

АҚПАРАТТЫҚ ЖҮЙЕЛЕР
ИНФОРМАЦИОННЫЕ СИСТЕМЫ
INFORMATION SYSTEMSDOI 10.51885/1561-4212_2023_2_31
MFTAA 28.23.15**Е.Н. Әмірғалиев^{1,3}, И.Н. Бөкенова^{1,2}, Ғ.С. Бөкенов², Ч.А. Кеншимов³**¹Халықаралық ақпараттық технологиялар университеті, Алматы қ., Қазақстан
E-mail: amir_ed@mail.ru*²Алматы технологиялық университеті, Алматы қ., Қазақстан

E-mail: ibukenowa@mail.ru

E-mail: gani1212@bk.ru

³Есептеуіш және ақпараттық технологиялар институты, Алматы қ., Қазақстан

E-mail: kenchimov.chingiz@gmail.com

**БЕЙНЕ АРҚЫЛЫ ЗОРЛЫҚ-ЗОМБЫЛЫҚ ҚОЗҒАЛЫСТАРЫН
ТАНУ БОЙЫНША БАҒДАРЛАМАЛЫҚ ШЕШІМДЕР****ПРОГРАММНЫЕ РЕШЕНИЯ ПО РАСПОЗНАВАНИЮ
НАСИЛЬСТВЕННЫХ ДВИЖЕНИЙ ПО ВИДЕО****SOFTWARE SOLUTIONS FOR THE RECOGNITION
VIOLENT MOVEMENTS BY VIDEO**

Аңдатпа: Бұл мақалада зорлық-зомбылықты танудың заманауи тәсілдері қарастырылып, оларды іске асыратын алгоритмдердің тану дәлдігі бағаланады. Қазіргі уақытта сенсорлық және визуалды технологиялардың дамуы арқасында адамның зорлық-зомбылық әрекеттерін тану жүйелері танымал болды. Агрессивтілік мәселесі, оның ішінде балалардың проблемасы көптен бері психологтардың назарын аударуда. Зорлық-зомбылықты қолдану әдістерінің арасында бірнеше жеке ерекшеліктер бар: машиналық оқытуды қолдана отырып зорлық-зомбылықты анықтау әдістері, SVM көмегімен зорлық-зомбылықты анықтау әдістері және терең оқытуды қолдана отырып зорлық-зомбылықты анықтау әдістері. Ең танымал модельдер – бақыланбайтын оқытуға арналған терең сенім желісі және бақыланатын оқытуға арналған конвульсиялық нейрондық желі. Бұл зерттеудің мақсаты – бейнедегі зорлық-зомбылық қозғалыстарын танудың заманауи бағдарламалық шешімдерін қарастыру. Жүргізілген жұмыс барысында агрессияны автоматты түрде анықтау үшін машиналық оқыту және жасанды интеллект алгоритмдеріне талдау жасалды.

Түйін сөздер: адамның іс-әрекетін тану; терең оқыту; тірек векторлық машина (SVM); жасанды нейрондық желі (ANN); кері байланысты нейрондық желілер.

Аннотация: В данной статье рассматриваются современные подходы распознавания насильственных действий человека, а также анализированы точности алгоритмов распознавания для данных методов. В современном мире благодаря развитию сенсорных и визуальных технологий, системы распознавания насильственных действий человека стали популярны. Проблема агрессивности, в том числе детской, давно привлекает внимание психологов. Среди методов с использованием насилия выделяется несколько отдельных: Методы обнаружения насилия с использованием машинного обучения, методы обнаружения насилия с использованием SVM и методы обнаружения насилия с использованием глубокого обучения. Самыми популярными моделями являются глубокая сеть доверия для неконтролируемого обучения и свёрточная нейронная сеть для контролируемого обучения. Цель данного исследования рассмотреть современные программные решения по распознаванию насильственных движений по видео. В ходе проделанной работы был сделан анализ существующих алгоритмов машинного обучения и искусственного интеллекта для автоматического выявления агрессии.

Ключевые слова: Распознавание действий человека. Глубокое обучение, машина опорных векторов (SVM), искусственная нейронная сеть (ANN), рекуррентные нейронные сети.

Abstract: This article discusses modern approaches to recognizing violent actions of a person. The recognition accuracies for these methods will also be presented. In the modern world, thanks to the development of sensory and visual technologies, systems for recognizing violent human actions have become popular. The problem of aggressiveness, including that of children, has long attracted the attention of psychologists. Among the methods using violence, a few stand out: Violence detection methods using machine learning, violence detection methods using SVM, and violence detection methods using deep learning. The most popular models are the deep belief network for unsupervised learning and the convolutional neural network for supervised learning. The purpose of this study is to consider modern software solutions for recognizing violent movements from video. In the course of the work done, an analysis was made of existing machine learning and artificial intelligence algorithms for the automatic detection of aggression.

Keywords: Recognition of human actions. Deep learning, Support vector machine (SVM), artificial neural network (ANN), recurrent neural networks.

Kіpіcne. Психоэмоционалды жағдайды анықтаудың автоматтандырылған жүйелерін жасау еңбек пен уақыт шығынын азайтудың өзекті қажеттілігіне айналды. Видеода агрессияны анықтау қиын міндет, өйткені агрессияны анықтау анық емес және белгісіз болуы мүмкін.

Бұл жұмыста агрессияны анықтаудың маңызды аспектілерінің бірі қарастырылады, ол зорлық-зомбылықты тану және анықтау болып табылады.

Бұл зерттеудің мақсаты – бейнедегі зорлық-зомбылық қозғалыстарын танудың заманауи бағдарламалық шешімдерін қарастыру.

Машиналық оқытуды қолдана отырып, қолданылатын зорлық-зомбылықты анықтау әдістерінде ең жоғары дәлдікке қол жеткізілді ме?

Әрбір үшінші қазақстандық оқушы қорлық пен зорлыққа тап болады. Еліміздің үш өңіріндегі бес мектепте жүргізілген мониторинг деректері осындай.

Сауалнамаға қатысқан балалардың 50 %-дан астамы мектепте және сыныпта қорқыту бар деп жауап берді. Оқушылардың 33,6 %-ы кибербуллингтің құрбаны болды деп мәлімдейді. Көбінесе бұл телефон арқылы, мессенджерлер немесе әлеуметтік желілер арқылы қорлайтын хабарламалар алған қыздар [14].

Балалардың жартысына жуығы орта мектеп оқушылары қорқытуға жүгінетінін мойындады. Екінші орында – құрдастар. Респонденттердің 10,3 %-ы мұғалімдерді қудалаудың бастамашысы, ал 5,6 %-ы ата-аналары мен аға-інілері деп атады. 4-7 сынып оқушылары жоғары сынып оқушылары тарапынан қорлауға көбірек ұшырайды.

Қорқытудың салдары жәбірленушілер үшін де, бұзақылар үшін де ауыр болуы мүмкін. Жәбірленуші депрессияны, соматикалық ауруларды, өмірге қанағаттанбауды және оқиғадан 3 жыл өткен соң қорқыту құрбандарының колледжге түсу жиілігін төмендетінін анықтады.

Канадалық зерттеу жасөспірімдер арасындағы суицид және қорқыту мен жыныстық қудалаудың байланысын зерттеді. Интернеттегі және оффлайн режиміндегі қайталанатын қорлау, бұл PTSD симптомдарын және ананың қолдауының жеткіліксіздігін тудырды, 12-18 жас аралығындағы жасөспірім қыздарда белсенді суицидтік ойларға ықпал етті, ал жыныстық қудалау суицидпен байланысты емес.

Сол сияқты, әйел жынысы мен суицидтік идеялар арасындағы байланыс Қытайдың 23 392 мектеп оқушыларының үлгісіндегі зерттеуінде расталды, онда суицидтік идеялар мен әрекеттермен агрессивті мінез-құлық пен жыныстың әсері арасындағы байланыс зерттелді. Жәбірленушілер, бұзақылар мен бұзақылардың құрбандарында бейтараптарға қарағанда суицидтік ойлар қаупі жоғары болды; ұқсас бірлестіктерді агрессивті мінез-құлық пен суицид әрекеттері арасындағы бірлестіктерден табуға болады. Әрі қарай

стратификациялық талдау көрсеткендей, басқаларды қорлау немесе қорқыту, суицидтік идеялар мен суицид әрекеттері ер балаларға қарағанда қыздарда біршама күшті болды.

Эмоцияны немесе зорлық-зомбылықты анықтау үшін қолданылатын модельдеу әдістерін таяз және терең деп жіктеуге болады.

Таяз модельдеу әдістері – бұл объектілерді өздігінен шығара алмайтын әдістер [1]. Қолмен жасалатын әдістермен алынған заттарды жіктеу үшін таяз модельге беру керек. Таяз модель ретінде тірек векторлық машина (SVM), бір жасырын қабаты бар жасанды нейрондық желі (ANN) және т.б. сол сияқты жіктеуіштерді қолдануға болады. Мұндай әдістердің басты кемшілігі – олар үлгінің өзгеруіне автоматты түрде бейімделмейді.

Таяз модельдерден айырмашылығы, терең үлгілердің көпшілігі бөлек функцияларды шығаруды қажет етпейді, өйткені олар берілген деректерден өз ерекшеліктерін үйренетін және олардың негізінде оларды жіктейтін мүмкіндіктерді оқыту әдісіне негізделген. Сонымен қатар, оқуды аяқтаудан басқа, алынған функцияларды SVM және басқа да таяз модельдер жіктеуіштері үшін кіріс ретінде беруге болады. Терең модельдерді жүзеге асырудың тағы бір тәсілі – қолмен жасалған сипаттамалардың дескрипторларынан белгілерді қолдану және оларды терең классификаторға беру. Бұл модельдер бақыланатын және бақыланбайтын оқыту әдістерімен жұмыс істейді. Белгіленбеген деректермен жұмыс істесе де, олар үлкен көлемде мәліметтер мен есептеу қуатын қажет етеді.

Компьютерлік көруде әрекетті тану зерттеудің маңызды саласына айналууда. Зорлық-зомбылық немесе төбелес сияқты тапсырмалар салыстырмалы түрде аз зерттелген, бірақ түрмелер, психиатриялық палаталар немесе мектептер сияқты көптеген бейнебақылау сценарийлерінде пайдалы болуы мүмкін. Олардың кең практикалық мүмкіндіктері зорлық-зомбылық немесе күрес детекторларының дамуына қызығушылық тудырады.

Әдеби шолу. Senst T. аз шығынды және табиғи жолмен төбелесті анықтау үшін қозғалысты талдауға негізделген тәсілді ұсынды [17]. Автор төбелесті анықтау үшін екі тәсілді қолданады. Біріншісі – белгілер жиынтығын тудыратын екі деңгейлі статистикалық агрегация. Оптикалық ағынның векторларын есептеу арқылы кадрлар сериясынан қозғалыс пиксельдері, содан кейін қозғалыс аймақтары алынады. Кедергілерді жойғаннан кейін ғана оптикалық ағынның суреттері векторлардың сипатына қарай жіктеледі. Осыдан кейін қозғалыс статистикасы тану үшін қолданылатын белгілер жиынтығын алу үшін жіктелген түрлерге сәйкес есептеледі. Қолданылатын екінші тәсіл – көрнекі сөздер жиынтығын жасау үшін қолданылатын «сөз сөмкесі». Содан кейін гистограмма ұрысты анықтау үшін бейнені сипаттау үшін вектор ретінде визуалды сөздердің бүкіл жиынтығында қолданылады. Эксперименттер нақты жекпе-жекке негізделген бейнежазбаларды қолдану арқылы жүзеге асырылады және эксперимент нәтижелері ұсынылған әдіс MoSIFT дескрипторларына негізделген қолданыстағы әдістерден асып түсетінін көрсетеді.

Келесі жұмыста [18] күдікті әрекеттерді анықтау мәселесін шешу үшін бейнеклиптердегі әртүрлі қалыптан тыс әрекеттерді автоматты түрде анықтайтын әдіс ұсынылады. Бұл жүйе үш негізгі кезеңді қамтиды: қозғалатын нысанды анықтау, нысанды қадағалау және әрекетті тану үшін мінез-құлықты түсіну. Алдын ала өңдеудің бірінші кезеңінде қозғалатын заттар анықталып, шу жойылады. Содан кейін объектілерді шығарып алу процесі бағыт, жылдамдық, өлшемдер және центроид сияқты негізгі объектілерді анықтау үшін қолданылады. Алынған мүмкіндіктер бейне кадрлардағы нысандарды бақылауға көмектеседі. Соңғы кезеңде ережелерге негізделген жіктеу әдісі бейнедегі әрекеттерді жіктеу үшін қолданылады және егер қандай да бір күдікті әрекет анықталса, ол дабыл шығарады.

Адамдар көп жиналатын жерлерде зорлық-зомбылықты бақылаудың күрделі мәселесін

шешу үшін нақты уақыт режимінде зорлық-зомбылықты анықтауға инновациялық тәсіл ұсынылады [19]. Ұсынылған әдіс уақыт бойынша ағын векторларының шамасының өзгеру статистикасын ескереді, екі өзара байланысты, бірақ әртүрлі міндеттер қарастырылады: зорлық-зомбылықты жіктеу және зорлық-зомбылықты анықтау. Негізгі мақсат – бұл өзгеріс болған сәттен бастап ең аз кідіріспен зорлық-зомбылықтан зорлық-зомбылықсыз мінез-құлыққа өзгерісті анықтау.

Бақылау жүйесінде күдікті әрекеттерді анықтау маңызды рөл атқарады. Ғалымдар арасында сенімді жұмыс істей алатын бақылау жүйесіне шұғыл қажеттілік бар. Сондықтан автоматты түрде анықтайтын нақты уақыттағы бейнебақылау жүйесі үшін жаңа құрылым ұсынылады [20]. Жұмыс үш кезеңге бөлінеді. Алдын ала өңдеу фазасы адамның қалыптан тыс белсенділігін анықтауды және мазмұнға негізделген кескіндерді шығару фазасын қамтиды. Алдын ала өңдеу кезеңінде барлық студенттер оқу курсы бастамас бұрын тіркелуі керек және тіркелу жеке деректерді жинауды қамтиды және студент студенттік билетті жасау үшін өз фотосуретін ұсынуы керек. Ұсынылған жүйеге студенттердің суреттері қажет, мысалы, ашу, қорқыныш, қайғы және т.б. мазмұнға негізделген кескіндерді іздеу (CBIR) тұрғысынан нақты сипаттама алу. Бұл жазбалар студенттердің қалыптан тыс белсенділігі анықталған жағдайда CBIR дерекқорында сақталады. Содан кейін келесі қадамда кескін кадрларға айналады.

Алғаш рет терең оқыту парадигмасы 3D CNN-ді қолдана отырып, кіріс ретінде толық бейне тізбегін қабылдайтын тапсырмаға қолданылады. Бірақ бұл тапсырма үшін адамның қозғалыс ерекшеліктері өте маңызды және толық бейнені кіріс ретінде пайдалану оқу процесінде шу мен артықтықты тудырады. Осы мақсатта Goyal R., Kahou S. E., Michalski V., Materzyńska J. «handcrafted/learned» гибриді фреймворкін ұсынды [15]. Бұл әдіс, біріншіден, белгілерді алу үшін кіріс ретінде алынған бейне тізбегінен иллюстрациялық бейнені алуға бағытталған, ал Хафа орманы классификатор ретінде қолданылады. Содан кейін 2D CNN осы кескінді жіктеу және реттілік туралы қорытынды жасау үшін пайдаланылады. Бұл әдіс қолдан жасалған ұсақ-түйектерді ұсынудың ең жақсы әдістерінің бірі үшін керемет нәтиже берді. Алайда, екі ағынды архитектура есептеулердің күрделілігіне байланысты нақты уақыттағы қосымшалар үшін жарамсыз болып қалуы мүмкін.

Нәтижелер ұсынылған әдіс дәлдік пен стандартты ауытқуларға негізделген қолмен және терең оқытудың әртүрлі әдістеріне қарағанда жақсы жұмыс істейтінін көрсетеді.

Материалдар және зерттеу әдістері. Зорлық-зомбылықты анықтау әдістері қолданылатын классификаторға байланысты үш санатқа жіктеледі: машиналық оқытуды қолдана отырып зорлық-зомбылықты анықтау әдістері, SVM көмегімен зорлық-зомбылықты анықтау әдістері және терең оқытуды қолдана отырып зорлық-зомбылықты анықтау әдістері. SVM және терең оқыту бөлек жіктеледі, өйткені бұл алгоритмдер компьютерлік көру қабілетінде кеңінен қолданылады. 1-кестеде зорлық-зомбылық оқиғаларын тану әдістерінің тізімі берілген. SVM (Support vector machine) – бұл күшейту жаттығуларын қолдана отырып, жіктеу есептерін шешу үшін қолданылатын алгоритм. Жіктеу міндеті мұғаліммен бірге оқуды білдіреді. SVM-мұғаліммен оқыту алгоритмі. SVM классификатор ретіндегі негізгі мақсаты бөлуші гипержазықтықтың теңдеуін табу болып табылады

$\omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n + \omega_0 = 0$ R^n кеңістігінде, бұл екі класты қандай да бір оңтайлы жолмен ажыратады.

Келесі шарт орындалса, алгоритм объектілерді дұрыс жіктейді:

$$y(\omega^T x - b) \geq 1$$

Ол Кун-Такер теоремасы арқылы аналитикалық жолмен шешіледі. Алынған есеп Лагранж функциясының ершік нүктесін табудың қосарлы есебіне тең.

$$\begin{cases} (\omega^T \omega) / 2 \rightarrow \min \\ y(\omega^T x - b) \geq 1 \end{cases}$$

1-кесте. SVM көмегімен зорлық-зомбылықты анықтау әдістері

Әдісі	Объектіні анықтау әдісі	Ерекшелікті шығару әдісі	Оқиға түрі	Дәлдік %
1	2	3	4	5
Нақты уақытта адам көп жиналған көріністердегі зорлық-зомбылықты анықтау [1]	ViF дескрипторы	Функциялар жиынтығы	Адамдардың көп шоғырлануы	88 %
Әрекеттерді анықтау үшін жеделдету арқылы Bag of words фреймворкі [2]	Фондық алу алгоритмдері	Тізбекті кадрлар үшін эллипсті бағалау әдісі	Адамдардың азырақ шоғырлануы	Шамамен 90 %
Бақылау және анықтау модулі бар генетикалық алгоритм фреймворкі [3]	Гаусс үлгісі	Оптикалық ағынды алу алгоритмі	Адамдардың көп шоғырлануы	82 %-89 %
Ішкі класс негізінде фреймворктың мультимодальдық ерекшеліктері [4]	CNN және ImageNet суреті	Функцияларды шығаруға арналған Google Net	Адамдардың азырақ шоғырлануы	98 %
Күштеп тағайындау әрекеттерінің жиілігін анықтау [5]	Кеңістіктік пирамидалар мен объектілерді анықтауға арналған торлар	Функцияларды алудың кеңістік-уақыттық тор әдістері	Адамдардың көп шоғырлануы	96 %-99 % эртүрлі деректер жиынын пайдаланады
Бағытталған зорлық-зомбылық ағынын қолдана отырып, зорлық-зомбылықты анықтау [6]	Оптикалық ағын әдісі	ViF және Onvif тіркесімі дескриптор	Адамдардың көп шоғырлануы	90 %
AEI және HOG визуалды қозғалыстардағы аномалды оқиғаны тану үшін фреймворкты біріктірді [7]	Фонды алу үшін AEI әдісі	HOG және белгілерді алудың кеңістік-уақыт әдістері	Адамдардың азырақ және көбірек шоғырлануы	94 %-95 %
Фреймворк алдын ала өңдеуді, белсенділікті анықтауды және суреттерді шығаруды қамтиды. Бұл жұмыс аномальды оқиғаны және	Объектіні анықтау үшін оптикалық ағын және уақыт айырмашылығы CBIR кескінді	Бейне файлдарды талдауға арналған Гаусс функциясы	Адамдардың азырақ шоғырлануы	97 %

деректерді анықтайды [8]	шығару әдісі			
--------------------------	--------------	--	--	--

1-кестенің соңы

1	2	3	4	5
Жоғары деңгейдегі белсенділікті анықтау үшін перцептивті уақытша қабаттарға арналған кеш біріктіру әдісі. 1-ден N-ге дейінгі бірнеше камераны пайдалану [9]	Екі өлшемдегі бірнеше камералардан сәйкестендіруге арналған қозғалыс векторы әдісі	SGT MtPL әдісі	Адамдардың азырақ шоғырлануы	98 %
Нақты уақытта анықтауға арналған қос арналы конволюциялық нейрондық желі [10]	Объектіні анықтауға арналған ImageNet	Белгілерді алу үшін VG-f моделі	Адамдардың көп шоғырлануы	91 %- 94 %
Анықтау мәселесін Connect арқылы мақсатты тереңдік және анық пішім бойынша бөлу арқылы шешіңіз [11]	Қозғалысты анықтау және Trof моделі	BoW тәсіл	Адамдардың азырақ шоғырлануы	96 %
Бейнедегі аномалияларды анықтау үшін кеңістік-уақыттық әдісті қолданатын сөздер қаптамасы әдісі [12]	Сегменттерді және ішкі бөлімдерді көрсету	Бейнекадрлар алу үшін HOF және HOG	Адамдардың азырақ шоғырлануы	84 %- 91 %
3D CNN көмегімен зорлық-зомбылықты анықтау [13]	3D сверткасы кеңістіктік ақпаратты алу үшін қолданылады	Кері таралу әдісі	Кісі көп жерде	91 %
Орынды тану үшін терең архитектура [14]	VGG VLAD кескінді іздеу әдісі	Функцияларды шығаруға арналған кері таралу әдісі	Кісі көп жерде	87 %- 96 %
CNN және терең дыбыс мүмкіндіктерін пайдаланатын зорлық-зомбылық көріністері [15]	MFB	CNN моделі	Кісі көп жерде	Шамамен 90 %
convLSTM көмегімен зорлық-зомбылық бейнелерін анықтау [15]	ConvLSTM-пен бірге CNN	CNN моделі	Кісі көп жерде	Шамамен 97 %

Зорлық-зомбылықты мойындау үшін терең білім берудегі зорлық-зомбылықты анықтаудың үш сатылы құрылымы ұсынылады [11]. Біріншіден, бейнебақылау ағындарында адамдар жарамсыз кадрларды өңдеуді еңсеру және азайту үшін жеңіл салмақты CNN үлгісін пайдаланып анықталады. Екіншіден, анықталған жеке тұлғалары бар 16 кадрдың тәртібі 3D CNN-ге беріледі, онда осы тізбектің кеңістіктік және уақыттық сипаттамалары алынып, Softmax классификаторына беріледі. Содан кейін CNN 3D моделі Intel жасаған нейрондық желілерді оңтайландыру құралдары мен ашық визуалды ШЫҒЫС

арқылы оңтайландырылған. Оқытылған модель аралық иллюстрацияға айналады және оны зорлық-зомбылықты түпкілікті анықтау үшін соңғы платформада орындау үшін өзгертеді. Зорлық-зомбылық анықталғаннан кейін дабыл әрекет ету үшін көрші қауіпсіздік бөліміне немесе полиция бөлімшесіне жіберіледі.

Дене қозғалысы да соңғы жылдары көбірек назар аударатын негізгі интерфейс болып табылады.

Зорлық-зомбылықты тану және анықтау сыртқы бақылау бейнелері үшін маңызды тақырыпқа айналады. Негізгі мақсат – зорлық-зомбылықтың болып жатқанын анықтау. Біріншіден, бейнеклиптер үшін жақсартылған Фишер векторларын (IFV) кеңейту ұсынылады [9]. Бейнені бейнелеуге мүмкіндік беретін жергілікті нысандар және олардың кеңістіктік-уақыттық позициялары қолданылады. Содан кейін зорлық-зомбылықты анықтау үшін танымал жылжымалы терезе әдісі зерттеледі. Тәсілді жеделдету үшін жиынтық аудан кестесінің деректер құрылымы қолданылады және ifv формулалары қайта құрылады. Екіншіден, жергілікті кеңістік-уақыт сипаттамалары жақсартылған тығыз траекторияларды (IDT) пайдаланып бейнелерден алынады. Содан кейін әр дескрипторға арналған бейне презентация ifv көмегімен бейнені ұсыну үшін HOG ретінде дербес есептеледі. Содан кейін сызықтық SVM классификаторы зорлық-зомбылықты тану үшін қолданылады және ақыр соңында жылдам жылжымалы терезе әдісін қолдана отырып, зорлық-зомбылық анықталады. Кең бағалау 4 заманауи зорлық-зомбылық ағындары, фильмдер мен хоккей ойындарының деректер жиынтығын қолдану арқылы жүзеге асырылады. Violence-Flow 21 деректер жиынтығы зорлық-зомбылықты анықтау міндеті үшін қолданылады. Нәтижелер ұсынылған тәсілдер қолданыстағы тәсілдермен салыстырғанда жақсы жұмыс істейтінін көрсетеді.

Іс-әрекеттерді визуалды тану қорлауды анықтаудың негізгі әдісі болып табылады. Іс-әрекетті тану. Іс-әрекетті визуалды тану агрессияны танудың негізгі әдісі болып табылады. Бұл тәсілді қолдана отырып, байқалған агрессияны бірнеше уақыт шкалаларында (MTRNN) екі басқарылатын қайталанатын нейрондық желілерді қолдану арқылы есептеуге болады [10]. Бірінші кезең адамның іс-әрекетін таниды, ал екінші кезең бірінші кезеңнің нәтижелері бойынша адам ниетін болжайды. Сонымен қатар, стихиялық эмоцияларды түсіну үшін әрекеттердің сипаттамалары иерархиялық түрде зерттеледі. Бұл тәсіл вербалды емес эмоцияларды, соның ішінде мимика мен дене қозғалыстарын мультимодальды тану мүмкіндіктерін біріктіру үшін көп арналы конволюционды нейрондық желіні (CCNN) қолданады.

Толық түсті RGB кескіндері негізінде адам денесінің позаларын анықтау алгоритмдерін 2 топқа бөлуге болады: жоғарыдан төмен алгоритмдер және төменнен жоғары алгоритмдер. Біріншісі адам денесінің кадрдың қай жерде орналасқанын анықтайды және осы шекте дененің буындарының орналасуын іздейді.

Бұл жұмыста бейнедегі нысандарды тану үшін жасанды интеллект моделін жасау қаңқа моделінде оқыту қолданылды. Бұл тәсіл есептеу шығындарын азайтуға мүмкіндік береді. Кадрдағы адам фигураларын анықтау үшін жеңіл, алдын ала дайындалған polenet нейрондық желісі таңдалды. Алдын ала үйретілген polenet желісін экстрактор функциясының ішінде пайдалануға болады, бұл кадрдың ерекшеліктері туралы білімді бір семантикалық кеңістіктен екіншісіне тасымалдауға мүмкіндік береді.

PoseNet бір позаны немесе бірнеше позаны бағалау үшін пайдаланылуы мүмкін, яғни алгоритмнің бір нұсқасы суреттегі/бейнедегі бір адамды ғана анықтай алады, ал басқа нұсқасы суреттегі/бейнедегі бірнеше адамды анықтай алады.

Бағдарламалық қамтамасыз етуді енгізу:

– бейнеде PoseNet моделінің көмегімен адам фигураларының қаңқалары анықталады;

– әрі қарай, негізгі нүктелердің координаттары агрессивті әрекеттер кластары бойынша классификатор үшін белгілер ретінде қолданылады.

Позаны бағалау моделі негізгі нүктелердің кіріс және шығыс ақпараты ретінде камера өңдеген кескінді пайдаланады. Анықталған негізгі нүктелер 0,0-ден 1,0-ге дейінгі сенімділік ұпайларымен бөлік идентификаторы бойынша индекстеледі. Сенімділік көрсеткіші негізгі нүктенің сол жерде болуы ықтималдығын білдіреді. Бұл әдістің маңызды сәттері төменде сипатталған.

Позаға деген сенімділік-бұл позаны бағалауға деген жалпы сенімділікті анықтайды. Бұл мән 0,0-ден 1,0-ге дейін. Оны сенімділігі жоқ позаларды жасыру үшін қолдануға болады.

Негізгі нүктелердің сенімділігін бағалау-бұл негізгі нүктенің болжамды орналасуының дәлдігіне сенімділікті анықтайды. Бұл мән 0,0-ден 1,0-ге дейін. Оны жеткілікті сенімділік жоқ негізгі ойларды жасыру үшін пайдалануға болады.

Негізгі нүктенің орналасуы-бастапқы кіріс кескінінде анықталған негізгі нүктелердің X және Y екі өлшемді координаттары.

Негізгі мәселе – адамның болжамды қалпының бөлігі: мұрын, оң құлақ, сол тізе, оң аяқ және т.б. Онда позициялар мен негізгі ойларға деген сенімділік ұпайлары бар. PoseNet қазіргі уақытта 17 негізгі нүктені анықтай алады.

Позаны бағалау үлгісімен анықталған дененің буындары 2-кестеде көрсетілген:

2-кесте. PoseNet моделімен табылған дененің әртүрлі буындары

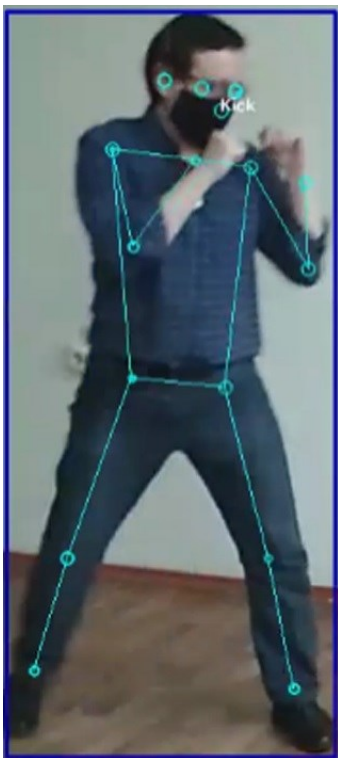
Идентификатор	Бөлім
0	мұрын
1	сол көз
2	оң көз
3	сол құлақ
4	Оң құлақ
5	сол жақ иық
6	оң иық
7	сол жақ шынтақ
8	оң жақ шынтақ
9	сол жақ білек
10	оң жақ білек
11	сол жақ жамбас
12	оң жамбас
13	сол жақ тізе
14	оң тізе
15	сол жақ сирақ
16	оң жақ сирақ

Нәтижелері және оларды талқылау. PoseNET нәтижесі адам фигурасын координаталар мен сенімділік ұпайларымен 17 негізгі нүкте арқылы көрсету болып табылады. Бұл 17 нүктеге мыналар кіреді: мұрын, көз, құлақ, иықтың жоғарғы жағы, шынтақ, білек, жамбас, тізе және тобық. PoseNET желісі арқылы 17 негізгі нүктені анықтау мысалы 1-суретте көрсетілген.

PoseNet әрбір анықталған адамның сенімділік мәнін және әрбір анықталған қимылдың негізгі нүктелерін қайтарады.

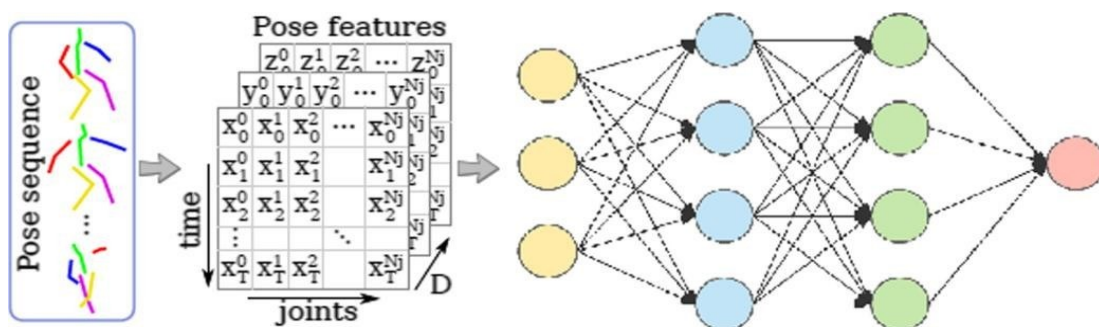
Жоғарғы деңгейден бастап позаны бағалау екі кезеңде жүреді:

1. RGB кескіні конволюциялық нейрондық желіге енгізіледі.
2. Бір немесе көп позициялы декодтау алгоритмдері позаларды декодтау, сенімділік ұпайларын құру, негізгі нүктелердің орналасуы және модельдің шығуына негізделген негізгі нүктелердің сенімділігін бағалау үшін қолданылады.



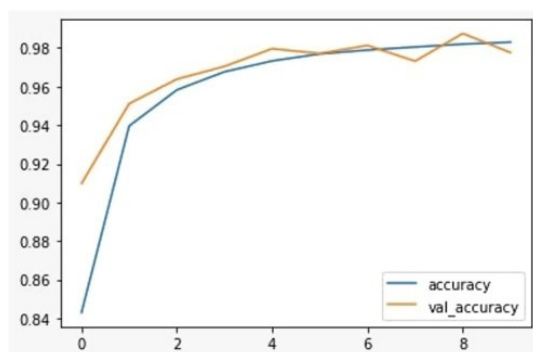
1-сурет. PoseNET желісі арқылы 17 негізгі нүктені анықтау

Адам фигурасының алынған көрінісін жіктеу үшін нейрондық желінің толық байланысқан қабаты қолданылды. Бұл нейрондық желі қондырмасы категориялық кросс-энтропиялық шығын функциясын азайту арқылы оқытылады. Жіктеу нәтижелері көп айнымалы логистикалық функциямен (Softmax) келесі қабатпен қалыпқа келтіріледі. Пайдаланылған posenet нейрондық желісінің архитектурасы 2-суретте көрсетілген.

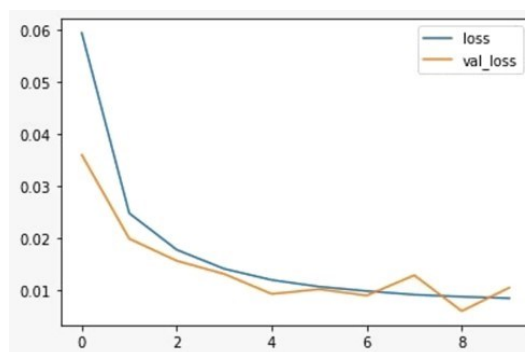


2-сурет. Пайдаланылған posenet нейрондық желісінің архитектурасы.
(ҚР ҒЖБМ АР14871625 жобасы)

Бейнедегі объектілерді тануға арналған жасанды интеллект модельдерін оқыту эксперименттері кеңінен қолданылатын схема бойынша жүргізілді: мәліметтер жиынтығы сәйкесінше 80 % және 20 % пропорцияда оқыту және тестілеу үшін 2 бөлікке бөлінді. 3-суретте оқу процесінде модель дәлдігінің өзгеруі көрсетілген. 4-суретте оқу процесінде модельдің жоғалту функциясының өзгеруі көрсетілген. *Зерттеу ҚР ҒЖБМ АР14871625 жобасының қолдауымен орындалды.*



3-сурет. Оқу процесінде модельдің дәлдігінің өзгеруі (ҚР ҒЖБМ АР14871625 жобасы)



4-сурет. Оқу процесінде модельдің жоғалту функциясының өзгеруі (ҚР ҒЖБМ АР14871625 жобасы)

Жоғарыда келтірілген графиктерден көріп отырғаныңыздай, бейнедегі объектілерді тану үшін жасанды интеллект моделінің дәлдігі 8 оқу кезеңінен кейін 98 % жетеді. Бірнеше эксперименттердің нәтижелері бойынша бейнедегі нысандарды тану үшін ең жақсы сапаны көрсететін модель тандалды.

Жалғыз позаны бағалау алгоритмі қарапайым және жылдамырақ. Оның идеалды сценарийі – кіріс суретінің немесе бейненің ортасында бір ғана адам болуы. Кемшілігі, егер суретте бірнеше адам болса, екі адамның негізгі нүктелерін бағалауға болады, яғни алгоритм арқылы бір позаның бөлігі ретінде анықтауға болады. Позаны біріктіруге болады. Енгізілген кескінде бірнеше адам болуы мүмкін болса, көп позицияны бағалау алгоритмін пайдалану керек.

Көп адамдық позаны бағалау алгоритмі бірнеше позаны (видеодағы адамдар) бағалай алады. Бұл жалғыз поза алгоритміне қарағанда күрделірек және сәл баяуырақ, бірақ оның артықшылығы бар, егер суретте бірнеше адам пайда болса, олар анықтаған негізгі нүктелер дұрыс емес позаға байланысты болуы екіталай. Осы себепті, қолданба сценарийі бір адамның позасын анықтауға арналған болса да, бұл алгоритм қажет болуы мүмкін.

Сонымен қатар, алгоритмнің тартымды ерекшелігі оның өнімділігіне кіріс кескіндегі адамдар санының әсер етпеуі болып табылады. 15 адам болсын, 5 адам болсын, есептеу уақыты бірдей.

Қорытынды. Компьютерлік көру кезінде іс-әрекетті тану зерттеудің маңызды бағытына айналады. Агрессивті мінез-құлық немесе төбелес сияқты міндеттер салыстырмалы түрде аз зерттелген, бірақ көптеген бейнебақылау сценарийлерінде пайдалы болуы мүмкін. Терең оқыту көптеген компьютерлік көру қосымшаларында дәстүрлі тәсілдерден асып түсетін Машиналық оқыту аясында өте танымал бағытқа айналды. Терең оқыту алгоритмдерінің өте пайдалы қасиеті-бұл қолмен жұмыс істейтін

дескрипторлар мен дескрипторлардың қажеттілігін жоятын шикі деректерден функцияларды үйрену мүмкіндігі.

Мақалада физикалық, әлеуметтік және басқа да зорлық-зомбылықты автоматты түрде анықтау және жіктеу, байқалған психоэмоционалды жағдайды анықтау үшін машиналық оқыту және жасанды интеллект алгоритмдері қарастырылды. Бұл алгоритмдерге шетелдік мамандардың басқа жұмыстары мен зерттеулерінің мысалдарымен салыстырмалы талдау жасалды.

Әдебиеттер тізімі

1. Marinou, E., Zafir, M., Olaru, V., & Sminchisescu, C. (2018). 3d human sensing, action and emotion recognition in robot assisted therapy of children with autism. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2158-2167).
2. El-Ghaish, H., Hussein, M.E., Shoukry, A., & Onai, R. (2018). Human action recognition based on integrating body pose, part shape, and motion. *IEEE Access*, 6, 49040-49055.
3. Song, S., Lan, C., Xing, J., Zeng, W., & Liu, J. (2016). An end-to-end spatio-temporal attention model for human action recognition from skeleton data. *arXiv preprint arXiv:1611.06067*.
4. Arici, T., Celebi, S., Aydin, A.S., & Temiz, T.T. (2017). Robust gesture recognition using feature pre-processing and weighted dynamic time warping. *Multimedia Tools and Applications*, 72(3), 3045-3062.
5. Ji, Y., Cheng, H., Zheng, Y., & Li, H. (2018). Learning contrastive feature distribution model for interaction recognition. *Journal of Visual Communication and Image Representation*, 33, 340-349.
6. Huynh-The, T., Le, B. V., Lee, S., & Yoon, Y. (2017). Interactive activity recognition using pose-based spatio-temporal relation features and four-level Pachinko Allocation Model. *Information Sciences*, 369, 317-333.
7. Charalampous, K., Kostavelis, I., Boukas, E., Amanatiadis, A., Nalpantidis, L., Emmanouilidis, C., & Gasteratos, A. (2019). Autonomous robot path planning techniques using cellular automata. In *Robots and lattice automata* (pp. 175-196). Springer, Cham.
8. H.S. Koppula, R. Gupta, and A. Saxena, "Learning human activities and object affordances from rgb-d videos," *International Journal of Robotics Research*, vol. 32, no. 8, pp. 951-970, 2017.
9. C. Granata, A. Ibanez, and P. Bidaud, "Human activity-understanding: A multilayer approach combining body movements and contextual descriptors analysis," *International Journal of Advanced Robotic Systems*, vol. 12, no. 7, 2019.
10. V. Dutta and T. Zielinska, "Predicting the intention of human activities for real-time human-robot interaction (hri)," in *ICRA*, 2020.
11. <http://hacs.csail.mit.edu>
12. Zhao, H., Torralba, A., Torresani, L., Yan, Z. Hacs: Human action clips and segments dataset for recognition and temporal localization // *Proceedings of the IEEE International Conference on Computer Vision*. – 2019. – С. 8668-8678.
13. М.Ю. Уздяев "Распознавание агрессивных действий с использованием нейросетевых архитектур 3d-CNN," *Известия ТулГУ. Технические науки*. 2020. Вып. 2.
14. Нелли Адаменко, ИА «NewTimes.kz»
15. Goyal R., Kahou S. E., Michalski V., Materzyńska J., Westphal S., Kim H., Haenel V., Freund I., Yianilos P., Mueller-Freitag M., Hoppe F., Thureau C., Bax I., Memisevic R. The "Something Something" Video Database for Learning and Evaluating Visual Common Sense // *ICCV*. – 2017. – Т. 1. – №. 4. – С. 5.
16. Ji, Y., Cheng, H., Zheng, Y., & Li, H. (2018). Learning contrastive feature distribution model for interaction recognition. *Journal of Visual Communication and Image Representation*, 33, 340-349.
17. Fu E.Y., Va Leong H., Ngai G., Chan S. Automatic fight detection in surveillance videos // *Proc. 14th Int. Conf. Adv. Mobile Comput. Multi Media*. – 2016. – P. 225-234.
18. Chaudhary S., Khan M.A., Bhatnagar C. Multiple anomalous activity detection in videos // *Procedia Comput. Sci*. – 2018. – № 125. – P. 336-345.
19. Hassner T., Itcher Y., Kliper-Gross O. Violent flows: Real-time detection of violent crowd behavior // *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*. – 2012. – P. 1-6.
20. Al-Nawashi M., Al-Hazaimeh O.M., Saraee M. A novel framework for intelligent surveillance system based on abnormal human activity detection in academic environments // *Neural Comput. Appl*. – 2017. – № 28(1). – P. 565-572.

References

1. Marinouiu, E., Zanzfir, M., Olaru, V., & Sminchisescu, C. (2018). 3d human sensing, action and emotion recognition in robot assisted therapy of children with autism. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2158-2167).
 2. El-Ghaish, H., Hussein, M. E., Shoukry, A., & Onai, R. (2018). Human action recognition based on integrating body pose, part shape, and motion. *IEEE Access*, 6, 49040-49055.
 3. Song, S., Lan, C., Xing, J., Zeng, W., & Liu, J. (2016). An end-to-end spatio-temporal attention model for human action recognition from skeleton data. arXiv preprint arXiv:1611.06067.
 4. Arici, T., Celebi, S., Aydin, A. S., & Temiz, T. T. (2017). Robust gesture recognition using feature pre-processing and weighted dynamic time warping. *Multimedia Tools and Applications*, 72(3), 3045-3062.
 5. Ji, Y., Cheng, H., Zheng, Y., & Li, H. (2018). Learning contrastive feature distribution model for interaction recognition. *Journal of Visual Communication and Image Representation*, 33, 340-349.
 6. Huynh-The, T., Le, B. V., Lee, S., & Yoon, Y. (2016). Interactive activity recognition using pose-based spatio-temporal relation features and four-level Pachinko Allocation Model. *Information Sciences*, 369, 317-333.
 7. Charalampous, K., Kostavelis, I., Boukas, E., Amanatiadis, A., Nalpantidis, L., Emmanouilidis, C., & Gasteratos, A. (2019). Autonomous robot path planning techniques using cellular automata. In *Robots and lattice automata* (pp. 175-196). Springer, Cham.
 8. H. S. Koppula, R. Gupta, and A. Saxena, "Learning human activities and object affordances from rgb-d videos," *International Journal of Robotics Research*, vol. 32, no. 8. – P. 951-970, 2017.
 9. C. Granata, A. Ibanez, and P. Bidaud, "Human activity-understanding: A multilayer approach combining body movements and contextual descriptors analysis," *International Journal of Advanced Robotic Systems*, vol. 12, no. 7, 2019.
 10. V. Dutta and T. Zielinska, "Predicting the intention of human activities for real-time human-robot interaction (hri)," in *ICRA*, 2020.
 11. <http://hacs.csail.mit.edu>
 12. Zhao, H., Torralba, A., Torresani, L., Yan, Z. Hacs: Human action clips and segments dataset for recognition and temporal localization // *Proceedings of the IEEE International Conference on Computer Vision*. – 2019. – S. 8668-8678.
 13. M.Ju. Uzdjaev (2020) Raspoznavanie agressivnyh dejstvij s ispol'zovaniem nejrosetevykh arhitektur 3d-CNN [Recognition of aggressive actions using 3d neural network architectures-CNN]. *Izvestija TulGU. Tehniceskie nauki.. Vyp. 2*.
 14. Nelli Adamenko, IA «NewTimes.kz»
 15. Goyal R., Kahou S. E., Michalski V., Materzyńska J., Westphal S., Kim H., Haenel V., Freund I., Yianilos P., Mueller-Freitag M., Hoppe F., Thureau C., Bax I., Memisevic R. The "Something Something" Video Database for Learning and Evaluating Visual Common Sense // *ICCV*. – 2017. – T. 1. – №. 4. – S. 5.
 16. Ji, Y., Cheng, H., Zheng, Y., & Li, H. (2017). Learning contrastive feature distribution model for interaction recognition. *Journal of Visual Communication and Image Representation*, 33, 340-349.
 17. Fu E.Y., Va Leong H., Ngai G., Chan S. Automatic fight detection in surveillance videos // *Proc. 14th Int. Conf. Adv. Mobile Comput. Multi Media*. – 2016. – P. 225-234.
 18. Chaudhary S., Khan M.A., Bhatnagar C. Multiple anomalous activity detection in videos // *Procedia Comput. Sci.* – 2018. – № 125. – P. 336-345.
 19. Hassner T., Itcher Y., Kliper-Gross O. Violent flows: Real-time detection of violent crowd behavior // *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*. 2012. – P. 1-6.
 20. Al-Nawashi M., Al-Hazaimeh O.M., Saraee M. A novel framework for intelligent surveillance system based on abnormal human activity detection in academic environments // *Neural Comput. Appl.* – 2017. – № 28(1). – P. 565-572.
-
-