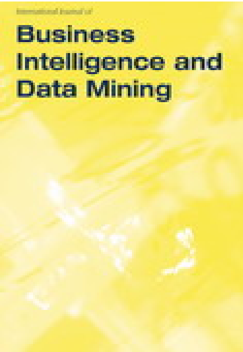


International Journal of Business Intelligence and Data Mining

ISSN online: 1743-8195 - ISSN print: 1743-8187
https://www.inderscience.com/ijbidm

Hybrid automation of student activity records in virtual learning environments in semi-dark scenarios

Manzoor Ahmed Hashmani, Mehak Maqbool Memon, Kamran Raza, Syed Sajjad Rizvi

DOI: 10.1504/IJBIDM.2023.10050463

Article History:

Received:	15 July 2021
Accepted:	17 September 2021
Published online:	30 November 2022

Hybrid automation of student activity records in virtual learning environments in semi-dark scenarios

Manzoor Ahmed Hashmani

Department of Computer and Information Sciences,
High-Performance Cloud Computing Centre (HPC3),
Universiti Teknologi PETRONAS,
Seri Iskandar, Malaysia
Email: manzoor.hashmani@utp.edu.my

Mehak Maqbool Memon*

Department of Computer and Information Sciences,
Universiti Teknologi PETRONAS,
Seri Iskandar, Malaysia
Email: mehak_19001057@utp.edu.my
*Corresponding author

Kamran Raza

Faculty of Engineering Science and Technology,
Iqra University,
Karachi, Pakistan
Email: kraza@iqra.edu.pk

Syed Sajjad Rizvi

Faculty of Computer Science,
SZABIST,
Karachi, Pakistan
Email: sshussainr@gmail.com

Abstract: Online learning is raging due to the offered advantages including the ease of information accessibility and flexible learning experiences remotely. However, it still lacks intelligent authentication and validation of users at the other end. Additional questions are raised for the effectiveness of transfer of knowledge, soft skills, and value of education as the student's engagement. Following the hype and crucial need for full-fledged virtual learning environments (VLEs) in recent times, the literature has witnessed a significant increase in the studies targeting the associated problem areas. The existing VLEs in place fail to provide aggregated functionalities of student authentication and activity tracking in dynamic visual environments. To solve the problem by providing interactive functionality, an automated hybrid information activity system (HIAS) based on facial biometrics, facial emotion recognition using computer vision techniques is proposed. The proposed system solves the problem intelligently by analysing the incoming online video from the tutor's computer screen.

Keywords: facial biometrics; emotion recognition; computer-vision; online videos.

Reference to this paper should be made as follows: Hashmani, M.A., Memon, M.M., Raza, K. and Rizvi, S.S. (2023) 'Hybrid automation of student activity records in virtual learning environments in semi-dark scenarios', *Int. J. Business Intelligence and Data Mining*, Vol. 22, Nos. 1/2, pp.16–33.

Biographical notes: Manzoor Ahmed Hashmani received his PhD degree in Communication Networks from the Nara Advanced Institute of Science and Technology (NAIST), Japan. He is currently with the Department of Computer and Information Sciences, Universiti Teknologi PETRONAS (UTP), Malaysia. His research interests include artificial intelligence, machine learning, blockchain technology, and communication networks.

Mehak Maqbool Memon received her ME (Masters of Eng.) in Information Technology from Mehran University of Engineering Technology (MUET) University Pakistan. She is currently pursuing a Doctorate degree in Information Technology from Universiti Teknologi PETRONAS (UTP), Malaysia. Her research interests include data analytics, machine learning, computer vision, and artificial intelligence.

Kamran Raza is a dedicated, energetic and devoted individual serving as a Professor and Dean at the Faculty of Engineering Science and Technology IQRA University Karachi, Pakistan. He is also the coordinator for Professional Engineering Body (PEB) established at IQRA University. His research interests are computer science, machine learning, and network security.

Syed Sajjad Rizvi is an eminent scholar and an academician. He is a certified professional computer system engineer with Masters degrees in Telecommunication and Business Administration. He earned his PhD in the area of image processing and information retrieval. In addition, he has also been awarded the status of approved PhD Supervisor by Higher Education Commission, Pakistan. He poses rich teaching experience and used to teach the courses in the domain of computer sciences, communication and managements both at undergraduate and postgraduate level.

1 Introduction

A major issue in virtual learning environments is user authentication and validation for information access of the educational activities (Hashmani et al., 2020), same concern in every virtual environmental setup such as transactions (Poongodi et al., 2020a; Poongodi et al., 2019). Another issue of VLEs is the tracking of student's performance by means of activities performed is crucial to resolve. The incentive behind finding the solution is to provide flexible and closed network information access to the students. User authentication by protecting the institution's private network, restricts the invalid user traffic over the network. Meanwhile, facial emotion recognition will provide considerable chances of enhancement of teaching and learning practices in online learning. Facial biometrics and emotion analysis is considered strongest for virtual learning environments for security authentication (Valera et al., 2015; D'Souza et al., 2019).

Following the COVID-19 pandemic, where all the countries have different approaches to handle the viral disease; one commonly witnessed practice to cease the spread of deadly disease is shifting of physical classroom settings to virtual classroom settings (Memon et al., 2020). In Malaysia, over 1.2 million students are enrolled in 20 public universities and 41 private universities (Statistik Pendidikan Tinggi, 2019). All these universities are striving to find the balanced measures to ensure the transfer of the knowledge. Even industries are now providing online trainings. Over 82% organisations conduct some training online and 28% organisations conduct fully online training (Stratistics Market Research Consulting Pvt. Ltd., 2019). Online learning challenges during pandemic in Malaysian Higher Learning Institutions for adaptation of new 'norm' include low internet access as reported by Malaysian students (Ismail and Wafa, 2020). The ever-growing online learning industry demands high quality imaging analysis tools and algorithms to fulfil user demands of enhanced experiences at minimal imposed requirements and efforts. Despite of switching to fast internet speeds and evolving fifth generation internet features (Poongodi et al., 2020b), currently limited bandwidth is one of the major issues. The low internet bandwidth and dynamic environmental conditions highly affects the functionality of intelligent computer vision (CV) systems. The notable mentions through which VLEs can be aided are facial recognition and emotion analysis technologies. The facial recognition and emotion analysis technologies based on digital image processing and CV techniques by comparing user facets from images and management database (Valera et al., 2015) proper recognition and monitoring interventions to online learners has become crucial element of VLEs for adequate delivery of knowledge among individuals. Alongside the recognition and monitoring systems automated sentiment analysis of students can be accomplished by analysing student's comments posted in closed groups using intelligent analysis tools (Poongodi et al., 2019b).

CV techniques provide different methods for analysis of gesture cues, postures, eye movements and facial expressions. Due to the non-intrusive nature of CV techniques, it has been widely used for automation industry (Memon, 2017). Similarly, due to this feature of CV techniques it is proposed to be incorporated in online classroom settings where students are to be monitored without disturbing their lecture and assessment activities. The readily available cell phones, tablets and computers with the built-in cameras at low cost have made the monitoring processes easier to incorporate in existing VLEs. However, the concerns related to digitally process the image persists. The processing of digital image is susceptible to the quality of incoming data, whereas in the VLEs students will have dynamic range of camera types, data bandwidths, lightning conditions, different image sizes, multiple user tracking and so on. For this reason, flexible online student management systems with relevant CV handlers for versatile user environments are needed (Dewan, 2019).

Despite the availability of potential solutions for interactive VLEs in terms of sophisticated CV techniques, the current VLEs provide constrained functionality in following aspects:

- Limited systems taking in account the relevance of human brain activity and change in facial features for processing the incoming videos over time.
- Absence of full-fledged system providing handlers for dynamic visual conditions in VLEs.

- Inexistence of aggregated intelligent system that duly recognises and analyses sentiments via facial videos to track activity in VLEs.
- No well-defined smart structure for automatic record-keeping VLEs.

For student authentication and monitoring to keep a check on student's responsiveness we propose a hybrid approach namely hybrid information activity system (HIAS) based on CV techniques. The proposed HIAS is exclusive in terms of the following features:

a Discrete analysis of facial feature analysis for the motivation level calculation based on the psychological aspects of the human brain.

b Identification of influential factors affecting the recognition process along with the automated handlers to minimise their effect.

c Detailed visual structural mechanics for integration of face recognition and face sentiment analysis to track student's activity in VLEs.

d Enhancement of teaching and learning (T&L) practices by providing no human intervention for student's record-keeping and quantifiable engagement levels in an online lecture.

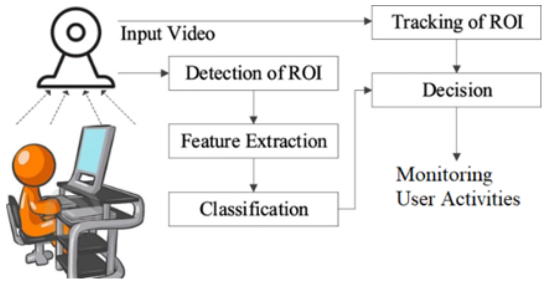
In this study, CV based hybrid student activity tracking for student authentication and record-keeping along with student's responsiveness is proposed. In Section 2, we briefly introduce CV research works relevant to the proposed method for monitoring student's activities in VLEs. In Section 3, the structure of our proposal is presented along with all its sub-modules. In Section 4, the preliminary data gathering protocol design is discussed along with the initial results. Section 5 concludes the entire study by highlighting the contributions. Finally, Section 6, points out the future direction of research.

2 Related work

CV methods for estimation of user activities are used similarly in the manner the activities are monitored and judged by the instructors in physical classroom settings. The entire analysis is dependent on the perceived information from the cameras mounted on the individual user's computer and transmitted to the instructor's computer over the internet and analysis takes place at the instructors' end. This automation of perceived student's activity is probable to help and revolutionise the academic management system which would be less prone to cheating practices in exams and false grades. This automation relies on the accurate predictions and it drives all its power from sophisticated machine learning and statistical base (Poongodi et al., 2020c; Poongodi et al., 2020d).

The generic framework for the student's tracking and recognition is based on the capturing the student's video, that video is processed for further detection of region of interest (ROI) which is face in the case of vision base recognition. Prominent facet is extracted by means of feature extraction and the features are classified and finally a decision is made for continuous monitoring and tracking of user activities (Dewan, 2019). Figure 1 shows the generic CV framework of for tracking activities.

Figure 1 Generic CV framework for monitoring user activities (see online version for colours)



Source: Dewan (2019)

2.1 CV based methods for monitoring student's activities in VLEs

Virtual learning environments providing distance learning are now the popular means of transfer of information to students which is supported to internet users around the globe. For effective and efficient learning educational institutes are facing problems like plagiarism and cheating (Fenu et al., 2018). For this reason, image processing and biometrics are expected to play a crucial part to save the E-learning processes from such activities. Face detection and recognition processes are being incorporated in online learning settings (Fenu et al., 2018; Wang et al., 2020). According to a literature survey Rabiha et al. (2018), similar image processing and vision-based algorithm are used for online image processing as well. The methods were critically reviewed and presented in the study for their implementation in VLEs. According to the critical analysis presented 3WPCA (Three level wavelet decomposition – principal component analysis (PCA)) was the fastest and accurate face-based recognition module with accuracy of detection up to 98% (Winarno et al., 2017). System proposed in Tantak et al. (2017), created automatic attendance for staff and students using recognition algorithm. The recognition system uses PCA and Eigen face approach for face recognition. They have analysed the system with different facial expressions and by adding salt and pepper noise in the facial image. The attendance gets updated in the lecturer's laptop every time student registers for attendance. One major limitation of this E-attendance system is it can only process on face at a time. Proposed image processing model in study (Rabiha and Ham, 2017), uses image processing techniques for student authentication in online environments. The system creates facial feature descriptor using Viola Jones algorithm and saves the recognised images into database as well. The study can be further extended for large scale datasets and to recognise multiple faces in a particular frame. A video based facial recognition module for online learning is proposed in Mian (2011). The proposed model uses local features for recognition, based on the similarity of feature the faces get clustered. The entire process of gathering and clustering faces continues until minimum number of faces and the distance of the farthest face from its cluster mean is below a threshold. The experiments are performed on dataset comprising information of 50 known subjects and 22 unknown subjects to the classifier model. The recognition rate of the proposed model is 97.8%. The system can be enhanced for incremental learning of the model and to tackle the variable environmental conditions. Another E-learning system proposed in Ayvaz et al. (2017), recognises facial emotions from the video captured from video conferencing app using the algorithm (Sagonas et al., 2013) for facial landmark detection and tracking. Further SVM and K nearest neighbours are used for emotion recognition. The system approach can be extended for facial authentication as well. However, some enhanced functionality is still required in it for handling various environmental settings.

2.2 Intuitive critical comparative analysis of student management system via CV

The studies related to VLEs referred in this section are critically compared based on the functionality mechanics behind the presented proposals. Moreover, only studies involving to topics such as CV and machine learning in context of VLEs have been referred for the detailed analysis and thereby further futuristic approaches are presented in next section to enhance the overall interactive user experiences. For the insightful comparative analysis, the research studies are gathered by querying Google Scholar. The search keywords included 'VLEs' along with combinations of 'attendance', 'sentiments', 'engagement evaluation', 'attendance management', 'CV' and 'machine learning'. The studies published over period of 2015 through 2021 are finalised for the later stage of analysis. The comparative analysis of shortlisted research studies is further divided based on the targeted research problem and its relevant solution.

Prior to the comparative analysis, it's worth awhile to shed some light on the working mechanics of the referred studies. Research in Ja et al. (2020) is backed by convolutional neural networks with multi-layered perceptron for facial recognition and relevant monitoring of student for recording attendance. The authors also presented Video databases presented in uncontrolled conditions of environment. However, no clear quantifiable performance numeric is reported. Research presented in Su et al. (2020), uses MoblileFaceNet-Deep based CNN approach for the face recognition task and student monitoring. The research is conducted in physical classroom environment settings. The reported accuracy is around ROI of 89%. Research presented in Filippidou and Papakostas (2020), uses MTCN face detector, along with Convolutional Neural Networks pre-trained with Imagenet (DenseNet121) for face recognition. This research is also conducted in physical classroom environment. The reported accuracy is around 99%, with no clear details of attainment of that high accuracy score. Research presented in Rusdi et al. (2020), is backed by CNN for face recognition. The research is limited in its course as it reports camera is placed in a strategic location for conducting the experiments. Moreover, two terms are ambiguously and interchangeably used, i.e., Raster-CNN or Region CNN. The reported accuracy is around 90%. Research referred in Dalwadi et al. (2020), uses Haar Cascade as face detector, Eigen face recogniser along with CNN for attendance keeping. Student is supposed to stand in front of camera for certain time.

The reported accuracy is 80% under normal lightning conditions. Research presented in Son et al. (2019), uses support vector machines and lazy learning approach of nearest neighbour search for face recognition based on four facial landmarks. Physical presence of subject/ student is needed. The research reported 80% F1 score for the controlled test experiments. Research presented in Harikrishnan et al. (2019), uses Haar Cascade for face detection and local binary patterns histogram for recognition. The prototype provides two modes of functionality single snap mode or real time surveillance mode. The subsequent reported accuracy is around 74%. Research presented in Tamimi et al. (2015), uses MathLib library for detecting face parts. The experiments are conducted in physical classroom setting with detection ratio 75% to 94.73%. For the engagement analysis of student's behaviour in lectures research presented in Liao et al. (2021), uses SE-ResNet-50 (SENet) and long-short-term memory (LSTM) network with global attention (GALN) for recognition and activity monitoring via facial expressions. Classifies student's activities in low or high engagement. No clear quantifiable performance numeric is reported. Research presented in Li et al. (2020), uses two different VGG-16 convolution neural networks for face expression recognition and mouse clicks to record student's activity. The study does not clearly mention if physical presence of student is required. The activity recording is based on facial expression and number of clicks. Strong correlation b/w mouse clicks and face expressions is observed. Whereas no clear quantifiable performance numeric is reported. Research presented in Jackson (2020), uses semi-automatic CV-based approach. The research experiments are conducted in physical classroom settings. Here again, no clear quantifiable performance numeric is reported. Research presented in Wang et al. (2020), uses CNN for activity tracking of students in physical classroom settings. Research presented in Olivetti et al. (2020), uses support vector machines for self-report classification and geometrical descriptors for facial expression tracking. The system evaluates engagement while interacting with a virtual environment by integration of self-report method and facial expression recognition. Physical presence of student is needed. Research presented in Karimah (2020), uses CNN for student's engagement assessment. The system provides functionality for extraction of images every 10 seconds snippet video. Training accuracy is 71% and the validation accuracy is 62%. Research presented in Sharma et al. (2019), is backed by Haar-cascade for face detection and binary classifier of CNN for recording eye activity. Classification is based on concentration index, a score between 0% and 100% is given to classify the student's level of engagement. Research presented in Sun et al. (2018), uses CNN and LSTM for facial expression tracking. The working functionality is constrained by Physical classroom settings. Engagement analysis of one person at a time. The reported accuracy is up to 72%.

The literary studies relating to student monitoring and activity analysis focuses on providing solutions for student attendance and engagement analysis separately. For this reason, the presented critical comparative analysis is classified into two categories based on the end solutions to the problem, i.e., student attendance management systems and engagement analysis. Table 1 presents the evaluated studies focusing attendance and engagement analysis in perspective of VLEs, respectively.

The analysis presented in Table 1., shows all the recent conducted research in domain of student activity recording are constrained by the controlled environment or the imaging sensors used for the working functionality of the proposed solutions. Moreover, only a few solutions work for processing or recording activities of more than one student at a time.

Table 1 Critical analysis of CV-based student's activity analysis

Student's attendance monitoring systems						
Reference	*Year*	*Controlled sensor*	*Handler for environmental dynamicity*	*Virtual setting*	*Parallel student processing*	*Critical analysis*
Ja et al. (xxxx)	2020	Logitech camera	✗	✗	✗	• Does not provide details of data gathering protocol.
Su et al. (2020)	2020	PTZ Logitech camera series	✗	✗	✓	• Works for different seating arrangements of students. • Takes in account only deep learning (DL) techniques.
Filippidou and Papakostas (2020)	2020	Fixed webcam for entire class	✗	✗	Not mentioned	• Benchmarking with 4 different networks. • Probable to high error rates due to image augmentation usage.
Rusdi et al. (2020)	2020	Fixed camera	✗	✗	✓	• Size of dataset is not mentioned which makes the achieved accuracy questionable. • Raster-CNN and Region CNN terms are used interchangeably which makes study ambiguous.
Dalwadi et al. (2020)	2020	Fixed camera	✗	✗	Not mentioned	• Works for limited lightning conditions.
Son et al. (2019)	2019	IP camera	✗	✗	Not mentioned	• Detailed system workflow diagrams. • Dataset updates of student images cannot be done.
Harikrishnan et al. (2019)	2019	Fixed camera	✗	✓	✓	• Provides in-app user flexibility to choose operating mode. • Constrained by camera placement and angles.
Tamimi et al. (2015)	2015	Fixed camera	✗	✗	✓	• Uses information of distance b/w facial features. (Easy yet robust inference) • Limited dataset of merely 12 students, 10 images each.

Table 1 Critical analysis of CV-based student's activity analysis (continued)

Ref.	*Year*	*Controlled sensor*	*Handler for environmental dynamicity*	*Virtual setting*	*Parallel student processing*	*Critical analysis*
Student's engagement analysis systems						
Liao et al. (2021)	2021	Microsoft lifecam wide-angle F2.0 camera	✗	✓	Not mentioned	• Sheds light on the future research aspect if engagement levels are defined by different categories or continuous values. • Limited to certain visual environments.
Li et al. (2020)	2020	IR imaging camera	✗	-	✗	Backed by two different conditions for recording activity, i.e., facial image and mouse movement data.
Jackson (2020)	2020	Fixed camera	✗	✗	Not mentioned.	• Reported experiment accuracy hold for different lightning conditions throughout the day. • Limited to certain camera angle, i.e., 180° view of the camera
Wang et al. (2020)	2020	Fixed webcam	✗	✗	✓	• Results in final classification of student's mood in 4 classes only.
Olivetti et al. (2020)	2020	SR300 structured-light depth camera	✗	✗	Not mentioned.	• Limited participants, i.e., 12 students.
Karimah (2020)	2020	No clear mentions.	✗ Mentions illumination normalisation for future work	✓	Not mentioned.	• Considers three engagement classes, namely, very engaged, normally engaged, or not engaged. • Overfitted model for prediction due to incoming feature data distribution.
Sharma (2019)	2019	Web-camera present in a laptop computer	✗	✗	✗	• Identifies student is focused or distracted then on top of the result add information of seven emotions. • Limited participants, i.e., 15 students.
Sun et al. (2018)	2018	Fixed camera	✗	✗	✗	• Quantifies evaluation of learners' learning status in numbers.

2.3 Literary deductions

The virtual learning environments are increasingly becoming widespread, along with the rising concerns of student's presence and activities performed online. The literary studies uncover the fact that there is absence of a system which takes in account the student's authentication and activity for automated report generation. Moreover, the existing systems are confined by the camera usage. However, for the implementation of student authentication and activity analysis every student will have different working device which means different cameras for the capturing student's activities while interacting with VLE. Another noteworthy aspect is that the existing research knowledge does not clearly define a line for the presence of student for the working environment of proposed solutions. It is not clearly mentioned if the student data is transmitted over internet and processed over instructor's device or not. Finally, as the VLEs are going to be dynamic in nature where every student will have different lightning conditions; respective intelligent handlers should be incorporated in proposed solutions.

3 Methodology

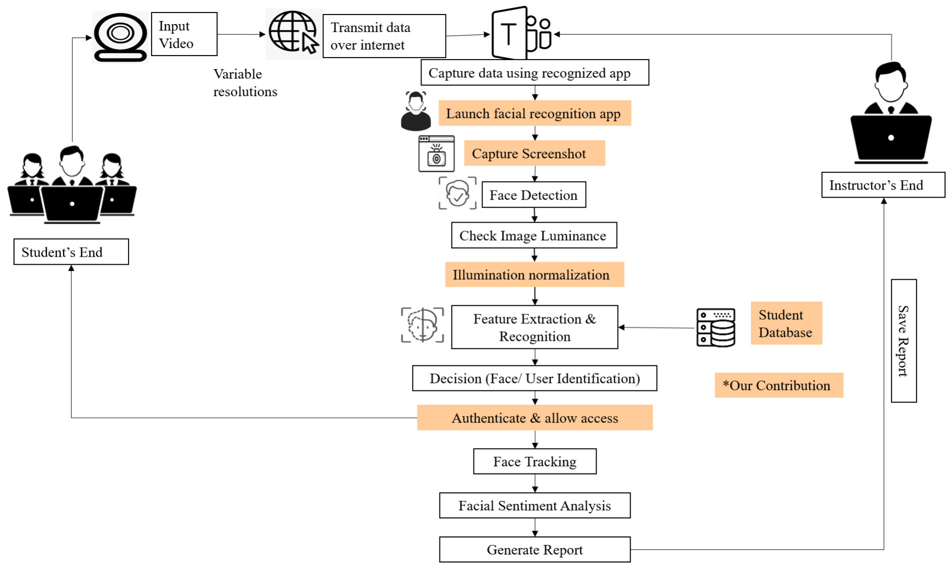
The rapid switch of entire teaching and learning practices from analogue to digital ones affects the overall learning behaviour of the students. Hence, enhanced learning management systems are required for offering better learning practices. The systems should be developed and enhanced in such a way that they should provide full-fledged functionalities of student teacher interaction in VLEs as the physical classroom counterparts.

The proposed system named HIAS will be covering up all the aspects of automated student management and information systems (MIS) along with the feedback mechanisms to ensure maximum transfer of knowledge in different geographical locations by tracking student's activity. The deployment details and the individual inference working details are comprehensively discussed in Section 3.1 and 3.2, respectively.

3.1 Deployed Infrastructure of HIAS

The proposed HIAS targets the online settings, so the key component is the communication between student's end and instructor's end. The communication usually is provided by means of MS teams, BBB, zoom these days because of the functional feasibility or some educational institutes also prefer their own online portals. Nevertheless, problem remains same, i.e., variant resolutions and variable lightning conditions of incoming image of every student. The reason could be the webcam specs used by each student or the internet bandwidth provided by the service provider. For that matter, the facial recognition system should be efficient enough to work with dynamic environments. Working mechanics of proposed student recognition system is shown Figure 2.

Figure 2 Infrastructure of student recognition activity tracking in VLEs (see online version for colours)



During the provided session, the student and instructor both are required to launch the usual app for virtual class communication. The input video from student's end is transmitted over internet. At the instructor's end facial recognition app is launched to capture the screenshots of instructor's computer screen involving frontal face images of students and are saved in .png or .jpeg format. The intervals between the screenshot capturing would be based on the psychological studies showing the timespans of brain activity and inactivity. The incoming screenshot would be further cropped based on the facial images of the student in the entire user interface of the communication application. The cropping operation takes place to eliminate the extra information of lecture presentation slides and other parts of image/screenshot containing no information. The cropped portions of the image having faces is further checked for illumination separately if the illumination of image turns out to be dark or semi-dark then illumination normalisation takes place else the image is passed to next step without normalisation. After those multiple faces are detected in the captured frames further these detected faces are recognised by matching the identities with student database. If the face gets validated information access is given to student and the attendance gets updated. Further processing is continued for the calculation of motivation level of student in online lecture using facial sentimental analysis and finally feedback reports are generated for teacher's reference.

3.2 Smart inference core of HIAS

The proposed student information activity module is designed in such a way that it provides exclusive handlers for following problems:

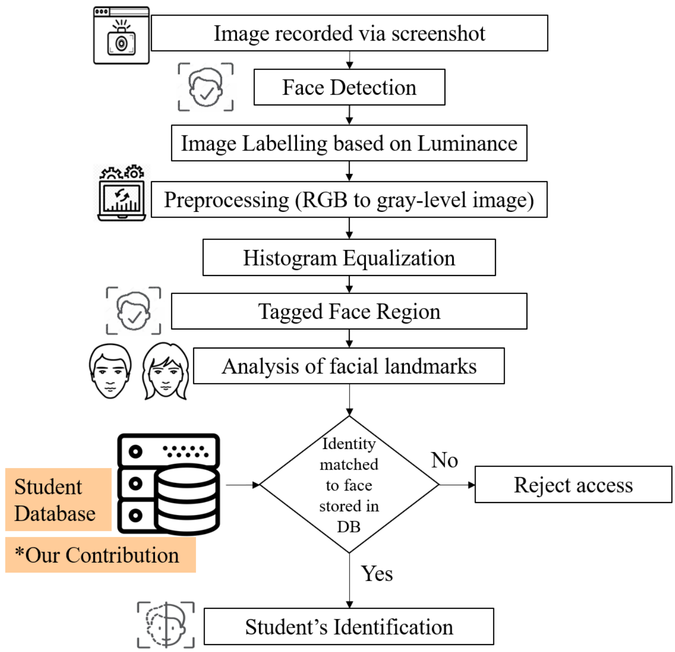
a Environmental dynamicity of lightning conditions of each student's incoming video.

b Parallel processing of each student's activity and relevant record-keeping.

c Hybrid system keeping attendance and student's feedback without intruding student's learning activities.

The entire HIAS infrastructure is divided into two small modules based on the working mechanics providing different end functionalities. The first module will cover the student's facial recognition task and providing the access to only the registered students or users. The working mechanics of facial recognition module as presented in Figure 3.

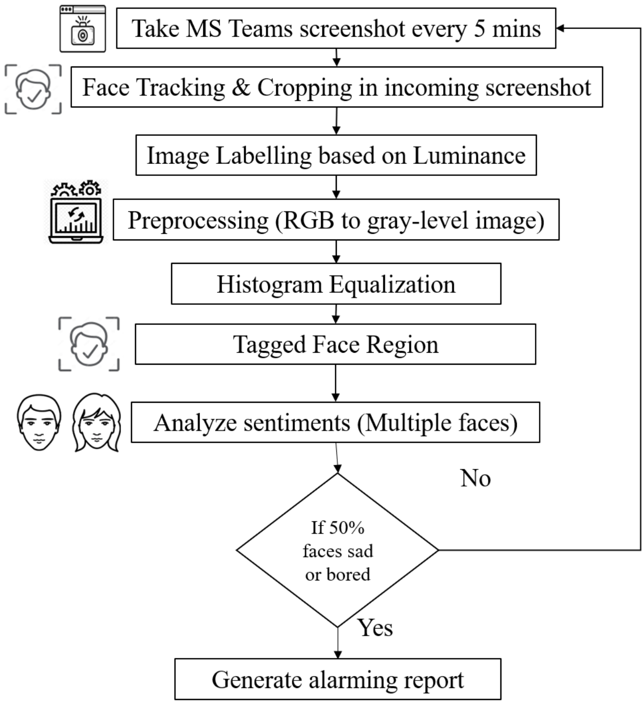
Figure 3 Working mechanics of student's authentication module (see online version for colours)



After successful connection is developed over internet between tutors and access requesting parties, HIAS starts recording the images of tutor's MS teams interface having images of different students. The recorded images are cropped based on the presence of faces in the captured the MS teams screen. The faces in the image are detected by the Paul Viola's object detector (Viola and Jones, 2004). Based on the detected faces bounding boxes around the faces are created. The bounding boxes then result in cropping of facial images residing in tutor's screen. The cropped facial images are further processed to calculate the luminance and accordingly histogram equalisation (Pizer et al., 1987) is employed to brighten up the image. The brighten images are passed to next step for identification the facial landmarks using histogram of oriented gradients (HoG) (Dalal and Triggs, 2005). The processing of cropped facial images includes identifying and locating the accurate facial landmarks using HoG for student's recognition. Finally support vectors machine or K- nearest neighbours' techniques is probable to be employed over the facial histograms for further identification of registered student. After the identification and authentication of registered student his/her record for the relevant session would be automatically updated. Finally, the access is intelligently granted to registered students else the access is rejected.

The second module will cover the student's sentiment analysis to help tutors in enhancing and providing interactive teaching and learning practices. The working mechanics of facial recognition module as presented in Figure 4.

Figure 4 Working mechanics of student's sentiments about lecture



The tutor's screen is continuously monitored to keep the track of student's activities throughout the lecture. The videos are divided into frames. These frames are separated by intervals of 5 minutes. Relevant luminance of facial images is calculated and histogram equalisation is performed to brighten the dark and semi-dark images. The Similar method of HoG and SVM is used to classify sentiments of each student in seven different categories. Finally, the sentiments of entire class are analysed if the more than 50% students are sad or bored negative report is generated. The negative report can then help teacher to change the teaching styles.

4 Preliminary enactment of interactive VLEs

For the data gathering phase of the presented research, initial trials of the intelligent prototype have been implemented. The prototype extracts the student's visual data from the verified application of university to conduct the online lectures. In this case, we have used MS teams to gather the facial data of students to conduct further recognition and sentiment analysis tasks. However, the implementation of the prototype is flexible enough to change it to any other virtual platform. Moreover, the same data gathering steps can be easily performed online or over the recorded sessions. The implementation of data gathering module is accomplished by using HoG (Dalal and Triggs, 2005) which allows the face detection in the incoming video data. Additionally, the prototype provides one more user flexibility feature in terms of the timer after which a single frame is extracted. Consequently, user can change the discrete interval time parameter depending on the available resources (disk memory) to save the frames in form of image data. The stepwise protocol of the implementation is presented in algorithm for data gathering.

Algorithm: data gathering module

Load *lecture session as video V*
Initialise *time delay parameter as TD*
Load *HoG model for detecting faces in V*
Faces = *HoG_detector* (*V*)
For each f_i *in V*
Begin
if faces is not empty *&&* *if* ((last_img_taken_time is empty) | ((system_time – last_img_taken_time) > = *TD*))
Set last_img_taken_time = system_time
Save f_i as image *i*
End

Following the algorithm for data gathering, the initial data gathering steps are performed and the data samples are presented in Table 1.

Table 2 Samples of student data captured in wild (see online version for colours)

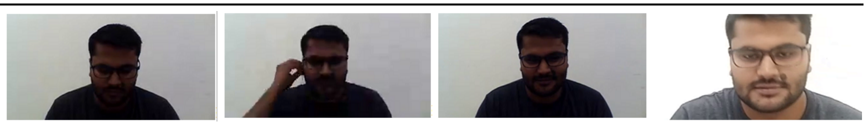
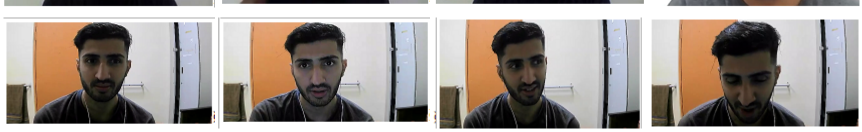
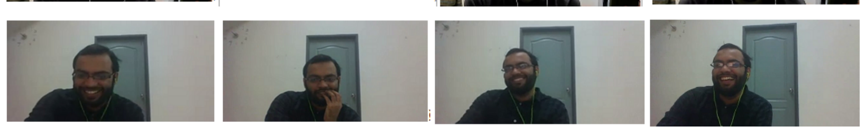
Student ID	*Extracted image samples*
Student ID 1	
Student ID 2	
Student ID 3	

Table 1 shows some of the samples captured from the recorded video of MS Teams. The dimension of the extracted frames depends on the resolution of the incoming video. Here, in these experiments we got the resolution of 1,920 by 972 depending on the video resolution. The samples show that while working with the virtual environment varying lightning conditions, facial angles and distances are expected. Moreover, the facial coverage would also be different for different devices. For the Student_ID 1, the fourth image in row is captured when student logged in session via the mobile device. Also, the lightning conditions are witnessed to be different by just changing the devices and keeping the environmental conditions (lightning) same. The proposed system HIAS is expected to tackle all the variant lightning conditions by providing sub-module which will manipulate the incoming image on pixel level. This manipulation of image on pixel level will result in the identification of luminance of the incoming image and will allow HIAS to perform relevant operations of brightening the low luminance images.

5 Conclusions

The presented research formulates a hybrid approach for tracking student activity without disturbing student or any extra requirement of efforts used by tutors. The hybrid approach is expected to add value to the existing teaching practices. Initially, a comprehensive critical review of research conducted in domain of VLEs for tracking student's activity is presented and classified into different categories. Following the literary deductions, a hybrid approach of student's recognition and activity tracking is proposed. The proposed framework of online activity monitoring makes sure to unify versatile incoming visual data and process it accordingly using different CV methods and enhancement techniques. The CV techniques are proposed in the mediatory layers of HIAS to normalise the illumination problems. The presented study also discusses the detailed mechanics of each singular component of HIAS separately and finally propose an integrated scheme. An additional feature of HIAS is presented, based on the detailed study of facial cues changes in relevance to the brain activity which determines the interval between extracted frames. This additional feature is expected to reduce the computational resources used for tracking activities.

6 Future directions

The future work includes the creation of organisation related master dataset for the student recognition and authentication process accompanied by the addition of proposed CV techniques in the implementation course. The CV techniques used for the recognition and tracking tasks will also be compared with the state-of-the art techniques of Machine Learning to redound the gains in terms of system accuracy. The techniques resulting in higher accuracy will become crucial component of HIAS. Finally, the integrated proposed framework would be implemented on an organisational level and further software testing sessions would be conducted to check overall user-flexibility of the framework.

Acknowledgements

This research study is a part of the funded project under a matching grant scheme supported by Iqra University, Pakistan and Universiti Teknologi PETRONAS (UTP), Malaysia.

References

Ayvaz U., Gürüler, H. and Devrim, M.O. (2017) 'Use of facial emotion recognition in e-learning systems', *International Technologies and learning tools.,* Vol. 60, No. 4, pp.95–104, ISSN: 2076-8184, DOI:10.33407/ITLT.V60I4.1743.

D'Souza, J.W., Jothi, S. and Chandrasekar, A. (2019) 'Automated attendance marking and management system by facial recognition using histogram', in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, IEEE, pp.66–69.

Dalal, N. and Triggs, B. (2005) 'Histograms of oriented gradients for human detection', in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, IEEE, Vol. 1, pp.886–893.

Dalwadi, D., Mehta, Y. and Macwan, N. (2020) 'Face recognition-based attendance system using real-time computer vision algorithms', in *International Conference on Advanced Machine Learning Technologies and Applications*, Springer, pp.39–49.

Dewan, M.A.A., Murshed, M. and Lin, F. (2019) 'Engagement detection in online learning: a review', *J. S. L. E*., Vol. 6, No. 1, p.1.

Fenu, G., Marras, M. and Boratto, L. (2018) 'A multi-biometric system for continuous student authentication in e-learning platforms', *J. P. R. L*., Vol. 113, pp.83–92, Special Issue: Integrating Biometrics and Forensics, edited by Nappi, M., Memon, N., Riccio, D. and Uhl, A.

Filippidou, F.P. and Papakostas, G.A. (2020) 'Single sample face recognition using convolutional neural networks for automated attendance systems', in *2020 Fourth International Conference on Intelligent Computing in Data Sciences (ICDS)*, IEEE, pp.1–6.

Harikrishnan, J., Sudarsan, A., Sadashiv, A. and Ajai, R.A. (2019) 'Vision-face recognition attendance monitoring system for surveillance using deep learning technology and computer vision', in *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, IEEE, pp.1–5,.

Hashmani, M.A., Junejo, A.Z., Alabdulatif, A.A. and Adil, S.H. (2020) 'Blockchain in education–track ability and traceability', in *2020 International Conference on Computational Intelligence (ICCI)*, IEEE, pp.40–44.

Ismail, N.M.A.B.N.S. and Wafa, S.W.W.S.S.T. (2020) 'Online learning challenges during pandemic COVID-19 in Malaysian higher learning institution', *Universal Journal of Educational Research,* Vol. 8, No. 12, pp.7151–7159.

Ja, J.R., Kb, K., Vc, V. and Moorthy, A. (2020) 'Computer vision based attendance monitoring system', *VDGOOD Journal of Computer Science Engineering*, Received: 14 March 2020; Revised: 30 March 2020; Accepted: 20 April 2020.

Jackson, S. (2020) Classroom engagement evaluation using multi-sensor fusion with a 180° camera view', *Urban Intelligence and Applications, Second International Conference, ICUIA 2020*, Taiyuan, China, 14–16 August, Revised Selected Papers.

Karimah, S.H.S.N. (2020) 'A real-time engagement assessment in online learning process using convolutional neural network', in *The Asian Conference on Education 2020 Official Conference Proceedings, IAF or The International Academic Forum*.

Li, Z. and Zhan, Z. (2020) 'Integrated infrared imaging techniques and multi-model information via convolution neural network for learning engagement evaluation, *Infrared Physics and Technology*, September, Vol. 109, Article Id. 103430.

Liao, J., Liang, Y. and Pan, J. (2021) 'Deep facial spatiotemporal network for engagement prediction in online learning', *J. A. I.*, pp.1–13, 2021-02-07, DOI: 10.1007/s10489-020-02139-8.

Memon, M.M., Hashmani, M.A., Jameel, S.M., Junejo, S. and Raza, K.J. (2020) 'Analysis of student procrastinatory behavior in virtual learning environments using machine learning', *J. o. H. U. N. S.*, Vol. 47, No. 1.

Memon, S., Memon, M., Bhatti, S., Khanzada, T.J. and Memon, A.A. (2017) 'Tracker for sleepy drivers at the wheel', in *2017 11th International Conference on Signal Processing and Communication Systems (ICSPCS)*, IEEE, pp.1–8.

Mian, A. (2011) Online learning from local features for video-based face recognition', *J. P. R.,* Vol. 44, No. 5, pp.1068–1075.

Olivetti, E.C., Violante, M.G., Vezzetti, E., Marcolin, F. and Eynard, B. (2020) Engagement evaluation in a virtual learning environment via facial expression recognition and self-reports: a preliminary approach', *J. A. S.*, Vol. 10, No. 1, p.314.

Pizer, S.M. et al. (1987) 'Adaptive histogram equalization and its variations', *Computer Vision, Graphics and Image Processing*, Vol. 39, No. 3, pp.355–368.

Poongodi, M. et al. (2020d) 'Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system', *Elsevier Computers and Electrical Engineering*, Vol. 81, p.10652.

Poongodi, M., Hamdi, M., Varadarajan, V., Rawal, B.S. and Maode, M. (2020a) 'Building an authentic and ethical keyword search by applying decentralised (blockchain) verification', in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, pp.746–753.

Poongodi, M., Hamdi, M., Vijayakumar, V., Rawal, B.S. and Maode, M. (2020b) 'An effective electronic waste management solution based on blockchain smart contract in 5G communities', in *2020 IEEE 3rd 5G World Forum (5GWF)*, IEEE, pp.1–6.

Poongodi, M., Vijayakumar, V., Al-Turjman, F., Hamdi, M. and Ma, M. (2019a) 'Intrusion prevention system for DDos attack on VANET with ReCAPTCHA controller using information based metrics', *IEEE Access*, Vol. 7, pp.158481–158491, doi:10.1109/access.2019.2945682.

Poongodi, M., Vijayakumar, V., Chilamkurti, N.J.I. and Transactions, S. (2020c) 'Bitcoin price prediction using ARIMA model', *J. o. I. T.*, Vol. 10, No. 4, pp.396–406.

Poongodi, M., Vijayakumar, V., Ramanathan, L., Gao, X-Z., Bhardwaj, V. and Agarwal, T.J. (2019b) Chat-bot-based natural language interface for blogs and information networks', *I. J. o. W. B. C.*, Vol. 15, No. 2, pp.178–195.

Rabiha, S.G. and Ham, H. (2017) 'Image processing model based e-learning for students authentication', in *2017 International Conference on Information Management and Technology (ICIMTech)*, IEEE, pp.187–191.

Rabiha, S.G., Kurniawan, A., Moniaga, J., Wahyudi, D.I. and Wilson, E. (2018) 'Face detection and recognition based e-learning for students authentication: study literature review', in *2018 International Conference on Information Management and Technology (ICIMTech)*, IEEE, pp.472–476.

Rusdi, J.F., Kodong, F.R., Indrajit, R.E., Sofyan, H. and Marco, R. (2020) 'Student attendance using face recognition technology', in *2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS)*, IEEE, pp.1–4.

Sagonas, C., Tzimiropoulos, G., Zafeiriou, S. and Pantic, M. (2013) '300 faces in-the-wild challenge: The first facial landmark localization challenge', in *2013 IEEE International Conference on Computer Vision Workshops*, IEEE, pp.397–403.

Sharma, P., Joshi, S.P., Gautam, S., Filipe, V. and Reis, M.C. (2019) *Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning*, ArXiv, abs/1909.12913.

Son, N.T., Chi, L.P., Lam, P.T. and van Dinh, T. (2019) 'Combination of facial recognition and interaction with academic portal in automatic attendance system', *in Proceedings of the 2019 8th International Conference on Software and Computer Applications*, pp.299–305.

Statistik Pendidikan Tinggi (2019) [online] https://www.mohe.gov.my/muat-turun/statistik/stat-2019/235-statistik-pt-2019-04-bab-2-universiti-awam/file (accessed August 2021).

Stratistics Market Research Consulting Pvt. Ltd. (2019) [online] https://www.researchandmarkets.com/reports/4827704/corporate-e-learning-global-market-outlook?utm_source=BW&utm_medium=PressRelease&utm_code=rld2dh&utm_campaign=1303824+-+Global+Corporate+ELearning+Market+Value+Nears+%2450+Billion+by+2026&utm_exec=joca220prd (accessed August 2021).

Su, Y., Cui, X., Xi, Q., Zhang, R. and Shen, J. (2020) 'An industry-ready single PTZ-camera based attendance management system', in *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, IEEE, pp.1–4.

Sun, C., Xia, F., Wang, Y., Liu, Y., Qian, W. and Zhou, A. (2018) 'A deep learning model for automatic evaluation of academic engagement', in *Proceedings of the 5th Annual ACM Conference on Learning at Scale*, pp.1–4.

Tamimi, A.A., AL-Allaf, O.N., Alia, M.A.J. (2015) 'Real-time group face-detection for an intelligent class-attendance system', *I. J. o. I. T. and Science, C.,* Vol. 7, No. 6, pp.66–73.

Tantak, A.S.A., Kale, A., Mehetre, R., Pophale, S.S. (2017) 'Face recognition for e-attendance for student and staff', *IOSR Journal of Computer Engineering (IOSR-JCE),* Vol. 19, No. 2, pp.89–94.

Valera, J., Valera, J. and Gelogo, Y. (2015) 'A review on facial recognition for online learning authentication', in *2015 8th International Conference on Bio-Science and Bio-Technology (BSBT)*, IEEE, pp.16–19.

Viola, P. and Jones, M.J. (2004) 'Robust real-time face detection', *J. I. J.O. C. V.*, Vol. 57, No. 2, pp.137–154.

Wang, W., Xu, K., Niu, H. and Miao, X. (2020) 'Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation', *Complexity*, Article ID 4065207, 9pp, https://doi.org/10.1155/2020/4065207.

Wang, Y., Kotha, A., Hong, P-H. and Qiu, M. (2020) 'Automated student engagement monitoring and evaluation during learning in the wild', in *2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, IEEE, pp.270–275.

Winarno, E., Hadikurniawati, W., Al Amin, I.H. and Sukur, M. (2017) 'Anti-cheating presence system based on 3WPCA-dual vision face recognition', in *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, IEEE, pp.1–5.