

ИНФОРМАТИКА

ӘОЖ 004.7;
МҒТАР 49.33.29

<https://doi.org/10.47526/2023-2/2524-0080.05>

Ж.С. ИСМАГУЛОВА¹, Б.Е. АЙТЖАН²

¹техника ғылымдарының кандидаты, Қожса Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті (Қазақстан, Түркістан қ.), e-mail: zhuldyz.ismagulova@ayu.edu.kz

²Қожса Ахмет Ясауи атындағы Халықаралық қазақ-түрік университетінің магистранты (Қазақстан, Түркістан қ.), e-mail: bayanaitzhan.ye@gmail.com

CNN КӨМЕГІМЕН СТУДЕНТТЕР БЕТ ӘЛПЕТИНЕН ЭМОЦИЯЛАРДЫ АНЫҚТАУ

Аңдатпа. Адамның эмоцияларын тану және түсіну, әсіресе білім беру ортасында, үлкен маңызға ие. Бұл зерттеуде оқушылардың бет әлпетіне негізделген эмоцияларын дәл анықтау үшін конволюциялық нейрондық желілерді (CNN) пайдалануға баса назар аударылады. Фациальды белгілерді талдау арқылы білім беру контекстіндегі эмоцияларды тиімді тану және түсіндіру үшін автоматтандырылған жүйе жасалады.

Бұл жұмыс үшін әртүрлі эмоциялармен оқушылардың бет-әлпетін бейнелейтін әртүрлі мәліметтер жиынтығы жасалады. Беттің әртүрлі аймақтарынан маңызды ақпаратты алуға мүмкіндік беретін фациальды бағдарлар алынады. Бұл кескіндер CNN моделін дәл оқыту мен бағалауды қамтамасыз ететін шынайы белгілерді қолдану арқылы мұқият түсіндіріледі.

CNN күрделі кеңістіктік заңдылықтар мен иерархиялық көріністерді тану қабілетіне байланысты белгілерді алу және эмоцияларды жіктеу үшін негізгі технология ретінде таңдалады. Модельді жалпылауға және эмоционалды өрнектердің кең ауқымына бейімделуге мүмкіндік беретін деректерді күшейту және оқытуды тасымалдау сияқты әдістерді қамтитын ауқымды оқыту жүргізілуде.

CNN моделінің өнімділігі дәлдік, болжау дәлдігі, толықтығы және F1 өлшемі сияқты көрсеткіштерді қолдана отырып бағаланады. Ұсынылған тәсілді бет әлпетіндегі эмоцияны танудың қолданыстағы әдістерімен салыстыратын мұқият эксперименттер жүргізіледі, бұл CNN моделінің оқушылардың бет әлпетіндегі эмоцияларды дәл анықтаудағы тамаша өнімділігін көрсетеді.

Кілт сөздер: Конволюциялық нейрондық желілер, мимиканы тану, эмоцияны тану, терең оқыту, білім беру ортасы.

Zh.S. Ismagulova¹, B.Y. Aitzhan²

¹Candidate of Technical Sciences Khoja Akhmet Yassawi International Kazakh-Turkish University (Kazakhstan, Turkistan), e-mail: zhuldyz.ismagulova@ayu.edu.kz

²Master's students of Khoja Akhmet Yassawi International Kazakh-Turkish University (Kazakhstan, Turkistan), e-mail: bayanaitzhan.ye@gmail.com

Determining emotions from students' facial expressions using CNN

Abstract. Recognizing and understanding human emotions, particularly in educational settings, is of great importance. This research focuses on utilizing Convolutional Neural Networks (CNNs) to accurately identify students' emotions based on their facial expressions. By leveraging facial cues, an automated system can be developed to effectively recognize and interpret emotions in educational contexts.

A diverse dataset of facial images featuring students expressing various emotions is carefully

curated for this study. Facial landmarks and action units are extracted to capture essential information from different facial regions. These images are meticulously annotated with ground truth labels, ensuring precise training and evaluation of the CNN model.

CNNs are chosen as the core technology for feature extraction and emotion classification due to their ability to learn intricate spatial patterns and hierarchical representations. Extensive training, including techniques like data augmentation and transfer learning, enables the model to generalize and adapt to a wide range of emotional expressions.

The performance of the CNN model is evaluated using metrics such as accuracy, precision, recall, and F1 score. Thorough experiments compare the proposed CNN approach with existing methods for facial emotion recognition, demonstrating the superior performance of the CNN model in accurately identifying students' emotions from facial expressions.

Keywords: Convolutional Neural Networks, Facial Expression Recognition, Emotion Recognition, Deep Learning, Educational Settings.

Ж.С. Исмагулова¹, Б.Е. Айтжан²

¹*кандидат технических наук Международный казахско-түрецький университет имени Ходжи Ахмеда Ясави (Казахстан, г. Туркестан), e-mail: zhuldyz.ismagulova@ayu.edu.kz*

²*магистрант Международного казахско-түрецького университета имени Ходжи Ахмеда Ясави (Казахстан, г. Туркестан), e-mail: bayanaitzhan.ye@gmail.com*

Определение эмоций по выражению лица студентов с помощью CNN

Аннотация. Распознавание и понимание человеческих эмоций, особенно в образовательной среде, имеет большое значение. В данном исследовании акцент сделан на использовании сверточных нейронных сетей (CNN) для точной идентификации эмоций учащихся на основе их выражений лица. Путем анализа фациальных признаков может быть разработана автоматизированная система для эффективного распознавания и интерпретации эмоций в образовательных контекстах.

Для этой работы составляется разнообразный набор данных, включающий изображения лиц учащихся с различными эмоциями. Извлекаются фациальные ориентиры и действия, позволяющие улавливать важную информацию из различных областей лица. Эти изображения тщательно аннотируются с использованием истинных меток, обеспечивая точное обучение и оценку модели CNN.

CNN выбраны в качестве основной технологии для извлечения признаков и классификации эмоций благодаря их способности распознавать сложные пространственные закономерности и иерархические представления. Проводится обширное обучение, включающее такие методы, как аугментация данных и перенос обучения, что позволяет модели обобщаться и адаптироваться к широкому спектру эмоциональных выражений.

Производится оценка производительности модели CNN с использованием таких метрик, как точность, точность предсказания, полнота и F1-мера. Проводятся тщательные эксперименты, сравнивающие предложенный подход с существующими методами распознавания эмоций по выражению лица, что демонстрирует превосходную производительность модели CNN в точном определении эмоций учащихся по выражению лица.

Ключевые слова: Сверточные нейронные сети, Распознавание выражений лица, Распознавание эмоций, Глубокое обучение, Образовательная среда.

Кіріспе.

Бет – біздің денеміздің керемет мәнерлі және коммуникативті бөлігі. Тіпті бір ауыз сөзбен айтпаса да сан алуан сезімдерді жеткізе алады. Бет әлпетін тану – адамның бет

бейнесінен эмоцияларды анықтау процесі. Ол адамның мінез-құлқы мен мінез-құлқы туралы көп нәрсені аша алады.

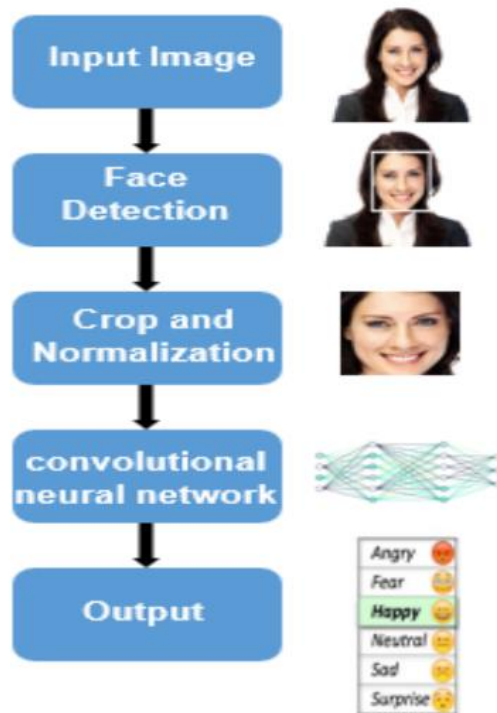
20-шы ғасырда американдық психологтар Экман мен Фризен барлық мәдениеттерге тән алты негізгі эмоцияны анықтады: ашу, қорқыныш, жиіркеніш, қайғы, таңданыс және қуаныш. Бет-әлпетті тану соңғы жылдары, әсіресе клиникалық тәжірибе, әлеуметтік робототехника және білім беру салаларында көп көңіл бөлді. Көптеген зерттеулер эмоциялардың оқуда маңызды рөл атқаратынын көрсетті, ал мұғалімдер дәстүрлі түрде оқушылардың қалай сезінетінін түсіну үшін емтихандарға, сауалнамаларға және бақылауларға сүйенеді. Дегенмен, бұл әдістер көбінесе өте тиімді емес.

Оқушылардың мимикасын талдау арқылы мұғалімдер олардың эмоционалдық күйі туралы құнды түсінікке ие болады және соған сәйкес стратегиялары мен оқу материалдарын реттей алады. Бұл мақаланың мақсаты студенттердің мимикасын талдау үшін конволюционды нейрондық желілерді (CNN) пайдаланатын автоматты жүйені құру арқылы білім беруге эмоцияны тануды енгізу болып табылады. CNN – кескіндерді жіктеуде кеңінен қолданылатын терең оқыту алгоритмінің бір түрі. Біздің жүйе үш кезеңнен тұрады: бетті анықтау, қалыпқа келтіру және эмоцияны тану, жеті эмоцияны таңдау: бейтарап, ашу, қорқыныш, қайғы, қуаныш, таңдану және жиіркеніш.

Келесі бөлімдерде біз байланысты жұмыстарды қарастырамыз, ұсынылған жүйемізді сипаттаймыз, енгізу мәліметтерін ұсынамыз және эксперимент нәтижелерін талқылаймыз. Қорытынды бөлімде біз дамудың болашақ бағыттарын талқылаймыз.

Зерттеу әдістері

Бұл бөлімде біз конволюциялық нейрондық желі (CNN) архитектурасын қолдана отырып, студенттердің бет әлпетін талдауға арналған ұсынылған жүйені сипаттаймыз. Алдымен жүйе кіріс кескініндегі бетті анықтайды, содан кейін анықталған беттер кесіліп, 48x48 өлшеміне дейін қалыпқа келтіріледі. Содан кейін бұл бет суреттері CNN үшін кіріс ретінде пайдаланылады. Ақырында, нәтиже-мимиканы тану нәтижелері (ашу, қуаныш, қайғы, жиіркеніш, таңдану немесе біржақтылық). 1-ші суретте біз ұсынған тәсілдің құрылымы көрсетілген.



1-сурет – Біздің мимиканы тану жүйесінің құрылымы

Convolutional Neural Network (CNN) – бұл басқа кескінді жіктеу алгоритмдерімен салыстырғанда минималды алдын ала өңдеумен визуалды кескін үлгілерін анықтай алатын терең жасанды нейрондық желілер [1]. Бұл желі дәстүрлі алгоритмдерде қолмен жасалған сүзгілерді зерттейтінін білдіреді. CNN қабаттарындағы маңызды бірлік – нейрон. Олар бір-бірімен байланысқан, сондықтан бір қабаттағы нейрондардың шығуы келесі қабаттағы нейрондардың кірісіне айналады.

Құн функциясының ішінара туындыларын есептеу үшін қатенің кері таралу алгоритмі (backpropagation) қолданылады. «Конволюция» термині функциялар картасын жасау үшін кіріс кескініндегі сүзгіні немесе ядроны пайдалануды білдіреді. Шын мәнінде, CNN моделі 2-ші суретте көрсетілгендей қабаттардың 3 түрін қамтиды:

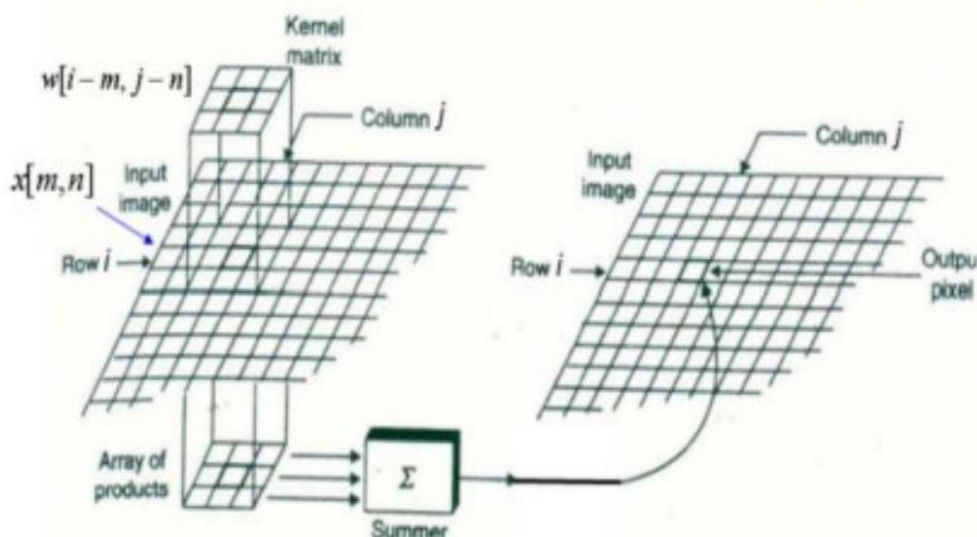


2-сурет – CNN архитектурасы

Конволюция қабаты (Convolution Layer) – кіріс кескінінен белгілерді шығаратын бірінші қабат. Convolutional Neural Network жағдайында конволюцияның негізгі мақсаты-кіріс кескінінен белгілерді алу [2]. Конволюция пикселдер арасындағы кеңістіктік қатынастарды сақтайды, кіріс деректерінің шағын квадраттарын пайдаланып кескін белгілерін зерттейді. Ол екі матрица арасында скаляр көбейтіндісін орындайды, мұнда біреуі кескін, екіншісі ядро. Конволюция формуласы 1-теңдеуде берілген:

$$\text{net}(t, f) = (x * w)[t, f] = \sum^m \sum^n x[m, n]w[t - m, f - n] \quad (1)$$

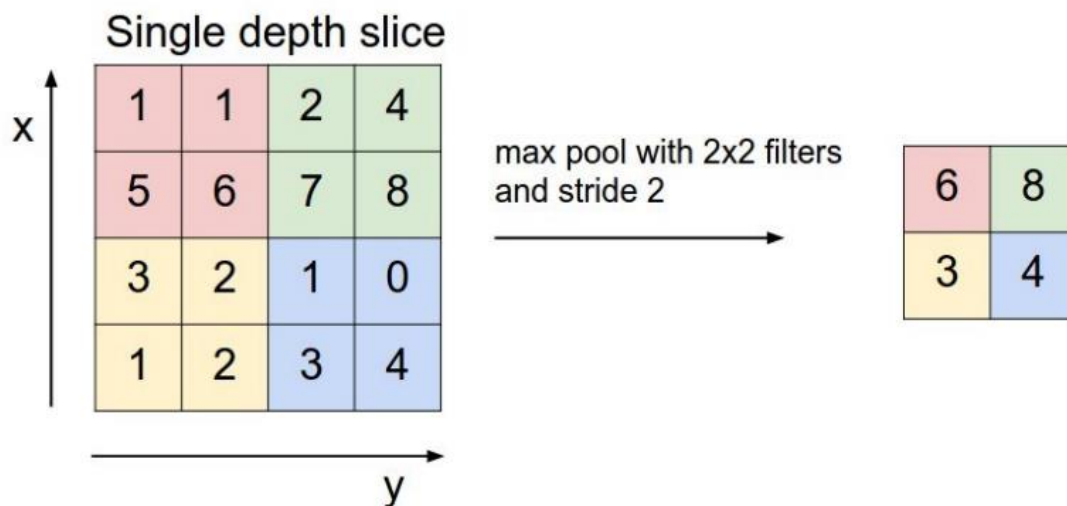
Мұндағы $\text{net}(t, f)$ келесі қабаттағы шығыс, x - кіріс кескіні, w - сүзгі матрицасы, $*$ - конволюция операциясы. 3-ші суретте конволюция қалай жұмыс істейтіні көрсетілген [3].



3-сурет – Конволюция қабаты туралы толық ақпарат

Біріктіру қабаты сүзгіге ұқсайды, ол әрбір нысан картасын конволюция қабатынан шығарады және маңызды ақпаратты сақтай отырып, оны қысады. Ол әр объект картасындағы

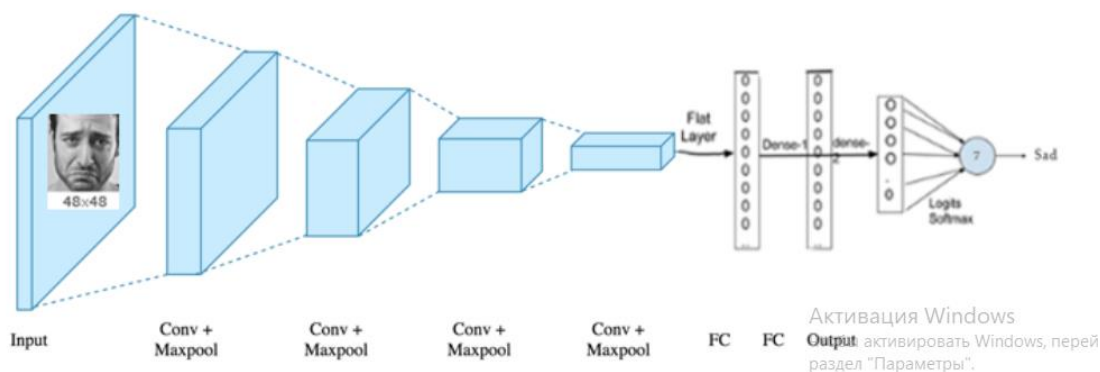
объектілердің шағын блогының максималды, орташа немесе қосындысын алу арқылы объектілерді қыса алады. Біріктірудің мақсаты – кіріс көрінісінің көлемін біртіндеп азайту және желінің кіріс кескініндегі шамалы өзгерістерге төзімділігін арттыру [4]. Бұл жұмыста біз 4-ші суретте көрсетілгендей біріктіру қабатының шығысы ретінде әр блоктың максималды мәнін ғана қолдандық.



4-сурет – Pooling layer туралы мәліметтер

Толық қосылған деңгей – жіктеу үшін пайдаланылатын нейрондық желі қабатының түрі. Ол «толық қосылған» деп аталады, себебі алдыңғы қабаттағы әрбір нейрон келесі қабаттағы әрбір нейронмен байланысқан [5]. Бұл қабаттың мақсаты конвульсиялық және жинақтау қабаттарынан шығысты алу және оны оқыту деректер жинағы негізінде кіріс кескінді әртүрлі санаттарға жіктеу үшін пайдалану болып табылады.

Негізінде конвульсия және біріктіру қабаттары кіріс кескінінен мүмкіндіктерді анықтайды және шығарады, ал толық қосылған қабат бұл мүмкіндіктерді алады және кескінді жіктеу үшін пайдаланады. Бұл кескінді талдау және оның нені бейнелейтінін анықтау процесінің соңғы қадамы сияқты.



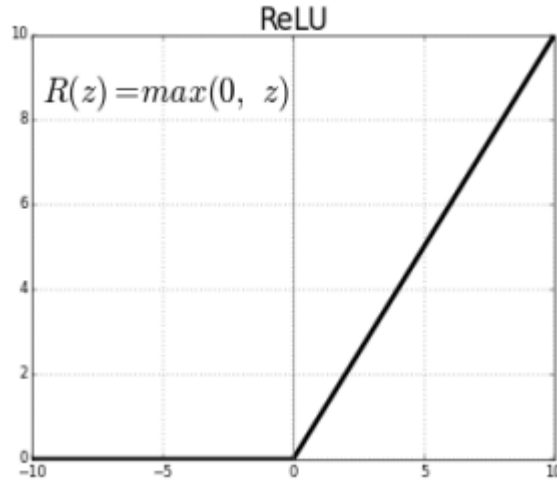
5-сурет – Біздің конволюциялық нейрондық желі моделі

5-ші суретте біздің конволюциялық нейрондық желі моделін ұсынады. Оның құрамында 4 қабатты пулинг бар 4 конволюциялық қабат, сондай-ақ 2 толық байланысқан қабат және 7 эмоция класы softmax қабаты бар. Кіріс кескіні 48×48 өлшемді ақ-қара бет кескіні. Әрбір конволюциялық қабат үшін біз 2-қадаммен 3×3 өлшемді сүзгілерді қолдандық. Пулинг қабаттары үшін біз максималды пулинг қабатын және 2-қадаммен 2×2 ядроны қолдандық. Осылайша, біздің модельге сызықтық емес енгізу үшін біз 2-теңдеуде

анықталған түзетілген сызықтық функцияны (ReLU) [6] қолдандық, бұл соңғы уақытта ең көп қолданылатын активтендіру функциясы.

$$R(z) = \max(0, z)$$

Суретте көрсетілгендей. 6, $R(z)$ нөлге тең, z нөлден аз болғанда, ал $R(z)$ z нөлден үлкен немесе оған тең болғанда z болады. Кесте 1-де біздің модельдің желілік конфигурациясы көрсетілген.



6-сурет – ReLU функциясы

1-кесте – CNN конфигурациясы

Layer type	Size	Stride
Data	48x48	-
Convolution 1	3x3	2
Max Pooling 1	2x2	2
Convolution 2	3x3	2
Max Pooling 2	2x2	2
Convolution 3	3x3	2
Max Pooling 3	2x2	2
Convolution 4	3x3	2
Max Pooling 4	2x2	2
Fully Connected	-	-
Fully Connected	-	-

Деректерді алу

CNN архитектурасын оқыту үшін біз 7-ші суретте көрсетілген FER2013 дерекқорын қолдандық. Ол Google кескін іздеу API көмегімен жасалған және ICML 2013 Challenges кезінде енгізілген. Деректер базасындағы адамдар автоматты түрде 48×48 пиксель өлшеміне дейін қалыпқа келтірілді. [7] FER2013 дерекқорында 7 өрнек белгісі бар 35887 сурет (28709 оқу суреті, тексеруге арналған 3589 сурет және тестілеуге арналған 3589 сурет) бар. Әр эмоцияға арналған суреттер саны 2-ші кестеде келтірілген.



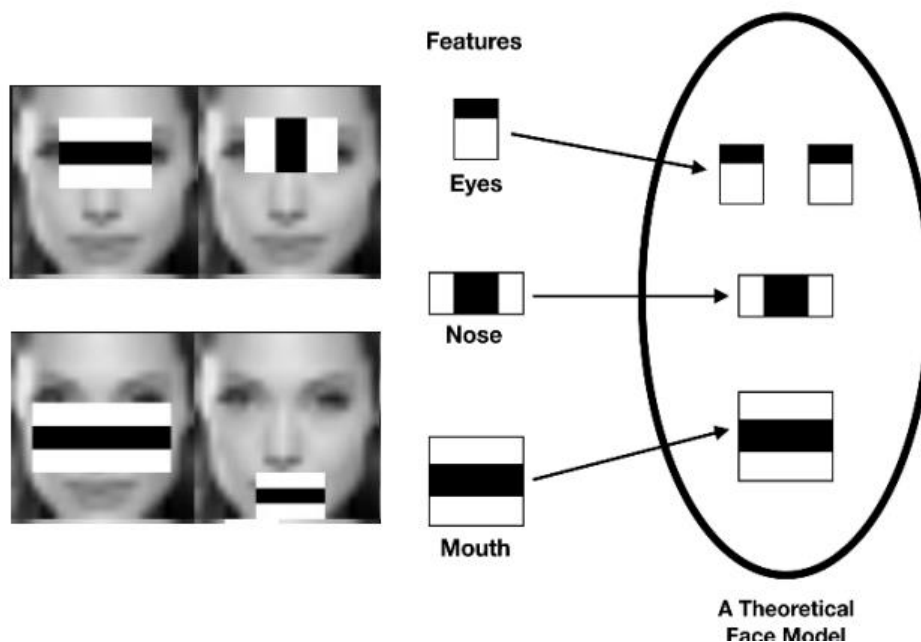
7-сурет – FER 2013 дерекқорынан алынған үлгілер

2-кесте – Fer 2013 дерекқорындағы әрбір эмоция үшін суреттер саны

Emotion label	Emotion	Number of image
0	Angry	4593
1	Disgust	547
2	Fear	5121
3	Happy	8989
4	Sad	6077
5	Surprise	4002
6	Neutral	6198

CNN іске асыру

Біз OpenCV кітапханасын веб-камерадан кадрларды түсіру және 8-ші суретте көрсетілген Haar cascades әдісі арқылы студенттердің бет-әлпетін анықтау үшін қолдандық. Haar Cascades әдісі Фрейд пен оның әріптестері әзірлеген Adaboost Машиналық оқыту алгоритмін пайдаланады, олар 2003 жылы жұмысы үшін Гodelь сыйлығына ие болды [8-9]. Adaboost тиімді жіктеуіштерді құру үшін үлкен жиынтықтан маңызды функциялардың аз санын таңдайды. Сонымен қатар, біз конволюциялық нейрондық желіні (CNN) енгізу үшін жоғары деңгейлі TensorFlow API және Keras көмегімен модель жасадық.



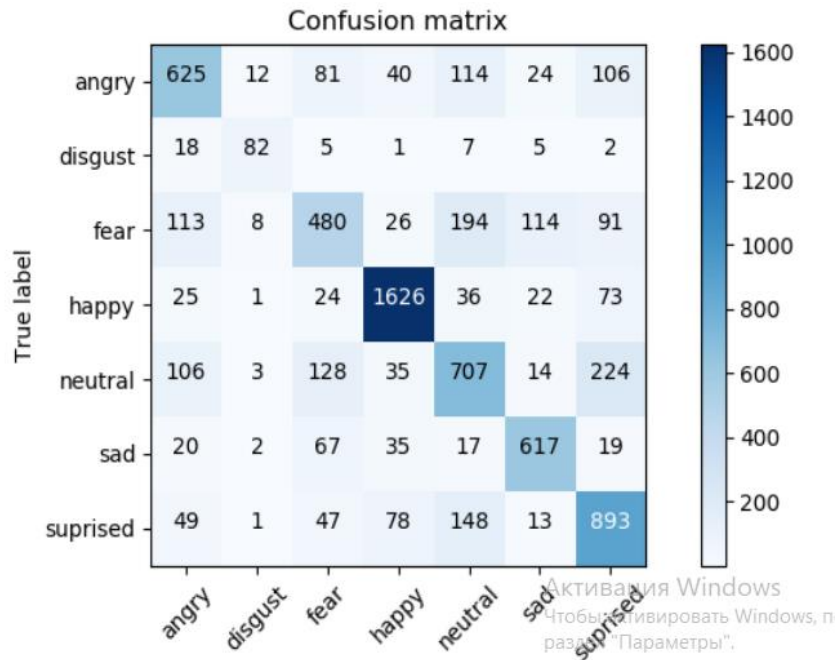
8-сурет – Haar Cascade арқылы бетті анықтау

Талдау мен нәтижелер

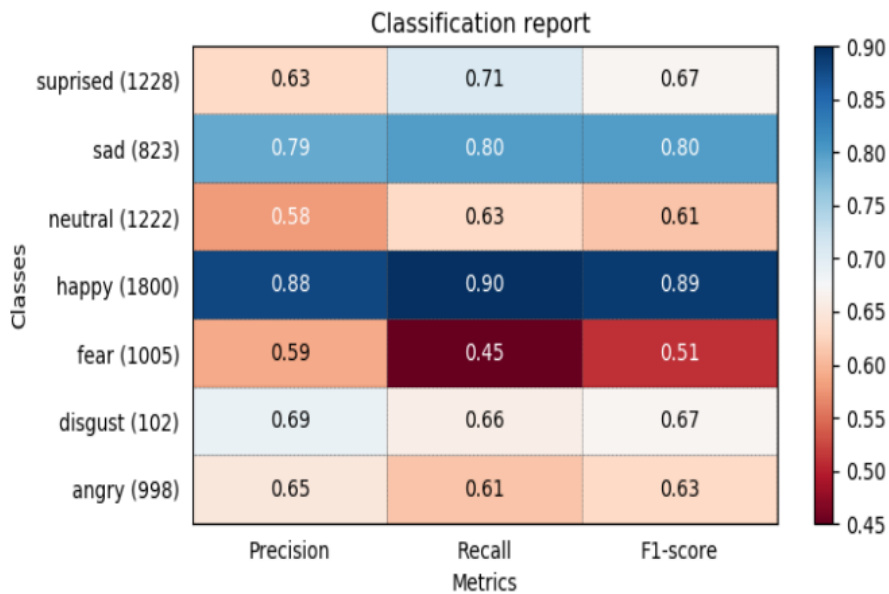
Біз конволюциялық нейрондық желі моделінің өнімділігін бағалау үшін эксперименттер жүргіздік. Модельді үйрету үшін біз FER 2013 дерекқорын қолдандық, онда жеті түрлі эмоцияны білдіретін беттердің суреттері бар: қуаныш, ашу, қайғы, жиіркеніш, бейтараптық, қорқыныш және таңданыс [10].

Суреттерді CNN үлгісіне енгізуге дайындау үшін біз анықталған бет кескіндерінің өлшемін 48x48 пиксельге өзгерттік және оларды сұр реңкке түрлендірдік. Экспериментке біздің факультеттің тоғыз магистранты қатысты, олардың арасында екі көзілдірік таққан. 11-ші суретте осы тоғыз оқушының эмоцияларының нәтижелері берілген. Болжалды эмоциялар қызыл мәтінмен көрсетіледі, ал қызыл жолақ болжанған эмоцияның ықтималдығын білдіреді.

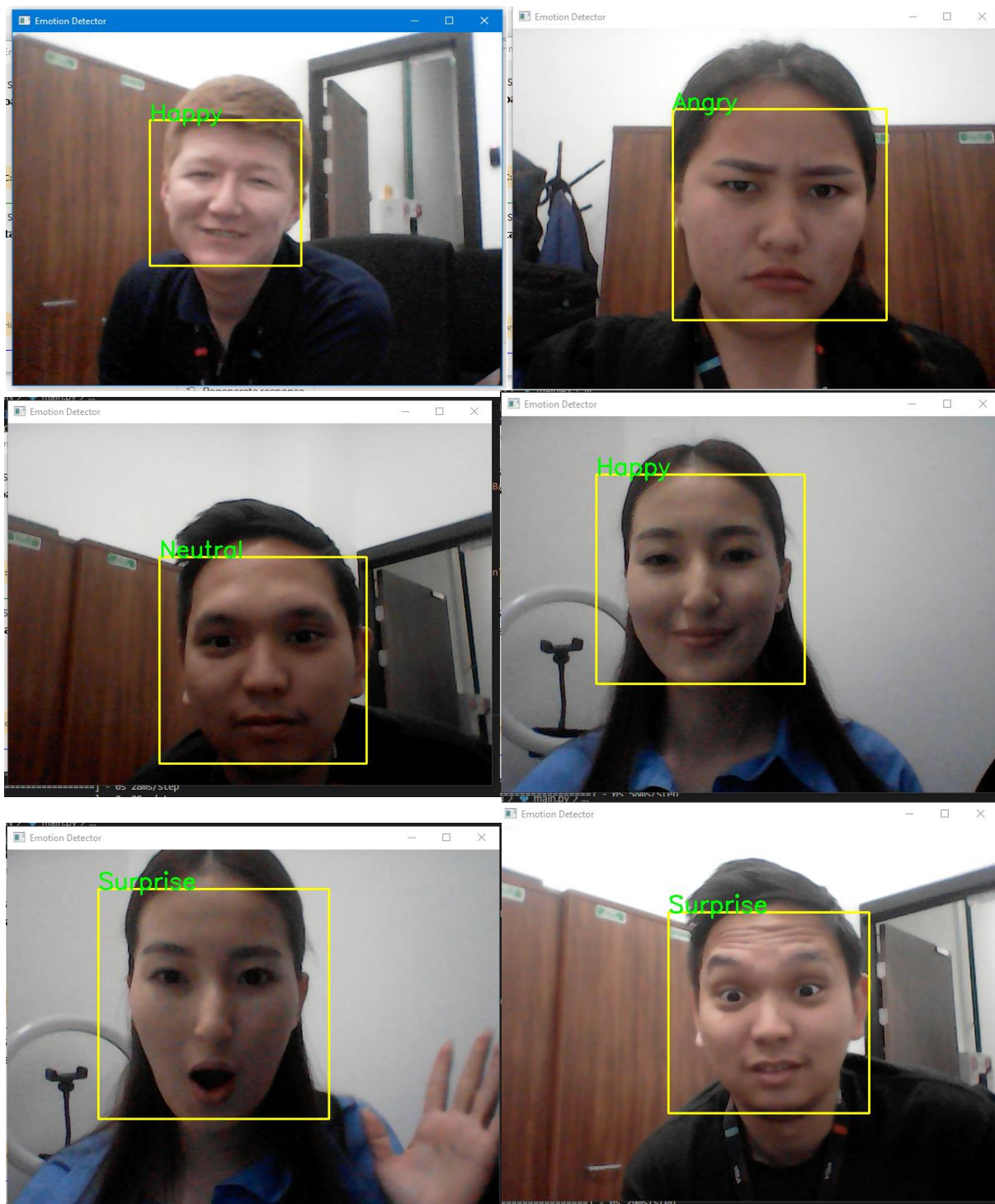
Біз 106 дәуірде 70% дәлдікке қол жеткіздік. Әдісіміздің өнімділігі мен сапасын одан әрі бағалау үшін 12 және 13-ші суреттерде көрсетілгендей қателік матрицасын, дәлдікті, еске түсіруді және F1 ұпайын есептедік. Біздің модель бақытты және таңданған беттерді болжауда өте дәл, бірақ қорқып кеткен беттерді анықтауда қиындық тудырады. және жиі оларды қайғылы нәрселермен шатастырады.



9-сурет – FER 2013 деректер базасында ұсынылған әдістің шатастыру матрицасы



10-сурет – FER 2013 дерекқорында ұсынылған әдістің жіктелуі туралы есеп



11-сурет – Студенттердің бетіндегі эмоцияларды тану нәтижелері

Қорытынды

Зерттеу жұмысымыздың бір бөлігі ретінде біз конволюционды нейрондық желі (CNN) үлгісі негізінде оқушылардың эмоцияларын тану жүйесін жасадық. Біздің ұсынылған модель беттерді анықтауға және оларды жеті эмоцияның біріне жіктеуге арналған бірнеше нейрондық желілерді қамтиды: таңдану, қорқыныш, жиіркеніш, қайғы, қуаныш, ашу және бейтараптық. Жүйе кіріс кескіндеріндегі беттерді табу үшін Нагг тәрізді детекторды

пайдаланады, содан кейін сәйкес эмоцияны анықтау үшін CNN үлгісін қолданады. Біз модельді FER 2013 дерекқорын пайдаланып сынап көрдік және 70% дәлдікке қол жеткіздік. Мұндай технология мұғалімдерге презентациялар немесе лекциялар кезінде студенттердің реакцияларын түсіну үшін пайдалы болуы мүмкін. Біз болашақ зерттеулерде эмоцияны егжей-тегжейлі талдау үшін 3D бет кескіндерін пайдалану мүмкіндігін зерттеу арқылы жүйемізді жетілдіреміз деп үміттенеміз.

REFERENCES

1. Tang, Chuangao, et al. «Automatic facial expression analysis of students in teaching environments». Biometric Recognition: 10th Chinese Conference, CCBR 2015, / Tianjin, China, November 13-15, 2015, Proceedings 10. Springer International Publishing, – 2015.
2. Kim, Yelin, Tolga Soyata, and Reza Feyzi Behnagh. «Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education». IEEE Access 6 (2018): 5308-5331.
3. Lv, Yadan, Zhiyong Feng, and Chao Xu. «Facial expression recognition via deep learning». 2014 international conference on smart computing. IEEE, – 2014.
4. Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001. Vol. 1. Ieee, – 2001.
5. Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi. «Understanding of a convolutional neural network». 2017 international conference on engineering and technology (ICET). Ieee, – 2017.
6. Freund, Yoav, and Robert E. Schapire. «A decision-theoretic generalization of on-line learning and an application to boosting». Journal of computer and system sciences 55.1 (1997): 119–139.
7. Harper, Robert G., Arthur N. Wiens, and Joseph D. Matarazzo. Nonverbal communication: The state of the art. John Wiley & Sons, – 1978.
8. Tang, Xiao-Yu, et al. «Classroom teaching evaluation based on facial expression recognition». Proceedings of the 2020 9th International Conference on Educational and Information Technology. – 2020.
9. Wu, Haobang. «Real Time Facial Expression Recognition for Online Lecture». Wireless Communications and Mobile Computing 2022 – (2022).
10. Guo, Xiaoxu, Juxiang Zhou, and Tianwei Xu. «Evaluation of teaching effectiveness based on classroom micro-expression recognition». International Journal of Performability Engineering 14.11 (2018): 2877.