

A REVIEW ON HUMAN ACTIVITY DETECTION USING DIVIDE AND CONQUER 1DCNN

Sehreen Mir¹, Er Simranjit Kaur²

¹Pg Scholar, Department of Electronics and Communication Engineering, Sri Sai College of Engineering and Technology, Badhani, Pathankot, Srinagar, J&K, India

²Assistant Professor, Department of Electronics and Communication Engineering, Sri Sai College of Engineering and Technology, Badhani, Pathankot, Pathankot, Punjab, India

ABSTRACT

There are various studies in the literature for identifying various human activities. In this study, we offer the most recent research that shows how deep learning algorithms have improved overall at identifying or detecting the type of human activities. We give a quick overview of the issues and difficulties associated in recognizing human activity. We try to create a framework that would predict six categories of human behavior through a division-based rule approach. This article examines six human activities, including walking, ascending stairs, sitting, standing, and sleeping/laying. These procedures include feature extraction, classification, activity analysis, and recognition. The purpose of this publication is to give this field's scholars a literature review of six different human activity recognition along with its general structure.

Keywords: deep, human, standing, classification, recognizing, recognition

1. INTRODUCTION

Human Activity Recognition (HAR) is the study of recognizing human movements or activities using sensor data. The majority of movements take place inside the home and include standing; sitting, talking, and walking. The kitchen and the office are two more rooms in the house where they come in handy. Sensor data can now be remotely recorded thanks to technologies like radar, video, and other wireless recording devices. Additionally, information about the individual may be obtained through the use of medical devices or smartphones that have gyroscopes and accelerometers built in. The process of gathering sensor data for event identification has historically been challenging, expensive, and required specialist equipment. Human actions like "walking" and "running," come extremely effortlessly to us in daily life. There are two categories of activities in HAR: simple and complex. Relatively few research have been done on recognizing complicated human activities, such dribbling a ball, brushing teeth, etc. complicated human activities entail completing a simple human activity together with a specific transition action (Vrigkas, Nikou, & Kakadiaris, 2015).

HAR with labeled data is a multivariate time series classification and supervised learning challenge in the machine learning arena. The task of activity recognition has been studied in a number of earlier research utilizing both conventional deep learning approaches (ANN, CNN, RNN, LSTM, etc.) and non-traditional ones (SVM, Random Forest, XGBoost, etc.). The problem with standard algorithms is that they require a lot of labor-intensive human feature extraction and feature engineering (Weiss et al., 2019). The primary goal of this work is to develop a 1D CNN deep learning model that can reliably categorize six human activities into static and dynamic categories. These data are gathered using a tri-axial accelerometer and gyroscope found in smartwatches and smartphones, which are able to measure the angular velocity and acceleration of human movements.

2. METHODOLOGY

The study's research methodology includes the essential techniques including data collection and acquisition related to sensor-based human activity recognition, pre-processing of collected data, segmenting raw sensor data into appropriate length segments using a sliding window, dividing the dataset into train, validation, and test datasets, developing models using a variety of deep learning algorithms, adjusting hyper-parameters, and assessing model performance using a range of performance metrics. Eventually, these stages will result in the supervised classification and identification of human activities using sensor-collected data.

2.1 Data Collection

The UCI -HAR dataset contains already-collected, openly accessible data that was used in this investigation. This dataset was compiled from 30 individuals who were wearing smartphones around their waists while engaging in a variety of activities (referred to as subjects in this dataset). The accelerometer and gyroscope sensors of the smartphone are used to record the data. This experiment was videotaped so that the data could be manually labeled.

2.2 Data pre-processing and transformation

The raw dataset must first be loaded into memory. The raw data consists of three primary signals: overall acceleration, body acceleration, and body gyroscope, each with three data axes (x, y, and z). Consequently, each time step has a total of nine variables. Additionally, each data series has been divided into lapping windows of 2.56 seconds, or 128 times. As a result, each row of data contains 1152 (128 * 9) elements. For the class number, the affair data is defined as an integer. For the data to be adequate for constructing a neural network multi class bracket model, these affair indicators were first one hot decoded. We'll load the train and test datasets. Using a Standard Scalar object, the data is further scaled.

1.3 Model development

The gauged data generates an estimate of the model's performance after fitting the successional model to the training dataset and testing it against the test dataset. Additionally softened are the uprooted characteristics. A powerful sub caste is then included to help prevent the model from being over fit to the training set of data. The features uprooted by the LSTM concealed subcaste are eventually interpreted by a thick fully linked subcaste before prognostications are made using a final affair subcaste. Since we require six issues as the outcome, Softmax activation function is employed in the final subcaste. . The effective Adam interpretation of stochastic grade descent is used to optimize the network, and the categorical cross entropy loss function is used to calculate the loss in the training process. Once the model is fit, it is estimated on the test dataset and the accuracy of the fit model on the test dataset is returned

1.4 Proposed model

In this work, different human activities have been categorized using deep learning-based feature extraction techniques. Two stages are involved in the learning of a one-dimensional (1D) convolutional neural network (CNN) model and classifier: the first stage involves training a binary 1D CNN model for abstract task recognition to distinguish between dynamic and static tasks; the second stage involves training two 3D-like 1D CNN models for event classification.

3. LITERATURE REVIEW

Smartphone human activity tracking apps have made it simpler for people to keep track of their daily routines. Numerous attempts have been made to assess the activities, but the outcomes have been unsatisfactory due to data inconsistencies or flaws in the various products. This study suggests a method for achieving peak performance. One category that might be used is logistic regression. The raw data for this analysis comes from smartphones, smartwatches, and other gadgets with in-built sensors like accelerometers and gyroscopes. A size reduction technique must be utilized since accelerometer sensors function in three dimensions. Principal component analysis is helpful because it enables us to extract from raw data the key characteristics that may be used to classify human behaviors. (Zaki, 2020) using logistic techniques obtained an accuracy of 96.19 percent for the UCI-HAR dataset and 94.5 percent for the HAPT dataset.

(Ahmed et al., 2020) study has suggested a filter-and-wrapper-method-based hybrid method feature selection approach. For enhanced activity recognition, the technique employs a sequential floating forward search (SFFS) to extract the desired features. The kernel method is then used to generate nonlinear classifiers by feeding features into a multiclass support vector machine (SVM) for training and testing. With the use of a benchmark dataset, we verified our model. Our suggested approach effectively identifies activities while utilizing minimal hardware resources.

(K. Chen et al., 2021) gave a survey of cutting-edge deep learning techniques for identifying human activity based on sensor data. The author first discussed the multi-modality of sensory data, offered details on open-access datasets that can be used to assess performance on various challenge activities, and then, provided a brand-new taxonomy to organize the deep methods according to problems.

There are a growing number of practical mobile sensing applications available. In order to better understand human behavior, these programs recognize human behaviors using mobile sensors built into smart phone.(Y. Chen et al., 2016) proposed a tri-axial accelerometer databased feature extraction method based on LSTM to identify human activities. His LSTM-based technique is workable and achieves 92.1% accuracy, according to the experimental results on the (WISDM) Lab public datasets.

In order to develop thorough classification algorithms for HAR employing wearable sensor data, a few Hybrid Learning Algorithms (HLA) were proposed by (Athota & Sumathi, 2022) .Utilizing the Convolution Memory Fusion Algorithm (CMFA) and Convolution Gated Fusion Algorithm (CGFA), which model learns local features as well as long-term and gated-term dependencies in sequential data, is the goal of this work. The use of different filter sizes has improved feature extraction. They are utilized to record various local temporal dependencies, and the improvement is thus put into practice.

In order to improve activity detection performance, the study done by (Mohd Noor et al., 2022) suggests a unique deep temporal Conv-LSTM architecture that takes advantage of both the interaction between sliding windows and temporal

properties from sensor data. The Smartphone-Based Recognition of Human Activities and Postural Transitions dataset, which includes a collection of transition activities, is used to assess the suggested architecture. The capacity to describe the temporal relationship of the activity windows, where the transition of activities is accurately captured, has been established for the proposed hybrid architecture with parallel features learning pipelines. In addition, the size of sliding windows is investigated, and it has been demonstrated that the choice of window size has an impact on the precision of activity recognition.

Deep learning methods, which effectively extract characteristics from the unprocessed data gathered by sensors like gyroscopes and accelerometers, are heavily utilized in the detection of human movement. The recommended model outperforms cutting-edge machine learning methods like Support Vector Machine (SVM) and k-Nearest Neighbors (KNN), according to experiments conducted and reported by (Murad & Pyun, 2017). By employing higher learning rates, which batch normalization makes possible, we can train networks more quickly.

Khushboo Banjarey (Banjarey et al., 2021) Due to the rapid advancements in artificial intelligence (AI), humans are more focused on new research goals to recognize things, comprehend the environment, evaluate time series, and predict outcome patterns. A domain called "human activity recognition" (HAR) focuses on identifying, analyzing, and rating human movement behavior. Deep learning has substantially benefited HAR. Deep learning models encounter tough obstacles despite their enormous promise, such the requirement for a sizable dataset for real-world training. To differentiate between static and dynamic behavior, however, the current study has to be enhanced by results that are more outstanding. The major objective is to develop a system that can recognize movements like sitting, standing, walking, and other similar ones using a one-dimensional convolutional neural network (1D CNN).

Human activity recognition (HAR), which extracts action-level information about human behavior from raw input data, improves people's lives. Human Activity Recognition has a wide range of applications, such as geriatric surveillance systems, anomalous behavior, and more. Convolutional neural networks and other deep learning models, however, have outperformed more conventional machine learning techniques. (Kumari et al., 2022) suggested that CNN makes it possible to extract features and cut back on computation costs. Leveraging CNN, a particular kind of artificial neural network, is used in transfer learning to describe the usage of pre-trained machine learning models that can be utilized to identify human behavior.

(Nafea et al., 2022) suggest combining gated recurrent uni with hierarchical multi-resolution convolutional neural networks. The authors tested the suggested model using the mHealth and UCI data sets, and the findings show that it is effective because it reached respectable accuracy levels of 99.35% in the mHealth data set and 94.50% in the UCI data (Chakraborty & Mukherjee, 2023) explored the unusual challenge of separating recurrent leg-swing activities while sitting from repetitive leg-swing activities whilst walking. A heterogeneous sensor system is put into place to achieve this goal, and it collects innovative multi-modal data from inexpensive leg-worn IMU sensors (m-module) and finger-tip-based pulse sensors (p-module). The features for a supervised learning framework are then extracted from this dataset through processing. The system performance is assessed using a 1-D Deep Convolutional Neural Network (DCNN) model, which has a maximum accuracy of 99% and an average accuracy of 97% in classifying walking activity.

.(Zebin et al., 2018) suggested a deep recurrent neural network (RNN) that uses a short-term model (LSTM) to recognize six commonplace behaviors using sensor data. Additionally, the batch normalization method enables greater precision with fewer epochs.

(Cho & Yoon, 2018) proposed a one-dimensional convolutional neural network (CNN) that employs a divide-and-conquer method based on categorization learning and test sharpening materials for the recognition of human activity. Several 1D CNN models are trained at two different levels using this method. In the first step, we create a binary classifier that can categorize actions into static and dynamic groups. Then, two distinct 1DCNN models for task recognition are employed. The authors advise information assessment in prediction to increase the accuracy of thought processes.

4. CONCLUSION

We thoroughly examined state-of-the-art methods for identifying human activities in this review. Using a confusion matrix to identify abstract activities, we created a two-stage HAR procedure. A stronger HAR resulted from our "divide and conquer" technique with 1D CNN. Once abstract tasks appropriate for the first step can be identified, our strategy is straightforward and simple to put into practice. We get to the conclusion that the accuracy of static activities on test data may have been improved while employing CNN architecture for a six-class classifier. Therefore, we may divide a difficult classifier into two steps by using a divide and conquer strategy. At one time, the class activities were divided into static and dynamic sections, and a model could anticipate things with a high degree of accuracy.

5. REFERENCES

- [1] Ahmed, N., Rafiq, J. I., & Islam, M. R. (2020). Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model. *Sensors (Switzerland)*, 20(1). <https://doi.org/10.3390/s20010317>
- [2] Athota, R. K., & Sumathi, D. (2022). Human activity recognition based on hybrid learning algorithm for wearable sensor data. *Measurement: Sensors*, 24. <https://doi.org/10.1016/j.measen.2022.100512>
- [3] Banjarey, K., Prakash Sahu, S., & Kumar Dewangan, D. (2021). A Survey on Human Activity Recognition using Sensors and Deep Learning Methods. *Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*. <https://doi.org/10.1109/ICCMC51019.2021.9418255>
- [4] Chakraborty, A., & Mukherjee, N. (2023). A deep-CNN based low-cost, multi-modal sensing system for efficient walking activity identification. *Multimedia Tools and Applications*, 82(11). <https://doi.org/10.1007/s11042-022-13990-x>
- [5] Chen, K., Zhang, D., Yao, L., Guo, B., Yu, Z., & Liu, Y. (2021). Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities. In *ACM Computing Surveys (Vol. 54, Issue 4)*. <https://doi.org/10.1145/3447744>
- [6] Chen, Y., Zhong, K., Zhang, J., Sun, Q., & Zhao, X. (2016). LSTM Networks for Mobile Human Activity Recognition. <https://doi.org/10.2991/icaite-16.2016.13>
- [7] Cho, H., & Yoon, S. (2018). Divide and Conquer-Based 1D CNN Human Activity Recognition Using Test Data Sharpening. *Sensors*, 18(4), 1055. <https://doi.org/10.3390/s18041055>
- [8] Kumari, Dr. N. V., Talha, M., Vinod, C. V., Kamitkar, M., & Pattar, V. (2022). Human Activity Recognition using CNN and Pretrained Machine Learning Models. *Emperor Journal of Applied Scientific Research*, 04(07). <https://doi.org/10.35338/ejasr.2022.4702>
- [9] Mohd Noor, M. H., Tan, S. Y., & Ab Wahab, M. N. (2022). Deep Temporal Conv-LSTM for Activity Recognition. *Neural Processing Letters*, 54(5). <https://doi.org/10.1007/s11063-022-10799-5>
- [10] Murad, A., & Pyun, J. Y. (2017). Deep recurrent neural networks for human activity recognition. *Sensors (Switzerland)*, 17(11). <https://doi.org/10.3390/s17112556>
- [11] Nafea, O., Abdul, W., & Muhammad, G. (2022). Multi-sensor human activity recognition using CNN and GRU. *International Journal of Multimedia Information Retrieval*, 11(2). <https://doi.org/10.1007/s13735-022-00234-9>
- [12] Zaki, Z. (2020). LOGISTIC REGRESSION BASED HUMAN ACTIVITIES RECOGNITION. *JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES*, 15(4). <https://doi.org/10.26782/jmcms.2020.04.00018>
- [13] Zebin, T., Sperrin, M., Peek, N., & Casson, A. J. (2018). Human activity recognition from inertial sensor time-series using batch normalized deep LSTM recurrent networks. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 1–4. <https://doi.org/10.1109/EMBC.2018.8513115>