

И.Г. Курмашев¹, А.К. Қосаяқова¹, Д.М. Калманова^{2*}, О.К. Абдирашев²

¹М.Қозыбаев атындағы Солтүстік Қазақстан университеті
150000 Қазақстан Республикасы, Петропавловск қ., Жумабаев көш., 114

²Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан
010000, Қазақстан Республикасы, Астана қ., Сатпаев көш., үй 2

*e-mail: dinara_kalmanova@mail.ru

ОБЪЕКТІНІ АНЫҚТАУ ТАПСЫРМАСЫНДА GPT ЖӘНЕ ӨЗІМІЗ ЖАСАҒАН НЕЙРОНДЫҚ ЖЕЛІ НЕГІЗІНДЕГІ МОДЕЛЬДЕРДІ САЛЫСТЫРМАЛЫ ТАЛДАУ

Аңдатпа: Соңғы жылдары нейрондық желілерді пайдалану компьютерлік көрудің әртүрлі тапсырмалары үшін Генеративті алдын ала дайындалған трансформатор (GPT) және конволюционды нейрондық желілердің (CNN) әртүрлі вариациялары сияқты модельдердің дамуымен айтарлықтай кеңейді. Бұл саладағы негізгі міндеттердің бірі кескіндердегі нысанды анықтау болып табылады. Соңғы жылдары объектілерді анықтау мәселелерін шешудің әртүрлі тәсілдері пайда болды, олар кең мағынада үш негізгі түрге бөлінеді: GPT негізіндегі модельдер, алдын-ала дайындалған модельдер және арнайы жасалған жасанды нейрондық желілер. Бұл тәсілдердің әрқайсысының бірегей артықшылықтары мен кемшіліктері бар, сондықтан әртүрлі сценарийлерде олардың сипаттамалары мен тиімділігін түсіну маңызды.

Бұл жұмыс нысанды анықтау контекстінде GPT негізіндегі модельдердің, алдын ала дайындалған модельдердің және құрастырылған жасанды нейрондық желілердің салыстырмалы талдауын ұсынады. Нысандарды анықтау компьютерлік көрудің негізгі міндеті болып табылады және қолданбалар автономды жүргізу, бақылау және медициналық бейнелеу сияқты әртүрлі өрістерді қамтиды. Зерттеу нысанды анықтау негіздерін және тиімді жүзеге асыру үшін дұрыс үлгіні таңдау маңыздылығын сипаттаудан басталады. Олардың ауқымды алдын ала дайындығының артықшылықтары жоғары есептеу талаптары мен шектеулі теңшеумен байланысты қиындықтармен салыстырылады.

Түйін сөздер: GPT, Chat-GPT, Neural networks, Chat-GPT, Object Detection, Machine Learning, Deep Learning.

Кіріспе

Нысандарды анықтау-суреттегі немесе бейнедегі нысандарды анықтау және локализациялаумен байланысты компьютерлік көру саласындағы ең маңызды міндет. Бұл технология автономды жүргізу мен бақылаудан бастап медициналық бейнелеу мен толықтырылған шындыққа дейінгі көптеген қосымшалардың негізінде жатыр. Объектілерді анықтаудың тиімді және дәл жүйелеріне сұраныс артқан сайын, осы міндеттерді шешуге қабілетті жетілдірілген модельдерге қажеттілік артып келеді [1].

Соңғы жылдары объектілерді анықтау мәселелерін шешудің әртүрлі тәсілдері пайда болды, олар кең мағынада үш негізгі түрге бөлінеді: GPT негізіндегі модельдер, алдын-ала дайындалған модельдер және арнайы жасалған жасанды нейрондық желілер. Бұл тәсілдердің әрқайсысының бірегей артықшылықтары мен кемшіліктері бар, сондықтан әртүрлі сценарийлерде олардың сипаттамалары мен тиімділігін түсіну маңызды.

GPT-3 және GPT-4 сияқты GPT негізіндегі модельдер тілдік деректерді өңдеудің әсерлі мүмкіндіктерінің арқасында айтарлықтай назар аударды. Олар ең алдымен табиғи тілге қатысты тапсырмалар үшін әзірленгенімен, олардың архитектурасы мен оқыту әдістері басқа салаларда, соның ішінде нысандарды анықтауда да қолданылуы мүмкін. Бұл модельдер жоғары өнімділік пен әмбебаптықты қамтамасыз ететін кең деректер жиынтығында Мұқият алдын ала дайындықтан пайда көреді. Дегенмен, олардың объектілерді анықтау үшін қолданылуы әлі де Есептеу ресурстары мен теңшеу қажеттіліктеріне байланысты мәселелер туындайтын дамып келе жатқан сала болып табылады.

GPT негізіндегі модельдер

Open air әзірлеген GPT (generative Pretrained Transformer) модельдері табиғи тілді өңдеуде (NLP) айтарлықтай ілгерілеуді білдіреді. Бұл модельдер transformer архитектурасына негізделген, ол адам сияқты мәтінді өңдеу және генерациялау үшін өзін-өзі оқыту механизмдеріне сүйенеді. GPT-2, GPT-3 және GPT-4 сияқты әртүрлі нұсқалардың

шығарылымы осы модельдердің әртүрлі салаларда қолдану мүмкіндіктері мен мүмкіндіктерінің артып келе жатқанын көрсетті [2].

Негізгі функциялар

1. Сәулет трансформаторлар:

– Сөйлемдегі әртүрлі сөздердің маңыздылығын бағалау үшін өзін-өзі бақылау механизмдерін қолданады.

– Модельді тиімдірек ететін параллельді деректерді өңдеуді қамтамасыз етеді.

2. Үлкен деректер жиынтығында алдын ала оқыту:

– Модельдер мәтіндік деректердің үлкен көлемінде алдын-ала дайындықтан өтеді.

– Бұл алдын-ала дайындық модельдерге грамматиканы, әлем туралы фактілерді білуге және кейбір ойлау қабілеттерін дамытуға мүмкіндік береді.

3. Көпжақты:

– Арнайы дайындықсыз көптеген тапсырмаларды орындай алады.

– Мысалдарға мәтін құру, аудару, жалпылау және сұрақтарға жауап беру жатады.

GPT модельдерінің мысалдары

1. GPT-3: – 2020 жылы шығарылған GPT-3-те 175 миллиард параметр бар, бұл оны сол кездегі ең үлкен тілдік модельдердің біріне айналдырады [4].

– Әр түрлі тақырыптар бойынша дәйекті және контекстке сәйкес мәтін құру қабілетімен танымал.

2. GPT-4:

– Соңғы нұсқасы, GPT-4, GPT-3-ті көбірек опциялар мен жетілдірілген мүмкіндіктермен жақсарттады.

– Аз қателіктер жіберіп, дәлірек жауаптар бере отырып, мәтінді түсіну және құру кезінде жақсы өнімділікті қамтамасыз етеді.

Нысанды анықтауға арналған қосымшалар

GPT модельдері, ең алдымен, тілді табиғи өңдеу тапсырмаларына арналған болса да, олардың негізінде жатқан трансформатор архитектурасы мен алдын ала оқыту әдістемелерін басқа салаларда, соның ішінде нысандарды анықтауда да қолдануға болады. GPT модельдері қалай пайдалы болуы мүмкін [5]:

1. Білімді беру:

– Кең мәтіндік деректер жиынтығымен жұмыс істеуді алдын-ала оқыту кезінде алынған білімді бастапқы негізді қамтамасыз ете отырып, объектілерді анықтау тапсырмаларына қолдануға болады.

2. Мультимодальды модельдер:

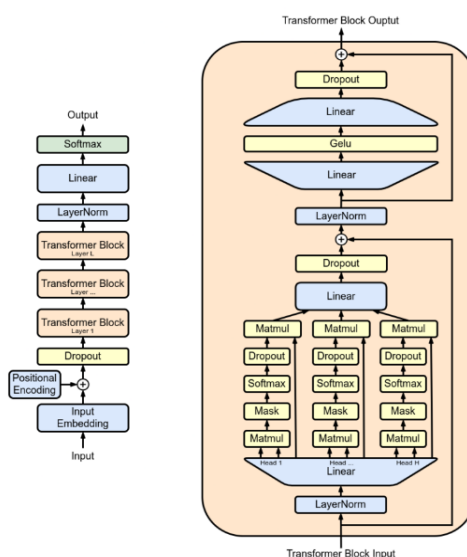
– GPT модельдерінің жаңа нұсқалары күрделі тапсырмаларды орындау үшін мәтінді, кескіндерді және басқа деректер түрлерін біріктіруге мүмкіндік беретін мультимодальды мүмкіндіктерді пайдаланады.

– Мұндай интеграция контексті түсінуді қамтамасыз ету және анықтау дәлдігін арттыру арқылы нысанды анықтауды жақсартта алады.

Генеративті жасанды интеллекттің мультимодальды модельдері технологияны дамытудың жаңа кезеңі болып табылады. Деректердің бір түрін (бір модальділік) қолдаумен шектелген дәстүрлі AI модельдерінен айырмашылығы, мультимодальдар әртүрлі форматтағы ақпаратты өңдеуге және синтездеуге бағытталған: мәтін, дыбыс, сурет. Мультимодальды AI модельдерінің дамуы технологиялық алыптармен ынталандырылады. Процестің алдыңғы қатарында OpenAI тағы да жақында шыққан ChatGPT-о шығарылымымен шықты, мұнда «o» «omni» дегенді білдіреді, яғни мультимодальды қолдау. Мультимодальды модельдердің негізгі артықшылығы-тапсырмалардың контексті мен нюанстарын толық түсіну мүмкіндігі. Мысалы, мұндай модель суреттегі нысанды танып қана қоймай, оны мәтінмен сипаттай алады, мәтінді синтезделген дауыстық дыбыстық файлға аударып алады немесе мәтіндік сұраныс бойынша бейне жасай алады. Яғни, мультимодальды шешімдер біздің өмірдің әртүрлі салаларында күрделі мәселелерді шешуге қабілетті интеллектуалды және адаптивті жүйелердің жаңа форматын құруға жақындатады. Қолданудың ең айқын бағыттары: мазмұн, сату, маркетинг, тұтынушы тәжірибесін жақсарту, бағдарламалық код жазу, зерттеу және әзірлеу (R&D), өндіріс, қаржы, персоналды басқару, заң қызметтері.

GPT-ге негізделген модельдің негізгі және әмбебап кемшіліктерінің бірі-жоғары есептеу ресурстары: GPT модельдерін оқыту және орналастыру айтарлықтай есептеу қуаты мен

жадты қажет етеді, бұл оларды кішігірім ұйымдар үшін қол жетімді етпейді. Барлық GPT модельдері Open Air есептеу ресурстарының үлкен көлемін қолдана отырып оқытылады (1 сурет) [6].



Сурет 1 – GPT негізіндегі модель құрылымы

Алдын ала дайындалған модельдер

Нысандарды анықтауға арналған алдын – ала дайындалған модельдер-бұл суреттердегі объектілерді тез және тиімді анықтауға және оқшаулауға арналған алдын-ала дайындалған нейрондық желі архитектурасы. Бұл модельдер үлкен деректер жиынтығында алдын-ала дайындалған, бұл оларды минималды қосымша оқытумен орналастыруға мүмкіндік береді. Міне, ең тиімді, пайдалануға дайын нысанды анықтау үлгілерінің кейбірі: Осы түрлердің ішінен ең жақсы модельді таңдау үшін қарапайым бәсекелестік талдау жасайық. Олардың бірнешеуін қарастырайық [7].

1. YOLO

YOLOv3 and YOLOv4:

YOLO-суреттегі нысандарды бөлектеуге мүмкіндік беретін тиімді алгоритм. Сипаттама: YOLO жылдамдығы мен дәлдігімен танымал.

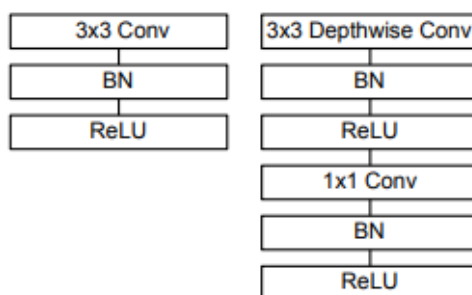
YOLOv3 және YOLOv4-анықтау жылдамдығы мен өнімділігін біріктіретін ең танымал нұсқалардың бірі. Өнімділік: жоғары дәлдікті сақтай отырып, секундына жоғары кадр жиілігін қамтамасыз етеді, бұл оны нақты уақыттағы қолданбаларға қолайлы етеді.

Артықшылықтары: нақты уақыттағы объектілерді анықтау, жоғары анықтау жылдамдығы. Пайдалану жағдайлары: автономды жүргізу, Нақты уақыттағы бақылау және робототехника.

2. MobileNet

Сипаттама: Mobile Net мобильді және ендірілген техникалық көру қолданбалары үшін оңтайландырылған CNN архитектурасын біріктіреді.

Сол жақта кәдімгі конволюциялық желі блогы, ал оң жақта MobileNet негізгі блогы салынған.



Бізді қызықтыратын желінің конволюциялық бөлігі басында 3x3 конволюциясы бар бір кәдімгі конволюциялық қабаттан және суретте оң жақта көрсетілген он үш блоктан тұрады, сүзгілер саны біртіндеп артып, тензордың кеңістіктік өлшемі төмендейді.

Бұл архитектураның ерекшелігі-max pooling қабаттарының болмауы. Оның орнына кеңістіктік өлшемді азайту үшін 2-ге тең stride параметрі бар конволюция қолданылады.

MobileNet архитектурасының екі гиперпараметрі- α (ені мультипликаторы) және өрнек (тереңдік мультипликаторы немесе ажыратымдылық мультипликаторы).

Ені мультипликаторы әр қабаттағы арналар санына жауап береді. Мысалы, $\alpha = 1$ бізге архитектураны береді, ал $\alpha = 0.25$ – әр блоктың шығысында төрт есе азайтылған арналар саны бар архитектура [8].

Ажыратымдылық мультипликаторы кіріс тензорларының кеңістіктік өлшемдеріне жауап береді. Мысалы, $\alpha = 0.5$ әр қабатқа кіруге берілген feature map биіктігі мен ені екі есе азаятынын білдіреді.

Екі параметр де желінің өлшемдерін өзгертуге мүмкіндік береді: α және этикті азайту арқылы біз тану дәлдігін төмендетеміз, бірақ сонымен бірге жұмыс жылдамдығын арттырамыз және тұтынылатын жадты азайтамыз.

Өнімділік: есептеу ресурстары шектеулі құрылғылар үшін оңтайландырылған жылдамдық пен дәлдік арасындағы тепе-теңдікті қамтамасыз етеді.

Артықшылықтары: жеңіл және тиімді, мобильді және перифериялық құрылғыларға жарамды.

Пайдалану жағдайлары: мобильді қосымшалар, интернет заттары және ендірілген жүйелер.

3. EfficientDet [9]

Сипаттама: тиімді диета EfficientNet-ті негіз ретінде пайдаланады және функцияларды біріктіру үшін жаңа bifpn желісін (Bi-directional Feature Pyramid Network) қамтиды.

Өнімділік: Аз параметрлермен және аз есептеу шығындарымен ең заманауи өнімділікті қамтамасыз етеді.

Артықшылықтары: тиімді есептеу кезінде жоғары дәлдік, масштабталатын модель архитектурасы.

Пайдалану мысалдары: жоғары өнімділік пен үнемділікті қажет ететін масштабталатын қолданбалар.

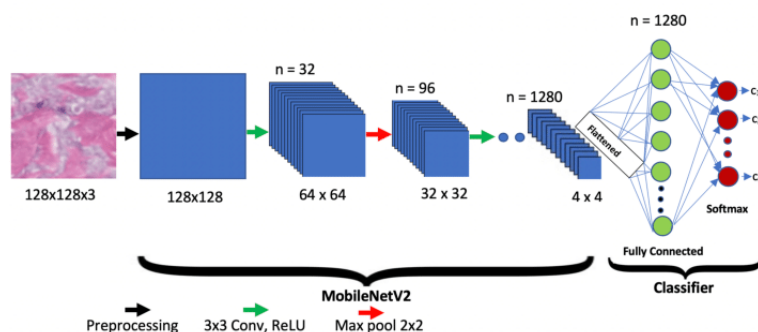
Бәсекелестік талдау

Жылдамдық пен дәлдіктің арақатынасы: YOLO және MobileNet сияқты модельдерде жылдамдыққа басымдық беріледі, бұл оларды нақты уақыт режимінде жұмыс істеуге өте ыңғайлы етеді. Олардан айырмашылығы, олар объектілерді егжей-тегжейлі және дәл анықтауды қажет ететін тапсырмалар үшін жарамды.

Ресурстарды пайдалану тиімділігі: MobileNet шектеулі есептеу ресурстары бар құрылғылар үшін оңтайландырылған, бұл оны мобильді және ендірілген жүйелер үшін тамаша шешім етеді. YOLOv4 сонымен қатар есептеу жылдамдығы мен тиімділігі арасындағы жақсы тепе-теңдікті қамтамасыз етеді [10].

Масштабтау: Efficient data масштабталатын архитектурасы оны аз қуатты құрылғылардан бастап жоғары өнімді жүйелерге дейін әртүрлі қолданбаларда пайдалануға мүмкіндік береді.

Осы талдау барысында біз мобильді желі моделі ең қолайлы екенін көреміз. Осылайша, пайдалануға дайын модельдердің ішінен Mobile Net ең өкілі ретінде таңдалды. Mobile net архитектурасы төмендегі 2 суретте көрсетілген.



Сурет 2 – Мобильді желі архитектурасының визуалды көрінісі

Меншікті жасалған модельдер

Нысандарды анықтауға арналған модельдердің өзі нейрондық желілерді нөлден бастап жобалау мен оқытуды немесе нақты талаптарға сәйкес бар архитектураларды дәл баптауды қамтиды. Бұл тәсіл белгілі бір қолданбаның бірегей қажеттіліктерін ескере отырып, максималды икемділік пен оңтайландыруды қамтамасыз етеді. Жеке модель жасау үшін бірнеше қадамдарды орындау қажет:

- 1) проблеманы анықтау (тиісті міндет қою үшін)
- 2) модель архитектурасын әзірлеу
- 3) модельді оқыту және тестілеу

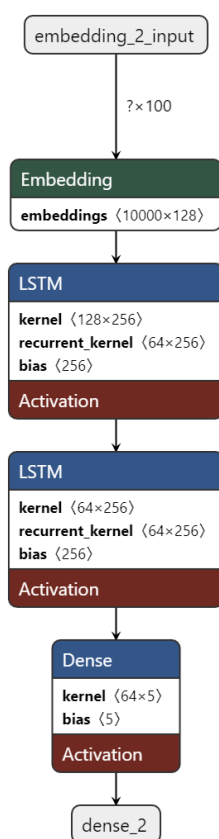
Мәселені анықтау

Соңғы жылдары объектілерді анықтау компьютерлік көру саласындағы маңызды міндетке айналды, оны қолдану автономды көліктерден бейнебақылау және медициналық бейнелеу жүйелеріне дейін. Нысандарды анықтау суреттегі немесе бейнедегі нысандарды анықтау мен орналастыруды, сондай-ақ оларды алдын ала анықталған санаттар бойынша жіктеуді қамтиды. Бұл тапсырманың мақсаты-суреттегі көптеген нысандарды дәл анықтауға және жіктеуге қабілетті сенімді нейрондық желі моделін жасау.

Нысандарды анықтау үшін нейрондық желіні құру міндеті бірнеше қиындықтарды тудырады:

Деректерді аннотациялау: әр объект үшін шектеу шеңберлері мен сынып белгілерімен үлкен деректер жиынтығына аннотация жасау көп уақытты қажет ететін және уақытты қажет ететін процесс.

Модельдің күрделілігі: YOLO (сіз тек бір рет қарайсыз) және жылдамырақ R-CNN (аймақтық конволюциялық нейрондық желілер) сияқты объектілерді анықтау модельдері өте күрделі және оқыту мен шығару үшін айтарлықтай есептеу ресурстарын қажет етеді (3 сурет).



Сурет 3 – Конволюциялық нейрондық желілерге негізделген іске асырылған модельдің құрылымы

Нақты уақыттағы өнімділік: модельдің суреттерді нақты уақыт режимінде өңдей алатындығына көз жеткізу автономды жүргізу және бейнебақылау сияқты қосымшалар үшін өте маңызды.

Дәлдік: әртүрлі жарық жағдайларында және Фонда шағын нысандарды, қабаттасатын нысандарды және нысандарды анықтауда жоғары дәлдікке қол жеткізу.

Нысандарды анықтау үшін нейрондық желіні құру-бұл әртүрлі факторларды, соның ішінде деректер сапасын, модель архитектурасын және есептеу тиімділігін мұқият есепке алуды қажет ететін күрделі, бірақ пайдалы тапсырма.

Құрылымдық тәсілге сүйене отырып, нақты уақыт режимінде дәлдік пен өнімділік талаптарына жауап беретін модель жасауға болады, бұл компьютерлік көру саласына айтарлықтай үлес қосады.

Модель архитектурасын әзірлеу.

Нысандарды анықтауға арналған нейрондық желіні жобалау сәйкес архитектуруны таңдауды, деңгейлерді анықтауды және оқу құбырын іске асыруды қамтиды. Бұл бөлімде есептеу дәлдігі мен тиімділігі арасындағы тепе-теңдікті қамтамасыз ететін танымал CNN архитектурасына негізделген нейрондық желі моделінің егжей-тегжейлі жобасы берілген. Архитектура-бұл бір бағалауда толық кескіндерге негізделген кластардың шектеу шеңбері мен ықтималдығын тікелей болжайтын конволюциялық нейрондық желі. Архитектура бірнеше конволюциялық қабаттардан тұрады, содан кейін толық байланысқан қабаттар. Төменде іске асырылған нейрондық желі архитектурасының егжей-тегжейлі жобасы берілген:

Салыстырмалы талдау

Нысанды анықтау саласында дұрыс модельді таңдау дәлдік пен есептеу тиімділігі арасындағы қажетті тепе-теңдікке қол жеткізу үшін өте маңызды. Бұл кестеде объектілерді анықтаудың кеңінен қолданылатын төрт моделінің салыстырмалы талдауы келтірілген: GPT негізіндегі модельдер, мобильді желі және іске асырылған модель. Салыстыру орташа дәлдікті (mAP), секундына кадрлар санын (FPS), шығу уақытын және дәлдік пен жылдамдық арасындағы жалпы тепе-теңдікті қоса алғанда, негізгі өнімділік көрсеткіштеріне негізделген. Кестеде дәлдік пен есептеу тиімділігінің арақатынасына негізделген нақты қажеттіліктер үшін ең жақсы үлгіні анықтауға көмектесетін төрт нысанды анықтау үлгісін көрнекі салыстыру берілген. Бұл талдау автономды көліктер мен бақылау жүйелерінен бастап нақты уақыттағы кескіндерді өңдеуге және толықтырылған шындыққа дейін әртүрлі қолданбаларда негізделген шешімдер қабылдау үшін қажет [11].

Кесте 1 – Модель түрлерінің көрсеткіштерін салыстыру

Модель	Орташа дәлдік (mAP)	Секундына кадр жиілігі (FPS)	Шығару уақыты (мс)	Дәлдік және Жылдамдықромасы
GPT негізіндегі модель	37-40%	32	200	Жоғары дәлдік, төмен жылдамдық
MobileNet	33%	25-30	40	Дәлдік пен жылдамдықтың жақсы балансы
Жасалған модель	31%	26	30	Жақсы жылдамдық, орташа дәлдік

Төмендегі 1 кестеде объектілерді анықтаудың әртүрлі модельдеріне қажетті есептеу қуатының салыстырмалы шолуы келтірілген. Осы талаптарды түсіне отырып, тәжірибешілер қолда бар Есептеу ресурстарына және оларды қолданудың нақты қажеттіліктеріне сүйене отырып, қандай модельді қолдану керектігі туралы негізделген шешімдер қабылдай алады. Бұл талдау модельдің күрделілігін практикалық іске асыру мәселелерімен сәйкестендірудің маңыздылығын көрсетеді (2 кесте).

Кесте 2 – Есептеу қуатын талдау кестесі

Модель	Есептеу талаптары	О уақыты	Шығуға арналған аппараттық құралдар
GPT based model	Жоғары	ұзақ (бірнеше күн)	Қуатты графикалық процессорлар (мысалы, NVIDIA Tesla V100)
MobileNet architecture	Орташа	Орташа (1-2 күн)	Тұтынушылық графикалық процессорлар (мысалы, NVIDIA GTX 1080 Ti)
Implemented model	Орташа	Орташа	Тұтынушы класындағы Орташа Орташа графикалық процессорлар

Бұл бәсекеге қабілетті талдау өнімділік көрсеткіштері мен өңдеу қуатына қойылатын талаптарға негізделген объектілерді анықтаудың әртүрлі үлгілерін жан-жақты салыстыруды қамтамасыз етеді. Талданған модельдердің қатарына жылдам RCN, YOLOv3, SSD, Retina Net, GPS негізіндегі модельдер және mobilenet архитектурасы жатады. Мұнда негізгі қорытындылардың қысқаша мазмұны берілген: сонымен қатар, бұл талдау таңдалған модельдің мақсатты қолданбаның нақты мақсаттары мен шектеулеріне сәйкестігін қамтамасыз ете отырып, өнімділік көрсеткіштері мен есептеу талаптарына негізделген нысандарды анықтаудың ең қолайлы үлгісін таңдау бойынша нұсқаулық ретінде қызмет етеді[12].

Кесте 3 – Бәсекелестік талдаудың қысқаша мазмұны

Модель	Лучшее решения для
GPT негізіндегі модель	Дәлдік өте маңызды және айтарлықтай есептеу ресурстарын қажет ететін қосымшалар. Нақты уақыттағы қосымшалар үшін қолайлы емес.
MobileNet	Дәлдік пен жылдамдық арасындағы жақсы тепе-теңдігі бар нақты уақыттағы қосымшалар. Орта деңгейлі жабдыққа орналастыру үшін қолайлы.
Жасалған модель	Нақты уақыттағы қосымшаларда жылдамдыққа сәл дәлдікпен басымдық беріледі. Тұтынушы класындағы графикалық процессорлар үшін және басқа көздерде пайдалану үшін өте қолайлы.

Бұл 3 кестелер объектілерді анықтаудың төрт танымал моделін олардың өнімділік көрсеткіштері мен есептеу талаптарына баса назар аудара отырып, жан-жақты салыстыруды ұсынады. Бұл талдау дәлдік пен жылдамдықтың, сондай-ақ оқыту мен нәтижелерді шығаруға қажетті аппараттық құралдардың арасындағы байланысты түсінуге көмектеседі, бұл қолданбаның нақты қажеттіліктеріне сүйене отырып, қандай модельді қолдану керектігі туралы негізделген шешімдер қабылдауға мүмкіндік береді.

Қорытынды

Нақты тапсырма мен талаптарға байланысты модельдердің әрқайсысының өзіндік артықшылықтары мен кемшіліктері бар. GPT модельдері жоғары икемділік пен масштабтауды қамтамасыз етеді, бірақ бейімделу үшін айтарлықтай ресурстарды қажет етуі мүмкін. Меншікті нейрондық желілер объектілерді анықтаудың нақты тапсырмаларында тиімдірек және дәлірек болуы мүмкін, бірақ теңшеу және оқыту үшін көп уақыт пен күш қажет.

Егжей-тегжейлі салыстыру үшін модельдер үшін бірдей деректер жиынтығы мен бағалау көрсеткіштерін қолдана отырып, эксперименттік зерттеу жүргізу ұсынылады, бұл олардың өнімділігі мен объектілерді анықтау тапсырмасындағы дәлдігін дәлірек бағалауға мүмкіндік береді.

Әдебиеттер тізімі

1. Improving Language Understanding by Generative Pre-Training / A. Radford et al // 2018. Available at: <https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf>.
2. Language Models are Few-Shot Learners / T. Brown et al // In: Advances in Neural Information Processing Systems. 2020. Available at: <https://arxiv.org/abs/2005.14165>.
3. Girshick R. Fast R-CNN / R. Girshick In: Proceedings of the IEEE International Conference on Computer Vision (ICCV) (2015). Available at: <https://arxiv.org/abs/1504.08083>.
4. YOLOv3: An Incremental Improvement / J. Redmon, et al. 2018. Available at: <https://arxiv.org/abs/1804.02767>.
5. SSD: Single Shot MultiBox Detector / W. Liu, et al // In: European Conference on Computer Vision (ECCV). 2016. Available at: <https://arxiv.org/abs/1512.02325>.
6. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale / A. Dosovitskiy et al // In: International Conference on Learning Representations (ICLR). 2021. Available at: <https://arxiv.org/abs/2010.11929>.
7. End-to-End Object Detection with Transformers / N. Carion, et al // In: European Conference on Computer Vision (ECCV) (2020). Available at: <https://arxiv.org/abs/2005.12872>.
8. A Comparative Study of Object Detection Algorithms in Various Image Datasets / Y. Liu et al // In: Journal of Visual Communication and Image Representation. 2019. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S1047320319300536>.

9. Object Detection with Deep Learning: A Review / Z. Zhao et al // In: IEEE Transactions on Neural Networks and Learning Systems. 2019. Available at: <https://ieeexplore.ieee.org/document/8662474>
10. Goodfellow I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville // MIT Press. 2016. Available at: <https://www.deeplearningbook.org/>.
11. Szeliski R. Computer Vision: Algorithms and Applications / R. Szeliski // Springer. 2010. Available at: <https://szeliski.org/Book/>.
12. Attention is All You Need / A. Vaswani et al // In: Advances in Neural Information Processing Systems (NIPS). 2017. Available at: <https://arxiv.org/abs/1706.03762>.

И.Г. Курмашев¹, А.К. Қосаяқова¹, Д.М. Калманова^{2*}, О.К. Абдирашев²

¹Северо-Казахстанский университет имени Манаша Козыбаева,
150000, Республика Казахстан, г. Петропавловск, ул. Магжана Жумабаева 114

²Евразийский национальный университет имени Л.Н.Гумилева,
010000, Республика Казахстан, г.Астана, ул.Сатпаева дом 2

*e-mail: dinara_kalmanova@mail.ru

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДЕЛЕЙ, ОСНОВАННЫХ НА GPT И СОБСТВЕННОЙ СОЗДАННОЙ НЕЙРОННОЙ СЕТИ В ЗАДАЧЕ ОБНАРУЖЕНИЯ ОБЪЕКТОВ

В последние годы применение нейронных сетей значительно расширилось за счет разработки моделей, таких как Generative Pre-trained Transformer (GPT) и различных вариаций сверточных нейронных сетей (CNN) для различных задач машинного зрения. Одной из ключевых задач в этой области является обнаружение объектов на изображениях.

В этой статье представлен сравнительный анализ моделей на основе GPT, предварительно обученных моделей и созданных искусственных нейронных сетей в контексте обнаружения объектов. Обнаружение объектов является ключевой задачей в компьютерном зрении, а приложения охватывают различные области, такие как автономное вождение, наблюдение и медицинская визуализация. Исследование начинается с изложения основ обнаружения объектов и важности выбора правильной модели для эффективной реализации.

Преимущества их обширной предварительной подготовки сопоставляются с проблемами, связанными с высокими вычислительными требованиями и ограниченной настройкой.

Ключевые слова: GPT, Chat-GPT, Neural networks, Chat-GPT, Object Detection, Machine Learning, Deep Learning.

I. Kurmashev¹, A.K. Kosayakova¹, D.M. Kalmanova^{2*}, O. Abdirashev²

¹Manash Kozybayev North Kazakhstan university,
150000, Republic of Kazakhstan, Petropavlovsk, Magzhan Zhumabayev St. 114

²L.N. Gumilyov Eurasian National University,
Republic of Kazakhstan, Astana, Satpayev street, building 2

*e-mail: kalmanova@mail.ru

A COMPARATIVE ANALYSIS OF MODELS BASED ON GPT AND ITS OWN CREATED NEURAL NETWORK IN THE PROBLEM OF OBJECT DETECTION

In recent years, the use of neural networks has expanded significantly through the development of models such as Generative Pre-trained Transformer (GPT) and various variations of convolutional neural networks (CNN) for various machine vision tasks. One of the key tasks in this area is the detection of objects in images. This article presents a comparative analysis of GPT-based models, pre-trained models and created artificial neural networks in the context of object detection. Object detection is a key task in computer vision, and applications cover various fields such as autonomous driving, surveillance, and medical imaging. The study begins by outlining the basics of object detection and the importance of choosing the right model for effective implementation. The advantages of their extensive pre-training are juxtaposed with the challenges associated with high computing requirements and limited customization.

Key words: GPT, Chat-GPT, Neural networks, Chat-GPT, Object Detection, Machine Learning, Deep Learning.

Авторлар туралы мәліметтер

Ильдар Гусманович Курмашев – PhD, қауымдастырылған профессор, М. Қозыбаев атындағы Солтүстік Қазақстан университеті, Қазақстан; e-mail: ikurmashev@ku.edu.kz. ORCID: <https://orcid.org/0000-0001-9872-7483>.

Ақнұр Көптілеуқызы Қосаяқова – М. Қозыбаев атындағы Солтүстік Қазақстан университеті. 2 курс докторанты, Қазақстан; e-mail: aknur_ast_enu@mail.ru. ORCID: <https://orcid.org/0000-0004-5435-5509>.

Динара Мирзабековна Калманова* – педагогика ғылымдарының кандидаты, Л.Н. Гумилев атындағы Еуразия ұлттық университетінің «Ғарыштық техника және технологиялар» кафедрасының доцент м.а., Астана, Қазақстан; e-mail: dinara_kalmanova@mail.ru. ORCID: <https://orcid.org/0000-0001-5977-8448>.

Өмірзақ Көптілеуұлы Әбдірашев – PhD, Л.Н. Гумилев атындағы Еуразия ұлттық университетінің «Ғарыштық техника және технологиялар» кафедрасының доцент м.а., Астана, Қазақстан; e-mail: omeke_92@mail.ru. ORCID: <https://orcid.org/0000-0001-7621-5444>.

Сведения об авторах

Ильдар Гусманович Курмашев – PhD, ассоциированный профессор, Северо-Казахстанский университет имени Манаша Козыбаева, г. Петропавловск, Казахстан; e-mail: ikurmashev@ku.edu.kz. ORCID: <https://orcid.org/0000-0001-9872-7483>.

Акнур Коптилеуовна Косаякова – Северо-Казахстанский университет имени Манаша Козыбаева, г. Петропавловск, Казахстан; e-mail: aknur_ast_enu@mail.ru. ORCID: <https://orcid.org/0000-0004-5435-5509>.

Динара Мирзабековна Калманова* – кандидат педагогических наук, и.о. доцент кафедры «Космическая техника и технологии», Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан; e-mail: dinara_kalmanova@mail.ru. ORCID: <https://orcid.org/0000-0001-5977-8448>.

Омирзақ Коптилеуович Абдірашев – PhD, и.о. доцент кафедры «Космическая техника и технологии», Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан; e-mail: omeke_92@mail.ru. ORCID: <https://orcid.org/0000-0001-7621-5444>.

Information about the authors

Ildar Gusmanovich Kurmashev – PhD, Associate Professor, Manash Kozybayev North Kazakhstan university, Kazakhstan; e-mail: ikurmashev@ku.edu.kz. ORCID: <https://orcid.org/0000-0001-9872-7483>.

Aknur Koptileuovna Kossayakova – Manash Kozybayev North Kazakhstan university, Kazakhstan; e-mail: aknur_ast_enu@mail.ru. ORCID: <https://orcid.org/0000-0004-5435-5509>.

Dinara Mirzabekovna Kalmanova* – Candidate of Pedagogical Sciences and Acting Associate Professor of the Department of «Space Engineering and Technology» at the Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan; e-mail: dinara_kalmanova@mail.ru. ORCID: <https://orcid.org/0000-0001-5977-8448>.

Omirzak Koptileuovich Abdirashev – PhD, Acting Associate Professor of the Department of «Space Engineering and Technology» at the Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan; e-mail: omeke_92@mail.ru. ORCID: <https://orcid.org/0000-0001-7621-5444>.

Редакцияға енуі 11.06.2025

Өңдеуден кейін түсуі 15.08.2025

Жариялауға қабылданды 18.08.2025

[https://doi.org/10.53360/2788-7995-2025-3\(19\)-12](https://doi.org/10.53360/2788-7995-2025-3(19)-12)

MPHTI: 20.23.27



А.А. Бондарович¹, А. Мәуліт^{2*}, И.А. Очередько³, Ж.З. Жантасова³

¹Алтайский государственный университет,
Российская Федерация, Алтайский край, г. Барнаул, пр. Ленина, 61

²Шәкәрім университеті,
071412, Республика Казахстан, г. Семей, ул. Глинки, 20 А

³Восточно-Казахстанский университет им. С.Аманжолова,
Республика Казахстан, г. Усть-Каменогорск, ул. Казахстан 55.

*e-mail: maulit.almas@gmail.com

ТЕХНОЛОГИИ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ В ОЦЕНКЕ ПОЧВЕННОЙ ВЛАЖНОСТИ НА ТЕРРИТОРИИ, ПОДВЕРЖЕННОЙ ПАВОДКАМ: ОБЗОР

Аннотация: Почвенная влажность является одним из ключевых гидрологических параметров, напрямую влияющих на формирование поверхностного стока, паводков и оползней. Особенно важную роль она играет в периоды весеннего снеготаяния и экстремальных метеоосадков, когда насыщенность почвы влагой достигает критических уровней. В условиях глобального изменения климата и участившихся экстремальных погодных явлений возрастает потребность в оперативном и точном мониторинге влажности почвы, особенно для регионов, подверженных паводкам, таких как Восточно-Казахстанская область. В статье представлен