ICEST-2K24 [International Conference on Engineering, Science and Technology]
In Association with
International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN - 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)

Predictive Human Activity Recognition

Satvik Patil, Abhijith Malayil, Shruti Wadile, Mayuresh Bhombe
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, SPPU, Pune, Maharashtra, India

ABSTRACT

Human activity detection is a critical application in both computer vision and machine learning, with widespread implications across healthcare, surveillance, sports analysis, and human-computer interaction domains. This review paper offers an extensive examination of recent advancements and methodologies in human activity detection utilizing machine learning techniques. It encompasses a diverse array of approaches, ranging from traditional machine learning algorithms to sophisticated deep learning architectures and hybrid models. The exploration extends to various sensor types and modalities commonly employed for human activity recognition, including wearable sensors, cameras, and depth sensors. Additionally, the review delves into the discussion of datasets and evaluation metrics utilized to assess and compare activity detection systems. Moreover, the paper sheds light on existing challenges and outlines future research directions in the field, such as enhancing robustness to environmental factors, achieving real-time performance, and improving model interpretability. Through the dissemination of insights into cutting-edge techniques and prevailing challenges, this review endeavors to offer guidance to researchers and practitioners in advancing the development of more precise and resilient human activity detection systems leveraging machine learning methodologies.

The recognition of human activities in the field of video surveillance is attracting more researchers. This has led to various approaches and proposals using different methods and techniques. The growing interest in the surveillance has also led researchers to give importance to abnormal human activities in order to propose appropriate and dedicated techniques to this type of activities. Unfortunately, the made proposals until now in this new field are ineffective and are adapted from those dedicated for normal human activities with minor modifications. They also suffer from several limitations and inadequacies and are very restricted because of the very limited number of works and syntheses. Therefore, this paper is an overview which provides a synthesis and an analysis of the existing works on the recognition of abnormal activities in order to provide researchers with a general view on the state of the art and be a help to propose new approaches.

I. INTRODUCTION

In intelligent imaging techniques, identifying human activities from static images or video sequences is a challenging task. In computer vision, recognition of human activity is a research area that has been studied extensively in recent years. HAR can be done using smartphones, sensors or images. The main purpose behind the cognitive function is to describe the perception and behaviour of various objects based on the observation of the object's behaviour and environmental conditions. Since HAR is difficult due to problems such as complexity of the background, scale change, brightness, and partial occlusion, many deep learning-based solutions have been proposed. Walking, standing, sitting, running, waving, jumping, clapping, driving, cooking, etc. Many activities carried out by people such as. It can be analysed with HAR technology. To verify the study, information can be collected through non-contact methods, such as images and video frames, or through communication methods that use devices that people used to carry (such as accelerometers and wearable sensors). HAR can be used to monitor the elderly, and this technology is widely used in diagnostic systems . Developing new visual models to detect and learn people's movements over time is a difficult task when there is not enough measurement data . According to the type of interaction and its complexity, human activities can be classified as human-to-human, human-to-object, event, group action, gestures, atomic actions, etc. can be divided into: > Identifying the simple steps involved. Input from the camera is the first step in improving input image quality. Human pose is estimated based on joints detected by the OpenPose human skeletal model. The recognition function is then performed with the help of a DNN based on an integrated system implemented by OpenPose:

- Human Gesture: It is focused on the action of a human's face, hand, or other body parts when walking, with no requirement of verbal contact.
- Action: it is just a series of gestures performed by humans such as running or sitting or walking.
- Interaction: It is also an important aspect that incorporates individual actions to be executed by human. Interaction can be with individual or single person.

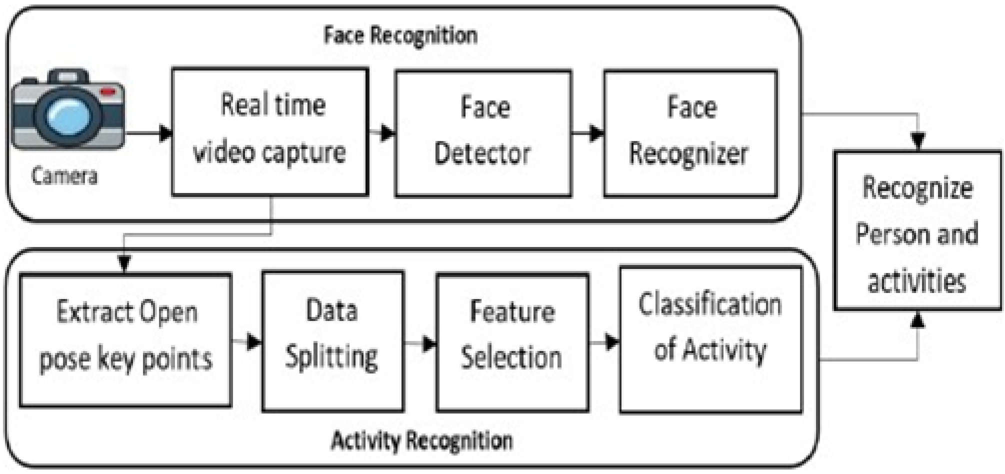


Fig.1

The activity recognition process using OpenPose involves the following steps:

Real-time pose estimation with OpenPose:

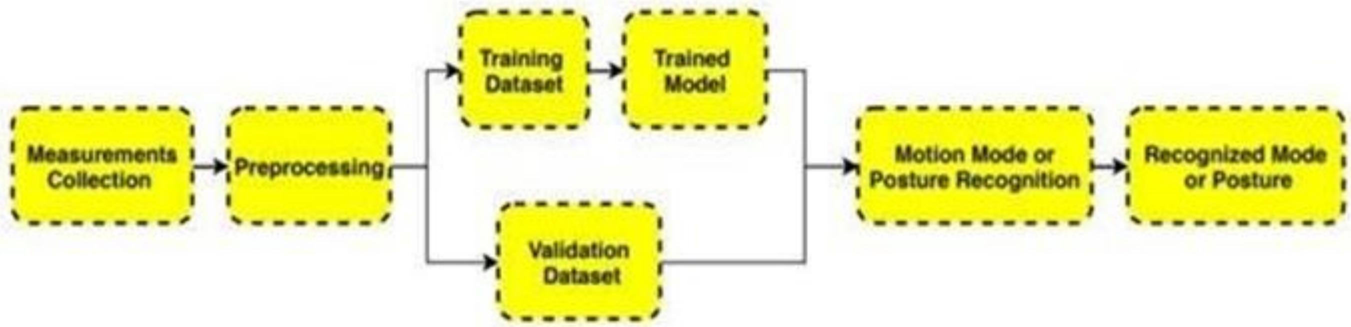
OpenPose offers a real-time system capable of detecting various key points of multiple persons within a single image. It can identify a total of 135 keypoints across different body parts, including hands, feet, body, and face. Open Pose's key point detection involves three main components: hand detection, face detection, and body and foot detection. This step begins with an input RGB image. The architecture utilizes a whole-body pose estimation network, incorporating Part Affinity Fields (PAFs) and confidence maps for face, body, and foot keypoints.The network undergoes training with a multi-task loss, amalgamating the losses from each individual keypoint annotation task. Convolutional layers are employed in model training. Bipartite matching is utilized to locate all whole-body parts belonging to the same individual, and subsequently, the whole-body poses of all individuals within the frame are generated.

II. RELATED WORK

In recent years, human activity recognition has become a prominent and dynamic area of research, attracting considerable attention across various methodologies. This technology finds widespread applications in real-time scenarios such as surveillance systems, healthcare, and human-computer interaction. For example, M. Milenkoski et al. devised a lightweight algorithm based on Long Short-Term Memory (LSTM) networks for activity detection, tailored to run efficiently on low-end devices like mobile phones. This algorithm stands out for its ease of implementation, improved accuracy, and resilience. Additionally, Zeng et al. proposed a CNN-based method for automatic extraction of discriminating features in action recognition, achieving an impressive 95% accuracy. Their approach adeptly captures scale invariance of signals and local dependencies. Experimental validation was conducted across three diverse datasets: Actitracker, covering activities like walking and jogging; Skoda, focusing on assembly line activities; and Opportunity, comprising various kitchen-related tasks.

Background Information and theory

The main goal of human action recognition in videos is to identify the action happening. This aim is achieved by analysing the frames of these videos to form and a set of data that can be classified successfully in terms of accuracy, speed, and simplicity. The main structure of human action recognition consists of two stages. human object tracking, and action classification. The first stage has to answer the question of how to detect and track the human object in each frame of the video sequences. The second stage has to answer the question of how to categorize the data from the first stage by applying an efficient classification algorithm.



Real-time human action recognition (HAR) systems using deep learning involve building complex models that can identify and classify various human actions over time based on input data such as video streams or text sensor. Deep learning, especially in convolutional neural networks (CNN) and recurrent neural networks (RNN), has made great progress in this field due to their ability to learn hierarchical and physical properties of raw materials. The project will start with data collection and then continue with the design and training of deep learning models. During training, the model learns to extract distinctive features and detect physical parameters from input data. Techniques such as transfer learning and data augmentation can improve model performance, especially when dealing with limited data. Once the model is trained, it can be deployed in a real-time environment to process incoming data and dynamically predict human performance. Integration with appropriate hardware or edge platform is required to achieve real-time performance. Continuous evaluation and improvement is important to ensure the model's robustness and adaptability to different scenarios and environments. Additionally, issues of privacy, ethics, and bias in data and predictive models must be addressed throughout the program.

III.METHODOLOGY

Human Activity Recognition (HAR) is a field of research that aims to identify activities performed by humans based on sensor data usually collected by wearable devices or environmental sensors. Below is a general guide to describing human activities:Define activities: Identify the activities you want to know about. This includes walking, running, sitting, standing, climbing stairs, etc. may include activities. This may include an accelerometer, gyroscope, magnetometer or other relevant sensor. Data can be collected from wearable devices, smartphones or environmental sensors. Size of window or short duration. This will help eliminate the features of small time. Possible features include scaling (mean, variance, etc.), frequency domain features (FFT), time domain features, and even raw materials. Characteristics: It can represent good work. Techniques such as principal component analysis (PCA) or factor analysis can be used. (k-NWS). mode. Use the validation process to optimize hyperparameters to avoid overfitting. Cross validation: Perform cross validation to ensure the robustness of the model. Find the best algorithm and architecture for the job. Since the application is required, make sure you complete it immediately or almost immediately. Replace it to have good performance.

IV. CONCLUSION

This methodology holds numerous future applications. Demonstrated through actual implementation, the approach accurately identified subjects engaging in a diverse range of activities. There is potential for further extension, encompassing larger datasets of activities, individuals, and outdoor scenarios. Additionally, addressing limitations such as pose variations, occlusions, and changes in lighting can enhance the efficacy of face recognition.

In conclusion, our study presents a comprehensive real- time human activity detection method based on Convolutional Neural Networks (CNN) using camera footage. We have developed a robust system capable of accurately identifying firearms under varying conditions, leveraging deep learning technology and advanced dataset preprocessing techniques. Our experiments in public spaces underscore the efficacy of our proposed methodology in enhancing security and safety for the broader community. Moving forward, further research and development efforts can focus on enhancing the scalability and practical applicability of the system for real-world deployment. Additionally, exploration of novel features and modalities can be pursued to enhance detection accuracy and efficiency.

Considering all aspects, our initiative marks a substantial advancement in the domain of activity detection technology, underscoring the critical importance of proactive security measures in protecting public areas. Through the utilization of real-time activity detection systems and deep learning technology, our objective is to improve the recognition of human activity in communities globally.

V. REFERENCES

[1]. Su, M.; Hayati, D.W.; Tseng, S.; Chen, J.; Wei, H. Smart Care Using a DNN-Based Approach for Activities of Daily Living (ADL) Recognition. Appl. Sci. 2020, 11, 10. [Google Scholar] [CrossRef].

[2]. C. Xu, D. Chai, J. He, X. Zhang, and S. Duan, “InnoHAR: A deep neural network for complex human activity recognition,” IEEE Access, vol. 7, pp. 9893–9902, 2019.

[3]. Wang, Q., Zhang, K. and Asghar, M. A. (2022). Skeleton-based st-gcn for human action recognition with extended skeleton graph and partitioning strategy, IEEE Access 10: 41403– 41410.

[4]. J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, “Deep learning for sensorbased activity recognition: A survey,” Pattern Recognit. Lett., vol. 119, pp. 3–11, Mar. 2019

[5]. Yu, Y.; Yang, X.; Li, H.; Luo, X.; Guo, H.; Fang, Q. Joint-level vision-based ergonomic assessment tool for construction workers. J. Constr. Eng. Manag. 2019, 145, 04019025. [Google Scholar] [CrossRef]

[6]. Mroz, S., Baddour, N., McGuirk, C., Juneau, P., Tu, A., Cheung, K. and Lemaire, E. (2021). Comparing the quality of human pose estimation with blazepose or openpose, 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART), pp. 1–4.

[7]. Gadhiya, R. and Kalani, N. (2021). Analysis of deep learning based pose estimation techniques for locating landmarks on human body parts, 2021 International Conference on Circuits, Controls and Communications (CCUBE), pp. 1– 4.

[8]. Alsawadi, M. S. and Rio, M. (2021). Skeleton- split framework using spatial temporal graph convolutional networks for action recognition, 2021 4th International Conference on Bio- Engineering for Smart Technologies (BioSMART), pp. 1–5.

[9]. P. Prabu, K. Amrutha, and J. Paulose, “Human body pose estimation and applications,” 2021.

[10]. A. Kumari, R. Gupta, and S. Tanwar, ‘‘Amalgamation of blockchain and IoT for smart cities underlying 6G communication: A comprehensive review,’’ Comput. Commun., vol. 172, pp. 102–118, Apr. 2021.