

## A Review on Multimodal Online Educational Engagement Detection System Using Facial Expression, Eye Movement and Speech Recognition

R. Angeline<sup>a</sup>, and A. Alice Nithya<sup>b</sup>

<sup>a</sup>  
Department of Computer Science and Engineering, SRMIST, India.  
ar1501@srmist.edu.in, alicenia@srmist.edu.in,

**Article History:** Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

**Abstract:** During this lockdown period, an online educational engagement system plays a vital role to enrich the knowledge of learners in various fields without interrupting their learning process. Online educational engagement systems include all the activities of a learner like listening, reading, writing, and so on. While participating in these activities, a participant may show various levels of engagements like fully engaged, partially engaged and completely not engaged. The participation of online learners has to be identified for an effective learning process. The existing literature could be classified depending upon the learners' participation as automatic, semi-automatic and manual. Further it could be sub categorised based on the data types used to identify the engagement system. In this paper, a review on computer based automatic online educational engagement detection systems is presented. Several educational engagement methods are applied for computer based online engagement detection systems. In these systems examining a participant's presence and attention with the modalities of facial expression, eye movement and speech are found to be a challenging task. In this work, it is also identified that there are few challenges like preparation and usage of proper datasets, identifying suitable performance metrics for different tasks involved and providing recommendations for future enhancement of online educational engagement detection by combining the modalities of facial expression, eye movement and speech are still unattended. Though there are several research gaps involved, an online educational engagement system will help the learners to engage themselves in a productive way of learning and getting evaluated efficiently and effectively during the lockdown period of pandemic disease COVID-19 without interrupting their learning process and gaining knowledge.

**Keywords:** automatic engagement, effective learning process, multimodal, eye movement, facial expression, online educational engagement detection, speech recognition.

### 1. Introduction

In an online educational engagement system the learner will attain knowledge by learning and their activities such as writing, reading, listening, partially listening and completely not listening, have to be monitored by the mentor. The monitoring of the learners activities could be done by various methods like fully automatic, partially automatic, and manual engagement detection systems. By using the online educational engagement system the performance of the learner can be measured to improve the teaching-learning process (Shoumy et al., 2020).

The engagement system of learners could be categorized in different forms. They are the affective engagement (Shoumy et al., 2020, Aluja-Banet et al., 2019, Kaur et al., 2018), the behavioural engagement (Aluja-Banet et al., 2019, Kaur et al., 2018, Littlewort et al., 2006, Matthews et al., 2002), the cognitive engagement (Aluja-Banet et al., 2019, Kaur et al., 2018, Matthews et al., 2002, Mach., 2005), the academic engagement (Booth et al., 2017), the emotional engagement (Mach., 2005), the psychological engagement and the agentic engagement (Bosch, 2016).

The affective engagement is having real interest and enjoys the learning (Kaur et al., 2018). The behavioural engagement is getting participant ideas like classroom interaction, activities, and assignments (Matthews et al., 2002). The cognitive engagement is how the learner is focussed on participation, implementation of creative ideas or skills for completing assignments (Mach., 2005). Academic engagement refers to the participants' on-task behaviour (Al-Hendawi, 2012). The emotional engagement refers to student-teacher reactions in the classroom, academics. It includes the learners' interest, boredom, happiness, sadness, and anxiety (Mach., 2005). Above all psychological engagement also plays a role in having relationships with teachers and learners (Johns et al., 2006). The agentic engagement is the constructive contribution of the learners with the full flow of instructions what they receive.

These different forms of learners' engagement are useful to find the interventions and helpful to modify the online engagement system. Analyzing the learners engagement level on different categories helps in measuring the performance of the learners' by using this engagement system (Bosch, 2014). The engagement theorists identify two types of data: a) internal to the individual, which includes affective and cognitive, and b) external observable factors, which includes gestures, postures, voice, movements (Kaur et al., 2018, Bosch et al., 2015, Bosch et al., 2016, Moeed & Anderson, 2018) factor analysis by descriptive latent variable (Cocea & Weibelzahl, 2011, Koydemir & Ozcan, 2018). The engagement level could be measured by (1) learners' valuable feedback, (2) self-reports, (3) engagement ratings, and (4) Automated measurements (Booth et al., 2017).

Since 1980, engaging learners in proper educational activities through online platforms is one of the major issues (Booth et al., 2017). Many researchers acknowledged that the online learning engagement system and its influence, increased its output by learning more, with less cost and without wasting much time in travelling. Designing and developing personalized pedagogical tools for an online engagement system could help in improving its success rate (Buscher et al., 2008). This kind of learning engagement system will change traditional learning environments and help in building a more effective and efficient educational system in a country like India (Aslan et al., 2014).

I am planning to write a review paper, which will help in improving and personalizing the online education system. Combining multimodal online educational engagement systems like facial expressions, eye movement and speech recognition will be helping more to improve the education system.

The remainder of the paper is organised as follows: Section 2 is about existing various educational engagement detection methods under the classification of fully automatic, partially automatic and manual. Section 3 is about online learning environments that could be examined by capturing and analyzing the video of the learners' facial expressions, audio and eye movements. So the online learning environment is promising, cost-effective, and nonintrusive in nature. Section 4 is about exploring existing datasets, performance measurement and evaluation metrics and techniques for an online educational engagement system. Section 5 provides conclusions and future recommendations.

## 2. The Review of Literature of Educational Engagement Detection System

In online educational engagement detection systems various engagement detection methods are used to study the learners activities. In this paper a detailed review on online educational engagement detection systems is performed to identify the research gaps involved in improving and personalizing the system. According to researchers, the engagement detection methods are classified as i) Automatic method, ii) Semi-automatic method and iii) manual method (Booth et al., 2017).

In the automatic educational engagement detection method, the activities of the learners are recorded and the performance metrics are evaluated. This automatic method could be recorded and evaluated using the computer vision-based method, sensor data analysis and log-file analysis (Chaouachi & Frasson, 2010). In the semi-automatic method, engagement tracing is done for detecting learners' educational engagement. In the manual method the mentor has to maintain the observational checklist and self reporting to measure the learners knowledge.

In this digital era, the learners are very much attracted to learn, submit assignments, assessments and evaluate through online mode of education. The learners could view many learning materials in the form of video, text, audio and various forms to enrich their knowledge compared to traditional educational systems such as black board and chalk method. The mentors could also deliver their knowledge in various forms pedagogical method in online learning.

While delivering materials in various methods, the mentor has to find out the different activities such as learning, reading, writing, and also engaging other than learning activities of the learners. This will improve the online educational system and to increase learner involvement rate.

### COMPUTER VISION BASED METHODS

In the automatic computer vision-based method for online educational engagement, performance metrics of various levels for learners' engagement are measured. By this, the mentor could observe and understand the learners in a conventional method of teaching environment. The automation method creates a virtual classroom setting which will reduce the cost, and the affective engagement system reduces frustration and dropout rates.

#### Generic Educational Engagement Detection Methods Using Computer Vision

Automatic computer vision based methods for learners' perceived educational engagement detection has five different modules (see Figure 1) (Anderson et al, 2004).

The learners' engagement will be captured by the camera. The video will be given as the input to detect and track the modules. The Region Of Interests (ROI) features will be extracted and passed through the classification module. From tracking trajectories and classification scores, decision will

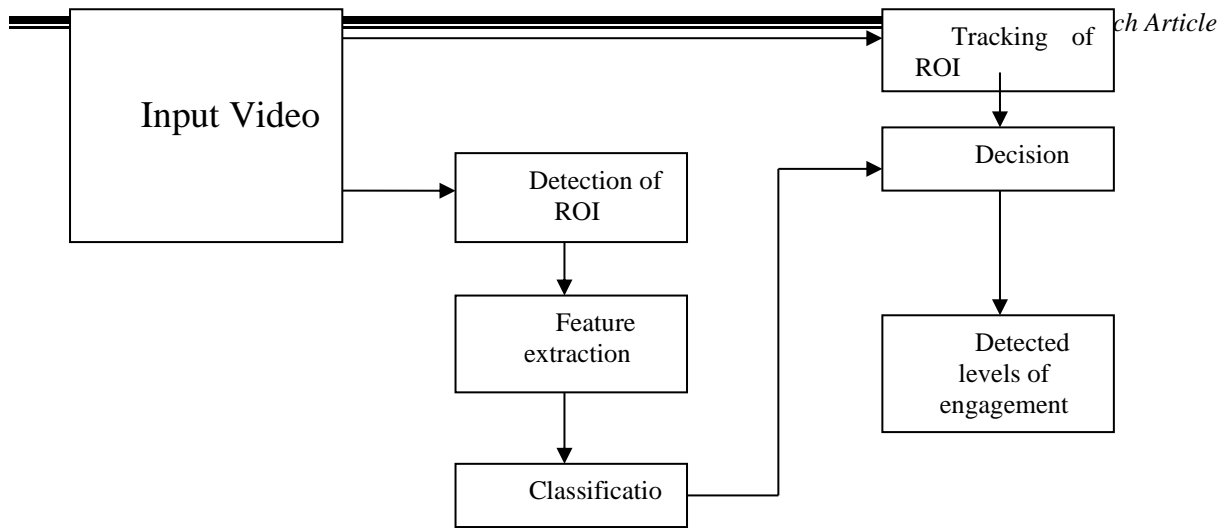


Figure 1. A generic engagement detection method using computer vision  
 be taken and detected the learners' level of engagement. We further review by combining the three modalities of facial expression, speech and eye movement for effective educational engagement detection.

#### Facial Expressions

In educational engagement facial detection, any inferential process such as happiness, sadness, confusion, anger, frustration, boredom have to be measured, which is a descriptive analysis of the learner. Facial expression of the learner is one of the modalities used to understand his/her educational engagement (Bosch, 2016, Sundar & Kumar, 2015, Dewan et al, 2018, Chen et al, 2013). Through camera facial images are captured in a continuous manner and without intrusion for understanding learners' activities. From the captured facial expressions the attitude of the learner could be understood.

Depending upon the information captured by the video camera, face expressions identifying methods are categorized into two: a) the Part-Based methods or Geometric-Based methods and b) the Appearance-Based methods (Bosch et al., 2016).

#### a)Part-Based Methods or Geometric-Based Methods

Dewan proposed the Part-Based methods or Geometric-Based methods in educational engagement detection technique is used to analyze the various parts of a face like eyes, eyebrows, nose, cheek, chin, mouth and so on (Bosch et al., 2016, Anderson et al, 2004).

Ekman and Friesen led the way for facial detection in developing the Facial Action Coding System (FACS) (Chen et al, 2013, Khalfallah & Slama, 2015). To design the theoretical measures of facial emotions, facial muscle movements known as Action Units (AUs) are used in FACS. For analysing facial expressions, neuroscientists and psychologists use FACS. Every AUs could be used as a single or any combination to map with facial expressions (Khalfallah & Slama, 2015).

The learners' inferential processes are captured, analysed and the result is correlated positive or negative actions (Moeed & Anderson, 2018). The learners' video was collected and the AUs were used to detect facial landmarks, eye gaze, emotion probabilities, average optical flow magnitude and direction, and head pose and size (Littlewort et al., 2006, Booth et al., 2017, Johns et al., 2006, Bosch et al., 2016, Anderson et al, 2004, Christenson et al, 2012, Chu et al., 2017, Cocca & Weibelzahal, 2009, D'Mello et al., 2014, D'Mello & Craig, 2009, D'Mello & Graesser, 2010, Ekman & Frisen, 1978). After classifying the positive and negative sides of the learners, ROI is tracked and decision was taken automatically. The classifiers were combined to find out the facial detection. C4.5 trees, and Bayesian classifiers were used for classification. The decisions were used to reduce or increase the speed of instruction, to fix the engagement category: engaged, not engaged, normally engaged and very engaged (Bosch, 2016, Bosch et al., 2016).

The intensity and frequency values could be combined and used to predict tutoring outcomes. The engagement level could be classified as confusion, frustration, boredom, neutral, and engaged. Again these engagement levels could be higher or lower depending upon the content of the lecture, learners' interest, and higher engagement and includes various conditions. In all the engagement levels, AUs might be combined to get facial expression but textures could not be extracted in part-based methods. 2D and 3D images received by Kinect cameras were combined with AUs for facial engagement detection.

#### b)Appearance-Based Methods

By noticing the learners' appearance from the video, the level of learner involvement is detected. In appearance-based methods, patterns were generated from the feature extraction for engagement classification. Some of the feature extraction techniques discussed by Moeed & Anderson, 2018 are,

- Local Binary Patterns (LBP)
- Local Directional Patterns(LDP)
- Histogram of Oriented Gradients (HOG)
- LBP in Three Orthogonal Planes (LBP-TOP).

For engagement detection system Deep Learning approaches Cocea & Weibelzahl, 2011 like

- LBP in Three Orthogonal Planes (LBP-TOP),
- Deep Multi-Instance Learning (DMIL).

were used. In this method texture extraction likes to be more sensitive in brightness, shadows. Also this method recognizes facial expressions by analysing the changes of the face surface in both static and dynamic. These appearance-based methods use different types of features, such as Gabor wavelet coefficients, optical flow, and active appearance models (Bosch et al., 2016).

The extracted features used to classify the learners' status- listening, boredom, frustration, head movement, hand movement, basic emotions (Jin Fei & Pavlidis, 2010), gestures, postures, and so on. The feedback generated from the classified learners' status (Frank & Fruchter, 2014). Learners' faces could be detected by fuzzy-based engagement detection methods which are used to evaluate the facial features, distance between facial features and edges, mouse operations and keyboard operations. Tuning the head, moving from the learning position, talking with someone was identified as less interesting in learning attitude (Fredricks et al, 2004).

In group engagement detection system, SVM classifier is used to classify six engagement levels of disengagement, the relaxed engagement, involved engagement, action, intention to act, and involved action (Krithika et al., 2016).

Remote laboratory was proposed for web-based intelligent tutorials where learners could import their knowledge from anywhere computing experiments in real laboratories using the internet. Using 70 small classifiers, the involvement of learners identified (Goldberg et al., 2011) frustration and serenity.

#### Eye Movement

Sinem Aslan used the eye tracker to track the learners' gaze and region of interests which is used to understand the engagement level in online educational activity. The eye tracker used for learners detects gaze and it is combined with statistical facial features and depth information. Nine pilot sessions among the learners were done feature selection and classification for engaged and not engaged learners by using five machine learning algorithms with increased accuracy up to 85-90% on the collected dataset (Chu et al., 2017).

The various concentration levels of the student are measured by the head movement and eye movement using the algorithms Viola Jones and Local Binary Pattern (Grafsgaard et al., 2013a). The reading of online tutorials by linear and segmented secure coding modules was used for eye tracking investigation. Linear modules were the large amount of content on a single screen and the segmented module was the same with broken contents. The investigation was done on the learners' content skipping and reading.

Performance metrics were done based on the reading scores and reading depths of eye tracking. Higher scores were received for segmented modules rather than linear modules (Grafsgaard et al., 2013a). To get the relevant feedback the read-to-skimmed ratio is measured (Mach., 2005). The main challenges faced in these methods are proper eye-calibration and restricting learners' movement to be within an eye-tracker range. This is not practically possible for the real life educational environment.

#### Speech Recognition

The speech of the learner could be combined with facial expression to judge the learners' engagement detection (Cocea & Weibelzahl, 2011, Grafsgaard et al., 2013b). The audio of the learner could be integrated with the measured facial expression and eye movement. By combining all these three modalities, Multimodal Sentiment Analysis has been done and the effectiveness of online engagement detection could be more accurate. This enhances the quality of the mentor and learners. Audio mining (Happy et al., 2017) could be done to find the interested area of the learners in educational engagement using frequently used keywords.

#### Datasets

A survey of available datasets, evaluation metrics and techniques for educational engagement detection systems are discussed in this section. These are the available datasets used for educational engagement activities in different levels of learners. All datasets have their own characteristics for learners' activities. Labeling the DAiSEE (Dewan et al, 2018) dataset could be done by crowd sourcing and HBCU by human experts. "In-the-wild" (Kaur et al., 2018) dataset is used to find facial expressions. SDMATH (Sathayanarayana et al., 2014) is the richly labelled dataset used for deictic gestures, which includes speech gestures, eye gaze and facial expressions.

Table 1. Datasets for educational engagement detection

Data set used	Number of videos	Number of learners	Activities	Input	Inference
DAiSEE (Dewan et al, 2018)	9068	112 (80 male and 32 female)	engaged, frustration, boredom, confusion 2 Level-engaged, not engaged 3Level-Not engaged, normally engaged, very engaged	Facial Expressions	First multi-label video classification dataset
HBCU (Whitehill et al., 2014)	122	34 (9 male and 25 female)	not engaged, nominally engaged, engaged, very engaged	Facial Expressions	Considering the facial expressions learners engagement is detected
in-the wild (Kaur et al., 2018)	195	78 (2 female and 53 male)	disengaged, barely engaged, normally engaged, highly engaged	Facial Expressions	Studying problem for unconstrained face recognition.
SDMATH (Sathayanarayana et al., 2014)	20	20 (10 male and 10 female)	deictic gestures	speech gestures, eye gaze, facial expressions	Study the engagement of learners considering speech gestures, eye gaze and facial expressions

Evaluation Metrics and Techniques

In this review widely used evaluation metrics and techniques such as human and automatic learners' perception (Whitehill et al., 2014), CERT results (Littlewort et al, 2011), Linear regression models (Grafsgaard et al., 2013a), Multi-instance learning based methods by (Kaur et al., 2018), Emote-aloud methodology for emotion learners by (D'Mello & Craig 2009), facial expressions using FACET method (Dewan et al, 2018) and Kinect face tracker, LBP-TOP, and Heart Rate - HR for eye, face and cardiac activity (Monkaresi et al., 2017) are tabulated in table 2.

Table 2. Performance metrics for educational engagement system

Investigation Method / Model	Metrics	First author and year
Human and automatic perception	Learners' pre and post-test evaluated	(Whitehill et al., 2014)
CERT results with manual annotations	Understanding every moment of learners	(Littlewort et al., 2011)
Linear regression models	Frustration and learning activities from facial expression	(Grafsgaard et al., 2013a)
Multi-instance learning-based method	Facial expression, body movements and gaze movements	(Kaur et al., 2018)

Emote-aloud methodology	Emotion of learners	(D'Mello & Craig, 2009)
FACET	Facial expression	(Dewan et al, 2018)
Kinect face tracker, LBP-TOP, and Heart Rate - HR	Eye, face and cardiac activity	(Monkaresi et al., 2017)

### 3. Conclusion

In this work, a review on an online engagement detection system using computer vision based methods for multimodality was done. From the review, it was observed that, combining facial expression, eye movement and audio were able to provide improved accuracy when the input data had less noise. It was also observed that of all the models used so far, combining the models for facial expression, eye movement, speech and heart rate were found to provide improved accuracy in educational engagement detection systems.

### Acknowledgment

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### References

1. Al-Hendawi, M. (2012). Academic engagement of students with emotional and behavioral disorders: Existing research, issues, and future directions. *Emotional and Behavioural Difficulties*, 17(2), 125-141. <https://doi.org/10.1080/13632752.2012.672861>
2. Aluja-Banet, T., Sancho, M., & Vukic, I. (2019). Measuring motivation from the virtual learning environment in secondary education. *Journal of Computational Science*, 36, 100629. <https://doi.org/10.1016/j.jocs.2017.03.007>
3. Anderson, A. R., Christenson, S. L., Sinclair, M. F., & Lehr, C. A. (2004). Check & connect: The importance of relationships for promoting engagement with school. *Journal of School Psychology*, 42(2), 95-113. <https://doi.org/10.1016/j.jsp.2004.01.002>
4. Aslan, S., Cataltepe, Z., Diner, I., Dundar, O., Esme, A. A., Ferens, R., Kamhi, G., Oktay, E., Soysal, C., & Yener, M. (2014). Learner engagement measurement and classification in 1:1 learning. 2014 13th International Conference on Machine Learning and Applications. <https://doi.org/10.1109/icmla.2014.111>
5. Bartlett, M. S., Littlewort, G. C., Frank, M. G., Lainscsek, C., Fasel, I. R., & Movellan, J. R. (2006). Automatic recognition of facial actions in spontaneous expressions. *Journal of Multimedia*, 1(6). <https://doi.org/10.4304/jmm.1.6.22-35>
6. Booth, B. M., Ali, A. M., Narayanan, S. S., Bennett, I., & Farag, A. A. (2017). Toward active and unobtrusive engagement assessment of distance learners. 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII). <https://doi.org/10.1109/acii.2017.8273641>
7. Bosch, N. (2016). Detecting student engagement. *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization - UMAP '16*. <https://doi.org/10.1145/2930238.2930371>
8. Bosch, N., Chen, Y., & D'Mello, S. (2014). It's written on your face: Detecting affective states from facial expressions while learning computer programming. *Intelligent Tutoring Systems*, 39-44. [https://doi.org/10.1007/978-3-319-07221-0\\_5](https://doi.org/10.1007/978-3-319-07221-0_5)
9. Bosch, N., D'Mello, S., Baker, R., Ocumpaugh, J., Shute, V., Ventura, M., Wang, L., & Zhao, W. (2015). Automatic detection of learning-centered affective states in the wild. *Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI '15*. <https://doi.org/10.1145/2678025.2701397>

10. Bosch, N., D'mello, S. K., Ocumpaugh, J., Baker, R. S., & Shute, V. (2016). Using video to automatically detect learner affect in computer-enabled classrooms. *ACM Transactions on Interactive Intelligent Systems*, 6(2), 1-26. <https://doi.org/10.1145/2946837>
11. Buscher, G., Dengel, A., & Van Elst, L. (2008). Eye movements as implicit relevance feedback. *Proceeding of the twenty-sixth annual CHI conference extended abstracts on Human factors in computing systems - CHI '08*. <https://doi.org/10.1145/1358628.1358796>
12. Chaouachi, M., & Frasson, C. (2010). Exploring the relationship between learner EEG mental engagement and affect. *Intelligent Tutoring Systems*, 291-293. [https://doi.org/10.1007/978-3-642-13437-1\\_48](https://doi.org/10.1007/978-3-642-13437-1_48)
13. Chen, J., Liu, X., Tu, P., & Aragonés, A. (2013). Learning person-specific models for facial expression and action unit recognition. *Pattern Recognition Letters*, 34(15), 1964-1970. <https://doi.org/10.1016/j.patrec.2013.02.002>
14. Chu, W., De la Torre, F., & Cohn, J. F. (2017). Selective transfer machine for personalized facial expression analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(3), 529-545. <https://doi.org/10.1109/tpami.2016.2547397>
15. Cocea, M., & Weibelzahl, S. (2009). Log file analysis for disengagement detection in E-LEARNING environments. *User Modeling and User-Adapted Interaction*, 19(4), 341-385. <https://doi.org/10.1007/s11257-009-9065-5>
16. Cocea, M., & Weibelzahl, S. (2011). Disengagement detection in online learning: Validation studies and perspectives. *IEEE Transactions on Learning Technologies*, 4(2), 114-124. <https://doi.org/10.1109/tlt.2010.14>
17. Dewan, M. A., Lin, F., Wen, D., Murshed, M., & Uddin, Z. (2018). A deep learning approach to detecting engagement of online learners. *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)*. <https://doi.org/10.1109/smartworld.2018.00318>
18. D'Mello, S. K., Craig, S. D., & Graesser, A. C. (2009). Multimethod assessment of affective experience and expression during deep learning. *International Journal of Learning Technology*, 4(3/4), 165. <https://doi.org/10.1504/ijlt.2009.028805>
19. D'Mello, S. K., & Graesser, A. (2010). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction*, 20(2), 147-187. <https://doi.org/10.1007/s11257-010-9074-4>
20. D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, 29, 153-170. <https://doi.org/10.1016/j.learninstruc.2012.05.003>
21. Ekman, P., & Friesen, W. V. (1978). Facial action coding system. *PsycTESTS Dataset*. <https://doi.org/10.1037/t27734-000>
22. Fairclough, S. H., & Venables, L. (2006). Prediction of subjective states from psychophysiology: A multivariate approach. *Biological Psychology*, 71(1), 100-110. <https://doi.org/10.1016/j.biopsycho.2005.03.007>
23. Frank, M., & Fruchter, R. (2014). Global teamwork: The influence of multiculturalism on project product and process success. *Computing in Civil and Building Engineering (2014)*. <https://doi.org/10.1061/9780784413616.175>
24. Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109. <https://doi.org/10.3102/00346543074001059>

25. Goldberg, B. S., Sottilare, R. A., Brawner, K. W., & Holden, H. K. (2011). Predicting learner engagement during well-defined and ill-defined computer-based intercultural interactions. *Affective Computing and Intelligent Interaction*, 538-547. [https://doi.org/10.1007/978-3-642-24600-5\\_57](https://doi.org/10.1007/978-3-642-24600-5_57)
26. Grafsgaard, J. F., Wiggins, J. B., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2013). Automatically recognizing facial indicators of frustration: A learning-centric analysis. 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction. <https://doi.org/10.1109/acii.2013.33>
27. Grafsgaard, J. F., Wiggins, J. B., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2013). Embodied affect in tutorial dialogue: Student gesture and posture. *Lecture Notes in Computer Science*, 1-10. [https://doi.org/10.1007/978-3-642-39112-5\\_1](https://doi.org/10.1007/978-3-642-39112-5_1)
28. Christenson, S., Reschly, A., Wylie, C., Handbook of research on student engagement. (2012). <https://doi.org/10.1007/978-1-4614-2018-7>
29. Happy, S. L., Patnaik, P., Routray, A., & Guha, R. (2017). The Indian spontaneous expression database for emotion recognition. *IEEE Transactions on Affective Computing*, 8(1), 131-142. <https://doi.org/10.1109/taffc.2015.2498174>
30. Harris, L. R. (2008). A phenomenographic investigation of teacher conceptions of student engagement in learning. *The Australian Educational Researcher*, 35(1), 57-79. <https://doi.org/10.1007/bf03216875>
31. Jeni, L. A., Cohn, J. F., & De La Torre, F. (2013). Facing Imbalanced data--recommendations for the use of performance metrics. 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction. <https://doi.org/10.1109/acii.2013.47>
32. Jin Fei, & Pavlidis, I. (2010). Thermistor at a distance: Unobtrusive measurement of breathing. *IEEE Transactions on Biomedical Engineering*, 57(4), 988-998. <https://doi.org/10.1109/tbme.2009.2032415>
33. Johns, J., Mahadevan, S., & Woolf, B. (2006). Estimating student proficiency using an item response theory model. *Intelligent Tutoring Systems*, 473-480. [https://doi.org/10.1007/11774303\\_47](https://doi.org/10.1007/11774303_47)
34. Kamath, A., Biswas, A., & Balasubramanian, V. (2016). undefined. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). <https://doi.org/10.1109/wacv.2016.7477618>
35. Kapoor, A., & Picard, R. W. (2005). Multimodal affect recognition in learning environments. Proceedings of the 13th annual ACM international conference on Multimedia - MULTIMEDIA '05. <https://doi.org/10.1145/1101149.1101300>
36. Kaur, A., Mustafa, A., Mehta, L., & Dhall, A. (2018). Prediction and localization of student engagement in the wild. 2018 Digital Image Computing: Techniques and Applications (DICTA). <https://doi.org/10.1109/dicta.2018.8615851>
37. Khalfallah, J., & Slama, J. B. (2015). Facial expression recognition for intelligent tutoring systems in remote laboratories platform. *Procedia Computer Science*, 73, 274-281. <https://doi.org/10.1016/j.procs.2015.12.030>
38. Koydemir, H. C., & Ozcan, A. (2018). Wearable and Implantable sensors for biomedical applications. *Annual Review of Analytical Chemistry*, 11(1), 127-146. <https://doi.org/10.1146/annurev-anchem-061417-125956>
39. Krithika L.B, & Lakshmi Priya GG. (2016). Student emotion recognition system (SERS) for E-Learning improvement based on learner concentration metric. *Procedia Computer Science*, 85, 767-776. <https://doi.org/10.1016/j.procs.2016.05.264>



40. Kuoan Hwang, & Chiahao Yang. (2008). Fuzzy fusion for affective state assessment in distance learning based on image detection. 2008 International Conference on Audio, Language and Image Processing. <https://doi.org/10.1109/icalip.2008.4589995>
41. Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., & Bartlett, M. (2011). The computer expression recognition toolbox (CERT). *Face and Gesture* 2011. <https://doi.org/10.1109/fg.2011.5771414>
42. Mach, M. (2005). Tracing legal knowledge evolution. Proceedings of the 10th international conference on Artificial intelligence and law - ICAIL '05. <https://doi.org/10.1145/1165485.1165527>
43. Martinez, B., Valstar, M. F., Jiang, B., & Pantic, M. (2019). Automatic analysis of facial actions: A survey. *IEEE Transactions on Affective Computing*, 10(3), 325-347. <https://doi.org/10.1109/taffc.2017.2731763>
44. Mason, S. J., & Weigel, A. P. (2009). A generic forecast verification framework for administrative purposes. *Monthly Weather Review*, 137(1), 331-349. <https://doi.org/10.1175/2008mwr2553.1>
45. Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., Grier, R., & Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, 2(4), 315-340. <https://doi.org/10.1037/1528-3542.2.4.315>
46. Moeed, A., & Anderson, D. (2018). The New Zealand context and research design. *Learning Through School Science Investigation*, 17-31. [https://doi.org/10.1007/978-981-13-1616-6\\_2](https://doi.org/10.1007/978-981-13-1616-6_2)
47. Monkaresi, H., Bosch, N., Calvo, R. A., & D'Mello, S. K. (2017). Automated detection of engagement using video-based estimation of facial expressions and heart rate. *IEEE Transactions on Affective Computing*, 8(1), 15-28. <https://doi.org/10.1109/taffc.2016.2515084>
48. Saneiro, M., Santos, O. C., Salmeron-Majadas, S., & Boticario, J. G. (2014). Towards emotion detection in educational scenarios from facial expressions and body movements through multimodal approaches. *The Scientific World Journal*, 2014, 1-14. <https://doi.org/10.1155/2014/484873>
49. Sathayanarayana, S., Satzoda, R. K., Carini, A., Lee, M., Salamanca, L., Reilly, J., Forster, D., Bartlett, M., & Littlewort, G. (2014). Towards automated understanding of student-tutor interactions using visual deictic gestures. 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops. <https://doi.org/10.1109/cvprw.2014.77>
50. Shoumy, N. J., Ang, L., Seng, K. P., Rahaman, D., & Zia, T. (2020). Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals. *Journal of Network and Computer Applications*, 149, 102447. <https://doi.org/10.1016/j.jnca.2019.102447>
51. Sundar, P., & Kumar, A. (2015). A novel disengagement detection strategy for online learning using quasi framework. 2015 IEEE International Advance Computing Conference (IACC). <https://doi.org/10.1109/iadcc.2015.7154784>
52. Whitehill, J., Serpell, Z., Lin, Y., Foster, A., & Movellan, J. R. (2014). The faces of engagement: Automatic recognition of student Engagement from facial expressions. *IEEE Transactions on Affective Computing*, 5(1), 86-98. <https://doi.org/10.1109/taffc.2014.2316163>
53. Wixon, N., Schultz, S., Muldner, K., Alessio, D., Burleson, W., Woolf, B., & Arroyo, I. (2016). Internal & external attributions for emotions within an ITS. Proceedings of the 2016

Conference on User Modeling Adaptation and Personalization - UMAP '16.  
<https://doi.org/10.1145/2930238.2930277>