IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

RECOGNITION OF HUMAN ACTIVITY USING MACHINE LEARNING

[1]Dr. S. Aruna, [2]T. Prapulla Naidu, [3]U. Rajasree, [4] G. Satya, [5]T. Divya Chandini, [6]K. Ashisha Vijalani

[1]Head of the Department, [2,3,4,5,6]Students
Electronics and Communication Engineering (ECE),
Andhra University College of Engineering for Women (AUCEW), Visakhapatnam, Andhra Pradesh, India

Abstract: The main purpose of this design is to achieve the certain exertion that is looking for by the system called Neural Networks which is like a heart in machine literacy. Device detectors give perceptivity into what persons are doing in real-time (walking, running, driving). Knowing the exertion of druggies allows, for case, to interact with them through an app. Artificial Neural Networks (ANNs) or Simulated Neural Networks are a subset of neural networks. Deep learning algorithms are built on the foundations of machine learning. Their name and structure are derived from the human brain, and they replicate how organic neurons communicate. The content of mortal exertion Recognition (HAR) is a prominent exploration area content in the field of computer vision and image processing area. It has empowered state- of- art operations in multiple sectors, including surveillance, digital entertainment, and medical healthcare. It is intriguing to observe and interesting to prognosticate similar kinds of movements. Several detector-grounded approaches have also been introduced to study and prognosticate mortal conditioning like accelerometer, gyroscope, etc., and it has their own advantages and disadvantages. In this paper, an intelligent mortal exertion recognition system is developed. Convolution Neural Networks (CNN) with spatiotemporal three-dimensional (3D) kernels are trained on a kinetics data set, which reflects people's actions in their daily lives and at work. As the notions of human activity recognition aid in understanding the concepts and issues of human action comprehension, which greatly aids in medication, management, learning patterns, and many circumstances involving video retrievals.

***IndexTerms* - Neural networks, ANNs, SNNs, HAR, deep learning, CNN.**

INTRODUCTION

Human activity recognition is important in human-to-human interaction and interpersonal relationships. It is difficult to interpret since it contains information on a person's identity, personality, and psychological state. One of the key topics of research in the scientific field of computer vision and machine learning is the human ability to recognize the activity of others. As a result of this research, a multiple-activity detection system is required for a variety of applications, including video surveillance systems, human-computer interfaces, and robots for the characterization of human behavior. It is difficult to construct a fully automated human activity recognition system capable of classifying a person's actions with low mistakes due to difficulties. Background clutter, partial occlusion, changes in scale, viewpoint, lighting and look, and frame resolution are some examples. Furthermore, annotating behavioral roles takes time and requires information on the specific occurrence. Furthermore, intra and inter-class commonalities make the challenge extremely difficult. That is, acts within the same class may be portrayed by various people using diverse body movements, whereas actions within classes may be difficult to differentiate due to comparable information. The way humans conduct an activity is determined by their habits, which makes determining the underlying activity challenging. Furthermore, building a visual model for learning and analyzing human actions in real-time with insufficient benchmark datasets for evaluation is difficult. Human activities are classified into three types based on their complexity:

- Gestures
- Human-to-object interaction
- Human-to-Human Interaction
- Group actions
- Behaviors
- Events

NEED OF THE STUDY

The primary goal of recognition is to:

- Implement a deep learning system to detect human behaviors from video on the document.
- To put in place a system for intelligent human activity recognition.
- To improve overall performance in order to recognize human activity

RESEARCH METHODOLOGY

Convolutional neural networks (CNNs) are a type of artificial neural network that uses a mathematical operation called convolution to analyze input having a grid-like architecture, such as photographs. Each neuron in a typical neural network is connected to all neurons inthe previous layer. Each neuron in a convolutional neural network, on the other hand, is only connected to a small local region of the preceding layer. This is known as a receptive field, and it enables the network to extract hierarchical characteristics from images. A typical convolutional neural network is made up of numerous layers, each with its own function. The first layer is often a convolutional layer that applies a collection of filters to the input image to extract features like edges.

A convolutional neural network (CNN) design typically consists of numerous layers that execute various types of operations on the input data. These layers collaborate to extract features from the input image and categorize it. Here is atypical CNN architecture:

1. Input Layer: This is the layer that accepts raw input data, which is often an image represented as a pixel value matrix.
2. Convolutional layer: This layer applies a series of learnable filters to the input image in order to extract essential properties such as edges, forms, and textures. Each filter is applied to a small local region of the input picture, and the output of each filter is sent via a non-linear activation function such as ReLU (Rectified Linear Unit).
3. Pooling layer: This layer decreases the dimensionally of the previous convolutional layer's feature maps. It accomplishes this by conducting operations on each feature map such as max pooling or average pooling, which essentially down-samples the data and makes it easier to process.
4. Remove the layer: This is an optional layer that removes a particular percentage of the neurons in the network at random during training. Driving the network to acquire more robust and generalizable characteristics, helps to prevent overfitting.
5. Fully-connected layer: This layer, like a standard neural network, connects every neuron in the previous layer to every neuron in the following one. It applies a nonlinear transformation to the input data and produces a vector of output values representing the network's prediction.
6. Output layer: This is the network's final layer, which generates the final preceding layer's output. For example, in a classification task, it might output a vector of probabilities for each possible class label.

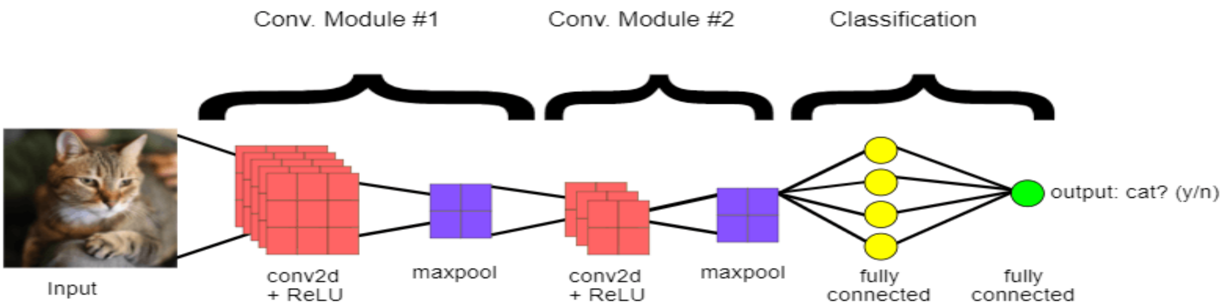


Figure 1: CNNs model

PROPOSED SYSTEM

In the future, interfaces of consumer electronics goods (for example, cell phones, games, and infotainment systems) will have more and more functionalities and be more complicated. How to develop a convenient human-machine interface (Human Machine Interaction/Interface, HMI) for each consumer electronics product has become an important issue. Using a bird's voice or videos to predict its species but this technique will not give accurate 3D results as the audio may contain background for other animal voices. Because of the mixed sounds in the surroundings, such as insects, real-world items, and so on, processing such information becomes more challenging. We now want to include some other changes which can make out properly applicable and useful soon. And applied to the new technologies and new generation.

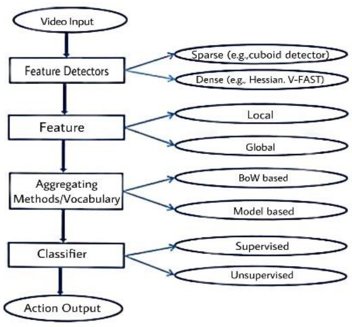


Figure 2: Input and Output flow diagram

Conceptual Design is a stage in the design process in which the broad outlines of a product's purpose and shape are articulated. Designing interactions, experiences, processes, and strategies in all parts of it. It entails understanding people's wants and how to address them through products, services, and processes. Concept sketches and models are two common conceptual design artifacts. A UML system is represented by five unique views that explain the system from distinct perspectives. A set of diagrams can be used to define each perspective. UML is specifically constructed through two different domains. They are UML analysis modeling, which focuses on the system's user model and structural model views. UML design modeling, which focuses on behavioral modeling, implementation modeling, and environment model views. A use case diagram, in its most basic form, is a representation of a user's interaction with the system that depicts the relationship between the user and the many use cases in which the user is involved.

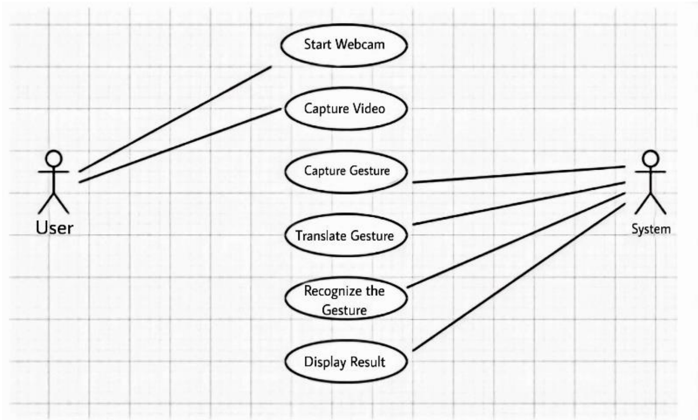


Figure 3: User case diagram

We fixed an activity class for actions based on movements in this activity detection match technique. Finally, if movement exists, they provide action that is recognized.

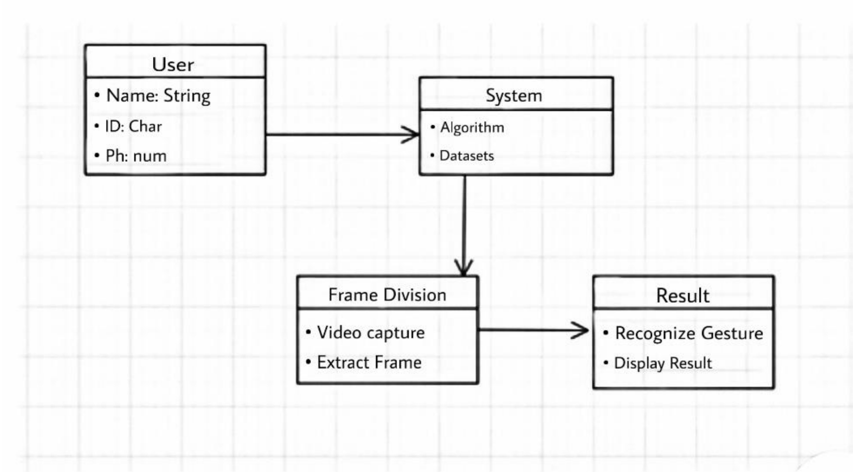


Figure 4: Class diagram

If, on the other hand, the action does not exist, the keyword class-based method is used to find it. Work Process: The work process steps from the commencement of the recognition process are listed below
Step-1: Start.
Step-2: Input video clip or live camera.
Step-3: Divide the video into frames.
Step-4: Foreground detection.
Step-5: Image enhancement.
Step-6: Image analysis.
Step-7: Action detection.
Step-8: Display the action.
Step-9: Name of the action.
Step-10: Stop.
Frame Division: It is the process of dividing a video into frames and storing the frames in a folder. We must offer video as input and get frames as output for these.
Detecting objects: The frames in the folder should provide input to a program that detects automobiles in each frame and counts them.

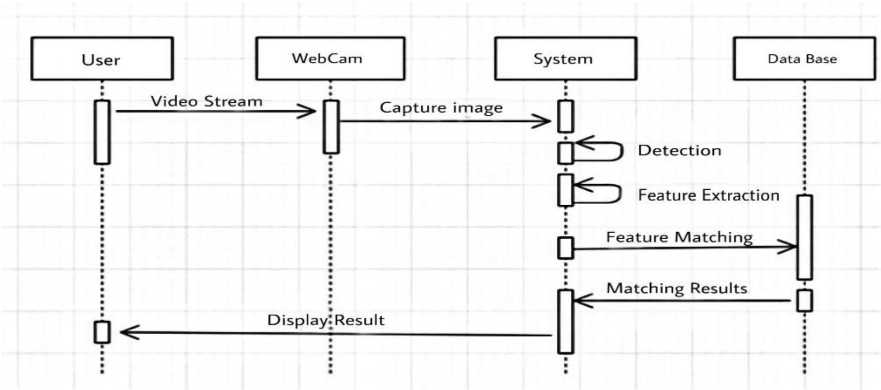


Figure 5: Process

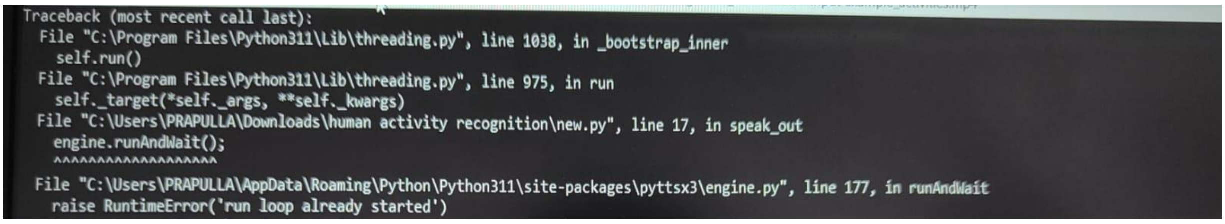
RESULTS AND DISCUSSION

Sample result 1

a. **Tuning parameters**

Parameters	Values
Algorithm	Classification algorithm
Sensors	Accelerometer, Gyroscope
Batch size	16

b. **Feature extraction**



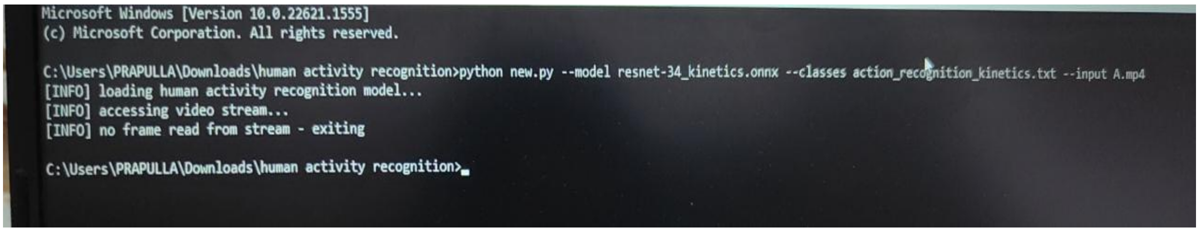
c. **Result screen**

Sample result 2

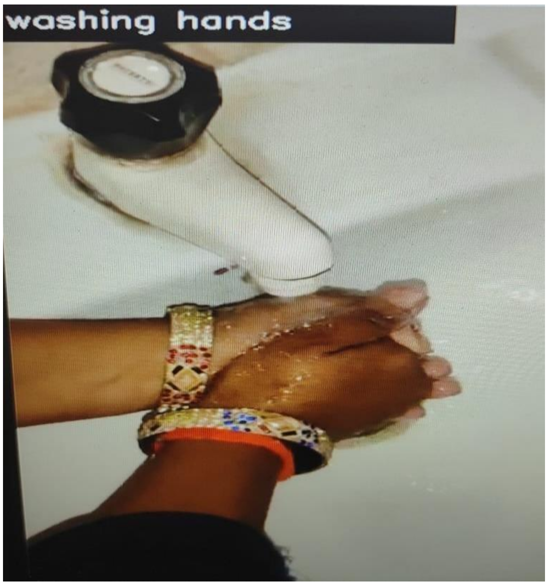
a. **Tuning parameters**

Parameters	Values
Algorithm	Classification algorithm
Sensors	Accelerometer, Gyroscope
Batch size	16

b. Feature extraction



c. Result screen



CONCLUSION

Using deep learning and neural networks a computerized method for recognizing human behavior has been suggested. The input video is read by using a convolutional neural network with a spatiotemporal three-dimensional kernel. These videos are preprocessed and a deep neural network is created in order to speed up the process. The proposed model yields promising results without any error. Human Activity Recognition (HAR) is a rapidly evolving field that has the potential to revolutionize many aspects of daily life, including healthcare, sports, entertainment, and safety. HAR systems use a combination of wearable sensors, machine learning algorithms, and signal processing techniques to recognize and classify human activities with high accuracy and reliability.

ACKNOWLEDGEMENT

We would like to express our sincere thanks to our Head of the Department (ECE) and our project guide **Dr. S. Aruna** for her encouragement. our project guide, whose supervision, inspiration, and insightful discussions have greatly aided us in completing our project. Their advice proved to be the most valuable in overcoming all obstacles in the completion of this major project **"RECOGNITION OF HUMAN ACTIVITY USING MACHINE LEARNING."**

REFERENCES

[1] Aggarwal, J. K., and Cai, Q. (1999). Human motion analysis: a review. Comput. Vis. Image Understand. 73, 428–440. doi:10.1006/cviu.1998.0744.

[2] Chen, L., Wei, H., and Ferryman, J. (2013b). A survey of human motion analysis using depth imagery. Pattern Recognit. Lett. 34, 1995–2006. doi:10.1016/j.patrec. 2013.02.006

[3] Ye, M., Zhang, Q., Wang, L., Zhu, J., Yangg, R., and Gall, J. (2013). "A survey on human motion analysis from depth data," in Time-of-Flight and Depth Imaging, Lecture Notes in Computer Science, Vol. 8200. eds M. Grzegorzek, C. Theobalt, R. Koch, and A. Kolb (Berlin Heidelberg: Springer), 149–187

[4] Atrey, P. K., Hossain, M. A., El-Saddik, A., and Kankanhalli, M. S. (2010). Multimodal fusion for multimedia analysis: a survey.

[5] Westerveld, T., de Vries, A. P., van Ballegooij, A., de Jong, F., and Hiemstra, D. (2003). A probabilistic multimedia retrieval model and its evaluation. EURASIP J. Appl. Signal Process. 2003, 186–198. doi:10.1155/S111086570321101X

[6] Fathi, A., Hodgins, J. K., and Rehg, J. M. (2012). "Social interactions: a first-person perspective," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Providence, RI), 1226–1233.

[7] Thurau, C., and Hlavac, V. (2008). "Pose primitive based human action recognition in videos or still images," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Anchorage, AK), 1–8.

[8] Candamo, J., Shreve, M., Goldgof, D. B., Sapper, D. B., and Kasturi, R. (2010). Understanding transit scenes: a survey on human behavior-recognition algorithms. IEEE Trans. Intell. Transp. Syst. 11, 206–224. doi:10.1109/TITS.2009. 2030963

[9] Turaga, P. K., Chellappa, R., Subrahmanian, V. S., and Udrea, O. (2008). Machine recognition of human activities: a survey. Proc. IEEE Trans. Circuits Syst. Video Technol. 18, 1473–1488. doi:10.1109/TCSVT.2008.2005594.

[10] Hardoon, D. R., Szedmak, S. R., and Shawe-Taylor, J. R. (2004). Canonical correlation analysis: an overview with application to learning methods. Neural Comput. 16, 2639–2664. doi:10.1162/0899766042321814

[11] Guo, G., and Lai, A. (2014). A survey on still image-based human action recognition. Pattern Recognit. 47, 3343–3361. doi: 10.1016/j.patcog.2014.04.018.
[12] Jaimes, A., and Sebe, N. (2007). "Multimodal human-computer interaction: a survey," in Computer Vision and Image Understanding, Vol. 108 (Special Issue on Vision for Human-Computer Interaction), 116–134.
[13] Pantic, M., and Rothkrantz, L. (2003). "Towards an affect-sensitive multimodal human-computer interaction," in Proc. IEEE, Special Issue on Multimodal Human-Computer Interaction, Invited Paper, Vol. 91 (IEEE), 1370–1390.
[14] Quattoni, A., Wang, S., Morency, L. P., Collins, M., and Darrell, T. (2007). Hidden conditional random fields. IEEE Trans. Pattern Anal. Mach. Intell. 29, 1848–1852. doi:10.1109/TPAMI.2007.1124
[15] Morariu, V. I., and Davis, L. S. (2011). "Multi-agent event recognition in structured scenarios," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Colorado Springs, CO), 3289–3296.