ELINVO (Electronics, Informatics, and Vocational Education), Mei 2022; Vol 7(1):48-62
ISSN 2580-6424 (printed), ISSN 2477-2399 (online,) DOI: https://doi.org/10.21831/elinvo.v7i1.51354

A Brief Study of The Use of Pattern Recognition in Online Learning: Recommendation for Assessing Teaching Skills Automatically Online Based

Pipit Utami[1*], Rudy Hartanto[2], Indah Soesanti[3]
[1,2,3] Department of Electrical Engineering and Information Technology, Gadjah Mada University
[*]E-mail: pipitutami@mail.ugm.ac.id

ABSTRACT

Online learning has become a trend for the current generation of students who have been exposed to advanced information and communication technology. Smart education can use pattern recognition. Manual assessments are subjective and inconsistent. To overcome these problems, pattern recognition can be used in the non-verbal aspect assessment system. This study describes pattern recognition in online learning about the functions, modalities, and algorithms and specifically related to the recognition system of non-verbal aspects of teaching skills. The literature study was carried out through the stages of planning, selection, extraction, and selection. There are 86 articles reviewed. The first result is the functions of implementing pattern recognition in online learning are engagement recognition, attention detection, emotion recognition, learning behavior, learning activity recognition, authentication, teaching training, etc. using four classifications of modality: visual, audio, biosignal, behavioral, and CNN as the most widely used learning algorithm. Secondly, all modalities (except behavioral) and CNN algorithm can be used for assessing teaching skills. Early development of the non-verbal aspect assessment system can use Facial Expression Recognition (FER) and Hand Gesture Recognition (HGR). The future analysis needs to focus on technology characteristics, the meaningfulness of the content, and the proper teaching mode. In the end, hopefully, prospective teachers will acquire technology that can make it easier for them to practice teaching and get objective assessments.

Keywords*:* pattern recognition, FER, online learning, assessing teaching skills

INTRODUCTION

The use of technology related to the Internet and interconnected devices deliver the Industrial Revolution (IR) 4.0 initiated by Germany and Society 5.0 by Japan [1]. Interconnecting in the smart manufacturing system is the core of IR 4.0. Society 5.0 integrates various advances and innovations from the 4.0 industrial revolution to meet multiple social needs to live entirely comfortably [2]. The spirit of IR 4.0 and Society 5.0 triggers the birth of Education 4.0 (Edu 4.0). On the other hand, educational institutions need to provide students with skills required in working life when they work both as entrepreneurs and workers. Technology skill is one of the learning experience that schools need to pay attention to help students face the future world of work [3]. One challenge that arises in Edu 4.0 is applying technology and capabilities to face the implications of using these technologies. Schools must provide space for students to be trained in technology and include values related to these technologies. The essential step is to get used to technology to help make it easier for students to learn. Embracing technology in education 4.0 is a must to achieve those experiences. Furthermore, Figure 1 shows the relationship between role in education with society development [4], [5], industrial revolution (IR) [1], web development [6], technology in education [7]–[10], and generations [11].

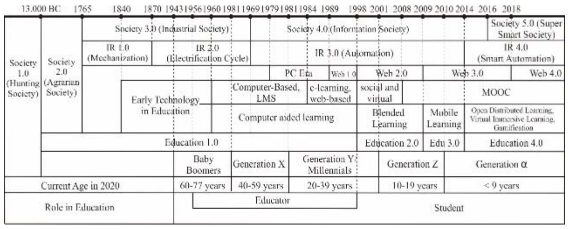


Figure 1. The relationship between the terminology of society, industrial revolution, web, education, and generation related to the role in education

Figure 1 shows that the Society started from the Hunting Society and continued with the Agrarian Society. Society 3.0 and IR 1.0 came together. IR 3.0 and Society 4.0 encourages the development of Website technology. Education 1.0 is a period of transmission of learning with a teacher-centered approach. Early technologies used in education are radio, projector, TV, headphones, etc. Computer equipment development started computer-based learning during the Education 1.0 era, which was also web-based with the emergence of the Learning Management System (LMS). Education 2.0 is social and virtual based on web 2.0 and applies blended learning. Massive Open Online Course (MOOC) is a transition to implementing Education 3.0. Technological developments in Education 4.0 are related to Smart Education.

The differences about teaching mode and learning theory in each Education era are: (1) Education 1.0 uses pedagogy with instructivism theory and behaviorism; (2) Education 2.0 uses Andragogy with Cognitivism theory; (3) Education 3.0 uses Heutagogy with Constructivism theory; and (4) Education 4.0 is Peeragogy and Cybergogy with Connectivism theory [8]. Several generations have grown in line with technological developments. Working-age and educational attainment is between the ages of 15 and 64 years [12]. This age as a generation boundary for only Baby Boomers. Baby Boomers, Millenials, and Generation Y are the age of educators. Simultaneously, students have a larger generation span from generation α to Generation X. Each generation has a different exposure to technology. Educational technology needs to pay attention to the characteristics of students. Education 4.0 as a response to IR 4.0 and Society 5.0 requires educators 4.0 to provide learning facilities that pay attention to student characteristics.

Smart education is the spirit of Education 4.0. One form of smart education is the application of Artificial Intelligence, such as Pattern Recognition. Pattern recognition is a system that facilitates detection, tracking, classification, recognition of specific modality by a computer system with a particular learning algorithm. There are various pattern recognition learning models that have been implemented in online learning, using both machine learning and deep learning approaches. This recognition uses multiple modalities that have many functions. This paper describes Pattern Recognition In Online Learning (PROL) by classifying the functions, modalities, and algorithms used. This can indicate several things: (1) opportunities to use certain modality and algorithms for other PROL functions; and (2) comparison of certain modality and algorithms for certain PROL functions. Furthermore, this study specifically points out opportunities for applying pattern recognition in assessing teaching skills relate to modalities and algorithms used. Pattern recognition can be applied in online learning, especially teaching training. Teaching practice is one of the skills that teachers must master. Therefore, prospective teachers must learn these skills.

In this case, prospective teachers are students in undergraduate study programs who receive micro-teaching in semester 7 at the age of about 21 years. This age shows that prospective teachers are part of generation Y. And in the next few years, prospective teachers are Generation Z. Both generations have interacted with web technology (web 1.0 to web 4.0 or today). Therefore, educational technology development for this generation must be updated to follow students' technological exposure characteristics. The teacher's socio-emotional competence is important to create a positive learning environment [13]. Non-verbal aspects function as emotional communication [14], [15]. In teaching, 82% are those aspects of communication [16]. Teaching practice performance was assessed manually using a scored questionnaire [17], [18]. The manual assessment presents subjectivity and inconsistency of assessment results [19]. Education 4.0 is related to an automated teaching practice assessment system that can provide objective and standardized assessments, especially regarding non-verbal aspects as emotional communication. There are various modalities and algorithms used in pattern recognition. It is necessary to study what modalities are most appropriate to describe non-verbal aspects of teaching skills and alternative algorithms that can be used. This article aims to explain the pattern recognition functions, modalities, and algorithms used in online learning; and describes the modality and algorithms in pattern recognition required for assessing teaching skills.

METHODS

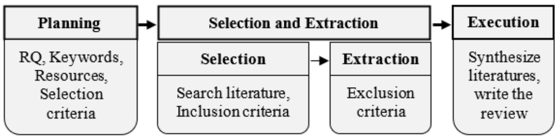


Figure 2. SLR method in reviewing articles

Figure 2 shows the literature review method in this paper adapted to the SLR method, consist of planning, selection, extraction, and execution [20]. The previous survey was about Artificial Intelligence in Education, which describes the application [21]. Unlike that review, this study describes the pattern recognition functions, modalities, and algorithms used in general online learning and specifically related to the non-verbal aspects of communication in a teaching skills assessment.

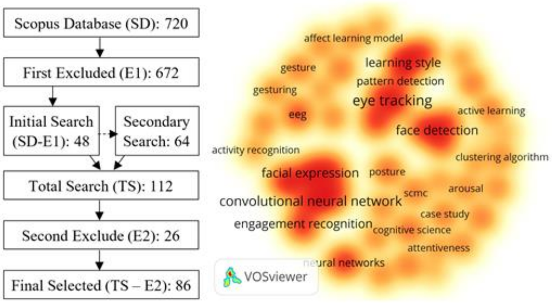
A. Planning Stages

In planning stages identifies research questions, keywords, resources to be searched, and selection article criteria in the planning step. By providing keywords, it can emphasize the topic.By giving a specific year span to ensure the novelty of the article. Selection of paper in English to ensure the readability of the study. The research questions in this study are: RQ1 - What are the pattern recognition functions, modalities, and algorithms used in online learning?; RQ2 - What are the modality and algorithms in pattern recognition required for assessing teaching skills? Research Keywords in this research using the key phrase "Pattern Recognition in Online Learning". The synonym of the keyword “pattern recognition” is "recognition system". The synonym of the keyword “e-learning” such as Electronic learning, MOOC, online learning, open course learning, and distance learning. Search string used in this review: (“e-learning” OR “electronic learning” OR “online learning” OR “learning online” OR “virtual learning” OR “e-education” OR “open course learning” OR “distance learning” OR “MOOC”) AND (“pattern recognition” OR “system recognition”).

B. Selection and Extraction Stages

The author carried out the selection stage by searching for literature in the Scopus database (www.scopus.com) and Google search (www.google.com) from August to October 2020. From these tools, there are articles from various publishers to produce comprehensive reviews. At the selection stage, inclusion criteria were applied. The inclusion criteria in this study are articles that discuss the application of PROL. The reduction of literature was carried out based on the exclusion criteria at the extraction stage. This stage ensures the article's relevancy related to the application of algorithms, modality, and the function description of PROL. Two stages of extraction in this study are by reading the abstract and full article. Those extractions to get an insight into the implementation of pattern recognition are suitable for teaching skills assessment.

First exclusion criteria in this study consisted of: (1) not meeting the inclusion criteria; (2) articles before 2015; (3) not a primary source; and (4) the item is not in English. These criteria are obtained by reading the title or abstract. The second exclusion criteria were: (1) unable to get the full article version; (2) does not apply the pattern recognition model/ method/ technique/ architecture/ algorithm; (3) does not inform the modality used; (4) do not provide information the modality function used; (5) non-online educational implementation. Figure 3(a) shows the selection and extraction process for literature in this literature review.



(a) (b)

Figure 3. (a) Selection and Extraction Process; (b) Density Visualization

The initial article recovery from the Scopus database were 720 literatures. The first exclusion reduced 672 study with the following details: (1) 224 unrelated topics; (2) publications before 2015 were 414; (3) apart from journal articles and proceedings totaling 23; and (4) the use of non-English languages as much as 11. The keyword online learning is why the large number of articles that are not related to the topic. These keywords lead to the items associated with specific programming approach algorithms that are not related to this study's case. The papers obtained in the first selection stage were 48, and the secondary search (google search) was 64. The number in secondary search is more than one. The keywords are less specific when searching directly in the database, for example, "Facial Expressions Recognition" AND "MOOC". The total articles from the two search techniques were 112, and by investigating all of the papers entered in the Scopus database, those papers have quality guaranteed. The second extraction contained 26 articles. From the selection and extraction process, 86 papers were obtained.

C. Execution Stages

The last stage in this SLR method is executed by synthesizing the data and writing the article. This study use VOSviewer [22], which reads data from bibliographic database files (Scopus). The display used is density visualization. The brighter the red color and the larger the text size, the more intense the words appear. From Figure 3(b), it can be seen that the most common variables appear in the following four aspects: (1) the function of pattern recognition is engagement recognition; (2) pattern recognition is "eye-tracking". Besides, if conduct a rotate analysis, "facial emotion recognition" or "facial expression recognition" is performed, it is also large and visible; (3) modalities are “eyes” and “facial expressions”; and (4) the algorithm is Convolutional Neural Network (CNN). However, these results still require further study by mapping modalities category (Figure 4), and algorithm classification (Figure 5).

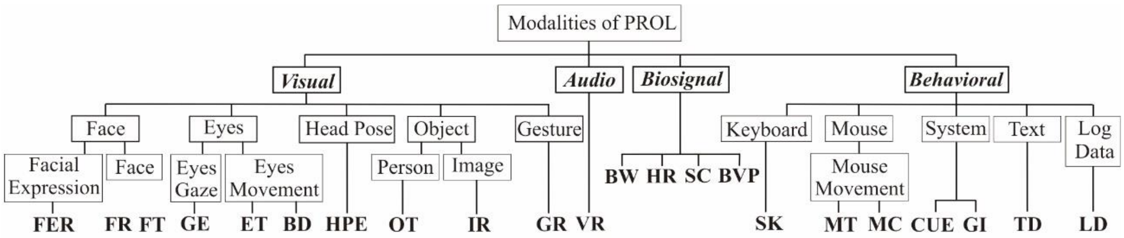
RESULT AND DISCUSSION

A. Pattern Recognition In Online Learning

1. Modalities and Algorithms Taxonomies for PROL

There are four classifications of modality, namely visual, audio, biosignal, and behavioral. Figure 4 shows the taxonomy of pattern recognition modalities in online learning obtained from the articles reviewed in this paper. Visual modalities use eyes, faces, head pose, hands gestures, and human activities images captured using camera sensors. Pattern recognition with visual modalities such as FR, FT, FER, GE, ET, BD, HPE, OT, IR, GR. Eye-tracking indicators such as pupil diameter, blink (count, rate), regression (count, rate, length), eye movement, etc. This eye movement such as fixation (count, rate, position and duration) and saccade (count, rate, amplitude, duration). Eyes movements pattern estimate learners' cognitive state [23]. Audio modalities use modality in a voice, which is captured using a microphone sensor. Pattern recognition with audio modality is VR. The main principle of voice recognition or speech detection is the comparison between the input signal with the threshold of certain voice/speech characteristics.

Biosignal modalities use modality with sensors that detect physiological signals in the user. The use of biosignal modalities requires supporting devices, such as brain waves (BW) with EEG; heart rate (HR) with ECG or heart flux (HF); skin conductance (SC) can use galvanic skin response (GSR); blood volume pressure (BVP) with photoplethysmograph (PPG). SC is a part of electrodermal activity. The rising cognitive workload and emotional arousal load can stimulate the brain to send signals to the skin's surface so that the sweat state increase. This affects the electrical changes on the surface of the skin. Behavioral modalities consist of user behavior when using the mouse (MT, MC), interaction with the system (CUE, GI), log data (LD), including data on the system in text (TD) and keyboard (SK). SK using features such as durations and latency, while GI means clicking the GUI menus [24]. Log datas store student learning activities related to time usage, work on assignments, searching and learning material, writing comments, etc.



Note:**FR**: Face Recognition; **FT**: Face Tracking; **FER**: Facial Expression Recognition; **GE**: Gaze Estimation; **ET**: Eyes Tracking; **BD**: Blink Detection; **HPE**: Head Pose Estimation; **OT**: Object Tracking; **IR**: Image Recognition; **GR**: Gesture Recognition; **VR**: Voice Recognition **BW**: Bain Waves; **HR**: HeartRate **SC**: Skin Conductance; **BVP**: Blood Volume Pressure; **MT**: Mouse Tracking; **MC**:Mouse Click-event; **DL**: DataLog **TD**: Text Detection; **SK**: Specified Keypress/Keystroke; **CUE**: Corresponding UNIX Epoch; **GI**: GUI Interactions

Figure 4. Modalities Taxonomy of PROL

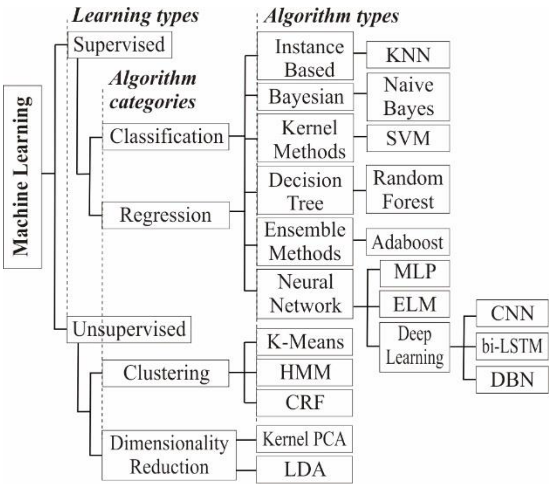


Figure 5. Algorithms Taxonomy of PROL

The applications of machine learning for PROL use both supervised and unsupervised learning types. Algorithm types that use PROL can be seen in Figure 5. This figure displays the machine learning algorithm used for PROL in a taxonomic format. The application of the pattern recognition algorithm type in online learning is adjusted to the characteristics of research needs in the form of a related environment, modality characteristics, the type of equipment used (low or high cost, poor or advanced technology, easy or difficult to assess, etc.), mastery of skills owned programming and so on. In general, the algorithm used for visual-based modalities for deep learning is CNN, while for machine learning, it is SVM. CNN can be used for the four types of modalities in PROL.

2. The Functions of PROL

During the recognition system, various learning settings are conducted, such as taking test activities, studying material, watching videos, etc. The application of the function of PROL varies widely. Some of these functions include engagement recognition, attention detection, emotion recognition, authentication, learning activity, etc. Furthermore, these functions are implemented for a more general purpose.

Table 1. Pattern Recognition of Authentication

Source	Pattern Recognition	Algorithm (accuracy)
[25]	FR	Feature Matching
[26]	FR	LDA (93%)
[27]	FR	CNN (99.28%)
[28]	FR	OpenFace Algorithm (natural lighting: 76.64%, artificial lighting: 80.4%)
[29]	SK, TD, FR, VR	Database matching
[30]	VR	NN: MLP (76.33%)
[24]	SK, MT, GI	BayesNet (99.39%); LibSVM (96.66%); Decision tree: J48 (86.64%)

Table 1 shows the Pattern Recognition for authentication, such as FR, SK, TD, VR, MT, GI. The algorithms used for the recognition are CNN, feature matching, LDA, OpenFace, database matching, NN. The unique modality as an authentication security key consists of modalities that are known to the user in the form of a password, modalities that are owned by the user which can be in the form of cards and a recognized face is one of the modalities that the user has (biometrics) [31]. This authentication is used to login to the e-learning system [25], [26]; to identify cheating in the online exams through the detection of student's faces [27]–[29]. Ensuring the e-assessment is done by registered students and the assignment is free from plagiarism multi-modality is used. These modalities are SK to recognize typing rhythm, TD to recognize writing style and check plagiarism, FR to recognize faces, VR to recognize voices during the exam [29]. As well as authentication, VR can be used to identify the exactness of student answers [30]. The fundamental difference regarding identification/ recognition is related to the number of class categories. Identification using multiclass, while authentication using binary class [24].

Table 2. Pattern Recognition of Engagement Recognition

Source	Pattern Recognition	Algorithm/ software/ devices/ sensors (accuracy)	Dataset/ State
[32]	FER	CNN (92.33%)	not/normally/highly-engaged
[33]	FER	DERN: the fusion of attention mechanism on BiLSTM and temporal convolution (60%)	DAiSEE
[34]	FER	CNN: C3D-FL (56.2%)	DAiSEE
[35]	FER	CNN: WSRGB-I3D Convnet (52.35%)	DAiSEE
[36]	FER	LDP-KPCA-DBN (two level: 90.89%; three level:87.25%)	two level: not/engaged; three level: not/normally/very-engaged
[37]	ET, BW, MT, CUE	ET: Gazepoint GP3; BW: OpenBCI Cyton Biosensing; MT, CUE: AutoHotKey	correct response
[38]	LD	K-means+FCA	minimal, regular, irregular

Table 3. Pattern Recognition of Attention Recognition

Source	Pattern Recognition	Algorithm/ software/ devices/ sensors (accuracy)	Dataset/ State
[39]	FR	CNN	Unidentified
[40]	FER, ET, SC	FER: OpenFace Software; ET: iView X Eye Tracker; SC: Shimmer3 GSR	Unidentified
[41]	ET	Eye tracker	Unidentified
[42]	ET	Eye tracker	monotone, non-monotone, and distraction
[43]	ET	Tobii X2-30 eye tracker (68.2%)	attention or inattention
[44]	GE	TabletGaze algorithm	Look or did not look the subtitles below and wandering
[45]	GE, VR	GE: geometric pattern matching; VR: pitch-detection algorithm	attentive or inattentive (not looking or speaking)
[46]	ET	Tobii eyetracker	Unidentified
[47]	BW	CFS-KNN algorithm (80.84%)	High, Neutral, Low
[48]	BW, HR	SVM (ECG:77,68%; HF: 59.64%; EEG: 86.86%)	Attention, Non-attention
[49]	MT	R: library mousetrap	focused, distracted, doing something else
[50]	MT	CRF (88.23%)	attention, ignored
[51]	FER	CNN (76.23%), LSTM (71.36%)	Attention: Happy, Angry, Surprise, Neutral; Non-attention: Sad, Fear, Disgust
[52]	GE, HPE,	GE: entropy-based segmentation; HPE: POSIT algorithms;	On, Off

Table 2 shows pattern recognition for engagement recognition. This table indicates that the engagement states vary. DAiSEE consist of boredom, confusion, engagement, and depression affective state. Patterns recognition for detecting engagement are FER, ET, BW, MT, CUE, LD. CNN, BiLSTM, DBN, K-means are algorithms that can be used to recognize the engagement status. Engagement recognition can be used to assist smart adaptive online learning [32]; to evaluate of studying experience [33]; to monitor when students viewing material on online learning [34]; to adjust learning outcome and reduce school dropout [35]; to give individual intervenes [36]; to evaluate e-learning [37]; to recognize teaching pedagogy [38].

Table 3 shows pattern recognition for attention recognition. Categorizing attention state based on the needs in the learning implementation. Patterns recognition for detecting attention are FR, FER, ET, SC, GE, VR, BW, HR, HPE. Algorithms applied for recognizing attention that are CNN, SVM, KNN, LSTM, etc. The individual cognitive approach can be indicated from patterns of visual attention. Attention recognition has many variations of implementation, such as: to enhance dynamic teaching [39]; related to the correctness of quiz answers [40]; to find usability problems in the system [41]; to enhance virtual learning (interaction and feedback) [42]; encourage self-regulated learning [43]; to place the subtitles in the video that impact learning outcomes [44]; to develop smart and flexible systems by monitoring learner learning activities [45]; as reflection by observing the previous online learning strategy [46]; for learner feedback [47]; to monitor cognitive activity [48]; to comprehend the self-produced cognition complexity [49]; to identify user tasks in the online system [50]. In addition, attention detection can be utilized to assess teaching strategy [51] and the combination of attention, arousal, and valence are utilized to determine the affective state used to adjust to personal learner requirements [52]. The two indicates that emotion recognition can be used for attention identification.

Table 4. Pattern Recognition of Emotion Recognition

Source	Pattern Recognition	Algorithm/ software/ devices/ sensors (accuracy)	Dataset/ State
[52]	FER, SC	FER: SVM; SC: wireless SC sensor	valence (positive and negative); arousal (high and low)
[53]	FER	CNN	neutral, happy, anger, sad
[54]	FER	CNN (64.77%)	FER2013
[55]	FER	PCA (71.05%)	Delighted, Confused, Concentrated, Distracted, Surprised, Thinking, Normal, Unknown, Note-taking
[56]	FER	EM-DeSTIN	Neutral, Smile, Anger, Scream
[57]	FER	CNN	CK+, JAFFE, NVIE
[58]	FER	LBP-PCA-SVM (80.79%)	JAFFE
[59]	FER, BD	FER: CNN; BD: SVM	FER2013: positive (happiness, surprise), neutral (peace), negative (anger, boredom, fear, sadness); BD: closing, opening eyes
[60]	FER	Face detection: Haar cascade	JAFFE
[61]	FER	CNN: VGG-RDA-ADA (91.6%)	confusion, distraction, enjoyment, neutral, fatigue
[62]	FER	SVM (92%)	happiness, interest, surprise, curiosity, concentration, attention, disappointment, bored, perplexity, disgust, frustration
[63]	FER	CNN (86%)	FER2013
[64]	VR	ESR Algorithm	positive, negative
[65]	VR	LSTM (85.32%)	angry, sad, happy, neutral, surprise
[66]	TD	DEI-TM	positive, negative
[67]	TD	A-CNN (89%), LSTM-ATT (71%)	boredom, shame, enjoyment, hope, disappointment, anger, anxiety, joy, relaxation
[68]	SC, BVP, HR	SVM (86.5%)	Engagement, Confusion, Frustration, Hopefulness
[69]	FER	Microsoft facial analyzer	happiness, surprise, anger, contempt, disgust, sadness, fear, neutral
[70]	FER	Emotient software	key performance indicators (sentiment, emotional engagement, attention); valence (negative, neutral, positive); emotion (joy, anger, fear, sadness)

Table 4 shows pattern recognition for emotion recognition. Patterns recognition for classifying emotions are FER, LD, TD, SC, BD, VR, BVP, HR. Algorithms applied for recognizing attention that are CNN, SVM, LSTM, etc. Academic emotion evaluation in online learning for boosting learner wellbeing [67]. Emotion recognition can be used for providing a personalized support as feedback to learners [54]–[57], [63], [64]; for enhancing efficient training [58]; for enhancing outcome learning and instruction quality [59]; for identifying learner comprehension [60]; for enhancing quality of teaching [65]; for extracting information from comment in online course learning [66]; for providing adaptable e-learning [68]; for presenting adaptive course content [69]; for supporting teacher training [70]. FER2013, CK+ and JAFFE dataset consist of six basic emotions (Angry, Disgust, Fear, Happy, Sad, Surprise), and Neutral. NVIE consist of six basic emotion.

Table 5. Pattern Recognition of Activity Recognition

Source	Pattern Recognition	Algorithm/ software/ devices/ sensors (accuracy)	State/ Indicator
[71]	FT, ET	SVM (98.5%); KNN (97.3%)	Reading words on the screen (concentrated); Looking the other with face movement (distracted); Thinking deeply with closing the vdozing); Sitting distant from the screen (doing something other than studying.).
[72]	ET	Tobii TX300 eye tracker	unconscious, conscious, strategic
[73]	ET	EyeTribe+Ogama Software	faster, slower
[74]	FT	AdaBoost+Cascade, CamShift (100%)	there is face, there is no face
[23]	ET	HMM (Idle 95.33%, Seeking 80.29%, Scanning 81.32%)	Idle, Seeking, Scanning
[75]	ET	EyeLink II; MJTL (95.09%)	good, poor
[76]	GE	Tobii TX300; pooled evidence algorithm	scanning, skimming, reading, re-reading
[77]	ET	Eye Tribe, Visualization software (Tableau Desktop 9.0)	the reader of (slow linear, fast linear, topic structure)
[53]	HPE	CNN	presence in class, facing the classroom, look left, look right, look center, raise right hand, raise left hand, raise two hands
[78]	OT	CalcOpticalFlowSF	-
[79]	FT	ViolaJones+Template Matching (93.3%)	right/ wrong/ not detected
[52]	LD	Evaluation process	positive, negative (answering duration, attempts quantity to answer, right/wrong answers quantity, the last score)
[80]	LD	KNN	active, reflective, sensing, intuitive, visual, verbal, sequential, global
[81]	LD	K-Means, PCA, Gaussian Naive Bayesian (Learning style prediction: 94%)	re-access, visiting time and accessing resources
[82]	LD	K-Means	perceiving, extroversion, judging, introversion, intuition, sensing, feeling, thinking

Table 5 shows the pattern recognition of activity recognition. Thus, activity recognition includes several activities, including body movement activities in front of the screen (face, head, hands, eyes), distance, learning activities in the online system (time, visits, order navigation and others). Patterns recognition for activity recognition are FER, LD, TD, SC, BD, VR, BVP, HR. Algorithms applied for recognizing activity are CNN, SVM, LSTM, etc.

Activity recognition can be used to improve the lecture quality [71]; to follow the movements of presenter [78], [79]; to outline learners' reading behavior and studying from explanatory essay, also to investigate the interaction between topic interest and learners' cognitive processing [72]; to analyze student characteristics (expert or novice) for e-learning improvement recommendation [73]; to improve learning [74]; to evaluate online learning [23]; to evaluate learners' study conditions and strengthen the interaction between teacher and learner [53]; to find out learner' learning style [80]–[82]; to analyze learner behavior of re-access the system [81]; to identify the progress of learning so that can be used to analyze feedback for student [52]. Recognizing learners' reading ability can be used to give a personalized online learning service [75]. To enhance learning material [76]. to analyze the needs for online learning instructional development that stimulating higher-order cognitive activities [77]. adaptive classroom about material and teaching strategy [83]. smart classroom with teaching interactivity [84] improving learning outcome[85].

Other Implementation of Pattern Recognition (FER) in online learning is intrinsic motivation recognition. That function can be used to analyze how to upgrade the learning process and outcomes [86]. CNN was used in this implementation for indicating high or low motivation.

3. Recommendation for the future research of PROL

The application of PROL in the future is the use of keywords that are specific to the literature review. Use of search more than once in the article database if it consists of several modalities or algorithms. This is to get articles that are contextual and comprehensive. To get additional information, you can look for the use of pattern recognition with certain modalities and algorithms outside of its application in online learning. Determination of recognition needs is adjusted to the problem. This study does not discuss the devices used in the recognition system, both in function, type and number. The discussion of each modality and algorithm is not discussed in detail one by one.

Several things need to be adjusted to the needs of system development, namely: (1) determining the status of the recognition of engagement, attention and emotion; (2) determination of pattern recognition and modality; (3) algorithm selection; (4) the use of supporting devices, especially for data acquisition; and (5) fulfillment of certain conditions for participants. Further studies need to focus on one particular specific topic, such as pattern recognition functions, modalities, algorithms or devices for data acquisition. It is hoped that future research can be applied and tested under actual conditions (not limited to laboratory-based). Generalization of lab findings is difficult because of the Hawthorne effect [84]. Collecting data in the lab with participants realizing that they are being watched results in unnatural activity. Data collection involving people needs to include a consent form from the participants. Data collection in the lab needs to describe the data collection procedures carried out, thus informing the existing limitations. To make comparisons, it is necessary to pay attention to the similarity of certain aspects as controlling factors.

B. Pattern Recognition for Assessing Teaching Skills

The future teacher and the novice need to mastery the teaching skills, so that teaching training is highly important [87]. Online teaching training has high urgency. Such experience requires an automatic assessment of teaching skills and providing feedback. The online training increases the number and quality of teaching experience for prospective teachers and novice teachers. The training system in question focuses on the assessment aspect. The advantage of an automated teaching training system is that it provides a consistent feedback element [19], [88]. Moreover, during the Covid-19 pandemic, this has further strengthened the importance of an online-based assessment automation system, in this case, for teaching training. Automated based systems can be an alternative to manual assessments to evaluate distance learning learners [89]. e-Assessment is a vital aspect of self-testing and formative assessment [90].

Communication skills play an essential role in teacher success. Mastery of communication skills as one of the 21st century skills needs to be integrated into learning to ensure graduate work readiness [91]. Communication is one of the assessment aspects of teaching skills [92]. Recognition systems need to be developed in online teaching practices to recognize non-verbal elements of communication, such as gestures and facial expressions [87]. Affective recognition systems, especially FER, need to be developed in online learning. Teachers' facial expressions have a positive effect on student achievement [93]. Therefore emotions are very relevant in online learning for both teachers and learners [94]. Deeper, the innovation of online exam activities offered by researchers is that there is an assessment of presentation skills, especially non-verbal. Student assessment of communication skills can be supported by a system of recognition. The achievement of communication skills needs to be known especially through assessment. Experiences of learners are the key to the development of MOOCs [66]. Teaching experience and feedback on the teaching skills assessment results are the primary keys to developing online-based learning for teaching skills.

1. *Patterns Recognition Function for Asessing Teaching Skills*

Reverse engineering thinking is applied in carrying out AATS development needs analysis. Characteristics of system users are users who practice teaching skills by displaying positive emotional, behavioral activities with good and appropriate attention and motivation. From Figure 6, it can be seen that the function of implementing PROL that is relevant for AATS is Teaching Training (FER and GR), Emotion Recognition (FER and VR), Learning Behavior (GR, GE, ET, HR, SC, BW), Learner Motivation (GE, FER), Affective Recognition (GE, BVP, SC), Learner's attention (VR, HPE, FER, BW, HR, SC, ET, GE), Learner activity recognition (ET, FT, OT). Thus, AATS can apply modalities to visual, audio, and biosignal aspects.

Teaching training is related to teaching emotions (and or things related to affective) and behaviors or activities that occur while practicing that experience. Implementing pattern recognition for AATS needs to pay attention to several technical aspects applied to users, such as ease, flexibility, and other technical matters. Several technical problems during the training, such as low-resolution image, occlusion, and lighting intensity [92]. The initial development of the AATS aims for a system capable of automatic assessment and feedback that students can use independently and flexibly.

Educator 4.0 shouldn't just talk about applying the latest technology. The teaching mode is not limited to pedagogy and andragogy, but rather a shift from heutagogy to cybergogy and peeragogy with an orientation of action learning and passion-based learning. AATS needs to facilitate students to be independent, flexible, and adapt to the abilities of students in practicing their teaching skills online. Students are able to self-direct their needs and learning time (heutagogy). The need for collaboration (peeragogy) and networking (cybergogy) is the limitation of the AATS's early development. The application of educational technology can be used as a learning medium as well as an assessment tool that can provide feedback. By conducting an in-depth analysis, the usage result of learning media will achieve the media's objectives to increase learning effectiveness.

AATS analysis needs to pay attention to the development objectives by taking into account the following aspects: (1) technological characteristics for an automatic assessment system and providing appropriate feedback; (2) the meaningfulness of the teaching training content to make sure the amount and quality of the teaching experience; and (3) the proper mode of teaching. The most suitable type of educational technology is obtained from an in-depth analysis of the most relevant aspects of technology, content, and teaching modes. This analysis can be done with TPACK [95] or TEPACK [96] or W2CPATK [97]. Furthermore, it is hoped that the development of a comprehensive AATS involving all related pattern recognition can be carried out in the future. This is to increase the number and quality of teaching skills using an automated assessment system with feedback based on online learning.

2. *Recommendation for Assessing Teaching Skills: FER and HGR using CNN*

Behavior modality can not be used because teaching activity does not control screen activity. Besides, the use of biosignal sensors is difficult and high cost. The delivery of messages through the visual component is greater than the voice and verbal aspects [98]. Therefore visual is chosen over the audio (voice) modality. The kinesic aspect is sufficient to identify teachers' non-verbal communication [99]. Kinesic aspect is body language. Teachers need to understand the proper gesture and facial expressions when teaching [100]. These show that facial expressions and hand gestures as part of visual kinesic modality can be used for early research in pattern recognition for assessing teaching skills. Figure 6 shows the FER process.

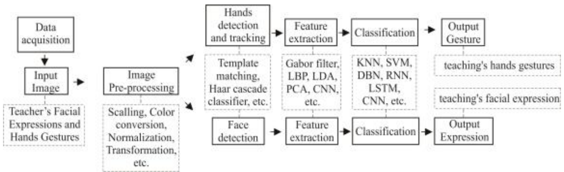


Figure 6. Framework of teaching's facial expressions and hands gestures recognition

Data acquisition needs to consider camera specifications according to system requirements. For example, the use of high or low cost can affect the type of image (RGB, Depth Image etc.). Another thing to pay attention to is the number of camera usage (single or multi). Preprocessing to enhance, noise removal, normalization and facilitate the next process. Segmentation related to lighting and background conditions (complex, dynamic). Gesture detection related to hand articulation. Gesture representation is related to static or dynamic gestures. Feature extraction to extract informative features for the classification stage. Classification needs to pay attention to the classifier. If supervised learning needs to pay attention to training and validation regarding computational, the number of parameters and gestures is not recognized for accuracy.

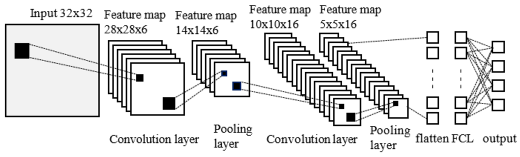


Figure 7. CNN algorithm structure

Figure 7 shows CNN algorithm structure. CNN architecture consists of an input layer, multiple hidden layers, and an output layer. Generally, this architecture has feature extraction and classification parts. There are a convolution layer and pooling layer in feature extraction, while in classification, there is a fully connected layer. Convolution produces a feature map to reduce parameters. The pooling layer summarizes information resulting from convolution to reduce dimensions. The pooling layer generally consists of average pooling and max pooling. The fully-connected layer is used for classifiers because it can define a particular matrix.

Table 6. FER in online learning

Source	Algorithm (accuracy)	Dataset/ State
[33]	DERN (60%)	boredom, confusion, engagement and depression
[34]	CNN: C3D-FL (56.2%)	boredom, confusion, engagement and depression
[35]	CNN: WSRGB-I3D Convnet (52.35%)	boredom, confusion, engagement and depression
[51]	CNN (76.23%), LSTM (71.36%)	Attention: Happy, Angry, Surprise, Neutral; Non-attention: Sad, Fear, Disgust
[54]	CNN (64.77%)	FER2013
[56]	EM-DeSTIN	Neutral, Smile, Anger, Scream
[57]	CNN	CK+, JAFFE, NVIE
[59]	FER: CNN	FER2013: positive (happiness, surprise), neutral (peace), negative (anger, boredom, fear, sadness)
[60]	Face detection: Haar cascade	JAFFE
[70], [101]	-	-
[86]	FER	63%
[61]	CNN: VGG-RDA-ADA (91.6%)	confusion, distraction, enjoyment, neutral, fatigue

Table 6 shows the use of FER and GR in online learning. CNN has become a widely used deep learning model. FER2013 is a facial expression dataset that can be used to train the CNN model for FER. The state types that become the classification output vary according to the needs of the FER application function. There are a variety of circumstances that can be identified from the use of GR (with CNN models), such as writing, reading, raising hand, concentrating, using telephone, looking around, standing, head down [83]; 17 types of teaching's static gesture [84]; and teaching's pointing gesture [85]. Researchers need to identify what kind of hand movements the teacher makes while teaching.

FER [70], [101], and GR [19], [88] as the implementation of pattern recognition in teaching training are recommended options for early development of AATS. FER is a modality that plays an essential role in the development of AATS [92]. Gesture recognition in teaching training is hand gestures, so in this case, HGR. Several types of facial expressions such as anger, fear, and sadness have low accuracy values [54]. The application of GR in teaching activities is largely related to its effectiveness in teaching. In the application of GR to determine the learner's behavior, high accuracy of 99.2862% uses the CNN architecture (GoogleNet) [53]. Further research needs to be carried out on a search that is more focused on FER and GR for teaching training with CNN. Moreover, researchers need to study the kinds of facial expressions and hand gestures relevant to teaching.

CONCLUSION

The most frequently used functions of implementing pattern recognition are engagement recognition, attention recognition, emotion recognition, activity recognition, authentication, and intrinsic motivation recognition. In this study there were four modalities of pattern recognition: visual, audio, biosignal, and behavioral, with the widely used are the eyes and facial expressions. The commonly used pattern recognition is eye tracking and facial expression recognition. The learning algorithm that is most widely used with the flexibility of modality coverage and the superiority of accuracy results from pattern recognition is CNN.

The functions of implementing pattern recognition for AATS include Teaching Training, Emotion Recognition, Learning Behavior, Learner Motivation, Affective Recognition, Learner's attention, and Learner activity recognition. Visual, audio, and biosignal aspects of modality can be used to develop a comprehensive AATS. However, several considerations related to ease, flexibility, and technical aspects of implementation by users need to be considered. AATS analysis needs to pay attention to the development objectives by taking into account the following features: technological characteristics for an automatic assessment system and providing appropriate feedback; the meaningfulness of the teaching training material content to support the amount and quality of the teaching experience; and the proper mode of teaching. The implementation of CNN for FER and GR in teaching training needs to be studied further. Besides, researchers need to study the types of facial expressions and hand movements relevant to teaching.

REFERENCES

[1] S. Phuyal, D. Bista, and R. Bista, "Challenges, Opportunities and Future Directions of Smart Manufacturing: A State of Art Review," *Sustain. Futur.*, vol. 2, p. 100023, 2020.

[2] M. Fukuyama, "Society 5.0: Aiming for a New Human-Centered Society," *Japan SPOTLIGHT*, vol. Special Ar, pp. 47–50, 2018.

[3] G. Elhussein, T. A. Leopold, and S. Zahidi, "Schools of the Future Defining New Models of Educ ation for the Fourth Industrial Revolution," Cologny/Geneva, 2020.

[4] M. E. Gladden, "Who Will Be the Members of Society 5.0? Towards an Anthropology of Technologically Posthumanized Future Societies," *Soc. Sci.*, vol. 8, no. 5, p. 148, May 2019.

[5] Keidanren, "Toward realization of the new economy and society." Japan, 2016.

[6] A. J. Balloni and P. H. de Souza Bermejo, "Governance, Sociotechnical Systems and Knowledge Society: Challenges and Reflections," in *International Conference on ENTERprise Information Systems*, 2010, pp. 42–51.

[7] M. Aparicio, F. Bacao, and T. Oliveira, "Trends in the e-Learning Ecosystem: A Bibliometric Study," in *AMCIS 2014 Proceedings*, 2014, pp. 1–11.

[8] S. Mokhtar, J. A. Q. Alshboul, and G. O. A. Shahin, "Towards Data-driven Education with Learning Analytics for Educator 4.0," in *Journal of Physics: Conference Series*, 2019, vol. 1339, p. 012079.

[9] M. Maria, F. Shahbodin, and N. C. Pee, "Malaysian higher education system towards industry 4.0 – Current trends overview," in *Proceedings of The 3rd International Conference on Applied Science And Technology (ICAST'18)*, 2018, p. 020081.

[10] S. Y. Tan, D. Al-Jumeily, J. Mustafina, A. Hussain, A. Broderick, and H. Forsyth, "Rethinking Oue Education to Face The New Industry Era," in *10th International Conference on Education and New Learning Technologies*, 2018, pp. 6562–6571.

[11] C. Sima, "Generations BB, X, Y, Z, α - the changing consumer in the hospitality industry," in *The Routledge Handbook of Hotel Chain Management*, M. Ivanova, S. Ivanov, and V. P. Magnini, Eds. Oxon: Routledge, 2016, pp. 471–479.

[12] M. Cristea, G. G. Noja, D. E. Dănăcică, and P. Ştefea, "Population ageing, labour productivity and economic welfare in the European Union," *Econ. Res. Istraživanja*, vol. 33, no. 1, pp. 1354–1376, Jan. 2020.

[13] M. Hen and M. Goroshit, "Social–emotional competencies among teachers: An examination of interrelationships," *Cogent Educ.*, vol. 3, no. 1, pp. 1–9, Feb. 2016.

[14] F. Bambaeeroo and N. Shokrpour, "The impact of the teachers' non-verbal communication on success in teaching," *J. Adv. Med. Educ. Prof.*, vol. 5, no. 2, pp. 51–59, 2017.

[15] R. W. Kaps and J. K. Voges, "Nonverbal Communications : A Commentary on Body Language in the Aviation Teaching Environment," *J. Aviat. Educ. Res.*, vol. 17, no. 1, pp. 43–52, 2007.

[16] Akinola, "The use of non – verbal communication in the teaching of english language," *J. Adv. Linguist.*, vol. 4, no. 3, pp. 428–433, 2014.

[17] H. Fawad and I. A. Manarvi, "Student feedback & systematic evaluation of teaching and its correlation to learning theories, Pedagogy & Teaching skills," in *Proceedings of 2014 IEEE International Conference of Teaching, Assessment and Learning (TALE)*, 2014, pp. 398–404.

[18] M. Sulaiman, Z. H. Ismail, A. A. Aziz, and A. Zaharim, "Lesson study: Assessing pre-service teacher's performance of teaching chemistry," in *2011 3rd International Congress on Engineering Education (ICEED)*, 2011, pp. 208–213.

[19] R. Barmaki and C. E. Hughes, "Embodiment analytics of practicing teachers in a virtual immersive environment," *J. Comput. Assist. Learn.*, vol. 34, no. 4, pp. 387–396, Aug. 2018.

[20] C. Okoli, "A Guide to Conducting a Standalone Systematic Literature Review," *Commun. Assoc. Inf. Syst. Vol.*, vol. 37, no. 43, pp. 879–910, 2017.

[21] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?," *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 1, p. 39, Dec. 2019.

[22] N. J. van Eck and L. Waltman, "VOSviewer Visualizing scientific landscapes." Universiteit Leiden and CWTS Meaningful metrics, 2020.

[23] J. Yi, B. Sheng, R. Shen, W. Lin, and E. Wu, "Real Time Learning Evaluation Based on Gaze Tracking," in *14th International Conference on Computer-Aided Design and Computer Graphics*, 2015, pp. 157–164.

[24] K. O. Bailey, J. S. Okolica, and G. L. Peterson, "User identification and authentication using multi-modal behavioral biometrics," *Comput. Secur.*, vol. 43, pp. 77–89, 2014.

[25] X. Song, S. Zhu, Y. Wei, Q. Sun, and L. Zhao, "Research on Information Engineering with Information Technology for E-Learning Based on Face Recognition," *Adv. Mater. Res.*, vol. 977, pp. 460–463, 2014.

[26] Z. Zainuddin and A. S. Laswi, "Implementation of The LDA Algorithm for Online Validation Based on Face Recognition," in *Journal of Physics: Conference Series. International Conference on Computing and Applied Informatics 2016*, 2017, pp. 1–6.

[27] R. Sivakumar and R. Sivakumar, "A Convolutional Neural Network based Fingerprint and Face Biometric Multi-modal for Educational Authentication System," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 3, pp. 3486–3497, 2020.

[28] O. Ruiz-Vivanco, A. Gonzalez-Eras, J. Cordero, and L. Barba-Guaman, "Monitoring for the Evaluation Process On-Line Prototype Based on OpenFace Algorithm," in *echnology Trends. CITT 2018. Communications in Computer and Information Science*, 2019, pp. 500–509.

[29] R. Tsankova, O. Marinov, M. Durcheva, and E. Varbanova, "Synergy Effect of the TeSLA Project in Management of Engineering Higher Education," in *Proceedings of the 9th International Conference on Management of Digital EcoSystems*, 2017, pp. 259–264.

[30] B. Akhmetov, I. Tereykovsky, A. Doszhanova, and L. Tereykovskaya, "Determination of Input Parameters of the Neural Network Model , Intended for Phoneme Recognition of a Voice Signal in the Systems of Distance Learning," *Int. J. Electron. Telecommun.*, vol. 64, no. 4, pp. 425–432, 2018.

[31] N. A. Mahadi, M. A. Mohamed, A. I. Mohamad, M. Makhtar, M. F. A. Kadir, and M. Mamat, "A Survey of Machine Learning Techniques for Behavioral-Based Biometric User Authentication," in *Recent Advances in Cryptography and Network Security*, InTech, 2018, pp. 43–59.

[32] M. Murshed, M. A. A. Dewan, F. Lin, and D. Wen, "Engagement Detection in e-Learning Environments using Convolutional Neural Networks," in *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, 2019, pp. 80–86.

[33] T. Huang, Y. Mei, H. Zhang, S. Liu, and H. Yang, "Fine-grained Engagement Recognition in Online Learning Environment," in *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, 2019, pp. 338–341.

[34] L. Geng, M. Xu, Z. Wei, and X. Zhou, "Learning Deep Spatiotemporal Feature for Engagement Recognition of Online Courses," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2019, pp. 442–447.

[35] H. Zhang, X. Xiao, T. Huang, S. Liu, Y. Xia, and J. Li, "An Novel End-to-end Network for Automatic Student Engagement Recognition," in *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, 2019, pp. 342–345.

[36] M. A. A. Dewan, F. Lin, D. Wen, M. Murshed, and Z. Uddin, "A Deep Learning Approach to Detecting Engagement of Online Learners," in *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, 2018, pp. 1895–1902.

[37] G. M. Notaro and S. G. Diamond, "Development and Demonstration of an Integrated EEG, Eye-Tracking, and Behavioral Data Acquisition System to Assess Online Learning," in *ICETC '18 Proceedings of the 10th International Conference on Education Technology and Computers*, 2018, pp. 105–111.

[38] L. Stanca, R. Lacurezeanu, and C. Felea, "Analysis for Visualizing Actors' Experiences in E-learning Spaces to Create an Active Learner's Profile," in *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2019, pp. 1882–1888.

[39] D. Canedo, A. Trifan, and A. J. R. Neves, "Monitoring Students' Attention in a Classroom Through Computer Vision," in *Highlights of Practical Applications of Agents, Multi-Agent Systems, and Complexity: The PAAMS Collection*, 2018, pp. 371–378.

[40] A. Edwards *et al.*, "Sensor-based Methodological Observations for Studying Online Learning," in *SmartLearn'17*, 2017, pp. 25–30.

[41] J. Wang, P. Antonenko, M. Celepkolu, Y. Jimenez, E. Fieldman, and A. Fieldman, "Exploring Relationships Between Eye Tracking and Traditional Usability Testing Data," *Int. J. Human–Computer Interact.*, vol. 35, no. 6, pp. 483–494, Apr. 2019.

[42] M. Fahimipirehgalin, F. Loch, and B. Vogel-Heuser, "Using Eye Tracking to Assess User Behavior in Virtual Training," in *Intelligent Human Systems Integration*, 2020, pp. 341–

347.
[43] T. Robal, Y. Zhao, C. Lofi, and C. Hauff, “Webcam-based Attention Tracking in Online Learning: A Feasibility Study,” in *IUI 2018*, 2018, pp. 189–197.
[44] H. Chen, M. Yan, S. Liu, and B. Jiang, “Gaze inspired subtitle position evaluation for MOOCs videos,” in *Proceedings of SPIE - The International Society for Optical Engineering*, 2017, vol. 10443, p. 1044318.
[45] Y. Uğurlu, “User Attention Analysis for E-learning Systems Using Gaze and Speech Information,” in *Information Science, Electronics and Electrical Engineering (ISEEE)*, 2014.
[46] U. Stickler and L. Shi, “Eyetracking methodology in SCMC : A tool for empowering learning and teaching,” *ReCALL*, vol. 29, no. 2, pp. 160–177, 2017.
[47] B. Hu, X. Li, S. Sun, and M. Ratcliffe, “Attention Recognition in EEG-Based Affective Learning Research Using CFS+KNN Algorithm,” *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 15, no. 1, pp. 38–45, Jan. 2018.
[48] A. Belle, R. Hobson, and K. Najarian, “A physiological signal processing system for optimal engagement and attention detection,” in *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, 2011, pp. 555–561.
[49] M. R. D. da Silva, M. Postma-Nilsenova, and F. Hermens, “Wandering Mice, Wandering Minds: Using Computer Mouse Tracking to Predict Mind Wandering,” in *22nd International Conference on Circuits, Systems, Communications and Computers (CSCC 2018)*, 2018.
[50] A. Elbahi and M. N. Omri, “Conditional Random Fields For Web User Task Recognition Based On Human Computer Interaction,” in *WSCG 2015. 23rd International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, 2015.
[51] N. Gerard *et al.*, “Detection of Subject Attention in an Active Environment Through Facial Expressions Using Deep Learning Techniques and Computer Vision,” in *Advances in Neuroergonomics and Cognitive Engineering. AHFE 2020. Advances in Intelligent Systems and Computing, vol 1201*, Switzerland: Springer, Cham, 2021, pp. 326–332.
[52] J. Chen, N. Luo, Y. Liu, L. Liu, K. Zhang, and J. Kolodziej, “A hybrid intelligence-aided approach to affect-sensitive e-learning,” *Computing*, vol. 98, no. 1, pp. 215–233, 2016.
[53] Y. Te Ku, H. Y. Yu, and Y. C. Chou, “A Classroom Atmosphere Management System for Analyzing Human Behaviors in Class Activities,” in *1st International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2019*, 2019, pp. 224–231.
[54] B. E. Zakka and H. Vadapalli, “Detecting Learning Affect in E-Learning Platform Using Facial Emotion Expression,” in *Proceedings of the 11th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2019)*, 2021, pp. 217–225.
[55] Y. Ling, *SmartPlayer : Inferring Learners ’ Emotions While They Are Watching Videos*. Springer Singapore, 2019.
[56] J. Xu, Z. Huang, M. Shi, and M. Jiang, “Emotion Detection in E-learning Using Expectation-Maximization Deep Spatial-Temporal Inference Network,” in *Advances in Computational Intelligence Systems. UKCI 2017. Advances in Intelligent Systems and Computing, vol 650*, 2018, pp. 245–252.
[57] A. Sun, Y.-J. Li, Y.-M. Huang, and Q. Li, “Using Facial Expression to Detect Emotion in E-learning System: A Deep Learning Method,” in *Emerging Technologies for Education. SETE 2017. Lecture Notes in Computer Science, vol 10676*, 2017, pp. 446–455.
[58] B. Xu, X. Li, H. Liang, and Y. Li, “Research on professional talent training technology based on multimedia remote image analysis,” *EURASIP J. Image Video Process.*, vol. 39, pp. 1–9, Dec. 2019.
[59] G. Li and Y. Wang, “Research on learner’ s emotion recognition for intelligent education system,” in *IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2018, pp. 754–758.
[60] D. Yang, A. Alsadoon, P. W. C. Prasad, A. K. Singh, and A. Elchouemi, “An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment,” in *Procedia Computer Science*, 2018, vol. 125, pp. 2–10.
[61] C. Bian, Y. Zhang, F. Yang, W. Bi, and W. Lu, “Spontaneous facial expression database for academic emotion inference in online learning,” *IET Comput. Vis.*, vol. 13, no. 3, pp. 329–337, 2019.
[62] F. D’Errico, M. Paciello, B. De Carolis, A. Vattanid, G. Palestra, and G. Anzivino, “Cognitive Emotions in E-Learning Processes and Their Potential Relationship with Students’ Academic Adjustment,” *Int. J. Emot. Educ.*, vol. 10, no. 1, pp. 89–111, 2018.
[63] C. Ma, C. Sun, D. Song, X. Li, and H. Xu, “A Deep Learning Approach for Online Learning Emotion Recognition,” in *13th International Conference on Computer Science & Education (ICCSE)*, 2018, pp. 1–5.
[64] B. D. Miljković, A. V. Petojević, and M. R. Žižović, “On-line Student Emotion Monitoring as a Model of Increasing Distance Learning Systems Efficiency,” *New Educ. Rev.*, vol. 47, no. 2, pp. 225–240, 2017.
[65] L. Jie, Z. Xiaoyan, and Z. Zhaohui, “Speech Emotion Recognition of Teachers in Classroom Teaching,” in *2020 Chinese Control And Decision Conference (CCDC)*, 2020, pp. 5045–5050.
[66] Z. Liu, W. Zhang, J. Sun, H. N. H. Cheng, X. Peng, and S. Liu, “Emotion and associated topic detection for course comments in a MOOC platform *,” in *International Conference on Educational Innovation through Technology Emotion*, 2016, pp. 15–19.
[67] X. Feng, Y. Wei, X. Pan, L. Qiu, and Y. Ma, “Academic Emotion Classification and Recognition Method for Large-scale Online Learning Environment—Based on A-CNN and LSTM-ATT Deep Learning Pipeline Method,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 6, p. 1941, Mar. 2020.
[68] L. Shen, V. Callaghan, and R. Shen, “Affective e-Learning in residential and pervasive computing environments,” *Inf. Syst. Front.*, vol. 10, no. 4, pp. 461–472, Sep. 2008.
[69] M. Soltani and H. Zarzour, “Facial Emotion Detection in Massive Open Online Courses,” in *Trends and Advances in Information Systems and Technologies. WorldCIST’18 2018. Advances in Intelligent Systems and Computing*, 2018, vol. 745, pp. 277–286.
[70] S. Park and J. Ryu, “Exploring Preservice Teachers’ Emotional Experiences in an Immersive Virtual Teaching Simulation through Facial Expression Recognition,” *Int. J. Human–Computer Interact.*, vol. 35, no. 6, pp. 521–533, Apr. 2019.
[71] H. Abe, T. Kamizono, K. Kinoshita, K. Baba, S. Takano, and K. Murakami, “Towards activity recognition of learners in on-line lecture,” *J. Mob. Multimed.*, vol. 11, no. 3–4, pp. 205–212, 2015.
[72] L. Catrysse, D. Gijbels, and V. Donche, “It is not only about the depth of processing: What if eye am not interested in the text?,” *Learn. Instr.*, vol. 58, pp. 284–294, 2018.
[73] F. Novita Sari, P. Insap Santosa, and S. Wibirama, “Comparison expert and novice scan behavior for using e-learning,” in *Proc. SPIE 10443, Second International Workshop on Pattern Recognition, 104430E*, 2017.
[74] B. V. Thiyagarajan, “Revolution in teaching/MOOCs by the use of real-time face detection,” in *2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 2015, pp. 1–7.
[75] Z. Zhan, L. Zhang, H. Mei, and P. S. W. Fong, “Online Learners’ Reading Ability Detection Based on Eye-Tracking Sensors,” *Sensors*, vol. 16, no. 1457, pp. 1–17, 2016.
[76] R. Moro and M. Bielikova, “Utilizing Gaze Data in Learning: From Reading Patterns Detection to Personalization,” in *Late-breaking Results and Workshop Proceedings of the 23rd Conference on User Modeling,*

Adaptation, and Personalization, 2015.

[77] B. The and M. Mavrikis, "A Study On Eye Fixation Patterns of Students in Higher Education Using an Online Learning System," in *ACM Proceeding of the Sixth International Conference on Learning Analytics & Knowledge*, 2016, pp. 408–416.

[78] Y. Bandung, K. Mutijarsa, and L. B. Subekti, "Design and Implementation of Video Conference System with Object Tracking for Distance Learning," pp. 6–11, 2017.

[79] B. Edwiranda, B. C. Purba, and Y. Bandung, "Design and Implementation of Real-time Object Tracking System based on Viola-Jones Algorithm for Supporting Video Conference," *2018 12th Int. Conf. Telecommun. Syst. Serv. Appl.*, pp. 1–6, 2018.

[80] C. Lwande, L. Muchemi, and R. Oboko, "Behaviour Prediction in a Learning Management System," in *2019 IST-Africa Week Conference (IST-Africa)*, 2019, pp. 1–10.

[81] B. Wu and J. Xiao, "Mining online learner profile through learning behavior analysis," in *Proceedings of the 10th International Conference on Education Technology and Computers - ICETC '18*, 2018, pp. 112–118.

[82] S. Fatahi, H. Moradi, and E. Farmad, "Behavioral Feature Extraction to Determine Learning Styles in E-Learning Environments," in *International Conference e-Learning*, 2015, pp. 66–72.

[83] Q. T. Nguyen, H. Tieu Binh, T. D. Bui, and P. D. N.T., "Student postures and gestures recognition system for adaptive learning improvement," in *2019 6th NAFOSTED Conference on Information and Computer Science (NICS)*, 2019, pp. 494–499.

[84] T. Xu, Z. Feng, W. Zhang, X. Yang, and P. Yu, "Depth based Hand Gesture Recognition for Smart Teaching," *2018 Int. Conf. Secur. Pattern Anal. Cybern. SPAC 2018*, pp. 387–390, 2018.

[85] J. Wang, T. Liu, and X. Wang, "Human hand gesture recognition with convolutional neural networks for K-12 double-teachers instruction mode classroom," *Infrared Phys. Technol.*, vol. 111, p. 103464, Dec. 2020.

[86] P. B. Santos, C. V. Wahle, and I. Gurevych, "Using Facial Expressions of Students for Detecting Levels of Intrinsic Motivation," *2018 IEEE 14th Int. Conf. e-Science*, pp. 323–324, 2018.

[87] P. Utami, F. Pahlevi, D. Santoso, N. Fajaryati, B. Destiana, and M. Ismail, "Android-based applications on teaching skills based on TPACK analysis," in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 535, p. 012009.

[88] R. Barmaki and C. Hughes, "Gesturing and Embodiment in Teaching: Investigating the Nonverbal Behavior of Teachers in a Virtual Rehearsal Environment," in *The Eighth AAAI Symposium on Educational Advances in Artificial Intelligence 2018 (EAAI-18) Gesturing*, 2018, pp. 7893–7899.

[89] Y. J. Lee, L.-Y. Chen, and M.-H. Lee, "Developing an Online System to Evaluate EFL College Students' Voice Performance in English Public Speaking," *Int. J. Technol. Learn.*, vol. 23, no. 2, pp. 1–13, 2016.

[90] E. Hettiarachchi, E. Mor, M. A. Huertas, and A.-E. Guerrero-Roldan, "Introducing a Formative E-Assessment System to Improve Online Learning Experience and Performance," *J. Univers. Comput. Sci.*, vol. 21, no. 8, pp. 1001–1021, 2015.

[91] P. Utami, G. P. Cikarge, M. E. Ismail, and S. Hashim, "Teaching Aids in Digital Electronics Practice through Integrating 21st Century Learning Skills using a conceptual approach," in *Journal of Physics: Conf. Series*, pp. 1–9.

[92] P. Utami, R. Hartanto, and I. Soesanti, "A Study on Facial Expression Recognition in Assessing Teaching Skills : Datasets and Methods," in *Procedia Computer Science*, 2019, vol. 161, pp. 544–552.

[93] Y. Wang, Q. Liu, W. Chen, Q. Wang, and D. Stein, "Effects of instructor's facial expressions on students' learning with video lectures," *Br. J. Educ. Technol.*, vol. 00, no. 00, pp. 1–15, 2018.

[94] S. Li, J. Zheng, and Y. Zheng, "Towards a new approach to managing teacher online learning : Learning communities as activity systems," *Soc. Sci. J.*, 2019.

[95] P. Mishra and M. J. Koehler, "Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge," *Teach. Coll. Rec.*, vol. 108, no. 6, pp. 1017–1054, 2006.

[96] S. Baddeley, "TPACK is missing an 'E,'" *Edtech in The Classroom*, 2020. [Online]. Available: https://simonbaddeley64.wordpress.com/2020/01/15/tpack-is-missing-an-e/. [Accessed: 20-Sep-2020].

[97] P. Sudira, "Model Pembelajaran Vokasional W2CPATK," in *Pembelajaran Vokasional Era Revolusi Industri 4.0*, Yogyakarta: UNY Press, 2020, p. 271.

[98] H. Tanaka, S. Sakti, G. Neubig, T. Toda, and S. Nakamura, "Modality and Contextual Differences in Computer Based Non-verbal Communication Training," in *4th IEEE International Conference on Cognitive Infocommunications*, 2013, pp. 127–132.

[99] M. Castañer and O. Camerino, "The Teacher's Body Communicates. Detection of Paraverbal Behaviour Patterns," in *The Temporal Structure of Multimodal Communication*, L. Hunyadi and I. Szekrenyes, Eds. 2020, pp. 159–159.

[100] R. I. Arends, *Learning to Teach*, 9th ed. NewYork: McGraw-Hill.

[101] C. Ma, C. Sun, D. Song, X. Li, and H. Xu, "A Deep Learning Approach for Online Learning Emotion Recognition," in *2018 13th International Conference on Computer Science & Education (ICCSE)*, 2018, no. Iccse, pp. 1–5.

[102] G. Du, M. Chen, C. Liu, B. Zhang, and P. Zhang, "Online robot teaching with natural human-robot interaction," *IEEE Trans. Ind. Electron.*, vol. 65, no. 12, pp. 9571–9581, 2018.

[103] S. D. Craig, J. Twyford, N. Irigoyen, and S. A. Zipp, "A Test of Spatial Contiguity for Virtual Human's Gestures in Multimedia Learning Environments," *J. Educ. Comput. Res.*, vol. 53, no. 1, pp. 3–14, 2015.