

Human Activity Recognition using CNN

Neha Nilesh Jadhav Sarnaik

nehajadhavsarnaik@gmail.com

DOI: 10.29322/IJSRP.10.02.2020.p9804

<http://dx.doi.org/10.29322/IJSRP.10.02.2020.p9804>

Abstract- Activity recognition systems are a large field of research and development, currently with a focus on advanced machine learning algorithms, innovations in the field of hardware architecture, and on decreasing the costs of monitoring while increasing safety. Human action recognition (HAR) research is hot in computer vision, but high precision recognition of human action in the complex background is still an open question. Most current methods build classifiers based on complex handcrafted features computed from the raw inputs, which are driven by tasks and uncertain. In this paper, type of deep model convolutional neural network (CNN) is proposed for HAR that can act directly on the raw inputs.

Index Terms- CNN, HAR, frame extraction

I. INTRODUCTION

The issue of programmed distinguishing proof of physical exercises performed by human subjects is alluded to as Human Activity Recognition (HAR). Human movement acknowledgment includes ordering times arrangement information, estimated at inertial sensors, for example, accelerometers or spinners, into one of pre-characterized activities.

In the most recent decade, human activity acknowledgment (HAR) is turning into an increasingly more appealing examination subject with a few applications, for example, video observation, augmented reality, astute human-PC communications, and so on. Nonetheless, precise acknowledgment of activities is a profoundly testing errand because of jumbled foundations, impediments, and perspective varieties. HAR comprise of a few phases, which portray the highlights that characterize exercises or low-level activities. CNN is a profound model that gets confounded various leveled highlights through convolutional activity rotating with sub-examining activity on the crude info pictures. It is affirmed that CNN can acquire astounding execution in visual objective acknowledgment assignments through fitting modification during the preparation. Furthermore, CNN has invariance for a specific posture, brightening, and dislocated ecological change.

A portion of the ordinarily utilized highlights for human activity acknowledgment are Histogram of Oriented Gradient (HOG), Histogram of Optical Flow (HOF), Motion Interchange Patters (MIP), Space-Time Interest Points (STIP), activity bank highlights and thick directions. Be that as it may, these methodologies are troublesome and tedious to stretch out these highlights to different frameworks. A huge piece of hand-plan

highlights is driven by task and various errands may utilize totally various highlights. In any case, as a general rule, it is difficult to tell what sort of highlight is imperative to a particular assignment, so the component determination is profoundly reliant on the particular issue. Particularly for human activity acknowledgment, various types of sports show an exceptionally enormous distinction in the appearance and movement model, it is difficult to get the basic element of activity in the uncommon difference in condition. In this way, a conventional component extraction strategy is should have been proposed to mitigate the requirement for hand-designed highlights and diminishes the figuring scale.

II. LITERATURE REVIEW

Late decade saw a decent number of productions in the field of visual reconnaissance to perceive the anomalous exercises. Moreover, a couple of studies can be found in the writing for the distinctive unusual exercises acknowledgment; however, none of them have tended to various anomalous exercises in an audit. In [1], Tripathi, R. K et al present the best in class which shows the general advancement of suspicious action acknowledgment from the observation recordings in the most recent decade. We incorporate a short presentation of the suspicious human action acknowledgment with its issues and difficulties. This Paper comprises of six irregular exercises, for example, surrendered object recognition, robbery location, fall identification, mishaps and illicit stopping discovery on street, brutality action identification, and fire recognition. By and large, we have talked about every one of the means those have been followed to perceive the human action from the reconnaissance recordings in the writing, for example, frontal area object extraction, object location dependent on following or non-following strategies, include extraction, characterization; movement examination and acknowledgment. The target of paper is to give the writing audit of six distinctive suspicious action acknowledgment frameworks with its general system to the analysts of this field.

Human action acknowledgment (HAR) plans to perceive exercises from a progression of perceptions on the activities of subjects and the ecological conditions. The vision-based HAR inquire about is the premise of numerous applications including video observation, social insurance, and human-PC collaboration (HCI). This survey features the advances of cutting edge action acknowledgment draws near, particularly for the action portrayal and characterization techniques. For the portrayal strategies, we sort out an ordered research direction from worldwide portrayals to neighborhood portrayals, and late profundity based portrayals. For the order techniques, we fit in with the arrangement of format based strategies, discriminative models, and generative

models and audit a few predominant strategies. Next, delegate and accessible datasets are presented. Planning to give an outline of those techniques and a helpful method for looking at them, we group existing literary works with a point by point scientific classification including portrayal and order strategies, just as the datasets they utilized. At last, we examine the bearings for future research.

Movement acknowledgment frameworks are an enormous field of innovative work, right now with an emphasis on cutting edge AI calculations, developments in the field of equipment design, and on diminishing the expenses of observing while at the same time expanding wellbeing. Paper [3] focuses on the uses of movement acknowledgment frameworks and reviews their best in class. We classify such applications into dynamic and helped living frameworks for keen homes, human services observing applications, checking and observation frameworks for indoor and outside exercises, and tele-drenching applications. Inside these classifications, the applications are arranged by the philosophy utilized for perceiving human conduct, in particular, in light of visual, non-visual, and multimodal sensor innovation. We give an outline of these applications and talk about the points of interest and confinements of each approach. Furthermore, we represent open informational collections that are intended for the assessment of such acknowledgment frameworks. The article closes with a correlation of the current strategies which, when applied to true situations, permit to figure examine inquiries for future methodologies.

The issue of programmed recognizable proof of physical exercises performed by human subjects is alluded to as Human Activity Recognition (HAR). There exist a few methods to gauge movement attributes during these physical exercises, for example, Inertial Measurement Units (IMUs). IMUs have a foundation position in this specific situation, and are described by utilization adaptability, minimal effort, and decreased protection sway. With the utilization of inertial sensors, it is conceivable to test a few estimates, for example, speeding up and rakish speed of a body, and use them to learn models that are prepared to do accurately arranging exercises to their comparing classes. In paper [4], Bevilacqua, A. et al propose to utilize Convolutional Neural Networks (CNNs) to group human exercises. This model utilized crude information acquired from a lot of inertial sensors.

Human action acknowledgment includes ordering times arrangement information, estimated at inertial sensors, for example, accelerometers or spinners, into one of pre-characterized activities. Convolutional neural system (CNN) has built up itself as an incredible strategy for human action acknowledgment, where convolution and pooling tasks are applied along the transient element of sensor signals. In a large portion of existing work, 1D convolution activity is applied to individual univariate time arrangement, while multi-sensors or multi-methodology yield multivariate time arrangement. 2D convolution and pooling tasks are applied to multivariate time arrangement, so as to catch neighborhood reliance along both worldly and spatial spaces for uni-modular information, so it accomplishes superior with less number of parameters contrasted with 1D activity. Anyway, for multi-modular information existing CNNs with 2D activity handle various modalities similarly, which cause impedances between qualities from

various modalities. In [5], Ha, S. et al present CNNs (CNNpf and CNN-pff), particularly CNN-pff, for multi-modular information. We utilize both halfway weight sharing and full weight sharing for our CNN models so that methodology explicit attributes just as normal qualities across modalities are found out from multi-modular (or multi-sensor) information and are in the end accumulated in upper layers. Analyses on benchmark datasets exhibit the superior of our CNN models, contrasted with condition of expressions of the human experience strategies.

Human action acknowledgment (HAR) is a characterization task for perceiving human developments. Techniques for HAR are of incredible enthusiasm as they have become instruments for estimating events and spans of human activities, which are the premise of shrewd assistive advancements and manual procedures examination. As of late, profound neural systems have been conveyed for HAR with regards to exercises of everyday living utilizing multichannel time-arrangement. These time-arrangement are gained from body-worn gadgets, which are made out of various kinds of sensors. The profound designs process these estimations for discovering fundamental and complex highlights in human corporal developments, and for characterizing them into a lot of human activities. As the gadgets are worn at various pieces of the human body, Fernando Moya Rueda et al [6] proposed a novel profound neural system for HAR. This system handles arrangement estimations from various body-worn gadgets independently. An assessment of the design is performed on three datasets, the Opportunity, Pamap2, and a modern dataset, beating the cutting edge. Furthermore, unique system arrangements will likewise be assessed. We find that applying convolutions per sensor channel and per body-worn gadget improves the abilities of convolutional neural system (CNNs).

Human activity acknowledgment (HAR) investigate is hot in PC vision, however high accuracy acknowledgment of human activity in the intricate foundation is as yet an open inquiry. Most present strategies assemble classifiers dependent on complex carefully assembled highlights figured from the crude information sources, which are driven by undertakings and questionable. In [7], sort of profound model convolutional neural system (CNN) is proposed for HAR that can demonstration straightforwardly on the crude sources of info. What's more, a productive pre-preparing system has been acquainted with decrease the high computational expense of bit preparing to empower improved certifiable applications. The proposed approach has been tried on the KTH database and the accomplished outcomes think about well against cutting edge calculations utilizing hand-structured highlights.

Action grouping is a difficult issue because of enormous sign dimensionality, high intra-and between subject changeability in action designs, nearness of transitional classes indicating blend of examples, and predominance of the invalid class. Administered learning has been the pervasive decision with profound neural systems (DNNs) indicating some encouraging potential. Profound adapting anyway requires an enormous number of named tests which is hard to get, particularly from defenseless more established individuals. In [8], Gerasimos Kalouris Evangelia actualized 3 diverse convolutional neural system models prepared on information from more established individuals, fusing Bayesian enhancement for effective hyper

parameter tuning. We abuse different enlargement techniques for time-arrangement to cause invariant forecasts and furthermore cross-to use information about physical action of more youthful people so as to improve speculation in our models intended for more established grown-ups.

Ha et al utilize the open datasets, and apply 2-D convolution over a solitary channel portrayal of the dynamic signs. This specific utilization of CNNs for the movement acknowledgment issue is additionally expounded by Ha et al. [9], with a multi-channel convolutional arrange that use both quickening and rakish speed signs to order day by day exercises from an open dataset of upper-appendage developments. The grouping task they perform

is customized, so the signs accumulated from every member are utilized to prepare singular learning models.

Alsheikh et al. [10] acquaint a first methodology with HAR dependent on profound learning models. They create a spectrogram picture from an inertial sign, so as to encourage genuine pictures to a convolutional neural system. This methodology beats the requirement for reshaping the signs in a reasonable arrangement for a CNN, be that as it may, the spectrogram age step essentially replaces the procedure of highlight remove however the spectrogram generation step simply replaces the process of feature extraction, adding initial overhead to the network training.

III. PROPOSED SYSTEM

Input video is converted to frames for better assessment. Activity recognition is continuously performed in system. Input image is first pre-processed to denoise it and then activity is recognized using CNN. Pre-processing is a common name for operations with images at the lowest level of abstraction both input and output are intensity images. The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing. Feature extraction is concentrated principally around the measurement of the geometric properties (size and shape) and surface characteristics of regions (color and texture).

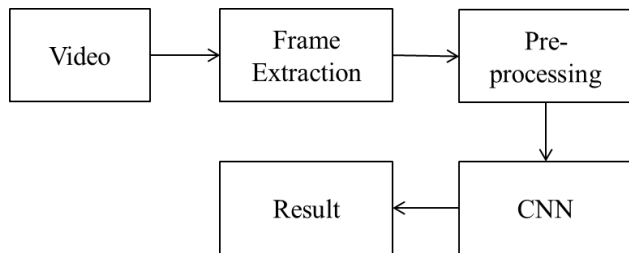


Fig block diagram of proposed system

The objective of image classification is to identify and portray, as a unique gray level (or color), the features occurring in an image in terms of the object or type of land cover these features actually represent on the ground. Classification (or recognition) of activity is performed using CNN.

IV. RESULT

Video of different activities is fed to system, classified result are given in fig 2-7. Human Activities like walking, reading, sleeping, tilting, standing etc. are recognized.

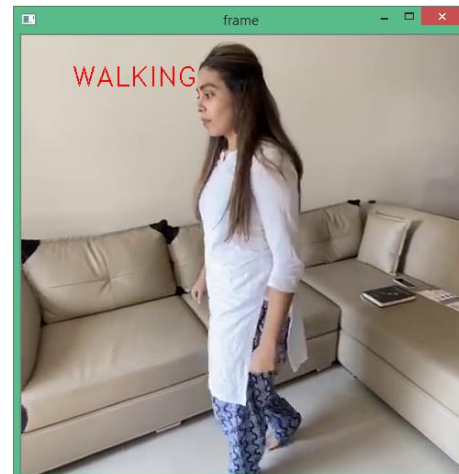


Fig 2 walking activity is recognized

Proposed system compares images form dataset (which are fed for training of CNN), as per similarity classified activity into particular group.

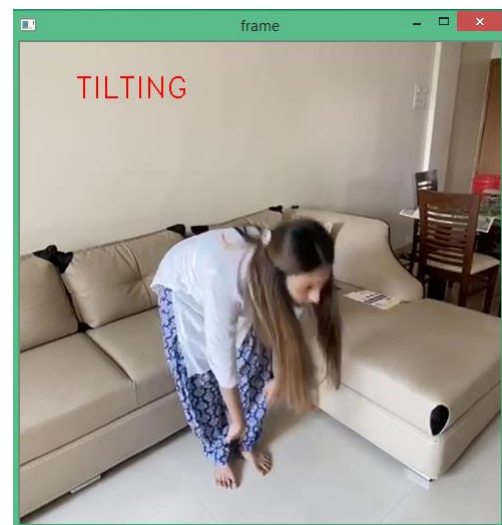


Fig 3 tilting activity is recognized

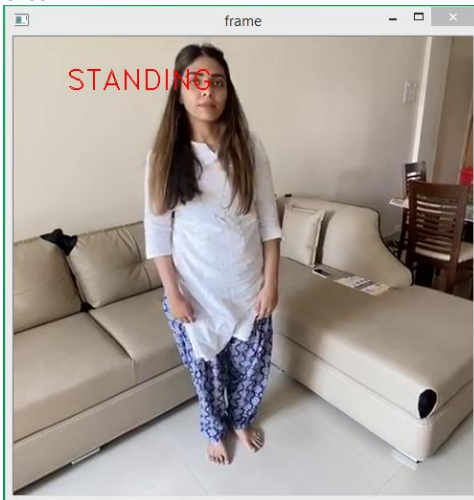


Fig 4 standing activity is recognized

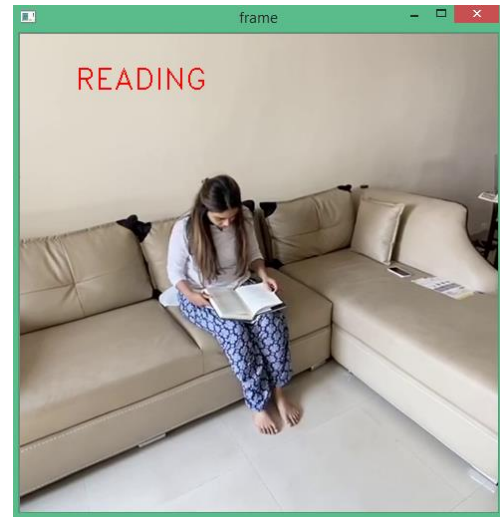


Fig 7 reading activity is recognized

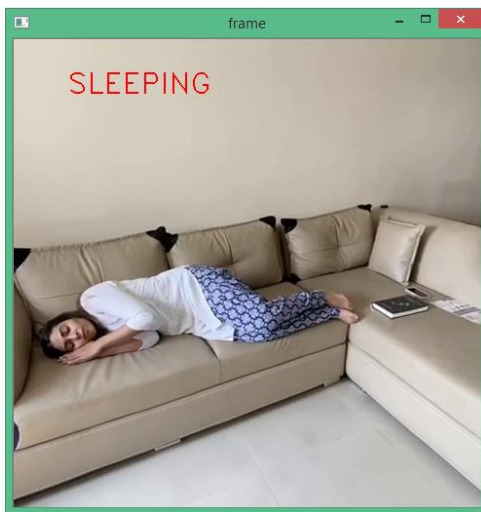


Fig 5 sleeping activity is recognized

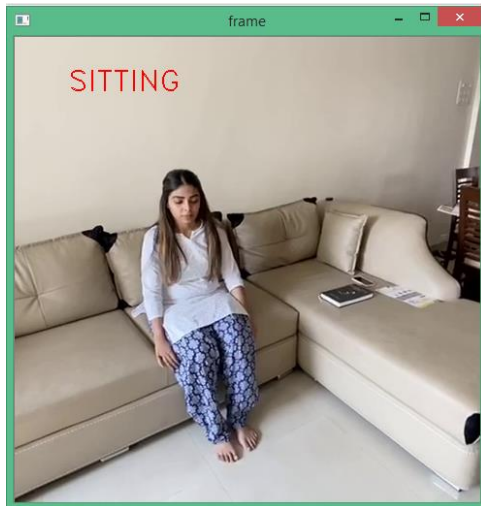


Fig 6 sitting activity is recognized

V. CONCLUSION

Human activity recognition (HAR) aims to recognize activities from a series of observations on the actions of subjects and the environmental conditions. In this paper, we presented a CNN model for the HAR problem. We focused on a set of activities extracted from a common exercise program for fall prevention. Further work on the application of convolutional models to real-world data is recommended. More activities could be included in the workflow, and different aggregations on the activities can be tested. In particular, it is recommended to diversify the population of participants, in order to validate the classification mechanism to wider age groups.

REFERENCES

- [1] Tripathi, R. K., Jalal, A. S., & Agrawal, S. C. (2017). Suspicious human activity recognition: a review. *Artificial Intelligence Review*, 50(2), 283–339. doi:10.1007/s10462-017-9545-7
- [2] Shugang Zhang,1 Zhiqiang Wei,1 Jie Nie,2 Lei Huang,1 Shuang Wang,1 and Zhen Li, "A Review on Human Activity Recognition Using Vision-Based Method", *Hindawi Journal of Healthcare Engineering* Volume 2017
- [3] Suneth Ranasinghe, Fadi Al Machot and Heinrich C Mayr, "A review on applications of activity recognition systems with regard to performance and evaluation", *International Journal of Distributed Sensor Networks* 2016, Vol. 12(8)
- [4] Bevilacqua, A., MacDonald, K., Rangarej, A., Widjaya, V., Caulfield, B., & Kechadi, T. (2019). Human Activity Recognition with Convolutional Neural Networks. *Lecture Notes in Computer Science*, 541–552. doi:10.1007/978-3-030-10997-4_33
- [5] Ha, S., & Choi, S. (2016). Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors. 2016 International Joint Conference on Neural Networks (IJCNN). doi:10.1109/ijcnn.2016.7727224
- [6] Article Convolutional Neural Networks for Human Activity Recognition Using Body-Worn Sensors Fernando Moya Rueda, René Grzeszick, Gernot A. Fink, Sascha Feldhorst and Michael ten Hompel
- [7] Human Action Recognition based on Convolutional Neural Networks with a Convolutional Auto-Encoder Chi Geng1, a, JianXin Song 5th International Conference on Computer Sciences and Automation Engineering (ICCSAE 2015)
- [8] Improving CNN-based activity recognition by data augmentation and transfer learning Gerasimos Kalouris Evangelia I. Zacharaki VVR

Group, Vasileios Megalooikonomou XXX-X-XXXX-XXXX-
X/XX/\$XX.00 ©20XX IEEE

- [9] Ha, S., Choi, S.: Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors. In: 2016 International Joint Conference on Neural Networks (IJCNN), pp. 381–388, July 2016. [https://doi.org/ 10.1109/IJCNN.2016.7727224](https://doi.org/10.1109/IJCNN.2016.7727224)
- [10] Alsheikh, M.A., Selim, A., Niyato, D., Doyle, L., Lin, S., Tan, H.P.: Deep activity recognition models with triaxial accelerometers. CoRR abs/1511.04664 (2015). <http://arxiv.org/abs/1511.04664>

AUTHORS

Author – Neha Nilesh Jadhav Sarnaik, Master of Science in Computer Science and nehajadhavsarnaik@gmail.com.