Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 18 No. 1 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijim.v18i01.46483

PAPER

Application of Computer Vision and Mobile Systems in Education: A Systematic Review

Rubaiya Tasnim Sholi, Md. Fouad Hossain Sarker(✉), Md. Salman Sohel, Md. Kabirul Islam, Maruf Ahmed Tamal, Touhid Bhuiyan, S.M. Khasrul Alam Shakil, Md. Foysal Ahmed

Daffodil International University, Dhaka, Bangladesh

fouadsarker@daffodilvarsity.edu.bd

ABSTRACT

The computer vision industry has experienced a significant surge in growth, resulting in numerous promising breakthroughs in computer intelligence. The present review paper outlines the advantages and potential future implications of utilizing this technology in education. A total of 84 research publications have been thoroughly scrutinized and analyzed. The study revealed that computer vision technology integrated with a mobile application is exceptionally useful in monitoring students' perceptions and mitigating academic dishonesty. Additionally, it facilitates the digitization of handwritten scripts for plagiarism detection and automates attendance tracking to optimize valuable classroom time. Furthermore, several potential applications of computer vision technology for educational institutions have been proposed to enhance students' learning processes in various faculties, such as engineering, medical science, and others. Moreover, the technology can also aid in creating a safer campus environment by automatically detecting abnormal activities such as ragging, bullying, and harassment.

KEYWORDS

computer vision (CV), mobile system, computer vision and mobile application, education, systematic review

1 INTRODUCTION

Education is vital for shaping society's structure, impacting employment, health, trade, income, family dynamics, and economic and political status. Enhancing education can lead to overall progress in a country [1]. Computer vision (CV) technology plays a significant role in national improvement by enabling humans to extract intelligence from their surroundings through computers. In the 21st century, CV is increasingly prevalent due to multimedia and image-based interactions on the web. It has evolved from a research notion to a widely used technology that dominates industries and enhances quality of life [2]. Despite the diverse

Sholi, R.T., Sarker, M.F.H., Sohel, M.S., Islam, M.K., Tamal, M.A., Bhuiyan, T., Alam Shakil, S.M.K., Ahmed, M.F. (2024). Application of Computer Vision and Mobile Systems in Education: A Systematic Review. *International Journal of Interactive Mobile Technologies (iJIM)*, 18(1), pp. 168–187. https://doi.org/10.3991/ijim.v18i01.46483

Article submitted 2023-09-04. Revision uploaded 2023-10-29. Final acceptance 2023-11-04.

applications and aspects of CV, it has diverse applications in education, including anomaly detection, attendance monitoring, educational robots, facial recognition, medical training, online proctoring, perception monitoring, and plagiarism checking. Anomaly detection, particularly in workplace sexual abuse studies, has shown significant progress using video sequence intrusion detection [3]. Identifying anomalies, such as violence, from surveillance camera streaming video is a common use of CV [4]. Convolutional neural networks (CNNs), like Keras (VGG16, DenseNet), and Support Vector Machine (SVM) models, are commonly employed for image processing and intrusion detection [5]. Another major CV application is monitoring student attendance using the face detection and recognition framework [6]. They have been widely applied in education to monitor student attendance. Researchers have utilized this technology to enhance educational quality through the development of attendance monitoring systems. These systems involve creating a database with student images linked to identification numbers, detecting students' faces in classroom videos, recognizing them from the database, and generating attendance reports. Numerous approaches have been employed to improve the accuracy of these systems. Many of the image processing algorithms included in OpenCV are now used in Advanced Robotic Systems (ARS) courses to provide support for teaching computer vision concepts using a mobile's camera and processing power to observe real-time effects generated on an image [7]. The integration of robots in education has sparked interest in recent years [8]. One notable aspect of educational robots is their use of computer vision (CV), particularly in object detection and recognition. Numerous studies have showcased the incorporation of CV technologies in educational settings, with popular methods including OpenCV and CNN. CV primarily focuses on image processing and facial recognition to identify individuals in different scenarios. Researchers in this field have made significant progress in analyzing facial landmark motions, head pose, face expressions, and eye gazing [9]. Advancements in CV are also reshaping traditional medical research. Medical trainees now familiarize themselves with medical imaging using machine learning (ML) algorithms. Simulation-based environments enable them to develop proficiency while ensuring patient safety [10]. With the rise of online learning, the demand for proctoring examinations has increased. Researchers have taken an interest in using CVs to monitor student activity during online tests, and several studies have emerged in this area [11–13] and [14–16]. Additionally, CV can assist in checking documents for plagiarism. As CV continues to advance, its applications in various disciplines, including education, are expanding. Vision technology is expected to become more prevalent in educational settings.

However, there is a big gap in the research that looks at how computer vision and mobile systems affect student learning outcomes, teaching methods, and accessibility when used together. This is because the use of these technologies in education has gotten a lot of attention and grown a lot in recent years. Although there are a plethora of studies highlighting the potential benefits of these technologies in isolation, there is a paucity of research that systematically investigates the synergistic effects and challenges that arise from their integration in educational settings. This research gap necessitates a thorough review that not only synthesizes the existing knowledge but also identifies areas where computer vision and mobile systems can be effectively harnessed in tandem to revolutionize the education landscape and, conversely, where potential pitfalls and limitations may impede their successful implementation. Such an analysis is crucial for educators, policymakers, and technology developers aiming to optimize the use of these tools for enhancing teaching and learning in diverse educational contexts.

2 LITERATURE REVIEW

2.1 Attendance monitoring

CV-based attendance monitoring is gaining popularity due to its affordability, efficiency, and accuracy. Singh et al. [17] developed a system that captures and stores student images for a CNN-based model. A recent study utilized the Haar Cascade algorithm for face detection and the LBPH model for recognition, with CNN and Microsoft Face APIs [18] also suggested as alternatives. Another study used the HAAR classifier for face detection and developed a server-based module, addressing the challenge of natural changes in faces by storing detected faces for model training [10]. The study [12] applied Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) using Fisherfaces and Eigenfaces projection, with PCA demonstrating faster execution time and LDAs achieving higher recognition rates. Mothwa et al. found that combining LDA and PCA yielded higher accuracy than other methods [13]. In a comparison study, the Mahalanobis distance-based system outperformed traditional Euclidean and Manhattan distances [19]. Another approach involved converting grayscale student images for attendance using LBPH and achieving high accuracy [20]. While many object detection systems rely on a CNN backbone network predefined for ImageNet, researchers [21] highlighted limitations in image classification and object detection using ImageNet alone. ImageNet's spatial resolution hampers the detection of large and small objects. To address this, the author of [22] proposed DetNet as a backbone for object detection, combined with FPN for feature extraction. Considering the importance of image quality, Bah, Serign [23] employed advanced image processing techniques such as light condition and contrast modification, bilateral filtering, linear image blend, and histogram equalizer to enhance image quality. An article shows a smart attendance monitoring system relying on facial recognition to track the attendance of students during classes, while another describes the same system using barcodes [24] and [25]. Table 1 summarizes CV and mobile system attendance monitoring.

Table 1. A summary of CV and mobile system attendance monitoring

Ref.	Functionality	Observed Features	Algorithms/ Methods/Tools	Results
[11]	Detects and records attendance and track facial changes over time	”	HAAR-Cascades	N/A
[13]	Implemented an attendance system and compared with others	”	LDA, PCA	83.57% (LDA), PCA (66.07%)
[18]	Track attendance, calculate present and absent percentages per subject/course, and provide results on a document/spreadsheet.	Backbone (DetNet)	CNN, HAAR-Cascades, LBPH, KNN, MFA	96% (LBPH), 97.35% (KNN), 99% (MFA)
[19]	Face recognition in real time with periodic attendance updates on register	”	LDA, LBP, PCA	90%
[20]	Detects and records attendance and track facial changes over time	”	HAAR-Cascades, Eigenface recognizer	80%
[22]	Identification of objects	”	CNN (DetNet, FPN)	N/A
[23]	Handle noisy image & detect faces for attendance monitoring	”	LBP+SVM+PS, Original, LBP, DCP+LBP+SVM	97.50% (DCP+LBP+SVM)
[24]	Face recognition for attendance monitoring	”	PCA, LDA, LBP	90%
[25]	Barcode system for attendance monitoring	”	AIDC, RFID	N/A

2.2 Educational robots

CVs in robots are used as educational tools and for assisting teachers in lessons. In a survey of 46 12-year-old students [26], one group was taught vocabulary using robotics-assisted language learning (RALL) by a teacher, which resulted in better performance compared to the traditionally taught group. The teaching activity was carried out by the social robot NAO [27], [28]. Jiménez et al. [29] conducted a study using NAO robots to teach children with Down syndrome to recognize colors through CV techniques. Esteban et al. [30] developed a smartphone-dependent educational robot by integrating CV algorithms into the smartphone's operating system (Android and iOS). The system utilized OpenCV code for ArUco marker detection, basic CV operations for lane detection, and a CNN model based on TensorFlow MobileNetv3 for object identification. Focusing a particular attention on the hands-on training sessions with Rovio mobile robots, another work outlines the most pertinent aspects of the module content and assessment methodology [31]. The authors of [32] proposed a CV-based system for teaching children the alphabet. Kusumota et al. [33] designed an educational robot that produces sounds, draws geometric shapes, displays mathematical questions on a screen, interacts with people by asking for answers, and expresses emotions through body movements and LCD screen emojis. For a summary of CV and mobile system educational robots, refer to Table 2.

Table 2. A summary of CV and mobile system educational robots

Ref.	Functionality	Observed Features	Algorithms/ Methods/Tools	Results
[27] [28]	To teach children with Down syndrome to recognize colors	microphone, sonar, speaker, pressure sensor, touch sensor	NAOqi	N/A
[30]	Real-time object identification, ArUcO marker detection, and lane detection	a smartphone-dependent educational robot	OpenCV, CNN	N/A
[31]	Hands-on training sessions, class lectures, and assessment	a mobile robot (Rovio)	OpenCV	N/A
[32]	For teaching children alphabet through robot	Robot	SVD	N/A
[33]	Works on real-time actions	Google Cloud Vision API, Real-time robot	CV	N/A

2.3 Medical training

Scientists are striving to enhance the applicability of computer vision (CV) in medical technology, as per research [34] that evaluated two common algorithms. CNN extracts feature vectors from images, while RCNN integrates CNN and region features for object detection. Selective search is employed to extract 2000 proposed regions, and each region undergoes CNN processing to calculate features. The study concludes that trained CNN and RCNN can replace many outdated classification methods in medical technology. Prior to actual surgeries, doctors acquire the necessary skills through CV-assisted Virtual Reality (VR). CV algorithms can analyze pre-recorded surgical videos to evaluate operation quality by assessing parameters such as speed, acceleration, needle rotation, and duration. Hisi, Rebecca et al. [35]

introduced a method for real-time detection and feedback during central venous catheterization training. Nugent et al. [36] developed a simulation-based laparoscopic training system called ProMIS, which combines haptic feedback and VR modules. Camera tracking systems capture equipment motion, instrument path length, acceleration, and movement uniformity from three different angles at a rate of 30 frames per second. Law, Hei, et al. [37] proposed a CV-based video analysis system for robotic surgery that reviewed information on the motion of the robotic device. They employed crowdsourcing on Amazon Mechanical Turk (AMT) to annotate robotic needles in the video. The annotated data was then used to train Stacked Hourglass Networks, and an SVM was trained using video recordings from 12 surgeries. Kil, Irfan, et al. [38] presented a method to assess surgeons' suturing skills by extracting relevant metrics from videos. With the help of a customized lens attached to the mobile's camera, a mobile app is now constantly updated with heart rate, blood pressure, and oxygen levels collected by sensors on the patient's wrist, and patients can use assistive technologies using mobile devices [39][40] (see Table 3).

Table 3. A summary of CV and mobile system medical training

Ref.	Functionality	Observed Features	Algorithms/ Methods/Tools	Results
[34]	Detects the workflow where a webcam	complex surgical video recordings, an electromagnetic tracking device	OpenCV	N/A
[35]	Investigation of medical images	classification and detection	CNN, RCNN	N/A
[36]	Collects equipment motion	a simulation-based laparoscopic training system	OpenCV	N/A
[37]	Track movement of two robotic needles from the video	Amazon Mechanical Truck (AMT)	ConvNet, SVM	91.67%
[38]	To assess suturing skill of surgeons	relevant metrics from video footage	OpenCV	N/A
[39]	Keep track of heart rate, blood pressure, and blood oxygen data	customized lens attached to a phone's camera	CNN, RNN, SVM	95%
[40]	Assistive techniques for patient's treatment	assist patient about their disorder through mobile application	CV, SLAM (self-localization and mapping)	N/A

2.4 Online proctoring

Surveillance of student activity during online activities is necessary due to the absence of physical invigilators. Sarthak et al. [41] proposed an automated system to monitor students' activity during online examinations. The system uses functions such as eye gaze tracking, facial movement, object detection, face detection, and audio-to-text conversion. It leverages the OpenCV Python package for computer vision, using the Dlib model for face detection and the YOLOV3 model for object detection. Similar work was done in another paper [42], which evaluated head pose, eye gaze, audio detection, and fraudulent activity prediction using the RealHePoNet CNN model. The study used the XGBoost and MLP classifiers, achieving 94% and 0.059% accuracy, respectively, in predicting cheating behavior. Another study suggested using a machine learning-based system [43] that was 96.04% accurate to help examiners spot cheating and malpractice in e-exams. An image-hashing-based proctoring system was proposed [44] to address privacy concerns in online proctoring.

The system protects student privacy by blurring faces or using masks. Real-time assessment during online testing is challenging, and a video summarization process can aid post-test review [45]. The proposed method involves detecting head pose, using the Hidden Markov Model (HMM) for behavioral modeling, and generating summaries for evaluation. The OpenCV algorithm is used for head tracking. Some works have described an online remote examination proctoring system for mobile devices that monitors the candidate's surroundings using a number of approaches, such as live video and audio streaming [46] and [47]. Table 4 provides a summary of CV and mobile system online proctoring.

Table 4. A summary of CV and mobile system medical training

Ref.	Functionality	Observed Features	Algorithms/Methods	Results
[42]	Automated online proctor	eye gaze tracking, mouth movement, object identification, face recognition, audio to text conversion	OpenCV, YOLOV3, Dlib	99.91% (person detection), 97.08% (object detection)
[43]	Automated online exam proctor	head pose, estimating eye gaze, detecting audio from a video, and predicting students' cheating behavior	CNN (RealHePoNet, OpenCV, Dlib)	94% (cheating behavior)
[44]	Privacy-preserving online proctor	image hashing, detecting face and body movement	MediaPipe, Dlib	N/A
[45]	Post-review video summarizer	head tracking, pose estimation	OpenCV (Hidden Markov Model)	79.8% (pose acc.), 79.3% (abnormal behavior recognition)
[46]	Remote examination proctoring system	real live audio, video streaming	OpenCV	N/A
[47]	Automated proctoring system	eye gazing, face detection, facial landmarks, head pose estimation, object and open mouth detection	OpenCV, Speech to text	N/A

2.5 Perception monitoring

The attention level of students in class has a significant impact on their academic success, as positive learning outcomes are strongly associated with student engagement [48]. Ngoc Anh et al. [49] proposed a system to monitor students' behavior in the classroom using CV methods like face identification, face embedding, gaze estimation, facial landmark identification, and face classification. However, their study did not track facial expressions, emotions, or head/body posing estimation, which are crucial factors in monitoring human behavior. Canedo et al. [50] addressed these limitations by introducing an agent to monitor students' attention, while Van der Har and Dustin [51] focused on recognizing students' emotions for the same purpose. The dynamic switch of a smartphone and gesture recognition allowed students to quickly engage with virtual content [52]. In another study [53], classroom involvement was evaluated using a semi-automated process that combined CV techniques with human observations. This model successfully recognized students' attention levels based on eye gestures but faced challenges in determining attention based on emotions. Savov, Teodor et al. [54] investigated how integrating IoT technology (e.g., PIR sensors, noise sensors) with CV enhances the accuracy of facial recognition for monitoring students' moods and assessing attention levels. In addition to students' perceptions, considering the teacher's point of view is crucial, although it is often overlooked by researchers. The authors [55] proposed a model that evaluates teacher perception in the classroom by integrating gaze data from a mobile eye tracker with egocentric vision. By using an eye tracking glass worn by the teacher, all students' faces can be recognized, and the model can also detect nonverbal activities that keep students engaged, such as moving around, gesturing, and making eye contact [56]. This model aims to automatically detect such teacher activities and establish their relationship with students' self-reported attention. Table 5 provides a summary of CV and mobile system perception monitoring.

Table 5. A summary of CV and mobile system perception monitoring

Ref.	Functionality	Observed Features	Algorithms/ Methods/Tools	Results
[49]	Student's classroom behavior monitoring	face identification, embedding, gaze estimation, landmark identification & classification	SSH, O-Net, L-Net, MTCNN, Hopenet, SVM, RF, DT, GB	N/A
[50]	Monitoring perception by detecting facial expression	facial expression, emotions, head or body posing estimation	MTCNN, OpenCV, Dilb, OpenPose, RestNet-101	N/A
[53]	Classroom involvement identification	facial expressions, note-taking devices, audio devices that allow students to connect with the teacher, eye posture towards the teacher or board, body motions, etc.	Not mentioned	N/A
[54]	Integration of IoT & CV to enhance perception	noise, fidgeting level, images of each student	OpenCV, LBPH, Fisher algorithm	N/A

2.6 Plagiarism checking of the handwritten document

Plagiarism, the act of stealing someone's work without proper credit, has become widespread in recent years. In a paper [57], a technique for extracting texture features and an enhanced SVM method are presented, utilizing four feature vectors to express features like Energy, Entropy, Moment of inertia, and Correlation. The classification method used is DAG_SVM (Directed Acyclic Graph SVM), with Minimal Hypersphere and achieves the highest accuracy at 93.32%. In a similar study by Yunyan Wang et al. [58], transfer learning-based convolutional neural networks were employed to classify images. The approach involved extracting features using the Histogram of Oriented Gradient (HOG) method from training data, which were then fed into SVM for pre-classification. By utilizing a transfer network in CNN, the results were compared to typical classifier algorithms, achieving an accuracy of 95%. Authors of article [59] aimed to determine the similarity of text images by calculating a similarity score based on text reuse patterns, slight differences in word morphology, word order, and content interpretation. The suggested CNN feature technique outperformed conventional hand-held design feature extraction when comparing similar images. However, this technique has a limitation as it only recognizes similarities in written texts, potentially overlooking graphs or diagrams present in documents. In a study [60], both OCR (Optical Character Recognition) and CNN were employed to convert images to text. OCR was used to convert digital images into machine-encoded text, and CNN provided additional refinement to the OCR output. It was observed that OCR alone produced satisfactory results compared to CNN. Oleg et al. [61] conducted a challenging task of checking for plagiarism in handwritten texts or documents. They demonstrated a process of recognizing such papers and determining if they are comparable to another document, known as a near-duplicate document. Dynamic time warping (DTW) was utilized as a function to compare similar documents. One drawback of the studies [60],[61] can be addressed by employing the strategy presented in the study [62]. Jithin S Kuruvila et al. [54] showcased a strategy to identify similarly shaped flowcharts in two documents. The operation consists of four modules: the first module applies the Canny edge detection (CED) algorithm to transform the image into binary form. The second module utilizes area detection algorithms to recognize various shapes present in the flowchart. A directed graph is formed to determine the orientation of the generated shapes. The final module compares the resulting graph with previously generated graphs in the repository to determine flowchart similarity. Moreover, detecting plagiarism in sentences using GLSA (Generalized Latent Semantic Analysis) has become noticeable in mobile learning [63]. Table 6 provides a summary of CV and mobile system plagiarism checking of a handwritten document.

Table 6. A summary of CV and mobile system plagiarism checking of the handwritten document

Ref.	Functionality	Observed Features	Algorithms/ Methods/Tools	Results
[57]	Classify educational images	texture feature extraction	DAG_SVM, CNN (Transfer network)	93.32%, 95%
[59]	Identifying near-similar document Image-to-text conversion	word segmentation	DTW (Dynamic time warping), FastDTW	87–96%
[60]	Image-to-text conversion	digital image into machine-encoded text	OCR, CNN	N/A
[61]	The similarity between images of handwritten documents	text reuse pattern, word morphology, word order, etc.	CNN	N/A
[62]	Identifying similar flowchart	pre-processing real flowchart, detecting shapes, constructing graphs	CED (Canny edge detection)	N/A
[63]	Detecting plagiarism on mobile learning	detect sentences with syntactic error or common words	GLSA (Generalized Latent Semantic Analysis)	N/A

2.7 Other potential applications

Identifying heritage sites, detecting cracks in concrete [64], and recognizing scholars and well-known individuals from images are significant tasks in image classification. In a research study [65], the authors employed transfer learning to incorporate a preexisting CNN model for heritage data classification. Fabric defect identification can also be addressed using computer vision (CV). P.R. Jayaraj et al. [66] proposed a multi-scaling deep CNN approach for fabric defect classification, examining six different defect classes. Sexual abuse is a serious violation of internal and external rights that is increasing in society. A study [67] suggested a CNN-based model with 95% accuracy to detect workplace harassment in social media videos. Additionally, the Social Force Model [68] has been introduced to detect abnormal behavior. For security purposes, a CNN-based helmet detection model linked to an Automated Teller Machine (ATM) can detect anomalies with a 95.3% accuracy, as demonstrated in [69]. The LRCN, a deep neural network model, was developed to detect five intrusion behaviors [70]. An updated version of the Gaussian Mixture Model has also been developed for object detection and trackingdynamic video sequences to identify intrusion [71]. Moreover, an IDS (Intrusion Detection System) implemented in [72] incorporates an object detection method (YOLO) and Real-time Tracking algorithm (SORT) with a 97% accuracy and an average frame rate of 30.

Another study [73] proposed an efficient framework for violence identification from surveillance camera streaming video using Linear SVM, Cubic SVM, Random Forest (RF), and Violent Flows (ViF) algorithms. OpenFace is an open-source real-time platform for facial behavior analysis, including landmark detection, head pose tracking, eye gaze, and facial action unit estimation. In OpenFace, CNN is trained using randomly distributed landmark positions [74]. OpenFace 2.0, which utilizes the SVR-HOG technique, was introduced to increase accuracy from 76.1% to 92.9% [75]. Researchers continuously strive to intelligently analyze and detect facial motions. A study [76] uses an ASM (Active Shape Model) and SVM classifier to identify front-view human faces in real-time with 93% accuracy. Facial Expression Recognition (FER) classifies facial features linked to six emotions using Adaboost and Haar Cascade classifiers [77]. Some FER investigations used LBP, LDP, and KNN classifications to reach 96.83% recognition [78]. STLMBP, SVM, and 3NN were used to analyze face expressions in [79]. A hybrid deep convolutional recurrent neural network can identify facial expressions based on landmark placements better than snipping [80]. Learning management systems (LMS) use CV traits for testing and performance assessment because CV is crucial in education [81]. Based on student time and features, some algorithms may detect boredom, bewilderment, enthusiasm, aggravation, and attentive involvement with 98% accuracy [82]. The Hidden Markov Model (HMM) and Independent Bayesian Classifier Combination (IBCC) recognized hand and finger movements with 96% accuracy for single-hand gestures and 94% accuracy for double-hand gestures [83]. Table 7 provides a summary of other potential applications for CV.

Table 7. A summary of CV-based other potential applications

Ref.	Functionality	Observed Features	Algorithms/ Methods/Tools	Results
[65]	Identifying heritage sites	transfer learning	CNN (VGG16)	95%
[66]	Fabric defect identification	six different defect classes	Multi-scaled CNN	96.55%
[70]	Detect intrusion	five types of intrusion behaviors	LRCN	N/A
[72]	Intrusion Detection System	intrusion from live video stream	YOLO, SORT	97%
[73]	Identifying violence	violence from surveillance camera streaming video	Linear SVM, Cubic SVM, Random Forest (RF), ViF	N/A
[75]	Facial behavior analysis	facial landmark detection, head pose tracking, eye gaze, and facial action unit estimation	CNN, SVR-HOG	92.9%
[76]	Face identification	real-time identification based on the front view human facial image	SVM	93%
[77]	Face identification	pleasure, sorrow, disgust, anxiety, anger, and surprise	Adaboost, Haar Cascade	N/A
[83]	Recognizing hand and finger gestures with facial expression	single hand gesture, double hand gesture	HMM, IBCC	94% (DHG), 96% (SHG)

3 DESIGN AND METHODS

The main purpose of this article is to identify the most current CV research trends in the education sector to fulfill the study objectives, which are listed below on the research questions. So, we collected published research papers on this topic from 2010–2021 and categorized them. Literature review sections find answers to those questions by analyzing some of the most important research papers. At the beginning of all our work, we studied numerous concepts of systematic review before proceeding with the popular guidelines provided by Okolio, Chitu et al. [84], where each stage is explained in depth and has also acquired popularity among researchers in recent years. Here, Figure 1 presents the workflow of our systematic review.

Fig. 1. Workflow of this systematic review

3.1 Research questions

- RQ1: What are the current research trends in CV and mobile systems in the educational sector?
- RQ2: How are these beneficial to the educational sector?
- RQ3: How can different aspects of these be applied in education?
- RQ4: How will these benefit from the perspective of future education?

3.2 Paper searching approach

Google Scholar was used to find research publications on this study's topic because of its high-quality relevance ranking. A search query was developed to find relevant articles, where the first 10 pages (100 research papers) of each search result is considered. See Box A for detailed search queries.

Box A. Search query

Search Query = ('*null*' **||** 'Attendance Monitoring' **||** 'Online Proctoring' **||** 'Image Classification' **||** 'Object Detection' || 'Face Identification' || 'Educational Robot' || 'Training' || 'Medical Training' || 'Plagiarism Checking' || 'Handwritten Document' || 'Perception Monitoring' || 'Anomaly Detection' || 'Behavior Monitoring' || 'Engagement' || 'Handwritten document') **&&** ('*null*' || 'Students' || 'Classroom' || 'Education' || 'Educational Institute' || 'Application') **&&** 'Computer Vision'

3.3 Screening of relevant paper

Many articles were found using our search criteria, but not all of these publications addressed our study's concerns and needs. So, from this list, we selected 318 papers related to CV applications in education and 30 articles on CV applications that have not yet been implemented in education but are likely to be used (see Table 8). These selection processes are carried out following identification, screening, eligibility tests, and inclusion-exclusion criteria. To determine current research trends, we graphically presented the year-wise distribution of those publications and classified all of those selected papers' relevant categories.

Table 8. Current research trends in computer vision and mobile system in education

Name of Category	Number of Studies	Percentages
Attendance Monitoring	100	31.44%
Perception Monitoring	75	23.58%
Handwriting Identification	64	20.12%
Online Proctoring	41	12.89%
Educational Robots	29	9.12%
Medical Training	9	2.83%

a) Inclusion criteria:

- CV-related research is aimed and has potential towards the educational sector.
- Articles published in the English Language.
- Research articles published between 2010 to 2021 (see Figure 2)

b) Exclusion criteria:

- CV research is not related to education and is unlikely to be used in the education sector.
- Short papers, review articles, less popular papers, unpublished research articles, Wikipedia content, and online blog.

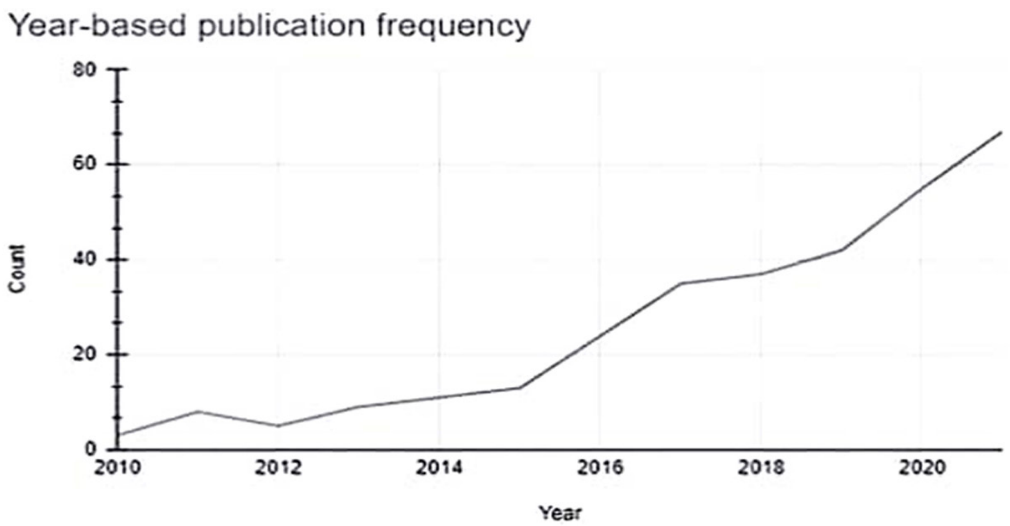

Fig. 2. Annual distribution of published articles on computer vision and mobile system in the education sector

3.4 Literature review execution

After reviewing all of the selected research papers, we independently selected some papers for discussion in the literature review section to answer some of the research questions raised.

4 RESULT AND DISCUSSION

4.1 Trends of CV and mobile system in the educational sector

Figure 3 illustrates the current research trajectory in the education sector regarding CVs and mobile systems. The domain of attendance monitoring is the most widely studied. Perception monitoring is the second step, followed by handwritten documents. The prevalence of online proctoring and pedagogical robots has decreased. Medical education, in conclusion, is the least researched field within education. The potential impact of integrating CVs and mobile systems into the education sector is a paradigm shift in the way students engage with and acquire knowledge from educational materials. An extensive selection of mobile applications pertaining to computer vision are presently available for free installation from the Google Play Store [7], enabling students to conveniently access them. A concentration on a variety of subjects, including attendance tracking, educational robots, medical training, online proctoring, perception monitoring, and OC, has been identified through our investigation of current research trends [8], [10], [11–14], [36–38], and [52]. However, it is crucial to acknowledge that the implementation of CV-assisted teaching may also present obstacles or ethical concerns that require attention. We have outlined the most recent developments in CV and mobile system research within the educational sector in the present paper. Almost certainly, our paper introduces these developments to comprehension by discussing how computer vision technology is presently being utilized or researched in the field of education.

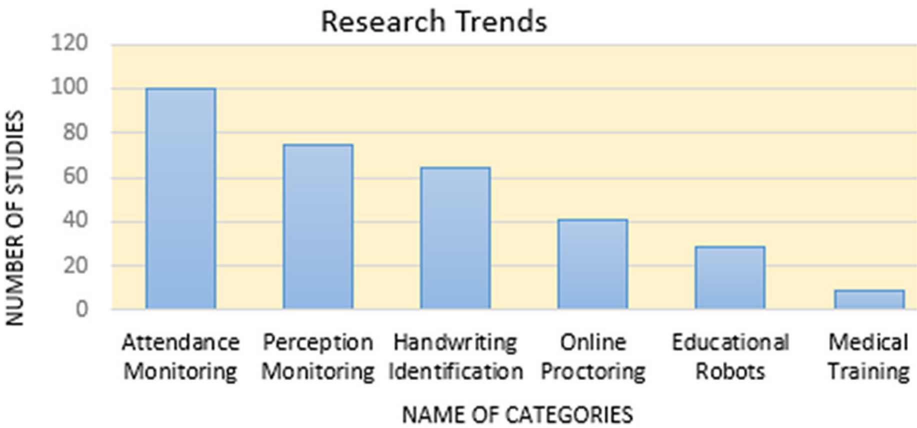

Fig. 3. Current research trends in computer vision and mobile systems in education

4.2 Benefits of CV with mobile system

To address our second research question, computer vision (CV) technology has the potential to improve the educational sector in a variety of ways. One of the primary advantages is the capacity to create tailored learning experiences for each student through the use of a small device known as a mobile device [43], [50]. This can help ensure that each student receives the necessary assistance and resources. Furthermore, CV technology can be used to make education more accessible to students with varied needs or abilities by providing images or other sorts of support. This can help to create a more inclusive and equitable learning environment.

Furthermore, our data imply that CV-assisted mobile teaching has the potential to significantly improve efficiency, accessibility, and student participation.

4.3 Applications of CV and mobile system in different avenues

We focused on the applications of CV and mobile systems in various avenues to answer our third research question, such as monitoring student attendance, engaging students in interactive learning experiences, enabling medical training, facilitating remote testing, grading exam papers, tracking student conduct, digitizing printed texts, object detection, face recognition, eye contact monitoring, emotion capturing, gesture recognition, picture categorization, and augmented reality. This study reveals that the use of CV and mobile technologies makes these applications possible [7], [9], [24–25], [30], [37], [39–40], and [49].

4.4 Future prospect

To answer the fourth research question, we found that integrating computer vision (CV) and mobile system in education can bring significant changes to how students learn and interact with educational material, potentially revolutionizing the sector. It is evident that CVs and mobile systems have the potential ability to enhance efficiency, accessibility, and student engagement that led to improved educational outcomes [51], [53], [78], [85]. It enables personalized learning experiences, immersive environments, improved accessibility for students with disabilities, automated assessments, and advanced research. Image classification and object identification algorithms can be applied across various fields such as medicine, ecology, civil engineering, and textile engineering. In the future, CV could detect abnormal activities like harassment and bullying, as well as identify strangers on educational premises.

5 CONCLUSION

The rapid growth of the computer vision industry and its integration with mobile applications have paved the way for transformative possibilities in education. Our systematic review of 84 research publications underscores the multifaceted advantages of utilizing computer vision technology, shedding light on its potential to revolutionize the educational landscape. By enabling the monitoring of students' perceptions, mitigating academic dishonesty, and automating administrative tasks like attendance tracking, computer vision and mobile technology offer substantial benefits to educators and institutions, optimizing the efficient use of classroom time. Moreover, its applications extend to diverse educational domains, from engineering to medical science, presenting opportunities to enrich the learning process across faculties.

Beyond academic enhancement, the technology's capacity to create safer campus environments by detecting and addressing abnormal activities such as ragging, bullying, and harassment is a testament to its broader societal impact. As we navigate the ever-evolving landscape of education, the fusion of computer vision and mobile systems opens up new horizons for both academic excellence and student well-being. These findings provide valuable insights for educational institutions, policymakers, and technology developers seeking to harness the full potential of this dynamic partnership in shaping the future of education.

6 REFERENCES

[1] Ozturk, Ilhan, "The role of education in economic development: A theoretical perspective," *SSRN,* 2008. https://doi.org/10.2139/ssrn.1137541

[2] Bebis, George, Dwight Egbert, and Mubarak Shah, "Review of computer vision education," in *IEEE Transactions on Education,* vol. 46, no. 1, pp. 2–21, 2003. https://doi.org/10.1109/TE.2002.808280

[3] Shahria, MD Tanzil, *et al.*, "Application of neural networks for detection of sexual harassment in workspace," in *2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, IEEE, 2021. https://doi.org/10.1109/ICAECT49130.2021.9392429

[4] Singh, Kuldeep, *et al.*, "Designing an efficient framework for violence detection in sensitive areas using computer vision and machine learning techniques," in *2018 Tenth International Conference on Advanced Computing (ICoAC)*, IEEE, 2018. https://doi.org/10.1109/ICoAC44903.2018.8939110

[5] Asep, Hadian SG, and Yoanes Bandung, "A design of continuous user verification for online exam proctoring on M-learning," in *2019 International Conference on Electrical Engineering and Informatics (ICEEI)*, IEEE, 2019. https://doi.org/10.1109/ICEEI47359.2019.8988786

[6] Meenpal, T., Aarti Goyal, and Ankita Meenpal, "Face recognition system based on principal components analysis and distance measures," *International Journal of Engineering & Technology,* vol. 7, no. 2, pp. 15–19, 2018. https://doi.org/10.14419/ijet.v7i2.21.11826

[7] Sigut, Jose, *et al.,* "OpenCV basics: A mobile application to support the teaching of computer vision concepts," *IEEE Transactions on Education*, vol. 63, no. 4, pp. 328–335, 2020. https://doi.org/10.1109/TE.2020.2993013

[8] Sophokleous, Aphrodite, *et al.*, "Computer vision meets educational robotics," *Electronics*, vol. 10, no. 6, p. 730, 2021. https://doi.org/10.3390/electronics10060730

[9] Indi, Chirag S, Varun Pritham, Vasundhara Acharya, and Krishna Prakasha, "Detection of malpractice in e-exams by head pose and gaze estimation," *International Journal of Emerging Technologies in Learning (iJET),* vol. 16, no. 8, pp. 47–60, 2021. https://doi.org/10.3991/ijet.v16i08.15995

[10] Gao, Junfeng, *et al.*, "Computer vision in healthcare applications," *Journal of Healthcare Engineering*, vol. 2018, 2018. https://doi.org/10.1155/2018/5157020

[11] Fuzail, Muhammad, *et al.*, "Face detection system for attendance of class' students," *International Journal of Multidisciplinary Sciences and Engineering,* vol. 5, no. 4, 2014.

[12] Raghuwanshi, Anshun, and Preeti D. Swami, "An automated classroom attendance system using video-based face recognition," in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, IEEE, 2017. https://doi.org/10.1109/RTEICT.2017.8256691

[13] Mothwa, Louis, Jules-Raymond Tapamo, and Temitope Mapati, "Conceptual model of the smart attendance monitoring system using computer vision," in *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, IEEE, 2018. https://doi.org/10.1109/SITIS.2018.00042

[14] Barba-Guaman, Luis Rodrigo, and Priscila Valdiviezo-Diaz, "The attention of students in online education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 17, no. 22, pp. 144–55, 2022. https://doi.org/10.3991/ijet.v17i22.34241

[15] Kadyrov, Bakhitzhan, Shirali Kadyrov, and Alfira Makhmutova, "Automated reading detection in an online exam," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 17, no. 22, pp. 4–19, 2022. https://doi.org/10.3991/ijet.v17i22.33277

[16] Chang, Xue-qin, Dao-hua Zhang, and Xin-xin Jin, "Application of virtual reality technology in distance learning," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 11, no. 11, pp. 76–79, 2016. https://doi.org/10.3991/ijet.v11i11.6257

[17] Singh, Akanksha Krishna, and Smita Kulkarni, "Attendance monitoring using computer vision," in *International Conference on Information and Communication Technology for Intelligent Systems*, Springer, Singapore, 2020. https://doi.org/10.1007/978-981-15-7078-0_72

[18] Gupta, Naman, *et al.*, "Automated attendance system using Open CV," in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, IEEE, 2020. https://doi.org/10.1109/ICRITO48877.2020.9197936

[19] Dalwadi, Darshankumar, Yagnik Mehta, and Neel Macwan, "Face recognition-based attendance system using real-time computer vision algorithms," in *International Conference on Advanced Machine Learning Technologies and Applications*, Springer, Singapore, 2020. https://doi.org/10.1007/978-981-15-3383-9_4

[20] Kothari, K. Viratkumar, "Automated image classification for heritage photographs using transfer learning of computer vision in artificial intelligence," *Turkish Journal of Computer and Mathematics Education (TURCOMAT),* vol. 12, no. 11, pp. 1940–1953, 2021. https://doi.org/10.17762/turcomat.v12i11.6149

[21] Dalwadi, Darshankumar, Yagnik Mehta, and Neel Macwan, "Face recognition-based attendance system using real-time computer vision algorithms," in *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020*, Springer Singapore, 2021. https://doi.org/10.1007/978-981-15-3383-9_4

[22] Li, Zeming, *et al.*, "Detnet: Design backbone for object detection," in *Proceedings of the European conference on computer vision (ECCV)*, 2018. https://doi.org/10.1007/978-3-030-01240-3_21

[23] Bah, Serign Modou, and Fang Ming, "An improved face recognition algorithm and its application in attendance management system," *Array,* vol. 5, p. 100014, 2020. https://doi.org/10.1016/j.array.2019.100014

[24] Mothwa, Louis, Jules-Raymond Tapamo, and Temitope Mapati, "Conceptual model of the smart attendance monitoring system using computer vision," in *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, IEEE, 2018. https://doi.org/10.1109/SITIS.2018.00042

[25] Elaskari, Salah, *et al.*, "Using barcode to track student attendance and assets in higher education institutions," *Procedia Computer Science,* vol. 184, pp. 226–233, 2021. https://doi.org/10.1016/j.procs.2021.04.005

[26] Alemi, Minoo, Ali Meghdari, and Maryam Ghazisaedy, "Employing humanoid robots for teaching English language in Iranian junior high-schools," *International Journal of Humanoid Robotics,* vol. 11, no. 3, p. 1450022, 2014. https://doi.org/10.1142/S0219843614500224

[27] Vrochidou, Eleni, *et al.*, "Social robot NAO as a self-regulating didactic mediator: A case study of teaching/learning numeracy," in *2018 26th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, IEEE, 2018. https://doi.org/10.23919/SOFTCOM.2018.8555764

[28] Olvera, Diana, Uriel Escalona, and Humberto Sossa, "Teaching basic concepts: Geometric forms and colors on a NAO robot platform," *Research in Computing Science,* vol. 148, no. 10, pp. 323–333, 2019. https://doi.org/10.13053/rcs-148-10-27

[29] Jiménez, Martha, *et al.*, "Recognition of colors through use of a humanoid nao robot in therapies for children with down syndrome in a smart city," *Research in Computing Science*, vol. 148, no. 6, pp. 239–252, 2019. https://doi.org/10.13053/rcs-148-6-18

[30] Esquivel-Barboza, Esteban A., *et al*., "Adapting computer vision algorithms to smartphone-based robot for education," in *2020 Fourth International Conference on Multimedia Computing, Networking and Applications (MCNA)*, IEEE, 2020. https://doi.org/10.1109/MCNA50957.2020.9264289

[31] Cielniak, Grzegorz, Nicola Bellotto, and Tom Duckett, "Integrating mobile robotics and vision with undergraduate computer science," *IEEE Transactions on Education,* vol. 56, no. 1, pp. 48–53, 2012. https://doi.org/10.1109/TE.2012.2213822

[32] Altin, Heilo, Alvo Aabloo, and Gholamreza Anbarjafari, "New era for educational robotics: Replacing teachers with a robotic system to teach alphabet writing," in *Proceedings of the 4th International Workshop Teaching Robotics, Teaching with Robotics & 5th International Conference Robotics in Education,* Padova, Italy, 2014.

[33] Kusumota, Victor Luis Pires, Rafael Vidal Aroca, and Felipe Nascimento Martins, "An open source framework for educational applications using cozmo mobile robot," in *2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE)*, IEEE, 2018.

[34] Yan, Haotian, "Computer vision applied in medical technology: The comparison of image classification and object detection on medical images," *Structure,* vol. 6, no. 7, p. 8, 2018. https://doi.org/10.2991/cecs-18.2018.17

[35] Hisey, Rebecca, *et al*., "Real-time workflow detection using webcam video for providing real-time feedback in central venous catheterization training," in *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*, International Society for Optics and Photonics, vol. 10576, 2018. https://doi.org/10.1117/12.2293494

[36] Nugent, Emmeline, *et al*., "Development and evaluation of a simulator-based laparoscopic training program for surgical novices," *Surgical Endoscopy,* vol. 27, no. 1, pp. 214–221, 2013. https://doi.org/10.1007/s00464-012-2423-0

[37] Law, Hei, Khurshid Ghani, and Jia Deng, "Surgeon technical skill assessment using computer vision based analysis," in *Proceedings of the 2nd Machine Learning for Healthcare Conference,* PMLR, 2017.

[38] Kil, Irfan, *et al*., "Development of computer vision algorithm towards assessment of suturing skill," in *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI),* IEEE, 2017. https://doi.org/10.1109/BHI.2017.7897197

[39] Gupta, Ankit, "StrokeSave: A novel, high-performance mobile application for stroke diagnosis using deep learning and computer vision," *arXiv preprint arXiv:1907.05358,* 2019.

[40] Leo, Marco, *et al*., "Computer vision for assistive technologies," *Computer Vision and Image Understanding,* vol. 154, pp. 1–15, 2017. https://doi.org/10.1016/j.cviu.2016.09.001

[41] Maniar, Sarthak, *et al*., "Automated proctoring system using computer vision techniques," in *2021 International Conference on System, Computation, Automation and Networking (ICSCAN),* Puducherry, India, 2021, pp. 1–6. https://doi.org/10.1109/ICSCAN53069.2021.9526411

[42] Ahmad, Istiak, *et al*., "A novel deep learning-based online proctoring system using face recognition, eye blinking, and object detection techniques," *System,* vol. 12, no. 10, 2021. https://doi.org/10.14569/IJACSA.2021.0121094

[43] Hossain, Zarin Tahia, *et al*., "Automated online exam proctoring system using computer vision and hybrid ML classifier," Diss., Brac University, 2021. https://doi.org/10.1109/RAAICON54709.2021.9929456

[44] Yaqub, Waheeb, Manoranjan Mohanty, and Basem Suleiman, "Image-hashing-based anomaly detection for privacy-preserving online proctoring," *arXiv preprint arXiv:2107.09373,* 2021.

[45] Cote, Melissa, *et al*., "Video summarization for remote invigilation of online exams," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, 2016. https://doi.org/10.1109/WACV.2016.7477704

[46] Kumar, Pandey Anoop, "e-Parakh: Unsupervised online examination system," 2020. https://doi.org/10.1109/TENCON50793.2020.9293792

[47] Maniar, Sarthak, *et al.*, "Automated proctoring system using computer vision techniques," in *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)*, IEEE, 2021. https://doi.org/10.1109/ICSCAN53069.2021.9526411

[48] Carini, Robert M., George D. Kuh, and Stephen P. Klein, "Student engagement and student learning: Testing the linkages," *Research in Higher Education,* vol. 47, no. 1, pp. 1–32, 2006. https://doi.org/10.1007/s11162-005-8150-9

[49] Ngoc Anh, Bui, *et al.*, "A computer-vision based application for student behavior monitoring in classroom," *Applied Sciences,* vol. 9, no. 22, p. 4729, 2019. https://doi.org/10.3390/app9224729

[50] Canedo, Daniel, Alina Trifan, and António JR Neves, "Monitoring students' attention in a classroom through computer vision," in *International Conference on Practical Applications of Agents and Multi-Agent Systems*, Springer, Cham, 2018. https://doi.org/10.1007/978-3-319-94779-2_32

[51] van der Haar, Dustin, "Student emotion recognition using computer vision as an assistive technology for education," in *Information Science and Applications*, Springer, Singapore, 2020, pp. 183–192. https://doi.org/10.1007/978-981-15-1465-4_19

[52] Sun, Minghui, Xinyu Wu, Zhihua Fan, and Liyan Dong, "Augmented reality based educational design for children," *International Journal of Emerging Technologies in Learning (iJET),* vol. 14, no. 3, pp. 51–60, 2019. https://doi.org/10.3991/ijet.v14i03.9757

[53] Duraisamy, Prakash, *et al.*, "Classroom engagement evaluation using computer vision techniques," in *Pattern Recognition and Tracking XXX*, International Society for Optics and Photonics, vol. 10995, 2019. https://doi.org/10.1117/12.2519266

[54] Savov, Teodor, Valentina Terzieva, and Katia Todorova, "Computer vision and internet of things: Attention system in educational context," in *Proceedings of the 19th International Conference on Computer Systems and Technologies*, 2018. https://doi.org/10.1145/3274005.3274014

[55] Sumer, Omer, *et al.*, "Teachers' perception in the classroom," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018.

[56] Bosch, Nigel, *et al.*, "Quantifying classroom instructor dynamics with computer vision," in *International Conference on Artificial Intelligence in Education*, Springer, Cham, 2018. https://doi.org/10.1007/978-3-319-93843-1_3

[57] Yujing, Shi, Bai Haijing, and Wang Xuejun, "Image classification of education resources based on texture features," *Procedia Engineering,* vol. 29, pp. 3281–3285, 2012. https://doi.org/10.1016/j.proeng.2012.01.480

[58] Wang, Yunyan, *et al.*, "Image classification based on transfer learning of convolutional neural network," in *2019 Chinese Control Conference (CCC)*, IEEE, 2019. https://doi.org/10.23919/ChiCC.2019.8865179

[59] Sharma, Atharva, *et al.*, "A patch-based convolutional neural network for remote sensing image classification," *Neural Networks*, vol. 95, pp. 19–28, 2017. https://doi.org/10.1016/j.neunet.2017.07.017

[60] Batomalaque, Mark B., *et al.*, "Image to text conversion technique for anti-plagiarism system," *International Journal of Advanced Science and Convergence,* vol. 1, no. 2, pp. 1–9, 2019. https://doi.org/10.22662/IJASC.2019.1.2.001

[61] Bakhteev, Oleg, *et al.*, "Near-duplicate handwritten document detection without text recognition," in *Computational Linguistics and Intellectual Technologies*, Russian State University for the Humanities, 2021.

[62] Kuruvila, S. Jithin, *et al.*, "Flowchart plagiarism detection system: An image processing approach," *Procedia Computer Science*, vol. 115, pp. 533–540, 2017. https://doi.org/10.1016/j.procs.2017.09.111

[63] Maniar, Sarthak, *et al.*, "Automated proctoring system using computer vision techniques," in *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)*, IEEE, 2021. https://doi.org/10.1109/ICSCAN53069.2021.9526411

[64] Silva, Wilson Ricardo Leal da, and Diogo Schwerz de Lucena, "Concrete cracks detection based on deep learning image classification," *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 2, no. 8, p. 489, 2018. https://doi.org/10.3390/ICEM18-05387

[65] Mathew, Alwyn, *et al.*, "An improved transfer learning approach for intrusion detection," *Procedia Computer Science*, vol. 115, pp. 251–257, 2017. https://doi.org/10.1016/j.procs.2017.09.132

[66] Jeyaraj, Pandia Rajan, and Edward Rajan Samuel Nadar, "Computer vision for automatic detection and classification of fabric defect employing deep learning algorithm," *International Journal of Clothing Science and Technology*, vol. 31, no. 4, pp. 510–521, 2019. https://doi.org/10.1108/IJCST-11-2018-0135

[67] Alawad, Mohammed, and Mingjie Lin, "Stochastic-based deep convolutional networks with reconfigurable logic fabric," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 2, no. 4, pp. 242–256, 2016. https://doi.org/10.1109/TMSCS.2016.2601326

[68] Mehran, Ramin, Alexis Oyama, and Mubarak Shah, "Abnormal crowd behavior detection using social force model," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2009. https://doi.org/10.1109/CVPR.2009.5206641

[69] Hanbay, Kazım, Muhammed Fatih Talu, and Ömer Faruk Özgüven, "Fabric defect detection systems and methods—A systematic literature review," *Optik*, vol. 127, no. 24, pp. 11960–11973, 2016. https://doi.org/10.1016/j.ijleo.2016.09.110

[70] Kim, Seung Hyun, and Su Chang Lim, "Intelligent intrusion detection system featuring a virtual fence, active intruder detection, classification, tracking, and action recognition," *Annals of Nuclear Energy*, vol. 112, pp. 845–855, 2018. https://doi.org/10.1016/j.anucene.2017.11.026

[71] Chen, Hang, Dongfang Chen, and Xiaofeng Wang, "Intrusion detection of specific area based on video," in *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, IEEE, 2016. https://doi.org/10.1109/CISP-BMEI.2016.7852676

[72] Nayak, Rashmiranjan, *et al.*, "Video-based real-time intrusion detection system using deep-learning for smart city applications," in *2019 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, IEEE, 2019. https://doi.org/10.1109/ANTS47819.2019.9117960

[73] Rambabu, Kalathiripi, *et al.*, "IoT based human intrusion detection system using lab view," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 8, pp. 557–560, 2019. https://doi.org/10.35940/ijitee.F1115.0486S419

[74] Baltrušaitis, Tadas, Peter Robinson, and Louis-Philippe Morency, "Openface: An open source facial behavior analysis toolkit," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, 2016. https://doi.org/10.1109/WACV.2016.7477553

[75] Baltrusaitis, Tadas, *et al.*, "Openface 2.0: Facial behavior analysis toolkit," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, IEEE, 2018. https://doi.org/10.1109/FG.2018.00019

[76] Shbib, Reda, and Shikun Zhou, "Facial expression analysis using active shape model," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 8, no. 1, pp. 9–22, 2015. https://doi.org/10.14257/ijsip.2015.8.1.02

[77] Gudipati, Vamshi Krishna, *et al.*, "Efficient facial expression recognition using adaboost and haar cascade classifiers," in *2016 Annual Connecticut Conference on Industrial Electronics, Technology & Automation (CT-IETA)*, IEEE, 2016.

[78] Kumar, Yogesh, and Sandeep Sharma, "A systematic survey of facial expression recognition techniques," in *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, IEEE, 2017. https://doi.org/10.1109/ICCMC.2017.8282636

[79] Huang, Xiaohua, *et al.*, "Multi-modal emotion analysis from facial expressions and electroencephalogram," *Computer Vision and Image Understanding,* vol. 147, pp. 114–124, 2016. https://doi.org/10.1016/j.cviu.2015.09.015

[80] M. Sivaram, *et al.*, "Detection of accurate facial detection using hybrid deep convolutional recurrent neural network," *ICTACT Journal on Soft Computing,* vol. 9, no. 2, 2019.

[81] Savov, Teodor, Valentina Terzieva, and Katia Todorova, "Computer vision and internet of things: Attention system in educational context," in *Proceedings of the 19th International Conference on Computer Systems and Technologies*, 2018. https://doi.org/10.1145/3274005.3274014

[82] Bosch, Nigel, *et al.*, "Detecting student emotions in computer-enabled classrooms," *IJCAI*. 2016. https://doi.org/10.1145/2930238.2930371

[83] Kumar, Pradeep, Partha Pratim Roy, and Debi Prosad Dogra, "Independent bayesian classifier combination based sign language recognition using facial expression," *Information Sciences,* vol. 428, pp. 30–48, 2018. https://doi.org/10.1016/j.ins.2017.10.046

[84] Okoli, Chitu, "A guide to conducting a standalone systematic literature review," *Communications of the Association for Information Systems*, vol. 37, 2015. https://doi.org/10.17705/1CAIS.03743

[85] H. Heflin, J. Shewmaker, and J. Nguyen, "Impact of mobile technology on student attitudes, engagement, and learning," *Computers & Education*, vol. 107, pp. 91–99, 2017, https://doi.org/10.1016/j.compedu.2017.01.006

7 AUTHORS

Rubaiya Tasnim Sholi is a Graduate Student of Daffodil International University, Bangladesh, major in Computer Science and Engineering. Currently she is working as a 'Research Assistant' and her research interests lie in computer vision, machine learning, artificial intelligence, and data analytics (ORCID: https://orcid.org/0000-0001-7905-3559).

Md. Fouad Hossain Sarker is an Associate Professor in the Department of Development Studies, under the Faculty of Humanities and Social Science, at Daffodil International University (DIU), Bangladesh. His research interests lie in the areas of e-learning, e-governance, and ESG (E-mail: fouadsarker@daffodilvarsity.edu.bd; ORCID: https://orcid.org/0000-0002-6884-837X).

Md. Salman Sohel is currently working as a Lecturer of Development Studies in Daffodil International University, Dhaka, Bangladesh. Mr. Salman received Master's degree from Hohai University, China. His research interest includes migration, climate governance, disaster management, displacement, resettlement.

Dr. Md. Kabirul Islam is the Professor and Dean of Faculty of Graduate Studies (FGS) at Daffodil International University (DIU), Bangladesh. At present, he is also performing his responsibility as the Director of the Division of Research at DIU, which has a greater impact on establishing a research culture in the university and publications in ISI and Scopus (ORCID: https://orcid.org/0000-0001-5554-3166).

Maruf Ahmed Tamal is currently working as a Research Fellow in the Faculty of Graduate Studies (FGS) at Daffodil International University (DIU), Bangladesh. His primary area of expertise is machine learning and data mining-based knowledge discovery. He also pursues interests in ICT-based pedagogy and mathematics anxiety (ORCID: https://orcid.org/0000-0001-8998-5739).

Dr. Touhid Bhuiyan is a Professor in the Department of Computer Science and Engineering at Daffodil International University (DIU), Bangladesh. His research expertise lies in the areas of cyber security, trust management, artificial intelligence, and e-learning. He also skilled in Database, Big Data Analytics, and has strong engineering professional with a Certificate focused in Cyber Security from University of Oxford.

S.M. Khasrul Alam Shakil is a graduate student of Daffodil International University, Bangladesh, major in Computer Science and Engineering. He has research interests in Machine Learning, Computer Vision and Image Processing (ORCID: https://orcid.org/0000-0003-0975-9892).

Md. Foysal Ahmed is a graduate student of Daffodil International University, Bangladesh, in Computer Science and Engineering department. He has research interests in Neural Network, Computer Vision and Image Processing (ORCID: https://orcid.org/0000-0003-1683-6315).