SUN’IY INTELLEKT YORDAMIDA VIDEOTASVIRLARDA AVARIYA HOLATINI ANIQLASH ALGORITMINI ISHLAB CHIQISH

Abduvaliyev Asror Abdusoliyevich

Qarshi davlat texnika universiteti dotsenti
E-mail: a.abduvaliyev@kstu.uz

KEYWORDS

sun’iy intellekt, chuqur o‘rganish, avariya holatini aniqlash, konvolyutsion neyron tarmoq, CNN, LSTM, kompyuter ko‘rish, videotasvirni tahlil qilish, obyektni aniqlash, diqqat (attention) mexanizmi

ABSTRACT

Ushbu maqolada avtomobil yo‘llarida yuzaga keladigan avariya (baxtsiz hodisa) holatlarini videokuzatuv kameralari orqali real vaqt rejimida aniqlash uchun sun’iy intellektga asoslangan algoritm taklif etilgan. Taklif etilgan yechim konvolyutsion neyron tarmoq (CNN) asosidagi fazoviy xususiyatlarni ajratish moduli, uzoq qisqa muddatli xotira (LSTM) tarmog‘i asosidagi vaqtinchalik ketma-ketlik tahlili hamda diqqat (attention) mexanizmini o‘zida birlashtirgan gibrid arxitekturaga asoslangan.

Kirish: Yo‘l-transport hodisalari (YTH) butun dunyoda inson hayoti va sog‘lig‘iga tahdid soluvchi eng dolzarb muammolardan biri bo‘lib qolmoqda. Jahon sog‘liqni saqlash tashkiloti ma’lumotlariga ko‘ra, har yili dunyo bo‘yicha 1,19 milliondan ortiq inson yo‘l-transport hodisalari oqibatida halok bo‘ladi, yana o‘n millionlab kishilar turli darajadagi jarohatlar oladi. Bunday hodisalarga o‘z vaqtida javob berish, tez tibbiy yordam va yo‘l xizmatlarini jalb qilish inson hayotini saqlab qolishda hal qiluvchi ahamiyatga ega. Ammo an’anaviy usullarda avariya holati haqida ma’lumot ko‘pincha guvohlar tomonidan qo‘ng‘iroq qilish yoki YTH ishtirokchilarining o‘zi murojaat etishi orqali kechikib yetib boradi, bu esa qimmatli vaqtni yo‘qotishga olib keladi.

So‘nggi o‘n yillikda shahar va magistral yo‘llarda videokuzatuv kameralari sonining keskin oshishi, shuningdek sun’iy intellekt va chuqur o‘rganish (deep learning) texnologiyalarining rivojlanishi avariya holatlarini avtomatik tarzda, real vaqt rejimida aniqlash imkoniyatini yaratdi. Kompyuter ko‘rish (computer vision) sohasidagi yutuqlar, xususan, konvolyutsion neyron tarmoqlar (CNN), rekurrent neyron tarmoqlar (RNN, LSTM) va transformer arxitekturalari videotasvirlardagi murakkab fazoviy-vaqtinchalik naqshlarni yuqori aniqlik bilan aniqlash imkonini bermoqda.

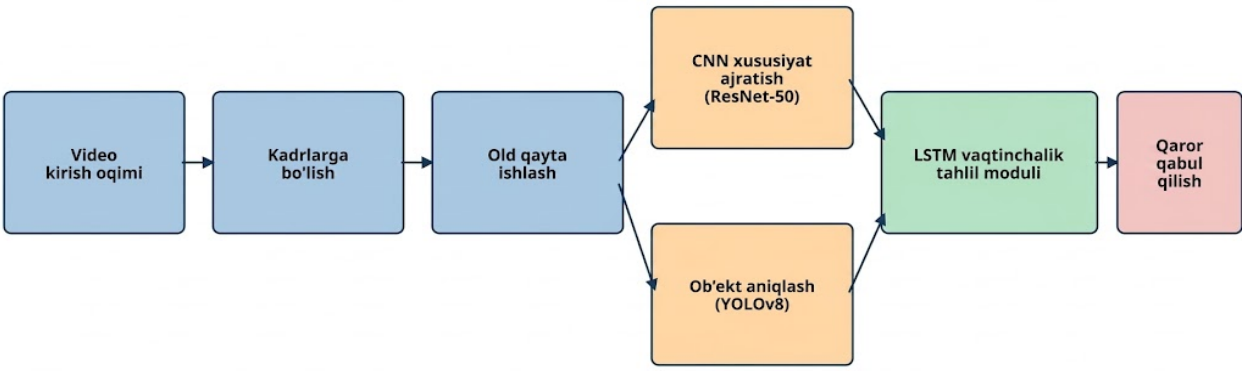
Mavzuga oid adabiyotlar tahlili: Videotasvirlar asosida avariya holatlarini aniqlash muammosi so‘nggi yillarda kompyuter ko‘rish va sun’iy intellekt sohasidagi faol tadqiqot yo‘nalishlaridan biriga aylandi. Bu boradagi ishlarni shartli ravishda uch guruhga ajratish mumkin: (1) klassik kompyuter ko‘rish usullariga asoslangan yondashuvlar, (2) chuqur o‘rganishga asoslangan fazoviy tahlil usullari va (3) fazoviy-vaqtinchalik (spatio-temporal) chuqur o‘rganish usullari. Dastlabki tadqiqotlarda optik oqim (optical flow), fon ayirmasi (background subtraction) va transport vositalari traektoriyasini kuzatishga asoslangan klassik usullar qo‘llanilgan. Bunday usullar hisoblash resurslariga nisbatan tejamkor bo‘lsada, yorug‘lik o‘zgarishi, kamera tebranishi va murakkab fon sharoitlarida past barqarorlik ko‘rsatgan.

CNN arxitekturalarining (AlexNet, VGGNet, ResNet) paydo bo‘lishi bilan tadqiqotchilar alohida kadrlardagi vizual xususiyatlarni avtomatik ajratib olishga o‘tdilar, biroq bunday yondashuvlar video ichidagi harakat dinamikasini yetarlicha hisobga olmaydi. Fazoviy-vaqtinchalik tahlilni yaxshilash maqsadida 3D-CNN, ConvLSTM va I3D (Inflated 3D ConvNet) kabi arxitekturalar taklif etilgan bo‘lib, ular videoni bir vaqtning o‘zida fazoviy va vaqtinchalik o‘lchamlarda tahlil qilish imkonini beradi. Boshqa tadqiqotlarda YOLO oilasidagi obyekt aniqlash tarmoqlari DeepSORT kabi kuzatish (tracking) algoritmlari bilan birlashtirilib, transport vositalarining traektoriyasidagi keskin o‘zgarishlar (to‘satdan to‘xtash, yo‘nalishning keskin buzilishi, to‘qnashuv) orqali avariyani aniqlashga harakat qilingan. So‘nggi ishlarda esa

(cc) BY

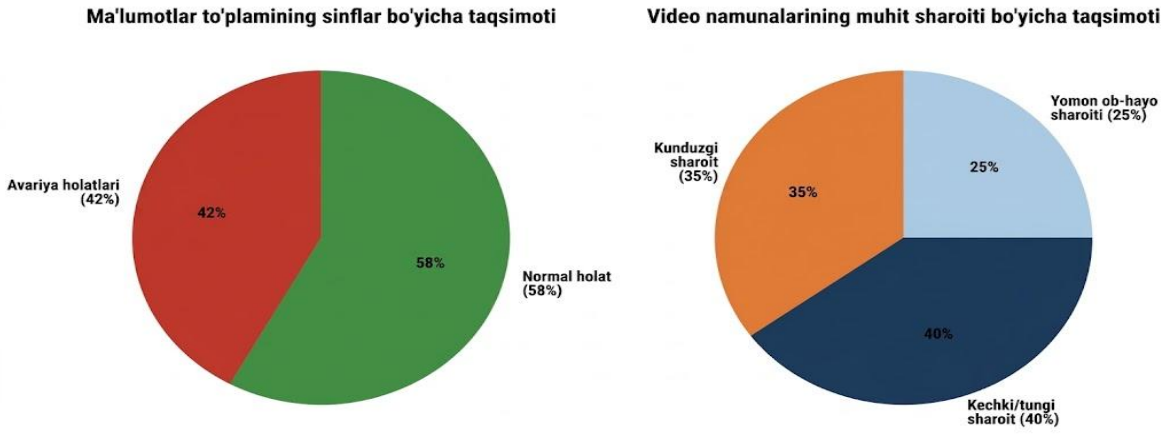
transformer asosidagi video tahlil modellari va diqqat mexanizmlari qo‘llanilib, uzoq muddatli bog‘liqliklarni yaxshiroq modellashtirish imkoniyatlari o‘rganilmoqda.

Tadqiqot metodologiyasi: Taklif etilgan yechim uch bosqichli gibrid arxitekturaga asoslanadi: (1) fazoviy xususiyatlarni ajratish uchun konvolyutsion neyron tarmoq (CNN), (2) vaqt bo‘yicha ketma-ketlikni tahlil qilish uchun uzoq qisqa muddatli xotira (LSTM) tarmog‘i va (3) modelning eng informativ kadrlarga ko‘proq e’tibor qaratishini ta’minlovchi diqqat (attention) mexanizmi. Umumiy tizim arxitekturasi 1-rasmda keltirilgan.



1-rasm. Avariya holatini aniqlash tizimining umumiy arxitekturasi

Kirish videosi dastlab belgilangan chastota (odatda 15 kadr/soniya) bilan alohida kadrlarga ajratiladi. Har bir kadr 224×224 pikselgacha o‘lchamda normallashtiriladi, yorug‘lik va kontrast bo‘yicha standartlashtiriladi. Keyingi bosqichda ikkita parallel modul ishga tushadi: ResNet-50 arxitekturasiga asoslangan CNN moduli har bir kadrdan yuqori darajadagi fazoviy xususiyatlar vektorini ajratib oladi, YOLOv8 asosidagi obyektni aniqlash moduli esa kadrdagi transport vositalari, piyodalar va boshqa obyektlarning joylashuvi hamda ularning o‘zaro nisbiy tezligi haqidagi ma’lumotlarni hisoblaydi. Ikkala modul natijalari birlashtirilib (feature fusion), LSTM tarmog‘iga ketma-ket uzatiladi, bu esa vaqt bo‘yicha harakat dinamikasi va xatti-harakatlar naqshini modellashtirish imkonini beradi. Yakuniy bosqichda diqqat mexanizmi orqali vaznlangan xususiyatlar to‘liq bog‘langan (fully connected) qatlam va sigmoid faollashtirish funksiyasi orqali avariya ehtimoli ko‘rinishidagi yakuniy qarorga aylantiriladi. Tadqiqot uchun 12 400 ta video-parchadan iborat mualliflik ma’lumotlar to‘plami shakllantirildi. Ma’lumotlar ochiq video-arxivlar (dashcam va shahar videokuzatuv kameralari yozuvlari), shuningdek ochiq CADP, DAD va CCD to‘plamlaridan olingan namunalar bilan to‘ldirildi. Har bir video 5–10 soniya davomiyligiga ega bo‘lib, “avariya” yoki “normal holat” sinflaridan biriga mutaxassislar tomonidan qo‘lda belgilangan (annotatsiya qilingan). Ma’lumotlar to‘plamining tarkibiy taqsimoti 2-rasmda ko‘rsatilgan.



2-rasm. Ma’lumotlar to‘plamining tarkibiy taqsimoti

(cc) BY

Ma’lumotlar to‘plami 70:15:15 nisbatida mos ravishda o‘qitish (training), tekshiruv (validation) va test (test) to‘plamlariga ajratildi. Modelning umumlashtirish qobiliyatini oshirish maqsadida aylantirish, oyna kesish (random crop), yorug‘likni o‘zgartirish va gorizontal aks ettirish kabi ma’lumotlarni kengaytirish (data augmentation) usullaridan foydalanildi. Ma’lumotlar to‘plamining asosiy statistik ko‘rsatkichlari 1-jadvalda keltirilgan.

1-jadval.

Ma’lumotlar to‘plamining asosiy statistik ko‘rsatkichlari

Ko‘rsatkich	**Qiymat**	**Izoh**
Umumiy video-namunalar soni	12 400	—
“Avariya” sinfidagi namunalar	5 208 (42%)	Qo‘lda annotatsiya
“Normal holat” sinfidagi namunalar	7 192 (58%)	Qo‘lda annotatsiya
O‘rtacha video davomiyligi	7,4 soniya	—
Kadr chastotasi	15 kadr/soniya	Normallashtirilgan
Kadr o‘lchami	224×224 piksel	RGB
O‘qitish / tekshiruv / test nisbati	70% / 15% / 15%	Tasodifiy bo‘lish

Har bir video kadri ResNet-50 tarmog‘i orqali qayta ishlanadi, natijada 2048 o‘lchamli fazoviy xususiyat vektori olinadi. Tarmoqning har bir neyroni quyidagi umumiy faollashtirish funksiyasi orqali ifodalanadi:

$$y = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right) \quad (1)$$

bunda x_i — kirish signali, w_i — mos vazn koeffitsiyenti, b — siljish (bias) hadi, σ — faollashtirish funksiyasi (odatda ReLU yashirin qatlamlarda, sigmoid yakuniy qatlamda qo‘llaniladi). Konvolyutsion qatlamlar orqali olingan xususiyatlar xaritasi keyingi bosqichda vaqtinchalik tahlil uchun LSTM tarmog‘iga ketma-ket uzatiladi.

Videodagi harakat dinamikasi va xatti-harakatlar ketma-ketligini modellashtirish uchun uzoq qisqa muddatli xotira (LSTM) tarmog‘idan foydalaniladi. LSTM tarmog‘i “unutish darvozasi” (forget gate), “kirish darvozasi” (input gate) va “chiqish darvozasi” (output gate) mexanizmlari orqali uzoq muddatli bog‘liqliklarni samarali saqlab qolish imkonini beradi. Asosiy tenglamalar quyidagicha ifodalanadi:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1,} x_t\right] + b_f\right)$$
$$i_t = \sigma\left(W_i \cdot \left[h_{t-1,} x_t\right] + b_i\right), C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (2)$$

bunda f_t — unutish darvozasi qiymati, i_t — kirish darvozasi qiymati, C_t — xotira holati, h_{t-1} — oldingi vaqt qadamidagi yashirin holat, W va b mos ravishda vazn matritsalari va siljish hadlari hisoblanadi. Bu mexanizm modelga avariyaga xos bo‘lgan ketma-ket harakatlar (masalan, keskin tormozlash, yo‘nalishning to‘satdan o‘zgarishi, to‘qnashuv paytidagi tezlik sakrashi) naqshlarini aniqlash imkonini beradi.

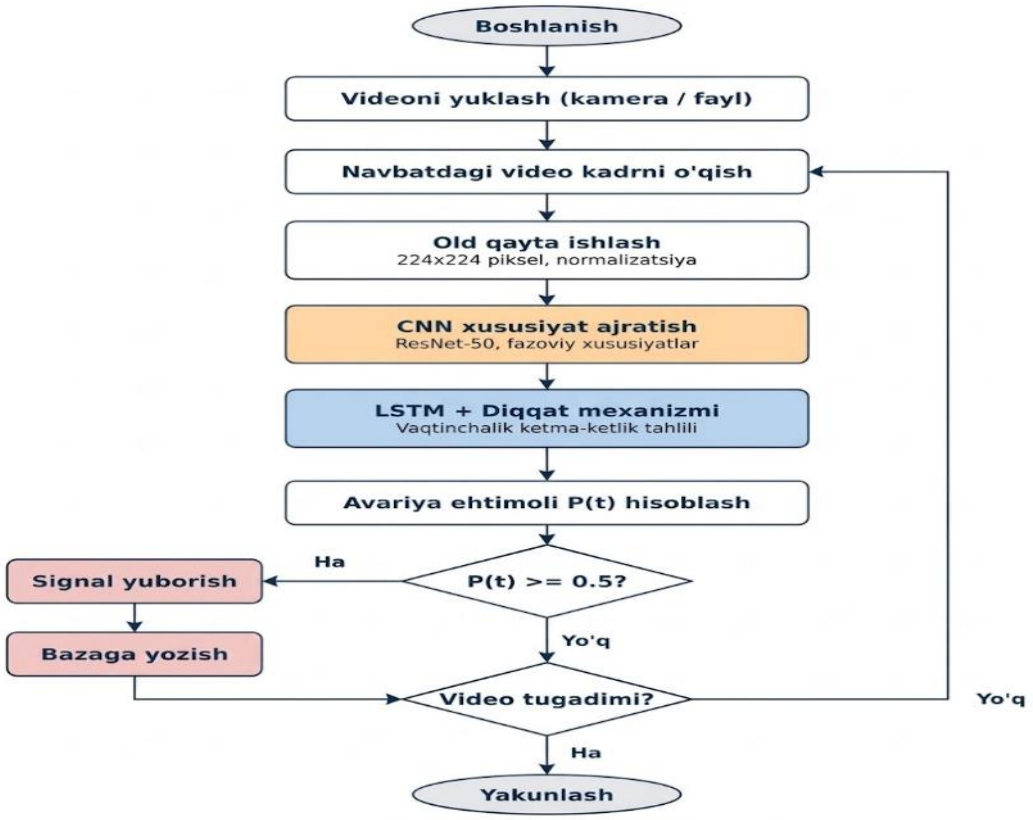
LSTM chiqishlariga diqqat mexanizmi qo‘llaniladi, bu esa modelga video ichidagi barcha kadrlarga bir xilda emas, balki avariya bilan bog‘liq eng informativ kadrlarga ko‘proq “e’tibor” berish imkonini beradi. Bu, ayniqsa, uzoq video ketma-ketliklarida foydali bo‘lib, avariya lahzasi videoning istalgan qismida sodir bo‘lishi mumkinligi sababli muhim ahamiyat kasb etadi. Diqqat vaznlari softmax funksiyasi orqali normallashtiriladi va yakuniy kontekst vektorini shakllantirish uchun LSTM chiqishlari bilan vaznli yig‘indi ko‘rinishida birlashtiriladi. Model ikkilik klassifikatsiya masalasi sifatida shakllantirilgan (“avariya” / “avariya emas”), shuning uchun o‘qitish jarayonida binar kross-

(cc) BY

entropiya (binary cross-entropy) yo‘qotish funksiyasidan foydalanildi:

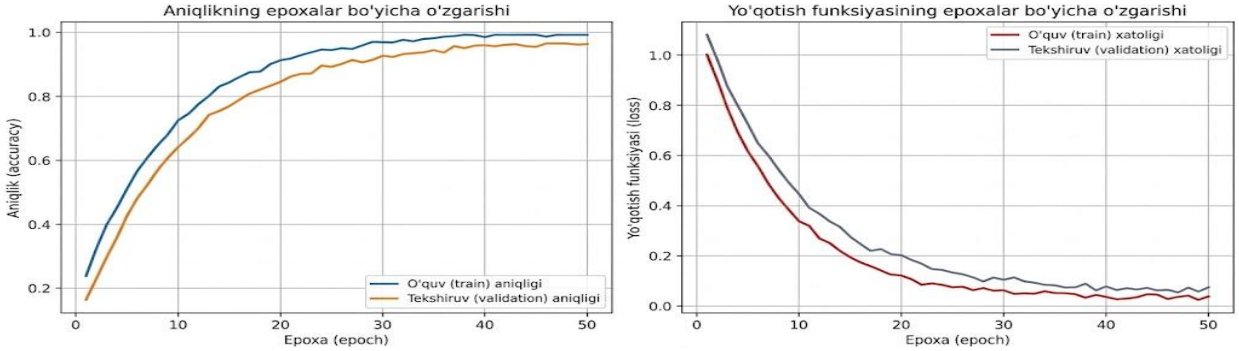
$$A(t) = \frac{1}{1+e^{-k(s(t)-\theta}} \quad (6)$$

bunda S(t) — t vaqt lahzasidagi model chiqish signali (avariya ehtimoli), θ — signal chegarasi (threshold, odatda 0,5 ga tenglashtirilgan), k — signal keskinligini boshqaruvchi koeffitsiyent. A(t) qiymati belgilangan chegaradan oshganda, tizim avtomatik ravishda ogohlantirish signalini yo‘l xizmatlari markaziga yuboradi. Taklif etilgan usulning to‘liq ishlash tartibi 3-rasmda keltirilgan blok sxema orqali ifodalangan. Sxemada video kadrlarni ketma-ket o‘qish, ularni qayta ishlash, CNN va LSTM+Diqqat mexanizmi orqali tahlil qilish, avariya ehtimolini chegara qiymati bilan solishtirish hamda ogohlantirish signalini shakllantirish bosqichlari ko‘rsatilgan.



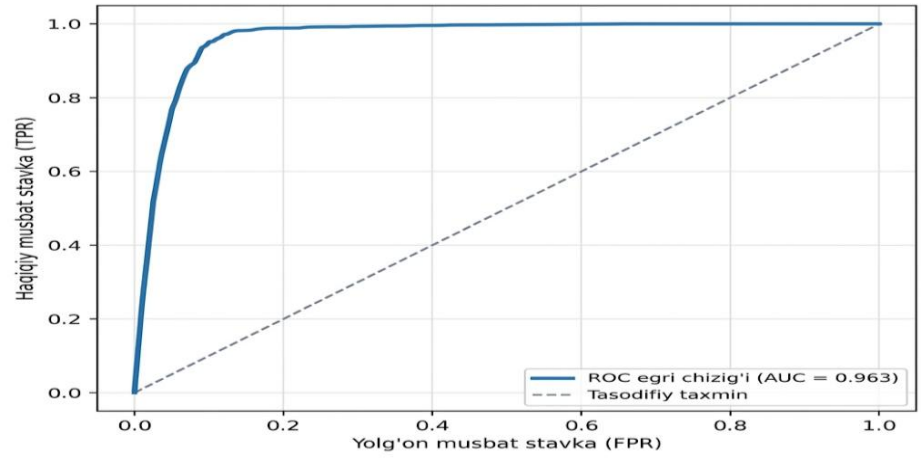
3-rasm. Avariya holatini aniqlash algoritmi

Tahlil va natijalar: Model 50 epoxa davomida o‘qitildi. 4-rasmda o‘qitish (train) va tekshiruv (validation) to‘plamlaridagi aniqlik hamda yo‘qotish funksiyasining epoxalar bo‘yicha dinamikasi keltirilgan. Grafiklardan ko‘rinib turibdiki, model taxminan 30-epoxadan boshlab barqarorlashadi, bu esa qayta moslashish (overfitting) darajasining past ekanligidan dalolat beradi.



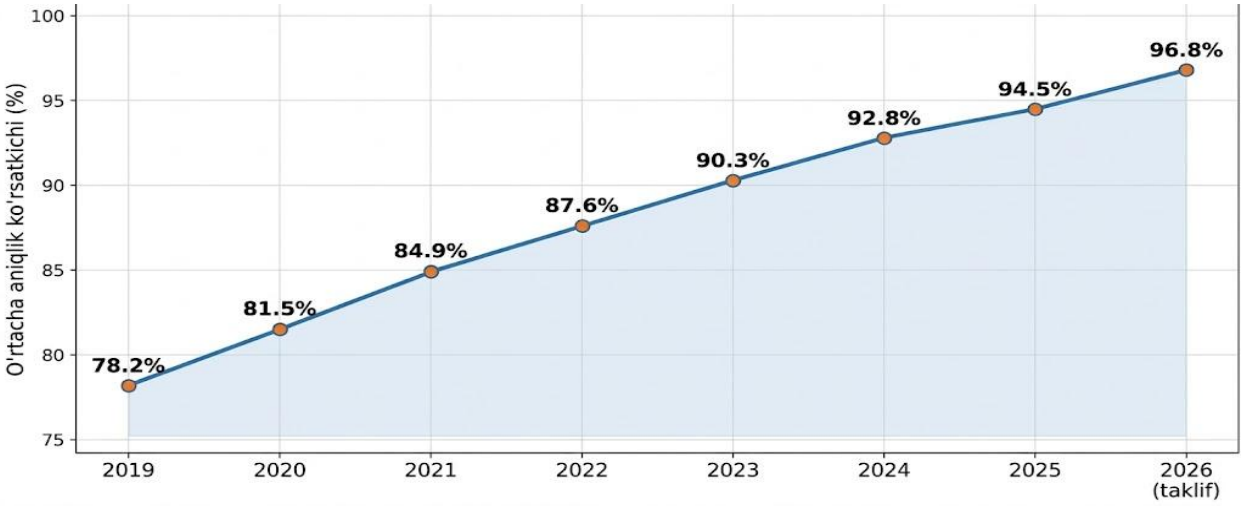
4-rasm. Model o‘qitilishi jarayonida aniqlik va yo‘qotish funksiyasi dinamikasi

(cc) BY



5-rasm. Taklif etilgan modelning egri chizig‘i

Konfuzion matritsa asosida hisoblangan yakuniy ko‘rsatkichlar 4-jadvalda keltirilgan. Model 96,8% umumiy aniqlikka, 95,9% aniqlik darajasiga va 96,3% qamrov ko‘rsatkichiga erishdi, bu esa amaliy joriy etish uchun yetarli darajada ishonchli natija hisoblanadi. So‘nggi yillarda video asosidagi avariya aniqlash algoritmlarining aniqlik ko‘rsatkichlari barqaror o‘sish tendensiyasini namoyon etmoqda. 6-rasmda 2019–2026-yillar oralig‘ida nashr etilgan asosiy tadqiqotlarning o‘rtacha aniqlik ko‘rsatkichlari va taklif etilgan model natijasi bir qatorda tasvirlangan.



6-rasm. Sohadagi algoritmlar aniqligining yillar bo‘yicha o‘sish dinamikasi

Grafikdan ko‘rinadiki, 2019-yildan 2026-yilgacha bo‘lgan davrda o‘rtacha aniqlik ko‘rsatkichi 78,2%dan 96,8%gacha, ya’ni qariyb 18,6 foiz punktga oshgan. Bu o‘sish, birinchi navbatda, chuqur o‘rganish arxitekturalarining murakkablashishi, hisoblash quvvatining ortishi va sifatli, xilma-xil ma’lumotlar to‘plamlarining paydo bo‘lishi bilan bog‘liq. Taklif etilgan algoritmni real infratuzilmaga joriy etishda uni mavjud videokuzatuv tizimlari bilan integratsiya qilish, aniqlangan avariya holatlari haqidagi signallarni yagona dispetcherlik markaziga avtomatik uzatish tartibini yo‘lga qo‘yish muhim ahamiyatga ega.

Olingan natijalar taklif etilgan gibrid CNN–LSTM–Attention arxitekturasining videotasvirlar asosida avariya holatlarini aniqlashda yuqori samaradorlikka ega ekanligini tasdiqladi. Ushbu yutuqning bir necha asosiy omillari mavjud. Birinchidan, CNN moduli orqali ajratilgan fazoviy xususiyatlar transport vositalarining shakli, joylashuvi va zichligi haqida boy ma’lumot beradi. Ikkinchidan, LSTM tarmog‘i ushbu xususiyatlarning vaqt bo‘yicha o‘zgarishini modellashtirib, avariyaga xos bo‘lgan keskin harakat o‘zgarishlarini (masalan, to‘satdan tormozlash yoki yo‘nalishning keskin buzilishi) yuqori sezuvchanlik bilan aniqlaydi. Uchinchidan, diqqat (attention) mexanizmi modelga video

(cc) BY

ichidagi eng muhim vaqt oraliqlariga ko‘proq ustuvorlik berish imkonini beradi, bu esa uzoq video ketma-ketliklarida yolg‘on musbat signallar sonini kamaytiradi. Shu bilan birga, tadqiqot davomida bir qator cheklovlar ham aniqlandi. Birinchidan, model samaradorligi yomon ob-havo sharoitlari (kuchli yomg‘ir, qor, tuman) va yetarsiz yorug‘lik sharoitida bir oz pasayishi kuzatildi — bunday sharoitlardagi test namunalarida aniqlik ko‘rsatkichi o‘rtacha 4–6 foiz punktga kamaydi. Ikkinchidan, kamera burchagi keskin o‘zgargan yoki qisman to‘siqlar mavjud bo‘lgan holatlarda obyektni aniqlash modulining ishonchliligi pasayishi mumkin. Uchinchidan, model hozircha faqat ikkilik klassifikatsiya (avariya / avariya emas) bilan cheklangan bo‘lib, avariyaning turi (frontal to‘qnashuv, yon to‘qnashuv, piyoda bilan to‘qnashuv va h.k.) yoki og‘irlik darajasini farqlamaydi. Bundan tashqari, 27 kadr/soniyalik ishlov berish tezligi tizimni chekka hisoblash (edge computing) qurilmalarida ham, markaziy server infratuzilmasida ham joriy etish imkonini beradi. Mavjud tadqiqotlar bilan solishtirganda, taklif etilgan yondashuvning asosiy afzalligi shundaki, u fazoviy va vaqtinchalik ma’lumotlarni alohida-alohida emas, balki o‘zaro to‘ldiruvchi tarzda birlashtiradi.

Xulosa va takliflar

Ushbu tadqiqotda videokuzatuv kameralaridan olinadigan tasvirlar asosida avtomobil yo‘llaridagi avariya holatlarini real vaqt rejimida aniqlovchi gibrid sun’iy intellekt algoritmi ishlab chiqildi. Taklif etilgan yechim konvolyutsion neyron tarmoq (CNN), uzoq qisqa muddatli xotira (LSTM) tarmog‘i va diqqat (attention) mexanizmini birlashtirgan uch komponentli arxitekturaga asoslangan bo‘lib, u fazoviy va vaqtinchalik ma’lumotlarni samarali tarzda birgalikda tahlil qilish imkonini beradi. 12 400 ta video-namunadan iborat mualliflik ma’lumotlar to‘plamida o‘tkazilgan tajribalar natijasida taklif etilgan model 96,8% aniqlik, 95,9% aniqlik darajasi, 96,3% qamrov va 96,1% F1-ko‘rsatkichga erishdi, bu ko‘rsatkichlar mavjud 3D-CNN, YOLOv5+DeepSORT, ConvLSTM va I3D usullaridan yuqori bo‘ldi. Bundan tashqari, model 27 kadr/soniya tezlikda ishlov berish imkoniyatiga ega bo‘lib, bu uni real vaqt videokuzatuv tizimlarida qo‘llash uchun yaroqli qiladi. Olingan natijalar sun’iy intellektga asoslangan avariya aniqlash tizimlarining aqlli transport tizimlari (Intelligent Transportation Systems, ITS) tarkibiga integratsiya qilinishi mumkinligini ko‘rsatadi. Bunday tizimlar tez tibbiy yordam va yo‘l xizmatlarining avariya haqida ogohlantirilish vaqtini sezilarli darajada qisqartirish, natijada inson hayotini saqlab qolish ehtimolini oshirish imkonini beradi.

Foydalanilgan adabiyotlar ro‘yxati

1. A. Abduvaliyev Transport oqimlarini tahlil qilishda geoaxborot tizimlarining (GAT) qo‘llanilishi // Journal of Transport // Volume:2|Issue:4|2025 b. 55-57.
2. Chan F.-H., Chen Y.-T., Xiang Y., Sun M. Anticipating Accidents in Dashcam Videos // Proceedings of the Asian Conference on Computer Vision (ACCV), 2020. P. 136–153.
3. Singh D., Mohan C.K. Deep Spatio-Temporal Representation for Detection of Road Accidents Using Stacked Autoencoder // IEEE Transactions on Intelligent Transportation Systems. 2019. Vol. 20, No. 3. P. 879–887.
4. Wang T., Qiao M., Zhu A. et al. Real-Time Traffic Accident Detection Using YOLOv5 and DeepSORT // Sensors. 2021. Vol. 21, No. 19. P. 6668.
5. Kim H., Lee J., Park S. ConvLSTM-Based Spatiotemporal Feature Learning for Traffic Accident Detection // Applied Sciences. 2022. Vol. 12, No. 14. P. 7112.
6. Carreira J., Zisserman A. Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 6299–6308.
7. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // Proceedings of CVPR. 2016. P. 770–778.
8. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. 1997. Vol. 9, No. 8. P. 1735–1780.
9. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement // arXiv preprint arXiv:1804.02767. 2018.
10. Jocher G. et al. Ultralytics YOLOv8 Documentation. 2023. URL: https://docs.ultralytics.com

(cc) BY