Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 5 (2023)

DEEP LEARNING FRAMEWORK FOR HUMAN ACTIVITY DETECTION BASED ON WIRELESS SENSORS.

Dr. G. ANANDHARAJ[1] **G.Bharathi[2]**

1. Assistant Professor, Department of Comput.er Science, Sri Vasavi College(Govt. Aided), Erode – 638316. India.

2. Assistant Professor, Department of IT & DCFS, Nehru College of Arts and Science, Coimbatore – 641105, India.

Abstract:- Recognizing human athletic behaviors has grown in importance as a component of analysis during the last few years. For upcoming training, FPGA-based recognition has been built to perfectly predict human actions. Intelligent environments where human athletic movements may be automatically recognized are required to ensure the viability of such long-term athletic training monitoring systems. The system discussed here employs an Artificial Neural Network (ANN) based Field-Programmable Gate Array (FPGA). In order to infer human behavior from the data collected by a smart environment, machine learning algorithms must first be trained on annotated datasets. To effectively extract characteristics from unprocessed data and develop a machine learning model for anticipating an individual's movement, conventional signal processing techniques and domain knowledge are required. The purpose of this work is to show how a hybrid deep learning model may be applied to recognize human behavior. Convolutional neural networks and continuous neural networks, for example, are deep learning techniques that will extract the features and help with categorization. The proposed model forecasts human activity using wireless sensor data mining datasets. Accuracy, training loss, testing loss, the confusion matrix, and other metrics have all been used to evaluate the model's performance.

Keywords: Action Recognition, Feature extraction, Classification, Field Programmable Gate Array, Internet of things.

1. Introduction

Actions, gestures, and sports are among of the most important ways that people communicate with one another. People frequently use their hands and minds to tell stories, which is more like using exercise equipment. Recognition of human behavior is the most active area of computer vision[1]. Numerous human behavior are evaluated using pattern recognition, computer vision, machine learning, and automated gesture analysis[2]. A lot of knowledge exists regarding human motion recognition methods, including motion recognition for various applications developed using video data, motion capture, depth data, or a combination of these methods[3].

For earlier activity recognition research, conventional algorithms like SVM and Random Forest have been integrated with more contemporary deep learning techniques like ANN, CNN, and RNN [4]. Feature engineering and feature extraction with standard algorithms are labor-intensive and time-consuming. Deep learning techniques, however, are more appropriate for categorizing human behaviors since they can automatically learn characteristics from data [5].

Understanding human behavior is a crucial subject for computer vision research and application. Current behavior is examined using human behavioural awareness. In a straightforward example, there is only one execution of the action. People's ability to communicate is well acknowledged. The focus of this study is mostly on these two actions. One of the many behaviour recognition algorithms, the state-based model method, is a commonly used application that describes human action as a collection of state-construction models. The model is statistically trained to match them using a set of feature vectors from that activity class. Statistical models are frequently made for each action.

CSIs are geographical and temporal signals of channel state data. This study suggested the two-stream structure as a crucial new mining teaching method. Many sound sub-activity cuts are used to isolate the full movement test in specific [6]. This study provides the profundity history picture (DHI), which was created by subtracting the profundity from the activity outline. The weight of deep learning structures that have undergone pre-tuning preparation is then taken into account when applying DHIs to Alex Net. DHI alone is insufficient to recognize closely linked activities. This is done using separate Alex Net training and 3D convex extraction [7].

Researchers enter the sensor-based HAR equipped with sophisticated understanding of sensing technology to choose the best high-quality sensors for their research settings. Following that, preparation work is needed to obtain consistent, high-quality bio-signals[8]. For video camera-based HAR, for instance, the calibration process is crucial since accurate distortion calibration is required for flawless activity recognition via external sensing. To optimize the camera model, conventional methods sometimes require hundreds or thousands of photos. The reprojection error was improved by 28.5% over the polynomial distortion model by Jin et al.'s novel point-to-point distortion calibration approach, which uses just a small number of pictures to produce a dense distortion rectification map [9].

2. Objectives

Using sensor-based time-series data, human activity recognition (HAR) attempts to identify people's everyday activities. In the past ten years, advances have been made in sensors, the Internet of Things (IoT), cloud computing, and edge computing. HAR research has switched to sensor technologies since detectors are accessible and simple to integrate or implant in both portable and non-portable systems. It is possible to quickly record a range of movements of the body for individual activity detection using IoT wearable devices with sensors.

3. Methods

This section will detail the proposed work's methodology for recognising human activities using a Wireless Sensor Data Mining (WISDM) dataset. Using CNN, the features will first be retrieved, and then an LSTM model will be used to perform the classification.

Data Pre-processing:

There will be discussion of data preparation techniques. The process of feature extraction is done first, and then a deep learning model is used for classification.

Feature Extraction:

We can better understand the feature by using an example. The next step after taking a photo is to identify it; however, to do so, the user must store a large number of photos and measure each one in terms of pixels per inch, which takes up a lot of storage space. Therefore, in order to store this image, features extraction is required. A feature extraction will result in a reduction in the image's dimensions. An important concept in computer vision is referred to as a "feature". Improved recurrent neural networks (RNNs) are the cause of improvements in long-term short-term memory.

Convolutional Neural Network:

A particular kind of neural network called a convolutional neural network is made to analyze multidimensional data, such images. and data on time series (CNN). The training procedure includes the computation of weights and feature extraction. These networks are known as convolutional networks because they use a convolution operator. The capability of CNNs to automatically extract features is their primary benefit. Figure 3 demonstrates how input data is initially given to a feature extraction network, and the resulting features are then sent, as shown, to a classifier network.

The feature extraction network consists of pairs of convolutional and pooling layers. A layer of digital filters known as a convolutional layer is used to convolution the input data.

Instead of RNN units, LSTM memory blocks can solve the issue when the gradient vanishes or bursts. Its key distinction from RNNs is the addition of a cell state to store long-term states. An LSTM network can keep track of and connect data obtained in the past with data gathered in the present. An input gate, a "forget" gate, and an output gate make up the LSTM's three gates. The "forget" gate refers to the past input, whereas the input gate refers to the current input. Figure 1 depicts the LSTM's internal architecture.

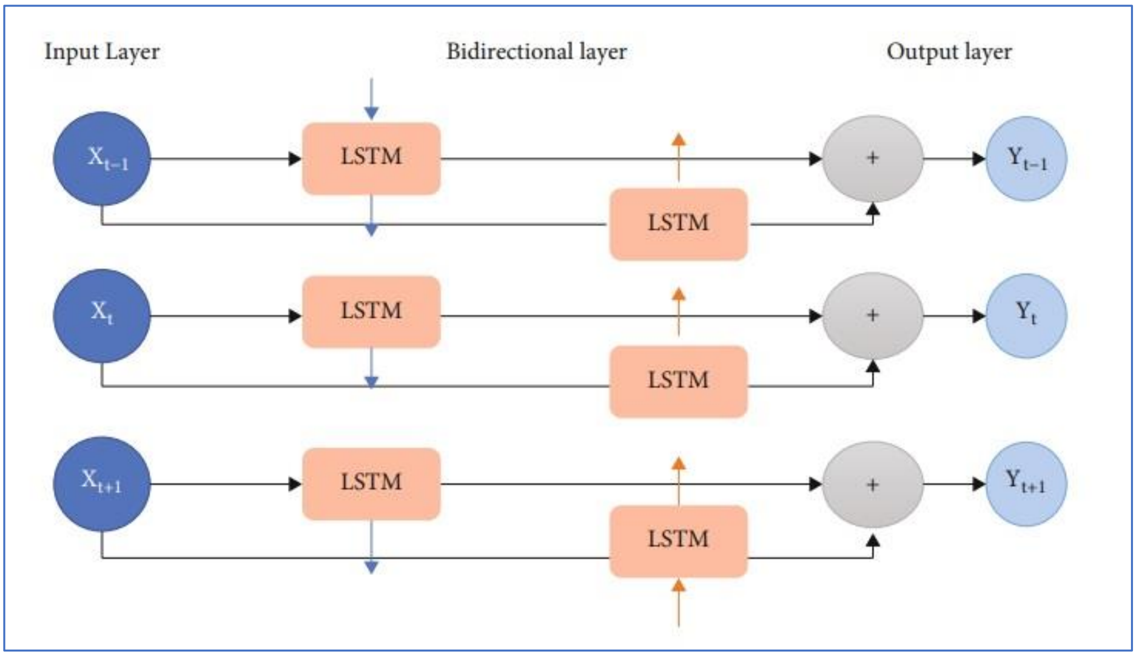

Figure 1: Architecture of Neural Network

There are 512 network units in total configured in this LSTM network layer. First, we use two blocks of closely linked layers, each of which is triggered by Rectified Linear Units (ReLUs) and has a size of 1024 network units. After the ReLU is activated, we use dropout regularization with a value of 0.8. Dropout regularization, which ignores randomly selected neurons, is employed during the training phase to prevent the complex co-adaptations of the fully linked layers. This stops over fitting from happening throughout the training period.

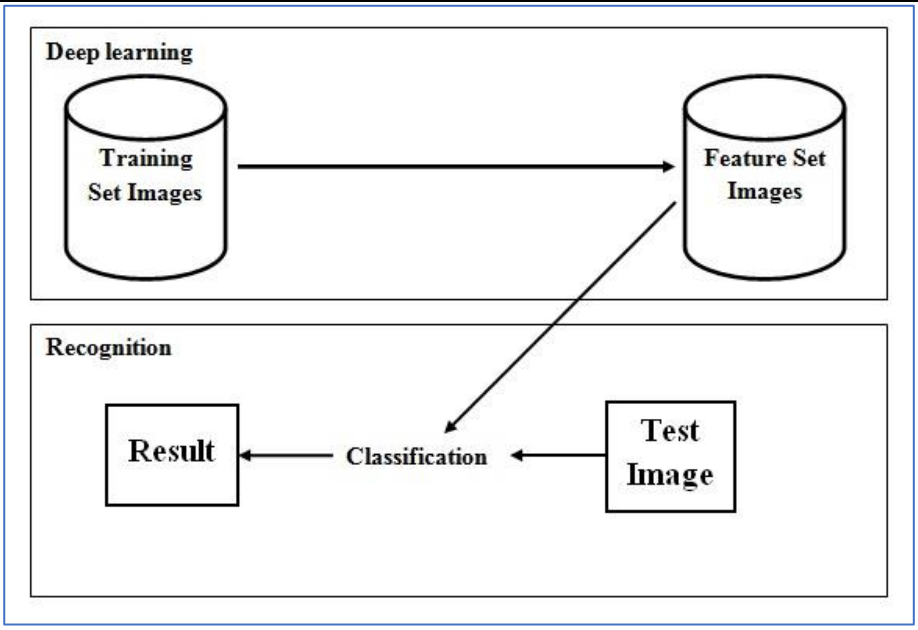

Figure 2: Recognition of Image.

The method of picture recognition using a machine learning approach is depicted in Figure 2. A sliding time window is utilised to split the data into distinct segments in order to supply the network with such temporal relationships. For increased accuracy, both the window width and the step size can be modified and tuned. The activity label that appears the most frequently is picked for each segment since each time step has an associated activity label. The time step is set to 100 and the time segment or window width is 200 in this example.

4. Results

To conduct the human activity recognition, the hybrid deep learning model has been applied to wireless sensor data mining datasets. The classifier's accuracy is 95%, while it could conceivably be marginally increased by reducing the sliding window's step size. The final confusion matrix is shown in the following graphs, which have been normalized such that each row adds to one.

5. Discussion

This method shows the train/test error/accuracy for each and every epoch. Python is the programming language used to carry out the suggested job. Confusion matrix, accuracy, and loss are the performance indicators used to assess the model. The model hyper-parameters are based on human actions including jogging, sitting, standing, going upstairs, and walking downstairs. These hyper-parameters are also used to build the confusion matrix. The confusion matrix's graph is shown in Figure 3.The hyper-parameters form the basis of the graph.

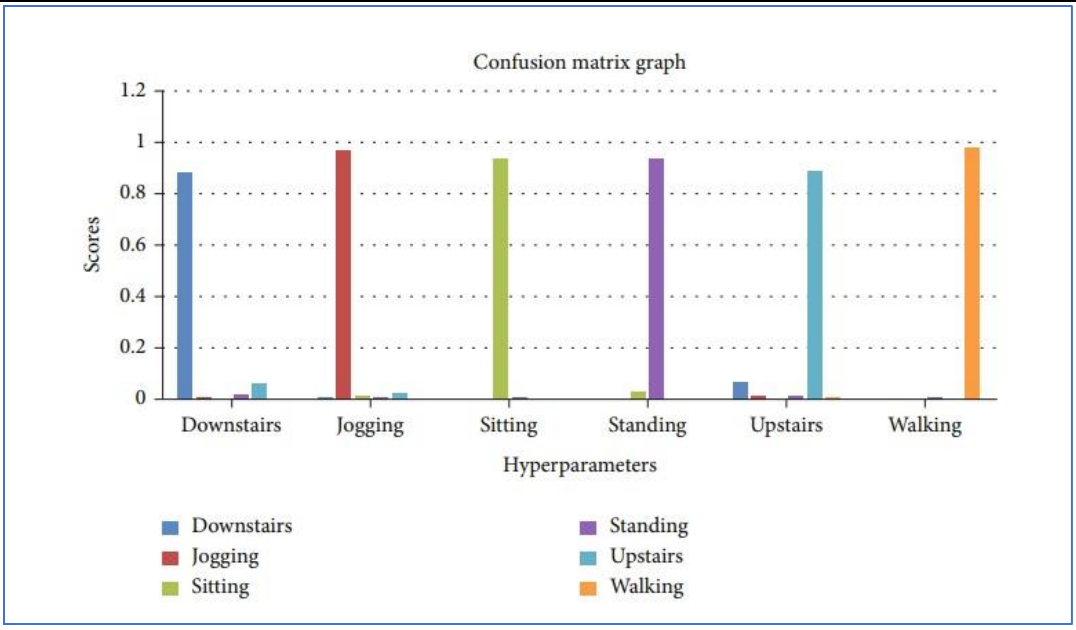

Figure 3: Confusion Matrix.

Recognizing our activities has beneficial consequences on our health and wellbeing, which makes it more and more in demand. In the current battle against overweight and taking care of the getting older, it is a crucial tool. The ability to identify activity among people is dependent on applying sensors to deduce human activity or activity from the body's movements and gestures. Many common tasks performed by humans can be automated or made simpler by activity recognition technologies. Human activity recognition systems may or may not be under supervision, depending on the situation. In this study, a hybrid deep learning model a convolutional neural network and a recurrent neural network—was used to recognize human activities. The confusion matrix, accuracy, and mistakes are the criteria used to assess the model. The proposed model has performed better than the current models, according to the results. Utilizing Capsule Network will enable the work to be improved in the future.

Refrences

[1] Q. Liu, L. Cheng, A. L. Jia, and C. Liu, "Deep Reinforcement learning for Communication flow control in wireless Mesh networks," IEEE Network, vol. 35, no. 2, pp. 112–119, 2021.

[2] H. Zhang, R. Nie, M. Lin, and M. Mokayef, "A deep learning based algorithm with multi-level feature extraction for automatic modulation recognition[J]," Wireless Networks, vol. 1, pp. 1–12, 2021.

[3] M. M. Hassan, S. Ullah, M. S. Hossain, and A. Alelaiwi, "An end-to-end deep learning model for human activity recognition from highly sparse body sensor data in Internet of Medical Things environment[J]," The Journal of Supercomputing, vol. 4, no. 4, 2020

[4] W. Cai, Y. Song, H. Duan, and Z. Xia, "Multi-feature fusion-Guided Multiscale Bidirectional Attention networks for Logistics Pallet segmentation[J]," Computer Modeling in Engineering and Sciences, vol. 131, 2022.

[5] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human activity recognition using inertial, physiological and environmental sensors: a comprehensive survey," IEEE Access, vol. 8, 2020.

[6] Murad, A., & Pyun, J. Y. (2017). Deep recurrent neural networks for human activity recognition. *Sensors*, *17*(11), 2556.

[7] Anguita, Davide, et al. "A public domain dataset for human activity recognition using smartphones." *Esann*. Vol. 3. 2013.

[8] Jin, Z.; Li, Z.; Gan, T.; Fu, Z.; Zhang, C.; He, Z.; Zhang, H.; Wang, P.; Liu, J.; Ye, X. A Novel Central Camera Calibration Method Recording Point-to-Point Distortion for Vision-Based Human Activity Recognition. *Sensors* *22*, 3524, 2022,.

[9] Ding, X.; Hu, C.; Xie, W.; Zhong, Y.; Yang, J.; Jiang, T. Device-Free Multi-Location Human Activity Recognition Using Deep Complex Network. *Sensors* 2022.