Journal of Advanced Research in Image Processing and Applications
Volume 6, Print Issue 1 - 2023, Pg. No. 9-15
Peer Reviewed Journal

Review Article

Nature-Inspired Multi-Stage Event Detection Model with Optimized Feature-Based Learning for Video Datasets

Amar Pal Singh

Student C.S Department Chandigharh University, Lucknow, Uttar Pradesh, India.

I N F O

E-mail Id:
ap19567@gmail.com
Orcid Id:
https://orcid.org/0009-0002-4771-9509
How to cite this article:
Singh AP. Nature-Inspired Multi-Stage Event Detection Model with Optimized Feature-Based Learning for Video Datasets *J Adv Res Image Proc Appl* 2023; 6(1): 9-15.

Date of Submission: 2023-03-23
Date of Acceptance: 2023-04-25

A B S T R A C T

Visual tracking is a crucial aspect of computer vision, with applications in various fields like robotics, video surveillance, human-computer interaction, autonomous cars, sports analytics. It involves estimating the state of an item being tracked, which is a challenging task due to non-visibility changes. The object tracking algorithm is effective due to its use of features such as Grayscale, Gradient, Texture, Color, Fusion. Current studies focus on identifying events in real-time, sports, traffic, natural catastrophes, more. The most researched topics in computer vision include object popularity and localization in 3-D devices. Visible tracking aims to adjust an object in any environment, either by monitoring a single item or tracking multiple objects simultaneously. The EFS-linear MSVM approach is proposed for identifying multiple events within video sequences. The research aims to provide an appropriate approach for feature selection and classification strategy for identifying multiple events in YouTube videos. The EFS method is recommended for selecting the most useful feature subsets from the extracted vectors, the MSVM is provided with the best feature subsets for multi-class classification. Euclidean distance is used to obtain relevant events and activities. Item tracking is used to accurately interpret an object's motion in a video.

Keywords: Visual Tracking, Event Detection, Video Event Detection, Object Tracking, Video Stream, Video Event

Introduction

'Object tracking' refers to the process of identifying or "tracking an object (or numerous objects) through a succession of photographs." This might refer to a single object or multiple items. The human eye has a tough time tracking any one item in real-time video streaming most of the time.[1] One of the most important senses for humans to use in order to extract information and meaning from their surroundings is the eye. According to the data, the sense of sight is responsible for the perception of roughly 80 percent of the ambient basics or components (such as color, brightness, form, movement, depth, etc.). In such a circumstance, cutting-edge technology known as computer vision (CV) proposes to equip computer systems with the capability to "see the world" like people, therefore resolving a difficult task that was previously assigned to humans. CV makes use of cameras to simulate the characteristic of the human eye, which recognizes the capabilities of extraction, popularity, monitoring of the item.[2] [Cameras] [Cameras] [Cameras] [Cameras] [In point of fact, visual object tracking is an essential component of many different computer vision systems. These systems include those

used for intelligent medical treatment, human-computer interaction, video analysis of unmanned aerial vehicle footage, dynamic behavior recognition, military navigation. Visual object tracking is a significant area of research in the field of computer vision. It is used in a broad variety of situations that include many types of input footage, such as a photo, a video, a real-time video, or a pre-recorded video. Some examples of these kinds of situations include: Both the categorization and the applications of object tracking are impacted by the kind of input.[1]

Figure 1.Object Tracking

The number of different videos that may be seen on the Internet has been growing at a fast rate. This expansion has been strongly driven by the act of publishing or sharing videos on social media platforms and on YouTube. As more people view videos online, it is becoming more crucial to be able to search for and retrieve movies. Today, in computer vision (CV) and multimedia uses, automated detection (AD) of compound actions in online videos is obtaining rising research consideration from the analyzers.[1] CV stands for computer vision, multimedia uses make use of both. Monitoring human activity as well as other video events is supported by the VED (video event detection).[12] It is useful in a variety of fields, including the military, public safety, law enforcement, among others.[3]

System for the Detection of Events and Videos in Events

The videos are either captured directly from the internet or accessed there, they include audio and visual information. The film is composed of multiple ongoing sequences of occurrences that are referred to as "Events." This technique, which is known as detection,[1] involves extracting some important information from movies that have been taken or found on the internet. In the topic of study known as event detection, several techniques, approaches, methodologies are used in order to identify the events that occur in films.[3] The event detection models extract contextual information about the occurrences from the videos in order to make their detections. The feature phase, the semantic phase, the preceding phase for event detection are the three stages that make up the contextual information or the usable data, respectively.[4] The dynamic information and the vast audience that may be found in sports game footage is explored. For the purpose of automatically detecting occurrences,[25] the contextual information and happenings in game recordings are utilized as input to various models and algorithms.

The Importance of Classification in the Process of Detecting Events in Video

In the field of video event identification, self-determination theory is used to do research on people's underlying motivations. According to the self-determination theory, a learner's motivation value, as demonstrated by the factors influencing their performance, is more important than their total level of motivation for predicting the intended learning outcome and level of processing. This is because the self-determination theory is based on the assumption that people have the ability to control their own behaviors. Presented a spectrum of self-determination ranging from fundamental motivation to motivation, with distinct types of extrinsic motivation (i.e., mixed, assessed, interjected, periphery) between that vary in the degree of qualified autonomy.[29] This spectrum extends from fundamental motivation to motivation.

Inspiration and the Role of Significance

The sport of cricket is gaining popularity across the globe, but computerized cricket video analysis is tough to do because of the many different metrics and the intricacy of the game. The video event detection provides a lot of different characteristics that are significant. This should motivate academics to participate in surveys so that video research may be improved.[25] The topic of video event detection is touched on just fleetingly in the context of event detection-based video. Video Event Systems is the provider of automated video analysis techniques and the implementations of such technologies. A low-level analysis is performed in order to sequentially segment the video into shots or play-break chunks.[30] This allows for the automatic extraction of key events. Categorize the video material and pull out the events that are semantically meaningful by using a system that detects events. Various techniques for detecting shot boundaries, classifying shots, identifying events all play significant roles in the process of event detection and have vital applications in a variety of professions, including medicine, sports, other areas.[15]

Literature Survey

The two most essential study subjects in modern computer vision technology are object popularity and video localisation. Both of these problems are becoming more relevant. The topics of item representation, detection, education, categorization all fall under the umbrella term of "object popularity." The bulk of the study focuses on the popularity of certain locations, which is different from the

localization of specific items. The popularity of things in three-dimensional devices, as well as their identification, present various challenges, such as the monitoring of the regions in which they may be extracted. The real-time component of three-dimensional object detection has been the subject of a great deal of investigation from a variety of research perspectives. The detection technique is at the center of the strategy, which is how it differentiates itself from other approaches. The many strategies that have been offered for detecting the object are mostly focused on visible aspects of its appearance. In order to extract distinctive descriptors from the body of a video or image, geometrical approaches are utilized.

An algorithm was presented by SusmithaAlamuru et al. (2021)[8] that can differentiate between events and movies. The histogram equalization approach was used on the movies in order to increase the contrast, the database UCF101 was used to flatten them. Without affecting the appearance of the frame in any way, histogram equalization was used to maintain the image's high quality. The speed of the training model was increased as a consequence of the features extraction process, which used a variety of methods to extract the feature vectors from the resulting structure. It was recommended that a selection method should be developed, extracted feature vectors were quickly measured in order to discover the best characteristics in the extracted frames. Multiple support vector machines were used in order to classify and separate the various occurrences based on their most salient characteristics (MSVM). With the help of the suggested method, the recognition of a variety of events taken from the frame was able to achieve the greatest possible performance, with a mean accuracy of 14.21% and a mean precision of 12.92%.

According to Bokun Wang et al. (2022),[12] the goal of security and safety may be found in a variety of settings, including hospitals, shopping malls, banks, airports, stations, streets, so on. There was a widespread implementation of various video surveillance techniques. The identification of anomalies in the video was a significant challenge for computer vision. A variety of strategies for detecting anomalies, all based on DL, for extracting features to use in training the system. The approach suggested deep neural network (DNN) hybrid with the support vector data description (SVDD), which was utilized to recognize the tiny sections of the data for modeling the SVDD and the retrieved data features. In the Avenue and Shanghai Tech datasets, the suggested technique achieved improved performance, with frame level AUC values of 86.84% and 73.20%, respectively.

Using the Random Forest (RF) approach, Xianjun Xia et al.(2018)[14] reported audio event identification as an application of the method. Taking into account the relevant facts and the Bottleneck characteristics. The areas of auditory event detection, the variety of event kinds, the edge information of events were investigated, although not in great depth. An extensive amount of time was spent with audio movements while the creation of a number of different pieces of frames was deployed with contextual event information. The bottleneck and global features suggested discriminated characteristics of the event, the database hired UPC-TALP and ITC-IRST and acquired the performance of the supplied model. The overall error rate for both approaches was 5.30%, with 4.44% being the difference between the two.

In their study published in 2010,[16] Vasileios Chasanis et al. offered event identification and sorting in the video event sequences. The contextual information was resolved by constructing the video frames using local invariant forms. This allowed for the films to be separated into a variety of modules. A visual word histogram was used in order to do the evaluation of the distance between the different modules of the films. The Machine Learning methods that were contracted were supposed to sort all of the events into their respective categories. According to the findings of the experiments, the suggested strategy achieved higher rates of event sorting and detection.

Table 1. (a) contains a summary of the relevant research papers that can be found in the event detection field. These research papers are based on a variety of machine learning techniques, within these techniques, multiple methods, models, parameters are defined for specific detection purposes. Various characteristics are taken into consideration in order to enhance the performance of the proposed model and draw attention to issues and gaps in the study domain.

Table 1.(a) Analysis of Machine Learning approaches in Video event detection

Author Name and Year	Proposed Model	Parameters	
		Accuracy	Precision
SusmithaAlamuru et al. (2021)[8]	EFS-MSVM		
Bokun Wang et al.(2022)[12]	SVDD	*	*
Xianjun Xia et al.(2018)[14]	Global and bottleneck feature method	*	*
VasileiosChasanis et al. (2010)[16]	Dynamic Time Warping(DTW), Support Vector Machine.		*

Author Name/Year	Pre-processing/ Segmentation	Classification Model	Description	Dataset	Parameters
SusmithaAlamuru et al. (2021)[8]	Hybrid feature extraction Histogram equalization	EFS-linear MSVM algorithm	Multiple detection events from video sequence	CCV and UCF101 databases	Accuracy, Errorrate Sensitivity, Specificity FOR, FDR
Bokung et al. 2022[12]	Nil	DSVDD	It is useful data representation and models can be learned.	Avenue ShanghaiTec public datasets	Frame level AUC ROC
Xia et al. 2018[14]	Nil	RF classifier	Classify based on acoustic event detection	UPC-TALP ITC-IRST	Absolute error
Vasileios et al. 2010[16]	Background substraction	SVM	To cluster and classify the events	Nil	High detection rate

Table 1.(b) gives a literature overview of the research articles that are based on different machine learning techniques. These research papers investigate the event detection area and specify diverse methodologies, tools, and databases for specific detection goals. Various characteristics are taken into consideration in order to enhance the performance of the proposed model and draw attention to issues and gaps in the study domain.

According to Chien-Hao Tseng et al.(2021),[11] the use of video detection is becoming an increasingly essential component in each sector for the goal of safeguarding the safety of public areas like airports, stations, malls, other businesses that are analogous to these types of locations. Due to the fact that the CCTV cameras collected many movies as data, the process of identifying the person of interest from the multiple camera system was a challenging and tough task. The fact that the videos had a wide variety of frame backgrounds and an erratic sequencing of those backdrops made it much more difficult to spot suspicious activity utilizing numerous cameras. The recommended FCN (Fully Convolutional Network) model was able to resolve all of those impurities in the multi-cameras and gain higher performance for person recovery from those multi-cameras thanks to the Deep Learning (DL) method that is based on instance segmentation.

MD Roman Bhuiyan et al. (2022)[13] indicated the crowd locations that should be used for event detection, the Hajj crowd was classified for the purposes of detection. It was proposed that crowd sensing may be used as a method for protecting public spaces and ensuring the safety of individuals. The widespread use of automatically deploying groups by means of computer vision led to the discovery that existing algorithms were unable to recognize video instances in crowded areas, such as those that exist during the Hajj pilgrimage. This was a significant limitation of the technology at the time. The intended research would concentrate on a number of different features of crowds, such as enumeration, analysis, thickness, other characteristics of crowds, among other crowd-related issues. The fundamental objective of this research was to use DL techniques in order to accomplish the task of visually identifying the hajj. Grigorios Tsagkatakis and colleagues (2017)[15] offered a concept for the event identification of intermittent video feeds in athletic venues. It was absolutely important to propose an event identification approach that was based on the score of football games because of the level of complexity, irregularity, lack of static information that was available in the video feeds. The proposed model was successful in eliminating all of these problems that were seen in the video sequences since it made use of a DL approach. For the training of the spatial and temporal data, the model used CNN, autoencoders were employed for the synthesis of feature information with a normalized output. They used the new dataset as well as the probable properties of the recommended model in order to conduct research and experiments on the hybridization of regular and irregular video sequences. The findings of the evaluation revealed that the proposed model was successful in attaining a greater degree of accuracy performance.

When collecting events from social media, ImadAfyouni et al.(2022)[17] suggested that the major emphasis should be placed on natural occurrences that are the consequence of environmental circumstances, weather, etc. The unique technique that was provided made use of a considerable

amount of data gathered from social media in order to hybridize the social events that were being held. The procedures of extraction and clustering, respectively, were carried out with the aid of supervised and unsupervised deep learning techniques, which were merged to build the hybrid model. These approaches were used to carry out the processes. The proposed machine learning model, which made use of natural language programming techniques, was responsible for handling the event detection process. The tool known as KeyBERT was used throughout the process of developing the datasets in addition to its role in the development of the words. CNN's categorization of the incident, as well as the fact that it occurred in both directions The locations of events that are identified by spatial clusters may be accessed via the use of long-term, short-term memory, which is abbreviated as LSTM. The hybrid technique that was recommended recognized and monitored features that were taken from social media for the purpose of doing a Spatio-temporal analysis of the real world. Table 2 (a). provides a literature evaluation of the research publications that are based on various deep-learning approaches. This review is relevant to the field of event detection. These publications provide a wide variety of strategies, models, parameters that have been tailored to distinct detection objectives. The performance of the suggested model is improved by taking into account a variety of features, flaws and holes in the research area are brought to light as a result.

Table 2.2 (b) gives a literature overview of research publications that are based on different deep-learning techniques. These papers specify diverse methodologies, tools, databases for specific detection goals within the event detection discipline. Various characteristics are taken into consideration in order to enhance the performance of the proposed model and draw attention to issues and gaps in the study domain.

The researchers now working on VE recognition developed a method to help enhance it. In most cases, the CCV and UCF101 datasets[22] are the ones that are used in order to gather the input video. After that, a method for improving frame differences and filtering the video known as Histogram Equalization (HE) was devised and used in the newly generated movies. Histograms of the images are essential to the functioning of this method. After the contrast of the frames has been brought up to par, a hybrid type of feature extraction called FE (feature extraction) using GLTP (gradient local term pattern), HoG (histogram of oriented gradients), Tamura features is carried out. This is done in order to extract the feature sets. After that, the EFS (ensemble feature selection) method is used in order to choose the most advantageous feature sets. EFS-MSVM[8] is used to categorize a variety of occurrences in order to

Author Name and Year	Proposed Model	Parameters	
		Accuracy	Precision
Chien-Hao Tseng et al.(2021)[11]	Multi-CNN classifier FCN model		*
MD Roman Bhuiyan et al.(2022)[13]	Context-CNN, Multitask-CNN, Scale-CNN, CNN	*	*
GrigoriosTsagkatakis et al.(2017)[15]	Spatial and Temporal methods used with CNN		*

Table 2.(b) Analysis of Deep Learning techniques in Video event detection

Author Name/Year	Pre-processing/ Segmentation	Classification Model	Description	Dataset	Parameters
Chien et al.(2021) [11]	Pixelwise Instance Segmentation	FCNN	Multi-CNN classifier is used to extract the presence attribute feature sets of a human masked image to represent a person.	DukeMTMCreID	Classification rate Accuracy Loss
Bhuiyan et al. 2022 [13]	Foreground Segmentation	PCANet CNN methods	Crowd classify base on hajj video	PASCAL COCO VOC	Nil
Grigorios et al. 2017 [15]	Spatial feature	CNN VGGNet	Pretrained CNNs for extracting both spatial and temporal features	YouTube.com	Classification accuracy

facilitate the recovery of dependable actions (Euclidean distance). The one-of-a-kind EFS-MSVM model was used in the simulation examination in the many manifestations of error, accuracy, specificity, sensitivity, amongst others. The EFS-MSVM model obtained an accuracy rate of 89.0% and a MAP (mean average precision) of 78.09% on the CCV dataset. On the UCF101 dataset, it produced an accuracy rate of 93.7% and a MAP (mean average precision) of 84.7%.

The Problematization of the Issue

The automated categorization of complicated behaviours in internet videos is obtaining increased research interest among investigators.[18] This classification method has many applications in computer vision (CV) and multimedia. The VED assists in the presentation of human and other visual actions that are described in the movies,[19] where it is used in fields such as the army, commercial, security, LE (law enforcement), other fields.[20] At the moment, machine learning techniques are quite advanced in the intelligence VA (video analysis). Because automation helps to reduce performance difficulties like expensive and frees security personnel from daily jobs,[21] the investigators focus on high-level intellectual tasks that better utilize the capabilities of the adopted approach. This is because automation helps to free security personnel from daily jobs. On VED, a number of suggested works are exhibited; these works are all cantered on the goal of finding answers to difficulties such as the simultaneous occurrence of several activities, bottlenecks, the large quantity of data that has to be processed. On the other hand, the approaches that came before have shown some effectiveness when used in controlled environments.

Conclusion

The future of object detection appears to be highly promising in a variety of settings and applications. We are proud to have played a role in the development of real-time intelligent vision, which was made possible by high-performance computing, artificial intelligence, machine learning, all of which helped us solve real-time challenges successfully. We will discuss the many uses of this technology, the necessity of computer vision, how and why things are identified, much more. The difficulty that arises as a result of noise, occlusion, split error is attempted to be solved in this piece of work by using a strategy for tracking several objects. To solve the problem, this was done. The goal of visible tracking is to reposition any object in any setting, given the past function and length of the target item. One of two methods—monitoring a single object or simultaneously tracking numerous objects—can be used to achieve the goal. In this work, a method for recognising numerous events inside video sequences is proposed using the EFS-linear MSVM methodology. The main goal of this study is to offer a suitable method for choosing characteristics and a classification strategy for identifying numerous events in YouTube videos. Histogram equalisation enhances the frames before GLTP, HOG, Tamura features are used to extract the feature vectors, which are frequently of a higher quality. The best feature subsets from among the extracted feature vectors should be chosen using an efficient feature selection technique called EFS. The best feature sub-sets are then given to the MSVM to perform multi-class classification (20 action scenarios on each database). In the end, the relevant events and activities are found using the Euclidean distance. As a result, object tracking is used to precisely identify the item in a video and interpret it as a set of trajectories. The appearance property of objects that function at various resolution levels can be used to overcome the split error and occlusion problems. The appearance model is created with a high level of resolution to better distinguish between people.

References

1. Jhuo, I., Guangnan Ye, ShenghuaGao, Dong Liu, Yu-Gang Jiang, D. T. Lee, Shih-Fu Chang. “Discovering joint audio–visual codewords for video event detection.” Machine vision and applications 25, no. 1 (2014): 33-47.
2. Yu, Junqing, Aiping Lei, Yangliu Hu. “Soccer video event detection based on deep learning.” In International Conference on Multimedia Modeling, pp. 377-389. Springer, Cham, 2019.
3. Lim, Mei Kuan, Szeling Tang, Chee Seng Chan. “iSurveillance: Intelligent framework for multiple events detection in surveillance videos.” Expert Systems with Applications 41, no. 10 (2014): 4704-4715.
4. Wang, Xiaoyang, QiangJi. “Video event recognition with deep hierarchical context model.” In Proceedings of the IEEE Conference on computer vision and pattern recognition, pp. 4418-4427. 2015.
5. Xu, D., & Chang, S. F. “Visual event recognition in news video using kernel methods with multi-level temporal alignment.” In 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8. IEEE, 2007..
6. Aggarwal, J. K., &Ryoo, M. S. “Human activity analysis: A review.” Acm Computing Surveys (Csur) 43, no. 3 (2011): 1-43..
7. Li, H., Achim, A., & Bull, D. «Unsupervised video anomaly detection using feature clustering.» IET signal processing 6, no. 5 (2012): 521-533.
8. Alamuru, S., & Jain, S. «Video event detection, classification and retrieval using ensemble feature selection.» Cluster Computing 24, no. 4 (2021): 2995-3010.
9. Luo, M., Chang, X., & Gong, C. «Reliable shot identification for complex event detection via visual-semantic embedding.» Computer Vision and Image Understanding 213 (2021): 103300.
10. Karbalaie, A., Abtahi, F., &Sjöström, M. “Event detection

in surveillance videos: a review." Multimedia Tools and Applications (2022): 1-39.
11. Tseng, C. H., Hsieh, C. C., Jwo, D. J., Wu, J. H., Sheu, R. K., & Chen, L. C. "Person Retrieval in Video Surveillance Using Deep Learning–Based Instance Segmentation." Journal of Sensors 2021 (2021).
12. Wang, B., Yang, C., & Chen, Y. "Detection Anomaly in Video Based on Deep Support Vector Data Description." Computational Intelligence and Neuroscience 2022 (2022).
13. Bhuiyan, M. R., Abdullah, J., Hashim, N., & Al Farid, F."Video analytics using deep learning for crowd analysis: a review." Multimedia Tools and Applications (2022): 1-28.
14. Xia, X., Togneri, R., Sohel, F., & Huang, D. «Random forest classification based acoustic event detection utilizing contextual-information and bottleneck features.» Pattern Recognition 81 (2018): 1-13.
15. Tsagkatakis, G., Jaber, M., &Tsakalides, P. «Goal!! event detection in sports video.» Electronic Imaging 2017, no. 16 (2017): 15-20.
16. Chasanis, V., &Likas, A. "Event detection and classification in video surveillance sequences." In Hellenic Conference on Artificial Intelligence, pp. 309-314. Springer, Berlin, Heidelberg, 2010.
17. Afyouni, I., Al Aghbari, Z., & Razack, R. A. "Multi-feature, multi-modal, multi-source social event detection: A comprehensive survey." Information Fusion 79 (2022): 279-308.
18. Jhuo, I.H., Ye, G., Gao, S., Liu, D., Jiang, Y.G., Lee, D.T., Chang, S.F «Discovering joint audio–visual codewords for video event detection.» Machine vision and applications 25, no. 1 (2014): 33-47.
19. Yu, J., Lei, A., Hu, Y "Soccer video event detection based on deep learning." In International Conference on Multimedia Modeling, pp. 377-389. Springer, Cham, 2019.
20. Lim, M.K., Tang, S., Chan, C.S, "iSurveillance: Intelligent framework for multiple events detection in surveillance videos." Expert Systems with Applications 41, no. 10 (2014): 4704-4715.
21. Xu, J., Denman, S., Reddy, V., Fookes, C., Sridharan, S, «Real-time video event detection in crowded scenes using MPEG derived features: A multiple instance learning approach.» Pattern Recognition Letters 44 (2014): 113-125.
22. Wang, T., Qiao, M., Zhu, A., Shan, G., Snoussi, H, "Abnormal event detection via the analysis of multi-frame optical flow information." Frontiers of Computer Science 14, no. 2 (2020): 304-313.
23. P. Over, G. Awad, M. Michel, J. Fiscus, G. Sanders, W. Kraaij, A. F. Smeaton, G. Quenot. "Trecvid 2013--an overview of the goals, tasks, data, evaluation mechanisms and metrics." (2013).
24. Lai, K. T., Yu, F. X., Chen, M. S., & Chang, S. F. "Video event detection by inferring temporal instance labels." In Proceedings of the ieee conference on computer vision and pattern recognition, pp. 2243-2250. 2014.
25. Jia, Y. "Automatic Detection Algorithm of Football Events in Videos." Computational Intelligence and Neuroscience 2022 (2022).
26. Holder, R. P., & Tapamo, J. R. "Improved gradient local ternary patterns for facial expression recognition." EURASIP Journal on Image and Video Processing 2017, no. 1 (2017): 1-15.
27. Song, Q., Jiang, H., & Liu, J. "Feature selection based on FDA and F-score for multi-class classification." Expert Systems with Applications 81 (2017): 22-27.
28. Choi, S., & Jiang, Z. "Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique." Computers in biology and medicine 40, no. 1 (2010): 8-20.
29. Wijnia, Lisette, Martine Baars. "The role of motivational profiles in learning problem-solving and self-assessment skills with video modeling examples." Instructional Science 49, no. 1 (2021): 67-107.
30. Raval, Khushali R., Mahesh M. Goyani. "A survey on event detection based video summarization for cricket." Multimedia Tools and Applications (2022): 1-29.