

Real Time Video Activity Detection Techniques In Machine Learning

Ms. Ashwini M. Bhugul¹

Research Scholar, Sipna COET, Amravati

Dr. Vijay S. Gulhane²

Professor, Sipna COET, Amravati

Abstract: Today, Real Time Video activity detection techniques are widely used for suspicious activity detection or for abnormal activity detection. This study paper has shed its light on different related work of real-time video activity detection technologies in the context of machine learning purposes. Moreover, this research study focused on video detection techniques such as face detection technology which is the most used technology of real time video activity detection with the help of CNN technology. In addition, it showed there are three different methods in human detection or video activity detection as Handcrafted motion features, Deep learning and Depth sensor. On the other hand, depth sensor is frequently using to develop the system of fall detection. The technique of depth sensor is help to indicate potential fall as well as Kinect sensor has been used for authenticates the form of fall alert.

Keywords: Machine Learning, Deep Learning

Date of Submission: 12-11-2021

Date of Acceptance: 28-11-2021

I. Introduction

Real-time video detection is an effective technology that is used in everyday applications in order to create a surveillance process with the help of a video camera. In the case of social interaction in real time, the video activity deception technique is used with the help of live video recording. Machine learning is another important factor in the recent business era as it is made up of *computer science* and a branch of *artificial intelligence (AI)* that helps to maintain data flow. Hence, in this paper, a piece of brief information about real time video activity detection technology has been discussed. Moreover, the main aim of this study is to identify different related work related to real time activity detection techniques in machine learning. On the other hand, several related works of real time video detection techniques in machine learning have also been included in this work with the help of some real time techniques.

This present study is going to critically discuss different techniques based on real-time video activity in the process of machine learning. On the other hand, most of the related work in the area of human activity analysis is largely focused on general classification regarding actual video-based action. Besides that, several forms of the actual video classification process have merely hacked due to the poor development of smart security. Different types of real-time techniques are highlighted in this study as they can effectively help to gather more information regarding different usable techniques in machine learning.

II. Related work

There are several similar works related to real time video activity technology used in machine learning with the help of AI and computer science. Real time video activity detection technique mainly used in the surveillance purpose of human resources also used in manufacturing and maintenance of robotic features. CNN technology is the most effective technique used in face detection with the help of CCTV footage. Therefore, *face detection* is a related work of real time video activity detection techniques in the case of machine learning [1]. Most organizations use real time video activity detection techniques to capture all necessary activity happening within an organisation. According to several studies, it showed companies mainly use real time video activity detection techniques to make business processes easier and effective.

Besides that, face detection work is helpful to detect any kind of misbehaviour or corruption happening within any organisation or institute. Real time video activity detection technique is also effective to build trust between two employees or between employees and company hierarchy as well. A *supervised convolutional neural network (S-CNN)* is an effective technique used in human detection through video play whereas it acts as a CNN feature extractor in case of human detection. Moreover, it showed *S-CNN technology* use in learning

machine management and produced an average accuracy of 95.6% with the help of SoftMax. In addition, pre-trained CNN produces an average accuracy of **98.09%** in the case of face detection technology [2]. The main related work which helps to focus on video detection techniques is the human activity recognition system that analyses and identifies the human activity with the help of various types of sensors. Different reviews of papers have been published that help to understand the multidisciplinary approach of human activity recognition to update the field of video detection in machine learning. The concept of human activity recognition mainly focuses on the movement of several parts of the personal body and enables the activity with the help of artificial intelligence. There is a need to focus on imperial identification as well as localisation of activity to boost up the understanding of events and video detection techniques [6]. Vision-based human activity recognition is a part of video detection which helps to focus on human vision activity e and understanding the concept of home monitoring as well as video surveillance. There is a need to understand the activity recognition as well as application of this system with the help of remote methods as well as contact-based systems. The concept of a contact-based system will help to generalize the knowledge of physical interaction and maintain comment acquisition with the help of video reduction techniques.

In the case of machine learning, **face detection technology** is the most used technology in the present business era. Several surveys showed in the case of both the retail and manufacturing industry all industries have focused to use real lime video detection technology that makes the payment process easier [4]. Apart from that, the use of face detection technology with the help of real time video activity detection helps to provide contact fewer services. Therefore, in the case of the Covid-19 pandemic, most of the industries have started to use machine learning features that help to maintain all covid-19 guidelines given by the country government.

However, real time video activity detection technology helps to improve the surveillance process with the help of CNN technology. Besides that, organizations or institutions use real time video activity detection techniques to capture all necessary activity happening within an organisation [5]. However, real time video activity detection techniques mainly used in the surveillance purpose of human resources are also used in the manufacturing and maintenance of robotic features. Therefore, all these are considered as the real time video activity detection technology used in machine learning.

One of the most developing works related to real-time video activity detection in machine learning is highly associated with the smart technology named artificial intelligence. On the other hand, the advanced system of modern technology such as artificial intelligence has been used in different sectors to increase their productivity rate. Besides that, the advanced technology named artificial intelligence is used for video surveillance as it specifically utilizes smart computer software programs [6]. Apart from this, smart technology is adversely beneficial in helping to evaluate both images as well as audio from the source of a video surveillance camera. The advanced security contractor's system is present in this software as it can accordingly help to define some of the major restricted areas within view of the camera.in addition, the smart function of artificial intelligence is accordingly preceded by the use of machine vision that is marked as a series of mathematical procedures or algorithms [17].

Based on the ongoing scenario of a globally competitive business market, each and individual industry are trying to develop their business construction depending on smart technologies. Besides that, some of the retail as well as other technology-based industries have adapted the techniques of real-time video activity to protect their delivery procedure. For example, "**behavioral analytics**" is marked as the adopted security that has been recently developed in different industrial activities from the generation of manufacturing processes to the area of delivering the service [7]. Besides that, big data applications are effectively consuming most of the space in the research area as well as industrial activity. In addition, among the standard examples of big data, the critical function of video streams from the generation of CCTV cameras is usually the same as other sources including medical data and agricultural data. Along with this, sensor data, as well as social media data, are effectively evolved from the research space [8].

Deep neural networks with deep learning are usually active with a much wider range of accuracy in the process of computer vision tasks rather than the general algorithm of classic machine learning. Due to the higher demand for both storage resources as well as computation, deep neural networks are promotionally often deployed in the cloud system [15]. Based on the ongoing global era, it can be evaluated that most of the related work is accordingly moving with the deep neural network [6]. As a result, the DNN is remarkably close to the end of data sources as it alternatively connects with the edge computing paradigm. According to the machine learning system, it is effectively marked as a standard approach in terms of identifying some of the critical issues in delivering or purchasing. Besides that, the lack of the area regarding open-source framework is associated with high-level API as it accordingly complicates the form of deployment. The deployment procedure regarding deep learning-enabled processes helps to deliver systematic work at the source of internet edge. On the other hand, Edge Eye is accordingly to help to deliver high level as well as task-specific API which is actively useful in terms of developers as they can easily concern themselves solely on a specific application logic [19]. Edge Eye is specifically helpful to enable developers to transform models through training with some

of the popular deep learning frameworks [7]. The chosen machine learning approach is actively beneficial in respect to creating advantages to deployable elements with the source of minimal effort. As a result, it can effectively leverage the area of optimized inference engines as those are belonging to different industries in order to achieve optimized interference efficiency and performance.

III. Study of real-time technique

Real time video activity recognition has become an important research area nowadays as it plays an important role in machine learning and robotics as well. Several methods help to detect different challenges faced by companies while using real time video detection technology for surveillance processes or machine learning processes as well. In the case of robotics AI technology is mainly used with the help of real time video activity detection techniques. The main methods that are used in real time video activity detection are **Handcrafted motion features**, **Deep learning** and **Depth sensor**. All these methods are helpful to detect human action via real time video activity detection technology by using a branch of AI and computer science [1]. All these methods hold both advantages and disadvantages in case of detecting video activity through AI and real time video detection technology as well.

Handcrafted motion features:

Handcrafted motion features are mainly used with CNN technology that helps to detect any kind of defects through video detection techniques. On the other hand, motion features provide **temporal information** via optical flow to obtain the best performance of any organisation related to human surveillance. In the case of machine learning, AI is the most effective technology that includes a motion feature that helps to record every small activity from the target substances [3]. This method also helps to collect features from the vocabulary of visual features in order to focus on the movement along with video recording.

Besides that, video recording helps to collect different data about the employees or the entire activity that happens during the maintenance of machine learning. Through this method, only motion-based approaches can be analysed which is considered as a disadvantage of this particular method. Apart from that, in order to maintain particular approaches in video recognition, many resources need to be used that make the entire process effective and innovative as well.

In addition, these methods run with the help of CNN technology which is an **artificial neural network** that is mainly used for face recognition. **Convolutional Neural Network (CNN)** technology is costly as compared to other **Artificial intelligence (AI)** technologies such as deep learning computer vision and others [3]. Hence, by using this handcrafted motion feature method operation costs started to increase that have an adverse impact on the performance of an organisation.

A large number of layers is used in this handcrafted motion learning model by which activity CNN technology started to increase and it created a positive impact in human surveillance process and machine learning as well. Apart from that, these CNN layers make the entire activity of this technique critical for new users and sometimes workers do not understand how to operate entire methods successfully in case of video detection [5]. Hence, the face detection technique is the most used technique of detecting real time video activity techniques in a machine learning activity.

Depth sensors:

The core use of depth sensors is effectively useful in terms of facilitating the prototyping of different ubiquitous computing applications regarding interactive space. On the other hand, the appropriate use of machine learning algorithms has effectively received real-time sensory input that is specifically input from the generation of three-axis accelerometers [9]. Along with this, gyroscopes are also strategically placed in the area of the human body as it is effectively useful to detect fall in progress.

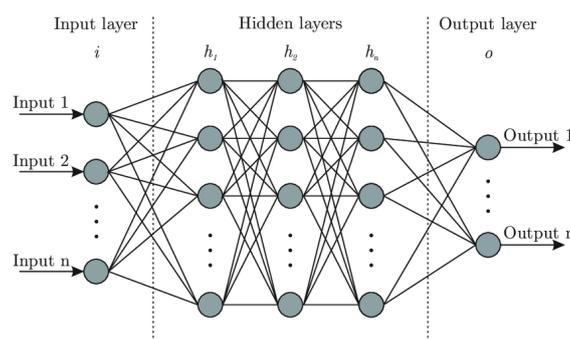


Figure 1: Artificial Neural Network
(Source: Liu et al. 2019)

According to the global business report, it can be easily determined that most of the top organizations have faced a serious issue of inactivity detection as it is specifically associated as a leading cause for slow down business activity. The serious issue has been generated through core detection in the area of smart video surveillance [9]. Due to solving the core issue, these types of applicants need a standard performance based on real-time face activity detection. In this specific area, depth sensors have played a fundamental role as it is usually very time-consuming in the case of detecting the form of actual activity [10]. Apart from this, depth sensors are adversely useful in terms of developing fall detection as it is marked as a potential cause to harm older people. As a result, depth sensors are a suitably appropriate method as it is specifically assigned a considerable demand regarding the lowest cost of fall detection systems. Based on the ongoing worldwide working scenario, advancements in the area of depth imaging technologies have been effectively made in a form named *human activity recognition (HAR)*.

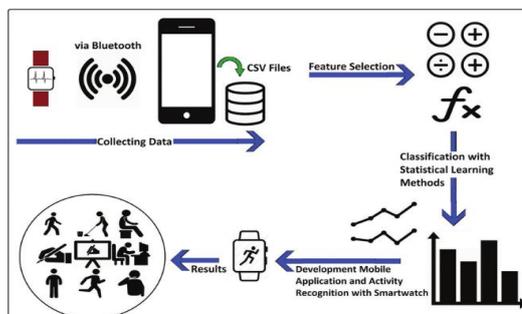


Figure 1: HAR
(Source: Qi and Aliverti, 2020)

The HAR is mostly relatable without attaching any optical markers as well as there is also no need for other quality motion sensors to the area of human body parts. On the other hand, real-time tracking of different working activities, orientation as well as the location of multiple people is raising the major importance regarding designing advanced computing applications based on novel ubiquitous. Top view cameras based on tracking specifically help to avoid the risk of occlusion while tracking people collaborating with the uses of smart technology of machine learning. although, it mostly required a complex tracking system as well as also needed advanced computer vision algorithms [10]. According to the ongoing smart technological scenario, depth sensors are routinely used in motion tracking and gesture recognition. Hence, it is specifically examined that the advanced real-time video activity recognition technique is beneficial for *advanced human-machine interfaces*. On the other hand, navigation systems, object counting, multi-point level sensing are also remarkably derived by the smart technical uses of depth sensors that are regulated through machine learning [18].

Deep learning:

Machine learning libraries and python-based computer vision need to be exploited in terms of improvement as well as an experiment on the ultimate uses of depth sensor technique. In addition, human activity is mostly placed in two major operations on a specific video [11]. On the other hand, for deep learning approaches, activity recognition from different videos in several CNN depending approaches is largely found. Along with this, most of the CNN-dependent approaches drastically developed the ultimate rate of standard efficiency in terms of detecting human activities. As a result, the core concern is remarkably derived through integrating the scope of adaptive comparison that is specifically associated with deep learning approaches. In addition, the technique of deep learning is effectively beneficial for video activity detection as it is specifically marked as a suitable technique to develop standard efficiency in the recognition of real-time activity [12].

Several ways of using deep learning are largely associated with text generation, computer vision, image caption generation, as well as the area of consumer experience. On the other hand, machine learning is effectively used through multiple forms in the ongoing business as it can specifically help to enhance the source of consumer experience. Besides that, deep learning techniques are placed through the smart representation learning methods that are widely associated with different levels of a video representation. Apart from this, the technique of deep learning is specifically obtained through composing simple although non-linear modules. *Backpropagation* is effectively listed as the most useful technique that is adversely used in the activity area of deep learning. Along with this, in the case of deep learning, the *black prop technique* or *Backpropagation* is ideally referred to as the term of central mechanism in respect to neural networks [13]. The chosen technique is moderately beneficial as it creates neural networks that are associated with learning about some of the

addressing errors which are present in data prediction. Most of the common applications of deep learning are speech recognition and image recognition as it is most effective for feature introspection [14]. Hence, it can be identified from the above discussion that deep learning is one of the most useful methods to manually video detection. In addition, it is also addressing that deep learning possibly shines while it comes to the form of most complex issues. For example, speech recognition, natural language processing as well as image classification are easily solved due to the uses of the machine learning approach by a significant form of deep learning technique.

IV. Conclusion

The entire part of this present study is concluded about different techniques that have been used in real-time video activity detection in the specific area of the machine learning process. Human detection in smart videos leads a major role in the area of different real-life applications. One of the most useful traditional approaches is largely depending on specific features through utilizing handcraft. The core concept of proposed features regarding machine learning approaches is quite easier as well as mostly cheaper as a result most of the industrial working processes are specifically regulated through face detection techniques. The present study is specifically highlighting similar works that are moderately associated with the form of real-time video activity technology. Based on the evaluation of ongoing working activity in a different industrial environment, it can be concluded that *a supervised convolutional neural network (S-CNN)* is an active technique that moderately using for human face detection for video activity detection. Apart from this, this present study is to help to analyses that *Handcrafted motion features, Depth sensors* as well as *Deep learning* are marked as the major methods which are frequently used in the area of real-time video activity detection.

Reference list:

- [1]. Haritaoglu, I., Harwood, D. and Davis, L.S. "W/sup 4: real-time surveillance of people and their activities". *IEEE Transactions on pattern analysis and machine intelligence*, 22(8), 2017, pp.809-830.
- [2]. Gaglio, S., Re, G.L. and Morana, M. Human activity recognition process using 3-D posture data. *IEEE Transactions on Human-Machine Systems*, 45(5), 2018, pp.586-597.
- [3]. Tango, F. and Botta, M. "Real-time detection system of driver distraction using machine learning". *IEEE Transactions on Intelligent Transportation Systems*, 14(2), 2017, pp.894-905.L. von Rueden et al., "Informed Machine Learning - A Taxonomy and Survey of Integrating Prior Knowledge into Learning Systems," in *IEEE Transactions on Knowledge and Data Engineering*. doi: 10.1109/TKDE.2021.3079836.
- [4]. N. Vaswani, A. K. Roy-Chowdhury and R. Chellappa, "Shape Activity": a continuous-state HMM for moving/deforming shapes with application to abnormal activity detection," in *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1603-1616, Oct. 2005, doi: 10.1109/TIP.2005.852197
- [5]. G. A. Tahir and C. K. Loo, "An Open-Ended Continual Learning for Food Recognition Using Class Incremental Extreme Learning Machines," in *IEEE Access*, vol. 8, 2020, pp. 82328-82346, doi: 10.1109/ACCESS.2020.2991810
- [6]. Liu, Y., Bi, S., Shi, Z. and Hanzo, L. "When machine learning meets big data: A wireless communication perspective". *IEEE Vehicular Technology Magazine*, 15(1), 2019, pp.63-72.
- [7]. Chelli, A. and Pätzold, M. "A machine learning approach for fall detection and daily living activity recognition". *IEEE Access*, 7, 2019, pp.38670-38687.
- [8]. Kumar, K., Shrimankar, D.D. and Singh, N, April. "Event bagging: A novel event summarization approach in multiview surveillance videos". In *2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC) 2017*, (pp. 106-111). IEEE.
- [9]. Qi, W., Su, H. and Aliverti, A. "A smartphone-based adaptive recognition and real-time monitoring system for human activities". *IEEE Transactions on Human-Machine Systems*, 50(5), 2020, pp.414-423.
- [10]. Sehgal, A. and Kehtarnavaz, N. A "convolutional neural network smartphone app for real-time voice activity detection". *IEEE Access*, 6, 2018, pp.9017-9026.
- [11]. Volkov, M., Hashimoto, D.A., Rosman, G., Meireles, O.R. and Rus, D, May. "Machine learning and coresets for automated real-time video segmentation of laparoscopic and robot-assisted surgery. In *2017 IEEE international conference on robotics and automation (ICRA)* " 2017, (pp. 754-759). IEEE.
- [12]. Tran, D., Do, H.M., Sheng, W., Bai, H. and Chowdhary, G. "Real-time detection of distracted driving based on deep learning". *IET Intelligent Transport Systems*, 12(10), 2018, pp.1210-1219.
- [13]. Volkov, M., Hashimoto, D.A., Rosman, G., Meireles, O.R. and Rus, D, May. "Machine learning and coresets for automated real-time video segmentation of laparoscopic and robot-assisted surgery". In *2017 IEEE international conference on robotics and automation (ICRA)* 2017, (pp. 754-759). IEEE.
- [14]. Xu, R., Nikouei, S.Y., Chen, Y., Polunchenko, A., Song, S., Deng, C. and Faughnan, T.R., May. Real-time human objects tracking for smart surveillance at the edge. In *2018 IEEE International Conference on Communications (ICC) 2018*, (pp. 1-6). IEEE.
- [15]. Zhang, X., Yu, Q. and Yu, H. "Physics inspired methods for crowd video surveillance and analysis: a survey". *IEEE Access*, 6, 2018, pp.66816-66830.
- [16]. Hsieh, Y.Z. and Jeng, Y.L. "Development of home intelligent fall detection IoT system based on feedback optical flow convolutional neural network". *Ieee Access*, 6, 2017, pp.6048-6057.
- [17]. Chelli, A. and Pätzold, M. "A machine learning approach for fall detection and daily living activity recognition". *IEEE Access*, 7, 2019, pp.38670-38687.
- [18]. Güera, D. and Delp, E.J., November. "Deepfake video detection using recurrent neural networks". In *2018 15th IEEE international conference on advanced video and signal based surveillance (AVSS) 2018*, (pp. 1-6). IEEE.
- [19]. Deeba, F., Ahmed, A., Memon, H., Dharejo, F.A. and Ghaffar, A. "LBPH-based enhanced real-time face recognition". *Int Adv Comput Sci Appl*, 2019, 10(5), pp.274-280.