Human Body Action Tracking and Detecting using Single Layer Neural Network

Prof. M. D. Ingle[1], Prashik Ramteke[2], Samarthraj Satbhai[3], Siddhi Pawar[4], Sanvidhan Mapare[5]

[1] Prof. M.D Ingle, Dept. Of Computer Engineering, Pune, Maharashtra, India

[2] Prashik Ramteke, Dept. Of Computer Engineering, Pune, Maharashtra, India,

[3] Samarthraj Satbhai, Dept. Of Computer Engineering, Pune, Maharashtra, India

[4] Siddhi Pawar, Dept. Of Computer Engineering, Pune, Maharashtra, India

[5] Sanbidhan Mapare, Dept. Of Computer Engineering, Pune, Maharashtra, India

ABSTRACT

Detecting human behavior growing in importance in recent years Wide range of applications spread across different applications fields. Awareness of human behavior plays an important role in this People-to-people exchanges and interpersonal relationships Since relationships provide information about identity, personality and mental state of the person. of the human ability to perceive the activities of others one of the major subjects in the natural sciences computer vision and machine learning This research applies to many applications such as video surveillance systems, human-computer interaction, Several duty-free game systems are required action recognition system. Recognition of human behavior It also has very important uses in security systems set up in public places to track down suspects activity or threat. The purpose of this project is Algorithms that can recognize actions such as jogging, Crouching, bowling, jumping, kicking, running, etc. Input video sequence. numerous actions in Video sequences are recognized in each pass Frames of a video sequence.

KEYWORDS:

Human action recognition, computer vision and machine learning.

1. INTRODUCTION

Human Action Recognition (HAR) is essentially The task of analyzing and recognizing video sequences an activity or action taking place in that particular thing video. Awareness of human activity in detail is very important Especially useful in human-centric fields Home care support, abnormal behavior, exercise, etc. fitness etc. Most of the men's daily chores are When such actions are detected by her HAR system, they are automated. Human behavior responds to complexity Classification: atomic actions, gestures, group actions, Human-to-object or human-to-human interactions, events and action. What is an atomic action a person who describes a particular movement that can be done Part of a more complex activity. gestures are basic Movement of the corresponding part human behavior. Activities for two people more people or things than those who object or are called out person-to-person interaction. its activity performed by a group of people is called group action. Events refer to high-level activities Accounts for social behavior between individuals. human Action means the physical action associated with it personality, emotions, and psychological state of Individually. HAR systems are typically based on: Unsupervised learning or supervised learning. Unmanned The system has a set of rules under development, Supervised system requires training with labeled data put. For supervised procedures, a computer will be provided sample input labeled with what you want Exit. In the unsupervised method, the data No label. As technology advances, The spread of the Internet and smartphones is progressing. action Recognition in personal videos became one It is an important research theme because of its wide range of applications. B. Automatic video tracking and video annotation. human Action detection also has broad applications in the following areas: security surveillance. Suspicious activity is Detecting human behavior and using it to identify Useful for security concerns.

2. LITERATURE SURVEY

we referenced our issue Research activities of the Human Body Action Project Tracking and detection by ELM algorithm. of The section brings together various research papers on humans Body Motion Tracking and Detection with ELM algorithm. A. Human activity detection survey and Classification: This paper was proposed by Abhay Gupta. A recent study focused on Kuldeep Gupta et al. Essays based on different methods of activity detection [1].

Conducting a survey centered on the three most popular Method of activity detection i.e. smartphone sensor, wearable, vision-based or pose-based evaluation. B. Skeleton-based human behavior detection and Convolutional Neural Networks: 2019, Yusi Yang, Zhuohao Cai et al. proposed a procedure. Based on data preprocessing using human skeleton Information that detects human behavior Convolutional neural networks [2].

the author has Convolutional Neural Network (CNN) based a method for automatically recognizing human behavior, Automatically learns spatially and temporally Data characteristics to improve Achievement of recognition. In this paper, intermediate frames A difference method is used to

extract the keyframes. Ha Subject identification by walking posture: in 2019 Suggested by Mihaela Natiuk, Mirel Paun et al. Posture recognition method when walking Leg tilt aid [3].

This system uses distractions sensor and mobile phone tilt sensor. mobile phone Attached to the top and bottom of the sole of the foot knees. A flex sensor is placed on the foot. of The system identifies subjects by their walking posture. D. Human action recognition using Deep Neural Network: This project was designed by Rashmi R. Koli. Tanveer I. Bagban develops hand platform in his 2020 Motion detection that recognizes hand gestures [Four].

In other words, it is human to perceive human behavior. gesture recognition. that's the only gesture body movements that convey meaning news. For this project, we used CNN Algorithm as an interpreter for interpreting gestures and Create a statement from your video. statement or text meaning of these gestures. E. Overview of Extreme Learning Machines: Extreme Learning Machines (ELMs) are one of the most important. A topic in the field of artificial intelligence in recent years. ELM is widely used for human action recognition, Multi-class classification and other fields. ELM Offer an efficient learning framework for regression, Classification, Feature Learning, and Clustering [5].

it has Much faster learning speed compared to traditional ones Support Vector Machine (SVM). In recent years, ELM Applications have surged.

3. ALGORITHMS USED:

A convolutional neural network (CNN) is a type of deep network. Neural network learning is particularly suitable for image and video recognition tasks. CNN gets inspired Through the visual system of animals and how they are processed information. CNN is automatically used and Adaptively learn spatial hierarchy of features from input picture. A CNN usually consists of an input layer. Multiple hidden and output layers. input layer This is where the image data is fed to the network. hidden things A layer consists of multiple layers of folds, Activation layer, pooling layer, normalization layer. These layers work together to extract features. Reduce the input image and its dimensions. Of A convolutional layer is the core of a CNN, A convolution operation is applied. fold The operation is Input image with a set of weights called kernels or filters Create a functional map. this process repeats several times using different filters to extract different Features from the input screen. The activation layer is Non-linear functions such as rectified linear units (ReLU) function, to the exit of the convolutional layer. Of Useful for introducing nonlinearity into the network. in short, Necessary for the network to learn complex representations Input data. The pooling layer is Spatial dimension of the feature map contributing to this It reduces the computational complexity of

the network. Finally, the output layer produces the final predictions. The output of CNN is usually set of classes. CNN has reached the cutting edge of technology Performance across a wide range of images and videos Image classification, object recognition tasks, etc. Discovery and semantic segmentation.

Support Vector Machine (SVM):

A support vector machine (SVM) is a supervised variant. Learning algorithms that can be used for classification regression task. The basic idea behind SVM is that hyperplane (either a straight line or a higher-dimensional plane) Separate data points of different classes within a feature Sky. The hyperplane that maximizes the scope, i.e. distance between hyperplane and closest datum A point for each class is chosen as the highest limit.

SVM is especially effective when the data is bad. Linearly separable, i.e. straight lines are not separable Drawn to separate classes. In such cases, the technique Using so-called kernel tricks, data high-dimensional space that becomes linear Separable. The most commonly used cores are radial Basis functions (RBF) and polynomial kernels. SVMs are can also be used for regression tasks using An optimization problem that minimizes the error.

Working:

Explain in detail how the SVM algorithm works With the following steps:

• Data preprocessing: the first step is data preparation By cleaning it, removing missing values and scaling This step is as important as SVM Sensitive to feature size.
• Kernel selection: The next step is kernel selection. function. This function maps input data to higher data A dimensional space in which the linear bounds can be found. Commonly used kernel functions are linear, polynomial and radial basis functions (RBF).
• Model training: The SVM algorithm is the best hyperplane that separates the data points into those points each class. This hyperplane is maximize the margin, i.e. The hyperplane and nearest data point for each class. of This process is known as a first-order optimization problem.
• Make predictions: Once the model is trained, Used to make predictions about new data. Input data is Map to higher dimensional space using selection kernel function. Data points are then categorized based on the following criteria: On which side of the hyperplane does it fall?
• Nonlinear separable case: when the data are nonlinear is separable, but SVM can still be used by introducing a slip Variables where some data points may be inaccurate border side.

These points are called supports vector. This process is known as Soft Margin SVM.

• use kernel tricks to map in non-separable cases Data in a higher dimensional space Linearly separable.

• Support vectors are a key element of SVM. method, and decision boundaries are complete Determined by the subset of training samples to support. vector.
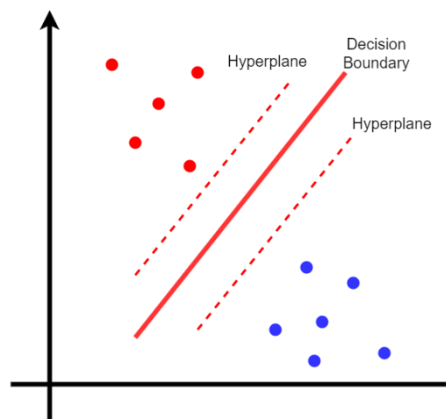
ELM algorithm:

An Extreme Learning Machine (ELM) is a type of Single Hidden Layer Feedforward Neural Network (SLFN). that is A variation of the traditional backpropagation algorithm, Used for training artificial neural networks. the most important Differences between ELM and traditional backpropagation ELM requires that the input weights and biases are random. Assigned prior to in-service training process backpropagation, they are learned through training procedure. ELM algorithms have several advantages over traditional algorithms. neural network. First, it can converge much faster than The input weights and biases are Randomly assigned, eliminating the need for multiple reservations Iterations during the training process. Second, ELM Algorithm can achieve high generalization performance Randomly assigned input weights and biases can lead to a better distribution of the input data for hidden layer. The ELM algorithm consists of the following steps:

• Input weights and biases are randomly assigned.

• training data is used for hidden layer computation Exit.

• output weights are calculated with a simple linear function regression algorithm.

• ELM model is ready for prediction. ELM algorithms can be applied to a wide variety of applications. Problems such as classification, regression, and functions approach.
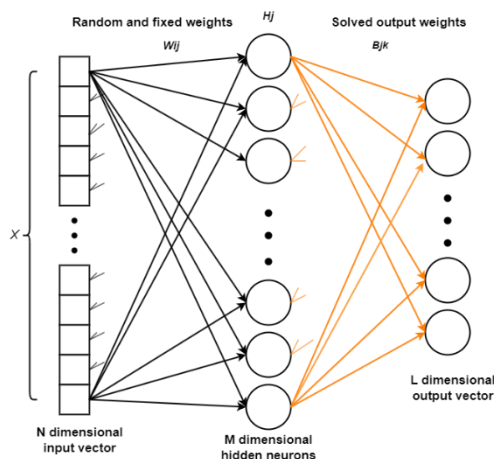
4. PROPOSED SYSTEM

1. Use SVM and ELM for action classification, ELM gives the best results. A method that can be used as a quick and reliable alternative About the method mentioned.

2. ELM has been successfully used in various fields, Ready to use despite short training period Accurate even when applied to large format video classification problem.

3. ELM is fast, efficient and economical in comparison When training large deep convolutions network. Currently used to classify pose-based methods.
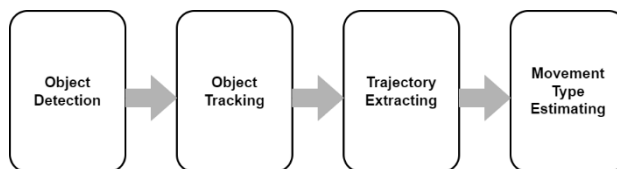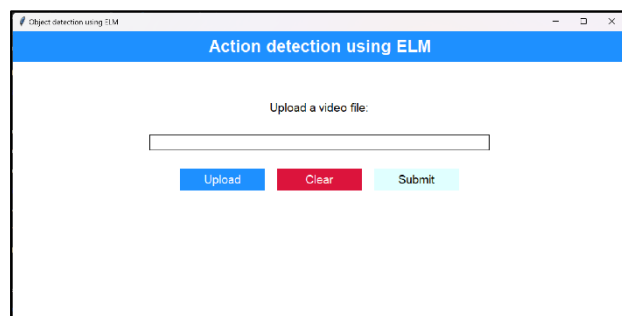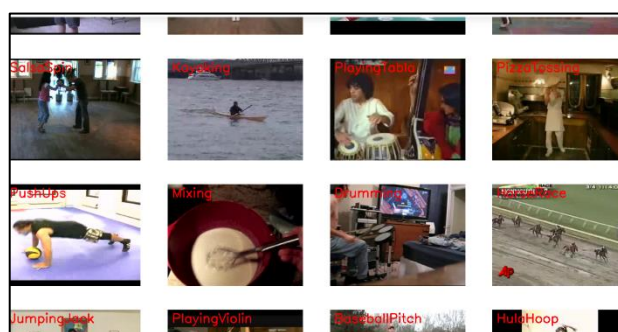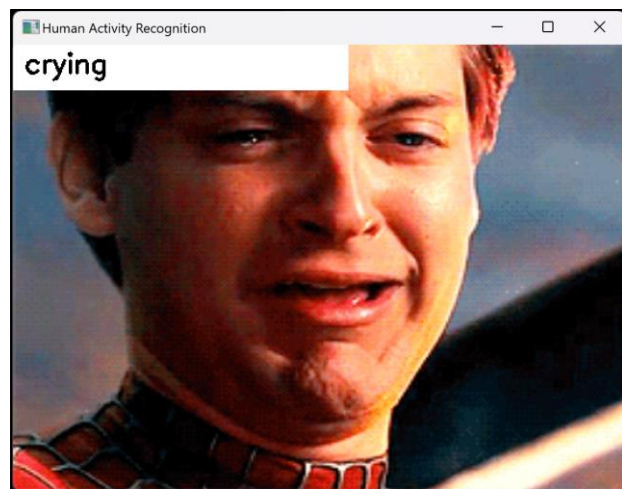

Fig.1: System Architecture

5. RESULTS

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 video analysis and human action recognition. After the emergence of neural networks, there has been a lot of research related to this topic in the past 5-6 years. The Frame-by-Frame application of CNNs helped in improving the accuracy as compared to the manual feature extraction techniques. After that, 3D-CNNs have further improved the ac- curacy of CNNs by processing multiple frames at a time. More recent architectures have started focusing on the Extreme Learning Machine (ELM) 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. Hu- man 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, viewpoint changes, high computational complexity and responsiveness to illumination changes.
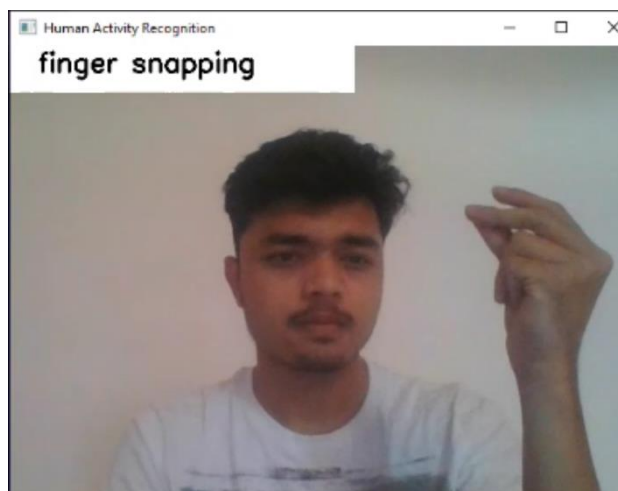
Graphical User Interface

Actual Results

Live Tracking

6. 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 the past 5-6 years. The Frame-by-Frame application of CNNs helped in improving the accuracy as compared to the manual

feature extraction techniques. After that, 3D-CNNs have further improved the accuracy of CNNs by processing multiple frames at a time. More recent architectures have started focusing on the 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, viewpoint changes, high computational complexity and responsiveness to illumination changes.

7. FUTURE SCOPE

In the future, systems can be made more precise, for example, it will make clear distinction between almost similar types of activities like StairCase 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 diseases 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 a 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.

8. REFERENCES

[1] Huang, Guang-Bin, Lei Chen, and Chee-Kheong Siew (2006). "Universal Approximation Using Incremental Constructive Feedforward Networks with Random Hidden Nodes"

[2] He, Qing, Xin Jin, Changying Du, Fuzhen Zhuang, Zhongzhi Shi (2014). "Clustering in Extreme Learning Machine Feature Space"

[3] Kasun, Liyanaarachchi Lekamalage Chamara, Yan Yang, Guang-Bin Huang, and Zhengyou Zhang (2016). "Dimension Reduction With Extreme Learning Machine"

[4] Barak, Omri; Rigotti, Mattia; and Fusi, Stefano (2013). "The Sparseness of Mixed Selectivity Neurons Controls the Generalization-Discrimination Trade-off"

[5] Kutlu, Yakup Kutlu, Apdullah Yayık, and Esen Yıldırım, and Serdar Yıldırım (2017). "LU triangularization extreme learning machine in EEG cognitive task classification".

[6] Yayık, Apdullah Yayık, Yakup Kutlu, and Gökhan Altan (2019). "Regularized HessELM and Inclined Entropy Measurement for Congestive Heart Failure Prediction"

[7]. 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.

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

[9]. 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.]

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

[11]. 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.

[12]. 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.

[13]. Jasvinder Pal Singh, Sanjeev Jain, Sakshi Arora, Uday Pratap Singh, Vision- Based Gait Recognition: A Survey, IEEE Access (Volume: 6), November 2018, pp. 70527-7049