Grenze International Journal of Engineering and Technology, June Issue

Human Activity Recognition using Deep Learning Techniques

Prachi Singh[1], Khushboo Kumari[2], Khushboo Chaudhary[3], Haripriya Srivastav[4], Daya Shankar Srivastava[5] and Arjun Singh[6]
[1-6]Greater Noida Institute of Technology, Noida, India
Email: singh.prachi.2004@gmail.com, Khushbu1958@gmail.com, haripriyasrivastava469@gmail.com, Khushboosingh1286@gmail.com, Srivastava.daya@gmail.com, innovativearjun-singh@gmail.com

***Abstract*— Human Activity Recognition using with new technologies and upcoming advancements such as deep learning techniques has emerged positive and improved results. With deep learning various enhancement has been shown in HAR. These HAR based systems are very useful in healthcare sector, smart homes and security purposes also. With the help of IoT devices, wearable devices using sensors the human activities such as walking, sitting, lying can be easily measured and diagnosed which lead to easily access the required information for identification and classification purpose. Using the techniques of deep learning and algorithms such as LSTM and CNN has emerged as advanced version in HAR and helpful for various performance increment like high accuracy, Precision etc.**

***Keywords*— Deep Learning, CNN-LSTM, Wearable Sensors, Accelerometer, Gyroscope.**

I. INTRODUCTION

In recent years Human Activity Recognition has developed as a crucial system because it has gain important roles such as monitoring in healthcare sector, smart access to environments, advanced security reasons, analysis of sports etc. The main purpose of HAR is identification and classification of human activities like walking, sleeping, lying etc. and analysis of these activities done by human for a valuable purpose. This can be done by the use of IOT de-vices, wearable sensors and vision-based systems. Early HAR systems has relied on traditional machine learning techniques but they had only manual feature extraction which was less efficient and time consuming [1]. Some of its algorithms are Support vector Machines, K-nearest neighbours, Decision trees, Random Forest. Overtime, we have achieved high advancements in computational technologies and power where deep learning played a major role. The new study presents a comparative analysis of traditional HAR techniques [2] and highlighting strength, limitations and performance differences. Recent advancements in computing capabilities and the availability of large amount of data have encouraged to use deep learning approaches. This approach includes various algorithms like Convolutional Neural network, Recurrent neural network, Long -short term memory etc. These models are capable of extracting and evaluating human activities which help to improve accuracy and scalability of the solution [3]. The comparison between the older technology and the newer is about the better advantages, better performance and less-er limitations [4]. The significant progresses are achieved through deep learning and shows its better performance while solving a problem and overcoming the challenges it faced with the manual extraction of the features in engineering and also in complex problem recognition and solving.

Grenze ID: 01.GIJET.12.2.67_20

It has become an important topic in research domain since it is more effective in various areas like healthcare systems, fitness tracking, camera surveillance so its main goal is to automatically identify the human activities performed on daily basis [5], analysing the data collected with it through sensors and processing it.[3].

II. Literature Review

In recent years, Human Activity Recognition (HAR) has developed because it has a vast range of applications and uses in different sectors specially in healthcare sector, smart homes, surveillances systems for security purpose also. The main focus of Human activity recognition includes identification and classification of various activities per-formed by human bodies such as running, walking sitting etc. [6] automatically using vast collection of data from IoT devices like sensors and wearable objects. The main focus is to provide accurate results in activity recognition to the sole problems.

A. Traditional Approaches to Har

Previous Har systems were depended upon handcrafted feature extraction methods integrated with traditional tech-niques of machine learning like KNN, Support Vector Machine (SVM), Decision trees etc. The extraction and classi-fication of Spatial and temporal feature were done manually from accelerometer and gyroscope signals [7]. These approaches were not highly accurate and applicable only for simple activities not complex ones.

B. Implementation of Deep Learning in Har

There were various limitations and constraints of Traditional methods used in HAR due to which it was not so adapt-able [8]. Hence Deep learning techniques were implemented with HAR which has the capability of learning discrim-inative features directly from sensor data without manual intervention.

C. CNN Convolutional Neural Network)

Mostly Used in Har Because Of Its Ability of Capturing Local Spatial Patterns from Data Collected from Sensors. One Dimensional Cnns Are Capable of Classifying Sensor Data of Time Series and the 2-Dimensional Cnns Are Able to Sense Images. Effective Learning and Representations of Motion Patterns.

D. LSTM and RNN Models

These Are Used Because Of Their Capability to Model Temporal Dependencies in Linear Data. Lstm Based Har System Are Considered to Be More Robust System and Are Able to Recognize Both Type of Activities Whether Short Term or Long Term [9]. Lstm Is Mostly Preferred Due To High Computational Cost of Rnn Models.Here is a comparison table represented in Fig 1.0 showing between traditional approaches and deep learning approaches in HAR.

TABLE I. COMPARISON TABLE BETWEEN TRADITIONAL METHODS AND DEEP LEARNING APPROACHES

Aspect	Traditional HAR Methods	Deep Learning–Based HAR Meth-ods
Feature Extraction	Manually done, handcrafted features	Automatic feature learning from raw data
Algorithms Used	SVM, Decision tree, RF, etc.	CNN, LSTM, GRU, Hybrid
Data Dependency	Proper data prepro-cessing and selection	Works on raw or less processed data
Temporal Modelling	Limited or indirectly done	Strong temporal modelling (LSTM, Transformer)
Spatial Pattern Learning	Not explicitly mod-elled	CNNs capture spa-tial and local motion patterns
Accuracy & Scala-bility	Moderate, Poor scalability for com-plex activities	High, Good scalabil-ity

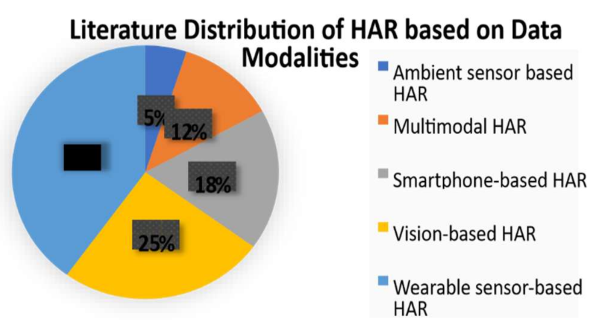

Fig 1. Literature Distribution of Human Activity Recognition Based on Data Modalities

In short, Deep learning has changed Human Activity Recognition system in a positive and improved way. Ongoing researches continues to aim on more improvement on its robustness and efficiency.
According to the application of deep learning models based on different human activities are gained from IoT devic-es. In Fig 2. Various types of HAR categorized on the basis of data modality.

III. PROPOSED MODEL

A. Overview

In this model, the purpose is recognition of various human activities by proper learning of machine to both spatial and temporal patterns. This can be achieved by the use of different kind of sensors or data containing videos [10], all these will perform using the concept of deep learning as shown in Fig 3 Various activities of human like walking, running or any movement in body should be recognized perfectly and captured using a hybrid approach [11] of Convolutional Neural Network and Long Short Temporal memory framework. (CNN-LSTM model).

B. Architecture Of Model

This HAR system contains various major components such as:

C. Data Accession

The Data Input Can Be Gained Through- Weara-ble Sensors Such as Accelerometer, Gyroscope Etc.
Videoframes like RGB or grayscale.

D. Data Preprocessing

The Preprocessing of Data Follows Various Steps for Betterment in Performance of Model Such As: Removing Noise Is One of The Major Processes Using Different Techniques of Filtering. Sensor Values Normalization Is Also Required for Proper Output [12]. Segmentation And Reshaping of Data So That It Is Suitable for Input in Deep Learning.

E. Extraction of Feature using Cnn

The Convolutional Neural Network (Cnn) Is, In Present, The Best Technology for Extracting Spatial Features Automatically from Data. The Convolution Layers Have the Capability of Capturing Spatial Patterns. Cnn Has the Functionality of Eliminating Features That Are Handcrafted and Also Increases the Strongness.

F. Lstm Used for Temporal Modelling

The Layers of Lstm Are Used to Capture Long Term Temporal Dependencies [13].

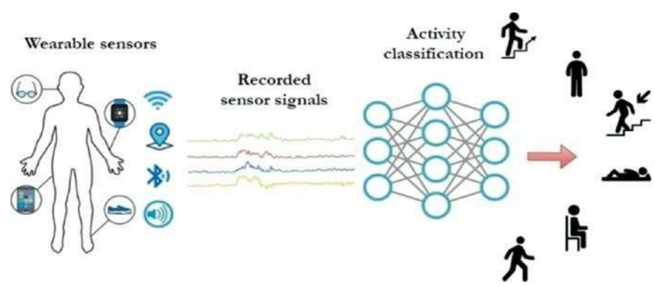

Fig 3. Showing Architecture of model

LSTM has the capability of storing past data as they have memory cells. It prevents from gradient issues from van-ishing.

G. Classification of Activities

The Last and Final Layers of Output Uses: Activation of Softmax for Multi-Class Classification.
Each of the specific activity of human is responded by neurons.

IV. RESULT AND DISCUSSION

The evaluation of HAR model using a dataset that is available to all public consisting of various activities of human like walking, running, moving etc. There was a division in the dataset into two different categories of training dataset and testing dataset in some ratio [14]. In this model deep learning was used for multiple epochs. The evaluation of performance was done using standard metrices such as precision, accuracy, F1 score etc. as they are commonly used for better configuration of results.

A. Comparison with Existing Methods

As compared to traditional methods and existing models such as SVM and RF, the proposed model in HAR using deep learning has an improved and increased performance. These improvements are happened because of various reasons such as:

The functionality of extracting features automatically, Effectiveness shown in temporal dependency modelling.

Lesser dependency of handcrafted feature. This proposed model based on CNN+LSTM techniques has a very strong performance and an effective recognition among all human activities as shown in Fig 4[15] The layers in convolutional network are effective in the extraction of various spatial features and the layers of LSTM are enough capable of temporal dependencies among activities.

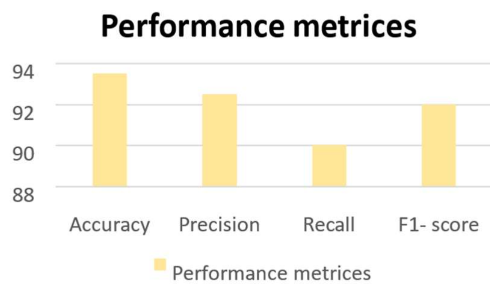

Fig 4. Overall performance metrices

B. Overall Classification Accuracy

This model has aimed at:

Overall accuracy of high classification performance High effectiveness in human activity recognition.

It can learn complex patterns from data.

V. CONCLUSION

This work has represented an effective approach of deep leaning for Human Activity recognition. This model used a hybrid CNN–LSTM framework for successful application and improved results from traditional methods. This model has capability of learning both features like spatial and temporal features very effectively [16], it leads to better understanding and classification of human activities like walking, running, lying etc. The increased accuracy of this model is result of using deep learning techniques. This model has achieved higher accuracy rates and recognition of various daily activities. This model achieves various standard performances from the experimental result it is con-cluded that it has high F-1 score, Precision, recall and Accuracy. This shows that it is robust and reliable. The combi-nation of CNN in which convolutional layers have capability of feature extraction automatically and LSTM layers can efficiently capture various dependencies such as temporal in linear data. On comparison to traditional methodol-ogies [17] which were used in previous decades with the recent methodologies using techniques of machine learning and approaches of deep learning has founded that it reduces errors due to handcrafted features and errors due to handcrafted features and enhances the identification and classification processes [18]. This results that the suggested framework is fit for actual world application like usage in healthcare sector, smart homes, etc.

FUTURE SCOPE

Several improvements and enhancement can be done in future research even if this model has promising results. These improvements can be done in:

Improvement of temporal feature learning and discrimination in activities.

Integration of information sourced from many sources like accelerometers, cameras, IOT devices can increase recognition. It may enable real-time activity recognition in lower costs by the deployment of this model into devices such as mobiles.

Diversion in datasets and increment in quantity of datasets can improve the training model with different users.

REFERENCES

[1] F. Xu, X. Gao, and W. Wang, "Human activity recognition using attention-enhanced deep neural net-works," IEEE Access, vol. 13, pp. 45521–45534, 2025.

[2] S. Dutta, T. Boongoen, and R. Zwiggelaar, "Deep learning-based human activity recognition: Recent ad-vances and challenges," IEEE Sensors Journal, vol. 25, no. 2, pp. 1345–1360, 2025.

[3] J. Shin, M. Hasan, and S. Nishimura, "Hybrid CNN–LSTM architecture for video-based human activity recognition," IEEE Access, vol. 12, pp. 118900–118915, 2024.

[4] A. Copetti, L. Silva, and M. Nunes, "Sequential image-based human activity recognition using CNNLSTM networks," IEEE Latin America Transactions, vol. 22, no. 11, pp. 1892–1901, 2024.

[5] M. Islam, M. Jannat, and S. Lee, "STC-NLSTMNet: Deep CNN-LSTM framework for sensor-based human activity recognition," IEEE Sensors Journal, vol. 23, no. 14, pp. 15672–15684, 2023.

[6] F. Duan et al., "Multi-task deep learning for sensor-based human activity recognition," IEEE Internet of Things Journal, vol. 10, no. 9, pp. 7821–7834, 2023.

[7] S. Suh and P. Lukowicz, "Transformer-based deep learning model for wearable sensor human activity recognition," IEEE Access, vol. 10, pp. 112345–112357, 2022.

[8] M. Khatun et al., "Deep CNN-LSTM with self-attention for human activity recognition using wearable sen-sors," IEEE Journal of Translational Engineering in Health and Medicine, vol. 10, pp. 1–11, 2022.

[9] Y. Wang, K. Yu, and H. Xue, "Multi-view deep learning framework for human activity recognition," IEEE Transactions on Industrial Informatics, vol. 18, no. 4, pp. 2567–2576, 2022.

[10] E. Ramanujam, T. Perumal, and S. Padmavathi, "Human activity recognition with smartphone and weara-ble sensors using deep learning: A review," IEEE Sensors Journal, vol. 21, no. 12, pp. 13029–13040, 2021.

[11] R. San-Segundo et al., "Human activity recognition using deep learning models on accelerometer data," IEEE Access, vol. 9, pp. 146012–146024, 2021.

[12] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," Pattern Recognition Letters, vol. 119, pp. 3–11, 2019.

[13] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," Expert Systems with Applications, vol. 59, pp. 235–244, 2016.

[14] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," Proc. IJCAI, pp. 1533–1540, 2016.

[15] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1192–1209, 2013.

[16] D. Roggen et al., "Collecting complex activity datasets in highly rich networked sensor environments," Proc. IEEE INSS, pp. 233–240, 2010.

[17] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," Proc. Pervasive Computing, pp. 1–17, 2004.