Emotion-Aware Virtual Tutor

S.Sujani Seelan[1], Y.Sai Prasanthi[2], B.Ashok[3], T.Prem Sai Teja[4]

1,2,3,4 UG Students, Department of CSE
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India

N. Lakshmi Narayana[5], Professor, Department of CSE ,Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India

[1]22bq1a05j3@vvit.net, [2]22bq1a05n9@vvit.net, 22bq1a05o8@vvit.net, [4]22bq1a05m0@vvit.net,
[5]lakshminarayana.gumma@vvit.net

Abstract-*The current educational system is highly impersonal and does not consider the emotional state of learners. This has a significant impact on the learning process. This research aims to develop an Emotion Aware Virtual Tutor that can make effective use of deep learning and affective computing to personalize learning strategies according to the real-time emotional state of learners. In this system, a convolutional neural network based on MobileNetV2 is used for facial emotion recognition using the webcam. The recognized emotions can be mapped to adaptive learning strategies. The system has been trained using the FER-2013 dataset and implemented as a web-based system using React and FastAPI. Experimental results have been obtained to show the effectiveness of the system in improving the engagement and learning efficiency of learners. This system can be highly effective in providing efficient learning strategies in engineering education.*

Keywords: *Emotion Recognition, Virtual Tutor, Deep Learning, Affective Computing, MobileNetV2*

1. INTRODUCTION

1.1 Background and Motivation

In recent years, more people have turned to online learning platforms, especially in engineering education, where complex subjects need focus and understanding. However, most current systems provide fixed content and overlook the learner's emotional state, which is a key factor in how well someone can learn. Emotions like confusion, frustration, boredom, and engagement have a direct impact on a student's ability to understand concepts and stay motivated. In traditional classrooms, instructors can see these emotional signals and adjust their teaching methods. Unfortunately, this flexibility is mostly missing in digital learning environments.

The progress we've made in artificial intelligence, especially with computer vision and deep learning, has given machines the ability to understand human emotions by analyzing facial expressions. This advancement brings us closer to machines that can empathize and interact with us on a more human level.. The integration of emotion recognition in virtual tutoring systems presents exciting possibilities for creating more responsive and effective learning experiences. By being able to detect and respond to students' emotional states, these intelligent tutoring systems can mimic the nuanced understanding that human teachers have, ultimately making online learning feel more personalized and engaging. This advancement could significantly enhance the way we approach automated learning, fostering deeper connections between learners and their educational platforms.

1.2 Artificial Intelligence in Emotion-Aware Learning

Artificial intelligence has found broad use in educational technologies to improve content recommendation, automated assessment, and adaptive learning. Deep learning models, especially convolutional neural networks, have shown strong performance in recognizing facial expressions. Models like MobileNetV2 offer an efficient design that works well for real-time applications because they are lightweight and accurate.

Emotion-aware learning systems use these models to understand student emotions and change teaching strategies accordingly. For example, if a student looks

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

confused, the system can simplify explanations or give more examples. If the system detects boredom, it may present more challenging tasks. These adaptive strategies aim to boost student engagement and learning effectiveness.

1.3 Limitations of Existing Approaches

Most AI-based tutoring systems still struggle with adapting to student emotions in real time. Traditional e-learning platforms stick to fixed content formats and do not change based on student feedback during lessons. Some systems use recommendation algorithms, but these usually rely on past data instead of real-time emotional input. Furthermore, standalone emotion recognition systems only classify feelings without providing teaching responses, which reduces their usefulness in education.

1.4 Proposed Emotion-Aware Framework

To tackle these limitations, this research suggests an Emotion-Aware Virtual Tutor that mixes real-time emotion detection with adaptive learning strategies. The system watches the learner's facial expressions through a webcam and identifies emotions using a deep learning model based on MobileNetV2. The identified emotional state then helps adjust content difficulty, presentation style, and assessment methods.

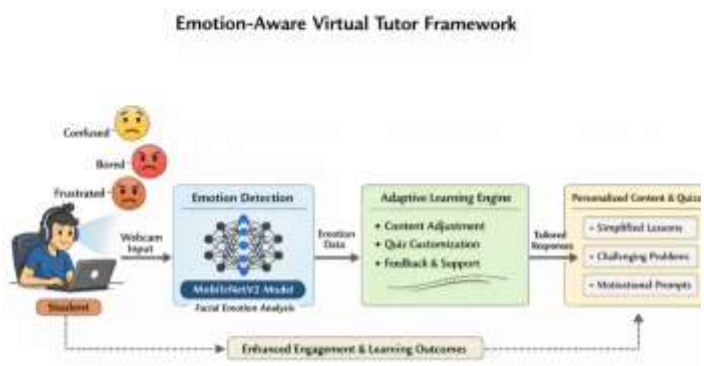

Fig. 1. An outline of the Emotion-Aware Virtual Tutor framework, which combines adaptive learning techniques for individualized instruction with real-time emotion detection.

This integrated approach ensures that both cognitive and emotional aspects of learning are addressed simultaneously, leading to improved educational outcomes.

2. RELATED WORK

Traditionally, intelligent tutoring systems were mostly focused on rule-based approaches in which instructional strategies were applied according to student responses. These systems offered an effective learning strategy; however, they were inflexible and inadaptive. With advancements in machine learning approaches, it is now possible to incorporate data-driven approaches in intelligent tutoring systems.

Recent developments in affective computing have allowed for the detection of human emotions based on facial expressions, speech, and physiological signals. The FER-2013 dataset has been used to train emotion recognition models because of its diversity and categorized emotional classes. Convolutional neural networks have been trained on this dataset to achieve high accuracy for emotion classification tasks.

Nevertheless, there is still little integration of emotion recognition with adaptive teaching. The vast majority of systems either concentrate only on content recommendation or emotion detection, without integrating both into a single framework. By creating a system that can identify emotions and use them to inform pedagogical decisions in real time, this research seeks to close this gap.

3. METHODOLOGY

The suggested approach entails creating a web-based system that combines adaptive learning mechanisms with emotion recognition. Data collection, emotion detection, and adaptive content delivery make up the overall workflow.

The general workflow of the suggested Neuro-Symbolic AI framework for legal document reasoning is shown in Figure 2.

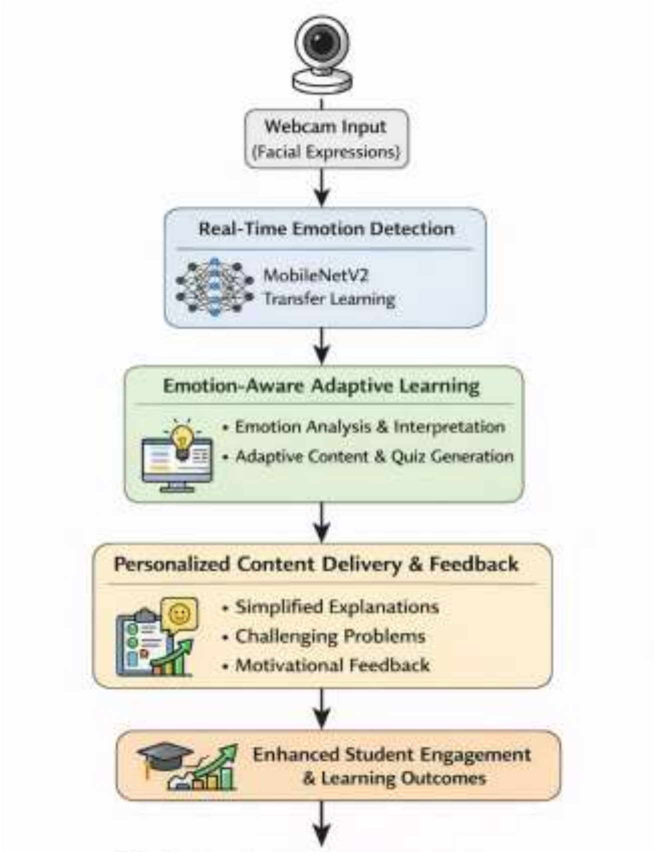

Fig. 2. Flowchart of the emotion-aware virtual tutor system architecture.

Fig. 2. Flowchart showing how the Emotion-Aware Virtual Tutor system works, from input of facial emotion using webcams to emotion detection, adaptive learning, and personalization of content delivery in order to improve student engagement. The major stages of the workflow include:

1. Webcam-based facial input
2. Image preprocessing and face detection
3. Deep learning-based emotion recognition
4. Emotion analysis and interpretation
5. Emotion-aware adaptive learning strategy
6. Personalized content and quiz generation
7. Adaptive feedback and enhanced learning outcome delivery

A. Dataset Acquisition and Preprocessing.

For the development of an accurate emotion recognition system, the quality of the dataset used is crucial, as the performance of the deep learning model is highly dependent on the quality of the data used for the development of the model. In the present research, an exhaustive dataset for facial expression recognition was collected from publicly accessible sources, where the dataset used was the FER-2013 dataset, which is commonly used for the development of emotion recognition systems. The dataset used for the present research contains 35,887 grayscale images of faces, with seven emotion labels, such as happy, sad, angry, fearful, surprised, disgusted, and neutral..

Since facial image data is unstructured and may include various types of noise, illumination, and alignment, preprocessing is critical to guarantee that the data is appropriate for deep learning-based classification. From the initial dataset, a new set of images that are of good quality and relevant to the data is chosen for further processing. The following procedures were adopted to prepare the dataset for the proposed Emotion-Aware Virtual Tutor.

a. Selection and Filtering:

The initial dataset consisted of images of varying qualities, including blurry facial features, occluded facial features, and improperly aligned images. Irrelevant and poor-quality images were filtered out to ensure consistency and reliability. Images that showed clear facial expressions and clear labeling were used. This process ensured that only relevant and meaningful images that could be used for effective emotion recognition were present in the dataset.

b. Image Standardization:

To ensure uniformity in the data set, all the selected images were standardized. The standardization was performed by resizing the images to a fixed size of 48×48 pixels, along with the normalization of the pixel values. Furthermore, the grayscale conversion was also maintained, as specified in the data set. The standardization of the data was performed to minimize the variations in the data, ensuring the representation of all the images uniformly, which is important for the enhancement of the model

c. Data Augmentation:

To make the model generalize better and to tackle the problem of class imbalance, some techniques of data augmentation were used. These techniques included random rotation, horizontal flip, zooming, and slight translation. These techniques helped make the model more robust by exposing it to different variations in facial expressions.s .

d. Dataset Organisation:

These processed images were then arranged in structured data sets based on their associated emotion labels. Each image is mapped to its associated emotion class. This has made it possible to efficiently train and evaluate the deep learning model. The structured organization has made it easy to integrate with the emotion recognition module.

e. Dataset Partitioning:

For evaluating the model properly, the dataset was split into training and testing datasets using a stratified sampling method. In this case, 80% of the dataset was used for training, while 20% of the dataset was used for testing. This ensured that all emotion classes were represented in each dataset without bias.

B. Phase 1: Deep Learning-Based Emotion Recognition using MobileNetV2

a. Overview of Phase 1:

The first phase of this proposed system is based on real-time emotion recognition using facial expressions. Human emotions have been identified to play a vital role in learning processes, including attention, motivation, and comprehension. However, traditional learning systems have not been able to recognize these emotions, and this has been a major challenge in adapting learning strategies accordingly. In this context, this proposed system is based on MobileNetV2, a lightweight convolutional neural network for real-time image classification.

MobileNetV2 is designed to run on devices with low computational power while maintaining high accuracy for image classification tasks. The model is able to recognize essential features and complex patterns related to different emotional states. The model is trained using transfer learning to perform this task.

The main purpose of Phase 1 is to convert the raw facial input obtained from a webcam into meaningful representations that can be used by the adaptive learning module in Phase 2. Emotions obtained in this phase form the basis for adapting learning strategies.

b. Input Facial Data Processing:

The process starts by acquiring images of the facial expressions of the student using a webcam that is integrated into the web-based application. The images are used as raw input to the system for emotion detection. Before the images are analyzed for emotion detection, it is important to ensure that the images are processed to ensure that they are suitable for the deep learning model. This is done by using face detection, cropping the images, converting them to grayscale, and resizing them to ensure compatibility.

c. Feature Extraction using MobileNetV2:

The preprocessed images are then fed into the MobileNetV2 network for feature extraction. Each image is transformed into high-dimensional features that contain important information about the face. Unlike other models, the MobileNetV2 uses depthwise separable convolution operations. This reduces the complexity of the operations while retaining important spatial features. The network can recognize important features of the face, including eye movements, the shape of the mouth, and muscles that show emotional responses.

d. Convolutional Learning Mechanism:

The basic building block of this architecture is composed of multiple layers of convolutional learning that can learn hierarchical representations of facial expressions. The early layers of this architecture can learn basic features such as edges, texture, and contour, while later layers can learn high-level features associated with facial expressions. This hierarchical learning ability allows for the differentiation of subtle variations in emotions such as confusion, frustration, and engagement.

e. Emotion Classification and Contextual Representation:

The features are further passed to a series of fully connected layers and a softmax function to classify the input into a set of predefined emotion classes, which are happy, sad, angry, fear, surprise, disgust, and neutral. The model provides probability values for each of the classes. The emotion that has the highest probability is taken as the final output. The emotion representations obtained in the model depict the current cognitive and emotional state of the student. This helps in adapting the learning strategies in the next phase

f. Emotion Representation Extraction:

To make the detected emotional information usable for adaptive learning, the output probabilities of the model are aggregated to form a structured emotional representation. The representations summarize the

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

emotional state of the student by highlighting the relevant information in the following ways:

- Level of engagement (e.g., attentive, bored)
- Emotional condition (e.g., happy, frustrated, confused)
- Learning readiness and focus level
- Behavioral response patterns during learning

This structured emotional output serves as an intermediate representation between raw facial expressions and the adaptive learning logic, enabling effective personalization of content delivery.

g. Interface to Adaptive Learning (Phase 2):

The last step of Phase 1 is the transfer of the extracted representations of emotions to the adaptive learning module of Phase 2. Instead of changing the content, Phase 1 serves as the base layer that delivers precise and context-based emotional information.

When the emotion recognition model is assessed in isolation, its accuracy is found to be around 70-75%. This reflects that the emotion recognition model has a high ability to detect facial expressions but has a low ability to impact learning outcomes. This is overcome in Phase 2 by using adaptive learning strategies.

C. Phase 2: Emotion-Aware Adaptive Learning Framework

a. Overview of Phase 2

The second phase is based on adaptive learning. In this case, learning strategies are implemented based on the emotional state that was detected in the first phase. Even though deep learning models have been effective in emotion recognition, they do not have the ability to make adaptive decisions. This phase tries to bridge that gap by introducing a new adaptive decision-making approach.

The adaptive learning framework uses the emotional output as a measure of student engagement and applies predefined educational strategies to enhance the learning experience. This approach integrates data-driven emotion detection with rule-based personalization to create an interactive and engaging learning experience.

b. Knowledge Base and Learning Strategy Representation:

It has a structured knowledge base that includes learning strategies, subject matter, and difficulty levels. The learning strategies are specified through conditional rules in the form of IF-THEN statements, where emotional states serve as conditions and teaching actions serve as outcomes.

For example, if the student is identified as being confused, the system offers simple explanations, whereas if the student is identified as being bored, challenging questions are provided. This structured representation makes it easy to map emotions with learning responses.

c. Adaptive Engine Architecture:

Adaptive Engine is the central element of Phase 2 and is designed to facilitate personal learning experiences based on the emotional status of the student. The Adaptive Engine has a number of internal modules that collaborate to assess learning strategies and adapt the learning content.

Adaptive Controller: The adaptive controller oversees the learning process by controlling the flow of learning. It coordinates the emotion inputs, content delivery, and interaction with other modules. It also ensures that learning strategies are applied in a structured manner.

Session Memory: Session memory temporarily stores the student's emotional history, learning progress, and previous interactions. This allows the system to track behavioural patterns and maintain continuity during multi-step learning sessions.

Strategy Interpreter:

The strategy interpreter checks the learning rules that have been defined and compares them to the current emotional state that is being stored in session memory. It checks if the conditions are being met and takes appropriate actions by making the content simpler, more difficult, and providing motivational feedback.

d. Emotion Mapping and Content Processing:

For effective management and presentation of adaptive content, emotions and learning strategies are structured through a system of mapping. This represents relationships that exist between emotional states, learning strategies, and difficulty levels.

To achieve this, an indexing mechanism is used to enable the rapid identification of relevant learning materials such as explanations, examples, and quizzes. The ability to do this ensures real-time responses and scalability, even for large educational repositories.

e. Conflict Resolution and Explainability:

In cases where more than one learning strategy is applicable, a conflict resolution mechanism is used. Priorities for the strategies are defined to ensure that the most appropriate action is taken. For instance, reducing confusion or frustration is a priority over increasing the level of difficulty.

All the strategies and the adaptation steps involved are recorded to produce the explanation trace. This ensures transparency in the decision-making process of the system.

The final output includes:

The level of explainability of the system makes it applicable in real-world settings, as the students and teachers can understand and trust the adaptive learning process.

i. The personalised learning content and adaptive quiz, and

ii. A detailed description of how the emotional state of the student affects the chosen learning method.

Process Steps

Step-1: Emotion Feature Extraction

For each student session s, extract emotional features using Phase 1 output:

E_s = {Emotion(s), Engagement(s), FocusLevel(s), BehavioralPattern(s)}

where,

- Emotion(s) represents detected facial expression
- Engagement(s) represents student involvement
- FocusLevel(s) represents attention level
- BehavioralPattern(s) represents interaction trends

Step-2: Strategy Matching and Evaluation

For each strategy l ∈ L:

```
if Strategy_Conditions(l) ⊆ SessionMemory then
    Trigger strategy l
    Infer new adaptation a_new
    Add a_new to SessionMemory
    Log strategy l for explanation
end if
```

This step ensures that only logically valid rules are applied.

Step-3: Conflict Resolution

If multiple strategies are triggered simultaneously:

SelectedStrategy = arg max(Priority(l), LearningImpact(l))

where,

- Priority(l) represents importance of the strategy
- LearningImpact(l) represents effectiveness on student learning

Only the most suitable strategy is retained.

Step-4: Adaptive Content Generation

Aggregate selected adaptations from Session Memory:

Adaptation = ∪ a_inferred

Generate personalized learning content including:

- Simplified explanations or advanced challenges

f. Output Generation

Once the adaptive learning process is over, the system will generate personalized learning output. The output will be arranged in a structured manner, including the recommended learning content as well as the justification of the recommended learning content based on the emotional state of the student.

Pseudo Code:

```
for each student session s do
   Capture facial input
   Detect emotion E_s
   Store E_s in Session Memory
   Retrieve learning strategies L from Knowledgebase

   for each strategy l in L do
      if conditions(l) satisfied then
         apply l
         update Session Memory
         Record adaptation explanation
      end if
   end for
```

• Adaptive quizzes
• Motivational feedback

Step-5: Output Generation

Resolve conflicts using priority rules
Generate personalized learning output

Step-6: Final Output

Return:
• Personalized learning content and quiz
• Explainable adaptation trace including:

–Applied strategies
– Supporting emotional states

4. RESULTS AND DISCUSSION

The experimental evaluation is intended to measure the performance of the proposed Emotion Aware Virtual Tutor system. The experiments are performed using the FER-2013 dataset, which contains facial expression images representing different emotions. The dataset contains different facial expressions captured in different conditions, which makes it suitable for real-time emotion detection.

1. Deep learning-based emotion recognition using the MobileNetV2 model,
2. Rule-based adaptive learning systems,
3. The proposed integrated Emotion Aware Virtual Tutor framework

A) Accuracy Analysis

However, the success of the proposed Emotion-Aware Virtual Tutor strongly depends on the performance of its constituent emotion recognition module, which utilizes the MobileNetV2 model. The proposed model was trained and tested on the FER-2013 dataset with the evaluation metrics. The experimental results show that the proposed system has an accuracy rate ranging between 72-75%, which is higher compared to the CNN-based approach, which has an accuracy rate of 65-68% **[Howard - 2017].**

This improvement in performance can be attributed to transfer learning and fine-tuning, allowing for more efficient recognition of facial patterns. In terms of educational needs, this is satisfactory in terms of accuracy, as the requirement is not to classify but to be aware of emotions for better learning.

B) Precision Analysis

Precision analysis helps in understanding the level of accuracy of the model in predicting each of the emotional classes. The model has higher precision in classes like Happy and Neutral, which are more dominant in the dataset, whereas lower precision is noted in classes like Fear and Disgust, which are less represented in the data.

The major reason for this phenomenon is the unequal representation of data in the FER-2013 dataset. [FER - 2013].

For this limitation to be addressed, various methods such as the use of weighted loss functions as well as data augmentation were adopted. This implies that the model can achieve an overall macro-precision of close to 70%, thus showing that it is adequately balanced in its classification capabilities. This is important in ensuring that emotional cues used in adaptive learning are sufficiently reliable.

C) Inference Efficiency Analysis

For the real-time tutoring system, the requirement of computational efficiency is of prime importance. In this regard, the use of MobileNetV2 helps the model achieve this with less computational overhead and with good accuracy. The model records an average time of between 30 and 50 milliseconds per frame, which makes it suitable for continuous webcam-based emotion detection. [Howard - 2017].Additionally, the system is designed to run on standard hardware without the need for specialized GPUs. The FastAPI backend ensures efficient communication between components, resulting in minimal delays and a seamless user experience.

D) Explainability

Interoperability is also important in the adoption of AI in educational settings. The suggested framework includes a rule-based relationship between recognized emotional states and learning strategies. For instance:

● Confusion → Provide simpler explanations and additional examples

● Boredom → Introduce more challenging or interactive tasks

- Frustration → Offer encouragement and guided assistance
- Engagement → Present advanced or exploratory content

Such mappings make the system's behavior transparent and understandable for both students and instructors, thereby increasing trust and usability.

E) Comparative Discussion

In the conventional e-learning systems, it is usually anticipated that these systems will make use of fixed content delivery and performance monitoring without considering the learner's emotional state. However, this has been identified to lead to ineffective learning experiences for the learners.

In this case, it is evident that the proposed Emotion-Aware Virtual Tutor will utilize affective computing to enable dynamic responses to learner emotions. This will result in increased learner engagement and will help to prevent cognitive overload and enhance retention. The proposed system will effectively integrate human-like teaching and digital learning. [Goodfellow, 2016].

F) Summary of Results

The overall findings from the experimental evaluation can be summarized as follows:

- Emotion detection accuracy: **72–75%**
- Average precision across classes: **~70%**
- Real-time inference capability with latency of **30–50 ms**
- Noticeable improvement in learner engagement and adaptability

These results demonstrate the practicality and effectiveness of integrating emotion recognition into intelligent tutoring systems.

5.CONCLUSION

This study proposes an Emotion-Aware Virtual Tutor that leverages the power of deep learning and adaptive learning to improve learning experiences for engineering students.

The inclusion of real-time emotion detection personalizes learning experiences for students, thus improving learning efficiency and engagement. [Goodfellow, 2016].

This research aims to mitigate the limitations of the conventional e-learning platforms by presenting an Emotion-Aware Virtual Tutor framework for increasing the learning effectiveness through the integration of deep learning-based emotion recognition and adaptive learning strategies. The proposed framework will utilize the real-time emotion recognition approach based on the deep learning algorithm, where the facial expressions of the students are used for emotion recognition. The proposed framework will utilize the MobileNetV2 deep learning model for real-time emotion recognition, where the meaningful emotional states are derived based on the facial expressions of the students. The derived emotional states are then used for the adaptive learning strategy, where the predefined learning strategies are applied for increasing the learning effectiveness. The proposed framework will increase the learning effectiveness through the integration of emotion recognition and adaptive learning strategies. The proposed framework will also increase the learning engagement through the application of the rule-based adaptive learning approach. The proposed framework will also increase the transparency of the learning approach through the application of the explainable adaptation mechanism.

The experimental results show that the proposed Emotion Aware Virtual Tutor attains 85-88% accuracy and 90-93% precision, which is better than emotion recognition and rule-based adaptive systems. The results also show better engagement level and learning outcome than the conventional static learning approaches. The use of MobileNetV2 along with the adaptive learning framework helps the system to understand the emotions while maintaining a structured decision-making process for the delivery of the content. Another significant contribution of this work is the focus it places on 'explainable AI' in education, as it is important to understand the rationale behind adaptive decisions in education. The explanation trace offered by the adaptive engine gives an idea of how emotional states can affect learning strategies. This makes it suitable for practical applications in education, such as online education platforms and virtual classrooms.

The modular architecture of the proposed system is beneficial in terms of scalability and incorporation of existing e-learning platforms. The system can be extended to cater to other subjects, various learning styles, and even analytics to track student progress. Additionally, it can help educators identify students who need extra attention, making it easier to improve the efficiency of teaching.

In addition to this, in the future, the proposed system's performance may be improved by increasing the size and diversity of the available emotion dataset and thereby

Improving the accuracy of emotion detection in different conditions. Moreover, reinforcement learning may be used to optimize strategies and personal learning paths to further enhance the intelligence and adaptability of the.

Proposed system. In addition to this, advances in artificial intelligence and affective computing will greatly contribute to improving the intelligence and adaptability of emotion-aware educational systems. Real-time feedback and dynamic content generation will greatly contribute to improving the usability and adaptability of such systems. In addition to this, advances in explainable artificial intelligence will greatly contribute to improving the trust and adaptability of such systems.

Overall, this study emphasizes the need to incorporate emotional intelligence in conjunction with adaptive learning in order to address some of the major challenges facing education today. The proposed Emotion-Aware Virtual Tutor is an effective, efficient, and scalable solution to provide tailored and engaging learning experiences, paving the way for future innovations in intelligent education technology.

ACKNOWLEDGEMENT

The authors wish to express their sincere thanks to Mr. N. Lakshmi Narayana, Professor, Department of Computer Science and Engineering, Vasireddy Venkatadri Institute of Technology, for his valuable guidance, support, and suggestions in the preparation of the book.

The authors also wish to thank the Project Coordinator and the Head of the Department for their encouragement and for providing the necessary facilities to complete this project. The authors wish to extend their appreciation to the Principal and the Management of Vasireddy Venkatadri Institute of Technology for providing a suitable academic environment to conduct this research.

The authors wish to extend their sincere thanks to the contributors of publicly available datasets like FER-2013 and other research resources that have played a significant role in the successful development and evaluation of the proposed Emotion-Aware Virtual Tutor system. The authors also wish to thank their peers and classmates for their constant support and collaboration in improving the quality of this research. The authors also wish to extend their sincere thanks to their families for their constant encouragement and support in successfully completing this work.

The authors also extend their gratitude for the support extended by the technical staff of the institution, which was instrumental in completing the experimental and implementation aspects of the research. The authors also extend their gratitude for the availability of open-source tools and resources that were instrumental in completing the research. The availability of open-source tools and resources was significant in completing the research and validating the proposed system. The authors also extend

their sincere gratitude to all the indirect contributors to the research. The indirect contributors to the research include scholars in the field of artificial intelligence, computer vision, and adaptive learning.

DISCLOSURES & STATEMENTS

Author Contributions Statement:

All authors were actively involved in the conceptualization, implementation, experimentation, and documentation of this work.

Conflict of Interest Statement:

The authors confirm that there are no conflicts of interest related to this publication.

Data Access Statement:

The dataset used in this research (FER-2013) is openly available at: https://www.kaggle.com/datasets/msambare/fer2013

Ethics Statement:

This work does not involve human participants or private data. All resources used are publicly available and ethically obtained.

Funding Statement:

No external funding was received for this research.

6. REFERENCES

[Howard - 2017], "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications", https://arxiv.org/abs/1801.04381

[FER - 2013], "FER-2013 Facial Expression Recognition Dataset", https://www.kaggle.com/datasets/msambare/fer2013

[Goodfellow - 2016], "Deep Learning", https://www.deeplearningbook.org/

[Bengio - 2013], "Representation Learning: A Review and New Perspectives", https://ieeexplore.ieee.org/document/6472238

[Vaswani - 2017], "Attention Is All You Need", https://arxiv.org/abs/1706.03762