

Жасанды интеллект негізінде фото және бейне деректердегі объектілерді тану әдістерін зерттеу

ЖАРИЯЛАНДЫ
28.05.2026СІЛТЕМЕ
<https://bilimger.kz/188993/>

Сыдан Ералы Дастанұлы

*Математика және Ақпараттық технологиялар факультеті**Академик Е.А Бөкетов атындағы Қарағанды ұлттық зерттеу университеті*

Аннотация

Бұл мақалада жасанды интеллект технологиялары негізінде фото және бейне деректердегі объектілерді тану әдістеріне жан-жақты зерттеу жүргізілген. Зерттеу барысында терең оқыту архитектуралары, атап айтқанда конволюциялық нейрондық желілер, жылдам анықтау желілері және трансформер негізіндегі модельдер талданды. Объектілерді тану жүйелерінің өнімділігін арттыруда деректер жиынтығының сапасы, модель архитектурасы және оқыту стратегияларының маңызы ерекше қарастырылды. Зерттеу нәтижелері осы технологиялардың нақты уақыт режимінде бейнені өңдеу, қауіпсіздік жүйелері, дәрігерлік диагностика және автономды көлік жүргізу салаларында кеңінен қолданылатынын көрсетеді. Мақалада сондай-ақ заманауи объектілерді тану жүйелерінің кемшіліктері мен болашақта дамудың негізгі бағыттары сарапталған.

Кілт сөздер: жасанды интеллект, объектілерді тану, конволюциялық нейрондық желі, терең оқыту, жылдам анықтау желісі, трансформер, бейнені өңдеу, компьютерлік көру, машиналық оқыту, нейрондық желі.

Кіріспе және тарихи шолу

Компьютерлік көру — жасанды интеллект саласының маңызды бөлімі болып табылады. Оның негізгі міндеті — машинаның кескін мен бейне ақпаратты адам сияқты түсініп, талдай алуына мүмкіндік беру. Фото және бейне деректердегі объектілерді автоматты түрде тану мәселесі 1960–70-жылдардан бастап зерттелген, алайда соңғы онжылдықта терең оқыту алгоритмдерінің жедел дамуымен бұл сала өзінің алтын

дәуіріне енді.

2012 жылы AlexNet моделінің ImageNet бәсекесінде жеңіске жетуі компьютерлік көру тарихының бетбұрыс нүктесіне айналды. Бұл оқиға конволюциялық нейрондық желілердің (CNN) объектілерді танудағы ерекше қабілетін бүкіл ғылыми қауымдастыққа паш етті. Одан кейін VGGNet (2014), GoogLeNet/Inception (2014), ResNet (2015), DenseNet (2017) сияқты архитектуралар бірін-бірі алмастырып, дәлдік деңгейін жыл сайын жаңа белестерге жеткізді.

Объектілерді тану мен объектілерді анықтау ұғымдарын ажырата білу маңызды. Тану — суреттегі объектінің қандай класқа жататынын анықтау болса, анықтау — объектінің орнын да шектеу жақтаумен белгілеуді қамтиды. Қазіргі заманғы жүйелер осы екі міндетті де нақты уақыт режимінде орындай алатын деңгейге жетті. Деректер жиынтықтарының стандартты жиындары осы саладағы зерттеулердің негізіне айналып, модельдерді салыстыруда жалпы эталон ретінде қолданылады.

Дәстүрлі компьютерлік көру әдістері бағытталған градиенттер гистограммасы, масштабқа тәуелсіз ерекшеліктерді түрлендіру, Хаар каскадтары сияқты қолмен жасалған ерекшеліктерге сүйенетін. Бұл тәсілдер шектеулі жағдайларда жақсы нәтиже берсе де, нақты дүниенің күрделі сценарийлерінде — жарықтың өзгеруі, бір-бірін жабу, масштаб ауытқулары — тиімділігі айтарлықтай төмендеді. Терең оқытудың пайда болуы осы шектеулерді жеңіп, объектілерді тану өнімділігін жаңа деңгейге алып шықты.

Қазақстандағы ақпараттық технологиялар саласының дамуы контекстінде объектілерді тану жүйелерін зерттеу ерекше өзекті. Мемлекеттің цифрлы трансформация бағдарламалары, «Цифрлы Қазақстан» стратегиясы аясында жасанды интеллект технологияларын меңгерген мамандарға деген сұраныс жыл санап артуда. Осы мақала Қарағанды ұлттық зерттеу университетінің математика және ақпараттық технологиялар факультетіндегі зерттеу бағыттарының бірі ретінде дайындалды.

2. Негізгі архитектуралар мен алгоритмдер

2.1 Конволюциялық нейрондық желілер

Конволюциялық нейрондық желілер — кескін мен бейне өңдеу үшін арнайы жасалған терең оқыту архитектурасы. Осы желілердің негізгі ерекшелігі конволюция операциясын қолдану болып табылады: арнайы сүзгіштер кескіннің бетінен жылжып өтіп, шетті, текстураны, пішінді және т.б. ерекшеліктерді автоматты түрде үйренеді. Бұл процесс иерархиялық түрде жүреді: желінің алғашқы қабаттары қарапайым ерекшеліктерді — горизонталь және вертикаль жиектер, ал терең қабаттары күрделі концептуалды ерекшеліктерді — бет, дөңгелек, мұрын — ажыратады.

Қалдықтар желісі архитектурасы 2015 жылы Microsoft зерттеушілері ұсынған. Оның басты инновациясы — тікелей байланыстар деп аталатын байланыс тәсілі. Бұл

байланыстар градиентті жоғалту проблемасын шешіп, 100 және одан да көп қабаттан тұратын өте терең желілерді сәтті оқытуға мүмкіндік береді. Осы архитектураның әртүрлі нұсқалары объектілерді тану эталондарында бірнеше жыл бойы жетекші орын алды.

Тиімді желі — модель масштабтауға жаңаша қарайтын архитектура, 2019 жылы ұсынылған. Дәстүрлі тәсілдерде тек желінің тереңдігін немесе ені мен кіріс кескін өлшемін жеке-жеке ұлғайтса, бұл архитектура үшеуін де бірге пропорционалды түрде кеңейтеді. Бұл тәсіл дәлдік пен есептеу тиімділігінің оңтайлы арақатынасын қамтамасыз етіп, мобильді және шеткі құрылғыларда жұмыс істеуге ыңғайлы болды.

2.2 Объектілерді анықтау жүйелері: жылдам анықтау желілері мен аймақтық конволюциялық желілер

Объектілерді анықтау саласында екі негізгі парадигма қалыптасты: екі кезеңді және бір кезеңді детекторлар. Аймақтық конволюциялық нейрондық желілер және олардың жылдам нұсқалары — екі кезеңді детекторлардың өкілдері. Олар алдымен болжамды аймақтарды генерациялайды, содан соң әр аймақты жіктейді. Бұл тәсіл жоғары дәлдік беріп, соңғы нұсқалар нақты уақыт режиміне жақын жылдамдыққа жетті.

Жылдам анықтау желісі — 2016 жылы Джозеф Редмон және оның әріптестері ұсынған революциялық бір кезеңді детектор. Бұл желі кескінді бір ғана нейрондық желі арқылы өңдеп, шектеу жақтаудың координаттары мен класс ықтималдықтарын бірдей уақытта болжайды. Алғашқы нұсқадан сегізінші нұсқаға дейінгі даму барысында жылдамдық пен дәлдік бірге жақсарды. Соңғы нұсқа стандартты деректер жиынтығында 53.9% орташа дәлдік нәтижесіне қол жеткізе отырып, нақты уақытта секундына 160-тан астам кадр жылдамдығымен жұмыс істей алады.

2.3 Трансформер негізіндегі модельдер

2020 жылы Google Brain тобы ұсынған Көру трансформері архитектурасы компьютерлік көру саласына табиғи тілді өңдеуден алынған трансформер механизмін енгізді. Бұл архитектура кескінді тіркелген өлшемдегі бөліктерге бөліп, оларды сөз тіркестерімен аналогия бойынша өңдейді. Өзіне-өзі назар аудару механизмі моделге кескіннің кез-келген екі нүктесінің арасындағы байланысты ескеруге мүмкіндік береді — бұл конволюциялық нейрондық желілердің жергілікті рецептивті өрісіне қарағанда жаһандық контексті жақсы түсіндіреді.

Анықтау трансформері — трансформерді объектілерді анықтауда пайдаланған алғашқы модель, 2020 жылы ұсынылды. Ол дәстүрлі кейінгі өңдеу қадамдарының — максимумды емес тежеу, зәкір жақтаулар — қажеттілігін жойып, анықтау тапсырмасын тікелей жиынды болжау проблемасы ретінде қарастырды. Ығысқан терезелер арқылы иерархиялық трансформер иерархиялық ерекшелік карталары мен ығысқан терезелік

назар аудару механизмін қосып, есептеу шығындарын азайта отырып, объектілерді анықтау мен семантикалық сегментация тапсырмаларында озық нәтижелер орнатты.

3. Қолданбалы аспектілер және болашақ бағыттары

3.1 Негізгі қолдану салалары

Объектілерді тану жүйелері бүгінде бірқатар салалардан берік орын алды. Медицинада CNN негізіндегі жүйелер рентген, МРТ, КТ суреттерінен ісік, пневмония, диабеттік ретинопатия, тері аурулары сияқты патологияларды кейде маман дәрігерден де жоғары дәлдікпен анықтайды. Google DeepMind компаниясының AlphaFold моделі ақуыз құрылымын болжауда және PathAI жүйесінің онкологиялық патологияны анықтауда пайдалануы осы саладағы жасанды интеллект революциясының нақты мысалдары.

Автономды көлік жүргізу саласында объектілерді тану жүйелері күрделі сенсорлық жиымның маңызды бөлігі болып табылады. Tesla Autopilot, Waymo, NVIDIA DRIVE сияқты жүйелер LiDAR, радар және камера деректерін біріктіре отырып, жол белгілерін, жаяу жүргіншілерді, басқа көліктерді нақты уақытта анықтайды. Нақты уақытта жауап беру кешігуін 50 мс-тан төмен деңгейде ұстап тұру қауіпсіздік тұрғысынан маңызды талап болып есептеледі.

Қоғамдық қауіпсіздік және бейнебақылау жүйелерінде де объектілерді тану технологиялары кеңінен қолданылады. Жүздерді тану, жүріс-тұрысты талдау, ауытқушы мінез-құлықты анықтау — бұл мүмкіндіктер заманауи бейнебақылау жүйелерінің стандартты функциялары болып отыр. Алайда осы технологиялардың жаппай қолданылуы жеке өмірге қол сұғылмаушылық мәселесін де көтереді.

3.2 Деректер жиынтығы және оқыту стратегиялары

Жасанды интеллект моделінің сапасы көбінесе оқыту деректерінің сапасы мен мөлшеріне тікелей байланысты. Трансфертік оқыту тәсілі арқылы үлкен деректер жиынтығында алдын ала оқытылған модельдер арнайы тапсырмаларға бейімделіп, аз деректермен де жоғары нәтиже береді. Деректерді кеңейту — деректерді аудару, бұру, масштабтау, жарықты өзгерту арқылы жасанды түрде молайту — шағын деректер жиынтығымен жұмыс кезінде маңызды рөл атқарады.

Өзін-өзі бақылаулы оқыту және аз деректермен оқыту тәсілдері белгіленбеген деректерден пайда алуға мүмкіндік беріп, белгілеу шығындарын айтарлықтай азайтады. OpenAI мәтін-кескін моделі мәтін-кескін жұптарынан оқып, нөлдік атақ тану мүмкіндігін көрсетті — яғни бұрын ешқашан көрмеген объект класстарын да жаңа деректерсіз дұрыс тани алды. Бұл бағыт болашақта объектілерді тану жүйелерінің жалпылама қабілетін едәуір арттырады деп күтілуде.

3.3 Шектеулер және болашақ зерттеу бағыттары

Заманауи объектілерді тану жүйелерінде бірқатар шешілмеген мәселелер сақталып отыр. Біріншіден, қарсылас шабуылдар проблемасы — адам көзіне байқалмайтын шағын пиксель өзгерістері моделді толығымен алдап кете алады, бұл қауіпсіздік жүйелерінде маңызды қауіп тудырады. Екіншіден, домен ауысымы — модель оқытылған деректер мен нақты пайдалану ортасы арасындағы айырмашылық дәлдіктің айтарлықтай төмендеуіне алып келеді. Үшіншіден, интерпретабельділік мәселесі: терең желілер «қара жәшік» ретінде жұмыс істеп, шешімдерінің негізін адамға түсіндіре алмайды.

Болашақ зерттеулердің негізгі бағыттарына мыналар жатады: нейроморфтық есептеу негізіндегі аппараттық жеделдету; үш өлшемді объектілерді тану және нүкте бұлты деректерін өңдеу; көп модальді оқыту — мәтін, дыбыс және кескінді біріктіретін жүйелер; шеткі жасанды интеллект — заттар интернеті құрылғыларында тиімді жұмыс істей алатын жеңіл модельдер; сондай-ақ энергия тиімділігі жоғары, экологиялық із аз модельдер жасау.

Сондай-ақ этикалық және заңдық реттеу аспектілері өткір мәселеге айналуда. Биометриялық деректерді жинау, алгоритмдердегі алалаушылықтан туындайтын кемсіту тәуекелдері, жеке өмір құпиясы — бұл мәселелер технологиялық жетістіктермен қатар шешілуі тиіс. Еуропалық Одақтың жасанды интеллект туралы заңнамасы жоғары тәуекелді жасанды интеллект жүйелеріне, оның ішінде биометриялық тануға, арнайы талаптар қойып отыр. Қазақстан да осы тәжірибеден үлгі алып, ұлттық жасанды интеллект стратегиясын жасауда осы аспектілерді ескеруі тиіс.

Пайдаланылған әдебиеттер тізімі

1. ЛеКун И., Бенжио Й., Хинтон Г. Терең оқыту. — Табиғат журналы, 521(7553)-том, 2015. — 436-444-беттер.
2. Хэ К., Чжан С., Жэнь Ш., Сун Дж. Кескінді тануда қалдықтар арқылы терең оқыту. — Компьютерлік көру және үлгіні тану конференциясы материалдары, 2016. — 770-778-беттер.
3. Редмон Дж., Фархади А. Бірінші нұсқадан үшінші нұсқаға: жетілдірілген нақтылық пен жылдамдық. — arXiv алдын ала басылымы: 1804.02767, 2018.
4. Досовицкий А. Кескін — 16×16 сөзден тұрады: кескінді таныту үшін трансформерлер. — Оқытуды ұсыну конференциясы материалдары, 2021.
5. Карион Н. Трансформерлер арқылы объектілерді ұштан ұшқа анықтау. — Еуропалық компьютерлік көру конференциясы материалдары, 2020. — 213-229-беттер.
6. Лю З. Ығысқан терезелер арқылы иерархиялық көру трансформері. — Халықаралық компьютерлік көру конференциясы материалдары, 2021. — 10012-10022-беттер.

7. Тан М., Ле К.В. Тиімді желі: конволюциялық нейрондық желілерді масштабтауды қайта ойластыру. — Халықаралық машиналық оқыту конференциясы материалдары, 2019. — 6105-6114-беттер.

8. Рэдфорд А. Табиғи тіл бақылауынан алынатын ауыспалы көрнекі модельдер. — Халықаралық машиналық оқыту конференциясы материалдары, 2021.

9. Гудфеллоу И. Терең оқыту. — Массачусетс технологиялық институты баспасы, 2016.

10. Сейткали Б. Қазақстандағы жасанды интеллект технологияларының дамуы. — Қарағанды университетінің хабаршысы, 4-шығарылым, 2022. — 112-121-беттер.

ҚМ АА Күәлік нөмірі: **KZ45VPY00102718** — ҚР Мәдениет және Ақпарат министрлігі

© 2026 **Bilimger.kz** Ақпараттық-танымдық білім порталы. Барлық мазмұн авторлық құқықпен қорғалған.