FMDB *Transactions on Sustainable Computer Letters*

Multimodal Facial Emotional and Micro Expression Analysis in Online Educational Environment Using CASME II and Image Feature Stabilization Algorithm

R. Angeline[1], A. Alice Nithya[2,*]

[1,2]Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, India.
angelinr1@srmist.edu.in[1], alicenia@srmist.edu.in[2]

Abstract: The digital education revolution affords students and teachers worldwide unprecedented opportunities. Student involvement, comprehension, and academic integrity can be challenging in online classrooms. Online education requires more effective ways to track and enhance student engagement. Teachers evaluate student attention and understanding with images and brief exchanges in traditional classrooms. Student-specific instruction in online classes is impossible without these identifiers. Online students may show emotion through microexpressions. This rapid, involuntary expression shows students' immediate emotions and attentiveness. In movies and blurred images, CASME II's facial emotional recognition shows student interest. Microexpressions help teachers identify distracted students and adjust their teaching strategies. Real-time feedback helped teachers better understand their students and enhance online education. Understanding microexpressions can change schooling. A teacher who can detect students' mild dissatisfaction, bewilderment, or interest. They can instantly support kids and establish a compassionate learning environment. Cheating and academic dishonesty are detected to ensure fair assessments and the value of the assessment. Teachers track student improvement with CASME II's microexpression analysis dataset. Similar to a crystal ball that predicts educational success. CASME II detects micro-expressions 43.51% accurately. To increase the accuracy and multimodal facial expression and micro expression analysis for 5-class classification to 84% Image Feature Stabilisation Algorithm accuracy, so that every student can succeed.

Keywords: Student Engagement; Involuntary Expression; Micro Expression; Emotional Language; Online Education Grows; CASME II Dataset; Visuals and Brief Exchanges.

Received on: 01/10/2024, **Revised on:** 05/12/2024, **Accepted on:** 09/01/2025, **Published on:** 03/06/2025

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSCL

DOI: https://doi.org/10.69888/FTSCL.2025.000424

Cite as: R. Angeline and A. A. Nithya, "Multimodal Facial Emotional and Micro Expression Analysis in Online Educational Environment Using CASME II and Image Feature Stabilization Algorithm," *FMDB Transactions on Sustainable Computer Letters*, vol. 3, no. 2, pp. 85–96, 2025.

1. Introduction

Identifying emotions among online learners is difficult because emotions are complex psychological processes that occur nonverbally. Similarly, a microexpression is a momentary facial emotion that happens in an instant. It occurs when someone

*Corresponding author.

experiences an emotional reaction that is both voluntary and involuntary simultaneously, often in opposition to one another. These momentary facial expressions reveal an individual's genuine emotions as the amygdala responds to stimuli, even when the person attempts to conceal them with a false emotional reaction [1]. Detecting microexpressions can be challenging, but analysing still images and videos can enhance the ability to perceive them [2]. Micro expressions are generally categorised according to their expression alterations and can be grouped into three main categories:

- **Simulated Expressions:** These microexpressions lack a genuine emotional basis and are often deliberately feigned.
- **Neutralised Expressions:** In this case, a genuine emotional expression is consciously suppressed, resulting in a neutral facial expression.
- **Masked Expressions:** Here, a fraudulent or fake expression completely conceals a genuine emotion.

In the digital age, online education has become a transformative force, breaking down geographical barriers and making learning accessible to people worldwide. However, this change in the educational environment brings unique challenges. One of the biggest challenges is assessing and ensuring student engagement and mental well-being in virtual classrooms. As educators strive to recreate the interactive dynamics of traditional face-to-face education, a revolutionary technology called microexpression recognition is emerging as a breakthrough solution. Microexpressions — the momentary, unconscious facial expressions that reveal our true emotions — have long fascinated psychologists and behavioural scientists. These brief facial expressions, often lasting less than a second, reveal a wealth of information about a person's emotional state and level of engagement. In the field of online education, this fascinating area of human behaviour will play a key role in reshaping the way we teach and learn. This comprehensive study examines the application and impact of integrating microexpression recognition systems with multimodal facial expression analysis in virtual classroom environments.

Explore how this technology is revolutionising the online education landscape by leveraging the Image feature stabilisation algorithm in the CASME II dataset. As the digital realm becomes a new frontier for learning, understanding, and utilising microexpressions will become an essential tool for educators and institutions alike. In this article, we embark on a journey to uncover the possibilities and opportunities of micro-expression recognition in the context of online education. We are considering a wide range of uses for it, from assessing student engagement and emotional state to providing real-time feedback and early intervention. Additionally, we are exploring ways to use microexpression analysis to create more personalised and effective online learning experiences. At a time when student success and participation are paramount, this paper proposes that microexpression recognition systems, along with multimodal facial recognition, can overcome the limitations of distance and digital interfaces, serving as a bridge between teachers and students. By leveraging the insights and capabilities of this cutting-edge technology, we aim to usher in a new era of online education that is emotionally intelligent, adaptable, and conducive to academic excellence.

2. Related Work

In daily life, various emotions can be recognised through e-technology, such as e-learning, e-health, and robotics. Due to the availability of various huge datasets, the research area expanded from unimodal to multimodal recognition. Computer scientists have expanded the field of multimodal facial emotion analysis by utilising various machine learning techniques and algorithms [3]. In speech emotion classification, feature and classifier-level fusion achieves the highest accuracy on the TESS dataset [4]. It primarily utilises three features—spectrogram, mel-frequency cepstral coefficients, and mel-spectrogram—as one-dimensional input vectors for deep neural networks and convolutional neural networks in speech emotion classification. In both classification methods, the combination of features outperforms the individual features. Micro-expressions are a special form of facial expression that reveals or conceals true emotions. Micro-expressions are so brief and delicate that they are either difficult or impossible for the average individual to see with the eye. The research was conducted with a limited number of data points, comprising three 2D and one 4D method for 4D micro-expressions [5].

Weekly supervised deep emotion detection can be trained on static images rather than expensive temporal videos [6]. The static images of a single person were used to detect the timing of facial emotion changes with their proposed framework across three video datasets: MMI, CASME II, and YouTube-ECD. In the medical field, microexpressions play a major role in detecting patients’ true emotions, enabling proper treatment [7]. LBP-TOP and SVM were used for basic evaluation on the CASME II dataset, achieving an accuracy of 61.43%. Two coders coded the Action Unit (AU), and their reliability was 0.846. This paper reported around 3,000 facial emotion samples from 247 micro-expression images of 26 participants. Among all the reference papers, blurred images cannot be accurately detected in the CASME II dataset for human facial emotions. To improve facial emotion detection among online learners, normal facial emotional images combined with microexpressions, along with the use of an Image stabilisation algorithm, increased the accuracy to 84%. The remainder of the paper outlines the proposed methods, the detection technologies employed, the CASME II dataset, the experimental results, applications and limitations, the conclusion, and future work.

3. Proposed Methods and Technologies Used

There is a wide range of methods for microexpression detection. In this paper, a deep learning algorithm, LSTM [8]; [9], is considered, as deep learning provides significantly better accuracy and more efficient results than regular Machine Learning Models. These neural networks can remember previous input and adjust the weights accordingly. Also, the Image Feature Stabilisation Algorithm undergoes two steps of (i) image fusion and (ii) differential image fusion. An image feature operator is added to the input image, referred to as positive image feature stabilisation, and subtracted from it, referred to as negative image feature stabilisation [10]. A larger dataset, CASME II [11], which contains all emotions, is considered for training. It is grouped into seven video categories: happiness, surprise, sadness, fear, repression, disgust, and others. This dataset is provided, and an ADL algorithm and a neural network are trained on it. The data is split into a test and a training set, and validation is used to evaluate the Network's accuracy. Once validation and training are complete, the network's weights will be adjusted to maximise accuracy at each epoch. Once the entire phase is completed, the model can be exported. This exported model can be used to predict the output emotion of new, unseen inputs. Figure 1 illustrates the flow of the input video into facial image detection using microexpressions.

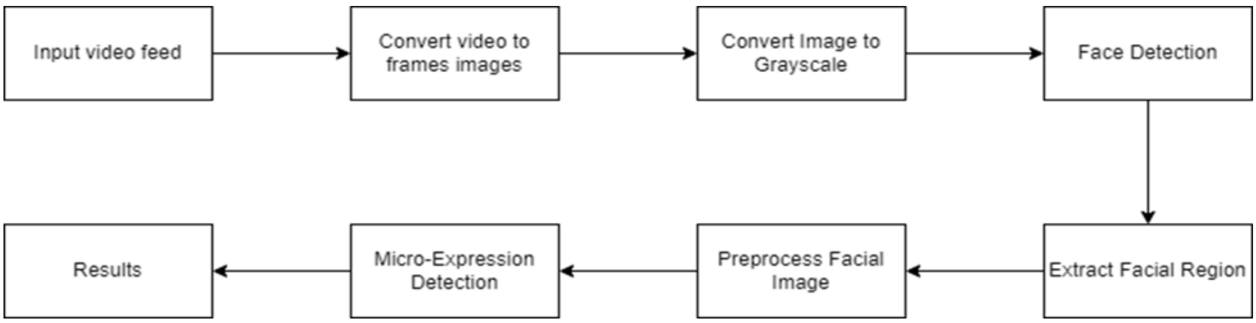

Figure 1: Facial emotion detection using micro expression

3.1. CASME II Dataset

A robust, automated micro-expression recognition system has the potential for widespread applications in national security, police interrogations, and clinical diagnoses. To develop such a system, the availability of high-quality databases with sufficient training samples is critical. However, there is a current lack of such databases. In response to this need, we conducted a comprehensive review of existing micro-expression databases and subsequently developed an enhanced database, CASME II. This new database offers superior temporal resolution at 200 frames per second (fps) and higher spatial resolution — approximately 280 x 340 pixels — specifically covering the facial area. Microexpressions were elicited from participants in a well-controlled laboratory setting, employing a rigorous experimental design and standardised lighting conditions. From nearly 3000 facial movements captured, we meticulously selected 247 micro-expressions, each labelled with the corresponding action units (AUs).

For the initial evaluation, we utilised Local Binary Pattern on Three Orthogonal Planes (LBP-TOP) for feature extraction and Support Vector Machines (SVM) for classification. Our evaluation using the leave-one-subject-out cross-validation method yielded an accuracy of 63.41% for 5-class classification. These unique features characterise the CASME II database:

- **Spontaneous and Dynamic Micro-Expressions:** The dataset contains samples of spontaneous and dynamic micro-expressions. These are very brief facial expressions that can reveal a person's true emotions, often occurring involuntarily and quickly.
- **Baseline frames:** Baseline frames, typically in a neutral facial expression, are captured before and after each microexpression. These frames serve as reference points for evaluating different detection algorithms.
- **High Temporal Resolution:** The recordings are captured at 200 frames per second (fps), enabling precise analysis of microexpressions.
- **Face Resolution:** The videos have a relatively high face resolution, with frames measuring 280 x 340 pixels. This higher resolution can aid in the analysis of microexpressions.
- **Micro expression labelling:** The labelling of micro expressions is based on the FACS (Facial Action Coding System) investigator’s guide and Yan et al.’s findings from 2013. This suggests a more detailed and specific approach to labelling microexpressions than traditional facial expression categorisation.

- **Proper Illumination:** The recordings have consistent, well-balanced illumination, avoiding flickering and minimising highlight regions on the face. This ensures that facial expressions are clearly visible and not affected by lighting.
- **Unequal Distribution of Categories:** Certain types of facial expressions are challenging to elicit in laboratory settings, resulting in an uneven distribution of samples across different categories. For example, there are 60 disgust samples but only seven sadness samples. This might be due to the natural rarity of certain micro-expressions.
- **CASME II Dataset:** The dataset is referred to as CASME II, which provides five classes of micro-expressions. These classes are likely based on the specific categories or emotional states captured in the dataset.

Some of the commonly available technologies are used in:

- Facial Expression Databases
- Criminal detection
- Suspect identification
- Person Identification
- Security

Here, the student's activities can be directly monitored from webcam footage. The input video is split into frames, and face detection is performed in each frame. This is done so that only the pixels covering the face can be cropped for further microemotion classification. The Fully Connected Long Short-Term Memory (FC_LSTM) neural network is employed to enhance accuracy and detection. Figure 2 shows a Long Short-Term Memory (LSTM) network, a type of recurrent neural network specifically designed to address the vanishing gradient problem commonly encountered in traditional RNNs. Its notable advantage over other RNNs, hidden Markov models, and sequence learning methods lies in its ability to handle varying gap lengths with relative insensitivity. The widely used open-source Python library TensorFlow is used to create and train the model.

A typical LSTM unit consists of a cell, an input gate, an output gate, and a forget gate. The cell is responsible for retaining values over arbitrary time intervals, while the three gates regulate the flow of information into and out of the cell. The forget gates determine which information from the previous state to retain by assigning a value between 0 and 1, where 1 indicates retention and 0 indicates discard. Similarly, the input gates determine which new information should be stored in the current state, using the same mechanism as the forget gates. The output gates control which information from the current state is output by assigning a value between 0 and 1, taking into account both the previous and current states. By selectively outputting relevant information from the current state, the LSTM network can maintain useful long-term dependencies, enabling accurate predictions in both the present and future time steps.

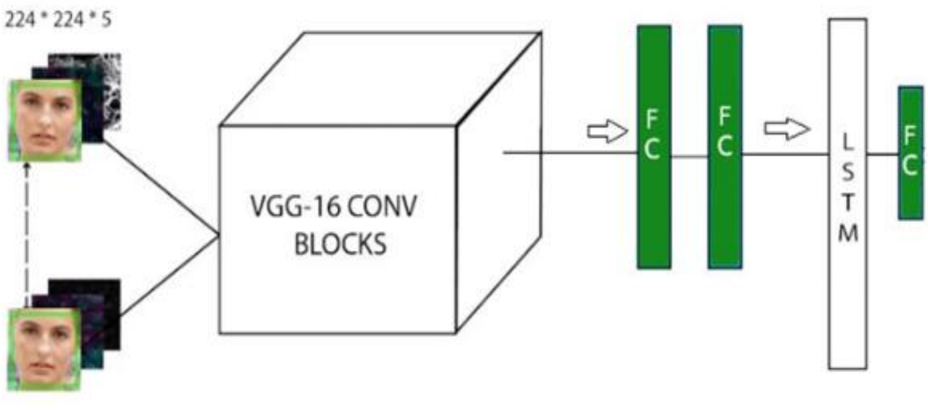

Figure 2: Enriched long-term recurrent convolutional network for facial micro-expression recognition

Figure 3 illustrates that facial features, including happiness, surprise, sadness, fear, disgust, and others, are extracted, processed, and provided to the LSTM network. Depending on the activation function and the weights, each neuron corresponding to the emotion is triggered and passed through the Image Feature Stabilisation algorithm for image fusion to retrieve positive and negative image values [10]. Positive values are added to the output phase, and negative values are removed. For real-time videos, images are captured, extracted, and processed, and the resultant model is trained for 1000 epochs. The microexpression of the learner’s image is detected and analysed by a deep learning LSTM algorithm with an image feature stabilisation algorithm.

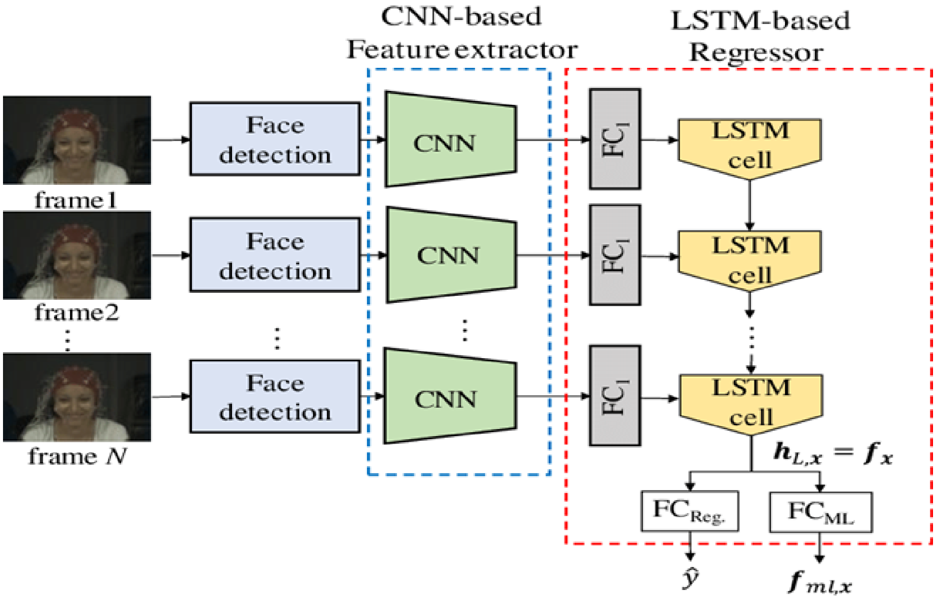

Figure 3: CNN-based feature extractor and LSTM-based repressor

4. Facial Emotion Detection vs. Micro Expression Detection

Microexpression detection is a powerful tool for assessing students' attention and emotional states in online classes. Its precision, real-time feedback, nuanced assessment, and ability to uncover concealed emotions make it superior to conventional emotion-detection methods, ultimately leading to improved learning outcomes. By capturing fleeting, subtle facial expressions, microexpression analysis provides educators with immediate feedback on students' engagement and comprehension, enabling them to make swift adjustments to their teaching methods. This capability surpasses traditional emotion detection, which provides aggregated and delayed emotional data. Furthermore, microexpressions reveal a broader spectrum of nuanced emotions, from interest to confusion, providing a granular understanding that traditional emotion categorisations lack. It also penetrates emotional concealment—a common occurrence in online learning—helping educators address concealed emotions and provide tailored support. Additionally, microexpression detection serves as a valuable tool for upholding academic integrity and identifying instances of cheating in online assessments. This precise, real-time, and insightful approach to student engagement and emotional states promises improved learning outcomes, making it a superior choice for detecting attention in online classes.

4.1. Facial Emotion Detection

Learner’s emotions can be detected using both speech and visual modalities. Emotional Detection is the process of identifying human emotion from a given facial image/video. This enables a machine to detect current emotion from facial features. Test cases were created for the labels listed in Figure 4.

```
LABELS = {
    0: "Angry",
    1: "Disgust",
    2: "Fear",
    3: "Happy",
    4: "Sad",
    5: "Surprise",
    6: "Neutral"
}
```

Figure 4: Labels used for learners’ facial emotions

The model was trained for 50 epochs with a batch size of 64 using TensorFlow and a neural network, achieving a final accuracy of 0.4351, which is much lower than the accuracy shown in Figure 5. This is far less accurate compared to microexpression detection. Microexpression detection has higher accuracy because it utilises higher-resolution images, where facial features are more readily detectable.

```
racy: 0.4377
Epoch 41/50
24402/24402 [==============================] - 3s 120us/sample - loss: 1.2470 - accuracy: 0.5266 - val_loss: 1.5504 - val_accu
racy: 0.4247
Epoch 42/50
24402/24402 [==============================] - 3s 120us/sample - loss: 1.2362 - accuracy: 0.5283 - val_loss: 1.5026 - val_accu
racy: 0.4344
Epoch 43/50
24402/24402 [==============================] - 3s 123us/sample - loss: 1.2277 - accuracy: 0.5317 - val_loss: 1.5371 - val_accu
racy: 0.4263
Epoch 44/50
24402/24402 [==============================] - 3s 119us/sample - loss: 1.2255 - accuracy: 0.5354 - val_loss: 1.5183 - val_accu
racy: 0.4363
Epoch 45/50
24402/24402 [==============================] - 3s 119us/sample - loss: 1.2113 - accuracy: 0.5404 - val_loss: 1.6092 - val_accu
racy: 0.4251
Epoch 46/50
24402/24402 [==============================] - 3s 121us/sample - loss: 1.2118 - accuracy: 0.5392 - val_loss: 1.5155 - val_accu
racy: 0.4332
Epoch 47/50
24402/24402 [==============================] - 3s 123us/sample - loss: 1.1985 - accuracy: 0.5491 - val_loss: 1.5454 - val_accu
racy: 0.4305
Epoch 48/50
24402/24402 [==============================] - 3s 124us/sample - loss: 1.1963 - accuracy: 0.5470 - val_loss: 1.5357 - val_accu
racy: 0.4319
Epoch 49/50
24402/24402 [==============================] - 3s 120us/sample - loss: 1.1875 - accuracy: 0.5532 - val_loss: 1.5324 - val_accu
racy: 0.4293
Epoch 50/50
24402/24402 [==============================] - 3s 117us/sample - loss: 1.1810 - accuracy: 0.5504 - val_loss: 1.5319 - val_accu
racy: 0.4351
```

Figure 5: Model trained with 50 epochs with an accuracy of 0.4351

The main disadvantage of using a Face emotion detection system over a microexpression-based system is that.

- Emotion detection focuses only on obvious emotions that are readily visible, which is insufficient for monitoring students during an online class/assessment.
- Emotion detection cannot determine whether a student understands a concept, as not everyone makes an obvious expression when they don’t understand, but microexpressions can detect the slightest difficulty.
- Normal emotions can be faked, but microexpressions cannot.

Therefore, for monitoring students during online classes, a combination of facial emotion and microexpression detection will be more efficient. The sample output shown in Figure 6 is provided for both happy and angry expressions.

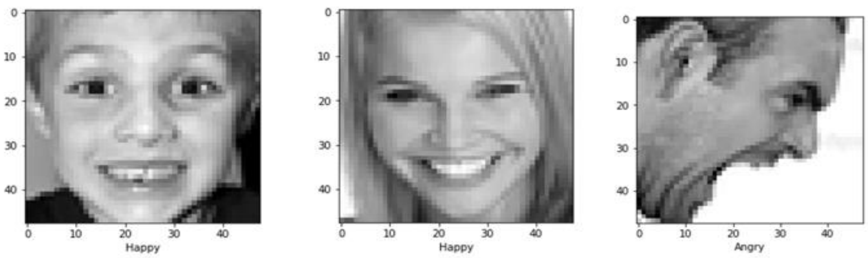

Figure 6: Sample output

Table 1 provides a brief note on the differences between micro-expression facial emotion detection and facial emotion detection in terms of real-time precision, academic integrity, monitoring learners, and learners' efficacy.

Table 1: Comparison of microexpression detection and facial emotion detection

Category	Microexpression Detection	Facial Emotion Detection
Real-Time Precision	Offers real-time insights into students' emotional states and engagement.	Often provides delayed and aggregated data, limiting real-time adaptability.
Improved Academic Integrity	Identifies cheating and dishonesty during online assessments to ensure fair evaluations.	Provides limited support for academic integrity.

Attention Monitoring	Monitors even momentary lapses in engagement with precision.	May not detect minor fluctuations in attention.
Increased Learning Efficacy	Assesses comprehension levels to improve learning outcomes.	May not offer insights specific to comprehension.

5. Applications and Benefits

There are various applications and benefits in micro-expression analysis and detection using the combined LSTM and Image Feature Stabilisation Algorithm, rather than the usual approach of facial emotion detection.

5.1. Student Engagement Assessment

Monitoring microexpressions can help determine students' levels of engagement in online classes. Disengaged pupils can be identified in real-time by educators, who can then adjust their teaching approaches accordingly. This can lead to improved learning outcomes.

5.2. Emotional Awareness

Unlike verbal communication or gestures, facial expressions are a universal system of signals that reflect the moment-to-moment fluctuations in a person's emotional state.

5.3. Emotional State Analysis

Detecting microexpressions enables the analysis of students' emotional states, including frustration, confusion, and interest. Educators can tailor their interactions to address students' emotional needs, providing timely support and encouragement.

5.4. Assessment

Analysing microexpressions can help assess whether students are comprehending the material, enabling teachers to adapt their teaching methods to clarify concepts or provide additional examples if needed, thereby improving learning outcomes.

5.5. Academic Tracking

Micro-expression detection can also help identify instances of cheating or academic dishonesty during online assessments, ensuring that students receive fair evaluations and that the value of their degrees is maintained.\

6. Challenges and Limitations

6.1. Technical Challenges

- The image must be of extremely high resolution to capture the smallest facial details, and the camera must also be able to capture a higher frame rate, as these micro expressions last only one-quarter of a second.
- This higher, larger volume of input data requires substantial computational power to process the images and produce the output.
- The algorithm's complexity also increases to account for all edge cases of facial expressions.

6.2. Drawbacks

- Requires high-quality hardware
- Might provide false output for smaller inputs
- Consumes Space and Bandwidth to transfer the video
- As the student strength increases, the hard requirement increases exponentially

6.3. Experimental Results by Combining Both Facial Emotion Detection and Microexpression

By combining both methods, a far more effective system for real-time use is created. It is more efficient and accurate than a single process, as it overcomes the drawbacks of the other system. Three sets of codes are executed for this purpose: LSTM for facial emotion detection, an LBP-CNN-SVM combination for microexpression detection, and an Image feature stabilisation algorithm to enhance images.

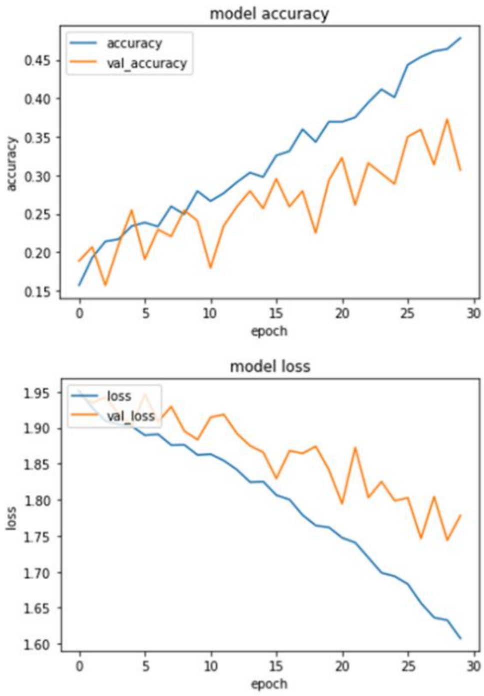

Figure 7: Model accuracy and model loss for 30 epochs

The graph shown in Figure 7 shows the model accuracy and loss over 30 epochs. From the graphs, we can interpret that the model becomes more accurate over time. Figure 8 shows the confusion matrix for the microexpressions of neutral, anger, contempt, surprise, fear, happiness, and sadness in microexpression facial emotion detection.

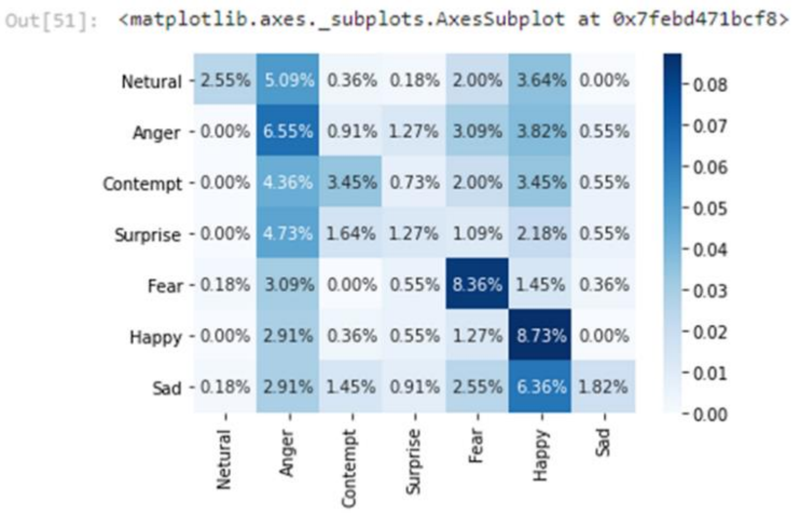

Figure 8: Confusion matrix for facial expressions with micro expressions like neutral, anger, contempt, surprise, fear, happy, and sad

Figures 9, 10, and 11 represent the confusion matrix and the heat map for emotions, including neutral, anger, contempt, surprise, and sadness. A confusion matrix in machine learning is a tabular representation used to evaluate the performance of a classification model. It provides a comprehensive summary of how well the model's predictions align with the actual class labels.

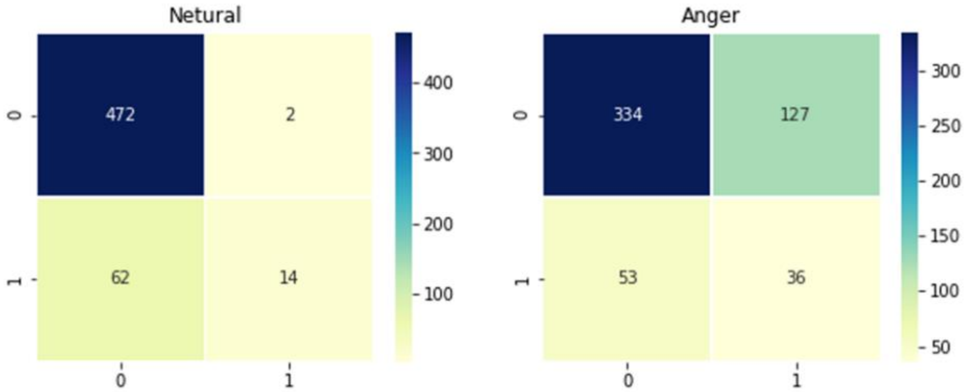

Figure 9: Confusion matrix for neutral and anger

The matrix consists of four key values: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). True positives are cases in which the model correctly predicted a positive class, while true negatives are cases in which the model correctly predicted a negative class.

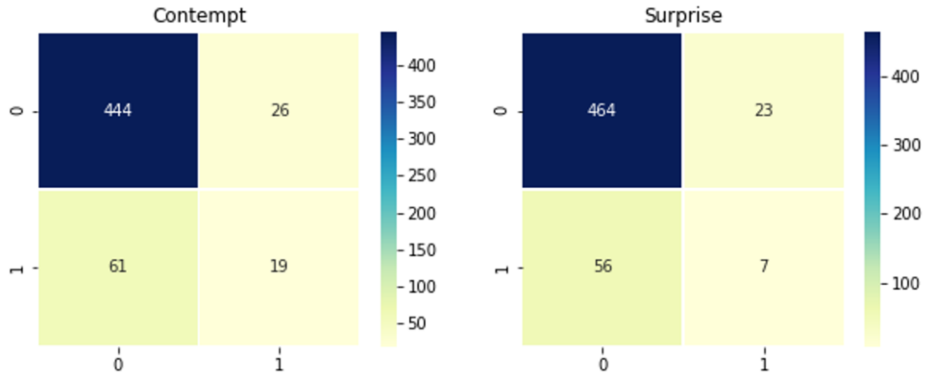

Figure 10: Confusion matrix for contempt and surprise

False positives occur when the model incorrectly predicts a positive class when it is negative, and false negatives occur when it incorrectly predicts a negative class when it is positive.

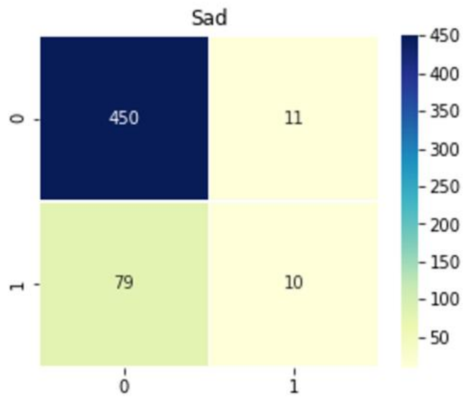

Figure 11: Confusion matrix for sad

The confusion matrix helps calculate various performance metrics, such as accuracy, precision, recall, F1 Score, and others, to assess the model's effectiveness. It is a valuable tool for understanding the strengths and weaknesses of a classification algorithm (Figure 12).

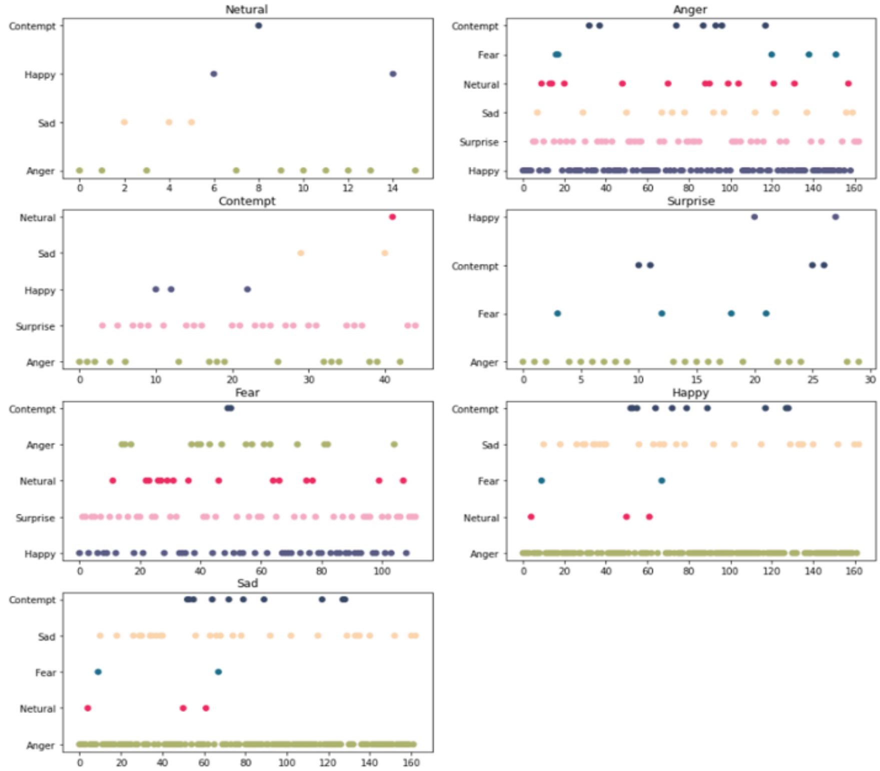

Figure 12: Heat map image for the micro expression neutral, anger, contempt, surprise, fear, happy, and sad

The data is split into individual emotions for easier visualisation. It is similar to the heatmaps used above. A heat map in machine learning is a data visualisation technique that uses a colour-coded matrix to represent the relationships and patterns within a dataset. It is particularly useful for understanding the distribution and correlations among different features or variables. Each cell in the heat map is colour-coded to represent the strength or magnitude of a particular value or relationship. Darker colours, such as red or blue, may indicate higher or lower values, respectively. Heat maps are widely used for various purposes in machine learning, including feature correlation analysis, anomaly detection, and model evaluation. For example, in feature correlation analysis, a heat map can help identify highly correlated features, which can lead to feature selection or dimensionality reduction. In anomaly detection, heat maps can highlight unusual patterns or outliers in data. Overall, heat maps offer a visual and intuitive way to gain insights into complex datasets and inform data-driven decisions in machine learning applications.

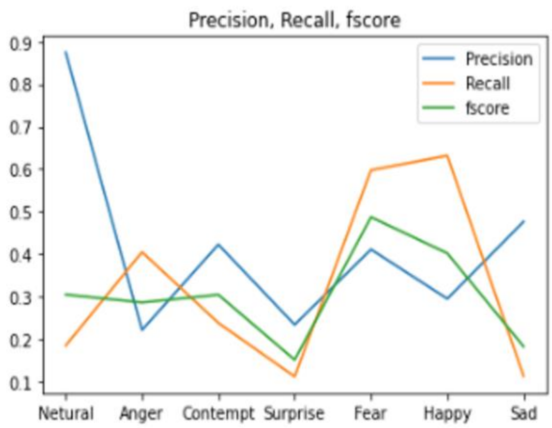

Figure 13: Precision, fscore, and recall graph for the micro expression neutral, anger, contempt, surprise, fear, happy, and sad

A precision, fscore, and recall graph is shown in Figure 13 for visual representation of the model. In machine learning, precision, recall, and F1-score are crucial metrics for assessing the performance of binary classification models, especially when the data is imbalanced. Precision measures the proportion of predicted positive instances that are correct, while recall assesses the proportion of true positive instances that the model correctly identifies. The F1-score, a harmonic mean of precision and recall, offers a balanced view of a model's performance. Precision-recall curves visualise how these metrics change as the classification threshold varies, helping you choose the right threshold for your specific needs. In these curves, a model with high precision and recall approaches the upper-right corner. F-beta curves are a more flexible version, allowing you to balance precision and recall as needed. Together, these metrics and curves guide you in optimising your model's performance based on the desired trade-off between precision and recall.

6.4. Micro-Expression Facial Emotion Detection Results

The final accuracy for facial emotion detection using microexpressions, combining an LSTM with an Image Feature stabilisation algorithm, is shown in Figure 14. The results for accuracy, macro average, and weighted average are far higher than those for normal facial emotion detection alone.

	precision	recall	f1-score	support
0	0.92	0.90	0.91	300
1	0.70	0.70	0.70	285
2	0.82	0.80	0.81	299
3	0.96	0.96	0.96	302
4	0.74	0.86	0.79	310
5	0.89	0.79	0.84	304
accuracy			0.84	1800
macro avg	0.84	0.83	0.84	1800
weighted avg	0.84	0.84	0.84	1800

Figure 14: Micro-expression facial emotion detection combining LSTM and image feature stabilisation algorithm

Figure 15 shows the graphical representation of the accuracy for facial emotion detection, microexpression emotion detection, and combined emotion detection. And we can conclude that combining both methods yield higher accuracy.

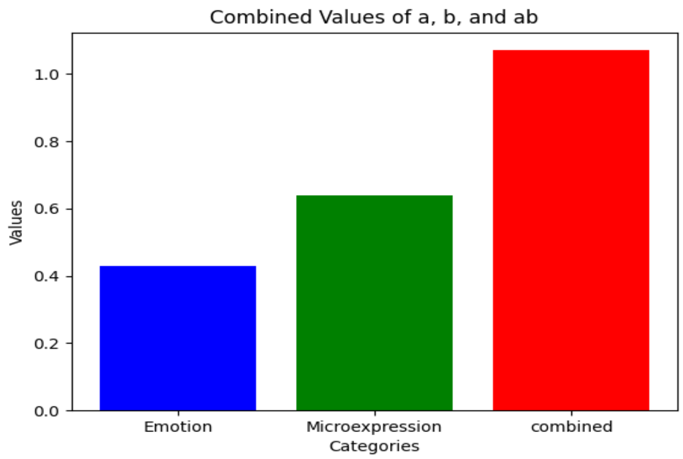

Figure 15: Accuracy of emotion detection, micro expression emotion detection, and combined emotion detection

7. Conclusion

Combining facial emotion detection and microexpression detection provides a holistic approach that significantly enhances the effectiveness of an online class student-monitoring system. By combining these two technologies, educators gain a comprehensive understanding of students' emotional states, from the subtle nuances revealed by microexpressions to broader emotional cues detected through facial emotion analysis. This insight enables educators to adapt their teaching methods in real time, offering personalised support that addresses both concealed emotions and overall engagement levels. This approach improves learning outcomes because teachers can tailor their interactions to meet students' individual emotional needs. Moreover, incorporating microexpression detection helps identify academic dishonesty during assessments, maintain academic integrity, and preserve the value of online degrees. By leveraging both technologies, online educators can gain a deeper understanding of their students' experiences and create an environment conducive to effective learning. The positive code results obtained from this combined approach affirm its practicality and potential for widespread adoption in online education. This synergy between microexpression detection and facial emotion analysis creates an environment that enables students to thrive, ensuring an enriching educational experience.

Acknowledgement: The authors gratefully acknowledge SRM Institute of Science and Technology for providing the resources and support necessary to complete this research work.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request to ensure transparency and reproducibility.

Funding Statement: This research work and manuscript preparation were carried out without any external financial assistance or funding from any organisation or agency.

Conflicts of Interest Statement: The authors declare that there are no conflicts of interest related to this study. All information, references, and citations have been properly acknowledged in accordance with academic standards.

Ethics and Consent Statement: The study was conducted in accordance with the ethical guidelines for research. Informed consent was obtained from all participants, and the necessary institutional permissions were secured prior to data collection.

References

1. F. Jiang, Q. Huang, X. Mei, Q. Guan, Y. Tu, W. Luo, and C. Huang, "Face2nodes: Learning facial expression representations with relation-aware dynamic graph convolution networks," *Information Sciences*, vol. 649, no. 11, p. 119640, 2023.
2. M. Shreve, S. Godavarthy, D. Goldgof, and S. Sarkar, "Macro- and micro-expression spotting in long videos using spatio-temporal strain," *in Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition (FG)*, Santa Barbara, California, United States of America, 2011.
3. K. Ezzameli and H. Mahersia, "Emotion recognition from unimodal to multimodal analysis: A review," *Information Fusion*, vol. 99, no. 11, p. 101847, 2023.
4. S. P. Mishra, P. Warule, and S. Deb, "Speech emotion classification using feature-level and classifier-level fusion," *Evolving Systems*, vol. 15, no. 11, pp. 541–554, 2023.
5. X. Li, S. Cheng, Y. Li, M. Behzad, J. Shen, S. Zafeiriou, M. Pantic, and G. Zhao, "4DME: A spontaneous 4D micro-expression dataset with multimodalities," *IEEE Transactions on Affective Computing*, vol. 14, no. 4, pp. 3031 – 3047, 2022.
6. B. Han, C. H. Yoo, H. W. Kim, J. H. Yoo, and J. Jang, "Deep emotion change detection via facial expression analysis," *Neurocomputing*, vol. 549, no. 9, p. 126439, 2023.
7. W. J. Yan, X. Li, S. J. Wang, G. Zhao, Y. J. Liu, Y. H. Chen, and X. Fu, "CASME II: An improved spontaneous micro-expression database and the baseline evaluation," *PloS One*, vol. 9, no. 1, pp. 1-8, 2014.
8. J. Zhao, F. Deng, Y. Cai, and J. Chen, "Long short-term memory-fully connected (LSTM-FC) neural network for PM2.5 concentration prediction," *Chemosphere*, vol. 220, no. 4, pp. 486–492, 2019.
9. D. Zhao, M. Chen, K. Shi, M. Ma, Y. Huang, and J. Shen, "A long short-term memory-fully connected (LSTM-FC) neural network for predicting the incidence of bronchopneumonia in children," *Environmental Science and Pollution Research*, vol. 28, no. 40, pp. 56892–56905, 2021.
10. R. Angeline and N. A. Alice, "Multimodal human facial emotion recognition using DenseNet-161 and image feature stabilization algorithm," *Traitement du Signal*, vol. 39, no. 6, pp. 2165-2172, 2022.
11. P. Choirina, U. D. Rosiani, I. M. Fitriani, and R. Baqi, "Facial micro expression recognition for feature point tracking using apex frames on CASME II database," *Sinkron: Jurnal dan Penelitian Teknik Informatika*, vol. 8, no. 1, pp. 81–89, 2023.