Harnessing Emotions for Effective Online Learning: A Systematic Review of Emotion Recognition Techniques

Mangaras Yanu Florestiyanto*[1], Herman Dwi Surjono[2], Handaru Jati[3]

[1,2,3] State University of Yogyakarta, Jl. Colombo No.1 Karangmalang Yogyakarta 55281, Indonesia

[1] Universitas Pembangunan Nasional Veteran Yogyakarta, Jl. Padjajaran, Sleman Yogyakarta 55283, Indonesia

mangaras.yanu@upnyk.ac.id

Abstract. The growing utilization of computer vision technology to enhance learning settings has captivated contemporary literature in the field of educational research. The integration of affective tutoring systems (ATS), student emotion recognition systems (SERS), sentiment analysis, and multi-agent systems (MAS) has dramatically improved the online learning environment. The absence of prior technologies hindered the progress of online learning environments since there was a deficiency in a corrective mechanism to address the disadvantaged perspectives of tutors in accurately documenting facts inside their instructional settings. Consequently, the integration of computer vision technology has been seen to augment the online learning milieu by equipping instructors with improved monitoring capabilities and supportive interventions. The present study employed a systematic review methodology, including a total of 33 scholarly works, to comprehensively examine the utilization of technology in educational settings. The study investigated several approaches for obtaining and interpreting emotional data, machine learning algorithms, adaptability to individual learners, camera resolution, and ethical considerations. This study aimed to elucidate the advantages and disadvantages of these systems. This study demonstrates that the implementation of emotion detection systems has the potential to enhance the effectiveness of online learning. These technologies enable the quantification and analysis of students' emotional responses, therefore facilitating instructors in improving their teaching strategies and instructional materials. By utilizing real-time feedback on students' emotions, educators have the ability to adapt their instructional strategies in order to enhance student engagement and improve academic achievement.

Keywords: Online Learning, Affective-Adaptive-Intelligent Tutoring System, Emotion Recognition.

1 Introduction

The shift towards online learning platforms has significantly impacted students' attention spans and necessitated teachers to adapt their traditional teaching approaches to

A. Putro Suryotomo and H. Cahya Rustamaji (eds.), *Proceedings of the 2023 1st International Conference on Advanced Informatics and Intelligent Information Systems (ICAI3S 2023)*,
Advances in Intelligent Systems Research 181,
https://doi.org/10.2991/978-94-6463-366-5_8

better suit the online learning environment [1]. The utilization of computer vision technology in the field of emotion identification represents a significant advancement, wherein digital images are analyzed, and the duration of students' focus is measured [2],[3]. The attention of students plays a crucial role in shaping the learning environment. Various educators have used strategies for experiential learning in order to sustain students' engagement and introduce approaches that enhance their ability to focus. Nevertheless, the presence of worry, boredom, lack of understanding, and an overpowering reaction to the learning environment can significantly impede the advancement of the educational setting [4]. Practitioners effectively identified pupils inside physical classroom settings by observing their limited capacity for active participation in class discussions. On the contrary, in online courses, the reliance on technology and external elements has notably diminished teachers' ability to ascertain students' inclination to engage in classroom activities.

The integration of Kohonen neural mapping into an affective tutoring system (ATS) for online education is easily achieved since it involves the examination of clustering patterns and the utilization of self-organizing maps [5]. The considerable potential of artificial intelligence (AI) to autonomously create applicant tracking systems (ATS) in an unsupervised environment is noteworthy. The contemporary implementation of the Automatic Target System (ATS) encompasses the consideration of students' attention span [6]. The correlation between student retention and their capacity to sustain attention throughout class has been observed [7]. Nevertheless, there existed a substantial requirement to quantitatively measure the attention span of students since previous studies predominantly regarded this element as qualitative and relied on students' bodily reactions, such as response rate, body language, and facial expressions, for assessment. In the context of online learning, it is observed that the instructor is limited to incorporating just facial expressions. However, the capacity to see pupils through the smaller screens of their gadgets sometimes leads to a disregard for their attention, therefore constraining their potential to enhance the educational setting.

The objective of this study is to examine the efficacy of emotion recognition systems within the context of online learning environments. Specifically, the research will concentrate on various aspects, including feature extraction, feature selection, emotion analysis, the specific emotions being evaluated (surprise, interest, confusion, and awe), as well as the impact of individual differences, camera quality, and ethical considerations. This study investigates the effects of these systems on student and teacher learning outcomes and levels of engagement while also addressing the challenges associated with their deployment in educational settings.

2 Methodology

2.1 Research Questions

The research questions (RQs) were formulated in order to establish and uphold the review's central objective. The entities' design was aided by employing the Population, Intervention, Comparison, Outcomes, and Context (PICOC) criteria, as described by Kitchenham and Charters [8]. The layout of the research questions utilizing the PICOC framework is depicted in Table I.

Table 1. PICOC.

PICOC	Description
Population	Online learning, Intelligent Tutoring Systems, Affective Tutoring Systems.
Intervention	Emotion recognition using computer vision, Computer vision algorithms, and datasets.
Comparison	Compared to standard online learning settings without emotion recognition.
Outcomes	Better online learning, student enthusiasm, increased engagement, enhanced academic performance, improved instructor performance, increased emotional well-being of online students, and user satisfaction with online learning platforms.
Context	Various online learning platforms and environments (e.g., MOOCs, LMS), different academic subjects and courses, ethical considerations related to privacy and data security, technical feasibility, and limitations of implementing computer vision systems in online learning.

The study goal has led to the formulation of the subsequent research questions as follows:

1) What kind of emotions are analyses based on facial expressions?
2) What kind of methods are used for extracting features of facial expressions?
3) What kind of methods are used for classifying emotion based on facial expression?

2.2 Inclusion and Exclusion Criteria

Table 2, shown below, displays the criteria for inclusion and exclusion that have been integrated into this study.

Table 2. Inclusion and exclusion criteria.

Criteria	Inclusion	Exclusion
Type of course	Emotional recognition system in the learning environment	Traditional learning systems
Language	English	Any language other than English
Publication date	2020 – 2023 (September)	Earlier than 2020
Publication Type	Scopus-accredited or Scopus-indexed journals	Grey literature

Keywords	Student Emotion Recognition System, Facial Expression* Tutoring System*, affective tutoring system, Micro Expression*, Facial Expressions Landmark*, Facial Expressions Recognition*, Facial Emotion Recognition*	Other keywords that are not included in inclusion criteria.

2.3 Data Source and Literature Research

This research conducted an extensive literature search, resulting in a total of more than 1,900 studies from reputable sources in the Scopus database. The investigation included examining previous scholarly articles focusing on the study of emotion recognition systems in virtual learning environments.

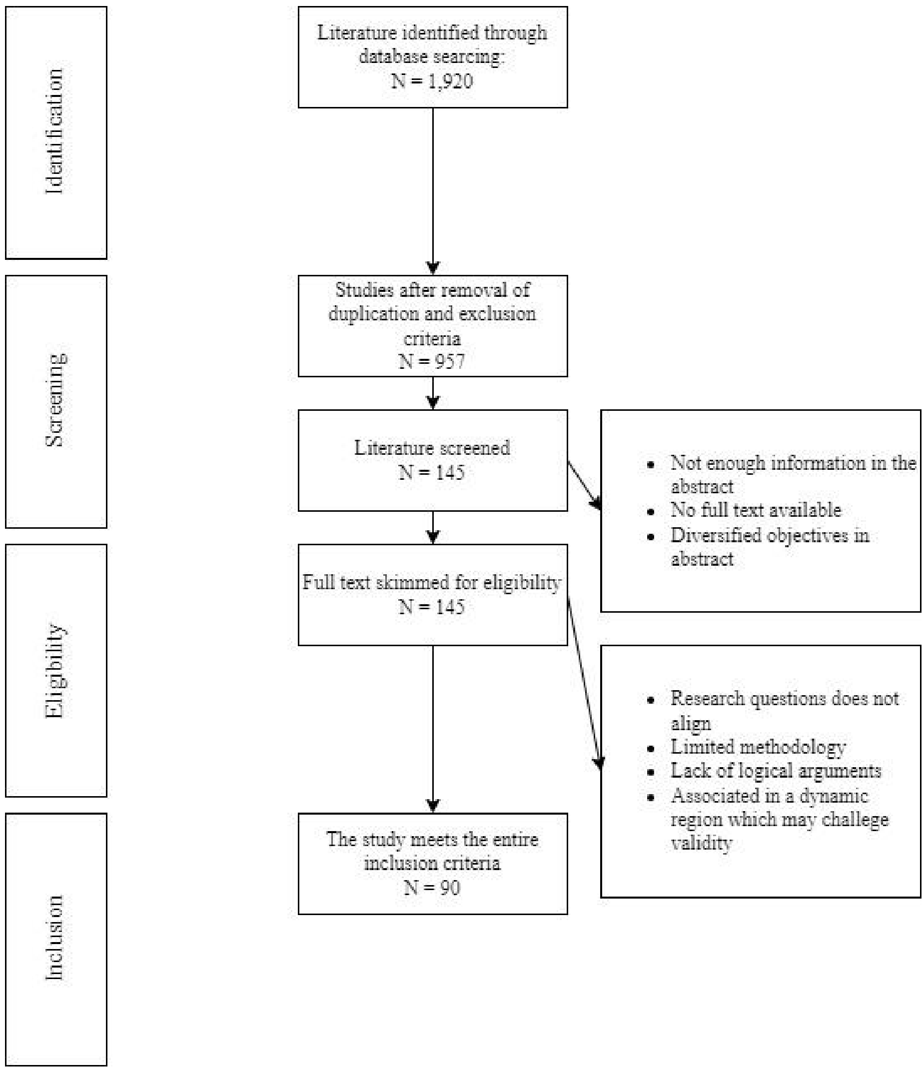

Fig. 1. PRISMA guidelines.

3 Research Results

3.1 Diversity of Student Emotion in Online Learning Environment

The emotional experiences of students in the process of learning may be categorized into four unique reflections pertaining to the learning environment, namely surprise, intrigue, bewilderment, and wonder [9]. The aforementioned analyses of the learning environment ascertain its efficacy. Educators in virtual learning settings endeavor to cultivate students' engagement and minimize the occurrence of uncertainty, given the students' adaptive capacity to resolve their tensions [10]. Furthermore, it has been observed that the unique attributes exhibited by students within the classroom setting serve as the primary obstacle preventing teachers from establishing a standardized and effective learning environment. Consequently, the present study was devised to investigate the potential impact of student emotions on attention span and learning outcomes within the context of online education. Multi-agent systems (MAS), affective computing, and sentiment analysis are often employed techniques for comprehending the range of emotions shown by students within an educational setting.

3.2 Affective Computing

The integration of technology in education has become pervasive in contemporary culture, characterized by its rapid speed and reliance on digital platforms. The implementation of emotional computing in the field of education signifies a substantial transformation. The integration of humans and machines across many digital platforms facilitates the field of affective computing in comprehending and reacting to human emotions [11]. Within this paradigm change in education, the traditional teacher-student interaction undergoes a transformation into a digital interface between the computer and the student. The utilization of affective computing technologies, which employ artificial intelligence to assess students' performance and attentiveness, has resulted in increased efficiency within contemporary learning settings [12]. This section provides an analysis of the advantages and disadvantages associated with the utilization of emotional computing in the field of education.

Affective computing has been found to enhance the efficiency of digital learning environments. The capacity to evaluate students' emotions and modify teaching materials is a notable advantage. The algorithms utilized in affective computing systems have the capability to identify various emotions such as discontent, engagement, boredom, and confusion [11]. These systems use the collected data to adjust the pace and complexity of learning content in real-time, ensuring that students remain challenged and interested.

The flexibility, as mentioned earlier, is particularly advantageous in the context of remote or online learning, as it necessitates educators to possess heightened levels of empathy and establish a more personalized connection with their students. Affective computing technologies assist digital teachers in effectively and empathetically responding by assessing students' emotions and behaviors [12]. The use of a tailored strategy enhances student happiness and motivation through the establishment of a connection between students and the digital learning environment. The flexibility, as mentioned earlier, confers advantages, but it also engenders concerns about excessive reliance on technology. Educators have consistently evaluated the emotional states of students and adapted their instructional methods accordingly. A potential consequence of affective computing is the possible displacement of human intuition and knowledge by computerized decision-making.

Affective computing pertains to the examination of pedagogical challenges in both traditional and digital learning contexts. Conventional teaching methods often need more variety and might become repetitive. Affective computing enhances the flexibility and engagement of the learning process. When a student experiences a decline in engagement, the system has the capability to seamlessly transition to multimedia, interactive activities, or gamified content [11]. The use of dynamic changes serves to mitigate feelings of boredom and cultivate student engagement in the learning process.

Affective computing aims to tackle the issue of reduced interaction between students and teachers, particularly in large class settings. Affective computing utilizes sentiment analysis and emotional recognition techniques to tailor and provide personalized responses. This enables the system to identify students who are experiencing difficulties and provide them with prompt assistance [13]. The outcome is a learning environment that is characterized by inclusivity and support, ensuring that every student is provided with equal opportunities for success.

Nevertheless, the utilization of technology to enhance pedagogy has several drawbacks. Affective computing systems can record and analyze the emotional data of students, which gives rise to concerns regarding privacy and data security. The management and safeguarding of this data give rise to significant ethical implications.

Emotional computing in education has been found to enhance student attention and engagement, which presents a notable challenge for educators in the era of digital technology. Maintaining student concentration is a formidable challenge in the age of pervasive digital distractions. Utilizing affective computing for real-time student emotion monitoring is a distinctive and innovative approach. The system can detect instances when students' attention diminishes and subsequently re-engage them through the provision of explanations or interactive activities [12].

By conducting historical assessments of students' emotional reactions and academic achievements, affective computing has the potential to uncover patterns in learning. Based on this acquired knowledge, the system can propose effective study plans and recommend relevant educational materials. If a learner encounters difficulty comprehending an idea, the system can provide further information or supplementary practice activities [11]. Individualized assistance enhances active participation and cultivates development by acknowledging the significance of exertion and progress.

In the realm of education, the integration of emotional computing has ultimately resulted in enhanced efficiency, flexibility, and pedagogical practices. Nevertheless, it is imperative to conduct a thorough critical analysis of the limits and ethical concerns associated with this technology. The use of affective computing could revolutionize the field of education by enhancing student engagement and learning outcomes. However, this advancement also raises concerns regarding teaching methodologies, safeguarding privacy, and ensuring the security of collected data. The effective management of the advantages and disadvantages associated with emotional computing in education will be of utmost importance, given the significant impact this technology has on the learning process.

3.3 Student Emotion Recognition System

Analysis of SERS in Digital Education. The Student Emotion Recognition System (SERS) is considered to be one of the pioneering technological advancements in the field of quantifying the effectiveness of the learning environment [4]. Traditionally, the assessment of the effectiveness of the learning environment and students' emotional responses has been based on qualitative feedback on the learning platform and its comparison with the students' end-of-year grades. However, this approach may result in a biased evaluation. On the contrary, the utilization of computer vision in developing Surface-Enhanced Raman Spectroscopy (SERS) has significantly enhanced the efficacy of the digital learning environment by allowing instructors to devote their attention to conceptual instruction during class sessions. The observation was made that tutors dedicated their time to elucidating the ideas while also considering the pupils' engagement in the classroom. Nevertheless, tutors needed to give more attention to introverted pupils and those excluded. Therefore, integrating surface-enhanced Raman spectroscopy (SERS) effectively enhanced the capabilities of online education. The implementation of SERS involves many steps, including system design, eye detection, and head rotation, as mentioned in reference [14]. Each of the preceding modules was covered separately, as outlined by the following points.

1) The system has been built based on the local binary pattern (LBP) algorithm, a robust method for evaluating students' engagement in the classroom. This approach integrates eye detection with head rotation analysis [15]. The observation was made that the system may effectively assess students' degree of knowledge by analyzing involuntary or voluntary motions of the head and eye.
2) Incorporating a facial recognition system can be utilized for eye detection [16]. However, the usefulness of the findings may need to be improved by the reliance on the camera quality of users to capture their face and eye rotation in the design. Hence, it is essential for the educational setting to establish a uniform practice of utilizing electronic devices among students to enhance the effectiveness of self-regulated learning strategies (SERS).
3) The head rotation module is responsible for detecting head movement based on the angle of the eyes and their visibility. Using Student Engagement and Response Systems (SERS) prompts the instructor to receive a message about the students' diminished focus during the lecture. This enables the instructor to assess their behavior based on the interaction. Moreover, a methodology has been proposed to facilitate the use of Surface-Enhanced Raman Spectroscopy (SERS) in educational environments.

Implementation Methodology. The Viola-Jones algorithm is a machine learning approach utilized for object recognition, as documented in the publication "Rapid object detection using a boosted cascade of simple features" [18]. The system's tendency was first designed for facial recognition, but it subsequently progressed to encompass instructional platforms utilizing Local Binary Patterns (LBP) [19]. The subsequent guidelines were created to assess the efficacy of the Viola-Jones algorithm and LBP in discerning the levels of high, medium, and low concentrations exhibited by students in the context of online learning.

1) The analysis of high concentration among students is based on the fixed posture of their head and eye on the screen, as shown by previous research [20]. When individuals are not oriented toward the screen, it may be observed that

their attention is diverted by background activities, potentially hindering their capacity to engage with the instructional content provided by the teacher fully.

2) The medium concentration identification is implemented using the Local Binary Patterns (LBP) and Viola-Jones algorithm [21], which are regarded with skepticism. The rationale behind the skewness observed in the decisive technique for assessing concentration levels is in its methodology of examining a singular eye directed towards the screen. On the contrary, the second eye remains concealed while the head is held in a stationary position.
3) The determination of low concentration is based on the periodic movement of the head and eye for an extended duration (22). The Viola-Jones algorithm effectively detects low attention levels by accurately identifying instances where students exhibit greater focus on background activities rather than on lectures.

There has been an argument made suggesting that the act of simplifying head and eye responses during classroom instruction may significantly diminish the efficacy of SERS, as stated in reference [23]. Nevertheless, research has indicated that students predominantly exhibit engagement behaviors through the use of their heads and eyes, regardless of their traits, such as introversion or extroversion [22]. Hence, it is imperative to acknowledge the significant efficacy of the Viola-Jones algorithm and Local Binary Patterns (LBP) in the contemporary educational framework.

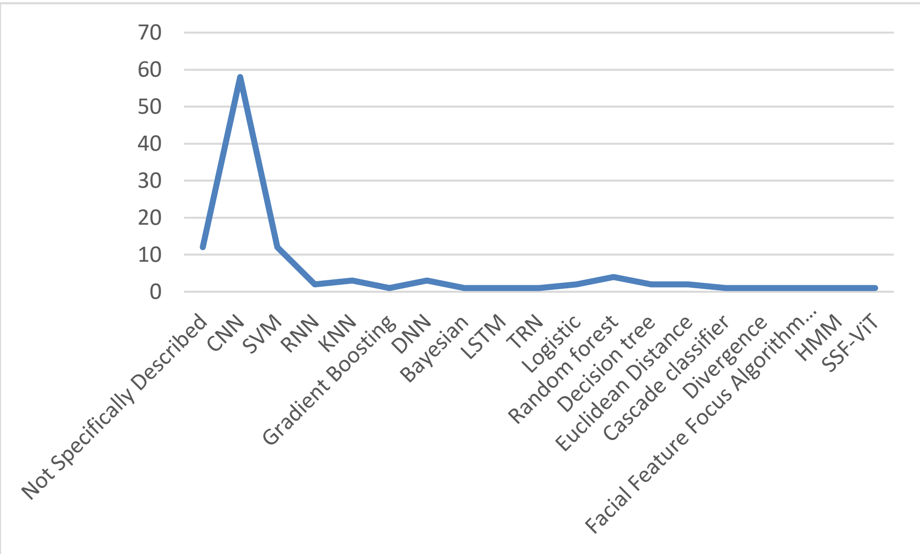

Fig. 2. Number of studies based on the Used Method for Classification.

Benefits of Emotion Recognition System for Online Learning Environment. Emotion recognition systems hold promising potential for enriching the online learning environment by understanding and responding to student emotions. Here are some key benefits they offer:

1) Enhanced Learning Environment. Teachers' capacity to effectively interact with pupils [24] has improved the traditional classroom learning environment. Educators engaged in introspection over the conduct of their pupils to optimize the educational milieu. Nevertheless, the proliferation of online learning platforms has resulted in an exacerbation of communication barriers between students and professors. Consequently, there emerged an enhanced need for technological solutions to foster transparency between the stakeholders, as mentioned earlier [25]. Implementing emotional recognition systems online has facilitated a more productive learning environment by measuring student reaction ratings [26]. Moreover, it enabled the instructors to broaden their pedagogical strategies to capture the students' desired engagement.
2) Better Decision-Making Strategies for Teachers. Traditionally, educators relied on immediate judgments to make decisions, which often led to misaligned choices that exacerbated the decline of the learning environment [27]. The emotional recognition system utilizes an AI and ML-trained model to provide precise outcomes based on the expressions and actions of students inside the digital learning environment. This enables teachers to adopt a standardized approach to improve the learning environment [28].
3) Engagement of Students and Teachers. The participation of both students and teachers in online learning settings was notably hindered due to heightened obstacles in communication. The emotional recognition system has been particularly significant in bolstering the efficacy of the affective computing approach in fostering increased levels of interaction between students and teachers [22]. Educators exhibited a growing consciousness of their students' conduct, facilitating a thorough refinement of their pedagogical strategies within the educational setting.

Challenges of Computer Vision for Education Technology. Implementing computer vision technology in educational settings engenders a conundrum wherein preserving data privacy and security is juxtaposed with the imperative of ensuring fairness and mitigating prejudice. To ensure ethical and equitable utilization of technology, institutions must adeptly manage the inherent contradictions posed by these opposing pressures. The protection of data privacy and security is of utmost importance in the contemporary era of digital technology, particularly when it comes to safeguarding sensitive information about students, teachers, and staff members [29]. The occurrence of data breaches, illegal access, and visual data exploitation carries significant consequences, necessitating robust protective measures. To safeguard data, it is advisable to employ encryption techniques, implement access restrictions, and assign appropriate permissions [29].

Acquiring diverse and inclusive training datasets poses challenges in mitigating bias within computer vision systems [30]. The implementation of data collecting and use protocols about privacy and security has the potential to restrict access to specific demographic groups, hence perpetuating inherent prejudices. This disagreement underscores the necessity of achieving a delicate equilibrium between safeguarding data privacy and promoting diversity and representation. The concepts of fairness and bias hold equal significance in algorithms. Using biased training data harms computer vision systems [30]. It is essential for institutions to carefully choose and manage training datasets that encompass a wide range of diversity. Additionally, it is crucial for these institutions to consistently evaluate algorithms for any potential biases that may arise [30]. The utilization of this technique raises concerns regarding possible infringements of data privacy rules, as it entails the sharing of data derived from various sources. The process of reconciling these conflicting pressures requires a nuanced approach. To ensure the integrity and ethical use of data, institutions must develop a comprehensive strategy encompassing many vital components such as data anonymization, access restrictions, detection of biases, and implementation of appropriate mitigating measures. The integration of privacy and fairness is necessary to uphold the principles of data security and algorithmic fairness.

Ensuring accessibility and inclusivity of computer vision technology poses challenges due to limitations in infrastructure and financing [31]. Including text-to-speech and speech-to-text technologies is of utmost importance in ensuring the accessibility of educational materials for a diverse range of individuals, including students with disabilities [32]. These technologies facilitate the accessibility of images by converting them into textual or auditory formats. Inclusion necessitates the presence of user interfaces and programs that are easily accessible [32]. Financial barriers provide a substantial challenge. Budgetary limitations might impede educational institutions from procuring and sustaining equipment and software [31]. Smaller or underfunded institutions have a more significant burden of these expenditures.

To address these difficulties, it is imperative to adopt balanced measures. The successful execution of computer vision projects necessitates meticulous planning and the deployment of resources in a systematic manner [31]. The utilization of cloud services and the establishment of technical collaborations have the potential to mitigate infrastructure costs. To provide equitable access to technological resources for all children, it is imperative to implement measures that promote accessibility. In general, digital education must place a high priority on ensuring accessibility and inclusivity. By employing strategic planning and effectively allocating resources, it is possible to overcome limitations in infrastructure and financial resources to provide equitable access to technology for all students.

The presence of technical capabilities inside educational institutions is a matter of considerable importance [33]. Numerous institutions require enhanced proficiency in the deployment and maintenance of computer vision systems. The presence of a skill gap imposes constraints on the potential of technology. According to existing literature, it is imperative to make substantial expenditures in staff and educator professional development [33]. The necessity of providing computer vision training to educators cannot be overstated. There are advantages to consulting with specialists in AI and computer vision, as well as technology businesses [33]. It is essential to acknowledge a disparity in technical proficiency among instructors. Educators needing more technical ability require user-friendly interfaces and tools to utilize computer vision systems effectively.

To address the educational skill gap, it is imperative to provide training, support, and accessible technological tools. To optimize the utilization of computer vision technology, institutions can enhance their outcomes by allocating resources toward the recruitment and professional development of personnel and instructors.

Usage of Computer Vision in the Education Sector. Computer vision, an area of study within artificial intelligence and computer science, has experienced a surge in popularity due to its extensive range of applications across several sectors. Computer vision is exerting a substantial influence in the field of education. The utilization of computer vision technology is significantly altering the educational landscape, impacting the learning experiences of children, instructional methodologies employed by instructors, and the operational procedures of educational institutions.

1) Customized Learning. Personalized learning is an instructional approach that is designed to accommodate the unique requirements, abilities, and preferences of individual students. The proposed solution employs computer vision techniques to study the activities and priorities of students in a personalized manner. The application of computer vision in customized learning involves using facial recognition technology to assess and quantify student involvement [34]. AI-powered systems can discern the level of engagement, disinterest, or confusion exhibited by students during a lecture by analyzing their facial expressions and body language. To optimize the learning experience for each student, educators can modify their instructional approaches by providing increased assistance or adjusting the pace of the curriculum. In the field of computer vision, it has been shown that the tracking of students' eye movements may be utilized as a means to ascertain the specific sections of a digital textbook or instructional film that capture their attention [34]. This data has the potential to assist content creators and educators in improving educational materials by identifying the sections that are most captivating and in need of improvement.
2) Automating Administrative Work. Numerous administrative responsibilities within the realm of education need a significant investment of time and resources. Computer vision technology has been implemented in several management operations, resulting in a reduction of worker burden and an improvement in overall efficiency (35). Educational institutions are responsible for administering several types of documentation, ranging from registration forms to academic transcripts. Computer vision to scan, digitize, and categorize these papers facilitates information management and retrieval. Implementing this technology has the advantage of time efficiency and decreasing human errors. The application of computer vision extends to the realm of school and university visitor control systems [35]. The present method can accurately identify and authenticate visitors, guaranteeing that only individuals who have been granted authorization are permitted entry. Additionally, it can monitor and record visitor check-ins and check-outs, enhancing the overall security measures on campus.
3) Improving Class Experience. Integrating computer vision technology can potentially enhance interactivity and engagement within educational settings. Smart whiteboards with computer vision technology can convert handwritten text and drawings into digital format. Educators can efficiently store and dis-

seminate notes, fostering student cooperation and active engagement. Educational applications of augmented reality (AR) and virtual reality (VR) leverage computer vision techniques to construct immersive learning environments [36]. Within the educational setting, students can engage in various academic activities, such as visiting significant historical sites, doing virtual dissections of different species, and exploring the intricate complexities of human anatomy. Hands-on and interactive techniques are more effective in comprehending and retaining complex information.

4) Student Feedback and Assessment. The utilization of computer vision plays a crucial role in automating evaluation processes and providing prompt feedback to students. Using computer vision enables the speedy and accurate scanning and grading of multiple-choice answer sheets with the implementation of OMR technology [37]. This practice minimizes the amount of time instructors spend on grading and provides students with prompt feedback. In addition, computer vision can assess student writing. Automated essay grading techniques assess the grammatical accuracy, coherence, and lexical proficiency of written compositions. Although these systems are unable to replace human evaluation fully, they do offer initial ratings and comments, which can save instructors time and ensure grading uniformity.
5) Campus Safeguard. Educational institutions give security priority. Implementing computer vision-based security systems has been shown to enhance the level of safety on campuses. Surveillance cameras equipped with computer vision algorithms can promptly identify and discern potentially suspicious behavior or individuals, hence promptly alerting security personnel [38]. Implementing a proactive approach may effectively mitigate security breaches and safeguard the well-being of both students and staff. Moreover, computer vision can effectively monitor attendance. Facial recognition technology can promptly identify students upon entering classes or campus buildings, therefore obviating the need for human attendance-taking procedures. Online learning platforms offer the advantage of time efficiency and enhanced accuracy in monitoring student attendance for educational institutions.
6) Special Education and Inclusion. Computer vision technology has the potential to provide crucial educational support to children with diverse requirements. Computer vision technology can convert printed text into audio format, enhancing accessibility to educational resources for visually impaired students [39]. Computer vision has the potential to assist children with learning disabilities through the provision of individualized suggestions derived from educational materials. If a student encounters difficulties with a particular subject matter, the educational system can offer additional resources or modify the level of complexity of the assigned tasks.
7) Use in Research Work and Analysis. Computer vision plays a pivotal role in higher education and research institutions by analyzing extensive datasets, executing experiments, and propelling advancements in the field of research. Within the field of biology, computer vision algorithms can examine microscope images to identify cellular structures and detect any anomalies that may

be present. Social science researchers can employ computer vision techniques to evaluate vast quantities of visual data, such as images and videos from social media platforms [40]. Examining public sentiment, cultural patterns, and human conduct by researchers may enhance the comprehensive understanding of society.

Limitation for the Study. To have a comprehensive understanding of the breadth and limitations of the study, it is imperative to assess its restrictions critically. The following are the fundamental restrictions that are identified:

1) Generalizability. The absence of generalizability is a notable aspect that requires enhancement in this work. This research investigates the functionality and effectiveness of emotion detection systems within online learning environments. Nevertheless, it is essential to note that the conclusions drawn from this study may need more generalizability and may only apply to specific educational contexts. The variability of outcomes across various online platforms, courses, and student groups poses challenges in drawing generalized conclusions.
2) Data Privacy and Ethics. Utilizing emotion detection algorithms in educational settings raises significant ethical and privacy concerns. This study acknowledges the difficulties mentioned earlier, but it may not comprehensively tackle issues about data privacy, permission, and the potential exploitation of emotional data. A comprehensive ethical evaluation necessitates doing independent research.
3) Technological Viability. The research examines the advantages of emotion detection systems. However, it may not comprehensively explore their technological feasibility in many educational settings. The potential barriers to widespread adoption of technical infrastructure include factors such as its availability, pricing, and accessibility for both students and educational institutions.
4) Limited Time Frame. The inclusion criteria of the study necessitate that publications fall between the time frame of 2017 to the present. The specified time frame may result in the exclusion of works published before 2017, hence overlooking significant advancements and valuable insights in school emotion identification systems.
5) Language Bias. The inclusion criteria of this study specifically exclude research conducted in languages other than English. The imposition of this limitation has the potential to result in the exclusion of significant studies and perspectives originating from non-English-speaking regions, influencing the overall research landscape.
6) Subject Specificity. This study investigates the efficacy of emotion recognition systems within online learning, with potential limitations in accounting for subject-specific variations. Different disciplines may possess distinct requirements and challenges when it comes to the development and utilization of these methodologies.
7) Limited Emotions. This study evaluates the emotions of surprise, interest, perplexity, and wonder within online learning. However, it only provides a comprehensive examination of some feelings students and instructors may encoun-

ter during online learning. A more extensive study would enhance the understanding of the relationship between emotions and learning outcomes.

4 Conclusion

The integration of computer vision-based emotion recognition technologies in online learning environments has significantly transformed the field of education. This study investigated the performance of these systems in several aspects, including feature extraction, emotion interpretation, individual variations, camera quality, and ethical considerations. The elements mentioned earlier have shed light on the transformative impact of emotion detection algorithms on online education. The investigation of emotion detection systems within the context of online learning environments was conducted through the utilization of research questions. The present study investigated several approaches to the extraction and analysis of emotional data, machine learning algorithms, adaptability to individual learners, camera quality, and ethical considerations. The objective of this study was to elucidate the advantages and disadvantages of these systems. This study demonstrates that the implementation of emotion detection systems has the potential to enhance the effectiveness of online learning. These technologies provide the quantification and analysis of students' emotional responses, aiding educators in improving their teaching strategies and instructional materials. By utilizing real-time feedback on students' emotions, educators have the ability to adapt their instructional strategies in order to enhance student engagement and improve academic achievement.

Integrating computer vision and emotion detection technologies can enhance the efficacy of student-teacher communication within the context of online learning. In a conventional classroom setting, educators can evaluate student responses and subsequently modify their instructional approaches. Emotion detection technology in virtual classrooms enables educators to provide enhanced tailored assistance and foster stronger student connections. While emotion detection technology in education offers advantages, it also presents various challenges. The ethical and responsible management of students' emotional data is crucial. Ensuring the security of sensitive information and maintaining transparency regarding its utilization are of utmost importance. The issue of bias and fairness in these systems requires further attention and consideration. Biased algorithms can exacerbate student inequality by discriminatory treatment of specific demographic groups. The utilization of a diverse and inclusive training dataset, as well as conducting frequent algorithm reviews, serves to mitigate prejudice. In addition, researchers need to consider factors such as accessibility and inclusivity. It is imperative that educational technology, such as emotion recognition systems, be made accessible to all students, including those with disabilities. To mitigate obstacles in the learning process, it is imperative for technology to offer interfaces that are easily navigable by users and incorporate diverse modalities. Implementing such systems might incur significant costs and require substantial infrastructure, particularly for schools with limited resources. To rationalize and maintain these investments, it is imperative to do a comprehensive cost-benefit analysis. Emotion detection systems driven by computer vision have the potential to enhance online education by augmenting learning environments, facilitating teacher decision-making, and fostering student engagement. It is imperative to employ these technologies responsibly, considering principles such as data privacy, fairness, accessibility, and cost-effectiveness. In the digital era, the advancement of education to cater to the needs of diverse learners necessitates the utilization of emotion detection technologies, which will serve as a pivotal factor in developing online learning environments that are both more efficient and inclusive.

References

1. O. P. Pinchuk and O. Y. Burov, “Digital transformation of learning environment: aspect of cognitive activity of students Institute of Information Technologies and Learning Tools of NAES of Ukraine,” in CTE 2018 – How cloud technologies continues to transform education, 2019. [Online]. Available: https://vchasno.com.ua
2. M.-H. Yang and N. Ahuja, FACE DETECTION AND GESTURE RECOGNITION FOR HUMAN-COMPUTER INTERACTION, vol. 1. 2001.
3. P. Panchal, R. Bhojani, and T. Panchal, “An Algorithm for Retinal Feature Extraction Using Hybrid Approach,” in Procedia Computer Science, Elsevier B.V., 2016, pp. 61–68. doi: 10.1016/j.procs.2016.03.009.
4. M. Imani and G. A. Montazer, “A survey of emotion recognition methods with emphasis on E-Learning environments,” Journal of Network and Computer Applications, vol. 147. Academic Press, Dec. 01, 2019. doi: 10.1016/j.jnca.2019.102423.
5. M. Hakimi-Asiabar, S. H. Ghodsypour, and R. Kerachian, “Multi-objective genetic local search algorithm using Kohonen’s neural map,” Comput Ind Eng, vol. 56, no. 4, pp. 1566–1576, May 2009, doi: 10.1016/j.cie.2008.10.010.
6. M. A. Hasan, N. F. M. Noor, S. S. B. A. Rahman, and M. M. Rahman, “The transition from intelligent to affective tutoring system: A review and open issues,” IEEE Access, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 204612–204638, 2020. doi: 10.1109/ACCESS.2020.3036990.
7. S. Hutt, K. Krasich, J. R. Brockmole, and S. K. D. Mello, “Breaking out of the lab: Mitigating mind wandering with gaze-based atention-aware technology in classrooms,” in Conference on Human Factors in Computing Systems - Proceedings, Association for Computing Machinery, May 2021. doi: 10.1145/3411764.3445269.
8. B. A. Kitchenham and S. Charters, “Guidelines for performing Systematic Literature Reviews in Software Engineering,” 2007.
9. B. Carrion, C. A. Gonzalez-Delgado, A. Mendez-Reguera, I. E. Erana-Rojas, and M. Lopez, “Embracing virtuality: User acceptance of virtual settings for learning,” Computers and Electrical Engineering, vol. 93, Jul. 2021, doi: 10.1016/j.compeleceng.2021.107283.
10. B. Thoma, A. Turnquist, F. Zaver, A. K. Hall, and T. M. Chan, “Communication, learning and assessment: Exploring the dimensions of the digital learning environment,” Med Teach, vol. 41, no. 4, pp. 385–390, Apr. 2019, doi: 10.1080/0142159X.2019.1567911.
11. M. Fteiha and N. Awwad, “Emotional intelligence and its relationship with stress coping style,” Health Psychology Open, vol. 7, no. 2. SAGE Publications Inc., 2020. doi: 10.1177/2055102920970416.
12. Z. Trabelsi, F. Alnajjar, M. M. A. Parambil, M. Gochoo, and L. Ali, “Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student’s Behavior

Recognition," Big Data and Cognitive Computing, vol. 7, no. 1, Mar. 2023, doi: 10.3390/bdcc7010048.
13. M. Cukurova, M. Giannakos, and R. Martinez-Maldonado, "The promise and challenges of multimodal learning analytics," British Journal of Educational Technology, vol. 51, no. 5. Blackwell Publishing Ltd, pp. 1441–1449, Sep. 01, 2020. doi: 10.1111/bjet.13015.
14. M. L. Barrón Estrada, R. Zatarain Cabada, R. Oramas Bustillos, and M. Graff, "Opinion mining and emotion recognition applied to learning environments," Expert Syst Appl, vol. 150, Jul. 2020, doi: 10.1016/j.eswa.2020.113265.
15. P. Sharma et al., "Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning," in International Conference on Technology and Innovation in Learning, Teaching and Education, 2022, pp. 52–68.
16. C. M. Chen, J. Y. Wang, and Y. C. Lin, "A visual interactive reading system based on eye tracking technology to improve digital reading performance," Electronic Library, vol. 37, no. 4, pp. 680–702, Sep. 2019, doi: 10.1108/EL-03-2019-0059.
17. L. B. Krithika and G. G. Lakshmi Priya, "Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric," in Procedia Computer Science, Elsevier B.V., 2016, pp. 767–776. doi: 10.1016/j.procs.2016.05.264.
18. P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.
19. P. Hu, H. Ning, T. Qiu, Y. Xu, X. Luo, and A. K. Sangaiah, "A unified face identification and resolution scheme using cloud computing in Internet of Things," Future Generation Computer Systems, vol. 81, pp. 582–592, Apr. 2018, doi: 10.1016/j.future.2017.03.030.
20. M. F. Peterson and M. P. Eckstein, "Looking just below the eyes is optimal across face recognition tasks," Proc Natl Acad Sci USA, vol. 109, no. 48, Nov. 2012, doi: 10.1073/pnas.1214269109.
21. L. Stark, "Facial recognition, emotion and race in animated social media," First Monday, vol. 23, no. 9, 2018.
22. C. Ma and P. Yang, "Research on Classroom Teaching Behavior Analysis and Evaluation System Based on Deep Learning Face Recognition Technology," in Journal of Physics: Conference Series, IOP Publishing Ltd, Aug. 2021. doi: 10.1088/1742-6596/1992/3/032040.
23. X. Miao, Z. Yu, and M. Liu, "Using Partial Differential Equation Face Recognition Model to Evaluate Students' Attention in a College Chinese Classroom," Advances in Mathematical Physics, vol. 2021, 2021, doi: 10.1155/2021/3950445.
24. R. French, W. Imms, and M. Mahat, "Case studies on the transition from traditional classrooms to innovative learning environments: Emerging strategies for success," Improving Schools, vol. 23, no. 2, pp. 175–189, Jul. 2020, doi: 10.1177/1365480219894408.
25. C. Müller and T. Mildenberger, "Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education," Educational Research Review, vol. 34. Elsevier Ltd, Nov. 01, 2021. doi: 10.1016/j.edurev.2021.100394.
26. W. Wang, K. Xu, H. Niu, and X. Miao, "Emotion Recognition of Students Based on Facial Expressions in Online Education Based on the Perspective of Computer Simulation," Complexity, vol. 2020, 2020, doi: 10.1155/2020/4065207.
27. A. F. Wise and Y. Jung, "Teaching with analytics: Towards a situated model of instructional decision-making," Journal of Learning Analytics, vol. 6, no. 2, pp. 53–69, Aug. 2019, doi: 10.18608/jla.2019.62.4.

28. E. B. Mandinach and K. Schildkamp, “Misconceptions about data-based decision making in education: An exploration of the literature,” Studies in Educational Evaluation, vol. 69, Jun. 2021, doi: 10.1016/j.stueduc.2020.100842.
29. M. Cote and A. B. Albu, “Teaching Computer Vision and Its Societal Effects: A Look at Privacy and Security Issues from the Students’ Perspective,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, IEEE Computer Society, Aug. 2017, pp. 1378–1386. doi: 10.1109/CVPRW.2017.180.
30. S. Caton and C. Haas, “Fairness in Machine Learning: A Survey,” ACM Comput Surv, 2023, doi: 10.1145/10.1145/3616865.
31. S. Gul et al., “A Survey on Role of Internet of Things in Education,” 2017. [Online]. Available: http://www.gartner.com/newsroom/id/3165317
32. A. Stefk, W. Allee, R. E. Ladner, and S. Mealin, “Computer science principles for teachers of blind and visually impaired students,” in SIGCSE 2019 - Proceedings of the 50th ACM Technical Symposium on Computer Science Education, Association for Computing Machinery, Inc, Feb. 2019, pp. 766–772. doi: 10.1145/3287324.3287453.
33. Ž. Bojović, P. D. Bojović, D. Vujošević, and J. Šuh, “Education in times of crisis: Rapid transition to distance learning,” Computer Applications in Engineering Education, vol. 28, no. 6, pp. 1467–1489, Nov. 2020, doi: 10.1002/cae.22318.
34. C. Pabba and P. Kumar, “An intelligent system for monitoring students’ engagement in large classroom teaching through facial expression recognition,” Expert Syst, vol. 39, no. 1, Jan. 2022, doi: 10.1111/exsy.12839.
35. A. Hemalatha, P. B. Kumari, N. Nawaz, and V. Gajenderan, “Impact of Artificial Intelligence on Recruitment and Selection of Information Technology Companies,” in Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 60–66. doi: 10.1109/ICAIS50930.2021.9396036.
36. S. Gargrish, A. Mantri, and D. P. Kaur, “Augmented reality-based learning environment to enhance teaching-learning experience in geometry education,” in Procedia Computer Science, Elsevier B.V., 2020, pp. 1039–1046. doi: 10.1016/j.procs.2020.05.152.
37. H. Habib and A. Jabeen, “Usability of Exam Sheet Grading Application,” in International Conference on Biological Research and Applied Science: Emerging Trends in Computing, 2020. [Online]. Available: https://bebyaz.com/ExamReader].
38. P. Baile, N. Sutar, S. Shinde, A. Brahmankar, S. Angadi, and P. Dhore, “A Survey on Suspicious Activity Detection in Examination Hall,” J Pharm Negat Results, vol. 13, no. 7, 2022, doi: 10.47750/pnr.2022.13.S07.912.
39. I. Eligi, “ICT accessibility and usability to support learning of visually-impaired students in Tanzania,” 2017.
40. Y. Chen, K. Sherren, M. Smit, and K. Y. Lee, “Using social media images as data in social science research,” New Media and Society, vol. 25, no. 4. SAGE Publications Ltd, pp. 849–871, Apr. 01, 2023. doi: 10.1177/14614448211038761.