

Python бағдарламалау ортасында роботтың адам бейнесі мен сөйлеуін тану жүйесін жасау

¹***БАЛТАБАЙ Дәурен Қуанышбекұлы**, докторант, dauren.baltabay.95@gmail.com,

¹**ЖҰМАШЕВА Жадыра Токановна**, т.ғ.к., қауымдастырылған профессор, Zhadyra_14@mail.ru,

¹«Әл-Фараби атындағы Қазақ ұлттық университеті» КеАҚ, әл-Фараби даңғылы, 71, Алматы, Қазақстан,

*автор-корреспондент.

Аңдатпа. Соңғы уақытта жеке тұлғаны анықтау мақсатында адам бейнесін танудың ақпараттық-іздеу жүйелерінің технологиясы кеңінен таралуда. Адам бейнесін тану және сөйлеу технологиясын қолдана отырып, жеке тұлғаны сәйкестендірудің ішкі жүйелері автоматтандырылған қауіпсіздік жүйелерінде жеке тұлғаны орнатумен айналысатын қызметтерге тиісті жедел ақпаратты бір уақытта бере отырып, іздестірілген тұлғаларды сәйкестендіру үшін қолданыла алады. Бұл жұмыстың мақсаты – Python бағдарламалық ортасында роботтың адамның бейнесі мен сөйлеуін тану жүйесін жасау. Сонымен қатар, «тану» сөзінің екі мағынасы бар. Біріншісі – ауызша сөйлеуді тиісті жазбаша мәтінге, сондай-ақ адамның бет-әлпетін, көзін ерекшелеп, оны сұр реңкке айналдыру. Екіншісі – талап немесе диктордың анықтамасы бар мәлімдемеге дұрыс реакция және бет пен көздің контурына тікбұрышты пішіндерді шығару. Сөйлеуді танудың синтезімен үйлесуі адам мен машина арасындағы байланыс жүйесін құра алады. Өнеркәсіп, ғылым және қорғаныс салаларындағы сөйлеу мен адам бейнесін тану бойынша қарқынды зерттеу мотиві компьютерлерді, кеңсені автоматтандыруды, байланысты, қорғаныс пен робототехниканы қамтитын ақпараттық индустрияның қарқынды дамуына байланысты туындады.

Кілт сөздер: жасанды интеллект, сөйлеуді тану, адам бейнесін тану, машиналық оқыту, компьютерлік көру, сәйкестендіру, процесті автоматтандыру, роботтар, python, OpenCV.

Кіріспе

Жасанды интеллект қазіргі заманғы технологиялар мен мәліметтерге негізделген әлемнің ажырамас бөлігіне айналуда. Ол іздеу жүйелері, үлгіні тану, сөйлеуді тану, робототехника, өздігінен басқарылатын машиналар және т.б. сияқты салаларда кеңінен қолданылады.

Жасанды интеллект (ЖИ) роботтарға адамның интеллектуалды мінез-құлқы мен ойлау қабілетін имитациялайтын мүмкіндіктермен қамтамасыз етуге мүмкіндік береді. Роботтар бағдарламалық жасақтамалармен басқарылады, сондықтан жасанды интеллект роботтардың әрекетін басқаратын интеллектуалды бағдарламалармен көп ұқсастыққа ие. Жасанды интеллект ғылымы роботтарға сыртқы ортаны бағалауға және әр түрлі жағдайларға оларға адам реакциясы сияқты жауап беруге мүмкіндік беретін теориялар мен әдістемелерді дамыту үстінде.

Соңғы жылдары роботтар біздің күнделікті өмірімізде жиі кездеседі. Тұтынушылар-

ға қызмет көрсетуге арналған автоматтандырылған чат-боттардан бастап зауыттарда жұмыс істейтін өнеркәсіптік роботтарға дейін олар біздің қазіргі қоғамымыздың ажырамас бөлігіне айналды. Дегенмен, олардың пайдалылығына қарамастан, роботтар адам қасиеттердің жоқтығы үшін жиі сынға ұшырады, бұл адамдарға олармен жеке деңгейде байланысуды қиындатты.

Бұл мәселені шешу үшін зерттеушілер роботтарға адамға ұқсас көрінуге, сондай-ақ адамның сөйлеуін тануға және оған жауап беруге мүмкіндік беретін жүйелерді әзірлеуде. Бұл мақалада біз роботтың адам бейнесі және сөйлеуін тану жүйесі қалай жасалатынын қарастырамыз. Біз осындай жүйені құруға қатысты мәселелер мен мүмкіндіктерді, зерттеушілер қолданған әртүрлі тәсілдерді және осы технологияны әртүрлі салаларда қолданудың әлеуетін талқылаймыз. Сайып келгенде, біз роботтардың адам бейнесі мен сөйлеуін тану жүйесін дамыту адамдар мен машиналар арасындағы табиғи және интуитивті өзара әрекеттесуге жол ашып, робот-

тардың қоғамға одан да пайдалы болуына мүмкіндік беретінін көреміз.

Адамның сөйлеуі және бейнесін тану жүйесін құрудың жаңалығы-адамдармен табиғи және интуитивті түрде қарым-қатынас жасай алатын машиналар жасау мүмкіндігі. Адамның бейнесін тану және сөйлеуді тану технологиялары ондаған жылдар бойы болды, бірақ машиналық оқыту мен табиғи тілді өңдеудегі және компьютерлік көру мен терең оқыту саласындағы соңғы жетістіктер адамның сөйлеуін түсінуге және оған нақты уақыт режимінде жауап беруге қабілетті сөйлеуді тану жүйелерін, адамның бет-әлпеті мен ерекшеліктерін жоғары дәлдікпен тануға қабілетті жүйелерді құруға мүмкіндік берді.

Сөйлеу мен бейнені тану технологияларын біріктірудің жаңалығы адамның сөйлеуін түсініп қана қоймай, адамның бет-әлпетін тани алатын роботтарды жасау мүмкіндігінде жатыр. Бұл денсаулық сақтау, өндіріс немесе тұтынушыларға қызмет көрсету саласында болсын, адамдармен табиғи және интуитивті түрде қарым-қатынас жасай алатын роботтарды жасаудың кең ауқымын ашады.

Мысалы, сөйлеуді және бейнені тану технологиясы бар роботты ауруханада қозғалғыштығы шектеулі немесе қарым-қатынас қабілеті шектеулі науқастарға көмек көрсету үшін пайдалануға болады. Робот пациенттің бетін тани алады және қажет болған жағдайда көмек пен қолдау көрсету арқылы оның сөйлеуіне жауап бере алады. Өндіріс жағдайында сөйлеуді және бейнені тану технологиясы бар робот қайталанатын тапсырмаларды автоматтандыру және адам жұмысшыларымен тиімдірек байланыс орнату үшін пайдаланылуы мүмкін.

Зерттеу әдістері

Жасанды интеллект туралы ғылымның даму бағыттарын мұқият талдау нәтижесі көрсеткендей, ғалымдар жасанды интеллектті анықтауда әр түрлі тәсілдерді қолданады. Қазіргі заманда жасанды интеллект көптеген әр түрлі нысандарға қатысады. Біз роботтардың сезініп, ой қорытып, ойланып, әрекет ете алатындығын қалаймыз. Сонымен қатар, біз роботтардың өзін ұтымды ұстағанын қалаймыз.

Жасанды интеллект саласының жұмысы адам миының қасиеттері мен жұмыс жасау принципін зерттеумен тығыз байланысты. Ғалымдар адам миінің жұмыс жасау принципін түсіну жасанды интеллекттің жасалуына мүмкіндік туғызады деп есептейді. Адамның миында жүзеге асырылатын оқу, ойлау және шешім қабылдау процестеріне еліктеу арқылы дәл сондай әрекет жасай алатын робот-

тар құрастыруға болады. Мұндай роботтар оқытуға қабілетті жүйелерді құрудың платформасы бола алады [1].

Жасанды интеллект жүйесін жасаудағы жетістіктер біздің өміріміздің барлық аспектілеріне тікелей әсер етуі мүмкін. Бұл сала бағытындағы зерттеулер объектінің мінез-құлқының заңдылықтары мен қасиеттерін зерттеуге негізделген. Жасанды интеллект жүйесінің көмегімен ақылды жүйелер құру мен роботтардың шығармашылық бағытындағы іс-әрекеттерді қалай жасауға болатындығын түсінуге болады.

Ғасырлар бойы қалыптасқан, бізге белгілі математика және физика сияқты ғылым салаларымен салыстырғанда, жасанды интеллект туралы ғылым салыстырмалы түрде жас. Бұл ғылым саласы соңғы жиырма жыл ішінде өзін-өзі басқаратын автомобильдер мен ақылды жүретін роботтар сияқты керемет жетістіктерді көрсетті. Қол жеткізілген нәтижелер жасанды интеллектуалды зерттеулердің алдағы жылдарда біздің өмірімізде түбегейлі өзгеріс туғызатын әлеуеті бар екенін айқын көрсетеді [2].

Жасанды интеллект әдістері қолданылып жүрген машиналарды, олар интелектісі артып, өз функцияларын жылдам және тиімді орындай алатындай етіп жетілдіруге белсенді қолданылады.

Негізінде, робототехникалық жүйелер саласы жасанды интеллект жүйесінің көптеген концепцияларын біріктіреді. Бұндай жүйелер көптеген әртүрлі есептерді шешуге қабілетті. Әртүрлі жағдайларға байланысты роботтар барлық әрекеттерді орындауға қабілетті жетектермен және сенсорлармен жабдықталуы мүмкін. Датчиктер өзінің көру аумағындағы заттарды тани алады, температурасын өлшей алады, олар шығаратын жылу мен олардың қозғалыстарына реакция жасай алады және т.б. Электрондық тақталарға қойылған процессорлар қажетті есептеулерді нақты уақыт режимінде орындайды. Сонымен қатар, роботтар өздерінің мінез-құлықтарын өзгертін сыртқы орта жағдайларына бейімдей алады.

Жасанды интеллект ғылымында «ЖИ – кешенді мәселелер» ұғымы бар. Бұл шешімдер жасанды интеллект құруға тең болатын мәселелер класы. Жалпы жағдайда компьютерлік көру мәселесі (суретте не көрсетілген және осы кескін туралы барлық сұрақтарға жауап) ЖИ – кешенді болып саналады.

Біз, суретте не бейнеленген деген сұраққа жауап беру үшін сізге суретке қарап қана қоймай, бүкіл өмірлік тәжірибеңізді, біліміңізді, кейде ішкі түйсігіңізді қолдану қажет екенін көрсеттік. Өкінішке орай, «күшті» жасанды интеллект жүйесі әлі жасалған жоқ. Сондықтан компьютерлік көзқарас туралы

ғылым төменде қарастырылатын белгілі бір тапсырмаларды ғана шешеді [3-4].

Компьютерлік көру мәселелерін шешуге арналған ең танымал кітапхана – OpenCV. Онда көптеген алгоритмдер және бағдарламалау тілдеріне, оның ішінде C++ және Python тілдеріне арналған интерфейстері бар.

OpenCV – бұл кескінді талдауға, жіктеуге және өңдеуге арналған компьютерлік визуалды кітапхана, сонымен қатар, компьютерлік көрініс және машинамен оқытудың ашық кітапханасы. Оған компьютерлік көру мен машиналық оқытудың классикалық және қазіргі алгоритмдерін қамтитын 2500-ден астам алгоритмдер кіреді. Бұл кітапханада әртүрлі тілдерде, соның ішінде Python (біз оны осы мақалада қолданамыз), Java, C++ және Matlab интерфейстері бар.

Python-да бағдарламалық жасақтама жасау кезінде қазіргі уақытта көбінесе компьютерде жоқ пакетті қолдану қажеттілігі туады.

Егер сіз Python бағдарламалау ортасы қамтамасыз ететін негізгі функционалдылықтан тыс мәселемен жұмыс істеуді шешсеңіз, қосымша пакетті орнату қажеттілігі өте тез пайда болады. Бұл жағдайда сізге қандай функционалды пакет керектігін анықтап, оны тауып, жүктеп, керекті каталогқа орналастырып, қолдануды бастау керек. Барлық осы әрекеттерді қолмен жасауға болады, бірақ бұл үрдіс автоматтандыруға мүмкіндік береді [5-7].

Python-да адам бейнесін тану жүйесі туралы айтатын болсақ, онда OpenCV – бұл жан-жақты функционалдылықты ұсынатын кеңінен қолданылатын кітапхана. Дегенмен, адам бейнесін тану тапсырмалары үшін де пайдалануға болатын басқа қолжетімді кітапханалар бар:

Dlib – бұл C++ тілінің әмбебап кітапханасы, ол Python-да компьютерлік көрудің әртүрлі міндеттеріне, соның ішінде бетті анықтау мен тануға мүмкіндік береді. Ол жоғары дәлдіктегі және тиімді бет-әлпетті тану алгоритмдерін, соның ішінде белгілі dlib бет-әлпетін тану моделін ұсынады. Дегенмен, dlib OpenCV-ге қарағанда орнату және конфигурациялау үшін қосымша қадамдарды қажет етеді.

PyTorch және TensorFlow-бұл машиналық оқытудың көптеген міндеттерін, соның ішінде бетті тануды шешу үшін кеңінен қолданылатын терең оқыту шеңберлері. Бұл құрылымдар FaceNet, VGGFace және arcface сияқты алдын ала дайындалған модельдердің бай экожүйесін қамтамасыз етеді, оларды бетті тану үшін пайдалануға болады. PyTorch және TensorFlow икемділікті қамтамасыз етеді және модельдерді бейімдеуге мүмкіндік береді. Дегенмен, олар OpenCV-мен салыс-

тырғанда терең оқыту тұжырымдамалары туралы терең білімді қажет етеді.

OpenFace-бет-әлпетті тану тапсырмалары үшін алдын ала дайындалған модельдер мен құралдарды ұсынатын ашық бастапқы бет-әлпетті тану кітапханасы. Ол бет-әлпетті тануға және оларды тексеруге арналған терең нейрондық желі архитектурасын ұсынады. OpenFace дәлдігі мен сенімділігімен танымал. Дегенмен, бұл OpenCV-ге қарағанда қосымша тәуелділіктер мен орнату қадамдарын қажет етеді.

Осы кітапханаларды салыстыра отырып, OpenCV бетті тануға арналған әмбебап және жан-жақты кітапхана ретінде ерекшеленеді.

Алайда, егер сіз бет-әлпетті танудың жетілдірілген мүмкіндіктерін қаласаңыз немесе күрделі сценарийлермен жұмыс жасасаңыз, dlib және face_recognition сияқты кітапханалар жоғары дәлдікті қамтамасыз ететін және күрделі жағдайларды жеңе алатын арнайы алгоритмдер мен модельдерді қолдануға болады.

Кітапхананы таңдау жобаның нақты талаптары, қажетті дәлдік деңгейі, бетті тану тапсырмасының күрделілігі, қол жетімді ресурстар мен тәжірибе сияқты факторларға байланысты. Осы факторларды бағалау бет-әлпетті тану қолданбасы үшін ең қолайлы кітапхананы анықтауға көмектеседі.

Python бағдарламалау ортасында сөйлеуді тану жүйесі туралы айтатын болсақ, әртүрлі мүмкіндіктер мен функционалдылықты ұсынатын бірнеше кітапханалар бар.

PyAudio-бұл portaudio кітапханасына Python сілтемелерін беретін кітапхана, бұл әзірлеушілерге аудионы оңай жазуға және ойнатуға мүмкіндік береді. Ол көбінесе аудио кірісті түсіру үшін сөйлеуді тану функциясымен бірге қолданылады. PyAudio аудио ағындарды төмен деңгейде басқаруды қамтамасыз етеді, бұл оны күрделі аудио өңдеу тапсырмалары үшін қолайлы етеді.

Google Cloud Speech-to-Text API – бұл Google ұсынатын бұлтқа негізделген сөйлеуді тану қызметі. Ол тілдердің кең ауқымын және диктормен сөйлесулерді жазу және нақты уақыт режимінде жіберу сияқты мүмкіндіктерді қолдай отырып, сөйлеуді танудың қуатты мүмкіндіктерін ұсынады. Бұл Интернет байланысы мен API кілтін қажет етсе де, сөйлеуді танудың дәл және сенімді нәтижелерін ұсынады. Дегенмен, бұл API-ді пайдалану пайдалану әдісіне байланысты шығындарды тудыруы мүмкін екенін ескеру маңызды.

CMU Sphinx – сөйлеуді тану тапсырмалары үшін әртүрлі құралдар мен кітапханаларды қамтамасыз ететін танымал ашық бастапқы сөйлеуді тану құралы. Ол сөйлеуді тану жүйелерін құру үшін пайдалануға болатын

акустикалық және тілдік модельдер жиынтығын қамтиды. Sphinx өзінің икемділігімен және офлайн режимінде жұмыс істеу қабілетімен танымал. Дегенмен, оның дәлдігі кейбір басқа бұлттық шешімдердегідей жоғары болмауы мүмкін және модельді үйрету және бейімдеу үшін көп күш қажет болады.

DeepSpeech – Mozilla әзірлеген сөйлеуді ашық бастапқы мәтінге түрлендіру қозғалтқышы. Ол сөйлеуді танудың нақты нәтижелеріне қол жеткізу үшін терең оқыту әдістерін қолданады. DeepSpeech-те алдын-ала дайындалған модельдер бар, оларды нақты тапсырмалар немесе тілдер үшін дәл реттеуге болады. Ол бұлтты және жергілікті орналастыру опцияларын ұсынады. DeepSpeech технологиясы дәлдігімен танымал, әсіресе ағылшын тілін тануда, бірақ оны пайдалану үшін оқу және шығару кезінде көбірек есептеу ресурстары қажет.

SpeechRecognition – сөйлеуді тану тапсырмаларын орындау үшін ыңғайлы және қолдануға оңай интерфейсті қамтамасыз ететін Python тіліндегі кітапхана. Бұл әзірлеушілерге әр түрлі көздерден, соның ішінде микрофондар мен алдын-ала жазылған аудио файлдардан сөйлеу кірісін бағдарламалық жасақтама өңдейтін және талдай алатын мәтіндік шығысқа түрлендіруге мүмкіндік береді. Кітапхана Google Speech Recognition, CMU Sphinx және Microsoft Bing Voice Recognition сияқты сөйлеуді танудың бірқатар механизмдерін қолдайды. Ол сондай-ақ WAV және MP3 сияқты жалпы аудио файл пішімдеріне жергілікті қолдауды қамтиды.

SpeechRecognition әзірлеушілерге сөйлеуді тану мүмкіндіктерін қолданбаларына жылдам біріктіруге мүмкіндік беретін сөйлеуді тану қиындықтарынан абстракцияланатын жоғары деңгейлі API ұсынады. Кітапхана кілт сөздерді табу, тілді анықтау және дыбыс көзін таңдау сияқты мүмкіндіктерді ұсынады, бұл оны сөйлеуді тану кезінде қолданудың кең ауқымы үшін жан-жақты құрал етеді. Тану механизмдері мен аудио файл пішімдерін кеңінен қолдаудың арқасында ол әзірлеушілерге сөйлеуді тану тапсырмалары үшін икемді және ыңғайлы шешім ұсынады.

Осы кітапханаларды салыстыра отырып, SpeechRecognition – бұл Google Cloud Speech-to-Text API, CMU Sphinx және басқаларын қоса, сөйлеуді танудың әртүрлі механизмдеріне қол жеткізу үшін қарапайым және біртұтас API ұсынатын әмбебап және ыңғайлы кітапхана. Бұл пайдаланушыларға әртүрлі тану механизмдері арасында оңай ауысуға мүмкіндік береді және офлайн және бұлттық шешімдерге қолдау көрсетеді. Кейбір басқа кітапханалардың жетілдірілген мүмкіндіктері болмаса да, ол Python-да

сөйлеуді тану тапсырмалары үшін ыңғайлы және қолжетімді опцияны ұсынады.

Зерттеу нәтижелері

Адам бейнесі мен сөйлеуін тану жүйелеріне Amazon Rekognition, DLib, Megvii Face++, BetaFace, BioID және басқалары. Аталған жүйелердің соңғысы адамның түрі мен дауысын тану жүйесі болса, қалған жүйелер тек бейне тану бағдарламалары ретінде әлем бойынша танымал болып келеді.

Адам бейнесі мен дауысын қоса тану жүйесін зерттеу мен құрастыру алгоритмдері, тәсілдері [11-14] еңбектерде қарастырылған. Ishwar S. Jadhav өз жұмысында аталған жүйені Matlab бағдарламасында жасаған. Өз жұмысында автор бейнені тану мен дауысты тану модульдерін жеке қарастырып, алынған нәтижелерді графикалық түрде бейнелеген.

Tresadern еңбегінде адамның жеке басын куәландыратын электронды MoBio жүйесін құрастыру ұсынылады. Жұмыста бұл жүйе де адамның бейнесі мен дауысын қатар тануы бойынша ғана жұмыс жасамай, сонымен қоса, жеке тек бейне тану және тек дауысты тану жүйесі ретінде жұмыс жасай алатындығы сипатталады.

Авторлармен құрастырылған адамның бейнесі мен сөйлеуін тану жүйесі Python бағдарламалау ортасында құрастырылған. Python ортасында бағдарлама жасау кезінде кескіндерді тану үшін OpenCV және Matplotlib кітапханалары, сонымен қатар дауысты тану жүйесін құру үшін SpeechRecognition, gTTS, PyAudio және Pytsx3 кітапханалары қолданылды.

Бағдарламаны әзірлеу кезінде операциялық жүйемен жұмыс істеуге және бағдарлама қабылдауы керек дыбыспен жұмыс істеуге арналған кейбір функциялар қолданылды:

```
def talk(words):
    os.system("say" + words)

def osrun(cmd):
    PIPE = subprocess.PIPE
    p = subprocess.Popen(cmd, shell=True,
        stdin=PIPE, stdout=PIPE, stderr=subprocess.
        STDOUT)
```

Сондай-ақ, бағдарламамен сұхбаттасуда қолданылатын микрофонды бақылауға қызмет ететін функция жасалды. Функцияны шақыру арқылы бағдарлама пайдаланушының айтқан сөздерінің бәрін қабылдайды, ал микрофон оны тыңдау үшін қолданылады. Алынған деректер алдымен жолдық түрге ауыстырылады, одан кейін тексеріледі.

```
def command():
    Нысан speech_recognition кітапханасы
    негізінде құрылады және деректерді анықтауға арналған әдіс шақырылады:
```

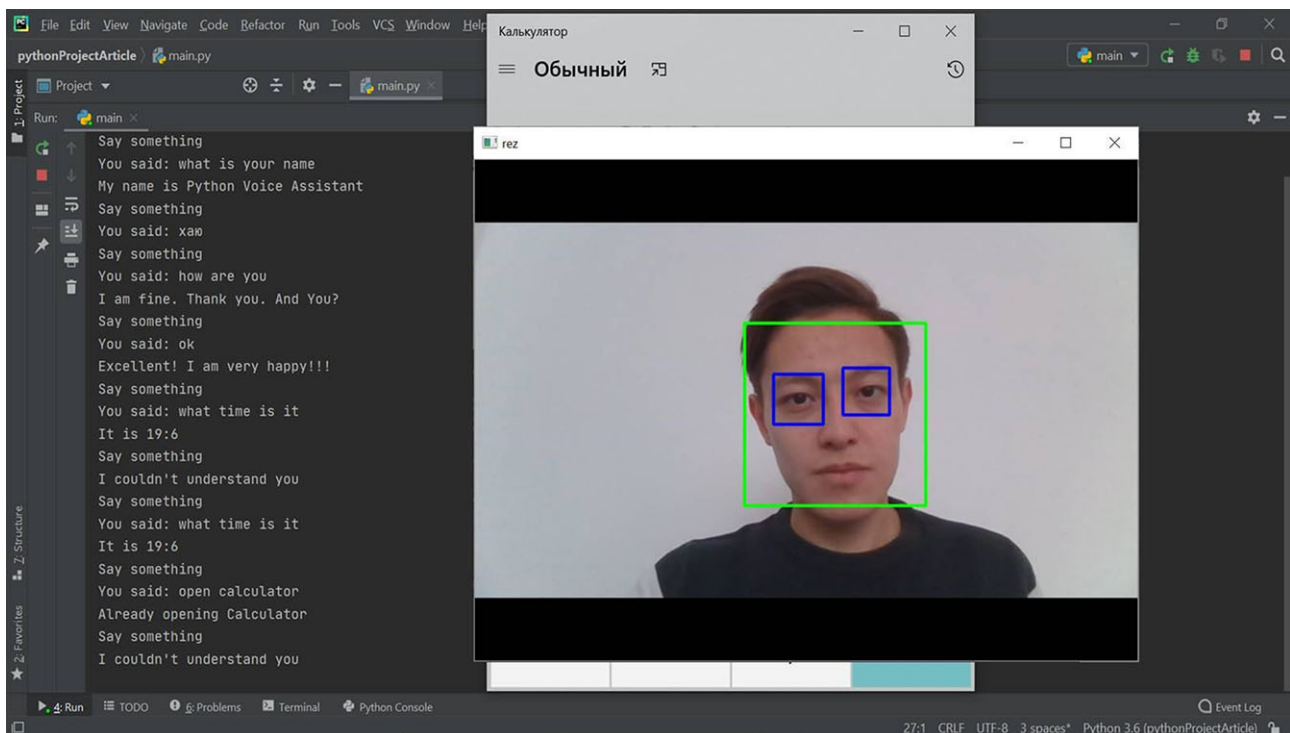
```
r = sr.Recognizer()
with sr.Microphone() as source:
    Дыбысты қабылдау кезінде дыбыстық
    жолдағы бөгде шуды жою үшін
    r.adjust_for_ambient_noise(source,
duration=1) жазылады.
    Lower () бәрін кіші әріпке айналдырады.
    Енді біз жолдың форматында қандай да бір
    деректерді алдық және оларды мына шарт-
    тарда оңай тексере аламыз:
    zadanie=r.recognize_google(audio,
language="uk - Uk"). Lower()
    Пайдаланушы айтқан мәтінді тексеру
    үшін жаңа функция жасалады (zadanie -
    қолданушыдан алынған мәтін):
    defmakeSomething(zadanie)
    importcv2 кітапханасын және осы кітап-
    хананың модульдерін жүктейміз:
    face_cascade_db = cv2.CascadeClassifier
    (cv2.data.harcascades + "haarcascade_
    frontalface_default.xml")
    face_cascade_db = cv2.CascadeClassifier
    (cv2.data.harcascades + "haarcascade_
    frontalface_default.xml")
    eye_cascade = cv2.CascadeClassifier (cv2.
    data.harcascades + "haarcascade_eye.xml")
    Орнатылған модульді қолданып, камера-
    ға жүгінеміз:
    cap = cv2.VideoCapture (0)
    Танылған бейнені сұр реңкке айналдыру
    керек:
```

```
img_gray = cv2.CvtColor (img, cv2.
COLOR_BGR2GRAY)
whileTrue:
    success, img = cap.Read( )
    img_gray = cv2.cvtColor (img, cv2.COLOR_
    BGR2GRAY)
    faces = face_cascade_db.detectMultiScale
    (img_gray, 1.1, 19)
    for (x,y,w,h) infaces:
        cv2.Rectangle(img, (x, y), (x+w, y+h), (0,
        255, 0), 2)
    img_gray_face = img_gray[y:y+h, x:x+w]
    eyes = eye_cascade.detectMultiScale (img_
    gray_face, 1.1, 19)
    for (ex, ey, ew, eh) in eyes:
        cv2.Rectangle (img, (x+ex, y+ey), (x+ex +
        ew, y+ey + eh), (255, 0, 0), 2)
        if cv2.waitKey(1) &0xff== ord ('q'):
            break
    cap.release ( )
    cv2.destroyAllWindows ( )
```

Жоғарыда аталған барлық функциялар-
ды қолдана отырып, Python бағдарламалау
ортасында роботтың адам бейнесі мен сөзін
тану жүйесі құрылды (1-сурет).

Қорытынды

Бізді болашақта не күтіп тұр? Есептеудің
бүкіл тарихында пайдаланушы интерфейс-
терін қолдану саласы табиғи болып келеді.
Экран мен пернетақта осы бағыттағы бір



1-сурет – Python бағдарламалау ортасындағы роботтың адам бейнесі мен сөйлеуін тану жүйесі

```
main x
Say something
You said: what is your name
My name is Python Voice Assistant
Say something
You said: хаю
Say something
You said: how are you
I am fine. Thank you. And You?
Say something
You said: ok
Excellent! I am very happy!!!
Say something
You said: what time is it
It is 19:6
Say something
I couldn't understand you
Say something
You said: what time is it
It is 19:6
Say something
You said: open calculator
Already opening Calculator
Say something
I couldn't understand you
```

2-сурет – Адамның сөйлеуін тану жүйесінің интерфейсі

қадамда болды, тінтуір мен пайдаланушының графикалық интерфейсі басқа қадамда болды. Сенсорлық экрандар – бұл дамудың соңғы кезеңі деп айтсақ та болады. Келесі қадам, кеңейтілген шындықтың, қимылдар мен дауыстық командалардың қоспасынан тұруы әбден мүмкін. Ақыр соңында, сұрақ қою немесе сұхбаттасу қандай да бір мәтінді теруден немесе бірнеше мәліметтерді желіге енгізуден гөрі әлде қайда оңайырақ.

Адам дауыспен жұмыс жасайтын құрылғылармен неғұрлым көп қарым-қатынас жасаса, жүйе алған ақпарат негізінде соғұрлым үрдістер мен заңдылықтарды анықтайды. Содан кейін бұл деректер пайдаланушының қалауы мен талғамын анықтауға пайдаланылуы мүмкін, бұл роботты ақылды ету үшін ұзақ мерзімді пайда әкеледі. Google және Amazon адам эмоцияларын талдауға және оларға жауап беруге қабілетті дауыстық жасанды интеллектті біріктіруді көздейді.

Бұл мақалада Python бағдарламалау ортасы кітапханасының қажетті бағдарламалық модульдерін жүктеу мен импорттаудан бастап, бағдарлама кодының өзі жазғанға дейін адамның бейнесі мен дауысын робот арқылы тануға арналған бағдарлама жасалғандығы туралы қарастырылды. Бағдарламаны тестілеудің нәтижелері бойынша жасалған бағдарламаның адамның сөйлеу қабілетін, беті мен көзін оларды төртбұрышқа қоршап, танып, оның бұйрықтарын орындайтынын және оған жауап беретіндігін көруге болады.

ӘДЕБИЕТТЕР ТІЗІМІ

1. James V Stone (2019). Artificial Intelligence Engines. Sebtel Press, 218.
2. Stuart J.Russell, Peter Norving (2015). Artificial Intelligence. Pearson Education India, 1164.
3. Поп Дейтел, Харви Дейтел. Python: Искусственный интеллект, большие данные и облачные вычисления. – СПб: Изд-во «Питер-Трейд», 2020. – 864 с.
4. Рассел Стюарт, Норвиг Питер. Искусственный интеллект. Современный подход. – М.: Изд-во «Вильямс», 2019. – 1408 с.
5. Anuradha Srinivasaraghavan and Vincy Joseph. (2019) Machine Learning. Wiley, 328.
6. Vinod Chandra S.S., Anand Hareendran S. (2014) Artificial intelligence and machine learning. Prentice Hall India Learning Private Limited; 1st edition, 368.
7. Джوشي Прадик. Искусственный интеллект с примерами на Python / Пер. с англ. – СПб: ООО «Диалектика», 2019. – 448 с.
8. U Dinesh Kumar and Manaranjan Pradhan. (2019) Machine Learning using Python. Wiley.
9. Alberto Artasanchez. (2020) Artificial Intelligence with Python: Your complete guide to building intelligent apps using Python 3.x, 2nd Edition. Packt Publishing, 618.
10. Андреас Мюллер, Сара Гвидо. Введение в машинное обучение с помощью Python: рук-во для специалистов по работе с данными. – М.: Изд-во «Вильямс», 2017. – 480 с.
11. K. Messer et al., «Face authentication test on the BANCA database», Proceedings of the 17th International

- Conference on Pattern Recognition, 2004. ICPR 2004., Cambridge, UK, 2004, pp. 523-532 Vol. 4, doi: 10.1109/ICPR.2004.1333826.
12. P. Tresadern et al., «Mobile Biometrics: Combined Face and Voice Verification for a Mobile Platform», in IEEE Pervasive Computing, vol. 12, no. 1, pp. 79-87, Jan.-Mar. 2013, doi: 10.1109/MPRV.2012.54.
13. Raghavendra, Rao & Kumar (2010) Raghavendra R., Rao A., Kumar GH. Multimodal person verification system using face and speech. Procedia Computer Science. 2010; 2: 181-187. doi: 10.1016/j.procs.2010.11.023.
14. Ishwar S. Jadhav et al, Human Identification using Face and Voice Recognition / (IJCISIT) International Journal of Computer Science and Information Technologies, Vol. 2 (3), 2011, 1248-1252.

Разработка системы распознавания образов и речи человека роботом в программной среде Python

¹***БАЛТАБАЙ Даурен Куанышбекулы**, докторант, dauren.baltabay.95@gmail.com,

¹**ЖУМАШЕВА Жадыра Токановна**, к.т.н., ассоциированный профессор, Zhadyra_14@mail.ru,

¹НАО «Казахский национальный университет имени аль-Фараби», пр. аль-Фараби, 71, Алматы, Казахстан,

*автор-корреспондент.

Аннотация. В последнее время широкое распространение получает технология информационно-поисковых систем распознавания лица с целью идентификации личности. Подсистемы идентификации личности с использованием технологии распознавания образа и речи могут быть использованы в автоматизированных системах безопасности с целью идентификации разыскиваемых персон с одновременной передачей соответствующей оперативной информации службам, занимающимся установлением личности. Целью настоящей работы является разработка системы по распознаванию роботом образа и речи человека в программной среде Python. Причем слово «распознавание» имеет два значения. Первое – пословное преобразование устной речи в соответствующий письменный текст, а также выделения лица, глаз человека и преобразование его в оттенок серого. Второе – правильная реакция на высказывание, содержащее в себе требование или справку диктора и выведение прямоугольных форм на контуре лица и глаз. Сочетанием распознавания речи с ее синтезом можно строить систему общения между человеком и машиной. Мотив интенсивного исследования по распознаванию речи и образа в сферах промышленности, науки и обороны возник в связи с быстрым развитием информационной индустрии, охватывающей компьютеры, автоматизацию офиса, связь, оборону и робототехнику.

Ключевые слова: искусственный интеллект, распознавание речи, распознавание лица, машинное обучение, компьютерное зрение, идентификация, автоматизация процесса, роботы, python, OpenCV.

Development of a System for Recognizing Human Images and Speech by a Robot in the Python Software Environment

¹***BALTABAY Dauren**, Doctoral Student, dauren.baltabay.95@gmail.com,

¹**ZHUMASHEVA Zhadyra**, Cand. of Tech. Sci., Associate Professor, Zhadyra_14@mail.ru,

¹NPJSC «Al-Farabi Kazakh National University», Al-Farabi Avenue, 71, Almaty, Kazakhstan,

*corresponding author.

Abstract. Recently, the technology of information and search systems for facial recognition for the purpose of identifying a person has become widespread. Personal identification subsystems using image and speech recognition technology can be used in automated security

systems for the purpose of identifying wanted persons with the simultaneous transmission of relevant operational information to identity identification services. The purpose of this work is to develop a system for recognizing the robot image and speech of a person in the Python software environment. Moreover, the word «recognition» has two meanings. The first is the word-by-word conversion of spoken speech into the corresponding written text, as well as the selection of the person's face, eyes and converting it to a shade of gray. The second is the correct reaction to a statement containing a request or a reference from the speaker and the drawing of rectangular shapes on the contour of the face and eyes. By combining speech recognition with speech synthesis, you can build a system of communication between a person and a machine. The motive for intensive research on speech and image recognition in the fields of industry, science, and defense arose in connection with the rapid development of the information industry, covering computers, office automation, communications, defense and robotics.

Keywords: artificial intelligence, speech recognition, face recognition, machine learning, computer vision, identification, process automation, robots, python, OpenCV.

REFERENCES

1. James V. Stone (2019). Artificial Intelligence Engines. Sebtel Press, 218.
2. Stuart J. Russell, Peter Norving (2015). Artificial Intelligence. Pearson Education India, 1164.
3. Pop Dejtrel, Harvi Dejtrel. Python: Iskusstvennyj intellekt, bol'shie dannye i oblachnye vychislenija. – Saint Petersburg: Publ. «Piter-Trejd», 2020. – 864 p.
4. Rassel Stjuart, Norvig Piter. Iskusstvennyj intellekt. Sovremennyj podhod. – Moscow: Publ. «Vil'jams», 2019. – 1408 p.
5. Anuradha Srinivasaraghavan and Vincy Joseph. (2019) Machine Learning. Wiley, 328.
6. Vinod Chandra S.S., Anand Hareendran S. Artificial intelligence and machine learning. Prentice Hall India Learning Private Limited; 1st edition (2014), 368.
7. Dzhoshi Pratik. Iskusstvennyj intellekt s primerami na Python / Per. s angl. – Saint Petersburg: OOO «Dialektika», 2019. – 448 p.
8. U Dinesh Kumar and Manaranjan Pradhan. (2019) Machine Learning using Python. Wiley.
9. Alberto Artasanchez. (2020) Artificial Intelligence with Python: Your complete guide to building intelligent apps using Python 3.x, 2nd Edition. Packt Publishing, 618.
10. Andreas Mjuller, Sara Gvido. Vvedenie v mashinnoe obuchenies pomoshh'ju Python: ruk-vo dlja specialistov po rabote s dannymi. – Moscow: Publ. «Vil'jams», 2017. – 480 p.
11. K. Messer et al., «Face authentication test on the BANCA database», Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., Cambridge, UK, 2004, pp. 523-532 Vol. 4, doi: 10.1109/ICPR.2004.1333826.
12. P. Tresadern et al., «Mobile Biometrics: Combined Face and Voice Verification for a Mobile Platform», in IEEE Pervasive Computing, vol. 12, no. 1, pp. 79-87, Jan.-Mar. 2013, doi: 10.1109/MPRV.2012.54.
13. Raghavendra, Rao & Kumar (2010) Raghavendra R., Rao A., Kumar GH. Multimodal person verification system using face and speech. Procedia Computer Science. 2010; 2: 181-187. doi: 10.1016/j.procs.2010.11.023.
14. Ishwar S. Jadhav et al, Human Identification using Face and Voice Recognition / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (3), 2011, 1248-1252.