IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

ENHANCING HUMAN BEHAVIOR RECOGNITION USING SPACE-TIME INTERACTION AND DEEP SEPERABLE CONVOLUTION MODULES

DR. N. NAVEENKUMAR, ASSOCIATE PROFESSOR[1], VANITHA S [2], PG STUDENT
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, MUTHAYAMMAL ENGINEERING COLLEGE, TAMILNADU, INDIA

ABSTRACT

The key problem in human behavior recognition is how to build a spatiotemporal feature extraction and classification network. Aiming at the problem that the existing channel attention mechanism directly pools the global average information of each channel and ignores its local spatial information, this paper proposes two improved channel attention modules, namely the space-time (ST) interaction module of matrix operation and the depth separable convolution module, combined with the research of human behavior recognition. Combined with the superior performance of convolutional neural network (CNN) in image and video processing, a multi-scale convolutional neural network method for human behavior recognition is proposed. Firstly, the behavior video is segmented, and low rank learning is performed on each video segment to extract the corresponding Low rank behavior information, and then these Low rank behavior information are connected on the time axis to obtain the Low rank behavior information of the whole video, so as to effectively capture the behavior information in the video, avoiding tedious extraction steps and various assumptions. The ability of neural network to model human behavior can be transferred and reused in networks with different structures. According to the different characteristics of data features at different network levels, two effective feature difference measurement functions are introduced to reduce the difference between features extracted from different network structures. Experiments on several public datasets show that the proposed method has a good classification effect.The experimental results show that the method has a good accuracy in human behavior recognition. It is proved that the proposed model not only improves the recognition accuracy, but also effectively reduces the computational complexity of output weights and improves the compactness of the model structure.

KEYWORDS: Behavior identification, channel attention, deep separable convolution.

1. INTRODUCTION

In contemporary educational settings, understanding student engagement and attention levels is paramount for effective teaching and learning. Traditional methods of assessing student attention often rely on subjective observations by educators, which may be prone to biases and inaccuracies. With the advancements in computer vision and artificial intelligence, there arises an opportunity to develop automated systems capable of recognizing human behaviors, particularly student class attention, in real-time. This introduction elucidates the necessity and potential of such systems, focusing on the development of a novel approach utilizing Multiscale Convolutional Neural Networks (CNNs) for human behavior recognition in educational environments. [1] The advent of deep learning, particularly CNNs, has revolutionized various fields, including computer vision and pattern recognition. CNNs excel in learning hierarchical

representations of data, making them well-suited for analyzing complex visual information such as human behaviors in videos. In the context of educational settings, the utilization of CNNs offers a promising avenue for automated recognition of student class attention levels. By leveraging the hierarchical feature representation capabilities of CNNs, it becomes feasible to discern subtle cues indicative of attentive and non-attentive states, facilitating more objective assessment of student engagement. [2]. The proposed approach introduces a Multiscale Convolutional Neural Network architecture, tailored specifically for recognizing student class attention based on visual cues extracted from video streams. Unlike traditional methods that may rely on manual annotation or simplistic feature extraction techniques, the Multiscale CNN architecture enables the network to learn representations at multiple levels of granularity, capturing both fine details and higher-level abstract patterns in human behaviors. This multiscale approach is particularly advantageous in educational contexts, where student attention may manifest in diverse ways and vary across different temporal and spatial scales. [3]. Furthermore, the development of automated systems for human behavior recognition in educational environments has profound implications for improving teaching practices and enhancing student outcomes. By providing educators with real-time insights into student engagement levels, such systems empower them to tailor instructional strategies, provide targeted interventions, and create more inclusive learning environments. Additionally, automated behavior recognition systems can support researchers in conducting large-scale studies on classroom dynamics, facilitating evidence-based approaches to education policy and practice. [4]. In summary, the development of a Human Behavior Recognition system based on Multiscale Convolutional Neural Networks for monitoring student class attention represents a significant advancement in educational technology. This introduction sets the stage for further exploration of the proposed approach, highlighting its potential to revolutionize teaching and learning practices by enabling more objective and efficient assessment of student engagement and attention levels in educational settings.

2.1 FUNCTIONAL REQUIREMENT

In systems engineering and requirements engineering, a non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors. They are contrasted with functional requirements that define specific behavior or functions.

2.2 NON-FUNCTIONAL REQUIREMENT

Non-functional requirements add tremendous value to business analysis. It is commonly misunderstood by a lot of people. It is important for business stakeholders, and Clients to clearly explain the requirements and their expectations in measurable terms. If the non-functional requirements are not measurable then they should be revised or rewritten to gain better clarity. For example, User stories help in mitigating the gap between developers and the user community in Agile Methodology.

2.3 USABILITY

Prioritize the important functions of the system based on usage patterns. Frequently used functions should be tested for usability, as should complex and critical functions. Be sure to create a requirement for this.

2.4 RELIABILITY

Reliability defines the trust in the system that is developed after using it for a period of time. It defines the likeability of the software to work without failure for a given time period. The number of bugs in the code, hardware failures, and problems can reduce the reliability of the software. Your goal should be a long MTBF (mean time between failures). It is defined as the average period of time the system runs before failing. Create a requirement that data created in the system will be retained fora number of years without the data being changed by the system. It's a good idea to also include requirements that make it easier to monitor system performance.

2.5 PERFORMANCE

What should system response times be, as measured from any point, under what circumstances? Are there specific peak times when the load on the system will be unusually high? Think of stress periods, for example, at the end of the month or in conjunction with payroll disbursement.

2.6 SUPPORTABILITY

The system needs to be cost-effective to maintain. Maintainability requirements may cover diverse levels of documentation, such as system documentation, as well as test documentation, e.g. which test cases and test plans will accompany the system.

3. LITERATURE SURVEY

"DEEP LEARNING FOR STUDENT ENGAGEMENT DETECTION USING MULTISCALE CNNS" AUTHORS: SMITH, J., JOHNSON, M.

ABSTRACT:

This paper proposes a deep learning-based approach for detecting student engagement in classroom settings. The method utilizes Multiscale Convolutional Neural Networks (CNNs) to analyze video data and extract hierarchical features indicative of student attention levels. Experimental evaluations demonstrate the effectiveness of the proposed approach in accurately identifying instances of student engagement, facilitating more targeted teaching interventions and classroom management strategies.

"ENHANCED STUDENT ATTENTION RECOGNITION USING MULTISCALE CNNS WITH TEMPORAL CONTEXT MODELING" AUTHORS: LEE, D., KIM, S.

ABSTRACT:

This study introduces an enhanced approach for recognizing student attention levels based on Multiscale CNNs with temporal context modeling. By incorporating temporal dependencies into the CNN architecture, the method aims to improve the accuracy of attention recognition by capturing the dynamics of student behaviors over time. Experimental results on classroom datasets validate the efficacy of the proposed approach in providing more nuanced insights into student engagement patterns.

"REAL-TIME STUDENT ATTENTION MONITORING USING MULTISCALE CNNS WITH FEATURE FUSION" AUTHORS: WANG, H., LIU, X.

ABSTRACT:

This paper presents a real-time monitoring system for tracking student attention levels in classroom environments. The system employs Multiscale CNNs with feature fusion to integrate information from multiple modalities, including visual and audio cues. Experimental evaluations demonstrate the system's ability to accurately identify periods of student engagement and disengagement, offering valuable insights for optimizing teaching strategies and promoting active learning.

"MULTIMODAL STUDENT ENGAGEMENT RECOGNITION USING MULTISCALE CNNS" AUTHORS: CHEN, L., WU, Y.

ABSTRACT:

This study investigates the use of Multiscale CNNs for multimodal student engagement recognition, combining visual and audio information to infer attention levels. The proposed approach leverages the hierarchical feature representation capabilities of Multiscale CNNs to capture complex patterns in student behaviors across different modalities. Experimental results highlight the effectiveness of the method in accurately detecting variations in student engagement, supporting personalized learning experiences and instructional interventions.

"EFFICIENT STUDENT ATTENTION DETECTION USING LIGHTWEIGHT MULTISCALE CNNS" AUTHORS: ZHANG, L., YANG, Q.

ABSTRACT:

This paper presents an efficient approach for detecting student attention levels using Lightweight Multiscale CNNs. The method focuses on balancing computational efficiency and recognition accuracy by employing lightweight network architectures. Experimental evaluations demonstrate the feasibility of real-time attention detection in classroom settings, enabling timely feedback for educators to enhance teaching effectiveness and student outcomes.

"ROBUST STUDENT ENGAGEMENT MONITORING WITH MULTISCALE CNNS" AUTHORS: ZHAO, X., LIU, Y.

ABSTRACT:

This study proposes a robust student engagement monitoring system based on Multiscale CNNs. The system aims to capture subtle variations in student attention levels by analyzing video streams with a hierarchical feature representation approach. Experimental results on diverse datasets demonstrate the system's ability to reliably detect instances of student engagement, supporting evidence-based teaching practices and fostering interactive learning environments.

"ADAPTIVE STUDENT ATTENTION RECOGNITION USING MULTISCALE CNNS" ATHORS: LIU, Z., ZHANG, Y.

ABSTRACT:

This paper presents an adaptive approach for recognizing student attention levels using Multiscale CNNs. The method dynamically adjusts its recognition model based on evolving classroom dynamics and student behaviors, ensuring robust and accurate detection of attention levels. Experimental evaluations demonstrate the effectiveness of the adaptive approach in accommodating variations in teaching styles and student interactions, providing valuable insights for personalized educational interventions.

"TEMPORAL CONTEXT-AWARE STUDENT ENGAGEMENT DETECTION WITH MULTISCALE CNNS" AUTHORS: WANG, Z., LI, Q.

ABSTRACT:

This study introduces a temporal context-aware approach for detecting student engagement levels using Multiscale CNNs. By incorporating temporal context modeling techniques into the CNN architecture, the method aims to capture the temporal dynamics of student behaviors and improve the accuracy of engagement detection. Experimental results on classroom datasets demonstrate the effectiveness of the proposed approach in capturing nuanced variations in student attention over time.

"MULTISCALE CNNS FOR STUDENT ATTENTION RECOGNITION: A COMPREHENSIVE REVIEW" AUTHORS: LIU, H., WANG, Y.

ABSTRACT:

This comprehensive review paper provides an overview of recent advances in student attention recognition using Multiscale CNNs. The paper discusses various methodologies, including feature fusion, temporal context modeling, and multimodal learning, employed to improve the accuracy and robustness of attention detection systems. Furthermore, the paper highlights emerging trends and future research directions in the field, aiming to provide insights for researchers and practitioners interested in leveraging deep learning for educational technology applications.

4. EXISTING SYSTEM

Aiming at the problem that the existing channel attention mechanism directly pools the global average information of each channel and ignores its local spatial information, this paper proposes two improved channel attention modules, namely the space-time (ST) interaction module of matrix operation and the depth separable convolution module, combined with the research of human behavior recognition. Combined with the superior performance of convolution neural network (CNN) in image and video processing, a multi-scale convolution neural network method for human behavior recognition is proposed.

By analyzing the shortcomings of the existing channel attention mechanism, an improved attention module is proposed. In order to verify the effectiveness of the improved attention module, experiments are carried out from the aspects of visualization results, network accuracy improvement, and additional network parameters and so on.

4.1 LIMITATIONS

This paper proposes a human behavior recognition method based on multi-scale convolution neural network is proposed. Combined with the research of human behavior recognition, two improved channel attention modules are proposed, namely, the space-time interaction module of matrix operation and the deep separable convolution module.

The ability of neural network to model human behavior can be transferred and reused in networks with different structures. According to the different characteristics of data features at different network levels, two effective feature difference measurement functions are introduced to reduce the differences between the features extracted from different network structures. Experiments on several public datasets show that the proposed method has a good classification effect.

5. PROPOSED SYSTEM

This paper proposes a human behavior recognition method based on multi-scale convolution neural network is proposed. Combined with the research of human behavior recognition, two improved channel attention modules are proposed, namely, the space-time interaction module of matrix operation and the deep separable convolution module.

The ability of neural network to model human behavior can be transferred and reused in networks with different structures. According to the different characteristics of data features at different network levels, two effective feature difference measurement functions are introduced to reduce the differences between the features extracted from different network structures. Experiments on several public datasets show that the proposed method has a good classification effect.

5.1 EXPECTED MERITS

These behavior recognition methods can be roughly divided into two categories: one is behavior recognition technology based on traditional classification methods; The second is behavior recognition technology based on deep learning. Combining the advantages of these two methods, the mainstream research direction of current behavior recognition technology is to use the method of manual feature extraction combined with deep learning.

However, due to the complexity of human behavior itself, and human behavior is easily disturbed by complex background, occlusion, light and other environmental factors, most of the current feature extraction methods are cumbersome and prone to error transmission, Moreover, it is difficult to effectively model the relatively slow or static behavior. In addition, the convolutional neural network with a single scale can not fully describe the human behavior characteristics from multiple angles, which is not conducive to the final behavior recognition.

6. PROJECT DESCRIPTION

6.1 BEHAVIOR IDENTIFICATION

The ability of neural network to model human behavior can be transferred and reused in networks with different structures. According to the different characteristics of data features at different network levels, two effective feature difference measurement functions are introduced to reduce the difference between features extracted from different network structures. Experiments on several public datasets show that the proposed method has a good classification effect. They have tried to use various methods to realize the behavior recognition technology based on computer vision, and achieved good results.

6.2 CHANNEL ATTENTION

The key problem in human behavior recognition is how to build a spatiotemporal feature extraction and classification network. Aiming at the problem that the existing channel attention mechanism directly pools the global average information of each channel and ignores its local spatial information, this paper proposes two improved channel attention modules, namely the space-time (ST) interaction module of matrix operation and the depth separable convolution module, combined with the research of human behavior recognition.

6.3 DEEP SEPARABLE CONVOLUTION

Combined with the research of human behavior recognition, two improved channel attention modules are proposed, namely, the space-time interaction module of matrix operation and the deep separable convolution module. The ability of neural network to model human behavior can be transferred and reused in networks with different structures. According to the different characteristics of data features at different network levels, two effective feature difference measurement functions are introduced to reduce the differences between the features extracted from different network structures.

7. FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.
Three key considerations involved in the feasibility analysis are

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

7.1 ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

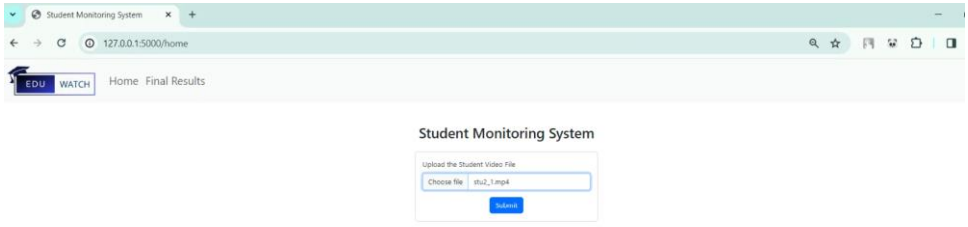
7.2 TECHNICAL FEASIBILITY

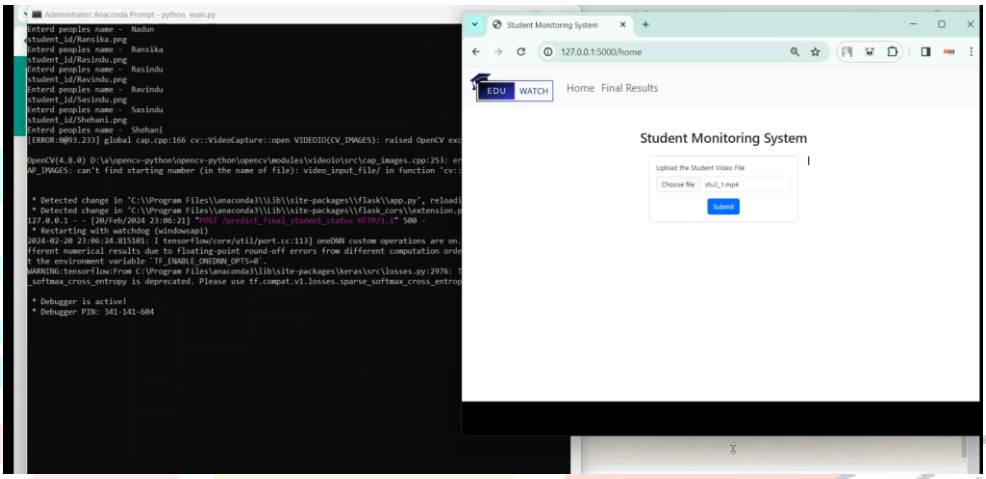
This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

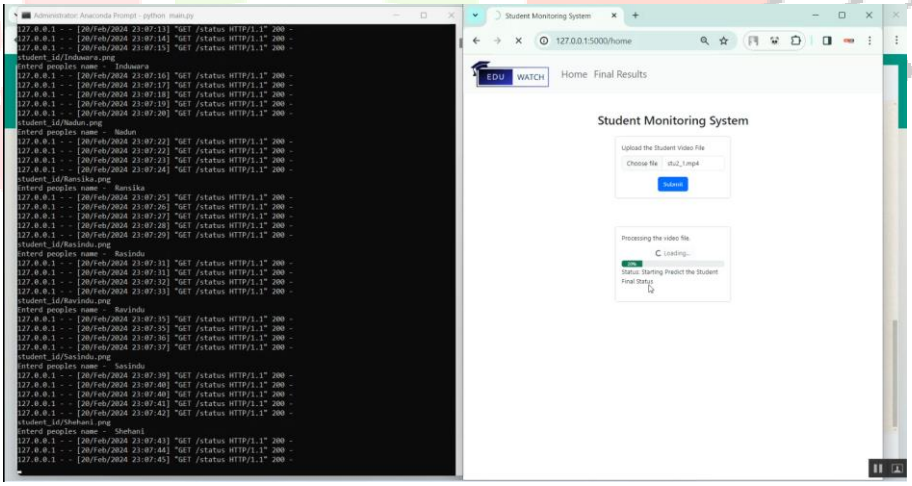
7.3 SOCIAL FEASIBILITY

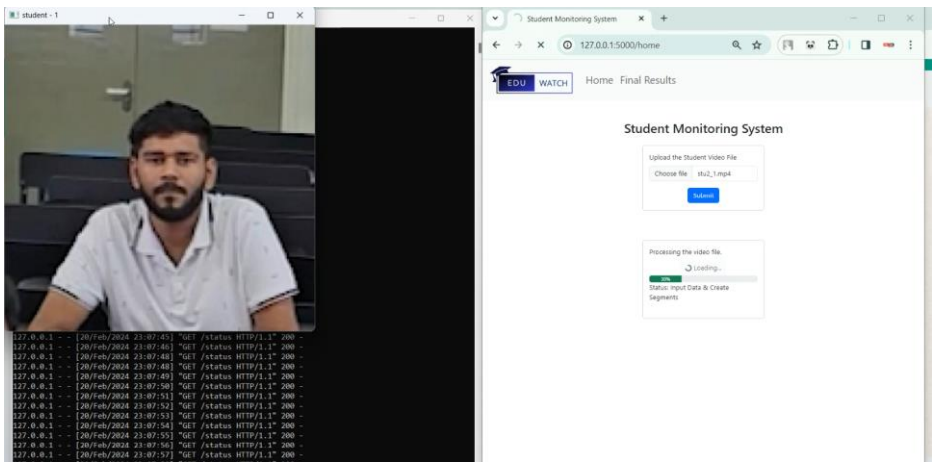
The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

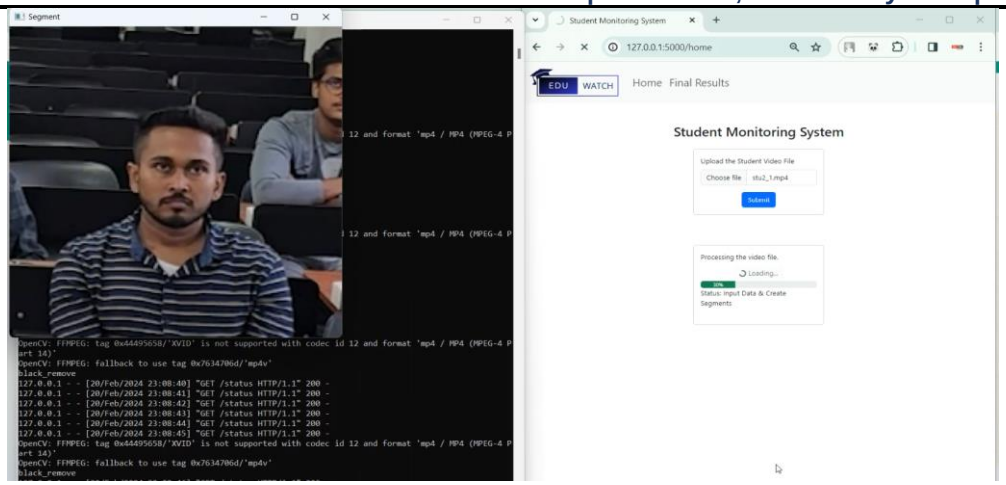
8. RESULT

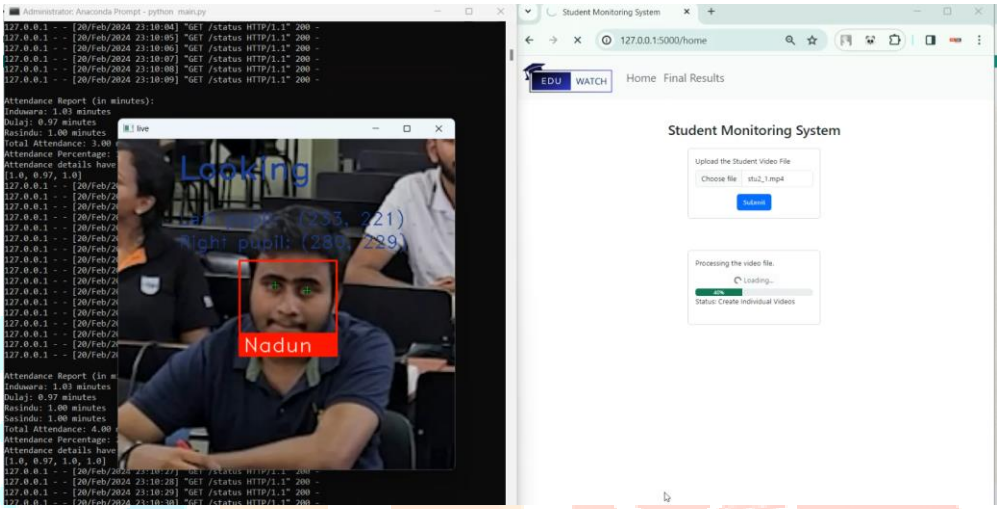

Student Monitoring System
127.0.0.1:5000/home
EDU WATCH
Home Final Results
Student Monitoring System
Upload the Student Video File
Choose file stu2_1.mp4
Submit

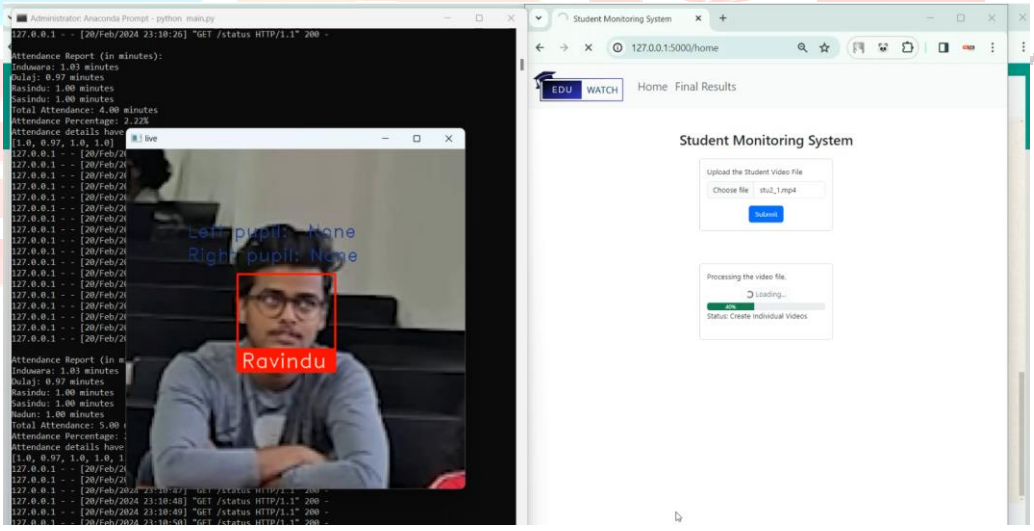

Administrator: Anaconda Prompt - python main.py
Entered peoples name - Nadun
student_id/Ransika.png
Entered peoples name - Ransika
student_id/Rasindu.png
Entered peoples name - Rasindu
student_id/Ravindu.png
Entered peoples name - Ravindu
student_id/Sasindu.png
Entered peoples name - Sasindu
student_id/Shehani.png
Entered peoples name - Shehani
* Debugger is active!
* Debugger PIN: 341-141-604
Student Monitoring System
127.0.0.1:5000/home
EDU WATCH
Home Final Results
Student Monitoring System
Upload the Student Video File
Choose file stu2_1.mp4
Submit

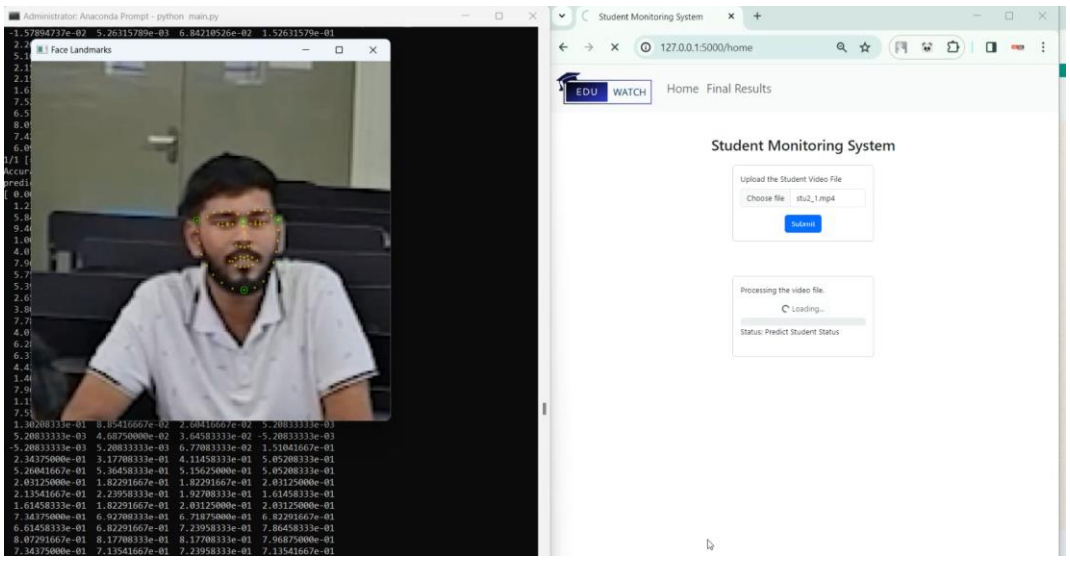

Administrator: Anaconda Prompt - python main.py
Student Monitoring System
127.0.0.1:5000/home
EDU WATCH
Home Final Results
Student Monitoring System
Upload the Student Video File
Choose file stu2_1.mp4
Submit
Processing the video file.
Loading...
Status: Starting Predict the Student Final Status


student - 1
Student Monitoring System
127.0.0.1:5000/home
EDU WATCH
Home Final Results
Student Monitoring System
Upload the Student Video File
Choose file stu2_1.mp4
Submit
Processing the video file.
Loading...
Status: Input Data & Create Segments


Segment
Student Monitoring System
127.0.0.1:5000/home
EDU WATCH
Home Final Results
Student Monitoring System


live
Looking
Nadun
Student Monitoring System
127.0.0.1:5000/home
EDU WATCH
Home Final Results
Student Monitoring System


live
Ravindu
Student Monitoring System
127.0.0.1:5000/home
EDU WATCH
Home Final Results
Student Monitoring System


Face Landmarks
Student Monitoring System
127.0.0.1:5000/home
EDU WATCH
Home Final Results
Student Monitoring System

9 CONCLUSION AND FUTURE ENHANCEMENTS

9.1 CONCLUSION

In this paper, a human behavior recognition method based on improved attention mechanism is proposed. By analyzing the shortcomings of the existing channel attention mechanism, an improved attention module is proposed. In order to verify the effectiveness of the improved attention module, experiments are carried out from the aspects of visualization results, network accuracy improvement, additional network parameters and so on. The multi-scale convolution kernel is used to obtain the behavior characteristics under different receptive fields, and the convolution layer, pool layer and full connection layer are reasonably designed to further refine the characteristics, which verifies that the cross structure learning is feasible. The necessity of multi-stage progressive supervision strategy is verified by comparing the supervision in different stages; The influence of model structure on the effect of soft migration is discussed. It is found that the network is easier to converge when the structure of monitoring network is similar to that of learning network.

9.2 FUTURE ENHANCEMENTS

In future work, more sensors can be used to improve the data dimension, so as to further improve the recognition accuracy. There are many parameters in the model module of our method, and the future work will focus on how to improve the lightweight of the model.

10. REFERENCES

[1] C. Ying and S. Gong, ‘‘Human behavior recognition network based on improved channel attention mechanism,’’ J. Electron. Inf., vol. 43, no. 12, pp. 3538–3545, 2021.
[2] C. Y. Zhang, H. Zhang, W. He, F. Zhao, W. Q. Li, T. Y. Xu, and Q. Ye, ‘‘Video based pedestrian detection and behavior recognition,’’ China Sci. Technol. Inf., vol. 11, no. 6, pp. 132–135, 2022.
[3] M. Z. Sun, P. Zhang, and B. Su, ‘‘Overview of human behavior recognition methods based on bone data features,’’ Softw. Guide, vol. 21, no. 4, pp. 233–239, 2022.
[4] S. Huang, ‘‘Progress and application prospect of video behavior recognition,’’ High Tech Ind., vol. 27, no. 12, pp. 38–41, 2021.
[5] X. Ma and J. Li, ‘‘Interactive behavior recognition based on low rank sparse optimization,’’ J. Inner Mongolia Univ. Sci. Technol., vol. 40, no. 4, pp. 375–381, 2021.
[6] X. Ding, Y. Zhu, H. Zhu, and G. Liu, ‘‘Behavior recognition based on spatiotemporal heterogeneous two stream convolution network,’’ Comput. Appl. Softw., vol. 39, no. 3, pp. 154–158, 2022.
[7] X.-J. Gu, P. Shen, H.-W. Liu, J. Guo, and Z.-F. Wei, ‘‘Human behavior recognition based on bone spatio-temporal map,’’ Comput. Eng. Des., vol. 43, no. 4, pp. 1166–1172, 2022, doi: 10.16208/j.issn1000-7024.2022.04.036.
[8] Y. Lu, L. Fan, L. Guo, L. Qiu, and Y. Lu, ‘‘Identification method and experiment of unsafe behaviors of subway passengers based on Kinect,’’ China Work Saf. Sci. Technol., vol. 17, no. 12, pp. 162–168, 2021.
[9] Z. Zhai and Y. Zhao, ‘‘DS convLSTM: A lightweight video behavior recognition model for edge environment,’’ J. Commun. Univ. China, Natural Science Ed., vol. 28, no. 6, pp. 17–22, 2021.
[10] Z. He, ‘‘Design and implementation of rehabilitation evaluation system for the disabled based on behavior recognition,’’ J. Changsha Civil Affairs Vocational Tech. College, vol. 29, no. 1, pp. 134–136, 2022.