Using CNN, GRU, and Bidirectional Multiscale Convolutional Neural Networks for Human Behavior Recognition

Mr. MV. Satyanarayana [1], K. Sai Bhavani [2], K. Keerthi [3], J. Sreeja [4]

[1] Assistant Professor, Department of CSE, Malla Reddy Engineering College for Women, Autonomous, Hyderabad

[2],[3],[4]Student, Department of CSE, Malla Reddy Engineering College for Women, Autonomous, Hyderabad

ABSTRACT:

The main challenge in recognizing human behavior is constructing a network for the extraction and categorization of spatiotemporal features. In order to address the issue that the current channel attention mechanism simply aggregates each channel's global average information while ignoring its specific spatial information, this work suggests two enhanced channel attention modules: the depth separable convolutions section and the time-space (ST) interaction section of matrices operation. These modules are also combined with research on the recognition of human behavior. Proposing a multiple habitats convolutional neural network technique for human behavior detection, it is combined with the excellent performance using convolutional neural network (CNN) for video and image processing. First, the behavior video is divided into segments. Next, low rank learning is applied to each segment to extract the associated low rank actions information. Finally, these minimal position behavior information are linked together in the time axis to get the low are behavior data for the entire video. This allows for the efficient extraction of behavior information from the video without the need for laborious extraction processes or assumptions. Neural networks can simulate human behavior in a variety of network topologies by transferring and reusing this capacity. To lessen the distinction between features derived from various network topologies, two efficient feature difference measurement methods are presented, taking into account the various properties of data features at various network levels. The suggested strategy has a decent categorization impact, according to experiments on a number of available datasets. The experimental findings demonstrate that the method's accuracy in identifying human behavior is excellent. It

has been shown that the suggested model increases recognition accuracy while simultaneously enhancing the compactness for the model structure and successfully lowering the computational cost of the output weights.

INTRODUCTION

Human behavior recognition research has the potential to advance computer vision engineering applications in addition to providing a foundation for relevant theory. The theoretical underpinnings of behavior recognition draw from a wide range of fields, including biology, artificial intelligence, computer vision, human kinematics, and image processing. Recognizing human behavior in video information is a crucial use of computer vision technology. This is a crucial line of inquiry.

Deep learning-based behavior identification techniques may be categorized into two groups depending on the kind of convolution kernels used: Numerous researchers have used deep learning to recognize motion using 2D and 3D convolution networks. They have made successful attempts to implement the computer vision-based behavior detection technology using a variety of techniques. Chapter 1 will provide a detailed analysis of the particular techniques and literature. Generally speaking, there are two types of behavior recognition technologies: deep learning-based behavior recognition technology and behavior identification technology based on conventional categorization techniques. The prevailing research path in behavior recognition technology nowadays is to integrate deep learning with manual feature extraction, therefore combining the benefits of these two approaches. However, because human behavior is so complex and can be easily influenced by complicated background, occlusion, light, and other environmental variables, the majority of the current methods for extracting features are laborious and susceptible to error transmission. In addition, it can be challenging to accurately model the comparatively slow and still behavior. Furthermore, a single-scale convolutional neural network is unable to adequately capture the aspects of human behavior from various perspectives, which hinders the ability to recognize behavior in its whole. Numerous effective network architectures, such C3R, eco, TSN, and others, have surfaced in domain study. Despite structural differences, all of these network models exhibit strong video data modeling capabilities and are capable of accurately identifying various human activities in real environments. Theoretically, (using the classification job as an example), feature description vectors produced from various network models are dependent on category information and become linearly separate at the network's output layer. The resulting feature vectors must to be comparable, albeit originating from distinct modeling procedures. The question of whether information obtained by various network architectures may be shared and learnt from is one that merits discussion. The original network's breadth and depth were expanded by Chen et al., who also achieved cross-structure transfer learning by initializing the weight parameters using the unit matrix or the decomposition corresponding to the original parameters. In order to make the 3D network suit to the output feature distributions provided by the 2D network, Ali et al. utilized a 2D network to monitor both the input and the output of a 3D network. This allowed them to indirectly achieve cross-structure learning. Motivated by this, this article further loosens the model structure's limitations and employs efficient measurement techniques across the two network with larger structural disparities in order to accomplish "soft transfer," a more broadly defined kind of transfer learning. Human behavior recognition techniques now fall into two primary categories: deep learning techniques and conventional manual feature extraction techniques. The three stages of manual feature extraction techniques are typically the extraction of features, local descriptor calculation, and feature extraction. Grouping. Following a match between the edge information and the critical posture and location of the mark, Sullivan et al. track across successive frames based on the contour information. The entropy properties of the cylinder neighborhood around a given space-time point are calculated in Oikonomo et al.'s proposed detector. In order to achieve improved action detection and recognition, Patrons et al. established automatic motion data and fluid data weighting that modified the importance of human information on the basis of action participation for highlighting motion features that represent various positions in the video.

In contrast to the 2D convolution approach, the 3D convolution network method boasts a more straightforward and effective network architecture. Due of noise, it is difficult to directly utilize the behavior data that the smart phone sensors (gyroscope and triaxial acceleration) obtained for the har study. In order to extract reliable human behavior traits from sensor data, feature engineering is thus often utilized in a variety of har models.The research presents the construction of a human behavior recognition framework, dubbed human, which provides the RF with greater recognition. A framework for recognising human behaviour based on environmental perception is proposed in this paper, which integrates data on human behaviour with environmental information. A behavior recognition architecture based on environmental information helps to increase the model's recognition performance, as shown by experiments utilizing decision trees (DT), support vector machines (SVM), and k-nearest neighbors (k-NN). A report developed a cascade integration-based learning model for human behavior recognition based on the needs of several areas. Extreme randomized trees (ERT), RF, softmax regression, and extreme gradient boost trees (egbt) make up each layer of the model. The probability vectors for each of the four models' distinct categories are then created once the models have been trained using sensor data in the first layer. Subsequently, the original input information and the probability vector are combined to serve as the input for the subsequent level classifier. Ultimately, the predicted outcomes are acquired in accordance with the final level classifier. According to experimental data, this approach achieves higher recognition accuracy than previous techniques while also requiring less effort and a more efficient model-training procedure. Human behavior recognition study using an expanded volume neural network model requires manual feature marking; moreover, the model's capacity for generalization, feature acquisition, and computation amount need to be further enhanced. The soft migration technique is used to acquire and inherit the network's capacity to model video features for the densenet's basic module. This results in a new MDN network structure. The structures of many network models vary. Recall that the mdn-i3d combo is a semi-supervised "learner supervisor" setup. A approach for recognising human behaviour based on multiple scales convolution neural networks is presented in this study. Two enhanced channel attention modules, the space-time interaction matrix operation module and the deep recoverable convolution module, are suggested in conjunction with the study on human behavior recognition. Neural networks can simulate human behavior in a variety of network topologies by transferring and reusing this capacity. To lessen the discrepancies between the features derived from various network topologies, two efficient feature difference measurement methods are presented, taking into account the various properties of data features at various network levels. The suggested strategy has a decent categorization impact, according to experiments on a number of available datasets. The outcomes of the experiment demonstrate this method's high accuracy.

RELATED WORK

"Human behavior recognition based on bone spatial-temporal map"

The main challenge in recognizing human behavior is constructing a network for the extraction and categorization of spatiotemporal features. In order to address the issue that the current channel attention mechanism simply aggregates each channel's global average information while ignoring its local spatial data, this paper suggests two enhanced channel attention modules: the length separable convolution module and the time-space (ST) communication module of matrix operation. These modules are also combined with research on the recognition of human behavior. Proposing a multiple scales convolutional neural network technique for human behavior detection, it is combined with the excellent performance using convolutional neural network (CNN) for video and image processing. First, the behavior video is divided into segments. Next, low rank learning is applied to each segment to extract the corresponding low rank behavior information. Finally, these low rank behavior information are connected on the time axis to obtain the low rank behavior information of the entire video. This allows for the efficient extraction of behavior information from the video without the need for laborious extraction processes or assumptions. Neural networks can simulate human behavior in a variety of network topologies by transferring and reusing this capacity. To lessen the difference among features extracted from various network topologies, two efficient feature difference measurement methods are presented, taking into account the various properties of data features across different network levels. The suggested strategy has a decent categorization impact, according to experiments on a number of available datasets. The experimental findings demonstrate that the method's accuracy in identifying human behavior is excellent. It has been shown that the suggested model increases recognition accuracy while simultaneously enhancing the compactness for the model structure and successfully lowering the computational cost of the output weights.

"Overview of human behavior recognition methods based on bone data features"

Due to the significant fluctuations in wind power generation, there is little influence from timing energy use for renewable energy, and the rate of consumption is low. Consequently, a novel scheduling model for renewable energy consumption is developed, drawing upon quantitative feedback theory. To determine the curve of response of wind energy production, first calculate the response of the output of wind power generation. The power generation handle is divided to three stages: the beginning phase, the peak stage, and the end stage. The response output throughout the peak period is then calculated to obtain the move rate of output throughout the peak period. Create a mathematical model of the production of electricity from renewable sources, determine the wind energy system's output condition, and examine the renewable energy storage system's output features; Second, using the quantitative feedback theory, the goal weight coefficients of renewable energy consumption is established; the objective function of renewable power consumption is constructed; the power injected by parameter the nodes is determined; the quantitative input controller is designed using loop shaping technology; the control structure with two degrees of freedom is established; the objective function of reusable energy consumption is constructed; and the capacity of renewable energy consumption is limited by static constraints, such as voltage and current constraints; Finally, the quantitative opinions theoretical controller is designed; both the input and the output transfer functions for the consumption system are determined; and the renewable energy consumption objective weight coefficient is established. Using a range of index parameters in the green energy consumption ascertained by the QFT controller, the scheduling model for renewable energy consumption is solved by optimization using particle swarms. The results of the experiments demonstrate that the suggested approach may efficiently raise the rate at which renewable energy is used and maximize the impact of scheduling consumption.

"Design and implementation of rehabilitation evaluation system for the disabled based on behavior recognition"

Abstraction Early intervention programs must be immediately accessible to all Malaysian children with impairments. It will include speech therapy treatments for a large number of these kids. These treatments are now very rare and concentrated in secondary medical facilities and wealthy metropolitan institutions. Through a student-led program for community-based rehabilitation (CBR), this project aims to create, implement, and assess a novel method to clinical education and early intervention delivery for Malaysia. It was carried out using action research methods and a pragmatic approach. According to Kempis and McTaggart (1982), the action research approach consisted of two cycles: planning, carrying out action, observing, and reflecting. In order to comprehend the requirements of the CBR members better, an assessment of needs was additionally carried out in the first cycle. Action research made sure that the opinions of many stakeholders were taken into account. In order to completely inform the procedure and answer the research objectives, a mixed method approach was used, enabling the gathering of a wide variety of data. The qualitative techniques of semi-structured conversations, group interviews, reflective diaries, and participant observations were combined with the quantitative techniques of surveys, visual analog scales, and child evaluations to collect the data. The execution of these services via a role-emerging placement proved to be correlated with the learning for all stakeholders engaged in this study, which is the first on student-led services at the CBR in Malaysia. The emerging paradigm of CBR learning found that learning facilitation requires at least two elements. The first component was the growth of certain qualities and social skills that supported efficient learning. The interpersonal skills of reciprocity, collaboration, and trust as well as the ability to compromise, negotiate, and engage in profound intersubjective communication were among these qualities. The second component illustrated how time, confidence, and a feeling of control all play important roles in fostering personal growth and learning engagement. However, there was eventually mixed results from the introduction of student-led activities on the CBR program. This study showed, on the one hand, how moms may help children improve their communication skills, how CBR practitioners can aid speech pathologists in their job, and how mothers can open up new avenues for rehabilitation. However, negative elements like unfulfilled demands and a lack of more significant changes to CBR practice coexisted with good IV impacts. The unfavorable circumstances and general lack of support this CBR initiative encountered were major contributing factors in this. This study also supports previous research's conclusions on the critical need for Malaysia's children with disabilities to get appropriate rehabilitative care. In order for local academics and government development officials to comprehend these results, the study pledges to make them public. For students, teachers, CBR professionals, children with disabilities, and families, the introduction of student-led programs in an urban CBR program has provided a ground-breaking educational experience. It is envisaged that these discoveries and initiatives would influence attempts to provide locally relevant and accessible speech pathology services in Malaysia.

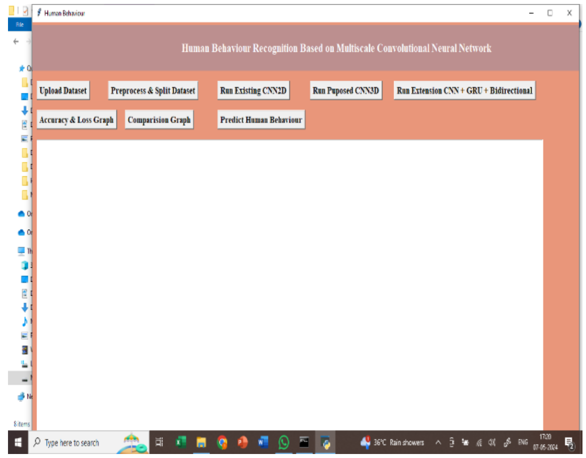
"Video based pedestrian detection and behavior recognition"

The main challenge in recognizing human behavior is constructing a network for the extraction and categorization of spatiotemporal features. In order to address the issue that the current channel attention mechanism simply aggregates each channel's global average information while ignoring its local spatial data, this work suggests two enhanced channel attention modules: the depth separable convolutions module and the time-space (ST) interaction module for matrices operation. These modules are also combined with research on the recognition of human behavior. Proposing a multiple scales convolutional neural network technique for human behavior detection, it is combined with the excellent performance using convolutional neural networks (CNN) for image and video processing. First, the behavior video is divided into segments. Next, low rank learning is applied to each segment to extract the associated low rank behavior information. Finally, these low rank actions information are linked together on the time direction to get the low rank behavior knowledge for the entire video. This allows for the efficient extraction of behavior information from the video without the need for laborious extraction processes or assumptions. Neural networks can simulate human behavior in a variety of network topologies by transferring and reusing this capacity. To lessen the difference between characteristics extracted from various network topologies, two efficient feature difference measurement methods are presented, taking into account the various properties of data features across different network levels. The suggested strategy has a decent categorization impact, according to experiments on a number of available datasets. The experimental findings demonstrate that the method's accuracy in identifying human behavior is excellent. It has been shown that the suggested model increases recognition accuracy while simultaneously enhancing the small size of the model's framework and successfully lowering the computational cost of the output weights.

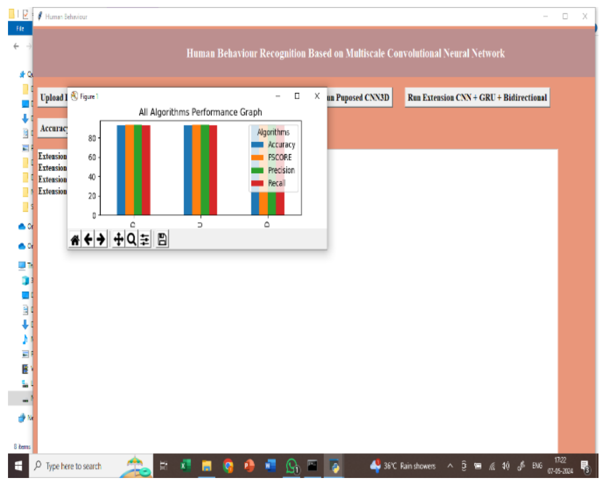
METHODOLOGY

1. Upload Dataset: In order to carry out this project, we must upload our dataset.
2. Pre-process & Split Dataset: Here, we must pre-process our dataset and divide it into two groups: 80% for education and 20% for testing.
3. Run Existing CNN2D: We must now execute our CNN2D model.
4. Run proposed CNN3D: We must now execute our CNN3D model.
5. Run Extension CNN + GRU + Bidirectional: Here We have to run our 3 Algorithms.
6. Graph: All of our algorithms' graphs are shown here.
7. Predict Human BehaviourHere, we forecasted the behavior of humans.

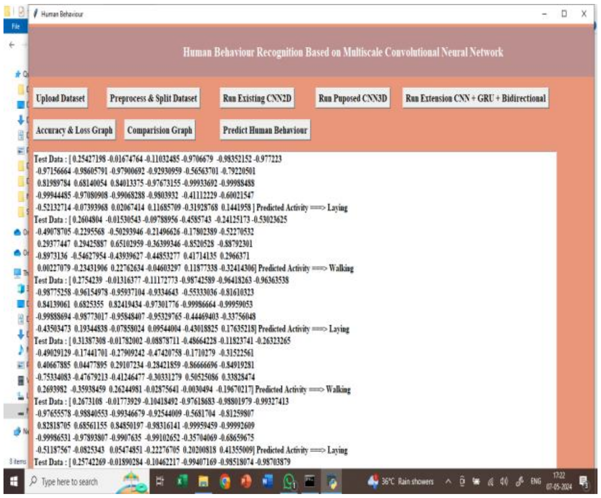
RESULT AND DISCUSSION



In the above result we got Output Window. Now Click On the "Upload Dataset" button to upload the dataset to the application.



In above result comparison graph x-axis represents algorithm names and y-axis represents accuracy and other metrics in different colour bars and in all algorithms extension got high performance.



In above result loaded test data and then predicting using extension model and in output in square bracket we can see Test Data Values and after arrow symbol can see predicted activity as Standing or any other activity.

CONCLUSION

This research proposes a novel attention mechanism-based approach for the identification of human behavior. A better attention module is suggested by examining the drawbacks of the current channel attention method. Experiments on visualization outcomes, network accuracy improvement, extra network parameters, and other topics are conducted to confirm the efficacy of the enhanced attention module. Utilizing a multi-scale convolution kernel, the behavior traits under various receptive fields are obtained. The feasibility of cross-structure learning is confirmed by the convolution layer, the pool layer, and full connection layer, which are all rationally designed in order to refine the characteristics. Comparing the supervision at various phases demonstrates the need for a multi-stage progressive supervision technique; the impact of model structure on the outcome of soft migration is also covered. It is discovered that when the monitoring network's structure resembles that of the learning network, the network will converge more easily. Future research may use more sensors to enhance the data dimension and raise the accuracy of identification even further. The model module for our approach has several parameters, and future research will concentrate on how to improve the lightweight of the model.

REFERENCES

1. 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.

2. 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.

3. 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.

4. C. Y. Zhang, H. Zhang, W. He, F. Zhao, W. Q. Li, T. Y. Xu, et al., "Video based pedestrian detection and behavior recognition", *China Sci. Technol. Inf.*, vol. 11, no. 6, pp. 132-135, 2022.

5. 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.

6. S. Huang, "Progress and application prospect of video behavior recognition", *High Tech Ind.*, vol. 27, no. 12, pp. 38-41, 2021.

7. 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.

8. 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.

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. 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.

11. Z. Duan, Q. Ding, J. Wang and W. Li, "Subway station lighting control method based on passenger behavior recognition", *J. Railway Sci. Eng.*, vol. 18, no. 12, pp. 3138-3145, 2021.

12. D. Liu, J. Yang and Q. Tang, "Research on identification technology of violations in key underground places based on video analysis", *Proc. Excellent Papers Annu. Meeting Chongqing Mining Soc.*, pp. 71-75, 2021.

13. Y. Ye, "Key technology of human behavior recognition in intelligent device forensics based on deep learning", Apr. 2021.

14. Y. Li, "Mining the spatiotemporal distribution law of CNG gas dispensing sub station and identifying abnormal behaviors based on machine learning", Apr. 2021.

15. W. Wang, "Research on behavior recognition based on video image and virtual reality interaction application", Mar. 2021.

16. J. Wang, "Design and implementation of enterprise e-mail security analysis platform based on user behavior identification", *J. Shanghai Inst. Shipping Transp. Sci.*, vol. 43, no. 4, pp. 59-64, 2020.

17. K. Han and Z. Huang, "A fall behavior recognition method based on the dynamic characteristics of human posture", *J. Hunan Univ. Natural Sci. Ed.*, vol. 47, no. 12, pp. 69-76, 2020.

18. F. Wang, "Research on attitude estimation and behavior recognition based on deep learning in logistics warehousing", 2020.

19. Y. Ying, "Analysis of prenatal behavior characteristics of Hu sheep and development of monitoring system based on embedded system", 2020.

20. J. Bao and H. Jin, "A semi supervised learning method for identifying intrusion behaviors in ship LAN", *Ship Sci. Technol.*, vol. 42, no. 24, pp. 136-138, 2020.

21. L. Zhang, Y. Zhang, M. Li, X. Shi, B. Zhai and W. Wang, "Identification method of downhole personnel behavior based on CSI", *J. Internet Things*, vol. 4, no. 4, pp. 26-31, 2020.

22. Y. Zhang and K. Jia, "Case study on identification and prevention of financial fraud", *Bus. Accounting*, vol. 17, no. 24, pp. 101-104, 2020.

23. Y. Li and L. Xie, "A behavior recognition algorithm combining RGB-D video and convolutional neural network", *Comput. Digit. Eng.*, vol. 48, no. 12, pp. 3052-3058, 2020.

24. H. Fang and Q. Lu, "Research on student classroom activity detection method based on behavior recognition", *Inf. Syst. Eng.*, vol. 25, no. 12, pp. 27-29, 2020.

25. X. Cao, "Design and implementation of traffic violation recognition algorithm based on vehicle video", 2019.

26. Z. Xu, "Research on human behavior recognition technology based on depth feature fusion and its application in video surveillance", 2019.

27. Y. Yang, "Research on key algorithms and platform development and application of human motion behavior recognition in unrestricted scenes Xinhua college", Mar. 2020.

28. X. Huang, "Human behavior recognition method based on point projection features of bone joints", *Mod. Comput.*, vol. 12, no. 36, pp. 3-7, 2019.

29. J. Chen, X. Xie, J. Li and G. Shi, "Behavior recognition method based on spatiotemporal attention mechanism", *China Stereol. Image Anal.*, vol. 24, no. 4, pp. 325-333, 2019.

30. X. Han and T. Wu, "Human behavior recognition algorithm based on deep learning", *Pract. Understand. Math.*, vol. 49, no. 24, pp. 133-139, 2019.