

DEEP LEARNING PROJECT BASED ON HUMAN ACTIVITY RECOGNITION

Sarumathi G¹, Susmitha T², Shivani S³, Mr. Sathish S⁴

¹Department of Computer Science and Business Systems & Bannari Amman Institute of Technology

² Department of Electronics and Communication Engineering & Bannari Amman Institute of Technology

³ Department of Computer Science and Engineering & Bannari Amman Institute of Technology

⁴ Department of Computer Science and Business Systems & Bannari Amman Institute of Technology

Abstract - The main objective of this paper is to implement human activities by performing some of the specified tasks. In order to classify a time series, the challenge of human activity recognition requires data from a number of timesteps. The methodology that would be followed in the paper is using deep learning and classification methods to make predictions with the model. To extract the local features of each frame, a convolutional neural network and Long Short Term Memory (LSTM) of the deep learning algorithm will be used. This implementation will help to monitor human activity and ensure their safety by sending a message notification through their mobile phones using Twilio packages that is a library. This is a programmable messaging software application that will send notifications from the prevention of danger and the monitoring process will be easy to know and convenient. The project will be developed as an application for the users to interact as well as analyze real-time data analysis. In this paper, we will discuss about how human activity recognition can effectively be employees to ensure safety and security of elderly people, women and also in burglary prevention. The complete implementation proposal along with the technical specifications have been designed so as to implement best possible solutions.

Keyword - CNN, LSTM, Deep Learning, Twilio packages, User Interface.

1. INTRODUCTION

Human Activity Recognition (HAR) is a type of deep learning task that applies to recognizing and identifying human activities based on sensor data. This task can be used to monitor and track people's physical activities, spot anomalies in activity patterns, and enhance the overall quality of life in a variety of industries, including healthcare, sports, and security. These models are trained and tested on large datasets from different sources including accelerometers, gyroscopes, and magnetometers.

The following phases make up the normal procedure for creating a HAR model using deep learning:

- Collecting and processing sensor data entails gathering information from the sensors that people wear while also extracting appropriate elements from the data.
- Data augmentation is the process of developing new versions of the existing data to enlarge the dataset. To do this, the sensor data is subjected to random modifications such as scaling, rotation, and noise.
- Model selection entails deciding which deep learning model architecture is best for the task.
- Using the pre-processed and enhanced datasets, the chosen model is trained.

- Using a different test dataset, the trained model's performance is assessed.
- To enhance the model's performance, it is adjusted and optimized.

2. LITERATURE SURVEY

Antonio Bevilacqua et al (2018) designed a study on "Human Activity Detection with CNN" investigates various activity-sensor combinations and demonstrates how alternative network designs can be used to modify motion signals for CNNs. When we analyse the performance of various sensor groups, the classification potential of single, double, and triple sensor systems is also explored.[2]

Artur Jordao et al (2019)researches numerous studies that represented the various activity categories using accelerometer, gyroscope, and magnetometer data showed impressive findings.Here,the author conducts a broad range of experiments to identify the weak points in wearable sensor-based human activity recognition by analysing various sample generation and validation methods.They implement and evaluate several top-performance methods, ranging from handcrafted-based approaches to convolutional neural networks.[4]

Lei Jin et al(2022) has introduced a system that solves the human activity data which greatly affect the interpersonal variability which brings the intra class similarity and the intraclass diversity.To solve this they developed a framework using a high-dimensional embedding layer and a loss function of Euclidean distance to improve the capacity of deep learning networks to mine discriminative features in order to address the issue.[7]

Jianyu Xiao et al(2021) introduces a general-purpose classifier that has been trained to identify the various activity patterns of most user types is what the HAR

baseline model basically is. It gives the information needed for the model personalization stages that follow.With the help of incremental learning based on the Broad Learning System (BLS) and Deep InfoMax (DIM), respectively.[8]

Jinyang Huang [9] et al (2020)studied a Anti-Interference Human Activity Recognition Based on WiFi Subcarrier Correlation Selection which states a problem statement that Modern WiFi-based HAR systems have recently been introduced due to their qualities of no invasion of privacy, high recognition accuracy rates, and no privacy leakage (RARs).[9]

Michalis Vrigkas et al (2015)reports the characteristics of future research directions and present some open issues on human activity recognition[11]

Nadia Oukrich et al(2017) states In order to identify human activities inside smart homes, the Auto-Encoder (AE) algorithm was used for unsupervised feature selection, and the Backpropagation (BP) algorithm was used for supervised learning to train reconstructed subsets.The effectiveness of the auto-encoder has then been assessed and contrasted with the conventional weighting method for selecting features.[12].

Thien Huynh et al (2021)introduces a physical Activity Recognition With Statistical-Deep Fusion Model Using Multiple Sensory Data for Smart Health that states the data on segmentation obtained from a multisensor system are encoded to an activity picture for deep model learning by converting the raw signal value to pixel intensity value.A multiclass support vector machine (SVM) classifier learns the hybrid feature combined from the handmade and deep features in the fusion model. [17]

Zhenguo Shi et al (2022)introduced a Environment-Robust WiFi-Based Human Activity Recognition Using

Enhanced CSI and Deep Learning and has developed to remedy some classification errors that do not follow the state transition of human behaviors, there is a prediction checking and correction system.[19]

From the above references, encouraged by the human activity recognition project is a type of time series classification problem where your data from the series of timesteps. It is one of the most important tasks for every business that relies heavily on user interface applications. The prediction and analysis will help us to understand how many users will become aware of a scenario or precaution that they might need to take and act accordingly.

3. PROBLEM STATEMENT

Human Activity Recognition is the problem of specifying a physical activity carried out by an individual dependent on a path of movement within an environment. This implementation will help to monitor human activity and ensure their safety by sending a message notification for prevention of danger. In this paper, we would have used the deep learning method by using classification methods to make predictions with the model.

4. VIDEO CLASSIFICATION METHOD

The task of labelling a video clip is known as video classification. This app is useful if you want to know what kind of activity is taking place in a video. A video is a collection of sequential frames or images that are played one after the other. Most videos we see in our daily lives have more than 30 frames per second. As a result, even for short videos, we must deal with a large amount of data in comparison to image classification. Because the images are highly correlated, it is common to skip the intermediate frames and process fewer frames per second. For example, a video may include a tree in

some frames, but the central label of the video may be something else (e.g., "hiking"). The task determines the granularity of the labels required to describe the frames and video. A typical task would be to assign one or more global labels to the video, as well as one or more labels to each frame within the video.

5. PROPOSED WORK

There are several deep learning approaches have been proposed for HAR, Some of which are as follows like Convolutional Neural Networks has own capacity to extract spatiotemporal features from sensor input, CNNs have been extensively used for HAR, Recurrent Neural Network is a type of neural network that is well-suited for time-series data, such as sensor data, because they can accept input sequences of varying length, LSTM is a type of RNN that may identify long-term dependencies in sequential data via selecting specific information. In this paper, After recognising the activities using LSTM algorithm, a message notification will sent to the nearby relatives, police station to ensure their safety from the prevention of danger using Twilio packages.

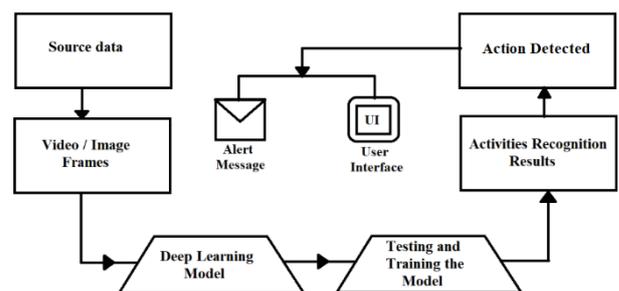


Fig-5.1 Flowchart of Working of the Application

6. IMPLEMENTATION PROCEDURES OF THE APPLICATION

There are some options for doing action classification. Users can view the action classification results by uploading their own input video and clicking the 'Submit' option. If a user uploads their own video, make sure it features a single person or multiple person performing the actions listed. Also, make sure to upload

a small video snippet to process it easily to find the results and actions being recognised.

7. TRAINING THE MODELS AND PREDICTION ANALYSIS

This application was created to demonstrate how 2D Pose Estimation and LSTM RNN machine learning models can be used to classify human actions. Detectron2 from Facebook AI Research is used to estimate 2D poses. The output of a 2D posture estimation algorithm applied to a series of consecutive video frames is used to categorize activities using an LSTM model. This application is trained on our own LSTM model from a data set. We have converted the dataset to map to Detectron2 output format since we are utilizing Detectron2 for pose estimation in order to train our LSTM model. In contrast to Detectron2, the original dataset was built using the Open Pose package, which generates 18 key points per human from the input image. The activity is divided into categories by the model, including jumping, jumping jacks, boxing, waving two hands, waving one hand, and clapping hands.

The activity recognition capabilities of LSTM have been successfully used in a number of industries, including entertainment, sports, and healthcare. However, a number of variables, including data quality and quantity, model architecture, and optimization strategies, affect how well LSTM performs. Hence, for accurate and trustworthy activity recognition, thorough consideration of these elements is essential.

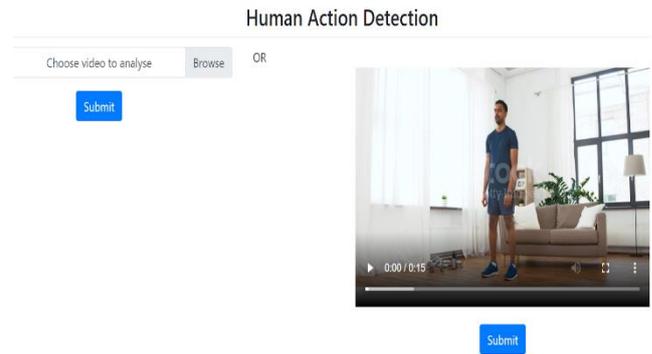


Fig-7.1 User Interface of Human Action Detection

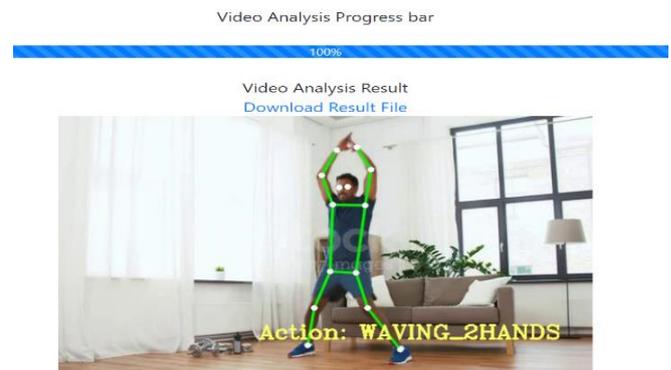


Fig-7.2 Video Analysis result of Activity Recognition

8. MESSAGE NOTIFICATION ALERT FROM TWILIO

With the help of the cloud communications platform Twilio, programmers may incorporate messaging into their applications. To make it simple for developers to incorporate communication features into their apps, Twilio provides a range of packages and APIs. Using the Twilio API, the Programmable SMS package enables developers to include SMS messaging functionality in their apps. Developers can send and receive text messages with Programmable SMS. By delivering a message notification for the purpose of preventing danger, this implementation will assist in keeping track of human behaviour and ensuring their safety.

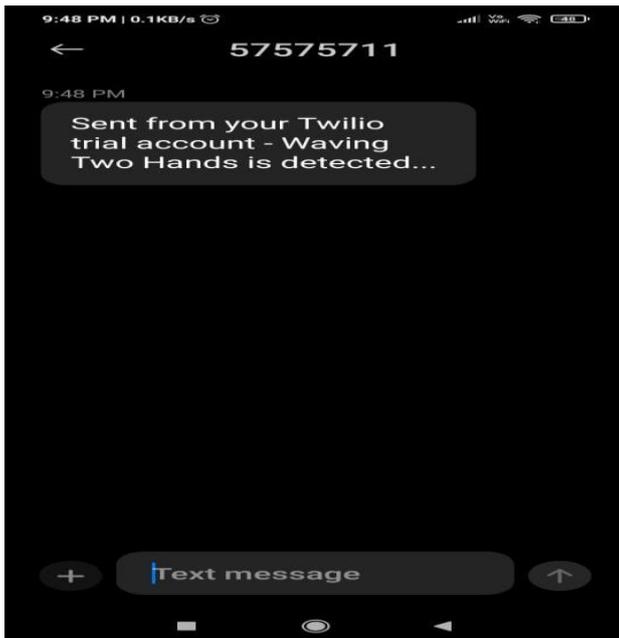


Fig-8.1 Notification Alert message through phone

9. REAL TIME USAGE OF THE APPLICATION

The development of numerous significant applications, including intelligent surveillance systems, human-computer interfaces, health care, security, and military applications, is made possible by the capacity to identify, comprehend, and predict complicated human actions. But, this application is specifically used for Emergency situations in hospitals for patients, Women's Safety, Accidents in Public Places, and Detection of breakage of law.

10. CONCLUSION

From this analysis and result, the unique feature of this application is that both Pose Detection and LSTM are implemented utilizing PyTorch modules. Python2 is used to create Detectron2. Using PyTorch, we created our LSTM model and trained it with PyTorch-lightning. The training code is made very simple and clear using pytorch-lightning. The libraries and the functionalities developed have been deployed in the UI and this is connected with the twilio software so that users can get timely notifications. These notifications can help people escape from danger and it will also help them stay safe

and secure. This application will benefit a wide number of people in specific areas of concern that include healthcare and other safety monitoring systems.

11. REFERENCE

- [1] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind K. Dey, Tobias Sonne, and Mads Møller Jensen, (2015) "Smart gadgets are unique in that: evaluating and reducing the heterogeneities of mobile sensing for activity identification", pp. 127–140.
- [2] Antonio Bevilacqua et al (2018) 'Deep Human Activity Recognition with Convolutional Neural Networks is designed for exploring several combinations of activities and sensors, investigating the classification potential of single, double and triple sensor systems'. Conference: ECML-PKDDAt: Dublin, Ireland.
- [3] Arnaud S. R. M. Ahouandjinou, Cina Motamed, Eugène C. Ezin, Antonio Pinti (2015) 'Credal human activity recognition based-HMM by combining hierarchical and temporal reasoning approach for video analysis by exploiting hidden Markov models'. International Conference on Image Processing Theory, Tools and Applications (IPTA) Electronic ISSN: 2154-512X.
- [4] Artur Jordao, Antonio Carlos Nazare, Jessica Sena, William Robson Schwartz (2019) 'Human Activity Recognition Based on Wearable Sensor Data implement and evaluate several top-performance methods, ranging from handcrafted-based approaches to convolutional neural networks'. Xiv:1806.05226v3 [cs.CV].
- [5] Daniele Ravi, Benny Lo, Charence Wong. (2016) 'A deep learning-based approach for human activity recognition is intended to provide accurate and immediate

classification', IEEE 13th International Conference
DOI:10.1109/BSN.2016.7516235.

[6] G. Batchuluun, J. H. Kim, H. G. Hong, J. K. Kang and K. R. Park, (2017) "Combining behaviour prediction with recognition in a fuzzy algorithm to recognise human behaviour", *Expert Syst. Appl.*, vol. 81, pp. 108-133.

[7] H. Basly, W. Ouarda, F.E. Sayadi, B. Ouni and A.M. Alimi, "Human activity recognition using CNN-SVM learning method" in *Image and Signal Processing. ICISP (2020). Lecture Notes in Computer Science, Cham:Springer*, vol. 12119.

[8] Jianyu Xiao, Haipeng Chen, Linlin Chen; Xuemin Hong (2021) 'Baseline Model Training in Sensor-Based Human Activity Recognition gives the information needed for the model personalization stages with help of incremental learning based on the Broad Learning System (BLS), Deep InfoMax (DIM)'. Published in: *IEEE Access (Volume: 9)*, pp(70261 - 70272)

[9] Jinyang Huang; Bin Liu; Chao Chen; Hongxin Jin; Zhiqiang Liu (2020) 'Human Activity Recognition Based on WiFi Subcarrier Correlation Selection introduced for qualities of no invasion of privacy, high recognition accuracy rates, and no privacy leakage (RARs)'. *IEEE Transactions on Vehicular Technology (Volume: 69)*, pp.6739 - 6754.

[10] M. Duong, T. Vu, H.H. Bui, H. Phung, Q. Dinh and V. Svetha, (2005) "Using the switching hidden semi-Markov model, activity identification and abnormality detection", *Proceedings of the IEEE Computer Society Conference CVPR*, pp. 838-845, 2005.

[11] Michalis Vrigkas Christophoros Niko, Ioannis A Kakadiaris (2015) 'A Review of Human Activity Recognition Methods for reporting the characteristics of

future research directions and present some open issues. *Frontiers in Robotics and AI* 2(28), DOI:10.3389/frobt.2015.00028 License: CC BY 4.0.

[12] Nadia Oukrich, EL Bouazzaoui Cherraqi (2017) 'Human activities recognition based on autoencoder pre-training and back-propagation algorithm Auto-Encoder algorithm (AE) has been used in unsupervised feature selection, Backpropagation (BP) algorithm has been used to train reconstructed subsets in supervised learning'. *Journal of Theoretical and Applied Information Technology* 95(19):5194-5202.

[13] Panos P. Markopoulos, "A limited training material for linear discriminant analysis," (2017) in *IEEE International Conference on Acoustics, Speech and Signal Processing*. pp. 4626–4630.

[14] Q. Li, W. Yang, X. Chen, T. Yuan and Y. Wang, "Network of connections between temporal segments for action detection", (2020) *IEEE Access*, vol. 8, pp. 179118-179127.

[15] Sara Ashry Mohammed, Reda Elbasiony Walid Goma (2018) 'An LSTM-based Descriptor for Human Activities Recognition using IMU Sensors consists of four types of 3D sensory signals: acceleration, angular velocity, rotation displacement, and gravity for 31 activities, ADL measured by a wearable smart watch.' *Conference: 15th International Conference on Informatics in Control, Automation and Robotics*, DOI:10.5220/0006902405040511.

[16] Susanna Pirttikangas, Kaori Fujinami, and Tatsuo Nakajima (2006) "A Wearable sensor used for the feature selection and activity detection," in *Ubiquitous Computing Systems*, pp. 516–527.

[17] Thien Huynh ,Cam-Hao Hua, Nguyen Anh Tu, Dong-Seong Kim(2021)'Physical Activity Recognition With Statistical-Deep Fusion Model Using Multiple Sensory Data for Smart Health The technique shows the superiority of the suggested intermediate fusion model in multisensor systems and outperforms several cutting-edge HAR approaches'..Published in: IEEE Internet of Things Journal (Volume: 8),pp 1533 - 1543.

[18] Yixiao Yun Keren Fu,Irene Yu-Hua Gu, Hamid Aghajan(2014)'Human Activity Recognition in Images using SVMs and Geodesics classification while concentrating on human-object interaction for feature representation of actions on Riemannian manifolds'.8th ACM/IEEE International Conference on Distributed Smart Cameras, ICDS-C 2014; Venezia; Italy; p. Art. no. a20-978-145032925-5 (ISBN).

[19]ZhenguoShi ,J.AndrewZhang,QingqingCheng,Richard Yi Da Xu(2022)' Environment-Robust WiFi-Based Human Activity Recognition Using Enhanced CSI and Deep Learning is developed to remedy classification errors'. IEEE Internet of Things Journal (Volume: 9),pp.24643 - 24654.

[20] Zhiqing Zhou Heng Yu,Hesheng Shi(2020)'Human Activity Recognition Based on Improved Bayesian Convolution Network Analysis of Healthcare Data The Bayesian network uses offloading strategy and an EDL design to handle the security concerns,IEEE Access (Volume: 8) ,pp.86411 - 86418.