ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

INFORMACION TECHNOLOGY. AUTOMATION

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ. АВТОМАТИЗАЦІЯ

UDC 004.932

A. Breskina,
S. Antoshchuk, DSc, Prof.
Odessa Polytechnic National University, 1 Shevchenko Ave., Odesa, Ukraine, 65044; e-mail: anastasia.breskina@gmail.com

EMOSTUDENT: A DATASET FOR COMPLEX STUDENT BEHAVIOUR EVALUATION

А. Брескіна, С. Антощук. **Emostudent: набір даних для комплексної оцінки поведінки студентів.** У цій статті розглянуті системи автоматизованого онлайн прокторингу, що основані на штучному інтелекті. Розглянута практична реалізація задачі аналізу поведінки учнів у процесі роботи з ними та набори даних, що використовуються для вирішення цієї задачі. Створена загальна модель обробки даних щодо аналізу активності людини у процесі онлайн-навчання, яка має за завдання аналізувати та описувати активність та емоційний стан учня. Для цього були виявлені різноманітні ознаки, які впливають на оцінку поведінки студента під час самостійної роботи над тестовими завданнями або іспитами. На основі аналізу існуючих наборів даних та з урахуванням проблем сучасних реалізацій автоматизованих систем онлайн прокторингу, зроблено класифікацію параметрів, що використовуються конкретно для аналізу відеопослідовностей в контексті вирішення задачі аналізу поведінки учнів. На основі цієї класифікації розроблено вимоги до необхідного набору даних, які б мінімізували проблеми дисбалансу класів, роблячи акцент не на кількості даних у вибірці, а на якості цієї самої вибірки. На основі розроблених вимог був запропонований набор даних. Джерелом даних для розробленого набору даних була платформа YouTube: використовувались відео з ліцензією Creative Commons. Для організації процесу лейблінгу даних та формування набору даних: використовувалась платформа Amazon SageMaker. Сформований набір даних був доданий до платформ Kaggle та hugging Face. Це дозволяє поширити роботу серед інших вчених та розробників програмного продукту та перевірити на практиці розроблений набір даних у навчанні різноманітних реалізованих моделей штучних нейронних мереж.
Ключові слова: комп'ютерний зір; нейронні мережі; набір даних; розуміння дій; розуміння відео; розуміння емоцій; системи прокторингу на основі штучного інтелекту

A. Breskina, S. Antoshchuk. **Emostudent: a dataset for complex student behaviour evaluation**. This article discusses artificial intelligence based automated online proctoring systems. The practical implementation of the task of analysing students' behaviour in the process of working with these systems and the datasets used to solve this task were considered. A general model for processing data on human activity in the process of online learning has been created, which is aimed at analysing and describing the activity and emotional state of the student. For this purpose, various features that affect the assessment of student behaviour during independent work on test tasks or exams were identified. Based on the analysis of existing datasets and the problems of modern implementations of automated online proctoring systems, a classification of features used specifically for analysing video sequences in the context of solving the problem of analysing student behaviour was made. In accordance with the developed requirements, a dataset was proposed. The data source for the developed dataset was the YouTube platform: videos with a Creative Commons licence were used. Amazon SageMaker platform was utilized to organise the process of data labelling and dataset formation. The generated dataset was added to the Kaggle and Hugging Face platforms. This allows us to utilized the work among other scientists and software developers and to test the developed dataset in practice in training of various implemented models of artificial neural networks.
Keywords: computer vision; neural networks; dataset; action understanding; video understanding; emotion understanding; artificial intelligence-based proctoring systems

Introduction

Distance learning technologies have a long history and are solely linked to the rise of computers and the Internet. The challenge of evaluating students' behaviour during exams and individual assignments has always been present. To address this issue, proctoring systems were introduced. Online proctoring systems serve as information systems designed to oversee the completion of tests or exams and to monitor and assess the honesty of students. These systems simulate the role of a teacher by observing and evaluating student conduct. Originally, synchronous proctoring systems relied on human observers, either the teachers themselves or hired staff, to monitor students. However, to streamline the process and reduce costs, asynchronous proctoring systems were introduced. These systems record the entire test-taking process, allowing teachers to review and analyse the recordings after the exam has concluded. However, the advancement of artificial intelligence methods and models has offered the possibility of automating the assessment of student integrity. The main emphasis of this article revolves around automated online proctoring systems that leverage information systems and artificial

DOI: 10.15276/opu.1.67.2023.07

intelligence to accomplish their objectives. Specifically, it delves into the implementation details of a video processing module that analyzes a student's actions and emotional state during their work.

Literature review and problem statement

There exists a wide range of automated online proctoring systems that rely on artificial intelligence. To categorize and organize these systems based on their implementation and usage in various distance learning platforms, they were divided into three groups [1]: plug-ins, that work in the context of specific distance learning systems, browsers, that provide more access to the student's operating system functionality and work like wrappers, and standalone projects that are separate platforms for testing.

Feedbacks from students and teachers and existing proctoring system implementations analysis have raised two main problems of these implementations: problem of handling personal data and video analysis inaccuracy.

While satisfactory results have been achieved in audio noise analysis and desktop monitoring, the computer vision module faces significant challenges. The computer vision module encountered difficulties and high level of both false positive and false negative responses. This situation was caused by two reasons: a lack of video analysis functionality (eyes tracking, working with random frames), and highly strict requirements for student behavior (e.g., student must look at the monitor all the time). Other developers and students have also reported issues with inadequate image processing and artificial intelligence algorithms, particularly in handling students from diverse racial backgrounds [2]. As a result, certain platforms, like ProctorU, have chosen to remove these problematic AI modules [1, 3].

The examination also revealed that the current systems fail to monitor and assess students' emotional well-being and health [1, 2]. For instance, the widespread expectation of continuously staring at the monitor during prolonged computer usage is not good for health [4, 5]. Consequently, students experience added stress, while teachers are inclined to decrease the number and complexity of exam tasks, compromising the level of knowledge being effectively evaluated.

Basing on the feedback and analysis of the current systems, it has been determined that there is a requirement to enhance the guidelines for student conduct by considering their emotional state and engagement during exams. Additionally, improvements are needed in the computer vision module to effectively monitor student behavior. This necessitates the creation of a comprehensive dataset specific to the development of automated online proctoring systems.

The purpose of the study is to examine the available datasets that focus on analyzing individuals' actions and emotions during computer-based work. It aims to establish the necessary criteria for developing a dataset specifically designed to address online proctoring tasks. Additionally, the study aims to create a prototype of this dataset to serve as a foundational resource for further research and development in the field.

Materials and methods of research

Prior to embarking on the dataset creation process, it is crucial to establish the key parameters that will be incorporated into the comprehensive model for analyzing student behavior. Additionally, it is essential to determine the specific characteristics that will be analyzed by the video processing module. These preliminary definitions serve as a foundation for ensuring a structured and focused approach to dataset formation. The complex analysis of the student's honesty during the performance of individual test tasks includes such components:

– desktop activity data. It includes checking the programs and Internet resources used;

– audio data. It includes checking for harsh background noises and conversations;

– description of student activity and type of emotions. This analysis of student behaviour and emotional state based on the data from the video camera is used to classify the student's rightful actions and at the same time are used to maintain the student's physical and mental health.

Based on this, it is proposed to make a comprehensive model of processing and analysis of student behavior, which will consist of two neural networks. One neural network processes and gives a description of the video obtained from the student's camera. Its input is a video segment of 5 minutes in size. The output is information about the student's emotions and a description of his or her actions.

Another neural network classifies the student's behavior in general, based on video analysis and other factors. Its input is the result of the first neural network, as well as such parameters as

– audio recording of the same time interval as the video,

 ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

– an array of logging entries of unauthorized actions on the desktop (working with unauthorized sites, unfocusing the browser, which is considered to be the use of third-party software). This array consists of a data structure that has three fields: type of unauthorized activity, timestamp, logging message. The output will be the classification student's activity: lawful, cheating, stressed, lost concentration. Since there are only four input parameters and a sequence of rules, the decision tree will be used to classify student's behavior.

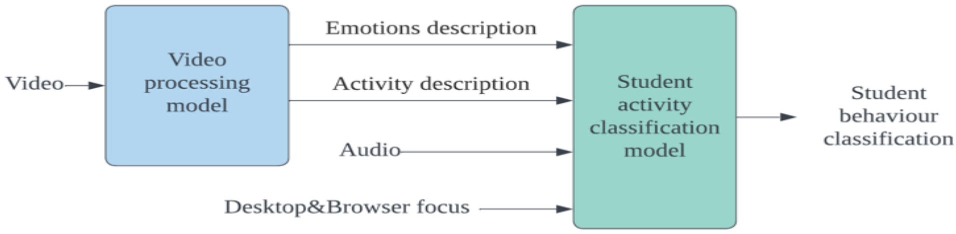

Fig. 1. The structure of complex students' behavior analysis model

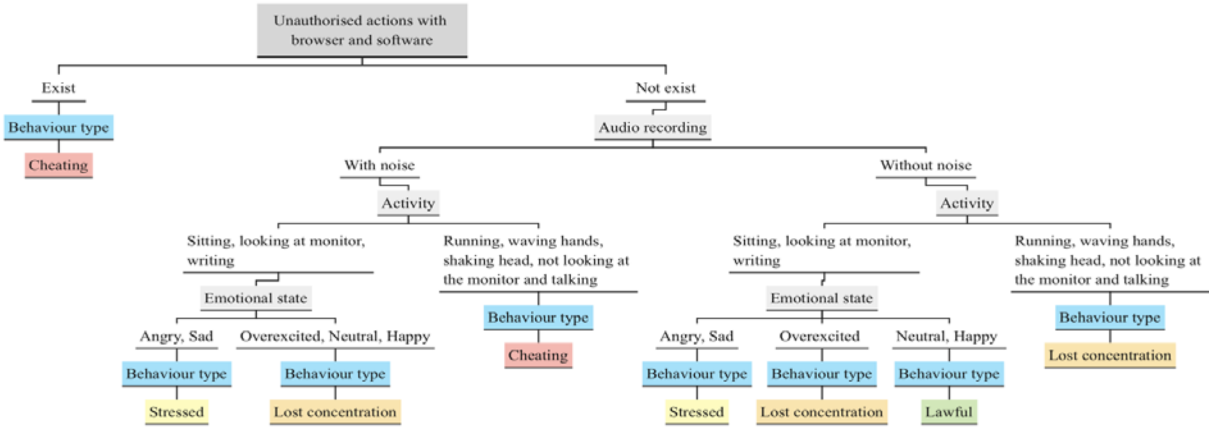

Fig. 2. The structure of behavior classification model

The video processing module is responsible for the comprehensive assessment of the student's behavior: description of movements and of his or her emotional state. For this purpose, a fundamental neural network will be used [6], which is trained using video sequences and text descriptions. The existing pre-trained InternVideo neural network used 39 datasets for its training [6]. However, all of them are created for solving the tasks of recognizing human actions and describing the video. It is proposed to adapt this neural network for detailed description of a person's emotional state. For this purpose, it is necessary to generate an accurate training dataset.

When it comes to data requirements, the primary goal is to have a large number of samples available. However, it is essential to recognize that quantity alone does not guarantee good results. The quality of the data plays a crucial role in achieving meaningful outcomes as well. Therefore, the data must undergo various processes, such as sampling, cleaning, balancing, tagging, and proper formatting. It is essential to investigate the task before assembling the dataset to ensure that the data collected meets the necessary quality standards. Taking these steps ensures that the data is reliable and accurately represents the information needed for the analysis.

The analysis of existing datasets [6] has shown that common modern datasets on the topic of human emotion analysis consist exclusively of photos, which does not allow detailed analysis of emotions in time (a frame can be recognized as one emotion, while it can be a transient change in facial expression). There are also datasets containing videos, but they have a small sample size (e.g., 100 videos [6]).

At the same time, besides the diversity of emotions, there is a problem of analyzing the emotions of people of different races.

To summarize, the following requirements for the dataset data are formed:

– dataset must contain data for people of different ethnic and racial backgrounds,

– dataset must contain examples of neutral, aggressive, and cheerful emotions,

– dataset must include examples of emotions of people of different ages (for the starting age taken in the acceptable age for using the YouTube account);

– dataset must consist of a set of videos and descriptions in them.

To classify the data based on ethnicity, the mix of the European standard classification of cultural and ethnic groups [7] and the List of ethnic groups formed by the UK government have been used [8]. These standards provided a systematic framework for categorizing and organizing the data according to different cultural and ethnic backgrounds. To minimize used data not all spectrum of ethnicities were mentioned, but the base ones that lead us to 9 variants of ethnical groups are shown on Fig. 3.

Based on the guidelines set by Google Class [9], which permits platform usage for individuals aged 13 and above, we propose initiating the data classification process based on age, starting from 13 years old and above. This approach ensures compliance with the platform's policies and facilitates age-appropriate analysis and categorization of the data. Based on the statistical data of secondary schools' students' average age and information about classification of samples of peoples' photos for face recognition [10], it was suggested to divide the sample into 5 groups displayed on Fig. 3.

Dataset data requirements

Ethnical groups classification	
Asian	Chinese
	Indian
	Others
Black, Caribbean or African	Caribbean
	African
	Others
White	North European
	South European
	Others
Others	Middle Eastern
	Latin American
	Native
	Mixed or multiple ethnic groups

Age types (years)
13-19
20-29
30-39
40-59
60 and over

Emotion types described with activity units (AU)	
Happy	AU6+12
	AU6+14
Sad	AU1+15
	AU4+24
Angry	AU9+10
	AU9+22
Overexcited	AU6+(26V27)
	AU7+(26V27)
Neutral	AU57
	AU7+26

Fig. 3. The requirements for student faces dataset

To effectively analyze emotions, we propose utilizing a comprehensive description based on the Emotion Facial Action Coding System (EmFACS) [11] and the Facial Action Coding System Affect Interpretation Dictionary (FACSAID) [12]. These systems encompass facial movements that are linked to various emotions. They provide a detailed breakdown of activity units (AU) and descriptors associated with facial expressions. By combining multiple movement units, we can identify and define specific emotions. Unlike existing solutions that primarily concentrate on gaze direction, this approach enables a more intricate and holistic evaluation of students' emotions. It allows us to offer a richer understanding of their emotional states during the learning process. In the current implementation of the dataset, only five basic emotions were considered, which could be described using a specific set of facial action units. Two types of combinations of action units represented each emotion. However, it is important to note that this area of the dataset will be further enhanced and expanded in future implementations. It is planned to include a broader range of emotions and utilize more comprehensive sets of facial action units to provide a more detailed and nuanced analysis of students' emotional states.

As the result, only for the one representation of persons with the specific ethnicity, age, and emotions there should be approximately 450 video clips in a dataset. In this particular implementation, each class consists of five examples, resulting in a dataset that comprises 2250 video sequences up to 1 minute. The data used to create the dataset was obtained from the YouTube platform, specifically utilizing videos with a Creative Commons license. To streamline the process of labeling the data and forming the dataset, the Amazon SageMaker platform [13] was used. Once the dataset was generated,

ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

it was added to the Kaggle [14] and Hugging Face [15] platforms, making it readily available for other researchers and developers to access and used in their work.

Research results

The structure of a model of complex student performance evaluation has been described. Requirements to the data set that would be optimal in solving the task of analyzing student's emotional state during independent work on test tasks were formed. Following the defined criteria, a dataset was formulated, taking into account the specific requirements. To gather the necessary data, videos from the YouTube platform were utilized, focusing on those with a Creative Commons license.

The data labeling and dataset formation process were efficiently managed using the Amazon SageMaker platform. Once the dataset was generated, it was made accessible on popular platforms such as Kaggle, Hugging Face. This enables the widespread distribution of the work to fellow researchers and software developers, facilitating practical testing of the dataset in training various artificial neural network models.

Conclusion

This article focuses on the implementation of artificial intelligence-based automated online proctoring systems and explores the datasets employed to address the task of analyzing students' behavior within these systems. A comprehensive model has been developed to process data related to human activity during online learning, specifically targeting the analysis and description of student activity and emotional state. Various factors that influence the assessment of student behavior during independent work on tests or exams have been identified as crucial features.

Drawing from an examination of existing datasets and the challenges faced by current implementations of automated online proctoring systems, a classification of features specifically employed for video sequence analysis in the context of analyzing student behavior has been established.

To fulfil the established requirements, a dataset has been proposed, utilizing videos from the YouTube platform that possess a Creative Commons license. The Amazon SageMaker platform was utilized to streamline the data labelling and dataset creation process. The resulting dataset has been shared on prominent platforms such as Kaggle and Hugging Face. This distribution enables the broader scientific community and software developers to access the work and facilitates practical testing of the developed dataset by training various artificial neural network models.

Література

1. Брескіна А. А. Розробка автоматизованої системи онлайн прокторингу. *Вісник сучасних інформаційних технологій*. 2023. Т. 6, № 2. С. 163–173. DOI: https://doi.org/10.15276/hait.06.2023.11.
2. Critten J. The problem with proctoring (part 1). *https://www.cu.edu/*. 2021. URL: https://www.cu.edu/blog/online-teaching-blog/problem-proctoring-part-1. (date of access: 15.12.2022).
3. Jaschik S. ProctorU abandons business based solely on AI Code. *https://www.insidehighered.com/*. 2021. URL: https://www.insidehighered.com/news/2021/05/24/proctoru-abandons-business-based-solely-ai. (date of access: 03.12.2022).
4. Safely using computers at work: guidelines for using computers. *https://www.worksafe.govt.nz/*. 2023. URL: https://www.worksafe.govt.nz/topic-and-industry/work-related-health/musculoskeletal-disorders/ergonomics/safely-using-computers-at-work/. (date of access: 30.07.2022).
5. Working with display screen equipment (DSE). *https://www.hse.gov.uk/*. URL: https://www.hse.gov.uk/pubns/indg36.PDF (date of access: 31.07.2022).
6. Антощук С. Г., Брескіна А. А. Моделі аналізу дій людини в системах прокторингу на основі штучного інтелекту та набори даних для них. *Прикладні аспекти інформаційних технологій*. 2023. Т. 6, № 2. С. 190–200. DOI: https://doi.org/10.15276/aait.06.2023.14.
7. Schneider S. L., Heath A. F. 1 Uncovering ethnic and cultural diversity in Europe: a new classification of ethnic and cultural groups. *3rd international ESS conference*. 2016. URL: https://www.researchgate.net/ publication/313870067. (date of access: 30.05.2023).
8. List of ethnic groups. *https://www.ethnicity-facts-figures.service.gov.uk/*. URL: https://www.ethnicity-facts-figures.service.gov.uk/style-guide/ethnic-groups (date of access: 05.05.2023).
9. About Classroom user accounts. *https://support.google.com/*. URL: https://support.google.com/edu/classroom/answer/7582372 (date of access: 04.03.2023).
10. Horng W.B., Lee C.P., Chen C.W. Classification of age groups based on facial features. *Journal of applied science and engineering*. 2001. Vol. 4, no. 3. 183–192. DOI: https://doi.org/10.6180/jase.2001.4.3.05.

11. The facial action coding system for characterization of human affective response to consumer product-based stimuli: a systematic review / E. A. Clark et al. *Frontiers in psychology*. 2020. Vol. 11. 920. DOI: https://doi.org/10.3389/fpsyg.2020.00920.
12. Farnsworth B. Facial action coding system (FACS) – A Visual Guidebook. *https://imotions.com/*. URL: https://imotions.com/blog/learning/research-fundamentals/facial-action-coding-system (date of access: 19.12.2022).
13. Amazon SageMaker. *https://aws.amazon.com/*. URL: https://aws.amazon.com/sagemaker/ (date of access: 05.05.2023).
14. Kaggle: your machine learning and data science community. *https://www.kaggle.com/*. URL: https://www.kaggle.com/ (date of access: 05.05.2023).
15. Hugging Face – The AI community building the future. *https://huggingface.co/*. URL: https://huggingface.co/ (date of access: 05.05.2023).

References

1. Breskina, A. (2023). Development of an automated online proctoring system. *Herald of advanced information technology*, 6, 2. 163–173. DOI: https://doi.org/10.15276/hait.06.2023.11.
2. Critten, J. (2021). The problem with proctoring (part 1). *https://www.cu.edu/*. Retrieved from: https://www.cu.edu/blog/online-teaching-blog/problem-proctoring-part-1 (date of access: 15.12.2022).
3. Jaschik, S. (2021). ProctorU abandons business based solely on AI Code. *https://www.insidehighered.com/*. Retrieved from: https://www.insidehighered.com/news/2021/05/24/proctoru-abandons-business-based-solely-ai. (date of access: 03.12.2022).
4. Safely using computers at work: guidelines for using computers. (2023). *https://www.worksafe.govt.nz/*. Retrieved from: https://www.worksafe.govt.nz/topic-and-industry/work-related-health/musculoskeletal-disorders/ergonomics/safely-using-computers-at-work/. (date of access: 30.07.2022).
5. Working with display screen equipment (DSE). *https://www.hse.gov.uk/*. Retrieved from: https://www.hse.gov.uk/pubns/indg36.PDF. (date of access: 31.07.2022).
6. Antoshchuk, S. G., & Breskina, A. A. (2023). Human action analysis models in artificial intelligence based proctoring systems and dataset for them. *Applied aspects of information technology*, 6, 2, 190–200. DOI: https://doi.org/10.15276/aait.06.2023.14.
7. Schneider, S. L., & Heath, A. F. (2016). Uncovering ethnic and cultural diversity in Europe: a new classification of ethnic and cultural groups. *3rd international ESS conference*. URL: https://www.researchgate.net/ publication/313870067. (date of access: 30.05.2023).
8. List of ethnic groups. *https://www.ethnicity-facts-figures.service.gov.uk/*. Retrieved from: https://www.ethnicity-facts-figures.service.gov.uk/style-guide/ethnic-groups. (date of access: 05.05.2023).
9. About Classroom user accounts. *https://support.google.com/*. Retrieved from: https://support.google.com/edu/classroom/answer/7582372. (date of access: 04.03.2023).
10. Horng, W.B., Lee, C.P., & Chen, C.W. (2001). Classification of age groups based on facial features. *Journal of applied science and engineering*, 4, 3, 183–192. DOI: https://doi.org/10.6180/jase.2001.4.3.05.
11. Clark, E. A., & et al. (2020). The facial action coding system for characterization of human affective response to consumer product-based stimuli: a systematic review. *Frontiers in psychology*, 11, 920. DOI: https://doi.org/10.3389/fpsyg.2020.00920.
12. Farnsworth, B. Facial action coding system (FACS) – A Visual Guidebook. *https://imotions.com/*. URL: https://imotions.com/blog/learning/research-fundamentals/facial-action-coding-system (date of access: 19.12.2022).
13. Amazon SageMaker. *https://aws.amazon.com/*. URL: https://aws.amazon.com/sagemaker/. (date of access: 05.05.2023).
14. Kaggle: your machine learning and data science community. *https://www.kaggle.com/*. URL: https://www.kaggle.com/ (date of access: 05.05.2023).
15. Hugging Face – The AI community building the future. *https://huggingface.co/*. URL: https://huggingface.co/ (date of access: 05.05.2023).

Брескіна Анастасія Аршавирівна; Anastasiia Breskina, ORCID: https://orcid.org/0000-0002-3165-6788
Антощук Світлана Григорівна; Svitlana Antoshchuk, ORCID: https://orcid.org/0000-0002-9346-145X

Received May 07, 2023

Accepted June 16, 2023