Educational Administration: Theory and Practice
2024, 30(10), 2599-2604
ISSN: 2148-2403
https://kuey.net/
Research Article

# Student Engagement And Disengagement Image Dataset For Educational Research

Parinita Chate[a,b] and Kailas Patil[a*]

[a]Department of Computer Engineering, Vishwakarma University, Pune, India
[b*]Bharati Vidyapeeth's College Of Engineering Lavale, Pune, India

***Corresponding Author:** Kailas Patil
*Email: kailas.patil@vupune.ac.in

**Citation:** Kailas Patil, et.al (2024). Student Engagement And Disengagement Image Dataset For Educational Research, *Educational Administration: Theory and Practice*, 30(10) 2599-2604
Doi: 10.53555/kuey.v30i10.10736

**ARTICLE INFO**

**ABSTRACT**

This study addresses the challenge of detecting student engagement and disengagement in online learning, a critical issue in the current era of virtual education. To support the development of machine learning models that can accurately identify student behavior during online sessions, we created the Student Engagement and Disengagement Image Dataset, comprising 16,000 images of students aged 5-21+ years. Images were systematically pre-processed, resized to 256 × 256 pixels, and organized into folders based on age and gender. Fine-tuning these models with our dataset improved their performance in classifying engagement and disengagement behaviors. The dataset's structured design and extensive labeling make it an essential tool for machine learning applications aimed at enhancing student monitoring in online education. Our findings highlight the dataset's potential to contribute to more personalized and adaptive e-learning environments, though future work incorporating multimodal data could further improve model accuracy and applicability.

**Keywords:** Online Education, Live Class, Attention, Human Behavior

## 1. Introduction

The COVID-19 pandemic has profoundly transformed global education systems, leading to a widespread transition toward online learning. With the continued growth of remote education, universities, schools, and digital platforms are offering an increasing number of virtual courses, certifications, and degree programs. Yet, one of the most critical challenges for educators remains ensuring student engagement in virtual classrooms—a factor strongly associated with learning outcomes and retention rates (Gray & DiLoreto, 2016; Martin & Bolliger, 2018). Student engagement in online learning has been examined through multiple approaches, from self-reported surveys to real-time assessments using artificial intelligence (AI) and machine learning (Hasnine et al., 2023; Dewan et al., 2019). Despite these advances, existing models often struggle to capture engagement levels accurately, especially in real-time and across varied age groups. Recent research highlights the significance of non-verbal signals, such as facial expressions and body posture, as key indicators of attentiveness or disengagement (Pabba & Kumar, 2022; Schiavo et al., 2024). Machine learning systems capable of recognizing these cues hold considerable promise for enhancing the adaptability of online learning environments.

High-quality, labeled image datasets play a vital role in developing such AI applications (Suryawanshi et al., 2022; 2023; 2024). However, the limited availability of student behavior datasets has constrained progress in this field. To address this gap, we introduce a novel image dataset specifically curated to train machine learning models for classifying engagement and disengagement during online sessions. The dataset encompasses images of students aged 5-21+, representing a wide spectrum of learning stages, and was captured using mobile technology before being standardized for machine learning applications.

The primary objective of this work is to provide researchers and developers with a valuable resource to advance real-time engagement monitoring in virtual classrooms. In doing so, it supports efforts to improve the quality, interactivity, and personalization of online education. This study not only addresses existing limitations but also establishes a foundation for future innovations in AI-driven educational tools.

## 2. Review of existing datasets

In this section, we review existing datasets and computer vision techniques used to analyse student engagement and disengagement in e-learning.
Most existing datasets that capture facial expressions, head position, and emotions are not directly linked to e-learning environments. Examples include FER-2013 MES dataset (Bhardwaj et al., 2021), EmotiW2019 (Buono et al., 2023), AffectNet, CK+ (Cohn-Kanade Plus), OULAD (Alruwais & Zakariah, 2023), the Affective-MIT Facial Expression (AM-FED) dataset (McDuff et al., 2013), and the Aff-Wild dataset (Kollias et al., 2019).

- **FER-2013 MES dataset**: It comprises a total of 35,000+ images, a collection of grayscale facial images with seven basic emotions: angry, disgust, fear, happy, sad, surprise, or neutral.
- **EmotiW2019**: It contains both facial expressions and audio to capture emotions in dynamic environments. Emotions include happy, sad, angry, surprised, neutral, fearful, disgusted, and more.
- **AffectNet**: A large-scale dataset of 1M+ images annotated with facial expressions and emotions: anger, contempt, disgust, fear, happiness, neutral, sadness, surprise.
- **CK+ (Cohn-Kanade Plus)**: 600+ facial expression sequences with labels. *Use Case*: Facial emotion recognition and deep learning.
- **OULAD**: Consists of structured tabular data and does not contain images. It includes numerical, categorical, and temporal data for student engagement analysis, student interactions, assessments, and course information.
- **Aff-Wild**: An in-the-wild collection of 500 YouTube videos where participants display emotions while watching videos or performing activities. It includes frame-by-frame labels for valence and excitement.
- **AM-FED**: Consists of frames from ~250 webcam videos of individuals watching commercials. It provides frame-by-frame annotations but exhibits limited variation in head pose.

A few publicly available datasets are more closely tied to student learning contexts. Among them, the **DAiSEE dataset** (Gupta et al., 2016) and the **Kaur dataset** (Kaur et al., 2018) are noteworthy.

- **The DAiSEE dataset** contains video recordings of students in a simulated online learning environment. Participants were shown two types of videos—educational and recreational—to represent focused and relaxed settings, thus enabling variation in engagement levels. Frame-by-frame annotations were crowdsourced for engagement, frustration, confusion, and boredom.
- **The Kaur dataset** includes ~200 videos collected from ~80 participants in a simulated educational setup. Learners were shown only instructional content, including *Learn the Korean Language in 5 Minutes*, *Tips to Learn Faster* (pictorial video), and *How to Write a Research Paper*. Each video received a single engagement label representing the overall engagement level.

Most of the datasets are generalized datasets without specific focus on e-learning environments. In contrast, our proposed dataset is designed to address the limitations of the above datasets. Our dataset contains engagement/disengagement classification considering age group and gender as factors along with facial expressions. This design makes it particularly suitable for detecting both engagement and disengagement of students of various age groups along with gender (see Table 1).

| Dataset | Type / Content | Size / Coverage | Emotions / Labels / Use Case |
|---|---|---|---|
| **FER-2013 MES** | Grayscale facial images | 35,000+ images | Seven basic emotions: angry, disgust, fear, happy, sad, surprise, neutral |
| **EmotiW2019** | Facial expressions + audio | Not specified | Emotions: happy, sad, angry, surprised, neutral, fearful, disgusted, and more |
| **AffectNet** | Facial images | 1M+ images | Emotions: anger, contempt, disgust, fear, happiness, neutral, sadness, surprise |
| **CK+ (Cohn-Kanade Plus)** | Facial expression sequences | 600+ sequences | Use Case: Facial emotion recognition and deep learning |
| **OULAD** | Structured tabular data | Not image-based | Student engagement analysis: numerical, categorical, temporal data; interactions, assessments, course info |
| **Aff-Wild** | In-the-wild video collection | 500 YouTube videos | Frame-by-frame valence and arousal (excitement) labels |
| **AM-FED** | Webcam video frames | ~250 videos | Frame-by-frame facial annotations; limited variation in head pose |
| **DAiSEE** | Videos of students in learning environment | Not specified | Frame-by-frame annotations for engagement, frustration, confusion, boredom |

| | | | |
|---|---|---|---|
| **Kaur dataset** | Videos in simulated educational setup | ~200 videos from ~80 participants | Single engagement label per video; learners watched instructional content |
| **Proposed dataset** | Facial images | 16000 labeled images | Hierarchical organization; balanced distribution, with 1,000 images per folder; enabling robust analysis across age/ gender |

**Table 1:** Available Dataset Comparison

## 3. Materials and Methods

### 3.1 Experimental setup

The experimental setup for creating the Student Engagement and Disengagement Image Dataset involved capturing high-resolution images of students exhibiting typical engagement and disengagement behaviors during online learning. Prior to the dataset creation, all participants and, for those under 18, their parents, were informed about the study's purpose and the nature of the dataset. Those who expressed interest in participating completed a consent form, agreeing to allow their facial images to be included in a globally accessible dataset. These consent forms are provided in the supporting documents.

The images were collected at Vishwakarma University, Pune (18°27′34.8″N 73°53′01.1″E) over a period from July 2024 to October 2024. A total of 17 participants, aged between 5-21+ years, contributed to the dataset. The participant's demonstrated common engagement and disengagement behaviors typically observed during e-learning sessions, such as attentive listening, focused facial expressions for engagement, and signs of inattention or distraction for disengagement. Additionally, the researchers provided instructions to the participants to exhibit specific physical behaviors that reflect these states, ensuring a comprehensive representation of various engagement levels.

### 3.2 Data Collection Setup

The image dataset for student engagement and disengagement was captured using a combination of mobile devices and an external webcam to ensure high-quality, diverse visual data. Three different devices were employed during the data collection process: the OnePlus 7 Pro (Model GM1911), the realme UI 5.0 (Model RMX3660), and the Lenovo 300 FHD Webcam (Model GXC1B34793). The OnePlus 7 Pro and realme UI 5.0 smartphones provided high-resolution images, while the Lenovo 300 FHD Webcam, equipped with a 2.1 MP CMOS sensor and capable of capturing Full HD (1920 x 1080) resolution, and was used for additional visual data collection, enhancing the variety of image sources in the dataset (Figure 1).

The captured images were systematically organized into folders, with 1,000 images per folder, ensuring a structured arrangement for easy access and analysis. The dataset as a whole contains a total of 16,000 images, providing a comprehensive set of visuals depicting various student engagement and disengagement behaviors during online learning sessions. This diverse collection of images supports a wide range of research applications, particularly in the training of machine learning models aimed at understanding student behavior in virtual education environments.

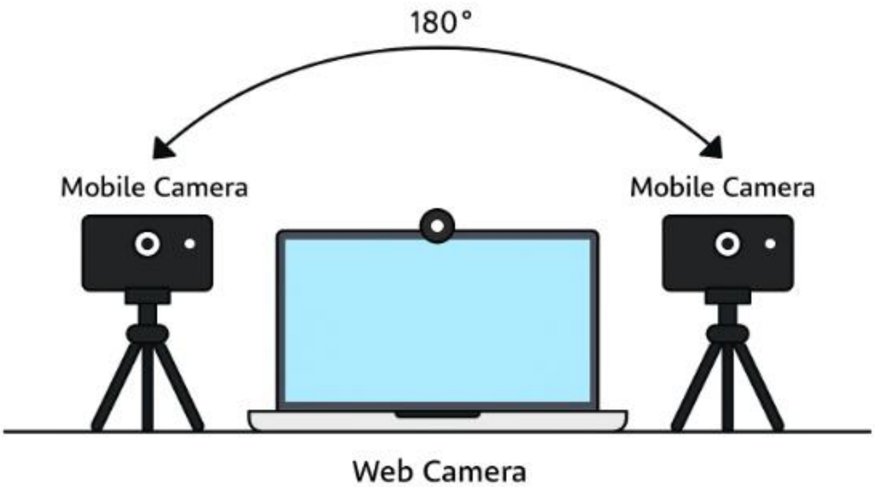


**Figure 1** – Data Collection Setup

### 3.3 Image Pre-Processing

The pre-processing of the Student Engagement and Disengagement Image Dataset was conducted to standardize the images for machine learning applications and to streamline data management. The initial step involved uniformly resizing all the captured images to 256 × 256 pixels using IrfanView software, ensuring consistency across the dataset. This resizing was crucial for aligning the images with common input dimensions

used in machine learning models. Along with resizing, the images were standardized to a resolution of 96 dpi with a 24-bit depth, maintaining high image quality while optimizing file size for computational purposes.
To further enhance the organization and manageability of the dataset, a systematic renaming process was applied to the images, ensuring a consistent naming convention across the entire collection. This step was essential for smooth integration with data processing pipelines and for simplifying the identification of files during analysis.
Following pre-processing, the images were systematically stored in hierarchical folders. The dataset was divided into two primary categories: Engagement and Disengagement, reflecting the behavioral states of the students. These two main folders were then subdivided into four age groups: 5-10 years, 11-15 years, 16-20 years, and 21+ years. To further categorize the dataset, each age group folder was divided into gender-specific subfolders, labeled as "Boy" and "Girl".
This structure resulted in a total of 16 folders, with 1,000 images in each folder, creating a balanced dataset of 16,000 images (Table 2). This organization not only supports detailed age- and gender-specific analyses but also enhances the usability of the dataset for machine learning model training and evaluation. The dataset is available with authors if anyone wish to use the dataset for machine learning purpose, they can email us. The dataset folder will be uploaded in future at zenodo if needed.

| Engagement (No. of Images) | | | | |
|---|---|---|---|---|
| Gender | **5-10 years** | **11-15 years** | **16-20 years** | **21+ years** |
| **Boy** | 1000 | 1000 | 1000 | 1000 |
| **Girl** | 1000 | 1000 | 1000 | 1000 |

| Disengagement (No. of Images) | | | | |
|---|---|---|---|---|
| Gender | **5-10 years** | **11-15 years** | **16-20 years** | **21+ years** |
| **Boy** | 1000 | 1000 | 1000 | 1000 |
| **Girl** | 1000 | 1000 | 1000 | 1000 |

**Table 2:** Structure and image count of Student Engagement Disengagement dataset.

## 4. Results and Discussion

The Student Engagement and Disengagement Image Dataset was successfully developed, comprising a total of 16,000 labeled images systematically categorized by engagement state, age group, and gender. The dataset is structured into two primary categories—Engagement and Disengagement—and further subdivided into four age ranges (5–10 years, 11–15 years, 16–20 years, and 21+ years), with gender-based folders for each group. This hierarchical structure enhances usability for machine learning applications by providing balanced and organized subsets for training and evaluation.

### 4.1 Dataset Characteristics

The dataset's systematic design ensures equal representation, with 1,000 images in each subfolder, thereby reducing class imbalance issues commonly observed in behavioral datasets. Images capture diverse behavioral cues such as focused attention, eye contact, distracted gazes, and signs of disinterest, which reflect real-world online learning conditions. The inclusion of participants from multiple age groups adds to its versatility, allowing researchers to explore engagement detection across different stages of education, from early schooling to university-level learners.

### 4.2 Significance for Educational Research

The availability of such a large-scale, structured dataset directly addresses the gap in labeled visual resources for student engagement studies. Existing literature often relies on small, context-specific datasets or indirect measures such as surveys and self-reports, which limit the reliability and generalizability of findings. By contrast, this dataset provides a visual, behavior-oriented benchmark that supports objective, data-driven research on online learning behaviors.
The structured nature of the dataset also facilitates the application of machine learning and deep learning approaches, enabling researchers to design and validate models for real-time engagement monitoring. Such tools can empower educators by offering automated feedback on student attentiveness, ultimately enhancing the interactivity and personalization of virtual classrooms.

### 4.3 Comparative Advantage

Unlike other publicly available educational datasets that often focus on limited contexts (e.g., classroom videos or single-age cohorts), this dataset incorporates a wide demographic range (5-21+ years) and maintains gender balance across categories. This diversity ensures broader applicability, making the dataset a reliable resource for training generalizable models rather than context-specific ones.Image dataset play important role in machine learning models (Meshram et al., 2023; Visvanathan et al., 2023; Visvanathan et al., 2024)

### 4.4 Limitations and Future Directions

While comprehensive in scale and structure, the dataset is currently limited to image-based data only. Engagement, however, is a multimodal construct influenced by facial expressions, body language, voice, and interaction patterns. Expanding the dataset to incorporate audio, eye-tracking, or physiological signals would provide a more holistic representation of student engagement.
Another limitation is the relatively small number of unique participants (17 individuals), which, despite producing a large number of images, may constrain diversity in natural variations of engagement behaviors. Future dataset expansions should aim to include a larger and more heterogeneous participant base across varied environments, devices, and lighting conditions.

### 4.5 Overall Contribution

The results highlight that the dataset provides a structured, balanced, and scalable foundation for research in educational technology. By making a large collection of labeled engagement and disengagement images available, this work enables researchers to explore new methodologies in attention detection, adaptive learning systems, and human–computer interaction in education. In doing so, it lays important groundwork for the development of intelligent systems capable of supporting personalized and effective online learning experiences.

### Acknowledgements

The authors thanks to participants who involved in this study.

### Funding statement

NA

### Competing interests

The author(s) have no competing interests to declare.

## References –

1. Gupta, A., D'Cunha, A., Awasthi, K., & Balasubramanian, V. (2016). Daisee: Towards user engagement recognition in the wild. arXiv preprint arXiv:1609.01885. https://doi.org/10.48550/arXiv.1609.01885
2. Alruwais, N., & Zakariah, M. (2023). Student-engagement detection in classroom using machine learning algorithm. Electronics, 12(3), 731. https://doi.org/10.3390/electronics12030731
3. Kaur, A., Mustafa, A., Mehta, L., & Dhall, A. (2018). Prediction and localization of student engagement in the wild. In 2018 Digital Image Computing: Techniques and Applications (DICTA) (pp. 1–8). IEEE. https://doi.org/10.1109/DICTA.2018.8615772
4. Bhardwaj, P., Gupta, P. K., Panwar, H., Siddiqui, M. K., Morales-Menendez, R., & Bhaik, A. (2021). Application of deep learning on student engagement in e-learning environments. Computers and Electrical Engineering, 93, 10727. https://doi.org/10.1016/j.compeleceng.2021.10727
5. Buono, P., De Carolis, B., D'Errico, F., Macchiarulo, N., & Palestra, G. (2023). Assessing student engagement from facial behavior in online learning. Multimedia Tools and Applications, 82(9), 12859–12877. https://doi.org/10.1007/s11042-022-14048-8
6. McDuff, D., Kaliouby, R., Senechal, T., Amr, M., Cohn, J., & Picard, R. (2013). Affectiva-MIT facial expression dataset (AM-FED): Naturalistic and spontaneous facial expressions collected. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 881–888). IEEE. https://doi.org/10.1109/CVPRW.2013.130
7. Dewan, M., Murshed, M., & Lin, F. (2019). Engagement detection in online learning: a review. Smart Learning Environments, 6(1), 1-20.
8. Kollias, D., Tzirakis, P., Nicolaou, M. A., Papaioannou, A., Zhao, G., Schuller, B., Kotsia, I., & Zafeiriou, S. (2019). Deep affect prediction in-the-wild: Aff-Wild database and challenge, deep architectures, and beyond. *International Journal of Computer Vision, 127*(6), 907–929. https://doi.org/10.1007/s11263-019-01158-4
9. Gray, J. A., & DiLoreto, M. (2016). The effects of student engagement, student satisfaction, and perceived learning in online learning environments. International Journal of Educational Leadership Preparation, 11(1), n1.
10. Hasnine, M. N., Nguyen, H. T., Tran, T. T. T., Bui, H. T., Akçapınar, G., & Ueda, H. (2023). A real-time learning analytics dashboard for automatic detection of online learners' affective states. Sensors, 23(9), 4243.

11. Martin, F., & Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. Online learning, 22(1), 205-222.
12. Meshram, V., Suryawanshi, Y., Meshram, V., & Patil, K. (2023). Addressing misclassification in deep learning: a merged net approach. Software Impacts, 17, 100525.
13. Pabba, C., & Kumar, P. (2022). An intelligent system for monitoring students' engagement in large classroom teaching through facial expression recognition. *Expert Systems*, *39*(1), e12839.
14. Schiavo, G., Cappelletti, A., & Zancanaro, M. (2014). Engagement recognition using easily detectable behavioral cues. *Intelligenza Artificiale*, *8*(2), 197-210.
15. Suryawanshi, Y., Meshram, V., Meshram, V., Patil, K., & Chumchu, P. (2023). Face mask wearing image dataset: A comprehensive benchmark for image-based face mask detection models. *Data in Brief*, *51*, 109755.
16. Suryawanshi, Y., Meshram, V., Patil, K., Testani, M., Chumchu, P., & Sharma, A. (2024). The image dataset of Indian coins: A machine learning approach for Indian currency. *Data in Brief*, *53*, 110098.
17. Suryawanshi, Y., Patil, K., & Chumchu, P. (2022). VegNet: Dataset of vegetable quality images for machine learning applications. *Data in Brief*, *45*, 108657.
18. Visvanathan, G., Patil, K., Suryawanshi, Y., & Chumchu, P. (2023). Sensor based dataset to assess the impact of urban heat island effect mitigation and indoor thermal comfort via terrace gardens. Data in Brief, 49, 109431.
19. Visvanathan, G., Patil, K., Suryawanshi, Y., Meshram, V., & Jadhav, S. (2024). Mitigating urban heat island and enhancing indoor thermal comfort using terrace garden. Scientific Reports, 14(1), 9697.