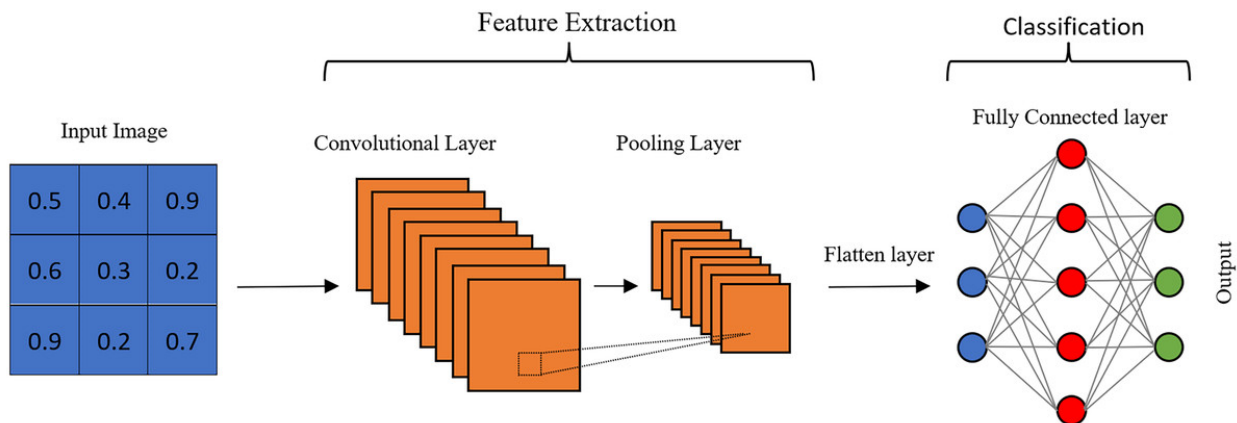
Mashinali o'qitish va chuqur o'rganish yondashuvlari asosida tasvirdan ob'yektni aniqlash algoritmlari[8]:



3-rasm. Ob'yektni aniqlash usullari

Chuqur o'rganish yondashuvi asosida tasvirdan ob'yektni aniqlash algoritmlari konvolyutsion neyron tarmoqlaridan (Convolutional Neural Network (CNN)) foydalanadi.

CNN uch turdagi qatlamlardan iborat. Bular konvolyutsion qatlamlar, birlashtiruvchi qatlamlar va to'liq bog'langan qatlamlardir. Ushbu qatlamlar yig'ilganda CNN arxitekturasini shakllanadi (4-rasm).



4-rasm. CNN arxitekturasini

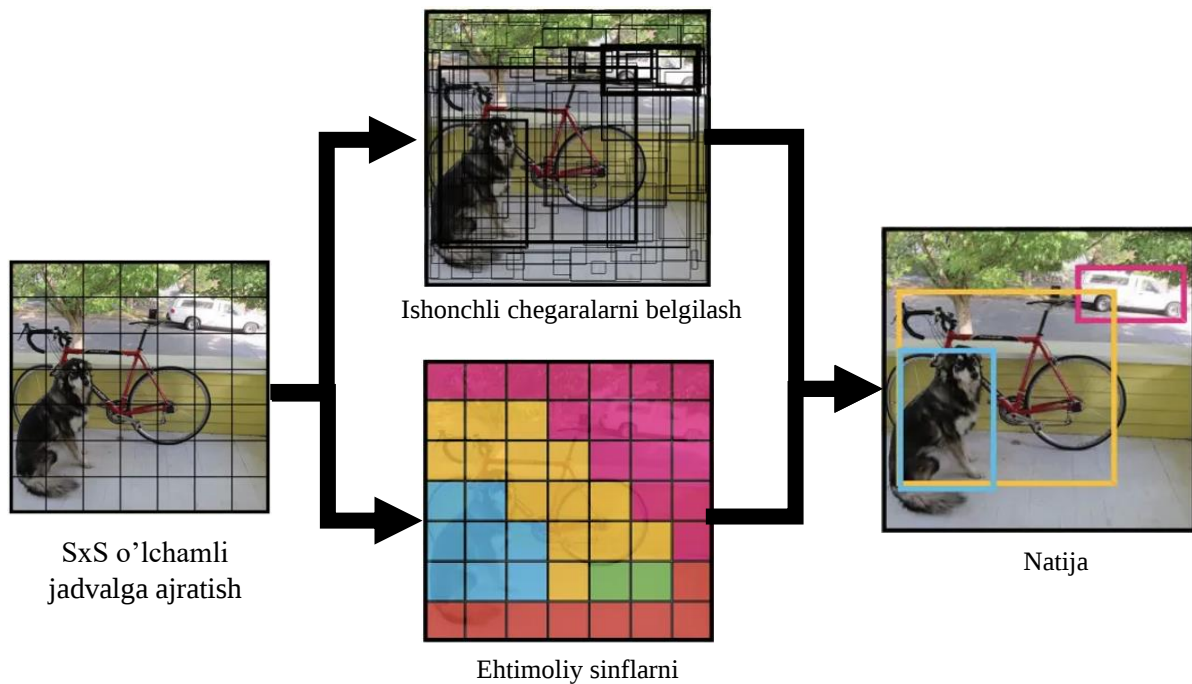
Faster R-CNN algoritmi. Bu algoritmi R-CNN algoritmiga o'xshaydi. Ammo, mintaq bo'yicha takliflarni CNNga yuborish o'rniga, biz konvolyutsion xususiyatlar xaritasini yaratish uchun CNNga kirish tasvirini beramiz. Konvolyutsion xususiyatlar xaritasidan biz takliflar hududini aniqlaymiz va ularni kvadratlarga aylantiramiz va RoI birlashtiruvchi qatlamidan foydalanib, biz ularni to'liq bog'langan qatlamga o'tkazilishi uchun qattiq o'lchamga aylantiramiz. RoI xususiyat vektoridan biz taklif qilingan hududning sinfini va chegara chizig'i uchun ofset qiymatlarini bashorat qilish uchun softmax qatlamidan foydalanamiz[9].

YOLO (You Only Look Once) algoritmi. YOLO - bu 2015-yilda Joseph Redmon va Ali Farhadi tomonidan ishlab chiqilgan real vaqt rejimida ob'yektni aniqlash algoritmi. YOLO kirish tasvirlaridagi ob'ektlarning chegaralangan qutilari va sinf ehtimolliklarini bashorat qilish uchun konvolyutsion neyron tarmog'idan (CNN) foydalanadigan bir bosqichli ob'jekt detektoridir. YOLO birinchi marta Darkent ramkasi yordamida amalga oshirilgan[10].

YOLO algoritmi kiritilgan tasvirni katakchalar panjarasiga ajratadi va har bir katakcha uchun ob'jektning mavjudligi ehtimoli

va ob'yektning chegaralovchi quti koordinatalarini hamda ob'yekt sinfini bashorat qiladi. R-CNN va uning variantlari kabi ikki bosqichli ob'yekt detektorlaridan farqli o'laroq, YOLO butun

tasvirni bir o'tishda qayta ishlaydi, bu esa uni tezroq va samaraliroq ishlashini bildiradi.

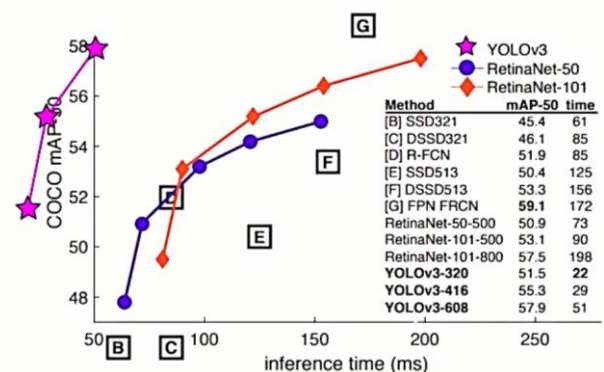


2-rasm. YOLO algoritmining ishlash prinsipi.

Yuqoridagi rasmda YOLO algoritmining ishlash prinsipi keltirilgan bo'lib, undan ko'rinib turibdiki, berilgan tasvir dastlab SxS jadvalga ajratiladi, jadval ichida m chegaralovchi qutilarni olamiz. Har bir cheklov qutisi uchun tarmoq ehtimollik sinfini va chegaralovchi quti uchun ofset qiymatlarini chiqaradi. Eshik qiymatidan yuqori sinf ehtimoliga ega bo'lgan chegaralovchi qutilar tanlanadi va tasvir ichidagi ob'yektni joylashtirish uchun ishlatiladi.

YOLO ob'yektni aniqlashning boshqa algoritmlariga qaraganda tezroq ishlaydi. YOLO algoritmining cheklanishi shundaki, u tasvir ichidagi kichik ob'yektlar bilan kurashadi, masalan, qushlar suruvini aniqlashda qiyinchiliklarga duch kelishi mumkin. Bu algoritmnining fazoviy cheklovlari bilan bog'liq.

Hozirgi kungacha YOLO algoritmining YOLOv1 dan YOLOv8 gacha talqini ishlab chiqilgan bo'lib, biz quyidagi grafikda YOLOv3 algoritmini boshqa algoritmlar bilan taqqoslanishini tasvirladik.



5-rasm. YOLOv3 algoritmini boshqa algoritmlar bilan taqqoslash.

Quyidagi jadvalda COCO ma'lumotlar to'plami orqali o'qitilgan YOLOv3, R-CNN, Faster R-CNN va SSD algoritmlarining taqqoslanishi keltirilgan.

1-jadval.

YOLOv3, R-CNN, Faster R-CNN va SSD algoritmlarini taqqoslash.

Algorit m	Tezli k	Aniqli k	FPS (Fram es Per	MAP (Mean Average
-----------	---------	----------	------------------	-------------------

			Second)	Precision)
YOLOv3	Judatez	Past	~30 FPS	58%
R-CNN	Sekin	Past	<1 FPS	52%
Faster R-CNN	Tez	Yuqori	5-10 FPS	59%
SSD	Sekin	Judayuqori	20-30 FPS	50%

YOLO avtonom boshqariladigan avtomobillar va kuzatuv tizimlari kabi turli xil ilovalarda keng qo'llaniladi. Bundan tashqari, real vaqt rejimida video tahlil va real vaqtda video kuzatuv kabi ob'yektlarni aniqlash vazifalari uchun keng qo'llaniladi.

XULOSA

Ushbu maqolada R-CNN, Faster R-CNN, single shot detector (SSD), YOLOv3 ob'yektlarni aniqlash algoritmlari muhokama qilindi va taqqoslandi. Muhokamalardan ma'lum bo'ldiki, model ishlab chiqilishi bilan tezlik va aniqlik yaxshilangan va ortgan. Fast R-CNN R-CNNga qaraganda yaxshilangan, ammo Faster R-CNN Fast R-CNNga qaraganda ancha yaxshilangan. Bundan tashqari, SSD Faster R-CNN dan yaxshiroq, YOLO esa SSDdan yaxshiroq va tezroqdir. YOLOv3 ishlab chiqilmaguncha, SSD eng yaxshisi edi. Xulosa qilib aytish mumkinki, YOLO v3 modelidan foydalanib, tensorflow yordamida bir nechta ob'yektlarni tezroq aniqlash mumkin va bu model tasvirdan ob'jektini aniqlash uchun bir bosqichli usul orqali ob'jektini to'g'ridan-to'g'ri aniqlay oladi.

FOYDALANILGAN ADABIYOTLAR

1. Muxamadiyev A.Sh., Abdullayev Y.Y. TASVIRDAGI OB'YEKTLARNI ANIQLASH USULLARI // Экономика и социум. 2022. №12-2 (103).
2. Xiao Y., Tian Z., Yu J. et al. A review of object detection based on deep learning. Multimed Tools Appl 79, 23729–23791 (2020). <https://doi.org/10.1007/s11042-020-08976-6>

3. Sampath V., Maurtua I., Aguilar Martín J.J. et al. A survey on generative adversarial networks for imbalance problems in computer vision tasks. J Big Data 8, 27 (2021). <https://doi.org/10.1186/s40537-021-00414-0>
4. Jiao Licheng, Zhang Fan, et al. (2019). A Survey of Deep Learning-based Object Detection.
5. Voulodimos Athanasios, et al. "Deep learning for computer vision: A brief review." Computational intelligence and neuroscience 2018 (2018).
6. Carbonell, Jaime G., Ryszard S. Michalski, and Tom M. Mitchell. "An overview of machine learning." Machine learning (1983): 3-23.
7. Chen, Liang-Chieh, et al. "Learning deep structured models." International Conference on Machine Learning. PMLR, 2015.
8. P. Rajeshwari, et al. Vinod "Object Detection: An Overview" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-3, April 2019, pp.1663-1665. <https://doi.org/10.31142/ijtsrd23422>
9. S. Ren, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, pp. 1137–1149, June 2017.
10. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified, real-time object detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788, June 2016.