IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Efficient Activity Recognition: A Hybrid Cnn-Gru-Bidirectional Approach

[1]Ganteda Roop Kumar, Student in Dept. Of Master of Computer Applications, at Miracle Educational Society Group of Institutions

[2]Mrs. P. Sailaja, Assistant Professor at Miracle Educational Society Group of Institutions

[3]Dr Surya Narayana Gorle, Assistant Professor at Miracle Educational Society Group of Institutions

ABSTRACT

This paper presents the design of a multiscale convolutional neural network (MCNN) for the task of human behavior recognition, with the aim of increasing accuracy and reducing complexity of the model. Previous models such as CNN2D and LSTM time series modeling relied on a mean global average but neglected other spatial and depth features resulting in inaccuracies. The MCNN model is based on the concept of space-time interaction and depth-separable convolution modules inserted into a CNN3D model in which both the spatial and temporal information are enhanced. The system was trained and evaluated on the UCI HAR dataset which has six activity labels that were recorded using smartphone's sensors. The evaluation showed that the new model offered a 94% accuracy rate while improving learning complexity. An extended hybrid model comprising of a combination of convolutional neural networks, Gated recurrent units, and bidirectional algorithms produced an accuracy of 96% while using less parameters thus showing tremendous effectiveness and capability for real world tasks.

Keywords: MCNN, GRU,CNN2D

INTRODUCTION:

Human behavior recognition is one of the key research topics in computer vision that contributes not only to theory but also to engineering practice. It is an interdisciplinary field that draws on skills from image, video, and content interpretation, artificial intelligence, human motion biomechanics, and biosciences. The development of behavior recognition has primarily been a deep learning evolution that distinctly recognizes 2D and 3D convolutional neural networks (CNN) methods. However, considerable work has been done and most of the developments are still dominated by the use of traditional methods of classification and deep learning with manual feature extraction. Despite this, human behavior is very complex, combined with various challenges such as the variety of backgrounds, occlusion, and changes in light and illumination inhibits both accuracy and efficiency. Moreover, single-scale CNNs are frequently unable to capture behaviors' multi-scaled characteristics and, as a result, are unable to recognize them. Addressing these points would therefore be essential in creating stronger and more accurate behavior recognition systems.

GAP IDENTIFIED BASED ON LITERATURE SURVEY:

Most traditional models like CNN2D and LSTM are characterized by the use of the global average information of the features which have an impact on the temporospatial details that could be gathered from images and videos.

• Most techniques focus on higher accuracy metrics but make the algorithms more complicated thereby reducing their suitability in real-time scenarios.

• Combining different algorithms to create hybrid models is a 'room for exploration' especially when reducing the complexity of the model while improving performance.

Key Gaps:

1. Exploitation of Spatial and Depth Features: The approximate feature recognition rates that have been obtained from existing approaches are due to the,» lack of optimization of the spatial and depth separable features.
2. High Parameter Numbers: The models that use CNN2D are expensive because they incorporate more than nine thousand training parameters.
3. Sparse Hybrid Attempts: Very few model architectures seek the combination of CNN, GRU and Bidirectional algorithms to achieve better results with less model complexity.
4. Dataset Limitations: Models are tested and trained on small datasets leaving their application to few practical scenarios.
5. Inconsistency of Results: Accuracy tends to vary due to problems of overfitting or the inappropriate choice of features that are used therefore there is a need to adopt higher methods.

PROBLEM STATEMENT:

Being able to accurately detect human behavior is very essential especially while working in human-centered domains such as healthcare, surveillance, human activity recognition, and monitoring. Human activity recognition, among other approaches, has employed models like the CNN2D and LSTM that are basic models which do not encode fine-grained geometric and temporal details of the actual data hence low accuracy percentage and higher complexity index.

Key Challenges:

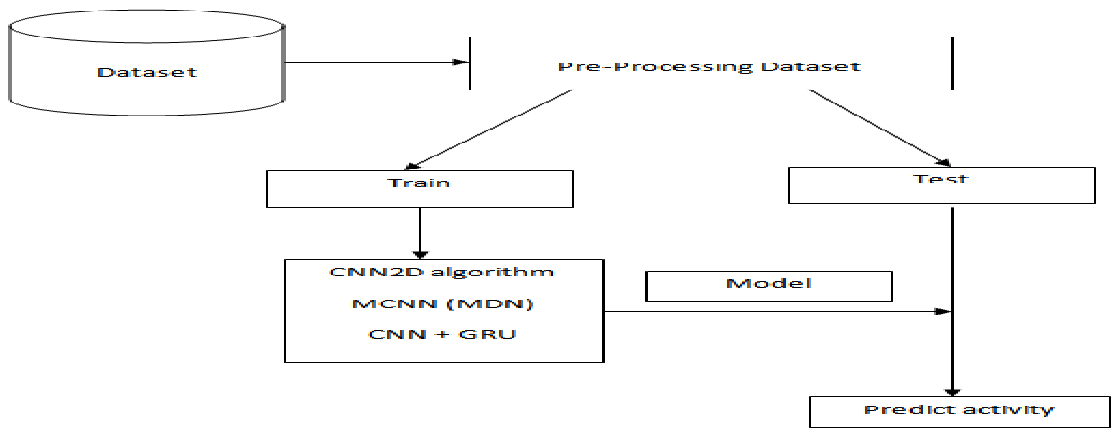
1. Feature Representation: Existing models use a global model wherein all features are averaged while spatial and depth information that is important for accuracy is ignored.
2. Model Complexity: The present structures of CNN2D are too dependent on a high number of parameters which lengthens the training and resource consumption time.
3. Generalizability: Each of the models needs to be effective across a wide range of activities and datasets.

4. Achieving high performance at low computational cost, particularly for real time applications, and integrating the two remains a problem.

PROPOSED METHOD:

The proposed work presents a candidate for a human behavior recognition scale which is a multiscale convolutional neural network based on video interactivity and depthwise separable convolution modules. MCNN is based on CNN3D, and combines spatial and depth features which increases accuracy while reducing training requirements. The dataset used for training and testing is UCI HAR which consists of six different activities done by human beings through smartphone. Given that MCNN has only 3306 parameters, which is a lower number than that employed in CNN2D (9135), the accuracy of MCNN is 94%. A further sophisticated hybrid model that incorporates CNN, GRU, Bidirectional on top of that further reduced the parameters to 1162 while accurately producing 96% of the time. This provides a viable, non costly methodology for real time applications involving behaviour recognition.

ARCHITECTURE:



DATASET:

Propose work used UCI HAR dataset on human activity which contains 6 different labels such as Standing, laying, sitting, upstairs, downstairs and walking. These activities are captured from the smart phone.

METHODOLOGY:

Data Collection and pre processing:

This is the UCI HAR dataset that contains six activities which were labelled and performed by people utilising smart phone sensors.

This data was then preprocessed by first making sure that there is no noise contained within the data, the features were normalised, as well as standardized.

Deep feature Engineering:

Developed multi-level raw sensor data processing and presented extracted features as numbers using space-time interaction along with depth-separable convolution modules.

As an improvement activity enhanced feature representation by optimizing data spatial and temporal information

First Baseline Model Implementation:

CNN2D model was trained on pre-processed dataset which needed 9135 training parameters and in return gave 93% accuracy.

However, as defined high-computational complexity together with accuracy were spotted.

Second Proposed MCNN (CNN3D) Implementation:

Constructed Multiscale Convolutional Neural Network (MCNN) that has a CNN3D architecture.

Utilization of space-time along with depth-separable modules and multi-scale and depth features.

Where 3306 training parameters were trained only achieving 94% accuracy thus complexity was greatly brought down.

Extension Model Development:

Utilization of all CNN, GRU and Bidirectional approaches into a singular model.

The number of training parameters were reduced this time round to 1162 and accuracy was at 96%.

Training optimized through the use of hybrid features developed for generalization.

Model Evaluation and Metrics:

Accuracy, precision, and recall with confusion matrices were the metrics used to evaluate all models.

Compared models in the CNN2D, MCNN, Hybrid model context to methods and accuracies and efficiencies improved.

Visualization and Results:

For all models training epoch acc versus each model was plotted to showcase steady constant success of advancements.

All algorithms were evaluated on accuracy and respective metrics, plots displaying such algorithms presented as bar graphs.

System Testing:

Finally, testing was done through viewed data giving a practical perspective to the efficacy of the model.

Demonstrated the ability to scale by implementing the hybrid model on bigger data sets and different types of activities.

EVALUATION:

Precision:

Formula: $\text{Precision} = \frac{\text{True Positives}}{\text{True Positives}+\text{False Positives}}$

Recall (Sensitivity):

Formula: $\text{Recall} = \frac{\text{True Positives}}{\text{True Positives}+\text{False Negatives}}$

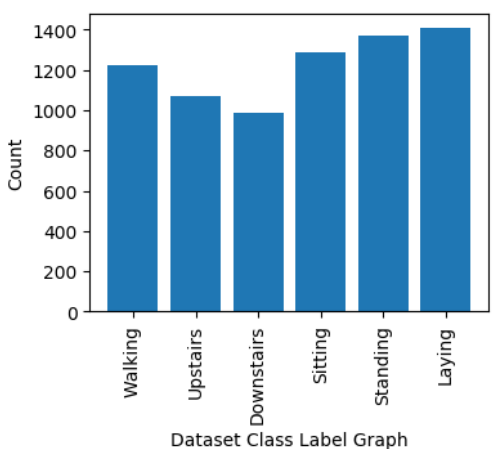
F1 Score:

Formula: $F1 = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$

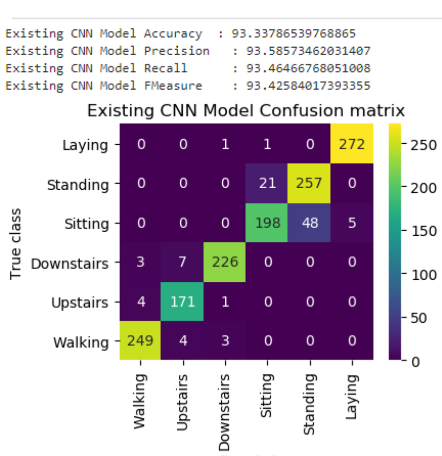
Accuracy:

Formula: $\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}}$

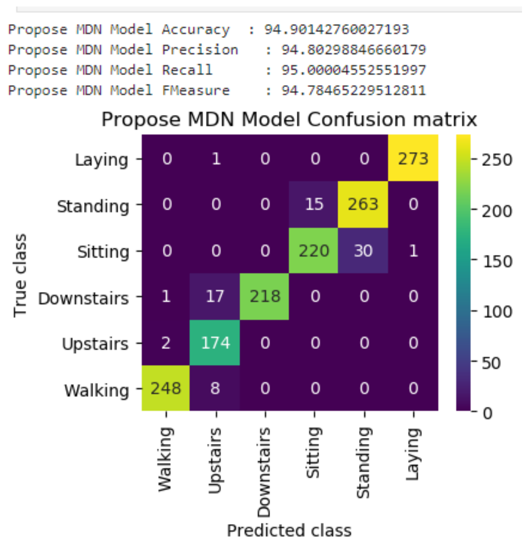
RESULTS:



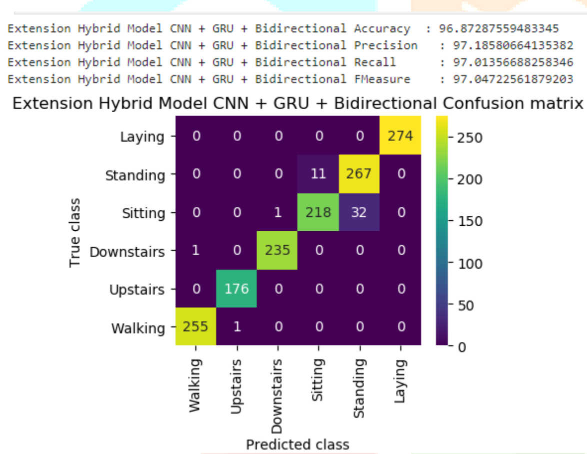
Activities found in dataset where x-axis represents ACTIVITY NAMES and y-axis represents count of those activities



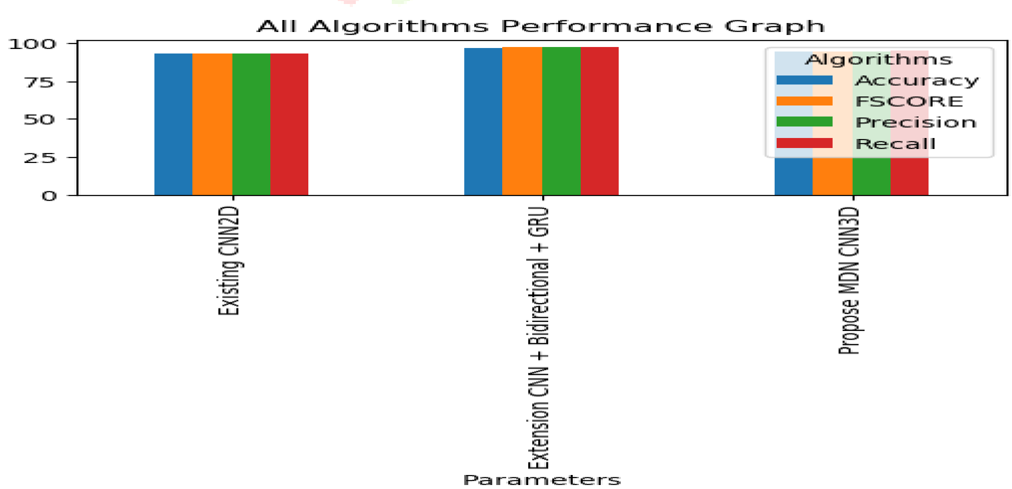
CNN2D model got 93% accuracy and can see other metrics and in confusion matrix x-axis represents predicted Labels and y-axis represents True Labels and all blue colour boxes represents incorrect prediction count and different colour boxes represents correct prediction count



MCNN MDN model got 94% accuracy and can see other metrics also



In above screen extension model got 96% 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

	Algorithm Name	Accuracy	Precision	Recall	FSCORE
0	Existing CNN2D Model	93.337865	93.585735	93.464668	93.425840
1	Propose MDN CNN3D	94.901428	94.802988	95.000046	94.784652
2	Extension Hybrid CNN + Bidirectional + GRU Model	96.872876	97.185807	97.013567	97.047226

Displaying all algorithms performance

Prediction:

```
Test Data : [ [illegible]
 -0.99542872 -0.98183758 -0.96535501 -0.93672818 -0.5694064  -0.80282975
  0.8461184   0.68774857  0.82947703 -0.9849843  -0.9999645  -0.99987755
 -0.99906323 -0.99537075 -0.98262514 -0.96306588 -0.54744428 -0.56341673
 -0.37817963  0.2738687  -0.0969558   0.09472723  0.0656154   0.21101292] Predicted Activity ===> Standing

Test Data : [ 0.25854868 -0.03295751 -0.07983694 -0.94612869 -0.884968   -0.91949304
 -0.96192131 -0.92194522 -0.93771352 -0.79892298 -0.47864893 -0.74335247
  0.761199    0.55993976  0.72176808 -0.9331412  -0.99825327 -0.99659775
 -0.99482814 -0.97354438 -0.96275079 -0.94524443 -0.37995875 -0.65145625
  0.02734375 -0.28082858  0.37673983 -0.09994524 -0.0038699   0.14812563] Predicted Activity ===> Laying
```

Predicted activity as Standing or any other activity

CONCLUSION

This project has proved that it is possible to successfully implement a Multiscale Convolutional Neural Networks (MCNN) by using various human behavior recognition techniques. The model significantly expedites training and system specification by attaining 94% accuracy with low levels space-time and depth-separable modules as compared to classic two-dimensional models of convolutional neural network. The hybrid extension model of optical neural networks with GRU and Bidirectional algorithm increases accuracy up to 96% with reduced parameters. This framework is scalable and many valued in the former model and operational cost which bridged the existing gaps most emphasizes real time application in healthcare, surveillance and activity monitoring. The proposed

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.