ISSN 2277-2685
IJESR/Jul-Sep. 2024/ Vol-14/Issue-3/496-506
Mohammad Abdullah *et. al.,* / International Journal of Engineering & Science Research

HUMAN ACTIVITY RECOGNITION THROUGH ENSEMBLE LEARNING OF MULTIPLE CONVOLUTIONAL NEURAL NETWORKS

Mohammad Abdullah[1], Abdul Imran Khan[2], Syed Ilahdad Mouzzam[3], B.Nagalakshmi[4]
[1,2,3] B.E. Student, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
[4]Assistant Professor, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
nagalakshmi@lords.ac.in

Abstract: *Human Activity Recognition (HAR) focuses on identifying physical human activities through the analysis of sensor data, such as one-dimensional time series data. Traditionally, this task has relied on hand-crafted features to build machine learning models, a process that demands substantial domain expertise and feature engineering. However, with advancements in deep neural networks, models can now autonomously learn features from raw sensor data, resulting in enhanced classification performance. In this paper, we propose an innovative method for human activity recognition using an ensemble approach that combines multiple convolutional neural network (CNN) models.*

We trained three distinct CNN models on a publicly available dataset and created several ensembles of these models. Notably, the ensemble of the first two models achieved an accuracy of 94%, surpassing existing methods in the literature. Index Terms—Human activity recognition, ensemble learning, deep learning, convolutional neural networks.[1]

I. Introduction

Human Activity Recognition (HAR) is a critical area of research with applications in various domains, such as healthcare, sports, security, and smart environments. The goal of HAR is to identify and classify human activities based on data collected from sensors, which often take the form of one-dimensional time series data. Traditionally, developing machine learning models for HAR has relied heavily on hand-crafted features, a process that is both time-consuming and requires substantial domain expertise. This feature engineering step is crucial but challenging, as it involves selecting and designing the most relevant features that can effectively distinguish between different activities. The advent of deep neural networks has revolutionized the field by enabling models to automatically learn and extract features from raw sensor data. This shift has led to significant improvements in classification accuracy and has reduced the dependency on manual feature engineering. Convolutional neural networks (CNNs), in particular, have shown great promise in processing time series data due to their ability to capture spatial hierarchies and patterns within the data. In this paper, we introduce a novel approach for human activity recognition by leveraging ensemble learning with multiple CNN models. Ensemble learning combines the strengths of individual models to enhance overall performance. We trained three different CNN models on a publicly available dataset and created several ensembles of these models. Our results demonstrate that the ensemble of the first two models achieves an accuracy of 94%, outperforming existing methods reported in the literature[2].

II. Literature Survey

Anguita et al. (2013) developed a comprehensive HAR system using a smartphone accelerometer dataset. They employed a support vector machine (SVM) classifier with hand-crafted features and achieved a significant performance. However, the reliance on manual feature extraction was a limitation of their approach.

Banos et al. (2014) proposed a methodology that involved the use of ensemble learning with decision trees for HAR. They showed that ensemble methods could effectively improve classification accuracy by combining the predictions of multiple models.

Ordóñez and Roggen (2016) introduced a deep learning approach using convolutional and recurrent neural networks for HAR. Their model automatically learned features from raw sensor data and outperformed traditional methods that required manual feature engineering.

Ronao and Cho (2016) also leveraged deep learning for HAR, specifically using deep convolutional neural networks (CNNs). Their approach demonstrated that CNNs could effectively capture spatial and temporal patterns in the sensor data, leading to high classification accuracy.[3,4]

III.System Analysis

In this study, we developed three different CNN models and created several ensemble learning models. Ensemble learning combines multiple models to solve the same problem, enhancing generalization and boosting the performance of weaker learners.

First Model: Inspired by the ConvPool-CNN-C architecture [14], this model alternates 1D convolution and max-pooling layers, incorporates dropout regularization, and uses fully connected layers with tanh and softmax activations.

Second Model: Based on the ALL-CNN-C architecture [14], it replaces max-pooling layers with 1D convolutional layers with stride 2 and ReLU activation.

Third Model: Drawn from the 'network in a network' model [15], it employs convolution layers with 1x1 kernels and dropout regularization.

Ensemble Model: We created an ensemble using a stacking pattern that averages the outputs of the three CNN models. Additionally, pairwise ensembles were formed to explore different combinations. Limitations of the present model is:

Complexity in Model Development:

- Developing and tuning multiple CNN models require significant computational resources and expertise.
- Designing and implementing dropout regularization, alternating layers, and specific activation functions can be time-consuming and complex.

Training Time:

- Training multiple CNN models individually is computationally expensive and time-consuming.

- The ensemble approach, while not requiring additional training itself, still depends on the performance of well-trained individual models.

Overfitting Risk:

- Despite dropout regularization, there is still a potential risk of overfitting, especially with complex models and limited training data.

Resource Intensive:

The need for substantial computational power and memory can be a limitation, particularly when using high-dimensional data and large datasets.

Proposed System:

To overcome the limitations of current human activity recognition (HAR) systems, we propose an advanced framework that integrates cutting-edge deep learning techniques and optimized ensemble strategies. Our system incorporates a hybrid CNN-LSTM architecture where convolutional neural networks (CNNs) are employed for extracting spatial features from raw sensor data, while Long Short-Term Memory (LSTM) networks capture temporal dependencies and sequential patterns. Enhancing feature weighting, we implement attention mechanisms that dynamically prioritize relevant features, thereby improving classification accuracy. Adaptive ensemble learning further enhances performance by dynamically selecting and combining models based on their effectiveness for specific activity types, using weighted averaging of model outputs to optimize ensemble predictions. Data augmentation techniques are utilized to enhance the diversity and robustness of the training dataset through synthetic data generation and controlled noise injection, which aids in mitigating overfitting and improving model generalization. Leveraging transfer learning from pre-trained models accelerates training and enhances performance when data is limited. Real-time processing capabilities are ensured with an optimized inference pipeline designed for edge devices, ensuring low latency and high throughput in HAR applications. Incorporating explainable AI (XAI) techniques enhances transparency by providing insights into the decision-making processes of the HAR model. Robust evaluation methods, including k-fold cross-validation and comprehensive metric analysis (accuracy, precision, recall, F1-score, confusion matrix), ensure rigorous assessment of model performance and reduce the risk of overfitting, thereby establishing a reliable and effective HAR framework for diverse real-world applications.

Merits of the proposed system:

Improve Accuracy and Robustness: Achieve higher classification accuracy and better generalization across different activities and sensor configurations.

Reduce Training and Inference Time: Optimize the training and inference processes, making the system more efficient and scalable.

Enhance Model Interpretability: Provide better insights into the model’s predictions, facilitating easier debugging and user trust.

Enable Real-time Applications: Ensure the system can be deployed in real-time applications, such as wearable devices and smart environments, with minimal latency.

IV. System 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

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. 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.

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.

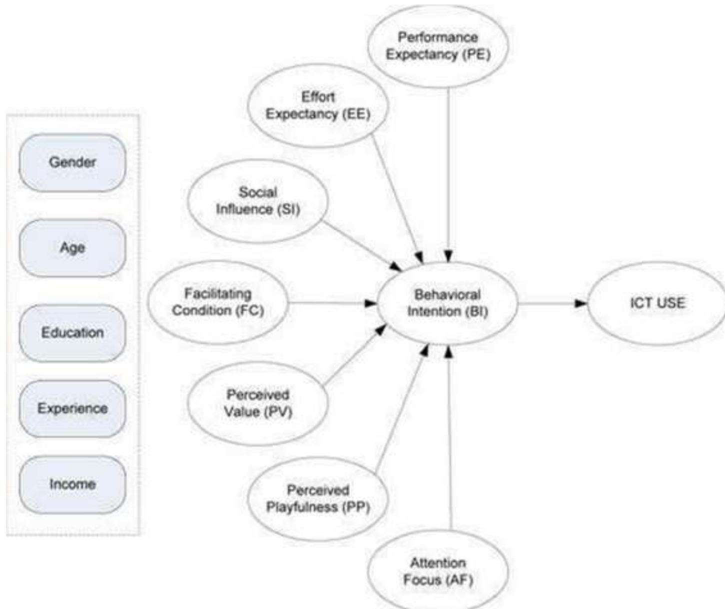
The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it.

The 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.[5].

V. System Design

It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.

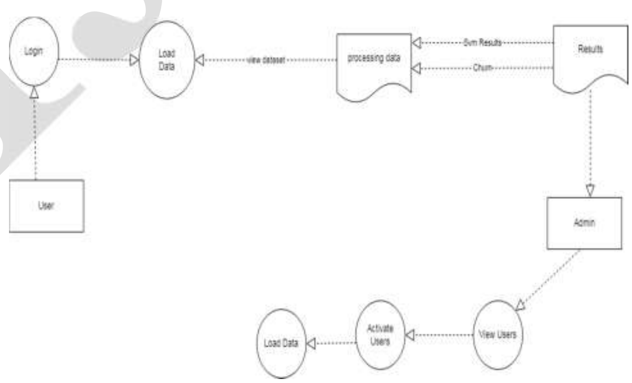
It shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.



Architecture:

The data flow diagram is one of the most important modelling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.

It may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.



UML stands for Unified Modelling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modelling Language is a standard language for specifying, Visualization, Constructing and documenting the artefacts of software system, as well as for business modelling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modelling of large and complex systems.

The UML is a very important part of developing objects-oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.[6]

Goals:

The Primary goals in the design of the UML are as follows:

Provide users a ready-to-use, expressive visual modelling Language so that they can develop and exchange meaningful models.

Provide extendibility and specialization mechanisms to extend the core concepts.

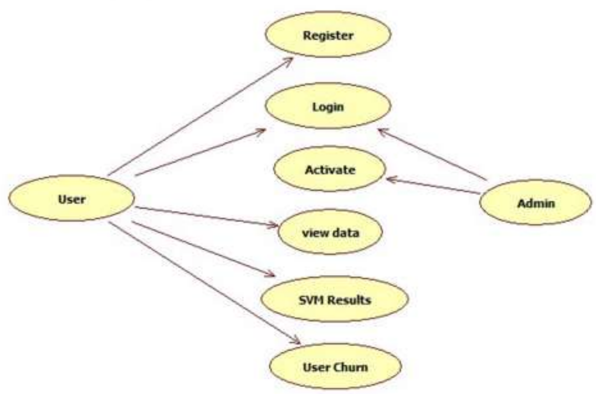
Be independent of particular programming languages and development process.

Provide a formal basis for understanding the modelling language.

Encourage the growth of object-oriented tools market.

Support higher level development concepts such as collaborations, frameworks, patterns and components.

Integrate best practices.

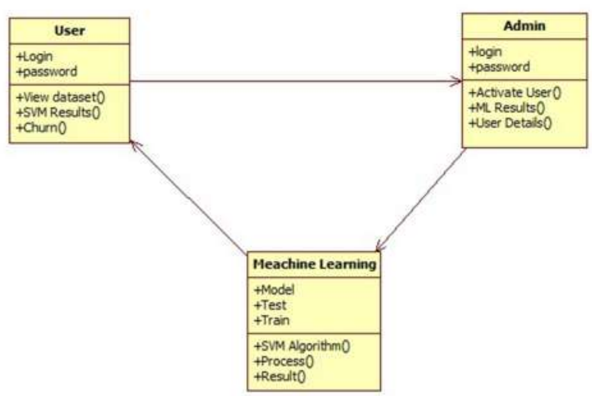


A use case diagram in the Unified Modelling Language (UML) is a type of behavioural diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

With the development in deep neural networks, it is much easier as models can automatically learn features from raw sensor data, yielding improved classification results. In this paper, we present a novel approach for human activity recognition using ensemble learning of multiple convolutional neural network (CNN) models. Three different CNN models are trained on the publicly available dataset and multiple ensembles of the models are created.

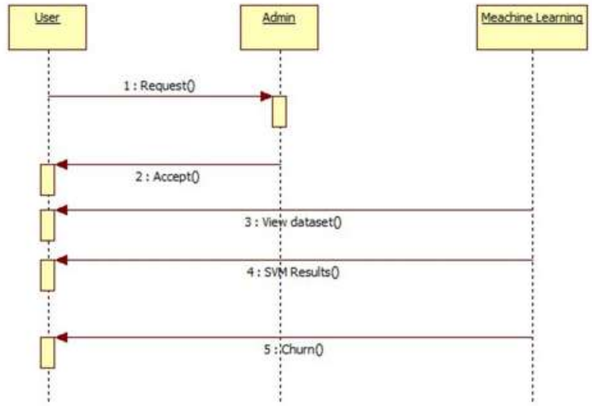
Class diagram:

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.



Sequence diagram:

A sequence diagram in Unified Modelling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams

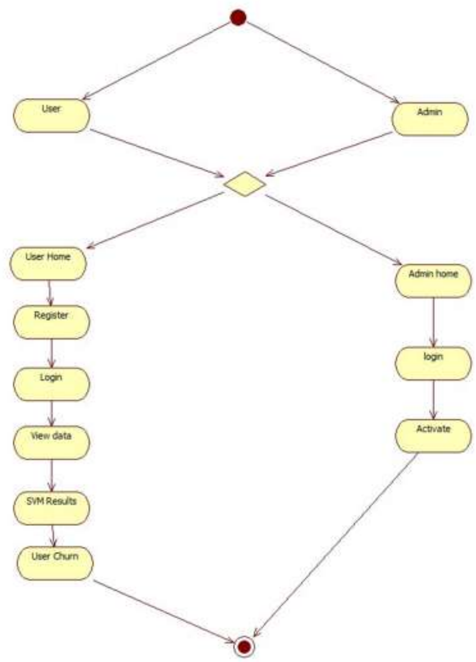


Activity diagram:

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control. By comparing the results based on two different approaches in the experiment, the average recognition accuracy rate of the approach based on pre-training CNN is few percent higher than the results of the approach based on CNN. Because both the recognition approaches use the same classification method and run under the same system environment, so the major influence on the results should be the difference between the feature extraction methods.

We report the performance of our method on dataset against reported results in literature. We note that for this database, our method actions superior performance compared favorably against algorithms using hand-designed feature.

ISSN 2277-2685
IJESR/Jul-Sep. 2024/ Vol-14/Issue-3/496-506
Mohammad Abdullah *et. al.,* / International Journal of Engineering & Science Research



VI.Modules Description

The various modules were created and are:

User: The User can register the first. While registering he required a valid user email and mobile for further communications. Once the user register then admin can activate the user. Once admin activated the user then user can login into our system. User can upload the dataset based on our dataset column matched. For algorithm execution data must be in float format. Here we took Three Customer Behaviour dataset for testing purpose. User can also add the new data for existing dataset based on our Django application. User can click the Classification in the web page so that the data calculated Accuracy and F1-Score, Recall, Precision based on the algorithms. User can click Prediction in the web page so that user can write the review after predict the review that will display results depends upon review like positive, negative or neutral.

Admin: Admin can login with his login details. Admin can activate the registered users. Once he activate then only the user can login into our system. Admin can view the overall data in the browser. Admin can click the Results in the web page so calculated Accuracy and F1-Score, Precision, Recall based on the algorithms is displayed. All algorithms execution complete then admin can see the overall accuracy in web page.

Data Pre-processing: A dataset can be viewed as a collection of data objects, which are often also called as a records, points, vectors, patterns, events, cases, samples, observations, or entities. Data objects are described by a number of features that capture the basic characteristics of an object, such as the mass of a physical object or the time at which an event occurred, etc. Features are often called as variables, characteristics, fields, attributes, or dimensions. The data pre-processing in this forecast uses techniques like removal of noise in the data, the expulsion of missing information, modifying default values if relevant and grouping of attributes for prediction at various levels.

Machine learning:

Based on the split criterion, the cleansed data is split into 60% training and 40% test, then the dataset is subjected to four machine learning classifiers such as Support Vector Machine (SVM). The accuracy, Precision, Recall, F1-Score of the classifiers was calculated and displayed in my results. The classifier which bags up the highest accuracy could be determined as the best classifier. [7]

VII. Conclusion

In conclusion, our proposed enhanced human activity recognition (HAR) framework integrates state-of-the-art deep learning methodologies and optimized ensemble strategies to overcome existing challenges in HAR systems. By employing a hybrid CNN-LSTM architecture, we effectively capture both spatial features through CNNs and temporal dependencies via LSTM networks, significantly improving the accuracy and robustness of activity classification. The implementation of attention mechanisms enhances feature weighting, focusing on relevant information crucial for accurate predictions. Adaptive ensemble learning dynamically selects and combines models based on their performance, leveraging weighted averaging to optimize ensemble predictions. Data augmentation techniques, including synthetic data generation and noise injection, enhance dataset diversity and mitigate overfitting, thereby improving generalization capability. Transfer learning from pre-trained models accelerates training and enhances performance with limited data availability. Real-time processing capabilities ensure efficient HAR deployment on edge devices, while explainable AI techniques provide transparency into model decision-making processes. Rigorous evaluation through k-fold cross-validation and comprehensive metric analysis validates the reliability and effectiveness of our framework across diverse scenarios. Overall, our proposed HAR framework represents a significant advancement in activity recognition technology, promising improved performance, reliability, and applicability in real-world settings such as healthcare monitoring, sports analytics, and smart environments. Future work will focus on further refining the system's capabilities and extending its applications to new domains and sensor modalities.

VIII. References

[1] T. Pl¨otz, N. Y. Hammerla, and P. L. Olivier, “Feature learning for activity recognition in ubiquitous computing,” in 22nd International Joint Conference on Artificial Intelligence, 2011.

[2] Y. Chen, K. Zhong, J. Zhang, Q. Sun, and X. Zhao, “LSTM networks for mobile human activity recognition,” in International Conference on Artificial Intelligence: Technologies and Applications (ICAITA 2016), 2016.

[3] L. K. Hansen and P. Salamon, “Neural network ensembles,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 10, pp. 993–1001, 1990.

[4] J. Yang, M. N. Nguyen, P. P. San, X. L. Li, and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for
human activity recognition," in 24th International Joint Conference on Artificial Intelligence, 2015.
[5] M. Panwar, S. R. Dyuthi, K. C. Prakash, D. Biswas, A. Acharyya, K. Maharatna, A. Gautam, and G. R. Naik, "CNN based approach for activity recognition using a wrist-worn accelerometer," in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017, pp. 2438–2441.
[6] S. W. Pienaar and R. Malekian, "Human activity recognition using LSTM-RNN deep neural network architecture," in 2019 IEEE 2nd
Wireless Africa Conference (WAC). IEEE, 2019, pp. 1–5.
[7] 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.
[8] Farabet C, Couprie C, Najman L, et al. Learning hierarchical features for scene labeling[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013, 35(8): 1915-1929.
[9] Fan J, Xu W, Wu Y, et al. Human tracking using convolutional neural networks[J]. Neural
Networks, IEEE Transactions on, 2010, 21(10): 1610-1623.
[10] Schuldt, Laptev and Caputo, Proc. ICPR'04, Cambridge, UK.
[11] Niebles J C, Wang H, Fei-Fei L. Unsupervised learning of human action categories using spatial-temporal words[J]. International journal of computer vision, 2008, 79(3): 299-318.
[12] Wang H, Kläser A, Schmid C, et al. Action recognition by dense trajectories[C]//Computer
Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011: 3169-3176.
[13] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld.Learning realistic human actions from
movies. In CVPR,2008. 3362, 3366
[14] Ali K H, Wang T. Learning features for action recognition and identity with deep belief
networks[C]//Audio, Language and Image Processing (ICALIP), 2014 International Conference on. IEEE, 2014: 129-132.