Integrated Fuzzy Atmosfield and Machine Learning to Evaluate Group Learning Status in Smart Classroom

Linyao Yang[1], Bemnet Wondimagegnehu Mersha[2][0000−0002−1000−8984], Yaping Dai[3][0000−0001−8795−5333], Kaoru Hirota[4][0000−0001−5347−6182], Wei Dai[5], and Yumin Lin[6]

[1] School of Automation, Beijing Institute of Technology, Beijing, 100081, China email: 1416898609@qq.com
[2] School of Automation, Beijing Institute of Technology, Beijing, 100081, China email: MershaBemnetW@bit.edu.cn
[3] Beijing Institute of Technology Zhuhai, Tangjiawan, Zhuhai City, 519088, Guangdong, China email: daiyaping@bit.edu.cn
[4] School of Automation, Beijing Institute of Technology, Beijing, 100081, China email: hirota@bit.edu.cn
[5] River Security Technology Co., LTD, Shanghai 200336, China email:daiwei@gmail.com
[6] River Security Technology Co., LTD, Shanghai 200336, China email:ymlin@riversecurity.com

Abstract. Individual and group classroom behavior evaluation is crucial for enhancing the quality of education in smart classrooms. The traditional classroom group behavior evaluations are labor-intensive, time-consuming and often biased. Current research has limited behavior categories and primarily concentrates on identifying individual student behavior, which has left a gap in the assessment of group behavior. To address this, an integrated Fuzzy Atmosfield and Machine learning method is proposed for group behavior evaluation. The proposed method is tested on a custom-made dataset suitable for group learning to showcase its potential for real-world application. The machine learning model achieved an average behavior classification accuracy of 94.7%. The integrated Fuzzy Atmosfield and machine learning method effectively classifies the group learning condition as poor with 92% accuracy, normal with 100% accuracy, and excellent with 92% accuracy. The test results from the custom-made dataset indicate that the proposed method offers significant advantages for evaluating group learning conditions in smart classrooms.

Keywords: Fuzzy Atmosfield · Action Recognition · Smart classroom · Machine learning.

1 Introduction

Education plays a crucial role in personal growth, societal development, and economic prosperity by establishing the foundation for a knowledgeable and skilled population. Therefore, it is essential to evaluate classroom students' group learning status to ensure the effectiveness of education. Such evaluation enables educators to adjust their methods, fostering a more engaging and supportive learning environment.

Current methods for evaluating teaching involve supervisors manually assessing the group learning process through in-person classroom observations or video analysis [1]. However, these procedures are labor-intensive, time-consuming and prone to bias. With advancements in artificial intelligence (AI), researchers have started applying AI technologies, such as human action recognition and facial expression recognition, to classroom learning environments [2–4].

The task of behavior recognition has been widely studied, and some research on behavior recognition in classroom environments has already been conducted. In this context, several datasets related to behavior recognition in classroom environments (The details about these datasets will be discussed in Section 2) have been created by researchers utilizing well-established techniques from the field of behavior recognition to recognize behaviors in classroom settings. This paper focuses on identifying the learning status of a group of students in a classroom environment through behavior recognition. This research targets two gaps in the current state of research. First, the classification of behaviors in existing classroom behavior datasets is not detailed enough; there are differences within the same category of behaviors, so using such behavior classifications for learning status analysis can only yield general and imprecise results. Second, current research primarily focuses on behavior recognition, with a lack of analysis regarding the learning status of student groups.

This research proposes a machine learning-based model with Fuzzy Atmosfield (FA) to evaluate the students' group learning status. FA, proposed by Kaoru Hirota's team in 2013, utilizes fuzzy theory to describe the "communication atmosphere" [5]. Originally, it was designed to help robots better understand human emotions through visual and auditory data, thereby improving interaction.

Our research aims to develop and implement a method for evaluating students' group learning status (see Figure 1). First, a machine learning model is employed to recognize individual students' behavior from video data. Then, individual and group learning statuses are modeled using the FA methodology. The behavior recognition results, combined with Fuzzy logic theories, facilitate the automatic evaluation of the group learning status. This study emphasizes the integration of advanced machine learning and decision analysis techniques to enhance the accuracy and efficiency of evaluating group learning dynamics.

The innovative points of the paper are the further classification of behaviors and the adaptation of Fuzzy Atmosfield (FA) theory. Firstly, this paper pays attention to the subtle differences in normal behaviors (e.g., listening). Secondly, this paper adapts FA theory to a specific environment (i.e., the classroom), which could provide a method for applying FA theory to particular environments.

The structure of this paper is organized as follows: Section 2 introduces the creation of the classroom student behavior dataset. Section 3 outlines the deep learning methods for behavior recognition. In Section 4, the modelling of individual and group learning status of students using the "Fuzzy Atmosfield" concept and the construction of the fuzzy inference system is discussed. Section 5 presents the experimental results of the deep learning models and the students' group learning statuses. Finally, Section 6 provides the conclusion of the paper.

2 The custom-made student classroom behavior dataset

Many student behavior datasets have been developed by researchers. Some notable examples are ActRec-Classrom [6], the Student Classroom Behavior Dataset [7], a Large-scale Student Behavior Dataset [8] and SCB-Dataset3 [9]. However, these datasets are not suitable for our research, which focuses on both individual and group behavior recognition in classroom settings. Therefore, more attention is given to the classification of common behaviors among students. Existing datasets do not align with our objectives, as they include broad categories that lack the specificity needed for our analysis. For instance, the standard "listening to the lecture" category is rather broad because students might be highly attentive or slightly distracted while listening to the lecture. Similarly, the commonly used category "communication" fail to differentiate between class-related or unrelated communication.

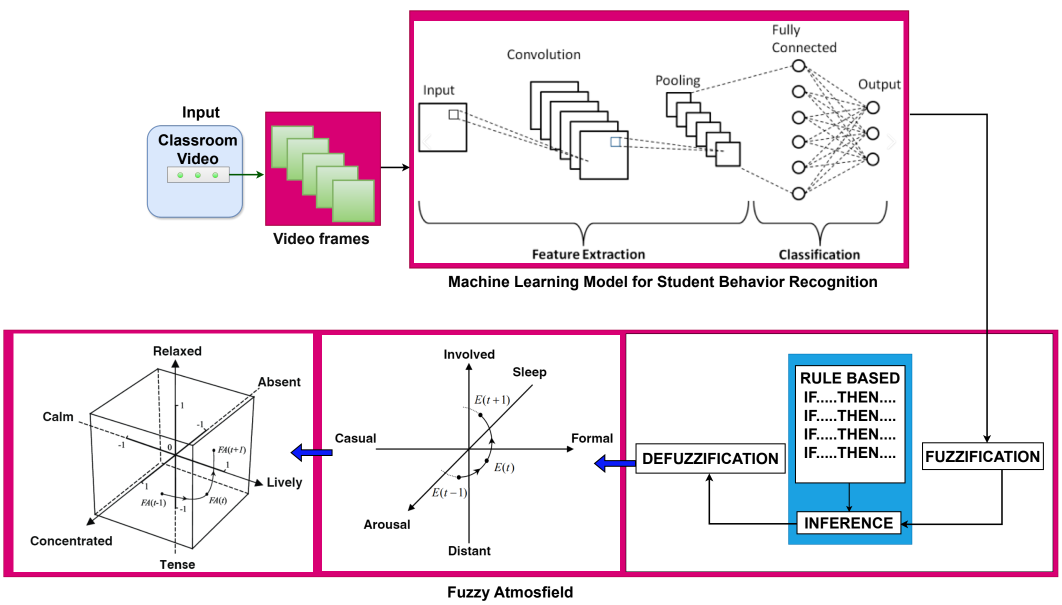


Fig. 1: Proposed classroom atmosphere evaluation method

To address the research needs for studying students' learning status, a custom-made dataset that finely classifies student classroom behaviors is developed. The detailed behavior categories in this dataset, tailored specifically for our research objectives, are presented in Table 1. During the development of the custom classroom student behavior dataset, six volunteers were seated in a classroom environment, each performing distinct classroom actions simultaneously. After recording each video segment, the volunteers altered their behavior, ensuring a diverse range of actions were captured. The process resulted in a total of 100 video segments, each approximately 8 seconds in duration. An example screenshot from these recordings is presented in Figure 2.

Table 1: Behavior Categories in the Custom Dataset

Behavior No.	**Behavior Type**
1	Listening (leaning forward)
2	Listening (leaning back)
3	Raising hand
4	Sleeping
5	Standing
6	Taking notes
7	Whispering
8	Using a laptop
9	Using a mobile phone

Fig. 2: An example image from the self-made dataset video

When developing the dataset, the focus was placed on including diverse data for each behavior category by having several student behavior performed in each video. For example, in Figure 2, multiple student behaviors such as using a laptop, using a mobile phone, standing, and raising a hand are observable within a single video segment. The dataset encompasses various commonly observed students' behaviors in real classrooms. Volunteers acted out these behaviors, and after the videos were recorded the videos, 30 frames of images were extracted from each segment. The single-person images were then cropped and categorized. The number of images for each behavior category in the dataset is detailed in Table 2.

Table 2: Scale of the Custom Classroom Student Behavior Dataset

Behavior Type	Number of Images
Listening (leaning forward)	180
Listening (leaning back)	186
Raising hand	168
Sleeping	186
Standing	162
Taking notes	186
Whispering	186
Using a laptop	180
Using a mobile phone	186

3 Deep learning methods for students' classroom behavior recognition

A convolutional neural network (CNN)-based model was developed to recognize student classroom behavior. The architecture of the CNN model, designed for behavior recognition, is depicted in Figure 3. The convolution blocks comprise convolutional layers, batch normalization, ReLU activation functions, and max pooling layers (see Figure 3). The depths of the convolutional layers in the three convolution blocks are 32, 64, and 128 sequentially, ensuring comprehensive feature extraction.

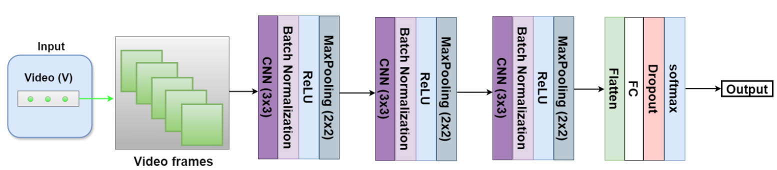


Fig. 3: Developed Action Recognition Model

The fully connected layer comprises 256 units, utilizes the ReLU activation function, and has nine output nodes. To mitigate overfitting, a dropout layer with a rate of 0.5 is incorporated. Additionally, the model is compiled using the Adam optimizer, with categorical cross-entropy as the loss function and accuracy as the evaluation metric.

4 Analysis of students' learning status based on Fuzzy Atmosfield (FA)

Kaoru Hirota's team proposed the "Fuzzy Atmosfield" methodology [5]. Fuzzy Atmosfield uses three-dimensional space to describe individual and group behavior. The original paper uses the American Psychological Association (APA) emotional space and the three-dimensional atmosphere space for group behavior. The original APA emotion space used to describe individual emotions consists of the axes "Affinity," "Pleasure-displeasure," and "Arousal-sleep." The original group emotion space used to describe the overall atmosphere consists of the axes "Lively-calm," "Friendly-hostile," and "Casual-formal."

In this research, the original APA emotional space is modified, and a three-dimensional atmosphere space suitable for a classroom setting is created. The modified Fuzzy Atmosfield is aimed at evaluating individual and group classroom learning status. For individual learning status, "arousal-sleep," "casual-formal," and "involving-distant" are used (see Figure 4). The individual learning status is defined as a three-dimensional vector, with each dimension corresponding to one of the axes of the individual learning status space coordinate system, as defined by

$$E = \begin{cases} (e_{involved}, e_{arousal}, e_{formal}), \\ \forall e_{involved}, e_{arousal} \ \ and \ \ e_{formal} \epsilon [-1, 1], \end{cases} \tag{1}$$

where E represents emotional state, with $e_{involved}$, $e_{arousal}$ and e_{formal} indicating precise positions along the respective axes. This framework allows for a detailed assessment of classroom dynamics, contributing to our understanding of individual learning experiences in educational contexts.

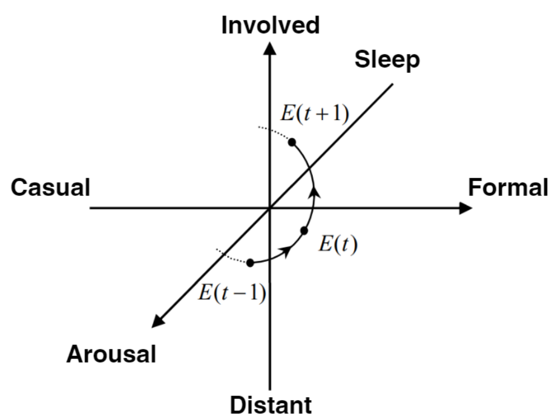


Fig. 4: Individual learning status

After establishing the individual emotion space, fuzzy theory is used to construct the group emotion space. For group learning status, "concentrated-absent," "lively-calm," and "relaxed-tense" are defined (see Figure 5). Each individual learning status and group learning status coordinate space value has a value from -1 to 1.

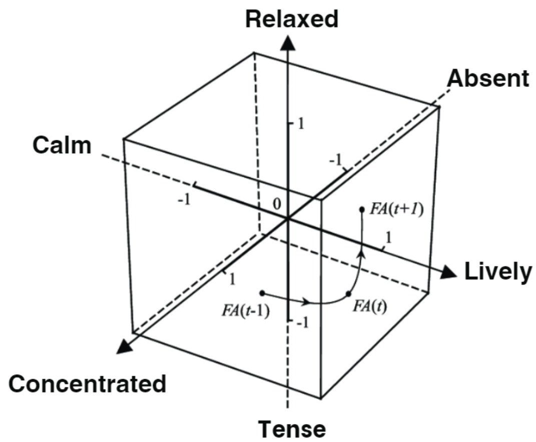


Fig. 5: Group learning status

The Fuzzy Atmosfield is calculated using

$$FA(t) = \begin{cases} f(E_1(t), \ldots, E_n(t)), \text{ if } t = 1, \\ (1-\lambda)FA(t-1) \cdot \gamma + \\ \lambda f(E_1(t), \ldots, E_n(t)), \text{ if } t = 2,\ 3,\ \ldots, m. \end{cases} \tag{2}$$

where f is a function regarding learning status at time t; n is the number of individuals in the learning scenario; γ is a defined forgetting factor that determines the rate of decrease of $FA(t-1)$; λ is a relevance factor, $0 \leq \lambda \leq 1$. When $t = 1$, the Fuzzy Atmosfield is in its initial state. The function f is defined as

$$f = \sum_{i=1}^{n} \varpi_i \cdot \text{defuzzy}(E_i(t) \circ R),\ i = 1,\ 2,\ \ldots,\ n \tag{3}$$

where ϖ is the weight of the i^{th} learning status with the sum of the weights equal to one, R represents the fuzzy relation between the Fuzzy Atmosfield and the individual learning status space, defuzzy is the defuzzification method, the defuzzification method used here is the centroid method.

To assign the parameters, we assumed that each student in the classroom has an equal influence on the group learning state; therefore, ϖ_i in Equation (3) is taken as $1/n$. The value of λ in Equation (2) reflects the degree to which the current group learning status is influenced by the previous learning status. The value of λ is set to 0.9. The selection of the forgetting factor is relatively complex. The original paper defined γ as $\gamma = \exp^{-kT}$, k is a positive parameter that determines the rate of fading of $CA(t-1)$ and T is the sampling time interval. After measuring the sampling interval during program operation, we set $\gamma = 0.2$.

After the fuzzy inference system is built and run, it initially outputs the levels in three membership functions for the group learning status. It then uses the centroid method for defuzzification calculation. The defuzzification results are values between -1 and 1. Then, the defuzzification value is used in Equation (2) to calculate the final group learning status results.

The centroid method utilized for defuzzification produces values derived from membership function levels, exhibiting specific characteristics, where the highest grade corresponds to values between -0.89 and 0.88. While this value may be reasonable for someone familiar with the system's operational principles, it may be confusing for most people. To address this issue, a normalization operation, which involves dividing all computed values by 0.9, is used. This adjustment aims to enhance the clarity and accessibility of the results for a broader audience.

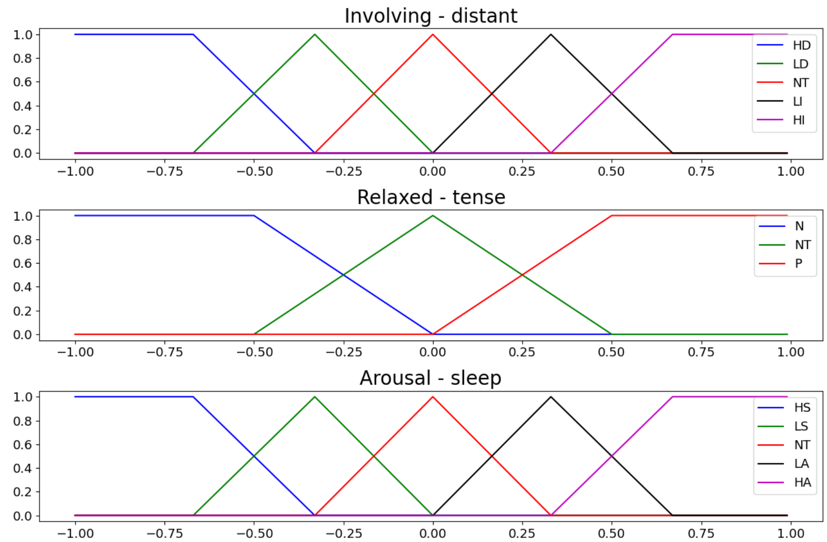
4.1 Membership Functions for classroom learning status

The FA methodology uses membership functions to mathematically represent individual and group learning status. Fuzzy rules are then employed to translate individual learning status into a group learning status. Once the Atmosfield is established, the research subjects' individual and group learning status are mathematically described. We used "arousal-sleep," "casual-formal," and "involving-distant" to describe individual learning status and "concentrated-absent," "lively-calm," and "relaxed-tense" to represent group learning status. Each evaluation criterion is assigned values between -1 and 1 for both individual and group learning status.

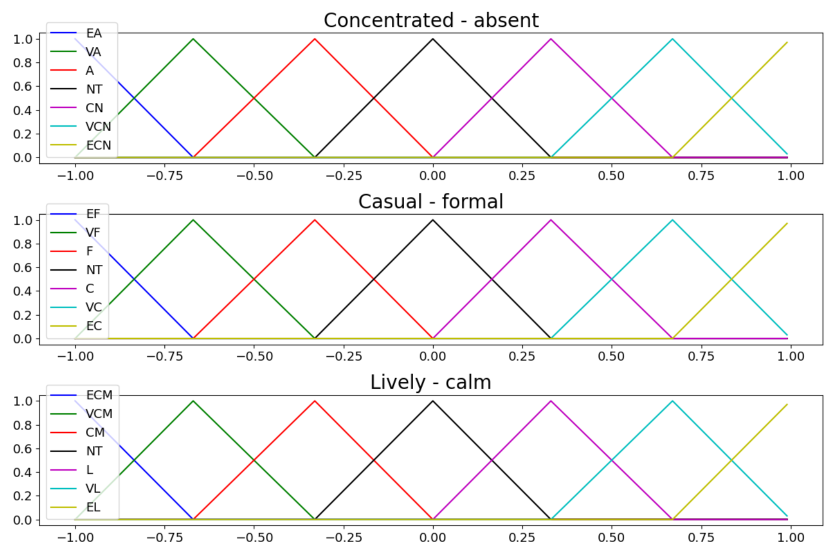
For individual learning status, the assessment includes considerations of students' relaxation or tension levels and their level of attentiveness, providing a comprehensive description of each student's learning status in the classroom. On the other hand, the designated area for group learning status focuses on the overall attentiveness, engagement, and relaxed or tense state of the classroom, appropriately describing the group learning status. The membership functions for individual and group learning status are depicted in Figures 6a and Figure 6b, respectively.

4.2 Fuzzy Rules for classroom learning status

After three-dimensional spaces and membership functions were established to describe individual and group learning statuses using fuzzy mathematics, a fuzzy inference system was developed. This system utilizes fuzzy mathematical variables to represent individual learning statuses, with rules established to translate behavior recognition results into these learning statuses.



(a) Membership function for individual learning status. In the figure "involving-distant" is labeled as "HD (Highly Distracted)," "LD (Lightly Distracted)," "NT (Neutral)," "LI (Lightly Involved)," and "HI (Highly Involved)." "Arousal-sleep" is labeled as "HS (Highly Sleepy)," "LS (Lightly Sleepy)," "NT (Neutral)," "LA (Lightly Active)," and "HA (Highly Active)." "Relaxed-tense" is labeled as "N (Tense)," "NT (Neutral)," and "P (Relaxed)."



(b) Membership function for individual learning status. In the figure, "Concentrated-absent" is labeled as "EA (Extremely Absent)," "VA (Very Absent)," "A (Absent)," "NT (Neutral)," "CN (Concentrated)," "VCN (Very Concentrated)," and "ECN (Extremely Concentrated)." "Casual-formal" is labeled as "EF (Extremely Formal)," "VF (Very Formal)," "F (Formal)," "NT (Neutral)," "C (Casual)," "VC (Very Casual)," and "EC (Extremely Casual)." "Lively-calm" is labeled as "ECM (Extremely Calm)," "VCM (Very Calm)," "CM (Calm)," "NT (Neutral)," "L (Lively)," "VL (Very Lively)," and "EL (Extremely Lively)."

Fig. 6: Membership Functions

Questionnaires were distributed to undergraduate and graduate students across various disciplines to formulate rules for translating behavior recognition results into individual learning statuses (see Table 3). Following the definition of membership functions and rules, fuzzy rules were developed aimed at translating individual learning statuses into group learning statuses. A total of 75 fuzzy rules were employed, a selection of which is presented in Table 4.

Table 3: Rules for converting behavior recognition results into individual learning states. (Look at Figure 6a for the description of the labels)

	Involving-distant	Relaxed-tense	Arousal-sleep
Listening (leaning forward)	LI	N	LA
Listening (leaning back)	LI	NT	LA
Raising hand	HI	N	HA
Sleeping	HD	P	HS
Standing	HI	N	HA
Taking notes	LI	N	LA
Whispering	LD	P	HS
Using a laptop	LI	NT	LA
Using a mobile phone	LD	P	NT

Table 4: Partial Fuzzy Rules of the Fuzzy Inference System. (Look at Figure 6a and Figure 6b for the description of the labels)

No.	Individual Status			Group Status		
1	Involving-distant	Relaxed-tense	Arousal-sleep	Concentrated-absent	Casual-formal	Lively-calm
2	HD	N	HS	EA	F	ECM
3	HD	N	LS	EA	F	VCM
4	HD	N	NT	VA	VF	VCM
5	HD	N	LA	VA	EF	CM
⋮	⋮	⋮	⋮	⋮	⋮	⋮

5 Experimental Results

5.1 Deep Learning Experimental Results

The custom-made behavior dataset was split into training and testing sets, with 80% allocated for training and 20% for testing. The developed CNN behavior recognition model was trained using the training data. During training, the model underwent 20 epochs, achieving an average training accuracy of approximately 94.72% and a loss value of about 0.17. To further analyze the model's performance, a confusion matrix was employed to visualize its results on the test set. Figure 7 displays the confusion matrix of the trained CNN behavior recognition model.

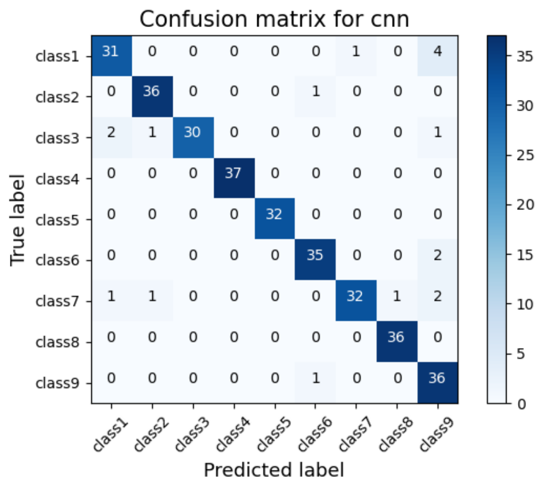


Fig. 7: Confusion matrix from the model trained by the CNN behavior recognition network: class1 = listening (leaning forward), class2 = listening (leaning back), class3 = Raising a hand, class4 = Sleeping, class5 = Standing, class6 = Taking notes, class7 = Whispering, class8 = Using a laptop and class9 = Using a mobile phone

The developed CNN model shows good performance when compared with Resnet18 and Resnet50, as indicated by its recall and precision metrics (see Table 5). Specifically, the model excels in accurately identifying behaviors such as listening (leaning back), sleeping, standing, taking notes, and using a laptop, underscoring its ability to effectively capture and comprehend these behaviors. However, there are some issues with other behavior categories. For instance, the model's ability to recognize the act of listening (leaning forward) shows recall and precision of 86.11% and 91.18%, which may be due to the lack of distinctive features in this behavior, leading to confusion with other behaviors (see Figure 7).

Moreover, instances of raising a hand sometimes get misclassified as listening (leaning forward or back) or using a mobile phone while whispering, which the model categorizes into other behaviors. This suggests a stringent interpretation of features within this category, resulting in high precision but varying recall rates. Furthermore, the model exhibits high recall but low precision for identifying mobile phone usage, indicating limitations in feature recognition, such as the phone's presence in the image. Overall, the deep learning model demonstrates robust accuracy and holds promise for effectively recognizing student behaviors within a classroom setting.

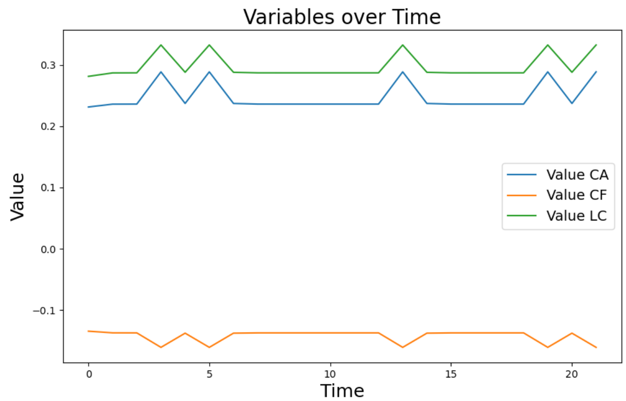
Table 5: Precision and Recall values for Our CNN based model and the benchmark models

Class	Metrics	Our CNN based Model	Resnet18	Resnet50
Listening (leaning forward)	Recall	86.11%	88.89%	88.57%
	Precision	91.18%	69.57%	93.93%
Listening (leaning backwards)	Recall	97.30%	91.89%	94.59%
	Precision	94.74%	97.14%	94.59%
Raising hand	Recall	88.24%	100%	100%
	Precision	100%	89.47%	82.93%
Sleeping	Recall	100%	75.70%	94.59%
	Precision	100%	100%	97.22%
Standing	Recall	100%	100%	100%
	Precision	100%	100%	100%
Taking notes	Recall	94.59%	100%	100%
	Precision	94.59%	100%	100%
Whispering	Recall	86.49%	75.68%	70.27%
	Precision	96.97%	93.33%	89.66%
Using laptop	Recall	100%	100%	100%
	Precision	97.30%	100%	100%
Using mobile phone	Recall	97.30%	100%	100%
	Precision	80%	90.24%	90.24%
Avg. Accuracy		94.72%	91.54%	94.16%

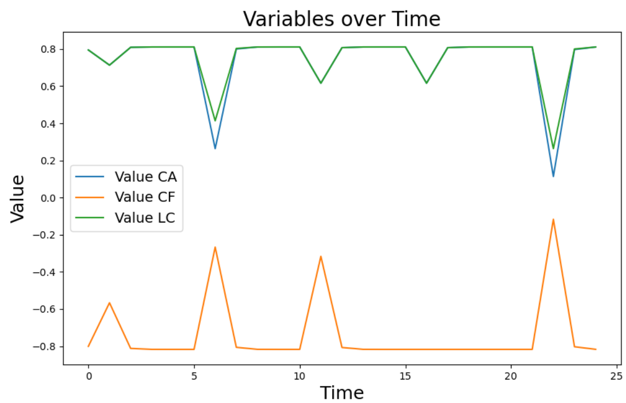
5.2 Students' Group Learning Status Analysis Experimental Results

The recorded video dataset categorizes students' classroom learning status into normal, excellent, and poor atmospheres. In normal classroom videos, half of the students are observed in an "engaged learning status" and the other half "disengaged from the classroom". Excellent classroom videos show all students in an "engaged learning status", whereas poor classroom videos depict all students as "disengaged from the classroom". The classification results are shown in Figure 8a, Figure 8b and Figure 8b. The horizontal axis of the chart represents the number of times images were analyzed, while the vertical axis shows the magnitude of different aspects of the student's learning status, specifically CA (Concentrated-Absent) and LC (Lively-Calm), used for classification.

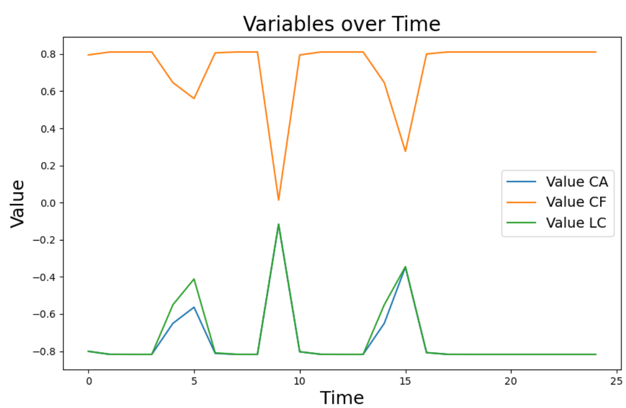
$$\text{Group learning} = \begin{cases} \text{CA} > 0.5 \text{ and LC} > 0.25\text{, Excellent,} \\ \text{CA} < 0.5\text{, Poor,} \\ \text{Others, Normal,} \end{cases} \tag{4}$$



(a) “Normal” classroom atmosphere



(b) “Excellent” classroom atmosphere



(c) “Poor” classroom atmosphere

Fig. 8: Group learning classroom status atmosphere experiment

1. **Normal classroom atmosphere experiment:** The accuracy value for the classification of “Normal” group learning status classification is 100%. Even though Figure 8a shows some fluctuations, the values are not more significant compared to the boundaries for classification as shown in Equation 4. A similar analysis of 'Casual-formal' and 'Lively-calm' values also demonstrated a strong alignment with the expected results (see Figure 8a).
2. **Excellent classroom atmosphere experiment:**The accuracy value for the classification of “Excellent” group learning status classification is 92%. The values for “Concentrated-absent,” “Casual-formal,” and “Lively-calm” suggest that the student group’s learning state was as anticipated: focused, lively, and formal (see Figure 8b).
3. **Poor classroom atmosphere experiment:** The accuracy value for the classification of “Poor” group learning status classification is 92%. The values for “Concentrated-absent,” “Casual-formal,” and “Lively-calm” indicated that the student group’s learning state was unfocused, inactive, and relaxed, as expected (see Figure 8c).

The experimental results above demonstrate that the behavior recognition of the proposed integrated Fuzzy Atomsfield can successfully analyze the learning status of student groups.

6 Conclusion

A method for evaluating classroom group learning status using integrated Fuzzy Atmosfield and machine learning is proposed. First, a custom-made dataset for classroom group learning evaluation is developed. Then, a machine learning model for behavior recognition is developed. Finally, the machine learning model is integrated with the Fuzzy Atmosfield for group classroom evaluation. The results indicate that the CNN-based model can effectively classify students’ classroom behavior. The average accuracy of the CNN-based model is 94.7%. Then, the Fuzzy Atmosfield is used to evaluate the group’s learning status, identifying normal, excellent, and poor classroom atmospheres. The integrated Fuzzy Atmosfield and machine learning method effectively classifies the group learning condition as poor with 92% accuracy, normal with 100% accuracy, and excellent with 92% accuracy. In the proposed method, this paper further classifies behaviors so that more detailed information about learning status can be extracted from behavior recognition. It also uses an adaptation of FA theory to calculate group learning status. The method proposed in this paper, if equipped with further developments like the enrichment of the dataset, should certainly be suitable for applications in real classrooms. Moreover, by changing emotion models and the fuzzy rules, this method can also be applied to study other statuses in different environments.

Bibliography

[1] P. Liuyong, H. Shuke, and C. Min, “Exploration of classroom teaching quality evaluation model based on OBE concept,” *Journal of Higher Education*, vol. 9, no. 29, pp. 102–107, 2023.

[2] J. Yijun and Y. Huiyao, “Overview of classroom behavior recognition methods based on computer vision,” *Automation & Instrumentation*, vol. 09, pp. 1 – 6, 2022.

[3] A. Kavitha, S. K, S. L G, V. J, and V. N, “Framework for detecting student behaviour (nail biting, sleep, and yawn) using deep learning algorithm,” in *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications (AIMLA)*, pp. 1–6, 2024.

[4] D. Yaping, Y. Fangfang, Z. Hanyi, J. Zhiyang, and H. Kaoru, “Auto analysis system of students behavior in mooc teaching,” *Journal of Higher Education*, vol. 4, no. 46, pp. 681–694, 2020.

[5] K. Hirota and F. Dong, “Concept of fuzzy atmosfield and its visualization,” *On Fuzziness: A Homage to Lotfi A. Zadeh–Volume 1*, pp. 257–263, 2013.

[6] R. Fu, T. Wu, Z. Luo, F. Duan, X. Qiao, and P. Guo, “Learning behavior analysis in classroom based on deep learning,” in *2019 Tenth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 206–212, IEEE, 2019.

[7] J. Zhou, F. Ran, G. Li, J. Peng, K. Li, and Z. Wang, “Classroom learning status assessment based on deep learning,” *Mathematical Problems in Engineering*, vol. 2022, pp. 1–9, 2022.

[8] H. Zhou, F. Jiang, J. Si, L. Xiong, and H. Lu, “Stuart: Individualized classroom observation of students with automatic behavior recognition and tracking,” in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, IEEE, 2023.

[9] F. Yang and T. Wang, “Scb-dataset3: A benchmark for detecting student classroom behavior,” *arXiv preprint arXiv:2310.02522*, 2023.