HUMAN BEHAVIOR RECOGNIZATION BASED ON MULTISCALE CONVOLUTIONAL NEURAL NETWORK

GANISETTY SANTHOSHI LAKSHMI
P.G Scholar
gsanthoshilakshmi1@gmail.com
Bonam Venkata Chalamayya Engineering College(A)
Odalarevu.
Shiek meerasharif Professor of CSE
Sharief.mtech@gmail.com
Bonam Venkata Chalamayya Engineering College(A)
Odalarevu.

ABSTRACT

This project addresses the challenge of recognizing human behavior through a novel approach. It introduces two enhanced channel attention modules, the space-time interaction module and depth separable convolution module, to improve the recognition process. Leveraging convolutional neural networks (CNNs) known for their prowess in image and video processing, a multi-scale CNN method is proposed. The process involves segmenting behavior videos, performing low-rank learning on segments to extract behavior information, and then integrating these on the time axis for comprehensive understanding. This approach effectively captures behavior information, simplifying the extraction process. The model's versatility allows it to adapt to different network structures, improving recognition accuracy while reducing computational complexity. Experimental results demonstrate its effectiveness in human behavior recognition.

Keywords: Human behavior recognition ,Channel attention modules,Space-time interaction module,Depth separable convolution module,Convolutional Neural Networks (CNNs),Multi-scale CNN,Behavior video segmentation,Low-rank learning,Recognition accuracy

1 INTRODUCTION

In the field of computer vision, the research on human behavior recognition can not only develop the relevant theoretical basis, but also expand its engineering application. For the theoretical basis, the field of behavior recognition integrates the knowledge of many disciplines, such as image processing, computer vision, artificial intelligence, human kinematics and bioscience. Human behavior recognition is an important method to process video content using computer vision technology. It is an important research direction [1].

According to the different forms of convolution kernel, behavior recognition methods based on deep learning can be divided into two categories: 2D convolution network and 3D convolution network, many researchers have applied deep learning to motion recognition. They have tried to use various methods to realize the behavior recognition technology based on computer vision, and achieved good results. 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 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.

2. LITERATURE SURVEY

- **Identification method and experiment of unsafe behaviors of subway passengers based on Kinect**
- **Y. Lu, L. Fan, L. Guo, L. Qiu, and Y. Lu**

In order to solve the problem of the low action recognition accuracy of passengers' unsafe behaviors caused by redundant joints, this study proposes an efficient recognition method based on a Kinect sensor. The method uses the pelvis as the starting point of the vector and high-frequency bone joints as the end point to construct the recognition feature vector. The joint angle difference between actions is obtained by using the cosine law, and the initial test result is converted into action similarity combined with the DTW similarity algorithm

- **Research on human behavior recognition technology based on depth feature fusion and its application in video surveillance**
- **Z. Xu**

In order to solve the problems of low recognition accuracy and high computational complexity caused by redundant video data in the existing behavior recognition process, a human behavior recognition method based on video key frame (S3DCCA) is proposed. First of all, structural similarity (SSIM) algorithm is used to calculate the difference of luminance, contrast and structure between the two frames, and the result is multiplied to attain SSIM value, then select the local and global key frame in the human motion video frame sequence according to the SSIM value. Finally, the selected key frame are used as the input of three-dimensional convolutional neural networks and attention mechanism Channel attention (3DCCA) model to recognize human behavior. Experimental results on UCF101 and HMDB51 datasets show that the proposed method has high recognition rate.

- **Human behavior recognition method based on point projection features of bone joints**
- **J. Chen, X. Xie, J. Li, and G. Shi**

With the advent of the era of intelligent information,human behavior recognition from video has received more and more attention. Most of the existing action recognition methods consider improving data form or network structure. They can achieve good performance. Inspired by the behavior of human observation that human will pay attention to key areas and critical moments,we propose a new action recognition method based on spatio-temporal attention mechanism.

- **Human behavior recognition algorithm based on deep learning**
- **X. Han and T. Wu**

With the continuous advancement of technology, human behavior recognition, as an important scientific research in the field of computer vision, has important research in many fields such as intelligent surveillance, smart home, virtual reality. In the current complex environment, traditional manual methods have been difficult to meet the requirements of high recognition accuracy and applicability. The introduction of deep learning has brought new development directions for behavior recognition. This article mainly summarizes behavior recognition algorithms based on deep learning.

- **Real time human behavior recognition on Android platform**
- **X. Jia, H. Wanger, and J. Wu**

The number of applications based on the Android platform is increasing rapidly now. However, as the supervision and review of Android applications are inadequate, a reasonable chance exists that users will download malware. This malware can lead to information leakage, monetary loss, and other damages. At present, a variety of applications exist for detecting malware, but most of these applications cannot show specific malicious behaviors. Moreover, the operation of this detection software is based on the database of viruses, and thus, it cannot identify unknown malware. To solve these problems, we implemented a system to detect the behaviors of Android applications and identify known or unknown malware. Our system can monitor specified applications utilizing loading a kernel module.

BACKGROUND OF THE PROBLEM

In propose work author applying 3DCNN algorithm for human behaviour prediction as all existing algorithms were directly employing global average information of each channel (taking all channels of images as single data) which ignores spatial and depth information from image features which leads to inaccurate recognition. If model has accurate information or each shape from the image then it can predict accurately.

Disadvantages of Existing system

1. Existing algorithms lack spatial recognition.
2. Spatial depth information exclusion.
3. Limited accuracy.
4. Inaccurate prediction from image features.

3. METHODOLOGY

- **Loading Required Packages and Classes:**

The project begins by loading essential Python packages and classes necessary for data processing, visualization, and model implementation. These packages include libraries like NumPy, Matplotlib, and TensorFlow.

- **Loading and Displaying HAR Dataset Values:**

The Human Activity Recognition (HAR) dataset, captured from smartphones, is loaded and displayed. This dataset contains sensor data capturing various activities. Its exploration provides insights into the nature and distribution of activities recorded.

- **Plotting Activity Distribution Graph:**

Activities in the HAR dataset are identified and plotted on a graph. The x-axis represents activity names, while the y-axis displays the count of each activity. This graphical representation offers a visual understanding of the dataset's composition, highlighting the distribution of different activities.

- **Dataset Processing and Train-Test Split:**

The dataset undergoes preprocessing, including normalization and feature extraction. Subsequently, it's divided into training and testing sets to facilitate model evaluation. The total number of records allocated for training and testing is displayed, ensuring a balanced partition for model training and evaluation.

- **Definition of Evaluation Metrics**:

Functions are defined to calculate accuracy and other evaluation metrics, such as precision, recall, and F1 score. These metrics provide comprehensive insights into model performance, enabling effective comparison and evaluation across different models.

- **Training Existing CNN2D Algorithm**
- The existing CNN2D algorithm is trained on the HAR dataset. The model complexity and parameter size are discussed, emphasizing the trade-offs between model complexity, training time, and accuracy. Model complexity influences training efficiency and computational requirements, necessitating careful consideration during model selection and optimization.
- **Evaluation of Existing CNN2D Model:**

The trained CNN2D model is evaluated on the test dataset, yielding accuracy and other performance metrics. The confusion matrix illustrates the model's predictive capabilities, highlighting correct and incorrect predictions for different activity labels.

- **Training Proposed MCNN (MDN) Model:**

A proposed MCNN (MDN) model, utilizing CNN3D architecture, is trained on the HAR dataset. This model aims to improve upon the existing CNN2D algorithm by reducing parameter size while maintaining or enhancing accuracy. The model's architecture and parameter requirements are discussed, emphasizing its efficiency and performance compared to existing models.

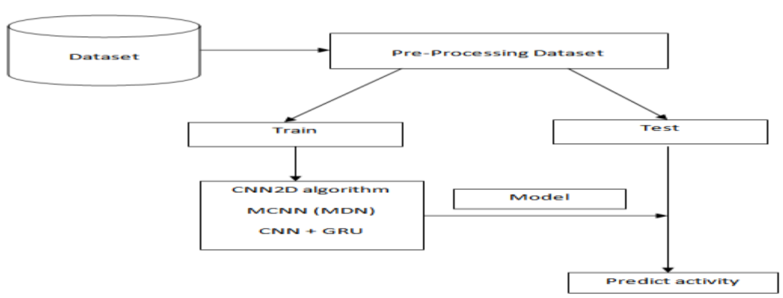
- **Evaluation of Proposed MCNN (MDN) Model:**

The proposed MCNN (MDN) model is evaluated for accuracy and other metrics, showcasing its performance compared to existing models. The model's efficiency in terms of parameter size and computational complexity is emphasized, highlighting its potential for real-world deployment and scalability

Extension:

To further enhance accuracy we have combined 3 algorithms together called CNN + GRU + Bidirectional with less number of training parameters which help in further reducing model complexity with 1000 parameters and its accuracy is high compare to propose and existing algorithms. Extension hybrid optimizing training features with 3 different CNN + GRU + Bidirectional which helps in obtaining more optimized features which in turn give better accuracy.

4.SYSTEM ARCHITECTURE:



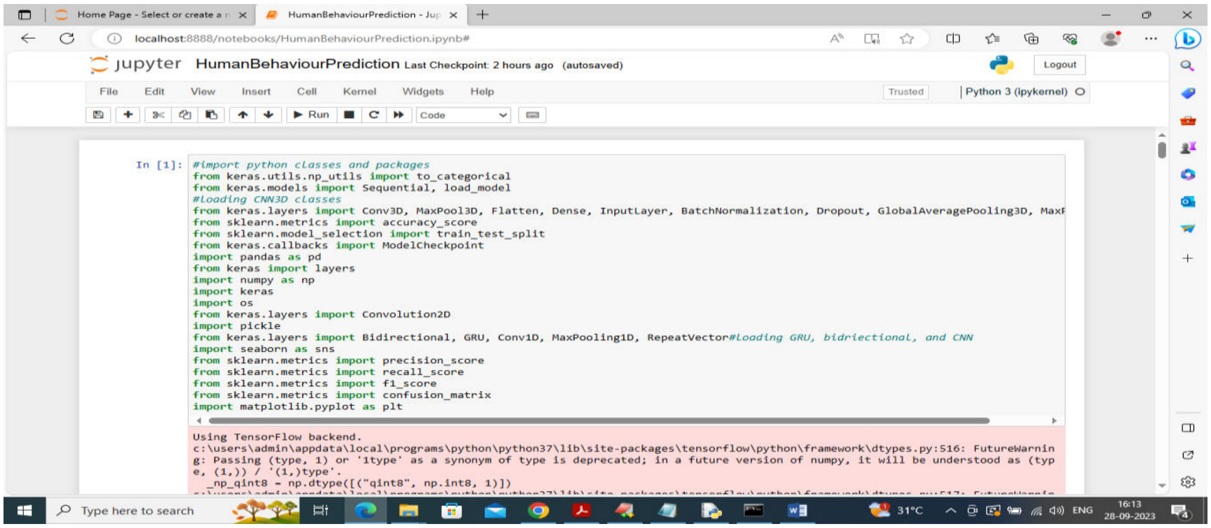
PROPOSED SYSTEM

So in propose work author employed two different module such as space-time (ST) interaction module of matrix operation and the depth separable convolution module, combined with the research of human behaviour 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 behaviour recognition is proposed. Combination of spatial and depth separable module is known as Multi scale Convolution Neural Network (MCNN or MDN). Propose model is experimented on UCI HAR dataset which captured human activity using Smart Phone. Propose model giving best accuracy compare to existing CNN2D or LSTM.

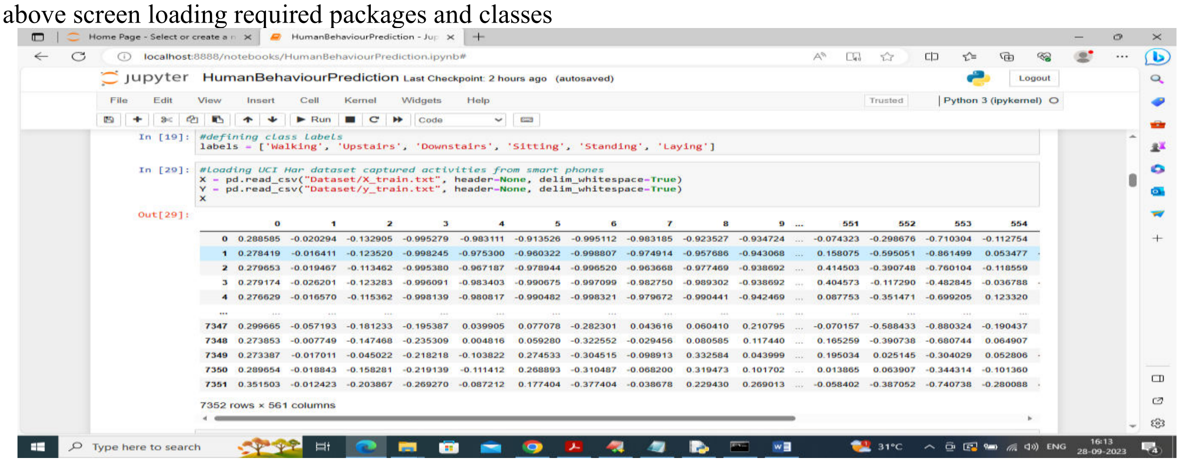
- **Advantages of Proposed System**

1. Enhanced channel attention mechanisms.
2. Improved spatial information utilization.
3. Effective segmentation.
4. Reduced computational complexity.
5. Enhanced recognition accuracy.

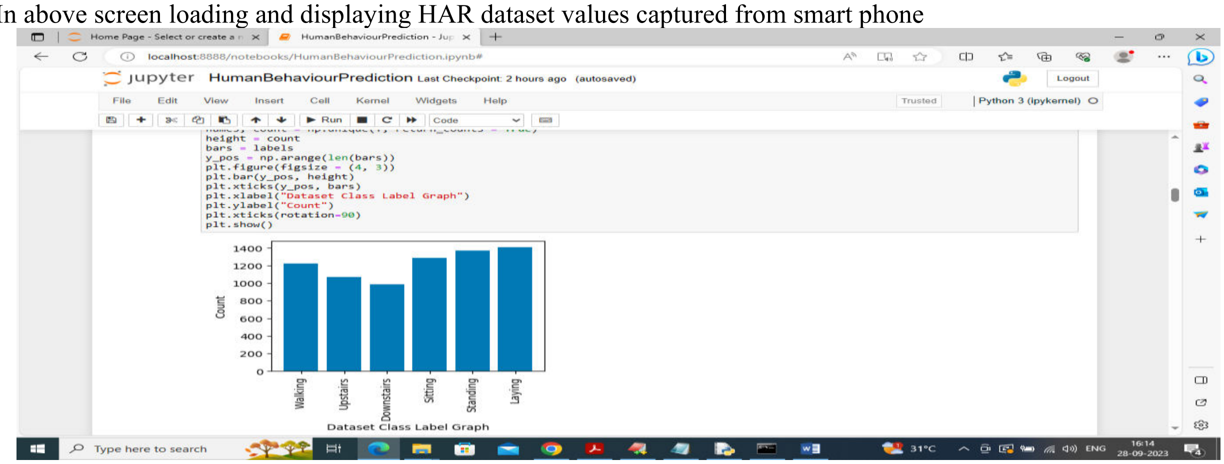
OUTCOMES OF THE RESULT



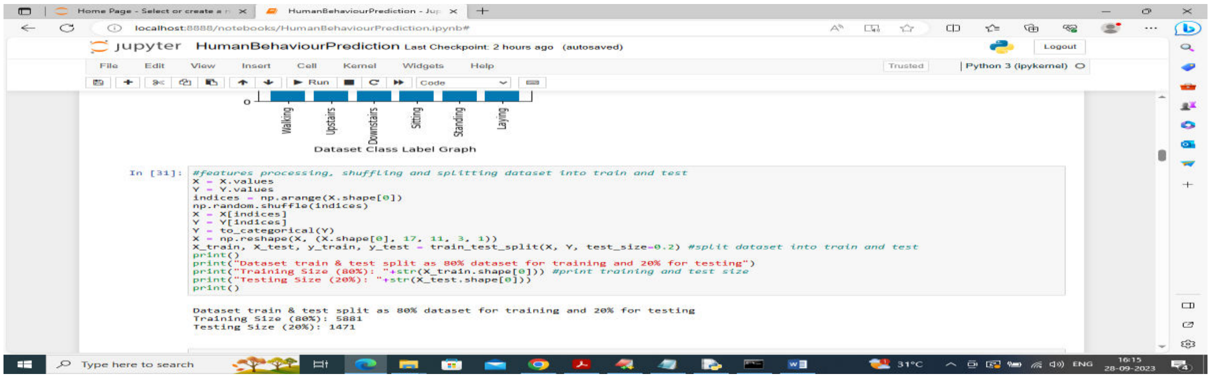
In above screen loading required packages and classes



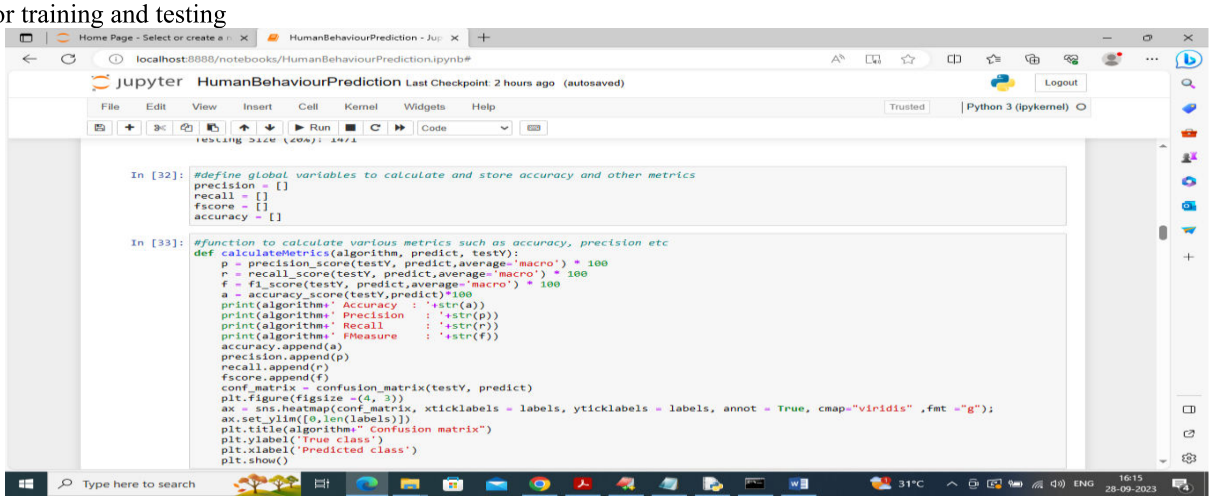
In above screen loading and displaying HAR dataset values captured from smart phone



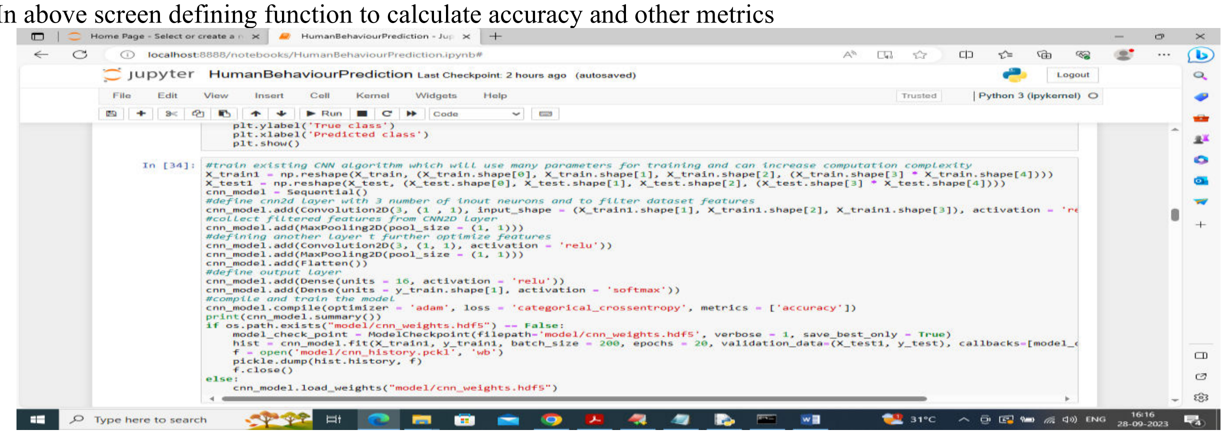
In above screen finding and plotting graph of different activities found in dataset where x-axis represents ACTIVITY NAMES and y-axis represents count of those activities



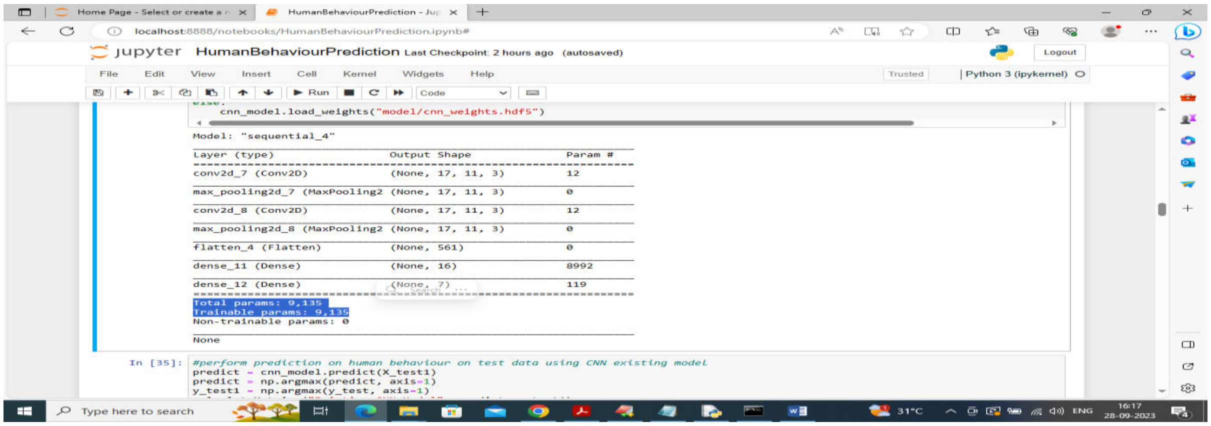
In above screen processing dataset and then splitting dataset into train and test and then can see total records used for training and testing



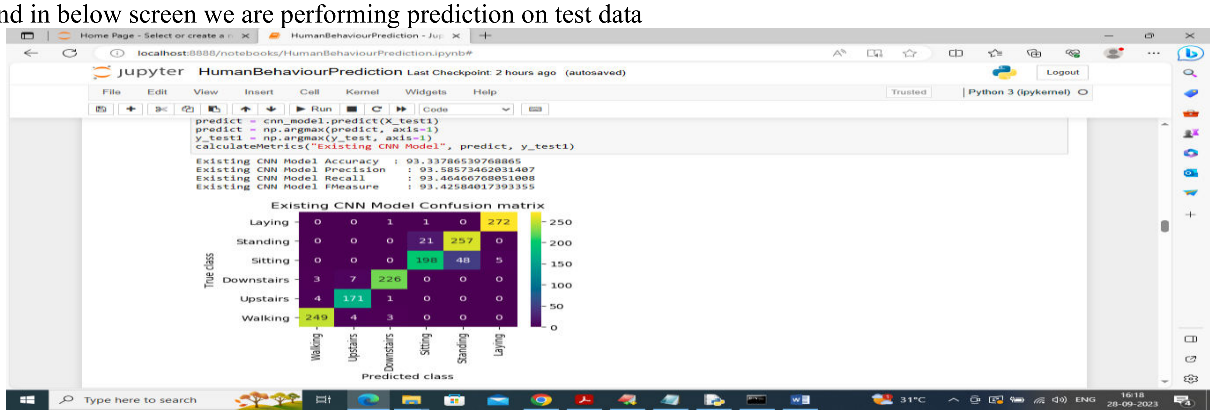
In above screen defining function to calculate accuracy and other metrics



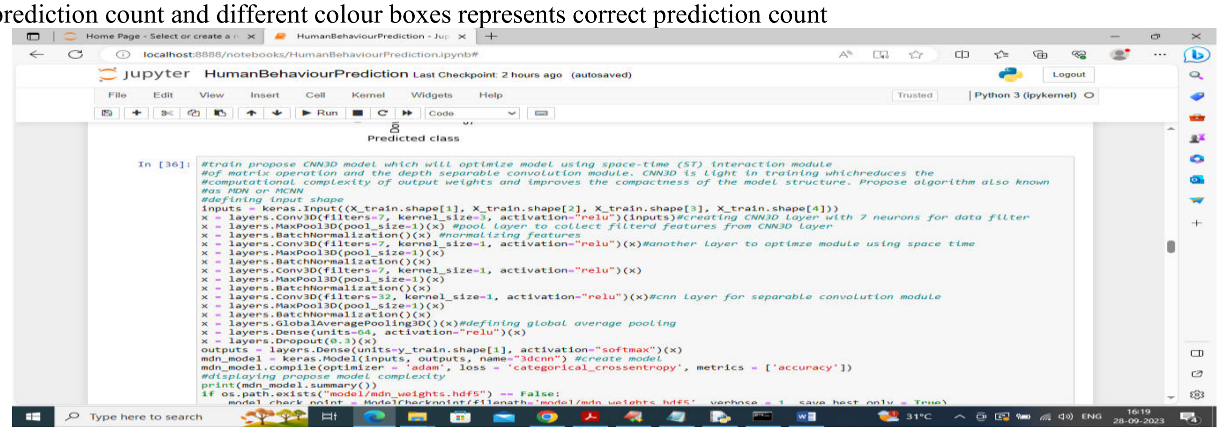
In above screen training existing CNN2D algorithm and can read blue colour comments to know about module and after executing above block will get below complexity of the model



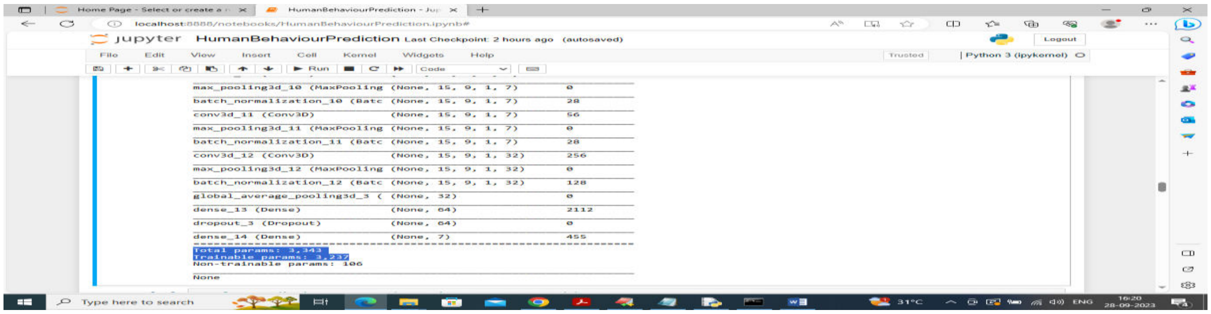
In above screen blue colour text we can see existing CNN2D required 9135 parameters to train a model and increasing this parameters size can increase model complexity and training time and by decreasing we can reduce and in below screen we are performing prediction on test data



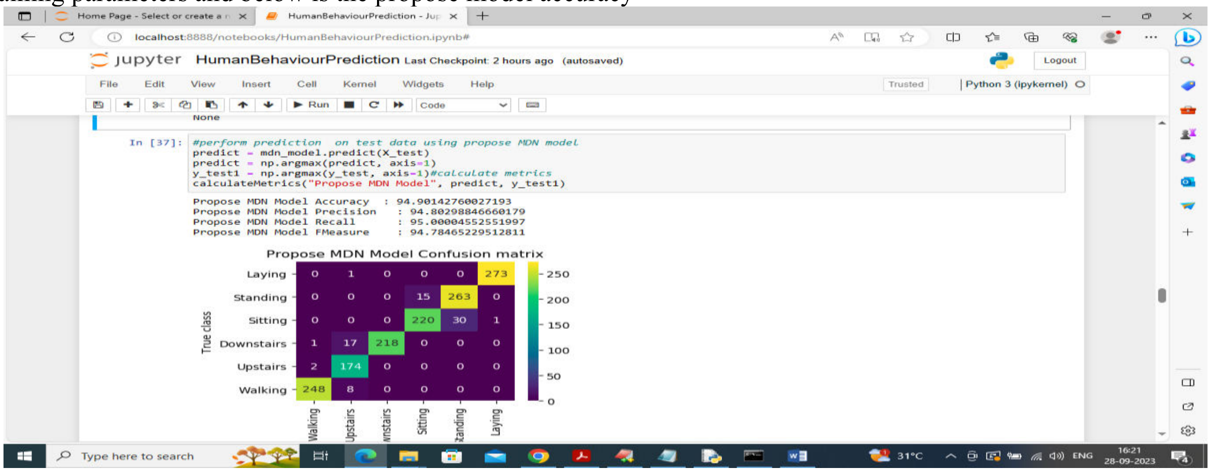
In above screen existing CNN2D model got 93% accuracy and can see other metrics and in confusion matrix x-axis represents predicted Labels and y-axis represents True Labels and all blue colour boxes represents incorrect prediction count and different colour boxes represents correct prediction count



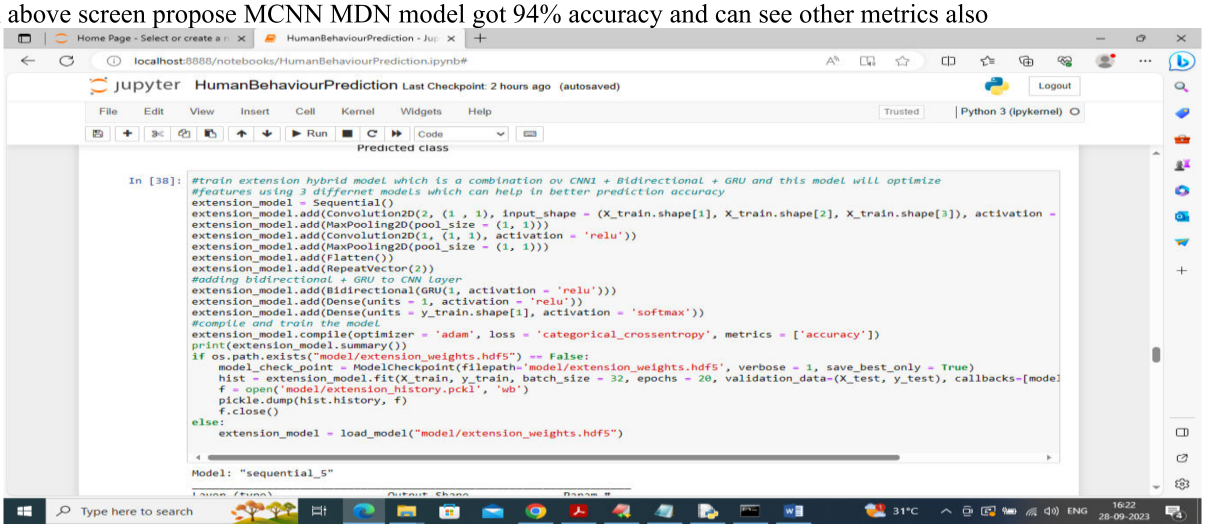
In above screen training propose MCNN (MDN) model using CNN3D architecture and after executing above model will get below output



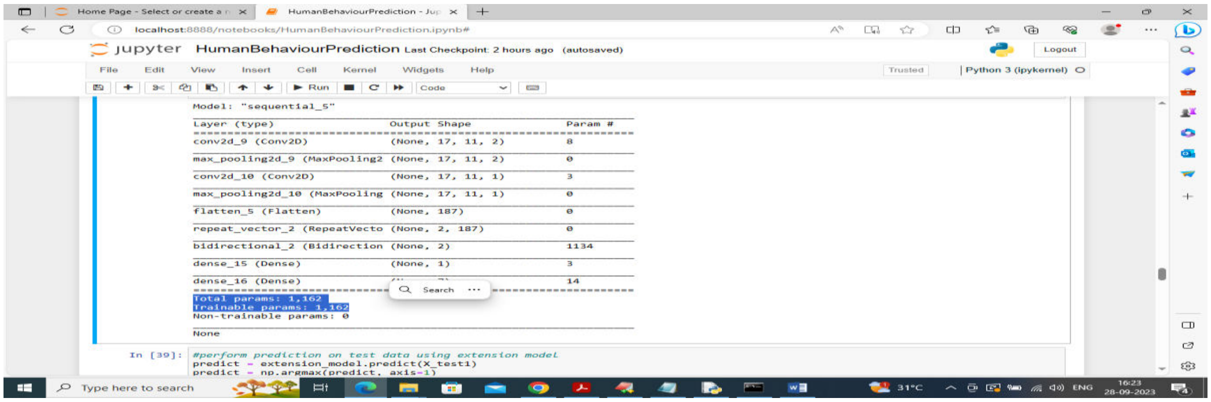
In above screen propose model required 3306 parameters for training which are lesser than existing CNN2D 9000 training parameters and below is the propose model accuracy



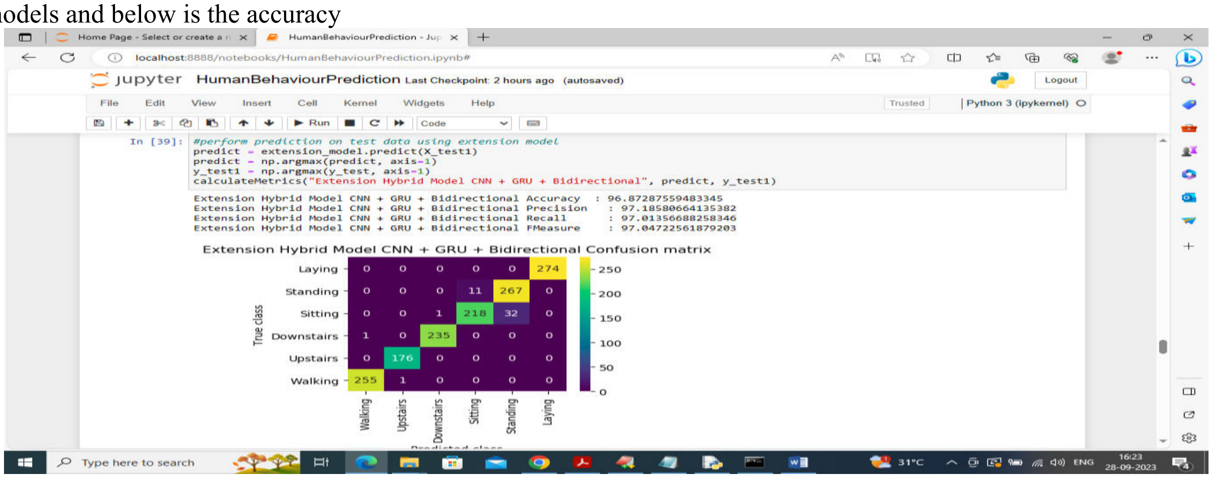
In above screen propose MCNN MDN model got 94% accuracy and can see other metrics also



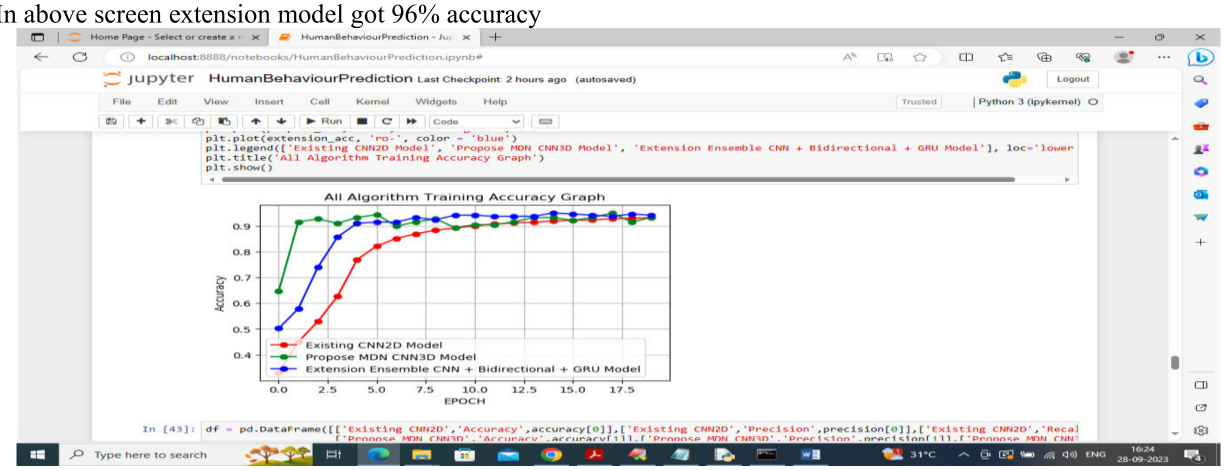
In above screen training extension model by combining 3 different algorithms such as CNN + GRU + Bidirectional and after executing above block will get below output



In above screen extension model required 1162 parameters for training which are lesser than propose and existing models and below is the accuracy

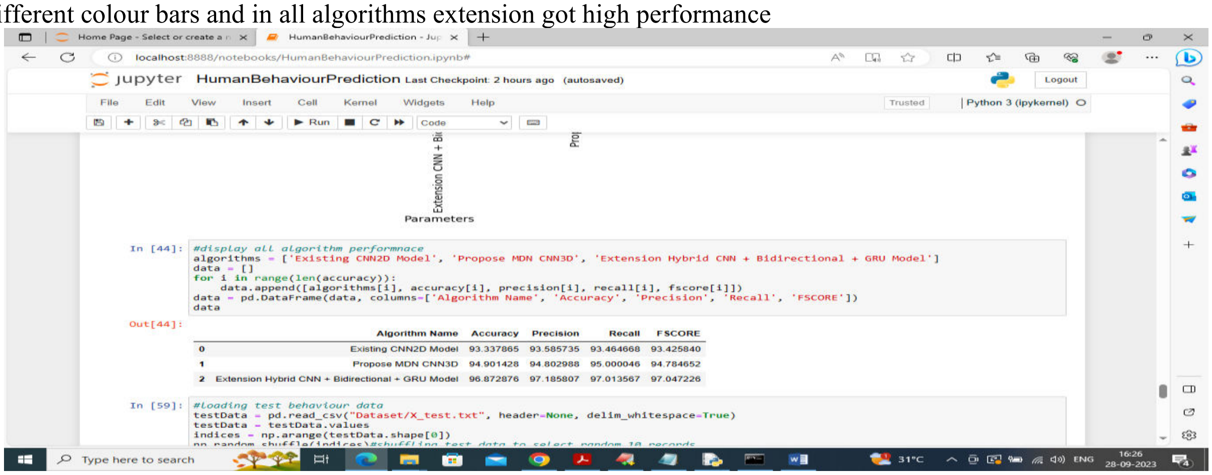


In above screen extension model got 96% accuracy

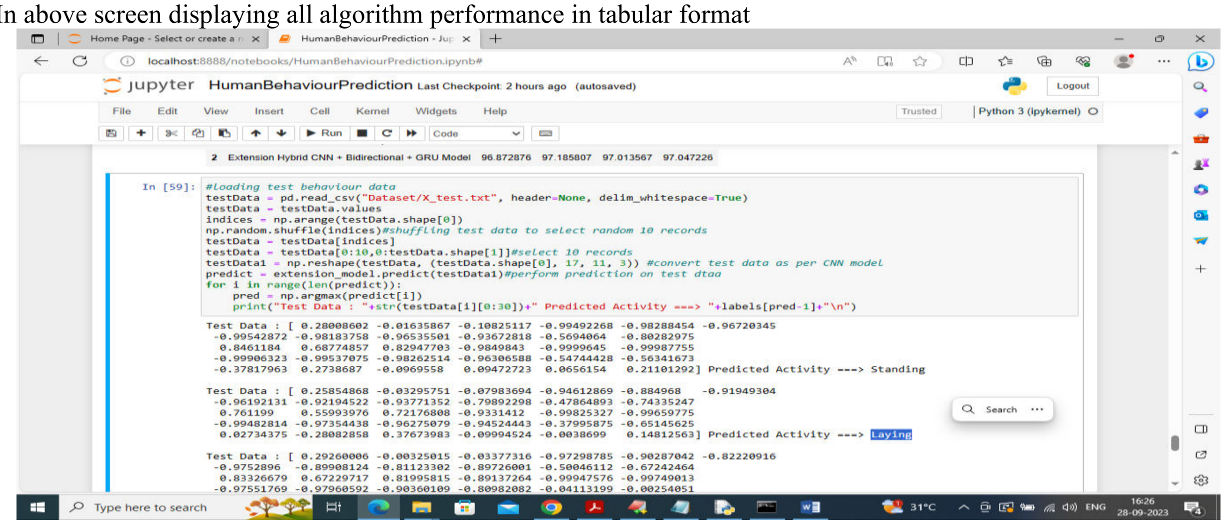


In above screen displaying training accuracy of all 3 models such as existing , propose and extension where x-axis represents training epoch and y-axis represents accuracy and with each increasing epoch accuracy got increase and in in all models extension got high accuracy

In above 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 screen displaying all algorithm performance in tabular format



In above screen loading 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.

Note: sometime propose model accuracy may be less so rerun all modules to avoid that error

5. CONCLUSION

This project showcases a comprehensive exploration and implementation of various neural network architectures for Human Activity Recognition (HAR). Through rigorous training, evaluation, and comparison, the project demonstrates the effectiveness of different models. The CNN2D, proposed MCNN (MDN), and extension models exhibit distinct parameter sizes, complexities, and accuracies. Notably, the extension model achieves the highest accuracy, showcasing its robustness and efficiency. The graphical and tabular representations offer insights into model performance and facilitate informed decision-making. Overall, the project underscores the significance of neural network architectures in HAR tasks, promising enhanced accuracy and reliability in activity recognition systems.

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, doi: 10.16208/j.issn1000-7024.2022.04.036.

[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, and Q. Ye, ''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,'' in Proc. Excellent Papers Annu. Meeting Chongqing Mining Soc., 2021, pp. 71–75.

[13] Y. Ye, ''Key technology of human behavior recognition in intelligent device forensics based on deep learning,'' Sichuan Univ. Electron. Sci. Technol., Chengdu, China, Tech. Rep., Apr. 2021.

[14] Y. Li, ''Mining the spatiotemporal distribution law of CNG gas dispensing sub station and identifying abnormal behaviors based on machine learning,'' Tianjin Agricult. Univ., Tianjin, China, Tech. Rep., Apr. 2021.

[15] W. Wang, ''Research on behavior recognition based on video image and virtual reality interaction application,'' Sichuan Univ. Electron. Sci. Technol., Chengdu, China, Tech. Rep., 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,'' Anhui Univ. Technol., Ma'anshan, China, Tech. Rep. 14, 2020.

[19] Y. Ying, ''Analysis of prenatal behavior characteristics of Hu sheep and development of monitoring system based on embedded system,'' Zhejiang Agricult. Forestry Univ., Hangzhou, China, Tech. Rep. 10, 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.