AJMAS
https://alqalam.utripoli.edu.ly/science/ eISSN 2707-7179

Original article

Performance Analysis of Human Activity Recognition

Urvashi Garg

Research Scholar M.Tech, College of Engineering, Roorkee, UTU, U.K, INDIA

ARTICLE INFO

DOI: 10.5281/zenodo.4052343

* **Urvashi Garg:** Research Scholar M. Tech, College of Engineering, Roorkee, UTU, U.K, INDIA. **Mobile phone**: +8874549734. urvashigargm@gmail.com

Received: *28-07-2020*
Accepted: *26-09-2020*

Keywords: *Convolutional, Neural, Networks, Classifier.*

ABSTRACT

Inside the past couple of years, human action Recognition (HAR) has become an extremely imperative and significant territory of exploration because of the progression of the computerized periods of gadgets like cell phones, smartwatches, and camcorders used in our day by day lives. Profound Learning (DL) has driven analysts to utilize HAR in different spaces including wellbeing and prosperity applications the most point of this paper is to Classify the exercises happening in the casings. HAR application fields produce a major measure of information and not every one of them gives the computational force that DL models in HAR require. Utilizing our Convolutional Neural Network and Keras, we had the option to get 97.07% precision.

Cite this articling *U. Performance Analysis of Human Activity Recognitional J Med App Sci. 2020;3(2):72-77.*

INTRODUCTION

Human Activity Recognition (HAR) has become a popular point inside the past couple of years because of its hugeness in contemplating numerous regions, including medicinal services applications, on the web and disconnected intelligent gaming, sports, and checking frameworks for general purposes [1]. HAR research has an abrupt increment in Deep Learning (DL) techniques as result high precision in acknowledgment [2,3].HAR has numerous application cell phones and smartwatches or is coordinated into garments or other explicit clinical hardware [2,4-6]. The activities are human-centered and spread a wide scope of classes including human-object cooperation, for example, playing instruments, just as human-human collaborations, for example, shaking hands.

Related Works

As to HAR approaches, most of the recently distributed examinations address two kinds of acknowledgment utilizing directed learning [7–12] or semi-administered learning [13,14]. Move learning has likewise been explored, whereby the examples or models for exercises in a single area can be moved to improve the acknowledgment precision in another space to decrease the requirement for preparing information [15–16].

Furthermore, HAR has been broadly announced in numerous fields utilizing sensor modalities, including surrounding sensors [21], wearable sensors [22], cell phones [23], and smartwatches [24]. Those sensors add to building up a wide scope of utilization spaces, for example, sport [25],human–PC communication [26], observation [27], video spilling [28], medicinal services framework [29], furthermore, PC vision zone [30]. Non-visual sensors since they are both difficulties allowed to introduce and privacy saving [31,32].

The corrupted complex picture and Video were influenced by an alternate type of unwanted signals like Brownian Noise (Fractal Noise) Rayleigh Noise, Gamma Noise, Poisson-Gaussian Noise, salt and pepper commotion, irregular esteemed drive clamor, dot clamor, Gaussian clamor, and Structured Noise, and so on as demonstrated as follows[37-40]. Clamor is extremely hard to expel it from the intricate pictures without the correct comprehension of the commotion model.

Table 2. Related Works in HAR approaches

Paper	**Approach**	**Method**	**Activity**	**Input Source**	**Performanc e**
[17]	Hybrid Deep learning for Activity and Action Recognition	GMM, KF, Gated Recurrent Unit	Walking, jogging, running.	Video	96.3%
[18]	Comparison study to classify human activities	SVM, MLP, RF, Naive Bayes	Sleeping, eating, walking, falling, talking on the phone	Image	86%
[19]	Active Learning to recognize human activity using Smartwatch	RF, Extra Trees, Naive Bayes, Logistic Regression, SVM	Running, walking, standing, sitting, lying down	Smart watch	93.3%
[20]	Recognizing human activity using smartphone sensors	Quadratic, k-NN, ANN, SVM	Walking upstairs, downstairs	Smart phone	84.4%
[33]	Zero-Shot activity recognition using visual and linguistic attributes	BGRU, GloVe	Drink, uncork, drool, lick	Image	42.17 %
[34]	Zero-shot activity-recognition based on a structured knowledge graph	Two-stream GCN method, self-attention mechanism	Biking, Skiing	Video	59.9%

DATASET AND METHODOLOGY

The dataset used in this paper is kinetics and link of dataset as below:https://deepmind.com/research/open-source/kinetics

The dataset contains 400 human activity classes, with at any rate 400 video cuts for each activity. Each clasp endures around 10s and is taken from an alternate YouTube video. The activities are human focused and spread a wide scope of classes including human-object cooperation, for example, playing instruments, just as human-human connections, for example, shaking hands. There are 400 movement pictures in the dataset and the model is prepared on every one of the 400 exercises.

Keras and Convolutional Neural Networks (CNNs)

The CNN engineering, we will use today is a littler, progressively smaller variation of the VGGNet network very deep convolutional networks for Large Scale Image Recognition[36]. The goal of the convolution operation is to extricate the elevated level highlights, for example, edges, from the info picture.

ConvolutionalNets need not be restricted to only one convolutional Layer and got high computational[35]. Customarily, the first convolutional Layer is liable for catching the Low-Level highlights, for example, edges, shading, inclination direction, and so on. With included layers, the engineering adjusts to the high-level highlights also, giving us a system, which has a healthy comprehension of pictures in the dataset. The pooling layer is utilized for lessening the spatial size of the convolved include. There are two essential sorts of pooling we used to lessen the spatial size one max pooling and another is mean pooling.

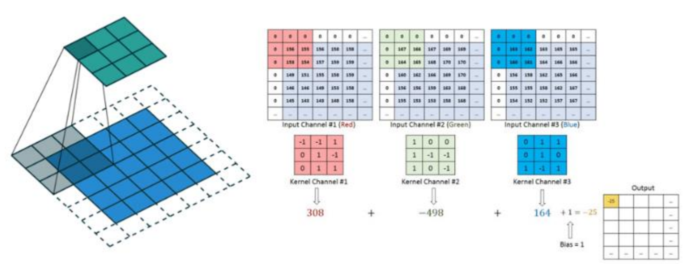

Figure1: Convolution Operation is to extract the high-level features. *Different parameters are used in proposed methodology like width,height,depth and classes.Width: The image width dimension.Height: The image height dimension.Depth: The depth of the image — also known as the number of channels.Classes: The number of classes in our dataset (which will affect the last layer of our model)*

So as to build exactness we increment our channel size and utilized max pooling. The more profound we go in the system, the littler the spatial elements of our volume as we increment size of channel and by diminishing the size of max pooling from 3 x 3 to 2 x 2 to guarantee we don't decrease our spatial measurements excessively fast.

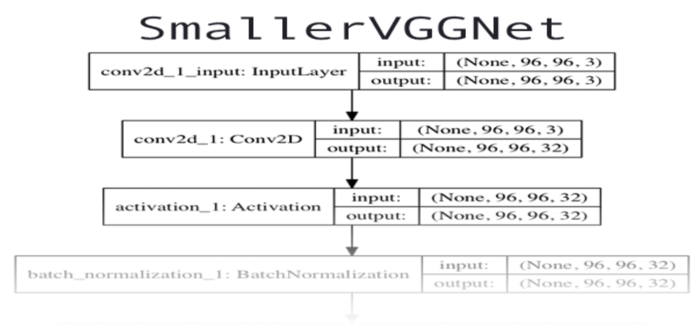

Figure 2: Architecture of the Model. *After training about 100 steps we got around Train Accuracy: 93.04 and Val Accuracy: 81.77.*

Optimize the network to improve the validation accuracy

We have expanded our channel size to 128 here. Dropout of 25% of the hubs is performed to diminish overfitting again lastly, we have a lot of FC => RELU layers and a SoftMax classifier. We balance the model with a SoftMax classifier that will restore the anticipated probabilities for each class label.A representation of the system engineering of initial scarcely any layers of SmallerVGGNet.

Table 2. Result Analysis

Epoch	Train Loss	Validation Loss	Train Accuracy	Validation Accuracy
0	1.3	2	0.7	0.5
20	0.8	0.8	0.75	0.8
40	0.7	0.8	0.8	0.85
60	0.6	1.5	0.85	0.5
80	0.52	0.75	0.9	0.8
100	-	0.5	0.95	0.95

CONCLUSION

To achieve this undertaking, we utilized a human action acknowledgment model pre-prepared on the Kinetics dataset, which incorporates 400-700 human exercises (contingent upon which adaptation of the dataset you're utilizing) and more than 300,000 video clips. Trained a Convolutional Neural Network (CNN) utilizing the Keras profound learning library. We are getting Good Accuracy with a large portion of the exercises covered. Using our Convolutional Neural Network and Keras, we had the option to get 97.07% precision.

Future Aspects

- Implementing a more robust network with more amount of data and a high system.
- We can implement a Resnet 3D model for the better extraction of features.
- Prediction of frames can be done in batch frames
- We can Implement it of many purposes like Activities in ATM's, Airports, Examination Surveillance, etc.

Conflict of Interest

The authors declare no conflict of interest.

REFERENCES

[1] Lara, O. D., and Labrador, M. A. A survey on human activity recognition using wearable sensors. IEEE communications surveys & tutorials 15, 3 (2012), 1192–1209.

[2] Kanjo, E., Younis, E. M., and Ang, C. S. Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection. Information Fusion 49 (2019), 46–56.

[3] Zeng, M., Nguyen, L. T., Yu, B., Mengshoel, O. J., Zhu, J., Wu, P., and Zhang, J. Convolutional neural networks for human activity recognition using mobile sensors. In 6th International Conference on Mobile Computing, Applications and Services (2014), IEEE,pp. 197–205.

[4] Liu, L., Peng, Y., Liu, M., and Huang, Z. Sensor-based human activity recognition system with a multilayered model using time series shapelets. Knowledge-Based Systems 90 (2015), 138–152.

[5] Neverova, N., Wolf, C., Lacey, G., Fridman, L., Chandra, D., Barbello, B., and Taylor, G. Learning human identity from motion patterns. IEEE Access 4 (2016), 1810–1820.

[6] Stisen, A., Blunck, H., Bhattacharya, S., Prentow, T. S., Kjærgaard, M. B., Dey, A., Sonne, T., and Jensen, M. M. Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition. In Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (2015), ACM, pp. 127–140.

[7] de Souza Júnior, A.H.; Corona, F.; Barreto, G.D.A.; Miché, Y.; Lendasse, A. Minimal Learning Machine:A novel supervised distance-based approach for regression and classification. Neurocomputing 2015, 164,34–44.

[8] Botros, M. Supervised Learning in Human Activity Recognition Based on Multimodal Body Sensing.Bachelor's Thesis, Radboud University, Nijmegen, The Netherlands, 2017.

[9] Nabian, M. A Comparative Study on Machine Learning Classification Models for Activity Recognition. J. Inf.Technol. Softw. Eng. 2017.

[10] He, J.; Zhang, Q.; Wang, L.; Pei, L. Weakly Supervised Human Activity Recognition from Wearable Sensorsby Recurrent Attention Learning. IEEE Sens. J. 2019, 19, 2287–2297.

[11] Kharat, M.V.;Walse, K.H.; Dharaskar, D.R.V. Survey on Soft Computing Approaches for Human ActivityRecognition. Int. J. Sci. Res. 2017, 6, 1328-1334

[12] Qian, H.; Pan, S.J.; Miao, C. Distribution-Based Semi-Supervised Learning for Activity Recognition; AAAI:Menlo Park, CA, USA, 2019.

[13] Zhu, Q.; Chen, Z.; Soh, Y.C. A Novel Semisupervised Deep Learning Method for Human Activity IEEE Trans. Ind. Informat. 2019, 15, 3821–3830.

[14] Chen, W.H.; Cho, P.C.; Jiang, Y.L. Activity Recognition Using Transfer Learning. Sens. Mater. 2017, 29,897–904.

[15] Cook, D.J.; Feuz, K.D.; Krishnan, N.C. Transfer learning for activity recognition: A survey. Knowl. Inf. Syst.2013, 36, 537–556.

[16] Hu, D. Transfer learning for activity recognition via sensor mapping. In Proceedings of the Twenty-SecondInternational Joint Conference on Artificial Intelligence, Barcelona, Spain, 16–22 July 2017.
[17] Jaouedi, N.; Boujnah, N.; Bouhlel, M.S. A new hybrid deep learning model for human action recognition.J. King Saud Univ. Comput. Inf. Sci. 2019, in press.
[18] Alex, P.M.D.; Ravikumar, A.; Selvaraj, J.; Sahayadhas, A. Research on Human Activity Identification Basedon Image Processing and Artificial Intelligence. Int. J. Eng. Technol. 2018, 7.
[19] Shahmohammadi, F.; Hosseini, A.; King, C.E.; Sarrafzadeh, M. Smartwatch Based Activity RecognitionUsing Active Learning. In Proceedings of the 2017 IEEE/ACM International Conference on ConnectedHealth: Applications, Systems and Engineering Technologies (CHASE), Philadelphia, PA, USA, 17–19 July2017; pp. 321–329.
[20] Bulbul, E.; Cetin, A.; Dogru, I.A. Human Activity Recognition Using Smartphones. In Proceedings of the2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT),Ankara, Turkey, 19–21 October 2018; pp. 1–6.
[21] Laput, G.; Zhang, Y.; Harrison, C. Synthetic Sensors: Towards General-Purpose Sensing. In Proceedings ofthe 2017 CHI Conference on Human Factors in Computing Systems, CHI ’17, Denver, CO, USA, 6–11 May2017; pp. 3986–3999.
[22] Chung, S.; Lim, J.; Noh, K.J.; Kim, G.; Jeong, H. Sensor Data Acquisition and Multimodal Sensor Fusion forHuman Activity Recognition Using Deep Learning. Sensors 2019, 19, 1716.
[23] Bulbul, E.; Cetin, A.; Dogru, I.A. Human Activity Recognition Using Smartphones. In Proceedings of the2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT),Ankara, Turkey, 19–21 October 2018; pp. 1–6.
[24] Balli, S.; Sa˘gba¸s, E.A.; Peker, M. Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm. Meas. Control. 2018, 52, 37–45.
[25] Hsu, Y.L.; Yang, S.C.; Chang, H.C.; Lai, H.C. Human Daily and Sport Activity Recognition Using a WearableInertial Sensor Network. IEEE Access 2018, 6, 31715–31728.
[26] Ilbeygi, M.; Kangavari, M.R. Comprehensive architecture for intelligent adaptive interface in the field ofsingle-human multiple-robot interaction. ETRI J. 2018, 40, 411–553.
[27] Dharmalingam, S.; Palanisamy, A. Vector space based augmented structural kinematic feature descriptor for human activity recognition in videos. ETRI J. 2018, 40, 499–510.
[28] Moon, J.; Jin, J.; Kwon, Y.; Kang, K.; Park, J.; Park, K. Extensible Hierarchical Method of Detecting Interactive Actions for Video Understanding. ETRI J. 2017, 39, 502–513.
[29] Zheng, Y.; Ding, X.R.; Poon, C.C.Y.; Lo, B.P.L.; Zhang, H.; Zhou, X.L.; Yang, G.Z.; Zhao, N.; Zhang, Y.T.Unobtrusive Sensing and Wearable Devices for Health Informatics. IEEE Trans. Biomed. Eng. 2014, 61,1538–1554.
[30] Jalal, A.; Kim, Y.; Kim, Y.J.; Kamal, S.; Kim, D. Robust human activity recognition from depth video usingspatiotemporal multi-fused features. Pattern Recognit. 2017, 61, 295–308.
[31] Stankovic, J.A.; Srinivasan, V. Non-Invasive Sensor Solutions for Activity Recognition in Smart Homes; University of Virginia: Charlottesville, VA, USA, 2012.
[32] Bhandari, B.; Lu, J.; Zheng, X.; Rajasegarar, S.; Karmakar, C.K. Non-invasive sensor based automatedsmoking activity detection. In Proceedings of the 2017 39th Annual International Conference of the IEEEEngineering in Medicine and Biology Society (EMBC), Jeju Island, Korea, 11–15 July 2017; pp. 845–848.
[33] Zellers, R.; Choi, Y. Zero-Shot Activity Recognition with Verb Attribute Induction. In Proceedingsof the EMNLP 2017: Empirical Methods in Natural Language Processing, Copenhagen, Denmark,7–11 September 2017.
[34] Gao, J.; Zhang, T.; Xu, C. I Know the Relationships: Zero-Shot Action Recognition via Two-Stream GraphConvolutional Networks and Knowledge Graphs. In Proceedings of the AAAI, Honolulu, HI, USA,27 January–1 February 2019.
[35] Binay Kumar Pandey,Sanjay Kumar Pandey,and Digvijay Pandey. “A Survey of Bioinformatics Applicaion on Parallel Architectures.”InternationalJournal of Computer Applications”,vol. 23, pp.21-25, June 2011.
[36] Simonyan, K. and Zisserman, A. Two-stream convolutional networks for action recognition in

videos. CoRR, abs/1406.2199, 2014. Published in Proc. NIPS, 2014.
[37] D. Pandey, Binay Kumar Pandey, Subodh Wariya, "An Approach To Text Extraction From Complex Degraded Scene", IJCBS, vol. 1, no. 2, May 2020.
[38] Digvijay Pandey., Binay Kumar Pandey, Subodh Wariya "Study of Various Techniques Used for Video Retrieval." Journal of Emerging Technologies and Innovative Research,vol. 6, pp.850-853, June 2019.
[39] Digvijay Pandey., Binay Kumar Pandey, Subodh Wariya "Study of Various Types Noise and Text Extraction Algorithms for Degraded Complex Image. " Journal of Emerging Technologies and Innovative Research, vol. 6, pp.234-247, June 2019.
[40] Pandey, D., Pandey, B.K. & Wairya, S. Hybrid deep neural network with adaptive galactic swarm optimization for text extraction from scene images. Soft Comput (2020). https://doi.org/10.1007/s00500-020-05245-4