Available online at: http://www.ijmtst.com/vol5issue07.html

International Journal for Modern Trends in Science and Technology

ISSN: 2455-3778 :: Volume: 05, Issue No: 07, July 2019

Activity Recognition Using Convolution Neural Networks: Exploring Various Features

P.P. Joshi [1] | Dr. P.B. Tamsekar[2] | Dr. P.R. Patil[3]

[1]Assistant Professor, Department of CSE, SSBES ITM College Nanded, pranavjoshi13@gmail.com
[2]Assistant Professor, Department of CSE, SSBES ITM College Nanded, pravin.tamsekar@gmail.com
[3]Assistant Professor, Department of CSE, SSBES ITM College Nanded, pritam.itm@gmail.com

To Cite this Article

P.P. Joshi, Dr. P.B. Tamsekar and Dr. P.R. Patil1, "Activity Recognition Using Convolution Neural Networks: Exploring Various Features", *International Journal for Modern Trends in Science and Technology*, Vol. 05, Issue 07, July 2019, pp.-79-82.

Article Info

Received on 15-June-2019, Revised on 22-July-2019, Accepted on 25-July-2019, Published on 28-July-2019

ABSTRACT

Healthcare, surveillance, and smart settings are just a few of the many fields that can benefit from activity recognition's ability to better understand human behavior. When it comes to computer vision tasks like activity detection, convolution neural networks (CNNs) have excelled. In this piece, we discuss the benefits and drawbacks of using different characteristics with CNN models for activity recognition. Here, we talk about the many kinds of features and the methods used to extract them from images and videos. We also look into the criteria used to judge the success of these models. Insights into utilizing CNN models with varying features for precise activity recognition are provided through a thorough analysis in this article.

Keywords:CNN,LSTM,Activity detection, Image based and video based features

I. INTRODUCTION

Using sensor data, like pictures or videos, activity recognition is used to identify and categorize human actions. Due to its superiority in learning hierarchical representations from visual data, CNN models have become widely used for activity recognition. The effectiveness of these models is, however, heavily dependent on the features that are used. Several methods exist for gathering HAR information from study participants; these methods can be roughly classified as either camera-based or sensor-based [3]. The former method involves setting up one or more video cameras to record a subject's behaviors over a period of time and then utilizing video analysis and processing techniques to identify the person. In the latter, different sensors are used to monitor the subject's motion. This method can be broken down even further according to whether internal or external sensors are employed [2]. Wearable sensors must be attached to the person at all times while collecting data, whereas external sensors are put at specific locations on the body. These methods all have their perks and drawbacks and specific uses. For even more data and to help machines better understand the activities they correspond to, some recognition methods integrate numerous recording techniques. Intelligent surveillance systems, haptics, HCI, motion or gesture-controlled devices, automatic health-care monitoring systems, prostheses, and robots are all examples of how HAR can be put to use. The articulated character of human activities, the participation of external objects in human relationships and the intricate spatiotemporal structures of the action signals [4] make HAR a difficult endeavor despite recent advances. To successfully identify these actions, high-level signal and image processing techniques and ML algorithms are required. Since perfect performance has not yet been attained, HAR is still a developing area of study. HAR is a real-world discipline that draws on Biomedical Engineering and Computer Science expertise. The realism of the task necessitates that the machines be trained using data from actual human behavior. Human motion data, in the form of analog signals, is available from a variety of datasets provided by universities and research labs throughout the world. The UCI HAR dataset [5] is one of the most well-known HAR datasets and is maintained by the University of California, Irvine's (UCI) Machine Learning Repository. The authors of the pioneering work on this human activity dataset (reference [6]) used a hardware-friendly version of the Support Vector Machine (HF-SVM) to achieve a classification accuracy of 89.35 percent. The dataset was made public in December 2012, although the study was already out when that happened. Mapping 561 statistical features from the raw data led the same group of scientists to report a 7% increase in classification accuracy in 2013 [7]. The data gathering technique, in-depth system design, data specification, and data processing details are all laid forth in the article. Another paper by the same authors, titled "Energy Efficiency of the Model," was published the same year [8], and it extracted a new set of features and classified them using several HF-SVMs with varying LUTs. As of early 2019, many researchers' efforts on the UCI HAR dataset have yielded numerous distinct approaches using a wide variety of feature extraction, feature selection, and ML techniques. Recent papers that have used the UCI HAR dataset include [9], which describes the authors' exploration of different neural network architectures for HAR signal classification, and [10], which details a strategy based on Semi-Supervised Active Learning (SSAL). The maximum accuracy is found in [11] and [12] among the research that used the cited dataset. In this paper, we will detail a classification model for human activities using a multilayer Convolution Neural Network (CNN). We fed a multichannel CNN model frequency and power information collected from the signals rather than the statistical data provided in [7] and [8]. Before applying the final classification, the outputs were combined. After each stage of the process, we've included the necessary figures, flowcharts, and tables to clarify the explanation and back up the methodology:

II. METHODOLOGY

The purpose of this research is to apply a two-channel CNN model to the UCI HAR dataset in order to categorize the HAR signals, as depicted in Fig. 1. This approach, like other supervised machine learning methods, consists of a training and evaluation phase. Data samples measuring a variety of attributes from subjects while they carry out a variety of predetermined activities are required for the training phase. The supervised learning method then attempts to "make sense" of the data by discovering the similarities and differences between samples of the same class and those of different classes, and by creating an internal model or models that emphasize the most important characteristics that can bring out these contrasts in order to complete the classification [2]. However, it may not be wise to simply feed the raw data collected from the sensors into the classifier. This is because time-domain signals often contain noise, interference, missing values, and, most importantly, time-domain attributes are simply not good enough to make the distinguishable properties perceptible to the classifiers. This is why featuring engineering [13] (the process of extracting useful characteristics from large amounts of data) is so popular amongst scholars. While papers [7] and [8] used similar statistical aspects to achieve satisfactory findings, we take a slightly different tack in the present investigation. We are using a two-channel convolution neural network (CNN) to process the raw time-domain accelerometer signals by first extracting frequency and power information (or features). The machine is taught on a subset of the dataset to create a workable model, which is then tested on the entire dataset.

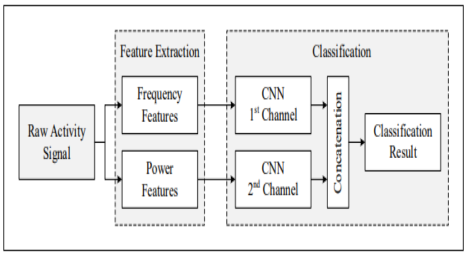

Figure 1: System Architecture

. **Image-Based Features:** Image-based features help identify activities by extracting static data from individual frames. We go over the following characteristics:

• Raw Pixel Intensities: Using the pixel values itself as features.

• Histogram of Oriented Gradients (HOG): Using photographs to record gradient information.

• Color histograms: Showing how color values are distributed throughout various channels.

• Local Binary Patterns (LBP): describing spatial linkages and textural patterns.

Video-Based Features: Video-based features improve activity recognition by capturing temporal information from video sequences. We investigate the following attributes:

• Calculating the difference between adjacent frames in order to capture motion.

• Optical Flow: Predicting how fast objects are moving between frames.

Features of 3D Convolution by utilizing 3D convolutions

capture both spatial and temporal information.

Performance Evaluation: It is crucial to use the right evaluation measures when evaluating the performance of activity recognition algorithms. We talk about the following widely-used metrics:

• Accuracy: The percentage of samples that were correctly categorized.

• Accuracy: The model's capacity to correctly detect positive samples.

• Recall: The model's capacity to identify positive samples with accuracy.

The harmonic mean of recall and precision is the F1 score.

• Confusion Matrix: A tabular display of the model's predictions in comparison to the actual data.

UCI dataset:

The UCI dataset, according to Anguita et al. [75], is derived from 30-person video recordings of people engaging in simple physical activities while wearing a waist-mounted smart phone with an inertial sensor. Three static positions—Standing, Sitting, and Lying—as well as three dynamic actions—Walking, Walking Downstairs, and Walking Upstairs—are included in this dataset.

MODEL	ACCURACY
SVM	84
LSTM	86
PCA+LSTM	78
PCA+CNN+LSTM	86
CNN+LSTM	92
CNN	90

Table 1: Accuracy

III. CONCLUSTION

: In this article, we looked into how to recognize activities using CNN models and a variety of characteristics. While video-based features like frame differences, optical flow, and 3D convolutions capture temporal dynamics, image-based features like raw pixel intensities, HOG, color histograms, and LBP capture static information. These features can be carefully chosen and combined to produce precise activity recognition using CNN models. The usefulness of the model can be determined by analyzing the performance using measures like accuracy, precision, recall, F1 score, and confusion matrix.

REFERENCES

[1] D. Spicer and M. Weber, "Computers | Timeline of Computer History | Computer History Museum." [Online]. Available: https://www.computerhistory.org/timeline/computers/. [Accessed: 10-May-2019].

[2] M. A. Labrador and O. D. Lara Yejas, Human activity recognition : using wearable sensors and smartphones. CRC Press, 2013. [

[3] Y. Fu, Human activity recognition and prediction. Springer, 2016.

[4] J. Wang, Z. Liu, and Y. Wu, Human Action Recognition with Depth Cameras. Cham: Springer International Publishing, 2014.

[5] J. Reyes-Ortiz, D. Anguita, A. Ghio, L. Oneto, and X. Parra, "UCI Machine Learning Repository: Human Activity Recognition Using Smartphones Data Set," 2012. [Online]. Available: DOI: 10.1109/ICAEE48663.2019.8975649 https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones. [Accessed: 20-Apr-2019].

[6] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7657 LNCS, pp. 216–223, 2012.

[7] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, A Public Domain Dataset for Human Activity Recognition Using Smartphones. 2013.

[8] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Energy efficient smartphone-based activity recognition using fixed-point arithmetic," J. Univers. Comput. Sci., vol. 19, no. 9, pp. 1295–1314, 2013.

[9] S. Bhattacharjee, S. Kishore, and A. Swetapadma, "A Comparative Study of Supervised Learning Techniques for Human Activity Monitoring Using Smart Sensors," in 2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC), 2018, pp. 1–4.

[10] P. Bota, J. Silva, D. Folgado, and H. Gamboa, "A Semi-Automatic Annotation Approach for Human Activity

Recognition," Sensors (Basel)., vol. 19, no. 3, pp. 1–23, 2019.

[11] W. Jiang and Z. Yin, "Human activity recognition using wearable sensors by deep convolutional neural networks," MM 2015 - Proc. 2015 ACM Multimed. Conf., pp. 1307–1310, 2015.

[12] B. Almaslukh, A. Jalal, and A. Abdelmonim, "An Effective Deep Autoencoder Approach for Online SmartphoneBased Human Activity Recognition," Int. J. Comput. Sci. Netw. Secur., vol. 17, no. 4, pp. 160–165, 2017.

[13] A. Zheng and A. Casari, Feature Engineering for Machine Learning and Data Analytics - Principles and Techniques for Data Scientists. 2018

[14] N. Sikder, K. Bhakta, A. A. Nahid, and M. M. M. Islam, "Fault Diagnosis of Motor Bearing Using Ensemble Learning Algorithm with FFT-based Preprocessing," 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), pp. 564–569, 2019.

[15] P. Stoica and R. Moses, "Spectral analysis of signals," 2005.

[16] M. Sewak, R. Karim, and P. Pujari, Practical Convolutional Neural. Packt Publishing, 2018.

[17] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. .

[18] C. C. Aggarwal, Neural networks and deep learning : a textbook. 2018.

[19] P. Golik, Z. Tüske, R. Schlüter, and H. Ney, "Convolutional neural networks for acoustic modeling of raw time signal in LVCSR," INTERSPEECH 2015, 2015.

[20] K. Bhakta, N. Sikder, A. Al Nahid, and M. M. M. Islam, "Fault Diagnosis of Induction Motor Bearing Using Cepstrum-based Preprocessing and Ensemble Learning Algorithm," in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1–6.

[21] G. Hackeling, Mastering Machine Learning with scikit-learn. Packt Publishing, 2014.

[22] J. L. Reyes-Ortiz, D. Anguita, L. Oneto, and X. Parra, "UCI Machine Learning Repository: Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set," 2015. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/SmartphoneBased+Recognition+of+Human+Activities+and+Postural+Transitions. [Accessed: 20-May-2019].