Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 14 Issue: 1s
Article Received: 02 March 2023 **Revised**: 05 April 2023 **Accepted**: 15 May 2023

Human Activity Recognition using WISDM Datasets

Pardeep Seelwal[1], Achyutha Prasad N[2], Chetana Srinivas[3], Rekha V S[4]
[1]Assistant Professor
Department of Computer Science
M.N.S. Government College Bhiwani
Pardeepseelwal@gmail.com
[2]Professor
Department of Computer Science and Engineering
East West Institute of TechnologyBengaluru, India
achyuth001@gmail.com
[3]Associate Professor Department of Computer Science and Engineering
East West Institute of TechnologyBengaluru, India
chetanasrinivas1@gmail.com
[4]Assistant Professor
Department of Computer Science and Engineering
Dayananda Sagar Academy of Technology and Management, Bengaluru, India
rekhasedam.vtu@gmail.com

Abstract— In this research, we describe a method for identifying human activity where we use a model that comprises of Convolutional Neural Network (CNN), using the datasets acquired from smartphones. Jogging, sitting, walking, standing, going downstairs and upstairs are included in the daily activities that taken into consideration. We use the WISDM datasets acquired from the three-dimensional raw accelerometer sensors. The performance of our CNN model showed 87.85% accuracy. This performed better than the Support Vector Machine model, which had an accuracy rate of 82.27 percent. Therefore, the results demonstrate that our suggested strategy can outperform current best practices for human activity recognition.

Keyword: - convolutional neural networks; human activity recognition using 3D accelerometer data; Support Vector Machine (SVM)

I. INTRODUCTION

Human Activity Recognition (HAR) has been around for generations, being widely [1]-[4] used in many applications like smarthome systems, healthcare etc. These smart applications have paved way for tremendous growth in Human Activity Recognition. The sensors embedded in the smart watches, smart phones and other smart appliances are accelerometer andgyroscopes. These sensors [20]-[24] that are embedded in these smart appliances are portable and does not contain large number of sensory devices, which makes it less complex and usable. The huge variety in wearables and mobile applications available has led to greater impact in recognition of human activity. We used the data that was accumulated from a health app, the activities that were detected include running, cycling and walking. We compared the recognized results with regard to the ground truth labels. The accuracy of data collected from them were low due to the individual differences in movement patterns. The recognition efficiency can be escalated if theHAR model [5]-[6] were to be trained by a particular user's personal activity information and is same for a particular user. HAR canbe implemented using various techniques and sensors. One of the ways in which data can be collected is by the vision-based approach where the data is collected using cameras, but this approach suffers from privacy issues. Another approach is through the Wi-Fi signals where the data is acquired by sending Wi-Fi signals where the patterns in which it is received determines [7]-[9] the human activity, the problems in this approach is the obstacles in the environments. These issuescan

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 14 Issue: 1s
Article Received: 02 March 2023 **Revised**: 05 April 2023 **Accepted**: 15 May 2023

be resolved using our proposed model. This model is a more efficient and accurate when compared with the SVM-based method and also used less computational cost in chronological [10] order to recognize human activity.

II. IMPLEMENTATION

This section contains the detailed explanation of our proposed architecture using CNN models. It includes a convolution layer is proceeded by a max-pooling & other CNN layer. After which the model contains a SoftMax layer that further classifies [18]-[19] the result into an activity class. Beforethe first CNN layer the normalization of the accelerometer datais employed to convert the datasets acquired into numerical form.

1. Input

As illustrated in Fig. 1.a. input, raw accelerometer data with a set time-length was employed. Utilizing a sliding window with a specific size of 90 samples and 50%-timedomain overlap, we divided each sensor value produced by theaccelerometer into the sensor's y, z, and x-axis. Then, to enhance efficiency and implement one-dimensional (1D) convolutional, the resulting data was rearranged so that it hada height of 1. The Z-score standardization approach was used to normalize the data; its formula is as follows:

$$Z' = \frac{Z - \mu}{\sigma} \qquad (1)$$

The equation (1) can be used to evaluate the mean value &standard deviation of Z, respectively, as well as the standard values of the raw y, z, and x acceleration data.

The dataset that was submitted to normal distribution was obtained after such basic data preprocessing.

2. Convolution

In our model, there are two CNN layers. Convolutional operations with a filter of size 60 and a stride size of 1 were used to create a feature mapping with a size of 31 in the first layer. A total of 60 different filters were applied to the input y, z, and x accelerometer raw information [14]-[16] before it was transmitted for normalization. The input from the max pooling layer was used asthe input for the second CNN layer, which was created by applying a filter with particular size and depth of 6 and a stride size of 1, which produced a single feature map that served as the source for the subsequent linked layer. All of the CNN's max- pooling layers are one-dimensional (1D).

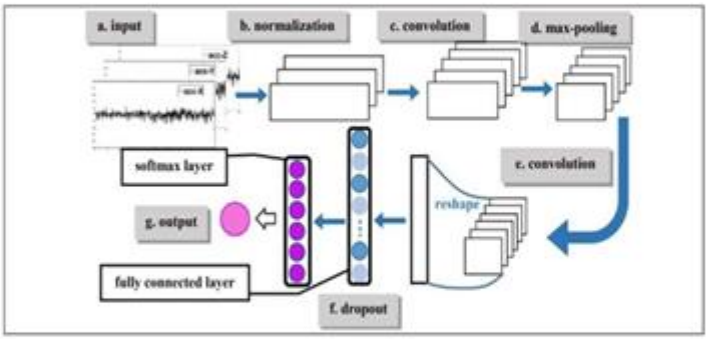

Fig.1. The network-basedarchitecture of the CNN-based model

3. Max pooling

After the data has passed through the first CNN layer, max pooling was performed, as illustrated in Fig. 1. d. The goal of using this layer was to choose the highest value for two reasons in particular: (1) Adding invariance to the processing, such as translation, rotation, and scale invariance; and (2) Reducing the parameter [11] and essential attributes to reduce computational complexity. The max-pooled featurevectors of size 6 can be produced by this filter size for a layer of 21, which was decided, along with a stride of 2, and was intended to serve as the second convolution layer's input.

Available @ http://onlineengineeringeducation.com

BY NC SA

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 14 Issue: 1s
Article Received: 02 March 2023 **Revised**: 05 April 2023 **Accepted**: 15 May 2023

4. Dropout

The output was flattened out after processing from the second CNN layer, as illustrated in Fig. 1. reshape to produce a lengthy feature unit vector of 1080 bytes, which was then utilized [12] as an input for the following connected layer. The dropout percentage's default value was 0.5, which was intended to prevent the model from over fitting and enhance generalization performance. There were 1000 neurons in the fully linked layer, and the network used here had a tan h activation function since it activated neurons more effectivelythan the sigmoid function.

5. Output

The SoftMax is used as an output layer towards the conclusion of this architecture, as highlighted in Fig. 1. Layer softmax. The SoftMax layer determines the output probabilityand assigns [13] a class of activities (i.e., jogging, standing, sitting, walking, downstairs, and upstairs). At the last stage ofthe design, as illustrated in Fig. 1.g. output, the activity that was classified having the greatest likelihood was selected as the acclaimed activity and the name was delivered as output.

III.EXPERIMENTS

In this section, we discuss the experiments performed using the datasets, including the hyper parameter settings in CNN model and feature extracting in SVM method.

1. Dataset

The ACTi tracker dataset that we used was made available by the Wireless Sensor Data Mining (WISDM) lab. This dataset includes six common human endeavors—jogging, standing, sitting, walking, going downstairs, and going upstairs—that were recorded using a particular accelerometer sensor in a well-organized lab setting.

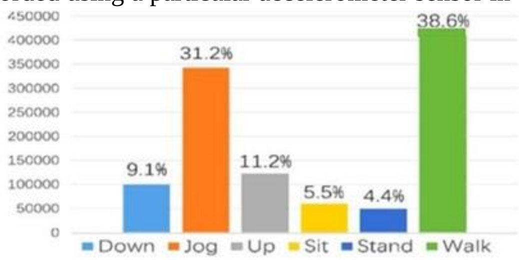

Fig. 2. Data distribution of WISDM dataset

The information was gathered using a 20Hz sample rate from 36 participants who put their smartphone in their pocket (20 datas per second). The dataset [25] has a total of 1,098,207 samples, and Figure 2 displays how they are distributed according to various activities (class labels).

Utilizing a moving window as mentioned in II A. input Section, the dataset obtained from the sensors in our study was first divided into 2,440,4 segments with a 50% overlap. Following that, the segments are separated using the 70/30 rule and the training and testing experiments. the volume of the train& test data, as shown in Table I.

TABLE I. TRAIN & TEST DATA

Train	Test	Total
1,707,9	7,325	2,440,4

2. Experiment Settings

The architecture of our CNN architecture technique is depicted in Figure 1 and was constructed using open source tools and TensorFlow. Additionally, CNN was tested and trained in the Python environment. Convolution window size = 60, 6, number of used convolution filters = 60, 6 (in the two convolution layers, respectively),

Available @ http://onlineengineeringeducation.com
BY NC SA

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 14 Issue: 1s
Article Received: 02 March 2023 **Revised**: 05 April 2023 **Accepted**: 15 May 2023

window stride = 1, pooling window volume = 21, window stride = 2, training batch volume = 16, training epoch volume = 20, and number of hidden neurons = 1000 were used to choose the values for the experiment's hyper parameters. We employed the Adaptive Moment Estimation optimization technique to reduce the adverse cost function in order to further optimize the CNN architecture.

The following are descriptions of some significant hyper- parameter settings in our CNN model:

- **Training epoch volume:** The training data loss (6.71 percent) and testing data loss (8.03 percent) varied particularly less when the training epoch size was set at 20, so it was deemed to be a suitable value.
- **Training batch size:** In general, a short training data batch can prevent slipping into the local minimum value, which will enhance the network's generalization capabilities. As a result, we choose a relatively low number, such as 16, for the training batch volume.
- **Weight initialization:** When initializing the weights of the convolutional kernels in the convolutional layer, tiny numbers are typically utilized, such as the raw data set subjected to a Gaussian distribution with a standard deviation (= 0.1) and mean value of 0 (= 0).

For the evaluation we compared the results of this architecture with a SVM-based activity recognition. SVM, the method of classification that mainly focuses on feature extraction. The six kindsof datasets acquired from WISDM were given as input forboth SVM and our CNN architecture. The free sourceLIBSVM, which functioned operating system for discs, was used to implement the SVM paradigm (DOS). As shown in equation (2), the y, z, and x acceleration data must first becomeinto vector immensity data by computing their Euclidean norms:

$$\|a\| = \sqrt{x^2 + y^2 + z^2} \qquad (2)$$

The six features are then briefly introduced in Table II after being extracted based on data from a vector immensity accelerometer as determined above.

TABLE II. FEATURES EXTRACTED FROM WISDM DATASET

Feature	Description
Mean	Average value of samples in window
Max	Maximum
Min	Minimum
STD	Standard deviation
Range	Maximum minus minimum
Mean crossing rate	Rate of signal crossing mean value

The moving window described in II.A. input section was used to extract all the attributes displayed in Table II from the vector immensity accelerometer sensor data.

Essentially it was seen that after the above process, the amount of data required for the trained and test data of an SVMmodel was particularly large when compared with that of our CNN model, in addition to that the computational complexity of the data processing was twice as much in SVM model.

IV.RESULT

The results of both the models when compared shows that, the accuracy and the efficiency of the classification based on human activity was more using the CNN-based model approach, which was comparatively less in SVM-based model approach. Depending on the features chosen, the SVM-based activity recognition's accuracy can vary. Three distinct types of sensors were used to capture the data for this system, then following feature extraction, 248 features in total could be seenin both the time domain and the frequency domain. The accuracy percentage for the SVM-based recognition systemwas 89.59 percent. This demonstrates that a significant level

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 14 Issue: 1s
Article Received: 02 March 2023 **Revised**: 05 April 2023 **Accepted**: 15 May 2023

of computational complexity in feature extraction is required for the SVM-based human activity recognition (Har) system toattain high precision.

Unlike SVM-based system, the CNN-based system for recognizing human activity can, as anticipated, achieve great accuracy at a cheap computing cost. The input for this model isgiven as per the dataset that was provided, this dataset is then converted to the balanced dataset. This balanced dataset is suitable for this model in order to accurately classify the activity class. As shown in Table III, the input of the following activity classes acquired after being passed through the proposed CNN model and the SVM-based method. This dataset is further optimized as shown in Table IV.

Table III. INPUT VALUES FOR THE PROPOSED CNN METHOD AND THE SVM METHOD

Walking	137375
Jogging	129392
Upstairs	35137
Downstairs	33358
Sitting	4599
Standing	3555
Name: activity, dtype: int64	

As shown in Table III, the input of the following activity classes acquired after being passed through the proposed CNN model and the SVM-based method. This dataset is further optimized as shown in Table IV.

Table III. INPUT VALUES FOR THE PROPOSED CNN METHOD AND THE SVM METHOD

Walking	3555
Sitting	3555
Upstairs	3555
Downstairs	3555
Standing	3555
Jogging	3555
Name: activity, dtype: int64	

Table IV shows the dataset values after they have beenfurther optimized into balanced data for an efficient processing using both the proposed CNN and the SVM- based method.

```
4/4 [==============================] - 0s 3ms/step - loss: 0.2489 - accuracy: 0.8785

   cat_crossentropy  ||   accuracy
  ________________________________
[0.2488994598388672, 0.8785046935081482]
```

Fig 3. The Resultant Values of Accuracy and Loss of the proposed CNN Method

CC BY NC SA

The above figure represents the resultant accuracy and the loss when the acquired dataset is passed through the proposed CNN model. It is seen that the accuracy when the data is passedthough CNN-based approach is **87.85%** and the loss encountered was of **0.25.** When the above results were in in contrast to outcomes seen in the SVM-based model approach, it was observed that the accuracy and the loss was significantly less and the model also required high computational cost.

The above graph shows Model accuracy and the Model loss against the epoch values for the respective properties likeaccuracy and loss. When the datasets were fed to the CNN- based model approach. The graph was obtained as an output for the CNN based approach. The first graph, shows how the resultant accuracy value was higher when compared to the training accuracy value. The second graph, shows how the resultant loss was less when compared with the training loss value.

V. CONCLUSION

This study uses raw tri-dimensional accelerometer data that was obtained from the smartphone of the user to create and compare a CNN-based model technique for human activity identification. We found that our proposed modeloutperformed the SVM-based model approach for human activity recognition. This shows how human activity recognition can be done more efficiently using the CNN- based model approach in comparison to other cutting-edge models. In addition, this architecture can be implemented in adevice like smartphone and wearable's that is portable and easy to use. This architecture can be used in devices with limited computational resources. Thus, our proposed model proves to be more accurate and efficient when compared with the SVM-based model approach. However, the proposed system still has some restriction such as the activity classes such as walking downstairs and upstairs provide less percentage accuracy when comparedwith other activity classes. Hence the future implementation of this project would be to increase the percentage of accuracy in these activity cases.

REFERENCES

[1] Zubair, Muhammad, K. Song, and C. Yoon. "Human activity recognition using wearable accelerometer sensors." IEEE International Conference on Consumer Electronics-Asia IEEE, 2016:1-5.

[2] Yang, Hua Cong, et al. "HARLib: A human activity recognition library on Android." International Computer Conference on Wavelet Active Media Technology and Information ProcessingIEEE, 2015:313-315.

[3] Siirtola, Pekka, and Juha Röning. "Recognizing Human Activities Userindependently on Smartphones Based on Accelerometer Data."International Journal of Interactive Multimedia & Artificial Intelligence 1.5(2012):38-45.

[4] Ha, Sojeong, and S. Choi. "Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors." International Joint Conference on Neural Networks IEEE, 2016:381-388.

[5] Yang, Jian Bo, et al. "Deep convolutional neural networks on multichannel time series for human activity recognition."International Conference on Artificial Intelligence AAAI Press, 2015:3995-4001.

[6] Lee, Song Mi, M. Y. Sang, and H. Cho. "Human activity recognition from accelerometer data using Convolutional Neural Network." IEEE International Conference on Big Data and Smart Computing IEEE, 2017:131-134.

[7] Sheng, Min, et al. "Short-time activity recognition with wearable sensors using convolutional neural network." ACM SIGGRAPH Conference on Virtual-Reality Continuum and ITS Applications in Industry ACM, 2016:413-416.

[8] Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." Journal of Machine Learning Research 15.1(2014):1929- 1958.

[9] Kwapisz, Jennifer R., G. M. Weiss, and S. A. Moore. "Activity recognition using cell phone accelerometers." Acm Sigkdd Explorations Newsletter 12.2(2011):74-82.

[10] Google TensorFlow, https://www.tensorflow.org/ (retrieved: Mar. 27, 2017)

[11] Kingma, Diederik P, and J. Ba. "Adam: A Method for Stochastic Optimization." Computer Science (2014)

[12] Hsu, C. W., C. C. Chang, and C. J. Lin. "A Practical Guide to Support Vector Classication." Taipei: National Taiwan University, Department of InformationEngineering 67.5(2016).

BY NC SA

[13] LIBSVM, http://www.csie.ntu.edu.tw/~cjlin/libsvm/ (retrieved: Dec. 22,2016)

[14] Chang, Chih Chung, and C. J. Lin. LIBSVM: A library for support vector machines. ACM, 2011.

[15] Tran, Duc Ngoc, and D. D. Phan. "Human Activities Recognition in Android Smartphone Using Support Vector Machine." International Conference on Intelligent Systems, Modelling and Simulation IEEE, 2017:64-68.

[16] N. A. Prasad and C. D. Guruprakash, "An ephemeral investigation on energy proficiency mechanisms in WSN," 2017 3rd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Tumkur, 2017, pp. 180-185.

[17] A. P. N and C. D. Guruprakash, "A Relay Node Scheme for Energy Redeemable and Network Lifespan Enhancement," 2018 4th International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Mangalore, India, 2018, pp. 266-274.

[18] Achyutha Prasad, N., Guruprakash, C.D., 2019. A relay node scheme of energy redeemable and network lifespan enhancement for wireless sensor networks and its analysis with standard channel models. International Journal of Innovative Technology and Exploring Engineering 8, 605–612.

[19] Achyutha Prasad, N., Guruprakash, C.D., 2019. A relay mote wheeze for energy saving and network longevity enhancement in WSN. International Journal of Recent Technology and Engineering 8, 8220–8227. doi:10.35940/ijrte.C6707.

[20] Achyutha Prasad, N., Guruprakash, C.D., 2019. A two hop relay battery aware mote scheme for energy redeemable and network lifespan improvement in WSN. International Journal of Engineering and Advanced Technology 9, 4785–4791. doi:10.35940/ijeat.A2204.109119.

[21] Prasad N. Achyutha, Sushovan Chaudhury, Subhas Chandra Bose, Rajnish Kler, Jyoti Surve, Karthikeyan Kaliyaperumal, "User Classification and Stock Market-Based Recommendation Engine Based on Machine Learning and Twitter Analysis", Mathematical Problems in Engineering, vol. 2022, Article ID 4644855, 9 pages, 2022. https://doi.org/10.1155/2022/4644855

[22] Rekha VS, Siddaraju., "An Ephemeral Analysis on Network Lifetime Improvement Techniques for Wireless Sensor Networks", International Journal of Innovative Technology and Exploring Engineering, vol. 8, issue 9, 2278-3075, pp. 810–814, 2019.

[23] R. V S and Siddaraju, "Defective Motes Uncovering and Retrieval for Optimized Network," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), 2022, pp. 303-313, doi: 10.1109/ICCMC53470.2022.9754109.

[24] N. G and G. C. D, "Unsupervised Machine Learning Based Group Head Selection and Data Collection Technique," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), 2022, pp. 1183-1190, doi: 10.1109/ICCMC53470.2022.9753995.

[25] Pooja Chopra, Vijay Suresh Gollamandala, Ahmed Najat Ahmed, S. B. G. Tilak Babu, Chamandeep Kaur, N. Achyutha Prasad, Stephen Jeswinde Nuagah, " Automated Registration of Multiangle SAR Images Using Artificial Intelligence & quot, Mobile Information Systems, vol. 2022, Article ID 4545139, 10 pages, 2022. https://doi.org/10.1155/2022/4545139.