

Human Activity Recognition by Edge Computing Based Convolutional Neural Network

Usha S Jange*, Rajeshwari B**

¹Department of ECE, PES University, Bangalore

*usha.jange1998@gmail.com, **rajeshwari@pes.edu

Abstract. Human activity recognition has become a hotly debated research topic due to its numerous important and futuristic applications, such as automated surveillance, automated vehicles, language interpretation, and human-computer interfaces. (HCI). Extensive and in-depth research has been conducted, and good progress is being achieved in this area of research. The proposed system is designed by taking health monitoring as an application and implemented on a Raspberry Pi board, which can be used for monitoring and surveillance purposes. The activities considered in our proposed work are standing, jogging, climbing upstairs, and climbing downstairs. The surveillance camera will watch people's movements to see if they are performing the assigned job. We propose a CNN-based technique to recognize different actions. An alert will sound and the system will report that the person's designated task has been disrupted if the camera determines that they are engaged in another activity.

Keywords- Video Surveillance, Raspberry Pi, Raspberry Pi Camera, Open CV

1 Introduction

Human activity recognition (HAR) is in extremely high demand right now due to the quick development of computer vision capabilities. Daily life activities including standing, jogging, and climbing up and down stairs are applications that can benefit from HAR. As a result, numerous ways were used for recognizing a variety of activities. One of the crucial processes in HAR is feature extraction (using a raspberry pi board), which may collect pertinent data to distinguish between diverse actions. Features taken from raw photos have a major impact on how well human activity recognition methods work. It takes domain expertise to create hand-crafted features for a particular application. It has been extensively explored in other research areas, such as picture recognition, where various sorts of features must be retrieved while attempting to recognize various human activities.

A few years ago, surveillance cameras were manually examined and monitored every day. Technology has advanced, though, and systems are now created so that no individuals are required to remain still for extended periods of time in order to check security cameras.

Since the development of video cameras and recorders, video surveillance has emerged as the most trustworthy method of monitoring individuals and events taking place in specified locations. The internet and wireless connections have made it possible to view surveillance from anywhere in the world. [1]

The remaining paper is sketched as below: Section II deals with the related works already done in this field of research. Section III explains the methodology adopted and further section explains the Software processing and working of RPi

2 Related work

With the rapid advancement of artificial intelligence, CNN-based human activity detection techniques, as well as machine learning techniques, becomes more popular. Ha and Choi proposed CNN algorithms based on recognizing human activities [2] for handling multivariate time series information collected at many diverse locales. The lower and top layers of the CNN model, respectively, used partial and complete weight sharing and full weight sharing.

They demonstrated that the CNNs handled multi- modal data more effectively than existing CNNs with 2D kernels and achieved a high performance of 91.94% on the benchmark. Mhealth dataset, which consists of approximately twelve daily activities captured by four different types of sensors.

Machine learning is often used for medical purposes to perform activities utilizing portable information from systems [3].

The authors of [4] discuss that sensor-enabled devices as accessories are altering the tracking of health. These gadgets have the potential to capture an enormous amount of analysis information regarding individuals, and machine training is anticipated to have a crucial factor in interpreting this fresh sort of knowledge. Still, because of technology constraints, using machine learning in medical sensors is equally difficult.

Lee, Sang, and Cho used a 1D CNN-based technique to recognize three human activities using information from triaxial accelerometers obtained through phones used by consumers [5]. To decrease the possibility of rotational interference, after converting 3-axis accelerated information to vector amplitude data, the network they developed obtained 92.71% activity detection precision, which was superior to the starting point of the random forest technique. Furthermore, they discovered that increasing the input vector dimension boosted activity detection performance.

3 Methodology

In this project, we are aiming to identify Human activity to achieve the desired objective of aiding professionals in the medical field. It is essential to make the machine understand what activity to be identified to complete the objective. Hence the machine is trained to identify the activities using the data set consisting of the images specific to the activities to be identified.

The size of the data set taken for the project is 712 i.e., there are 712 images. Out of which, the data split into different proportions for the training and testing purpose. So, 75% of the data set is used for training and 25% of the data set is used for testing, i.e., 75% of 712 is 534 images for training the model and 25% of 712 is 178 images for testing the model.

Table 1. Size of the dataset Training the model using the training dataset.

Train	Test	Total
583	178	712

A. Dataset

The daily logging of physical human activities makes up the dataset. The dataset has four classes. They are climbing stairs, jog, standing, and climbing upstairs. Before the training started, the data set is resized from the initial size of 600x450 to 224x224, which aims to speed up the computational process. The dataset is divided into 75% data train and 25% data test. For the data train is taken 10% of validation data is used to determine whether the required model can classify images that have not been seen before in training.

B. Data Augmentation

It is an image manipulation technique without lessening any data information. Augmentation can improve the accuracy of the CNN model since the model will get additional information for generalizing the data [10].

Table II displays the data augmentation parameter for this investigation.

Table 2. Table of data augmentation

Type	Factor
Rotation Range	30
Zoom Range	0.15
Width Shift Range	0.2
Height Shift Range	0.2
Shear Range	0.15
Horizontal Flip	True

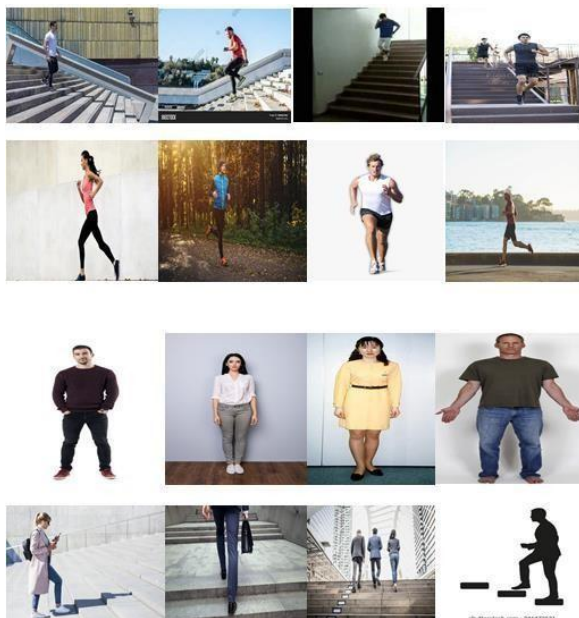


Fig 1. Dataset

C. CNN Method Training

This model uses a technique that can identify different types of Human activities from a digital form of the training data i.e., photos.

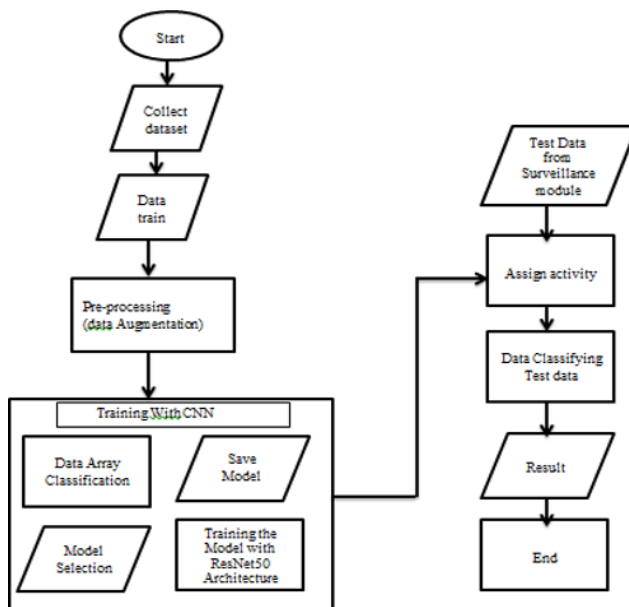


Fig 2. Flowchart of classification of system

The primitive stage in the training is to initialize the labels which have to be trained and later identified. Those labels are Downstairs, Upstairs, standing, and jogging. As mentioned in the previous section the RESNET-50 model will be trained using one custom training dataset of the predefined size. Initially, the images are resized from the regular size into the standard 224*224 pixels and later they are converted into RGB CHANNEL. Post image-resizing, the input data set is converted into a series of arrays called NumPy arrays.

Train the data segmentation object to rotate, zoom, and identify the height of the image, Width of the image. This process of making the machine familiar with the data set is called an epoch. Such many epochs are run to achieve greater accuracy yielding a confusion matrix. And as the number of epochs increases the resultant confusion matrix containing Recall, F1-score gets closer to greater accuracy. The Confusion matrix talks about the predicted class and actual class of the activities.

Testing: In the testing part we used healthcare as an application and implemented a number of investigations that have taken into account the medical advantages of frequent exercise monitoring and recognition. There is convincing evidence that regular physical activity monitoring and recognition can help control and lower the risk of numerous diseases, including diabetes, cardiovascular disease, and obesity. Several studies have been conducted to create efficient human activity

recognition systems. In this project, testing is implemented using Raspberry Pi 4.

In this project, we used a surveillance system where it picks up the movements. To implement the surveillance system, we used a Raspberry Pi 4 and a Raspberry Pi camera to monitor, record, and stream live-feed video. We have used Python as the programming language. When the motion is detected, the system records the image and detects the activity assigned to them. If the person is doing an activity other than the assigned activity then there will be a beep sound.

4 Software simulation results

To evaluate the performance of the optimized CCN model presented above, some tests have been carried out. The parameters used to compare performance are Accuracy, Precision, Recall and F Measure.

1. Accuracy =
$$\frac{[TP + TN]}{[TP + TN + FP + FN]}$$
2. Precision = $[TP/TP] + FP$
3. Recall = $[TP/TP] + FN$
4. F Measure =
$$2x - \frac{[Precision \times Recall]}{[Precision + Recall]}$$

They are defined in the below mentioned formulae

Where TP are the True Positives, TN the True Negatives, FP the False Positives & FN the False Negatives.

For each dataset accuracy has been evaluated to assess the global CNN performance (i.e., considering the whole set of classes), while F- Measure, Precision and Recall have been computed to give a more precise indication on how the CNN behave in recognizing a particular class.

The original dataset was initially split into two halves of with the ratio of 75:25 for training and testing respectively. Among this, 25% of the data set is being used here for testing the model to assess ability to identify the expected activities.

The images are assigned individually while compilation to test the efficacy of the trained model. The confusion matrix related to this test is reported in Table 3, together with Recall (RCL) Precisions (PRC) and F-Measure (FM) per class.

Table 3. Confusion Matrix for Proposed CNN model (Activity Names on Top are Abbreviated)

CNN Model	D	J	S	U	Recall%
Downstairs	13	2	1	9	52.0%
Jogging	0	24	7	0	77%
Standing	0	0	71	0	100%
Upstairs	8	1	5	37	70.0%
Accuracy	57%	83%	92%	76%	81%



Fig 3. Human activity recognition

The system built till the software simulation, is used to achieve a specific objective of identifying patients' activity in the medical field. A surveillance system with pre-decided objective is built for the patient's activity monitoring, where it picks up the movements. This is referred as health care surveillance system. To implement the surveillance system, we used a Raspberry Pi 4 and a Raspberry Pi camera to monitor, record, and stream live feed video.

The system is rebuilt with small tweaks to receive video feed as an input for the surveillance system using Python only on which the machine learning system was built. The video is feed is used as an input for the system, which identifies the kind of motion when detected. Once the motion is detected, the system records the image and detects the activity assigned to them. If the person is found doing any activity other than assigned activity then there will be beep sound.

Below diagram talks about the hardware setup using Raspberry pi4.

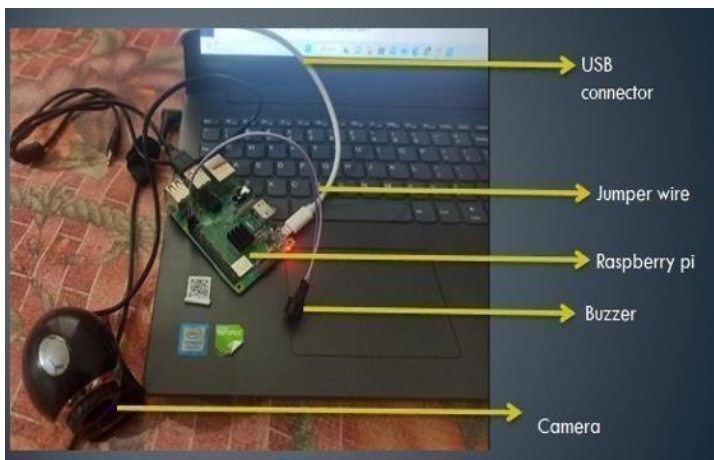


Fig 4. Hardware Setup

The below Figure depicts the front End of Human Activity Recognition. This activity is assigned to a person to perform that particular activity.

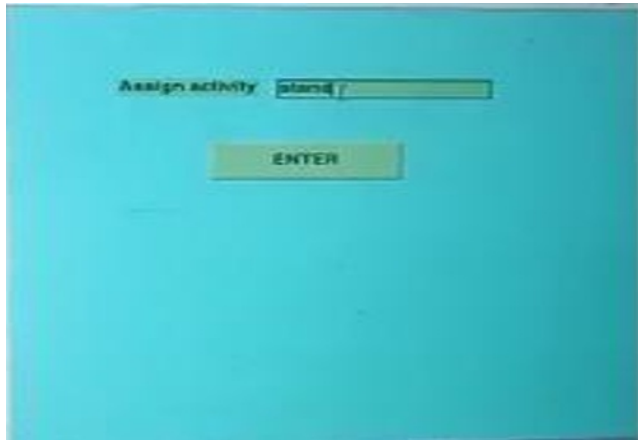


Fig 5. Front-end Assigned

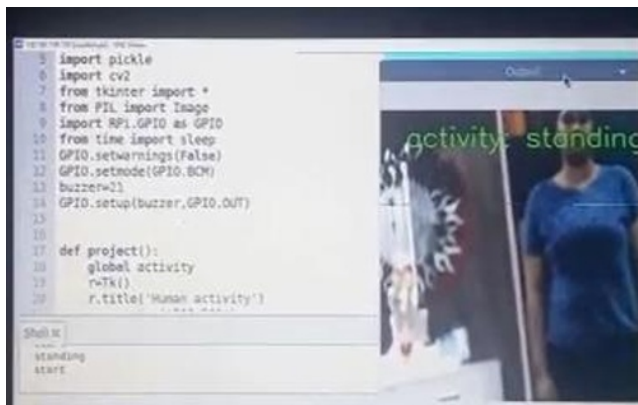


Fig 6. Output

5 Conclusion

In this proposed work, Human activity recognition using edge computing has been proposed. Our proposed model has presented a deep convolutional neural network model for the classification of four activities. The model is applied with ResNet 50 with an accuracy of 81%. The proposed model can be used in the healthcare sector using the Raspberry Pi4 model. Here the person is assigned to do particular work for physical activity which can be monitored without any human intervention, implemented using a surveillance camera. Future enhancements can be done to this model for better accuracy in prediction.

References

1. C. Yoon, K. Song, Muhammad, Zubair. IEEE International Conference on Consumer Electronics-Asia IEEE, *Human activity recognition using wearable accelerometer sensors*, 1-5 (2016).
2. Hua Cong, Yang. International Computer Conference on Wavelet Active Media Technology and Information Processing IEEE, *HARLib: A human activity recognition library on Android*. 313-315 (2015)
3. Juha Röning, Siirtola, Pekka. International Journal *Recognizing Human Activities User independently on Smartphones Based on Accelerometer Data*
4. S. Choi, Ha, Sojeong International Joint Conference on Neural Networks IEEE *Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors* 381-388 (2016)
5. Jian Bo, Yang. International Conference on Artificial Intelligence AAAI Press, *Deep convolutional neural networks on multichannel time series for human activity recognition*, 3995-4001 (2015)
6. H. Cho, M. Y. Sang, Song Mi, Lee. IEEE International Conference on Big Data and Smart Computing IEEE, *Human activity recognition from accelerometer data using Convolutional Neural Network*, 131-134 (2017)
7. Sheng, Min. ACM SIGGRAPH Conference on Virtual-Reality Continuum and ITS Applications in Industry ACM, 131-134 (2017)
8. Nitish, Srivastava. Journal of Machine Learning Research, *Dropout: a simple way to prevent neural networks from overfitting*, **15** 1 1929-1958 (2014)
9. S. A. Moore, G. M. Weiss, Jennifer R, Kwapisz. Acm Sigkdd Explorations Newsletter, *Activity recognition using cell phone accelerometers*. **12** 2 74-82 (2011)
10. C. J. Lin, Chih Chung, Chang. *LIBSVM: A library for support vector machines*. ACM, (2011)
11. D. D. Phan, Duc Ngoc, Tran. International Conference on Intelligent Systems, Modelling and Simulation IEEE, *Human Activities Recognition in Android Smartphone Using Support Vector Machine* 64-68 (2017)