INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Development of Face Expression Recognition Model to Support Learning Feedback in Higher Education

Muhammad Firdaus Mustapha[1*], Siti Haslini Ab Hamid[2]

[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara Cawangan Kelantan, Bukit Ilmu, 18500 Machang, Kelantan, Malaysia

[2]Department of Information Technology, FH Training Center, 16800 Pasir Puteh, Kelantan Malaysia

***Corresponding Author**

DOI: https://dx.doi.org/10.47772/IJRISS.2025.91100070

Received: 14 November 2025; Accepted: 20 November 2025; Published: 29 November 2025

ABSTRACT

Face expressions offer a non-verbal channel for understanding student engagement and feedback in higher education learning environment. With the rise of affective computing, face expression recognition (FER) applications have gained attention for their ability to the recognize and respond to learners' emotional cues in real time. Nevertheless, developing a stable FER model often involves complex deep learning architectures and large-scale annotated datasets. Therefore, this study presents the development of a FER model using Google Teachable Machine (GTM) to support learning feedback in higher education. The proposed FER model can classify five categories of face expressions. A dataset comprising 600 face images was collected and divided into 85% for training and 15% for validation/testing. Model performance was evaluated using accuracy, precision, recall and F1-score metrics. The confusion matrix showed reliable performance for all face expression categories, validating the effectiveness of GTM for accessible FER model.

Keywords: Face Expression Recognition, Google Teachable Machine, Higher Education, Learning Feedback

INTRODUCTION

The emotional state of learners plays a crucial role in their ability to acquire, retain, and apply knowledge. Emotions such as frustration, boredom, interest, or confusion can impact concentration and motivation in significant ways [1]. Face expressions, as a direct and observable indicator of emotion, offer a non-verbal channel through which student feedback can be interpreted in real time [2]. Face expression is one of the most informative non-verbal cues in human communication and emotional recognition [3]. Mehrabian [4] claimed that 93% of the emotional meaning is transmitted as follows: 55% come from facial expression, 38% come from vocal expression and 7% come from verbal expression.

In the context of higher education, monitoring the evolving face expressions of students over time is crucial for gaining insights into their engagement, emotional states, and learning responses. Nevertheless, this task is difficult due to the intricate and highly variable nature of face expressions. Traditional approaches often fall short in accurately detecting and interpreting subtle facial cues, resulting in limitations in effectively monitoring student learning. The growing interest in affective computing has led to the development of systems capable of recognizing and responding to emotional cues [5]. Among such applications is Face Expression Recognition (FER), which has become increasingly relevant in educational contexts. FER provides feedback that can enhance instructional adaptation, improve learner satisfaction, and support personalized learning [6], [7]. In recent years, FER also has gained traction in applications such as education, security systems, healthcare diagnostics, and customer experience analysis [7], [8].

Developing a robust FER model often involves complex deep learning architectures and large-scale annotated datasets. One of the solutions is by applying Google Teachable Machine (GTM) to create the FER model. These challenges can be solved comprehensively specifically for educator. GTM simplify this process by offering an intuitive, no-code interface that leverages pre-trained models for transfer learning [9]. GTM provides a simple and efficient approach for developing a stable FER model that can continuously monitor students' emotional states and engagement levels in real time. This approach helps close the gap in face expression analysis by providing educators with meaningful insights into student learning behaviors, enabling timely interventions and personalized support to improve the online learning experience. GTM offers an accessible alternative, enabling non-technical users to create classification models with ease. This makes it particularly attractive for educators and researchers looking for rapid deployment in real-world learning environments. GTM accepts three types of input from users which are image, audio and pose. The inputs can be gathered via webcam, microphone or upload [9].

Therefore, this study aims to develop a FER model using GTM to support learning feedback in higher education. The desired output is building an accurate, fast model that can precisely categorize face expression into five different classes which are happiness, sadness, surprise, anger and neutral. The dataset will be obtained from an online repository. The developed model has potential applications in virtual classrooms, online assessments, and self-paced e-learning environments.

The remainder of this paper is organized as follows. Section II reviews related work on FER while section III describes the methodology of the research starting from data preparation until deployment. Section IV presents the experimental result and discussion. Finally, section 6 concludes the research together with limitations and directions for future research.

LITERATURE REVIEW

FER systems traditionally rely on feature extraction and machine learning classification. Early approaches used hand-crafted features such as Scale-Invariant Feature Transform (SIFT) and Local Binary Pattern (LBP) [10]. The rise of deep learning led to widespread use of convolutional neural network (CNN) [11]. Ko (2018) provided a review of FER technologies and noted that machine learning techniques such as CNN and Long Short-Term Memory (LSTM) are highly effective in emotion classification tasks. There are many researches related to FER have been done. For example, Whitehill et al. [12] developed an automated FER system to monitor student engagement. Their system used a CNN to predict attention levels and found a strong correlation with academic performance. Minaee et al., [13] conducted experiments on four different datasets using their proposed end-to-end deep learning framework based on an attentional convolutional network. From their experiment, by using the FERG dataset, they attained the highest accuracy rate of around 99.3% compared to using other datasets.

More interestingly, many researchers using hybrid techniques or combined several techniques for FER. In example, Abinaya et al. [14] proposed Hybrid Adaptive Kernel based Extreme Learning Machine (HAKELM) scheme on their research and achieved 95.5% of accuracy, 90.12% of sensitivity, and 95.1% specificity compared to the previous existing algorithm. Rahul et al. [15] proposed a hybrid approach for emotion recognition by combining CNN and Recurrent neural networks (RNN) and tested using three datasets. FER-13 dataset achieved 94.08% of accuracy rate meanwhile EMOTIC dataset attained 72.64% and the lowest accuracy rate is 68.10% for FERG dataset. Moreover, Kong et al. [16] introduced a real-time FER method utilizing iterative transfer learning and an Efficient Attention Network (EAN), specifically designed for edge environments with limited resources. This approach effectively addresses issues related to server overload and the risk of privacy breaches.

Developing a robust FER model often involves complex deep learning architectures and large-scale annotated datasets. Recent studies have explored using MobileNet, Visual Geometry Group (VGG), and Residual Network (ResNet) for lightweight deployment in resource-constrained environments. MobileNet has shown good trade-offs between performance and efficiency [17]. For example, Haslini et al. [18] developed android application for FER using Personal Image Classifier, where their backend engine using MobileNet. Besides, Aly [19] utilized ResNet50+CBAM+TCNs to track student engagement in online classrooms. The proposed techniques achieved accuracies of 91.86% for RAF-DB, 91.71% for FER2013, 95.85% for CK+, and 97.08% for KDEF dataset. Moreover, Huang et al. [20] used six emotion categories related to classroom teaching and learning and their results showed that MultiEmoNet achieved a classification accuracy of 91.4% on a homemade classroom student emotion dataset. Gao et al. [8] highlighted classroom expression recognition systems using spatial, channel, self-attention for teaching feedback. They constructed five category of classroom expressions and the proposed method got 88.34% of accuracy in expression recognition tasks and offers strong support for smart classrooms.

Table I summarized several current researches on FER that consist information on number of face expression, dataset and techniques. Based on the table, many researchers used custom FER dataset and also several benchmark datasets for FER research including FER2013, Radboud Faces Database (RAF-DB), and Extended Cohn-Kanade (CK+). These datasets have enabled researchers to develop models with high accuracy [13], [16]. FER2013 is a large-scale dataset collected from the internet, consisting of face images extracted from YouTube videos. It comprises of seven emotion categories: anger, disgust, fear, happiness, sadness, surprise, and neutral [19]. RAF-DB is a dataset that consists of static images of face expressions collected from online sources. It comprises a variety of emotion categories such as happiness, surprise, fear, disgust, sadness, and anger. CK + dataset consists of posed face expressions of emotions such as happiness, sadness, anger, surprise, disgust, and fear. All three datasets are popular in FER research. In term of techniques, most of the researchers using deep learning technique to classify face expression such as CNN, RNN, ResNet and YOLO. In addition, although Ekman's emotion theory is a landmark in the field of emotion recognition, its categorization may not fully capture the complexity of students' emotions in a specific classroom setting [20]. However, due to publicly available databases of students' emotions are limited [21] and student privacy issues, many researchers still used seven basic categories of face expression which are anger, disgust, fear, happiness, sadness, surprise, and neutral [19], [22].

The importance of accessible artificial intelligence (AI) tools like GTM is rising. GTM provides an intuitive way to build image classifiers without coding and widely used in education to teach machine learning (ML) concepts [23]. Carney et al., [23] emphasized several benefits of GTM, including its user-friendly interface, the absence of any requirement for coding or prior ML experience, and its potential to offer interactive tools and simplified concepts that make teaching and learning ML more accessible. In other words, it enables individuals from various backgrounds to use ML without requiring specialized knowledge or technical skills. In addition, Wong & and Fadzly [24], highlighted that while GTM does not offer deep customization, it is highly effective for rapid proof-of-concept models and training with small datasets. Such tools are becoming increasingly relevant as interest in low-code and no-code AI development grows.

Table I: Several researches on face expression recognition

Researcher, Year	Number of Expression	Dataset	Technique
[8], 2025	5	Custom FER dataset	Multi-attention fusion network (MAF-ER)
[19], 2024	7	RAF-DB, FER2013, CK + and KDEF	ResNet50+CBAM+TCNs
[22], 2024	7	FER2013, FERPlus, RAF-DB, AffectNet, real smart classroom facial expression dataset (SCFED)	Multi-scale and deep fine-grained feature attention enhancement (MDFAE)
[20], 2024	6	Custom FER dataset	Enhance YOLOv8
[18], 2022	3	Custom FER dataset	Personal Image Classifier (CNN)
[16], 2022	7	FER2013, RAF-DB	EAN
[15], 2022	7	EMOTIC, FER-13, FERG	CNN and RNN
[14], 2021	7	AT&T, YALE FACE B	HAKELM
[13], 2021	7	FER2013, CK+, JAFFE, FERG	End-to-end deep learning framework based on an attentional convolutional network

Therefore, this paper proposes development of FER model using GTM to perform multiclass student face expression classification in higher education context.

METHODOLOGY

This section consists of information related to the methodology for developing FER model using GTM which are hardware requirement, data preparation, model configuration, training process and finally model testing and validation.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Hardware and Software Requirement

To develop the FER model using GTM, a personal computer (PC) was utilized with the following specifications: an AMD Ryzen 5 4500 6-core processor, 8 GB of RAM, and Windows 10 Pro as the operating system. Google Chrome, along with a stable internet connection was used to access the GTM platform and perform the model training process.

Data Preparation

The model training process begins with data collection, where images representing different face expressions are collected from RAF-DB dataset [25] and five classes has been created: happiness, sadness, surprise, anger and neutral. Each class is represented by multiple images to improve the model's ability to generalize and recognize expressions across numerous conditions. Preprocessing involves in this step which is cleaning activity to remove blur images and also children’s images. A total of 600 images were selected and then uploaded into GTM website. Fig. 1 represents some sample images from RAF-DB dataset.

Fig. 1. Sample images of dataset

For this study, we used a single holdout method in GTM to split data for training and testing/validation. This split is standard machine learning practice to prevent overfitting. 85% of the total data is used for training (510 images), while 15% is reserved for internal testing and validation (90 images) as shown in Table II. This ratio is commonly used to balance model training and validation [26].

TABLE II: DATA SPLITTING USING HOLDOUT METHOD

Class	**Training (85%)**	**Testing / Validation(15%)**	**Total**
Happiness	102	18	120
Sadness	102	18	120
Surprise	102	18	120
Anger	102	18	120
Neutral	102	18	120
Total	510	90	600

Model Configuration

In supervised learning, particularly deep learning, several hyperparameters control the learning process which are epoch, learning rate and batch size. Table III shows the parameter setting for training the FER model using GTM. These parameter values are usually used in small image classification tasks using MobileNet [17], [27].

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Table III: Training parameter

Parameter	Value
Epoch	50
Learning rate	0.001
Batch size	16

Epochs are defined as one complete forward and backward pass of all training samples. Each epoch allows the model to learn and refine its parameters through weight updates. Too few epochs may result in underfitting, while too many may lead to overfitting [26]. In this study, suitable value for epoch is 50.

Learning rate controls the size of the steps the model takes to update its weights during the training process. A high learning rate may lead to fast convergence but overshooting, while a low rate results in slower, more stable learning [28]. In this study, a low learning rate, 0.001 is used to make sure the stable learning process for creating FER model.

Batch size refers to the number of data samples processed before the model's parameters are updated. Smaller batches provide more updates but may be noisy, while larger batches are computationally efficient but less flexible [29]. For this study, a smaller value of batch size which is 16 is applied for training purpose.

Training Process

GTM uses Tensorflow.js, an open source machine learning library in JavaScript to train and run the training result in a model in a web browser [30]. GTM also leverages the concept of transfer learning where instead of training a neural network from scratch, it uses a pre-trained MobileNet model. Transfer learning has proven effective, allowing pre-trained models to be fine-tuned on emotion classification tasks. MobileNet is a CNN with a smaller model size with less trainable parameter and calculation amount [18].

Therefore, the training process in GTM consists of the following steps:

1. The uploaded face expression images were pre-processed which included image resizing and normalization. Images were automatically resized to 224x224 pixels by GTM.
2. A CNN backbone is used internally to learn distinguishing features of each face expression.
3. The model is trained using transfer learning, where a pre-trained model (MobileNet) is fine-tuned on the new dataset.
4. The model continues training for a set number of epochs until it achieves satisfactory accuracy, evaluated via loss and accuracy metrics.

Model Testing and Validation

Once the training is completed, the model is validated using the test dataset (15%). This helps assess the model's generalization ability. A confusion matrix as shown in Fig. 2 is a table that summarizes the classification results, showing the number of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) for each class. Confusion matrix displays the number of predictions made by the algorithm compared to the actual true values in the test dataset. The Y axis (Class) denotes to the class of sample, while the X axis denotes to the predicted class.

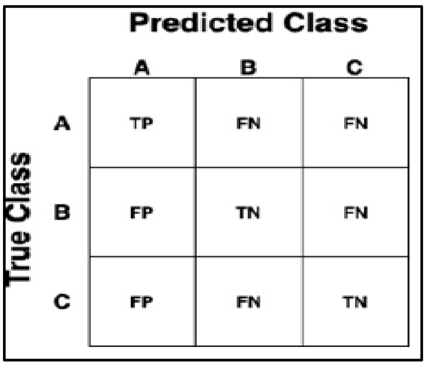


Fig. 2 Multiclass classification in confusion matrix [31]

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Moreover, based on the result from confusion matrix, four key metrics that commonly utilized to assess the model's effectiveness are Accuracy, Precision, Recall, and F1-Score [32], [33]. Accuracy refers to overall correctness of the model or the proportion of correctly classified face expressions to the total number of expressions [19]. Precision measures how many predicted positives are actually correct. Recall, also known as Sensitivity measures how many actual positives were correctly predicted. F1-Score is the harmonic mean of Precision and Recall, providing a balanced assessment of the system's performance. Therefore, the formula for Accuracy, Precision, Recall, and F1-Score are displayed in Eq.1, Eq.2, Eq.3 and Eq.4 respectively. TP denotes True Positives (correctly recognized expressions), TN signifies True Negatives (correctly ignored expressions), FP represents False Positives (incorrectly recognized expressions), and FN indicates False Negatives (incorrectly ignored expressions) [19].

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = 2 \text{ x } \frac{Precision \text{ X } Recall}{Precision+Recall} \quad (4)$$

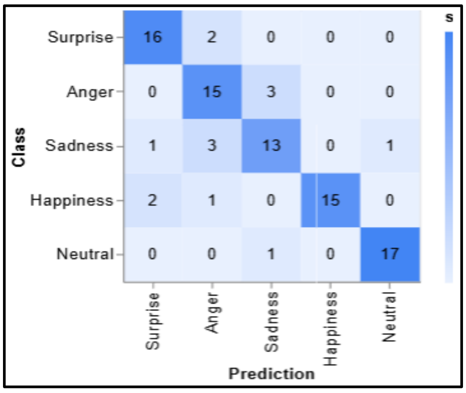
RESULT AND DISCUSSION

Table IV shows the accuracy result per class. Based on the table, the highest accuracy result is Neutral (0.86 or 86%) followed by Surprise (0.89 or 89%). Both Happiness and Anger have the same accuracy of 0.83 or 83%. The lowest accuracy class is Sadness (0.72 or 72%). This result shows that Neutral face expression is the easiest to be recognized compared to Sadness face expression that is the most difficult to be identified.

TABLE IV: ACCURACY PER CLASS

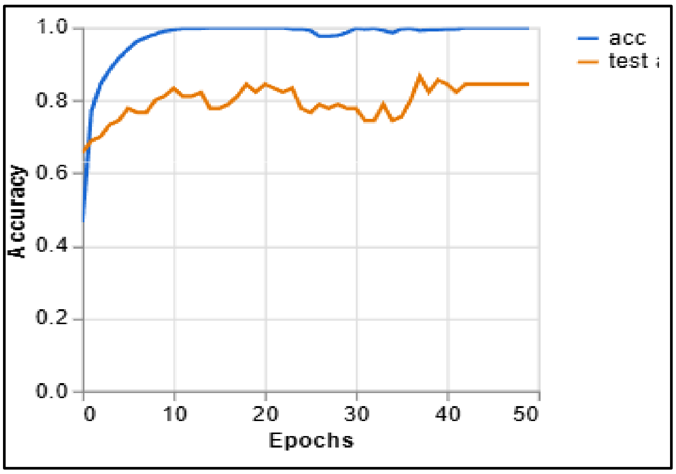
Class	Accuracy (%)	#Samples
Happiness	83	18
Sadness	72	18
Surprise	89	18
Anger	83	18
Neutral	94	18

Fig. 3 shows the result for confusion matrix, accuracy per epoch and loss per epoch from GTM internal analysis. Based on the confusion matrix in Fig. 3(a), in general result, the FER model has performed strongly with most predictions along the diagonal are correct classification. Misclassifications occur in the FER model but are relatively small in number. The best recognized classes are Neutral and Surprise while the most confused class in Sadness. In example, class Neutral is very high accuracy because this expression is consistently recognized with one misclassified as Sadness.

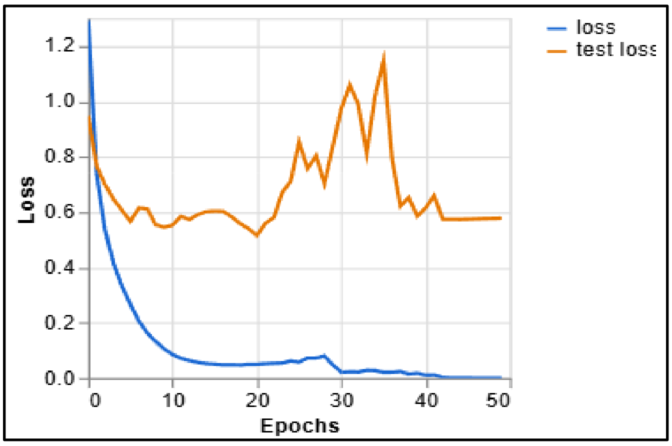


(a) Confusion Matrix

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025



(b) Accuracy per epoch



(c) Loss per epoch

Fig. 3 Result for confusion matrix, accuracy per epoch and loss per epoch

Based on confusion matrix in Fig. 3(a), the value of accuracy, precision, recall and F1-Score can be calculated. Table V displays the summary result for precision, recall and F1-Score for each class. Based on the calculation, the FER model achieved an overall accuracy of 84.44%, demonstrating good performance for five classes. The model showed its strongest results for Neutral and Happiness, both achieving high precision and F1-scores. In contrast, Anger and Sadness were more challenging for the model, with lower precision and recall due to confusion between the two classes. This is expected in FER studies, as negative emotions often share similar facial cues. Despite this, the model still reached acceptable performance levels for these classes.

TABLE V: RESULT FOR PRECISION, RECALL AND F1-SCORE

Class	Precision (%)	Recall (%)	F1-score (%)
Happiness	100	83.33	90.91
Sadness	76.47	72.22	74.29
Surprise	84.21	88.89	86.49
Anger	71.43	83.33	76.92
Neutral	94.44	94.44	94.44

The result for accuracy per epoch is depicted in Fig. 3(b). Accuracy reflects how correctly the prediction model performs. It represents the percentage of correct classifications made during training. A perfect prediction yields an accuracy of one, while any errors result in a value less than one. A good accuracy is shown by the intercept line between actual accuracy and the test accuracy. Moreover, the result for lost per epoch is shown in Fig. 3(c). Loss per epoch represents the number of errors during each training cycle (epoch); generally, a lower loss value indicates better model performance.

Moreover, the performance comparison between our proposed FER model and several other researchers using the RAF-DB dataset [16], [19], [22] is shown in Table VI. The proposed FER model's accuracy is comparable to more complex deep learning-based models, but achieved using a simpler tool with no coding involved and smaller size dataset.

TABLE VI: PERFORMANCE COMPARISON OF OTHER RESEARCHERS ON RAF-DB DATASET

Researcher	Method	Accuracy (%)
[22]	Multi-Scale and Deep Fine-Grained Feature Attention	92.93
[16]	EAN	85.30
[19]	ResNet50, CBAM, and TCNs	91.86
Proposed	GTM	84.44

Furthermore, each face expression has the specific learning feedback interpretation as summarized in Table VII. For example, expression Happiness means positive learning engagement, learner understands the learning content or learner feels satisfaction and motivated. Sadness expression interprets that learner feels boring with the content, learner needs emotional support or fatigue during the classroom session.

TABLE VII: RELATIONSHIP BETWEEN FACE EXPRESSION AND LEARNING FEEDBACK

Class	Learning feedback interpretation
Happiness	Engagement, understanding, satisfaction
Sadness	Boredom, demotivation, fatigue
Surprise	Attention, curiosity, cognitive shift
Anger	Frustration, cognitive overload
Neutral	Focused attention, passive engagement

This FER model can be integrated into online learning platforms or physical classrooms via webcams to provide real-time emotion monitoring in higher education environment. For instance, if a learner shows Sadness emotions over time, the system could prompt a motivational message or offer help materials. Educators could use the insights to adjust content delivery dynamically.

CONCLUSION

This research demonstrates the successful development of a multiclass face expression classification model using GTM. The proposed model reliably distinguishes between five expressions. With an overall accuracy of 84.44%, this model is effective in distinguishing key facial expressions for multi-class setup. The proposed FER model can be exported in multiple formats such as TensorFlow.js (for web-based deployment), TensorFlow Lite (for mobile and embedded systems) and downloadable Keras model (for further development). The FER model can be integrated into applications or systems that provide real-time learning feedback, classroom monitoring, or affective computing systems, making it valuable in educational technology research.

One of the primary contributions of this research is the use of GTM to simplify FER model training. This approach enables educators to create custom feedback models without any programming knowledge, low setup requirement, and intuitive interface making it a suitable solution for real-world educational applications.

Some limitations of the model include limited generalization across diverse datasets and model overfitting due to small dataset. Future work will focus on expanding the model using larger datasets with real student face expressions and embedding the model in Learning Management Systems (LMS).

REFERENCES

1. R. Pekrun, "The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice," Educ. Psychol. Rev., vol. 18, no. 4, pp. 315–341, 2006.

2. P. Ekman and W. V Friesen, “Constants across cultures in the face and emotion.,” Journal of Personality and Social Psychology, vol. 17, no. 2. American Psychological Association, US, pp. 124–129, 1971.
3. Y. Huang, F. Chen, S. Lv, and X. Wang, “Facial Expression Recognition: A Survey,” Symmetry (Basel)., vol. 11, no. 10, 2019.
4. A. Mehrabian, “Communication without words,” in Psychology Today, 1968, pp. 51–52.
5. R. W. Picard, Affective Computing. MIT Press, 2000.
6. R. A. Calvo and S. D’Mello, “Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications,” IEEE Trans. Affect. Comput., vol. 1, no. 1, pp. 18–37, 2010.
7. S. D’mello and A. Graesser, “AutoTutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back,” ACM Trans. Interact. Intell. Syst., vol. 2, no. 4, Jan. 2013.
8. Y. Gao, L. Zhou, and J. He, “Classroom Expression Recognition Based on Deep Learning,” Appl. Sci., vol. 15, no. 1, 2025.
9. “Google Teachable Machine,” 2025. [Online]. Available: https://teachablemachine.withgoogle.com/. [Accessed: 20-May-2025].
10. C. Shan, S. Gong, and P. W. McOwan, “Facial expression recognition based on Local Binary Patterns: A comprehensive study,” Image Vis. Comput., vol. 27, no. 6, pp. 803–816, 2009.
11. A. Mollahosseini, D. Chan, and M. H. Mahoor, “Going deeper in facial expression recognition using deep neural networks,” in 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016, pp. 1–10.
12. J. Whitehill, Z. Serpell, Y. C. Lin, A. Foster, and J. R. Movellan, “The faces of engagement: Automatic recognition of student engagement from facial expressions,” IEEE Trans. Affect. Comput., vol. 5, no. 1, pp. 86–98, 2014.
13. S. Minaee, M. Minaei, and A. Abdolrashidi, “Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network,” Sensors, vol. 21, no. 9, 2021.
14. D. Abinaya, C. Priyanka, M. Rocky Stefinjain, G. K. D. Prasanna Venkatesan, and S. Kamalraj, “Classification of Facial Expression Recognition using Machine Learning Algorithms,” J. Phys. Conf. Ser., vol. 1937, no. 1, p. 12001, Jun. 2021.
15. M. Rahul, N. Tiwari, R. Shukla, D. Tyagi, and V. Yadav, “A New Hybrid Approach for Efficient Emotion Recognition using Deep Learning,” Int. J. Electr. Electron. Res., vol. 10, no. 1, pp. 18–22, 2022.
16. Y. Kong, S. Zhang, K. Zhang, Q. Ni, and J. Han, “Real-time facial expression recognition based on iterative transfer learning and efficient attention network,” IET Image Process., vol. 16, no. 6, pp. 1694–1708, 2022.
17. A. G. Howard et al., “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” CoRR, vol. abs/1704.0, 2017.
18. S. Haslini, A. Hamid, N. Mustapa, and M. F. Mustapha, “An Android Application for Facial Expression Recognition Using Deep Learning,” J. Appl. Math. Comput. Intell., vol. 11, no. 2, pp. 505–520, 2022.
19. M. Aly, “Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model,” Multimed. Tools Appl., vol. 84, no. 13, pp. 12575–12614, 2024.
20. Y. Huang, W. Deng, and T. Xu, “A Study of Potential Applications of Student Emotion Recognition in Primary and Secondary Classrooms,” Appl. Sci., vol. 14, no. 23, 2024.
21. B. Fang, X. Li, G. Han, and J. He, “Facial Expression Recognition in Educational Research From the Perspective of Machine Learning: A Systematic Review,” IEEE Access, vol. 11, pp. 112060–112074, 2023.
22. Z. Shou et al., “A Student Facial Expression Recognition Model Based on Multi-Scale and Deep Fine-Grained Feature Attention Enhancement,” Sensors, vol. 24, no. 20, 2024.
23. M. Carney et al., “Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification,” in Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–8.
24. J. J. N. Wong and N. and Fadzly, “Development of species recognition models using Google teachable machine on shorebirds and waterbirds,” J. Taibah Univ. Sci., vol. 16, no. 1, pp. 1096–1111, Dec. 2022.
25. Dev-ShuvoAlok, “RAF-DB DATASET,” 2023. [Online]. Available: https://www.kaggle.com/datasets/shuvoalok/raf-db-dataset?select=DATASET. [Accessed: 01-May-2025].

26. I. Goodfellow, Y. Bengio, and C. Aaron, Deep Learning. MIT Press, 2016.
27. M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proceedings of the 36th International Conference on Machine Learning, 2019, vol. 97, pp. 6105–6114.
28. L. Bottou, "Stochastic Gradient Descent Tricks," in Neural Networks: Tricks of the Trade: Second Edition, G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 421–436.
29. D. Masters and C. Luschi, "Revisiting Small Batch Training for Deep Neural Networks," CoRR, vol. abs/1804.0, 2018.
30. E. Malahina, R. Hadjon, and F. Bisilisin, "Teachable Machine: Real-Time Attendance of Students Based on Open Source System," IJICS (International J. Informatics Comput. Sci., vol. 6, no. 3, p. 140−146, 2022.
31. K.-A. Tait et al., "Intrusion Detection using Machine Learning Techniques: An Experimental Comparison," in 2021 International Congress of Advanced Technology and Engineering (ICOTEN), 2021, pp. 1–10.
32. A. S. A. Mohammed Aly, "EMU-Net: Automatic Brain Tumor Segmentation and Classification Using Efficient Modified U-Net," Comput. Mater. \& Contin., vol. 77, no. 1, pp. 557–582, 2023.
33. M. H. Behiry and M. Aly, "Cyberattack detection in wireless sensor networks using a hybrid feature reduction technique with AI and machine learning methods," J. Big Data, vol. 11, no. 1, p. 16, 2024.