International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 9 Issue 8 ǁ 2021 ǁ PP. 84-89

Human Body Action Tracking and Recognition Using ELM Algorithm

Prashil S. Taksande
UG student, Department of Electronics & Telecommunication Engineering, Government college of Engineering, Amravati, India

Piyush H. Balani
UG student, Department of Electronics & Telecommunication Engineering, Government college of Engineering, Amravati, India

Aditya A. Chavan
UG student, Department of Electronics & Telecommunication Engineering, Government college of Engineering, Amravati, India

Abstract
Human action recognition is the field of research that has gained importance in the recent years because of its widerange of applications spread across different fields. Human action recognition plays a crucial role in human-to-human interaction and also interpersonal relations as it provides the information about the identity, personality and psychological state of the person. The human ability to recognize other person's activities is one of the major subjects of study of the scientific areas of computer vision and machine learning and because of this research many applications such as video surveillance systems, human-computer interaction, control free gaming systems, etc. require multiple action recognition system. Human action recognition also has a very significant use in security systems installed in public places to track any suspicious activity or threat. This project aims to develop an algorithm which can recognize actions such as jogging, bending, bowling, jumping, kicking, running etc. from the input video sequences. The multiple actions in the video sequence will be detected with each passing frame of the video sequence.

Keywords: *Human action recognition, computer vision and machine learning.*

Date of Submission: 03-08-2021 Date of acceptance: 16-08-2021

I. INTRODUCTION

Human Action Recognition (HAR) basically refers to the task of analyzing the video sequence in order to recognize the activity or action that is taking place in that particular video. Detecting human activities in detail is very beneficial in areas which are particularly human centric such as home care support, abnormal activities, exercise and fitness, etc. Most of the human's daily task can be automated if such actions are recognized by HAR system. According to the complexity, the human actions are categorized into: atomic actions, gestures, group actions, human to object or human to human interactions, events and behaviors. Atomic actions refer to the movements of a person which describes certain motions that may be the part of more complex activities. Gestures are basic movements of the body part that corresponds to particular action of the person. The activities that involve two or more persons or objects are called as human to object or human to human interactions. The activities that are performed by group of persons is called as are called as group actions. Events refers to the high-level activities that describes social actions between individuals. Human behavior means the physical action that are related with the personality, emotions and psychological state of the individual. HAR systems are usually based on unsupervised or supervised learning. Unsupervised system has set of rules during its development whereas supervised system requires training using labelled data sets. In supervised method, the computer is provided with example inputs that are labelled with their desired outputs. In the unsupervised method, the data is unlabeled. As the technology is rising up the use of internet and smart phone are increased. The action recognition in the personal videos has become an important research topic due to its wide applications such as automatic video tracking and video annotation. Human action recognition has also its wide use in the area of security surveillance. Any suspicious activity can be identified with the use of human action recognition and it can help in security concerns.

II. LITERATURE SURVEY

With respect to our problem statement, we have referred research papers related to our project human body action tracking and recognition using ELM algorithm. This section summarize the various research papers on human body action tracking and recognition using ELM algorithm.

A. ***A Survey on Human Activity Recognition and Classification:*** This paper was proposed by Abhay Gupta, Kuldeep Gupta, et al., which focused on recent research papers based on various methods of activity recognition [1]. They mainly conducted their survey on three popular methods of activity recognition, namely smartphone sensors, wearable devices and vision based or using pose estimation.

B. ***Human Action Recognition Based on Skeleton and Convolutional Neural Network:*** In 2019, Yusi Yang, Zhuohao Cai, et al., have proposed the method which is based on data preprocessing using human skeleton information to recognize human action through Convolutional Neural Networks [2]. The authors have presented a Convolutional Neural Network (CNN) based automatic human action recognition method, which automatically learns the spatial and temporal characteristics of the data in order to improve the performance of recognition. In this paper interframe difference method is used to extract the key frames.

C. ***Subject Identification using Walking Posture:*** In 2019, Mihaela Hnatiuc, Mirel Paun, et al., have proposed a method to recognize the posture during walking with the help of leg inclination [3]. This system uses deflection sensor and mobile phone tilt sensor. The mobile phones are attached at the back side of the foot above and below the knee. The flex sensor is positioned on the foot. The system identifies the subject after the walking posture.

D. ***Human Action Recognition using Deep Neural Networks:*** This project was designed by Rashmi R. Koli, Tanveer I. Bagban in 2020 to develop a platform for a hand movement recognition which recognizes the hand gestures [4]. Human action recognition in other words is human gesture recognition. Gestures are nothing but the movement of body part that convey some meaningful message. In this project, they have used CNN algorithm as an interpreter that interprets the gestures and it builds a statement from the video. The statement or text is the meaning of those gestures.

E. ***An Overview of Extreme Learning Machine:*** Extreme Learning Machine (ELM) is one of the most important topics in the field of artificial intelligence in recent years. ELM has been widely used in human action recognition, multiclass classification and other fields. ELM provides efficient learning framework for regression, classification, feature learning and clustering [5]. It has much faster learning speed compared to traditional Support Vector Machine (SVM). In recent years the ELM applications have increased rapidly.

F. ***Human Activity Recognition Based on Evolution of Features Selection and Random Forest:*** In 2019, Christine Dewi and Rung-Ching Chen have proposed a study on dataset for Human Activity Recognition (HAR) which is based on four methods Support Vector machines (SVM), K-Nearest Neighbors (KNN), Random Forest (RF) and Linear Discriminant Analysis (LDA) with the different features to select the best classifier among the models to test the dataset [6]. From the experiments and analysis on different dataset it was concluded that among this four the RF is the best classifier method.

III. METHODOLOGY

1. Video Processing:

By identifying the actions of the video input, a specific video play is analyzed. The entire video is framed independently at a certain frame rate. Actions on video are framed by frame and appropriately labeled based on a pre-trained data set used for this project. Video processing is a form of signal processing, especially image processing, in which input and output signals are video or video streams. We can also process video processing with the help of image processing if we are able to increase the speed or rate at which the frames in the video are processed. If we process images in batches, the same processing power can be used and we can speed up the process.

2. Image Processing:

Image processing is a method of performing specific tasks or analyzing an image. In this process, we get an improved image or extract some useful information from a particular image. Image processing is a type of signal processing where we provide input such as an image and output that we receive can be an image or features associated with that image. Two types of image processing methods are analogue and digital image processing. Analogue image processing is often used for hard copies such as printing and photography. Digital

image processing techniques create the illusion of digital images with the help of computers. Other common categories that all types of data are required to use while using the digital process are pre-processing, development, display and retrieval of information.

2.1 OpenCV:

OpenCV (Open Source Computer Vision Library) is a library of program functions that focuses on real computer vision. OpenCV-Python is a library of Python bonds designed to 9 solve computer vision problems. OpenCV-Python uses Numpy, a highly optimized library of numerical performance with the MATLAB syntax. All properties of the same OpenCV members are converted and removed from the same Numpy members. This also makes it much easier to integrate with other Numpy libraries such as SciPy and Matplotlib.

3. Artificial Neural Networks:

A neural network consists of artificial neurons or layered nodes, which take a specific input vector and convert it to output. In this process, each neuron picks up input and activates a function that is usually a non-linear activity and transmits the result to the next layer.

4. Python:

Python is a goal-oriented programming language developed by Guido van Rossum that is becoming increasingly popular, especially because of its easy-to-read and code-readability. It enables the program editor to display ideas in a few lines of code without reducing readability. Python is a very simple and easy to use language compared to other languages such as Java, C, etc. The functionality in python is also useful compared to working in MATLAB. Also, one of the disadvantages of MATLAB is that it is not an open source and license is very expensive. Since python is an open source, in the event of an error, we can search it directly online and fix it. Some benefits of python are Easy to read, learn and write, Portability, Free and Open-Source, etc.

5. Extreme Learning Machines (ELM):

Extreme Learning Machines (ELM) is a feedforward neural network used for division, retrieval, integration, feature reading and compression by a single layer or multiple layers of hidden modes. ELM has advantages such as countless functions, pattern matching, automatic code formatting, good error correction and overcoming slow learning speed. ELM makes managing the flow of data between objects much easier as it offers new design code editing. With the use of ELM, we do not need to compare between the various components of the system so we can focus on problem solving.

6. Tools Used:

Anaconda is a distribution of Python and R computer programming languages for computer science including data science, machine learning applications, big data processing, forecasting statistics, etc. It aims to simplify package management and shipping. This distribution includes data science packages suitable for Windows, Linux, and MacOS. Anaconda distribution has more than 250 automated packages, and more than 7,500 open-source packages can be downloaded from PyPI as well as a conda package with a visual environment manager. Anaconda is an industry standard for growing, testing and training on a single machine.

7. Flowchart

At first the input image is taken, this input image is taken after converting to frame-by-frame from video input. The image will be checked through neural network, for this purpose the unlabelled data is used. The image that matches with the dataset is used. Feature extraction is used to recognise action and after recognising the action the movement is recognised in feature extraction. Classifier is used to select the particular class from the multiple classes to give the data. The data with which the input image matches the most is chosen. Finally, the image tag is displayed which gives the information about the action being performed in the video sequence.

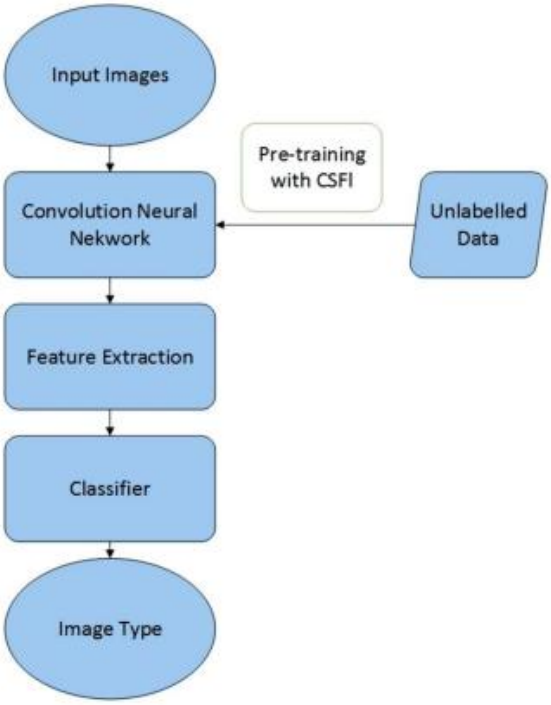

Fig.1: Flowchart

IV. RESULT

The project has been tested on 30 videos with a range of 200-1800 frames. According to visual analysis, labelling (HAR) seems to have accuracy of 80%. The below images show the result of project with some of the recognised actions.

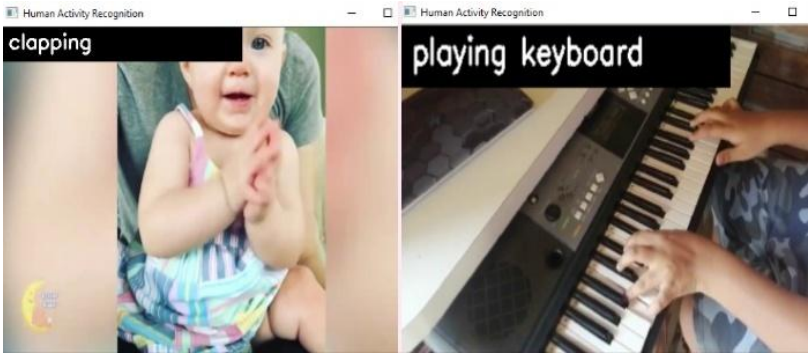

Fig.2: Output 1

Fig.3: Output 2

Fig.4: Output 3

V. APPLICATIONS

The major applications of human body action tracking and recognition using ELM algorithm are: to track suspicious actions from CCTV footage's, violence detection on the road and public places, to track the most wanted human's action and walking pattern so police can recognize them on a fake face, in sports for recognizing player's action while playing, in tracking motion patterns of humans that could be used to help identify people experiencing dangerous conditions such as a seizure, heart attack or serious fall, in the exercise and fitness field, in gaming and animations.

VI. CONCLUSION

In this paper we have presented the technique by which we can identify the human actions being performed in the video input. The literature survey on human action recognition shows that there has been plenty of research in the area of video analysis and human action recognition. After the emergence of neural networks, there has been a lot of research related to this topic in past 5-6 years. The Frame-by-Frame application of CNNs helped in improving the accuracies as compared to the manual feature extraction techniques. After that, 3D-CNNs has further improved the accuracies of CNNs by processing multiple frames at a time. More recent architectures have started focusing on Extreme Learning Machine (ELM) in order to factor in the temporal component of the videos. The most recent architectures have started developing attention mechanisms to focus on the important parts of the videos. Human action recognition is still an active research area, and new approaches are being presented to solve the issues with the current approaches. Some of the existing issues with human action recognition are background clutter or fast irregular motion in videos, view point changes, high computational complexity and responsiveness to illumination changes.

VII. FUTURE SCOPE

In future, system can be made more precise, for example, it will make clear distinction between almost similar types of activities like Stair Case Down-and-Walking and Jogging-and-Running. We can also make advancements in our system so that it can carry out different kinds of human physical analysis like heartbeat, pressure, specific disease like asthma and other medical issues. Constant monitoring and implementation for better analysis can be done by improving the system further. Implementation of wireless client-server architecture can be developed. Moreover, Innovations such as implementation of complete automated on-chip system for data collection and analysis can be made and accounting other prospective of data classification, application of digital filters with variable filter size.

REFERENCES

[1]. [1]. Abhay Gupta, Kuldeep Gupta, Kshama Gupta, Kapil Gupta, "A survey on human activity recognition and classification", International Conference on Communication and Signal Processing, July 2020, pp. 0915 – 0919.

[2]. [2]. Yusi Yang, Zhuohao Cai, Yingdong Yu, Tong Wu and Lan lin, "Human Action Recognition Based on skeleton and Convolutional Neural Network", 2019 Photonics Electromagnetics Research Symposium - Fall (PIERS - Fall), December 2019, pp.1109-1112.

[3]. [3]. Mihaela Hnatiuc, Mirel Paun, Ambroise lafeuille, "Subject Identification Using Walking Posture", 2019 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), November 2019, pp. 01-06.

[4]. Rashmi R. Koli, Tanveer I. Bagban, "Human Action Recognition Using Deep Neural Networks", 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), July 2020, pp. 376 – 380.

[5]. Bohua Deng, Xinman Zhang, Weiyong Gong, Dongpeng Shang, "An Overview of Extreme Learning Machine", 2019 4th International Conference on Control, Robotics and Cybernetics (CRC), September 2019, pp. 189 – 195.

[6]. Christine Dewi, Rung-Ching Chen, "Human Activity Recognition Based on Evolution of Features Selection and Random Forest", 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), October 2019, pp. 2496- 2501.

[7]. Jasvinder Pal Singh, Sanjeev Jain, Sakshi Arora, Uday Pratap Singh, "Vision-Based Gait Recognition: A Survey", IEEE Access (Volume: 6), November 2018, pp. 70527-70497.