DOI: https://doi.org/10.46632/daai/2/4/23

Data Analytics and Artificial Intelligence
Vol: 2(4), 2022
REST Publisher; ISBN: 978-81-948459-4-2
Website:http://restpublisher.com/book-series/data-analytics-and-artificial-intelligence/

Human activity recognition and Alert system

***P Velvizhy, K S Vinu Varshith, Ajay Sellappan, Abdul Kareem A**
Anna University, Chennai, Tamil Nadu, India.
*Corresponding author Email: velvizhi@gmail.com

Abstract: Monitoring and taking care of people has been a necessary factor. Since it is very vital to monitor the movements of the people left alone in our homes and people in unattended scenarios, this project focuses on Human Activity Recognition (HAR). This project showcases a smart human action recognition method to automatically identify the human activities from skeletal joint motions and combines the competencies. This is a low-cost method and has high accuracy. An independent mobile application is also developed to monitor the activities of the people and their surroundings when they are alone. A mobile notification is sent during abnormal activities. Thus, our project provides a way to help senior citizens from any kind of mishaps.

1. Introduction

Human activity recognition, or HAR for short, is a broad field of study concerned with identifying the specific move- ment or action of a person based on data. Movements are often typical activities performed indoors, such as walking, sleeping, standing, and sitting. They may also be more focused activities such as those types of activities performed in a kitchen or on a factory floor. The data may be remotely recorded, such as video, or other wireless methods. Alternately, data may be recorded directly on the subject such as by carrying custom hardware or smart phones that have accelerometers and gyroscopes. Historically, data for activity recognition was chal- lenging and expensive to collect, requiring custom hardware. Now smart phones and other personal tracking devices used for fitness and health monitoring are cheap and ubiquitous. As such, data from these devices is cheaper to collect, more common, and therefore is a more commonly studied version of the general activity recognition problem. The problem is to predict the activity given a snapshot of data. Generally, this problem is framed as a univariate or multivariate time series classification task. It is a challenging problem as there are no obvious or direct ways to relate the recorded data to specific human activities and each subject may perform an activity with significant variation, resulting in variations in the recorded data. The intent is to record data and corresponding activities for specific subjects, fit a model from this data, and generalize the model to classify the activity of new unseen subjects from their data.

2. Literature Survey

Due of the position and temporal fluctuations included in action movies, human action recognition (HAR) is a difficult undertaking.In this study, the innovative idea of a depth history image is combined with sequence and shape learning to overcome these issues (DHI).To describe the temporal relationship between the action frames, a deep bidirectional long short term memory (DBiLSTM) is built for sequential learning. A trained convolutional neural network is used to extract the action information from each frame (CNN). The DHI is created by estimating and projecting the depth information from each action frame onto the X-Y plane. When learning shapes, a deep pre-trained CNN network is taught using the shape data obtained from DHI. A deep pre-trained CNN network is trained during shape learning using the shape data obtained by DHI. The overfitting problem is resolved by utilising the trained knowledge of the pre-trained network. The fine-tuned network is employed to identify actions in the analysed DHI images. By essentially expanding the training set, data augmentation is used to prevent overfitting of the network. Assistive robots would need to recognise the current human activity and offer assistance with the future in some scenarios, for example. These scenarios also call for the algorithm to have both recognition and prediction capabilities. We must think about two questions in order to come up with a joint solution for activity recognition and prediction: What is a decent inference algorithm to deal with 1) what is an appropriate representation for the structure of human activities/tasks, , what would be a decent inference algorithm to handle such a representation, and secondly. The Markov model family is a well-liked group of event representations (e.g., hidden Markov Model). Markov models, however, lack sufficient expressiveness since human tasks frequently display non-Markovian and compositional features. Different human occupations can be distinguished by the geometric dynamic arrangements of bodily joints.However, a lot of current methods for recognising human movement are unable to automatically learn these configurations from sets of joints in four-dimensional space (spatio and temporal). This study suggests a method for automatically learning joint configurations that is based on dictionary learning and sparse representation. The suggested technique accomplishes the following attributes: 1) It quickly and easily learns dynamic spatio-temporal geometric configurations of joints in the body involved in activities; 2) It eliminates the need for manually designing features and introduces a new technique for arranging joint coordinate data into fixed length column vectors that are appropriate for dictionary learning; The learnt dictionary words record sub-activity properties, and the frequency of various words appearing in different activities characterizes the categories of global activity. 3) It substitutes the traditional bag of words model with the sparse coding method joint motions and combines the competencies. This is a low- cost method and has high accuracy. An independent mobile application is also developed to monitor the condition of the people and its surroundings when they are alone.

A Noti- fication API integration facilitates sending alert notification during abnormal condition is also implemented in the mobile application. Thus, our project provides a way to help the senior citizens and children from any kind of mishaps and health issues.

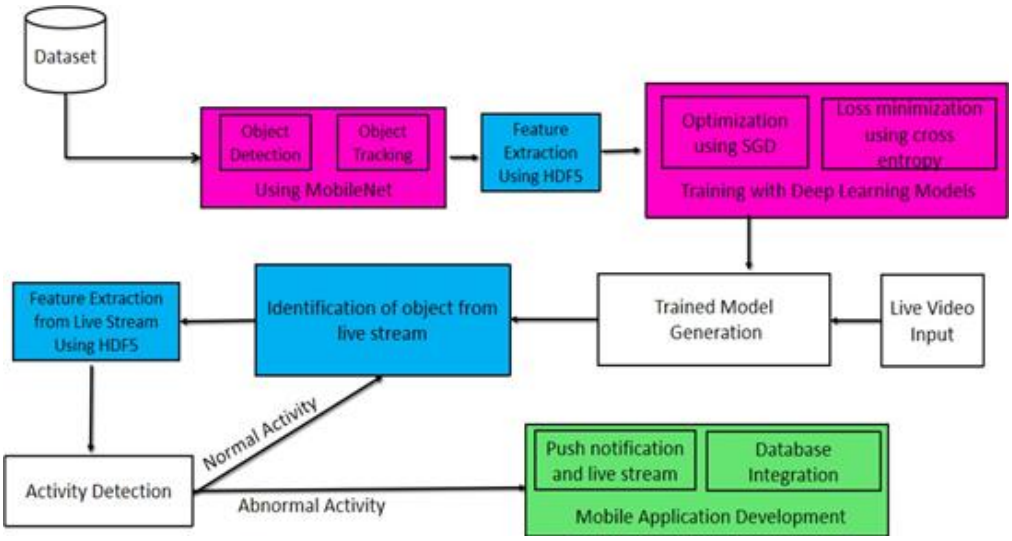

FIGURE 1. Overall Architecture Diagram

3. System Analysis

Existing System: Human action recognition (HAR) is a challenging task due to the presence of the pose and temporal variations in the action videos. To address these challenges, HAR-Depth is proposed in this paper with sequential and shape learning along with the novel concept of depth history image (DHI). A deep bidirectional long short term memory (DBiLSTM) is constructed for sequential learning to model the temporal rela- tionship existing between the action frames. Action informa- tion in each frame is extracted using pre-trained convolutional neural network (CNN). The depth information of each action frame is estimated and projected onto the X-Y plane to form the DHI. During shape learning, the shape information through DHI is used to train a deep pre-trained CNN network. By leveraging the trained knowledge of the pre-trained network, overfitting issue is handled. The finetuned network is used to recognize actions from query DHI images. Data augmentation is adopted to avoid overfitting of the network by virtually increasing the training set. The proposed work is evaluated on publicly available datasets like KTH, UCF sports, JHMDB, UCF101, and HMDB51 and the results on these datasets suggest that the proposed work of this paper performs better in terms of overall accuracy, kappa parameter and precision compared to the other state-of-the-art algorithms present in the earlier reported literature.

Proposed System: In the present scenario one of the common diseases that is found in elderly people is dementia. Dementia is a brain disorder that most often affects the elderly. It's caused by the failure or death of nerve cells in the brain. Alzheimer's disease is the most common cause. By some estimates, about one-third of people ages 85 and older may have Alzheimer's. Although age is the greatest risk factor for dementia, it isn't a normal part of aging. Some people live into their 90s and beyond with no signs of dementia at all. And monitoring and taking care of these people is been a major factor. Since it is very vital to monitor the movements of elderly and people in coma in unattended scenarios, this project focuses on HAR. This project showcases a smart human action recognition method to automatically identify the human activities from skeletal

4. System Design

Object Detection: Object detection can be defined as a branch of computer vision which deals with the localization and the identification of an object. Object localization and identification are two different tasks that are put together to achieve this singular goal of object detection. Object localization deals with specifying the location of an object in an image or a video stream, while object identification deals with assigning the object to a specific label, class, or description. We are using MobileNet model in this project for the purpose of object detection. This architecture uses proven depth-wise separable convolutions to build lightweight deep neural networks.

Labelling Of Dataset: Dataset labeling is the process of identifying raw data (images, text files, videos, etc.) and adding one or more meaningful and informative labels to provide context so that a deep learning model can learn from it. Data labeling is the process of identifying raw data (images, text files, videos, etc.) and adding one or more meaningful and informative labels to provide context so that a machine learning model can learn from it. For example, labels might indicate whether a photo contains a bird or car, which words were uttered in an audio recording, or if an x-ray contains a tumor. Data labeling is required for a variety of use cases including computer vision, natural language processing, and speech recognition.The frames of each pose

collected from the live camera are effectively are tagged to with the representative labels corresponding to the pose. The accuracy of the trained model will depend on the accuracy of the labeled dataset, so spending the time and resources to ensure highly accurate dataset labeling is essential.

5. Feature Extraction With Hdf5 Dataset Gen- Erator

HDF5 is a binary data format created by the HDF5 group to store gigantic numerical datasets on disk (far too large to store in memory) while facilitating easy access and computation on the rows of the datasets. Data in HDF5 is stored hierarchically, similar to how a le system stores data. Data is first defined in groups, where a group is a container-like structure which can hold datasets and other groups. Once a group has been denied, a dataset can be created within the 6group. A dataset can be thought of as a multi- dimensional array (i.e.,a NumPy array) of a homogeneous data type (integer, oat, unicode, etc.).

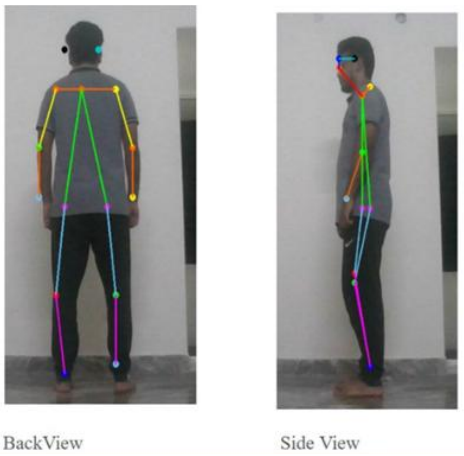

FIGURE 2. This figure shows the detection of single human and part detection and association

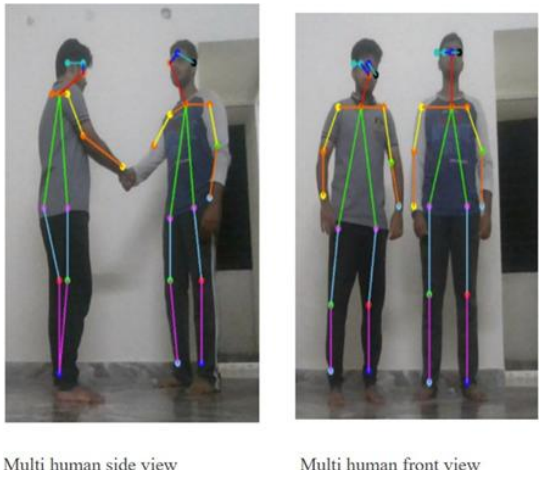

FIGURE 3. This figure shows the detection of multiple human and part detection and association

Feature extraction includes several convolution layers followed by max-pooling and an activation function. For extracting the features we will be using VGG16 architecture.

Live streaming: The internet of things, or IoT, is a system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers and the ability to transfer data over a network without re quiring human-to-human or human-to-computer interaction. Here IoT is used for live streaming. Streaming media is multimedia that is constantly received by and presented to an end-user while being delivered by a provider. The verb ”to stream” refers to the process of delivering or obtaining media in this manner. The term refers to the delivery method of the medium, rather than the medium itself, and is an alternative to file downloading, a process in which the end-user obtains the entire file for the content before watching or listening to it.

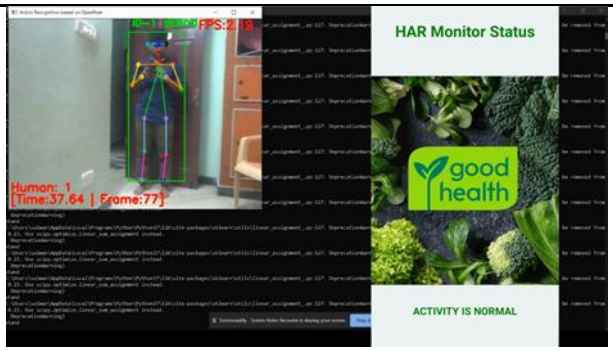

FIGURE 4. Normal Activity Screen

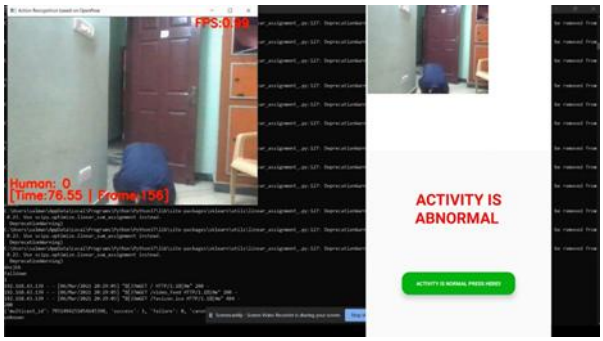

FIGURE 5. Abnormal Activity Screen with Live streaming

Mobile App Development: React Native is a framework that builds a hierarchy of UI components to build the JavaScript code. It has a set of components for both iOS and Android platforms to build a mobile application with a native look and feel. React Native seems to be a viable solution for building high-quality apps in a short time with the same performance and user-experience standards that native apps provide. React Native uses different mechanisms to create an efficient, consistent and reusable visual identity for the applications.

Performance Analysis: The performance analysis is been done in comparison with the existing system which listed out certain differences and our system proved to be efficient in comparison to the existing system. This can be seen in table1. Thus, from the above results and discussion, it is clear that we have efficiently made a project for identifying human activity and also we have developed a react native based mobile application where we can view the live streaming of the surroundings with a notification alert when an abnormal activity occurs. Thus we have successfully implemented the scope of the project.

TABLE 1. Comparison with the Existing system and Proposed syste

EXISTING SYSTEM	PROPOSED SYSTEM
Identifying the human activity image based	Identifying the human activities
Depth videos are used and frames are combined to find out the motion which results in errors	Efficient deep learning VGG-16 architecture is used to train the dataset of the human activities
Accuracy is less	Highly accurate
Research on the dataset is carried out	Real time implementation of the project
No application to view the surroundings and no alert message	Automatic notification alert when an abnormal activity occurs.

6. Metrics

Based on the result Training involves making a prediction based on the current state of the model, calculating how incorrect the prediction is, and updating the weights or parameters of the network to minimize this error and make the model predict better. We repeat this process until our model has converged and can no longer learn. We trained the project with number of epochs as 40 and the summary of our result and we obtained 0.9991 accuracy.

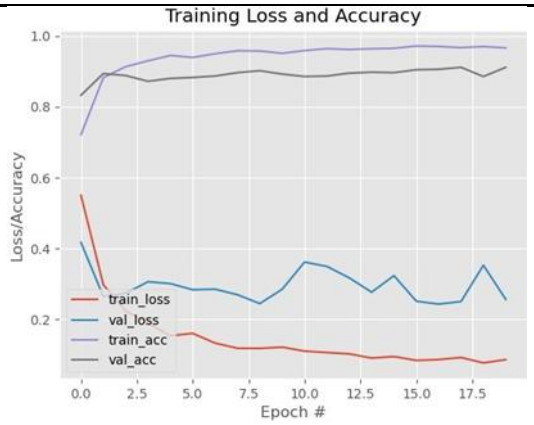

FIGURE 6. Obtained result in Graphical representation

Conclusion

This project a human action recognition and human gesture recognition system which can automatically recognize the human daily activities using the currently prevailing deep learning approach. We develop an effective skeleton infor- mation based HAR. It will recognize the human activities effectively. We have developed a react native application so that we can view the live streaming with a notification when an abnormal activity is detected. So, we can save our grandparents from sudden health issues and can also help the old age homes for taking care of the elder people.

References

1. "Chhavi Dhiman, Dinesh Kumar Vishwakarma","View-invariant Deep Architecture for Human Action Recognition using Two-stream Motion and Shape Temporal Dynamics", Vol No:1057-7149, 2020
2. "Ulysse Cˆ ot ´e-Allard, Cheikh Latyr Fall, Alexandre Drouin, Alexandre Campeau-Lecours, Cl´ement Gosselin, Kyrre Glette, Franc¸ois Laviolette†, and Benoit Gosselin", "Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning", vol. 1534- 4320, Mar. 2019.
3. "Jin Qi , Zhangjing Wang, Xiancheng Lin, and Chunming Li ", "Learn- ing Complex Spatio-Temporal Configurations of Body Joints for Online Activity Recognition", IEEE 2018, Vol No: 2168-2291.
4. "Wanru Xu, Zhenjiang Miao, Xiao-Ping Zhang, Yi Tian", "A Hierarchi- cal Spatio-Temporal Model for Human Activity Recognition", Vol:1520- 9210, 20177.
5. "Dapeng Tao, Lianwen Jin, Yuan Yuan, Yang Xue", "Ensemble Manifold Rank Preserving for Acceleration-Based Human Activity Recognition",Vol: 2162-237X, 2016
6. "Lei Wang, Xu Zhao, Yunfei Si, Liangliang Cao, Yuncai Liu", "Context- associative Hierarchical Memory Model for Human Activity Recognition and Prediction",Vol:1520-9210 , 2016.