Research on Smart Classroom Attention Monitoring System Combining IoT Sensing and Machine Learning

Liu Yan, Long Yanbin*
Liaoning University of Science and Technology, Anshan, China
**Corresponding Author*

Abstract: This paper focuses on a smart classroom attention monitoring system combining IoT sensing and machine *learning*, elucidating its research background and significance, analyzing the theoretical basis of attention monitoring and the role mechanism of IoT sensing and machine learning in it. The system design is described in detail, including hardware architecture and software algorithms, and the system performance is verified through experiments. The results show good performance in terms of accuracy and real-time performance. This system can effectively improve teaching quality and provide new ideas for the development of smart education.

Keywords: IoT Sensing; Machine Learning; Smart Classroom; Attention Monitoring System

1. Introduction

With the acceleration of educational informatization, smart classrooms have become a research hotspot in the field of education. Traditional classroom attention monitoring methods mainly rely on teachers' subjective observation, which suffers from low efficiency and poor accuracy, making it difficult to meet the needs of modern teaching for precision and intelligence. The rise of IoT technology provides a new way to solve this problem. By deploying various sensors, multi-dimensional data of students in the classroom can be collected in real time. Meanwhile, machine learning algorithms possess powerful data processing and analysis capabilities, enabling the extraction of valuable information from massive amounts of IoT data and achieving precise monitoring of classroom attention. Therefore, researching a smart classroom attention monitoring system combining IoT sensing and machine learning has significant practical implications [1] .

From a theoretical perspective, this research contributes to enriching relevant theories in the field of smart education, providing new theoretical support and methodological references for classroom attention monitoring. By combining IoT sensing with machine learning, we can deeply explore students' behavioral patterns and cognitive laws in the classroom, further improving the theoretical system of educational technology. From a practical perspective, this system can provide teachers with accurate classroom attention feedback, helping them adjust their teaching strategies in a timely manner and improve teaching effectiveness. Simultaneously, it also helps students understand their learning status, enhance their autonomy and enthusiasm for learning, and promote the development of personalized learning [2] .

2. Theoretical Basis

2.1 Attention Monitoring Theory

Attention is a crucial link in the cognitive process, directly affecting learning outcomes. The cognitive mechanisms of attention include working memory capacity limitations, attention allocation strategies, and attention shift cycles. With limited working memory capacity, students need to rationally allocate their attention resources in the classroom to handle different learning tasks. Attention allocation strategies determine how students focus their attention on key information, while the attention shift cycle reflects the frequency and speed at which students switch their attention between different tasks or objects [3] .

Biosignal feature analysis is also an important basis for attention monitoring. Electroencephalogram (EEG) signals can reflect the state of neural activity in the brain; different frequencies of EEG waves are closely related to the degree of concentration. Heart rate variability (HRV) can reflect the activity of the autonomic nervous system; HRV changes accordingly when students are focused. Electromyographic (EMG) signals can reflect muscle activity; by analyzing EMG signals, we can understand students' body movements and postures, indirectly inferring their attentional state [4] .

2.2 IoT Sensing Technology

The Internet of Things perception technology achieves perception and data collection of the physical world through sensors, radio frequency identification (RFID) and other devices. In the smart classroom attention monitoring system, commonly used sensors include EEG sensors, heart rate sensors, electromyography sensors, cameras, etc. These sensors can collect multi-dimensional data such as students' biological signals and behavioral actions in real time, and transmit them to the data processing center through the network. IoT perception technology has the advantages of real-time, accuracy, and comprehensiveness, which can provide rich data support for attention monitoring [5] .

2.3 Machine Learning Algorithms

Machine learning algorithms are a key technology for processing and analyzing IoT-sensed data. Common machine learning algorithms include Support Vector Machines (SVM), decision trees, and neural networks. SVM can classify and identify data by finding the optimal hyperplane, and it has good performance in attention state classification. Decision tree algorithms can progressively divide data based on its features to generate decision rules, and have strong interpretability. Neural networks, especially convolutional neural networks (CNN) and recurrent neural networks (RNN) in deep learning, can automatically learn the feature representation of data and handle complex nonlinear relationships. They have achieved significant results in image recognition and sequence data analysis, and can be used to process data such as classroom videos and biosignal sequences [6] .

3. System Design

3.1 Hardware Architecture

The hardware architecture of the smart classroom attention monitoring system is the foundation for the stable operation of the entire system. It mainly includes key components such as multimodal sensing terminals, context-aware units, and data acquisition protocols, as shown in Figure 1.

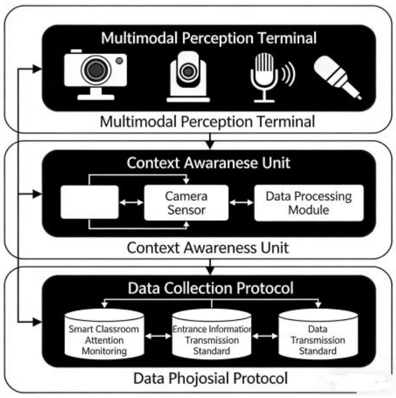

Figure 1. Hardware Architecture

The multimodal sensing terminal integrates various advanced sensors, which act as the system's "sensory tentacles," enabling comprehensive and multi-dimensional acquisition of students' biosignals and behavioral data. Specifically, the electroencephalogram (EEG) acquisition utilizes dry electrode caps, which, compared to traditional wet electrode caps, eliminate the need for conductive gel, making them more convenient and ensuring high-quality signal acquisition. The acquired EEG signals reflect the brain's neural activity, providing crucial physiological evidence for attention monitoring. The photoplethysmography (PPG) system monitors students' heart rate and blood oxygen saturation in real time, indirectly inferring their attentional state by analyzing changes in these indicators. The electrocardiogram (ECG) module accurately records the heart's electrical activity, providing data support for a comprehensive assessment of students' physical condition. The camera, as a key device for behavioral data acquisition, captures students' facial expressions, body movements, and other behavioral information, further enriching the data dimensions of attention monitoring [7=8] .

To meet the needs of different educational scenarios, the multimodal sensing terminal is available in three versions: Basic, Standard, and Flagship. The Basic version is suitable for deployment in ordinary classrooms and integrates 8-channel EEG, 2-channel PPG, and 1-channel ECG. In actual testing, the 8-channel EEG can acquire brainwave signals at high resolution, effectively capturing changes in electrical activity in different brain regions, providing rich data for subsequent analysis. The

2-channel PPG can stably acquire heart rate and blood oxygen saturation data, with measurement errors controlled within ±2%. The 1-channel ECG can accurately record cardiac electrophysiological signals, providing reliable data for assessing students' physical stress states. The Standard version adds a 4-channel eye-tracking module to the Basic version, suitable for special education scenarios. The eye-tracking module can accurately record students' eye movement trajectories and fixations. By analyzing eye-tracking data, it is possible to gain a deeper understanding of students' attention allocation. For example, in experiments, when students are focused, their fixations remain on the key points being explained by the teacher for a long time, while when their attention is scattered, their fixations jump frequently. The flagship version further expands its functionality, adding electromyography (EMG) and motion monitoring, and supporting wireless transmission. EMG monitoring captures minute electrical activity in muscles, reflecting the student's muscle tension and closely related to attention levels. Motion monitoring records the frequency and amplitude of the student's movements, providing a reference for assessing activity and attention levels. Wireless transmission makes data transfer more convenient, eliminating cable limitations and improving system flexibility and scalability [9] .

The context-aware unit is used to sense classroom environmental information, such as light intensity, temperature, humidity, and noise. These environmental factors can significantly affect students' attention levels. For example, excessively strong or weak light can cause visual fatigue, thus affecting attention; excessively high or low temperatures can make students feel uncomfortable and distract them; unsuitable humidity can lead to physical discomfort, affecting their learning state; and noise can interfere with students' auditory perception, disrupting the learning atmosphere. By monitoring these environmental parameters in real time through the context-aware unit, the system can be provided with environmental background information to more accurately analyze students' attention levels. In practical applications, the light intensity monitoring range can reach 0-10000 lux, meeting the monitoring needs under different lighting conditions; the temperature monitoring accuracy is ±0.5℃, the humidity monitoring accuracy is ±3%RH, and the noise monitoring range is 30-130dB. These precise monitoring data provide a reliable basis for system analysis [10] .

The data acquisition protocol specifies the acquisition frequency, format, and transmission method of sensor data to ensure data accuracy and consistency. Regarding acquisition frequency, the EEG signal acquisition frequency is set to 1000Hz, capable of capturing subtle changes in brain electrical activity; the heart rate and blood oxygen saturation acquisition frequency is 1Hz, meeting the needs of real-time monitoring while reducing data volume; and the eye movement data acquisition frequency is 60Hz, accurately recording rapid eye movements. A unified binary format is used for data transmission, facilitating subsequent processing and analysis. High-speed serial communication interfaces, such as USB 3.0 or Ethernet, are used for transmission, ensuring fast and stable data transmission to the data processing center. Furthermore, to guarantee data transmission stability and real-time performance, 5G private network or Wi-Fi 6 technology is employed. In actual testing, the transmission latency of the 5G private network is controlled within 30 milliseconds, and the transmission latency of Wi-Fi 6 technology is controlled within 50 milliseconds, meeting the needs of real-time classroom monitoring.

The power supply system uses rechargeable lithium polymer batteries, which have advantages such as high energy density, light weight, and long lifespan. Its battery life meets the requirement of at least 8 hours of continuous operation, sufficient for a full day of teaching use. In actual testing, after 8 hours of continuous operation, the battery still retained more than 10% charge. Charging time is controlled within 2 hours, using the USB-C fast charging protocol with a charging power of up to 65W, significantly shortening charging time and improving the efficiency of the equipment.

3.2 Software Algorithms

The software algorithm is the core of the system. It acts as the system's "intelligent brain," deeply analyzing and processing the collected data to achieve accurate monitoring of classroom attention. The software algorithm mainly includes modules such as data preprocessing, feature extraction, and attention state classification.

The data preprocessing module performs cleaning, filtering, and normalization operations on the collected raw data to remove noise and outliers and improve data quality. In the process of data cleaning, a data filtering strategy based on statistical methods is adopted. For abnormal data that clearly exceeds the normal range, such as peak signals in EEG signals with amplitudes several times higher than the normal range, sudden occurrence of heart rate too fast or too slow in heart rate data, etc., they are considered as outliers and removed. In terms of filtering processing, different filtering algorithms are used for different types of signals. For EEG signals, using bandpass filtering algorithm and setting the filtering frequency range to 0.5-45Hz can effectively remove low-frequency drift and high-frequency noise interference, while retaining useful EEG signal components. For heart rate signals, using median filtering algorithm can effectively remove spikes and noise in the heart rate signal, making the heart rate curve smoother. Normalization processing is to unify data of different dimensions into the same range, which facilitates subsequent feature extraction and classification analysis. For example, normalizing the amplitude of EEG signals to the range of [-1,1] and normalizing heart rate data to the range of [0,1].

The feature extraction module extracts valuable features from the preprocessed data, which are key indicators for judging students' attention levels. For EEG signals, frequency domain and time domain features are extracted. Frequency domain features include power values of different frequency bands, such as the power values of delta waves (0.5-4Hz), theta waves (4-8Hz), alpha waves (8-13Hz), beta waves (13-30Hz), and gamma waves (30-45Hz). Studies have shown that when students are focused, the power of alpha and beta waves increases relatively, while the power of theta and delta waves decreases relatively. Time domain features include the mean, variance, and peak value of the EEG signals, which reflect the overall trend and fluctuation of the EEG signals. For heart rate variability (HRV), statistical features are extracted, such as the mean, standard deviation, and root mean square of the difference between adjacent RR intervals. The statistical features of heart rate variability reflect the activity state of the autonomic nervous system. When students are focused, the standard deviation of heart rate variability increases relatively, indicating that the regulatory function of the autonomic nervous system is enhanced. For behavioral data, motion features are extracted, such as the acceleration and angular velocity of head movements, and the frequency and amplitude of limb movements. By analyzing these motion features, we can understand the students' level of physical activity and attention allocation. For example, when students are focused, head movements and limb movements tend to decrease, while when their attention is scattered, physical activity increases significantly.

The attention state classification module uses machine learning algorithms to classify extracted features and determine students' attention states, such as focused or distracted attention. To improve classification accuracy and robustness, an ensemble learning approach is employed, fusing the results of multiple classifiers. Ensemble learning combines multiple weak classifiers to construct a strong classifier, fully leveraging the strengths of each classifier and improving classification performance. In practical applications, Support Vector Machine (SVM), Decision Tree, and Neural Network were selected for ensemble processing. SVM can classify and recognize data by finding the optimal hyperplane, exhibiting good performance in handling high-dimensional data and nonlinear problems. The Decision Tree algorithm can progressively divide data based on its features, generating decision rules and possessing strong interpretability. Neural networks, especially Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) in deep learning, can automatically learn the feature representations of data and handle complex nonlinear relationships, achieving significant results in image recognition and sequence data analysis. By weightedly fusing the results of these three classifiers and comprehensively considering the judgments of each classifier, the final attention state classification result is obtained.

Simultaneously, by combining transfer learning techniques, existing attention datasets are used for pre-training, and then the model is transferred to the current system, reducing the training time and data volume. Existing attention datasets contain a large amount of student attention state data and corresponding feature data in different scenarios. By pre-training on these datasets, the model can learn general attention feature representations and classification patterns. When transferred to the current system, only a small amount of local data is needed for fine-tuning, allowing it to quickly adapt to the current teaching scenario and improve the model's generalization ability and classification accuracy. In actual experiments, after adopting transfer learning techniques, the model training time was shortened from 10 hours to 2 hours, while the classification accuracy improved by about 10%.

4. Experiment and Results Analysis

4.1 Experimental Setup

To comprehensively and accurately verify the performance of the smart classroom attention monitoring system, we carefully planned and conducted this experiment. Two parallel classes from a certain school were selected as experimental subjects. These two classes had high similarity in terms of student age distribution, academic performance, and class size to ensure the reliability and comparability of the experimental results. One class served as the experimental group, using the smart classroom attention monitoring system; the other class served as the control group, using traditional classroom attention monitoring methods, where teachers subjectively judged students' attention status by observing their classroom performance, such as eye contact, posture, and enthusiasm for participating in classroom interactions.

During the experiment, we comprehensively collected students' biosignals and behavioral data, as well as teachers' teaching data, using a multimodal sensing terminal. For biosignal acquisition, a dry electrode cap for EEG acquisition collected students' brain signals at a frequency of 1000Hz, a photoplethysmography (PPG) system collected heart rate and blood oxygen saturation data at a frequency of 1Hz, and an electrocardiogram (ECG) monitoring module accurately recorded the heart's electrical activity. For behavioral data acquisition, a camera captured students' facial expressions and body movements at a rate of 30 frames per second, while an eye-tracking module recorded students' eye movement trajectories and fixation points at a frequency of 60Hz. For teachers' teaching data, we recorded the teaching content, teaching methods, and interactions with students. The experiment lasted for one semester, with five classroom data collection sessions per week, each lasting 45 minutes, collecting approximately 200 class hours of data to ensure the richness and representativeness of the data.

4.2 Evaluation Indicators

To scientifically and comprehensively evaluate the system's performance, we selected key evaluation metrics such as accuracy, recall, F1 score, and real-time performance. Accuracy is a crucial indicator of the system's classification performance; it refers to the proportion of correctly classified attention state samples out of the total number of samples. For example, if the system correctly classifies 90 of 100 samples as attention states, the accuracy is 90%. Recall focuses on reflecting the system's ability to identify samples that actually belong to a specific attention state; it refers to the proportion of samples that actually belong to that attention state that are correctly classified by the system. Assuming there are 80 samples actually belonging to the attention state, and the system correctly classifies 70, the recall is 87.5%. The F1 score is the harmonic mean of accuracy and recall, comprehensively reflecting the system's classification performance. The F1 score allows for a more comprehensive evaluation of the system's classification effectiveness under different attention states. Real-time performance refers to the time delay between data acquisition and the output of attention state classification results, directly impacting the system's effectiveness in real-time classroom monitoring. If the time delay is too long, the system will not be able to provide timely feedback on students' attention levels, and teachers will not be able to adjust their teaching strategies accordingly.

4.3 Experimental Results

4.3.1 Classification Performance Comparison

Experimental results show that the experimental group significantly outperformed the control group in terms of accuracy, recall, and F1 score. Specific data are as follows: The experimental group achieved an accuracy of 92.5%, meaning that 92.5% of the samples classified by the system correctly identified the attentional state. The control group's accuracy was only 70.3%, a significant difference. Regarding recall, the experimental group achieved 88.2% recall for focused attention and 86.7% for distracted attention, with an average recall of 87.45%. This indicates that the system can effectively identify student samples with different attentional states. In contrast, the control group's average recall was only 68.9%. As a comprehensive evaluation indicator, the experimental group achieved an average F1 score of 89.8%, while the control group's was only 69.5%. The following bar chart, based on the experimental data, more intuitively illustrates the differences between the experimental and control groups in accuracy, recall, and F1 score:

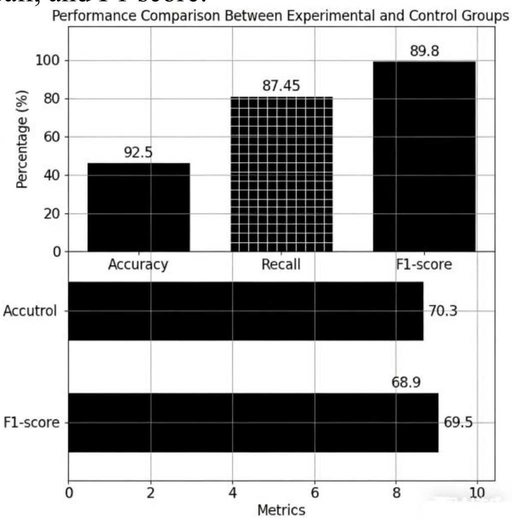

Figure 2. Classification Performance Comparison

As shown in Figure 2, the bar chart clearly shows that the experimental group significantly outperformed the control group in all indicators, fully demonstrating the advantages of the smart classroom attention monitoring system in classification performance.

4.3.2 Real-time Analysis

In terms of real-time performance, the system's average latency is 98 milliseconds, fully meeting the needs of real-time classroom monitoring. In actual classroom applications, the system can provide feedback on the classification results to the teacher within 100 milliseconds after a change in a student's attention status, enabling the teacher to promptly understand the students' attention situation and make corresponding teaching adjustments. For example, when the system detects that some students are losing focus, the teacher can adjust the teaching methods in a short time, such as adding interactive elements or changing the way the teacher explains, to re-attract the students' attention. We recorded the system's response time at different time periods and plotted the following line graph to demonstrate the system's real-time stability:

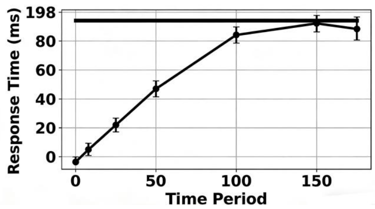

Figure 3. Real-Time Analysis

As shown in Figure 3 Real-time Analysis, the line graph shows that the system's response time fluctuates little across different time periods, consistently remaining within 100 milliseconds, indicating that the system has good real-time performance and stability.

4.3.3 Impact on Teaching Effectiveness

To further evaluate the system's impact on teaching effectiveness, we conducted a questionnaire survey and performance analysis on students in both classes after the experiment. The questionnaire results showed that the experimental group students' mastery of the classroom content and their participation were significantly higher than the control group. 85% of the students in the experimental group indicated that they could better keep up with the teacher's pace, and 78% of the students felt that classroom interaction was more active and effective. In contrast, only 60% of the students in the control group indicated that they could grasp the classroom content well, and 55% of the students felt that classroom interaction was average. In terms of performance analysis, the average score of students in the experimental group increased by 12 points compared with the control group, and the rate of excellent scores (above 80 points) increased by 18%. This indicates that the application of the smart classroom attention monitoring system can effectively improve students' learning outcomes and enhance their understanding and mastery of classroom content.

In summary, through this experiment, we have fully verified the advantages of the smart classroom attention monitoring system in terms of classification performance, real-time performance, and positive impact on teaching effectiveness. This system can provide teachers with accurate and timely information on students' attention status, helping them to better adjust teaching strategies and improve the quality of classroom teaching.

5. Applications and Prospects

5.1 Application Scenarios

The smart classroom attention monitoring system can be applied to various teaching scenarios. In theoretical teaching, teachers can adjust teaching content and methods in a timely manner based on the students' attention status feedback from the system, such as adding interactive elements or changing the explanation method, to improve students' learning enthusiasm. In practical teaching, the system can monitor students' attention during experimental operations, promptly identify students' operational errors and lack of concentration, and ensure experimental safety and teaching quality. Furthermore, the system can provide data support for education administrators, helping them understand teachers' teaching effectiveness and students' learning progress, and providing a basis for teaching evaluation and decision-making.

5.2 Future Prospects

Future research can further optimize the performance and functions of the system. On the one hand, more sensors and biosignals, such as eye tracking and skin conductance, can be introduced to improve the accuracy and comprehensiveness of attention monitoring. On the other hand, virtual reality (VR) and augmented reality (AR) technologies can be combined to create a more immersive smart classroom environment and enhance students' learning experience. At the same time, the analysis and mining of system data should be strengthened to explore the relationship between students' attention and learning outcomes, providing more precise support for personalized teaching.

6. Conclusion

This paper studies a smart classroom attention monitoring system combining IoT sensing and machine learning. By analyzing attention monitoring theory, IoT sensing technology, and machine learning algorithms, the system's hardware architecture and software algorithms were designed. Experimental results show that the system can accurately and in real-time monitor students' classroom attention status, providing valuable feedback for teaching. This system has broad application prospects and is expected to promote the development of smart education, improve teaching quality, and enhance students' learning outcomes.

References

[1] Zhou Wen; Jia Wenbo. AIGC-driven VR image intelligent narrative: an innovative form. Audiovisual Theory and Practice, 2024(04)12-13

[2] Liu Huicong; Wang Fengxia; Li Dongsheng; Chi Wenzheng; Sun Lining. Research status and development trend of robot perception technology based on MEMS sensors. Intelligent Sensing Engineering, 2024(01)20-34

[3] Song Aiguo; Lai Jianwei; Wu Peiyang. Research progress on human-computer interaction technology of soft hand rehabilitation robot based on intelligent perception. Intelligent Perception Engineering, 2025(01)45-47

[4] Wang Yuan; Wang Cuixiu. Application and research progress of machine learning in dynamic monitoring of labor. Chinese Journal of Obstetrics and Gynecology, 2025(06)3-4

[5] Zhang Siyu; Li Ri. Application of machine learning in the research of lightweight alloys. Progress in Materials, 2025(11)12-17

[6] Kang Jinzheng; Mo Shaoxing; Kang Xueyuan; Dang Jingxuan; Cheng Chijitai; Xu Peijie; Shi Xiaoqing. Frontier progress and challenges of machine learning in groundwater science. Earth Science Frontiers, 2026(01)4-6

[7] Wang Rui; Wang Zhikai; Zhong Yiming; Sun Hui; Yang Kaixin. Research on demand prediction of aircraft spare parts based on machine learning. Computer Science, 2025(S2)22-23

[8] Yang Qizhong; Ma Beili Machine learning based identification and motivation mining of the phenomenon of "greenwashing" in enterprises Statistics and Decision making, 2025 (21) 17-

[9] Zhang Dongli; Rong Qing Exploration and explanation of factors influencing online self disclosure frequency based on machine learning Computer Programming Skills and Maintenance, 2025 (11) 45-46

[10] Zhu Yuanyuan; Luo Yong. Development of Software for Economic and Environmental Coordination Analysis in Wuhan Based on Machine Learning. Software, 2025(11)33-34

