Irish Interdisciplinary Journal of Science & Research (IIJSR)
Vol.4, Iss.3, Pages 66-78, July-September 2020

Prediction and Classification of Fouls in Soccer Game using Deep Learning

T.Thamaraimanalan[1], D.Naveena[2], M.Ramya[3] & M.Madhubala[4]

[1]*Assistant professor, Sri Eshwar College of Engineering, Coimbatore. Email: thamaraimanalan.t@sece.ac.in*
[2]*Sri Eshwar College of Engineering, Coimbatore, India. Email: naveenadurairaj1999@gmail.com*
[3]*Sri Eshwar College of Engineering, Coimbatore, India. Email: ramyaraghav372919@gmail.com*
[4]*Sri Eshwar College of Engineering, Coimbatore, India. Email: madhubalam969@gmail.com*

Article Received: 11 June 2020 Article Accepted: 09 August 2020 Article Published: 06 September 2020

ABSTRACT

Background: *Event detection within the games in the challenging task in the field of object detection system. Event of the certain games represents the various foul factors, player action as well as the object tracking. The proposed work implies the tracking of the football which is followed by the players tracking that implies in the action of the event within the football game. Various games can be analyzed using the deep learning methods which yield the discovery of fouls and penalties in certain games in Olympics. This can lead towards change in results in games with the accurate prediction.*

Method: *The object detection within the football games depends upon the basic image processing techniques. The image acquisition stage involves the collection of database as in the video format or in the live data stream. The pre-processing stage is carried out using the background subtraction technique using the binary conversion technique. Histogram optical flow and Local Binary pattern is carried out for the segmentation of movements of players as well as ball tracking. Features are extracted and stored from the segmentation process. Training process is carried out using Convolutional Neural Network (CNN) classifier.*

Results: *The trained and the tested phase of the proposed work maintain the accuracy rate of 87.63% with the reduction in the sensitivity and the specificity range of 72.5 and 86.2%. The experimental results are carried out using the Python domain which uses the tensorflow platform that yields the lesser error rate and higher accuracy of classification.*

Keywords: *Image acquisition, Background subtraction technique, Optical flow and Local binary pattern.*

1. Introduction

In our daily lives, computers are an integral part. Although it is productivity, the web is accessed for gathering data, or even for playing games. In the past few periods, computer growth had contributed to great strides and development in fields such as Mathematics, Finance, and Biology. Meanwhile the arrival of Machine learning, computers have been able to process further like the quality of humans such as image, form, speech, and even in visual object recognition [19]. The accessibility of a huge quantity of information in Machine learning is the key reason for success. In this field, the research had been prosperous due to the recent abundance of information. So that it contributes towards algorithmic and computing performance. The machine had been allowed to examine and study from this data.

Computer Vision is one of the most fascinating areas of Machine Learning. It includes the derivation of a high-level understanding of low-level raw pixel data [40]. To study this field is more difficult due to the complication in an understanding of what these pixels represent. For example: Given a picture of a sunset and asked to explain the image, a human, below usual circumstances will be capable of quickly recognizing the sunset. A machine, though, is not quite so talented. The raw pixel data is processed by the image but the machine does not know what this data signifies. Developments in Computer Vision helps to bridge the gap between the machine's and the human brain's abilities to interpret visual signs. Computer Vision has recently been in advance power. Image classification and object detection are the subdivision of Computer Vision which is being changed to fields such as engineering and medicine. In the automotive industry, the study has been proposed using both for an application like autonomous cars with favor to pedestrian detection [42] and road detection [25]. Video games in Computer Vision are another interesting area. A study published in 2015 proposed that with the current beginning of the e-sports industry and growth of inexpensive gaming, live streaming of sports will rise by 140.6% among 2014 – 2019 [29]. The study also forecasts the growth concerning the viewing audience for sports around 305.5% within the same period [29]. This increased evolution of competitive tournaments opens new paths for Computer Vision exploration.

Examples of possible benefits of this type of examination include: gathering information on player habits to find faults or enhancements, development of more complex game AI to assist players to learn faster, or executing ideal stability updates. Computer Vision research in sports is presently lacking. There are few studies which seem to attention specifically on image classification combined with reinforcement learning or Q-learning and producing agents that can play the game [18], [26]. Further subdivisions of Computer Vision such as object detection don't seem to have been familiar by the gaming community as a useful area of study. By considering this, we decided to push the partial examination done in this area by applying object detection to video games.

For our study, we chose to concentrate on object detection as a total with numerous core classification perceptions. We applied a real-time detection model to organize and track video game characters from the popular fighting game Super Smash Brothers Melee. By this, we constructed a basic bot, capable of measure based on tracked locations of a secondary character on screen. This bot was accomplished using labeled frame data from the game and successfully achieve basic actions in real-time.

2. Related Works

Till the detection of deep learning, sports video observation, particularly in football video analysis, had been classified into two group pattern recognition [21, 9] and object tracking. In case of object tracking personalized camera was used results in estimation cost, however the low-level events were extracting by pattern recognition and then high-level events were detected by the classifier.

Qian et al [17] proposed a limited method, that is used with clear success for soccer movement recognition contains classification of events into a discrete category like the goal, shoot, and so on. This approach contains heuristic rules and feature extraction for spotting actions. They accomplish simple event analysis to detect symbols (arcs, goalmouth, fields, lines, and logo), ball and player position etc., and it originates between complex event and simple event features using the signs.

Finally, they built up a standard based framework to distinguish applicable events like the corner, the goal, and so on. Jin et al [10] proposed a Hidden Markov-based calculation for detection of video event dependent on signs blend and mix. In soccer video analysis, the detection higher-level from low-level events is highly difficult problem.

The exposure reveals, e.g., the movements of the ball and the player on the pitch, which might be useful to classify definite movements ("passing the ball", “shot on goal”, etc.) or too well known the overall trend of the game. Later in 2012, a deep learning approach such as Restricted Boltzmann Machines and Convolution Neural Networks have been efficiently used for activity and event detection. Convolution Neural Networks have exposed superior performance in object detection, modeling high-level and image classification pictorial semantics [6], [11], [8]; Temporal dynamics video is exhibiting by the recurrent neural network [12].

Normally utilized activity confinement procedures, for example, fast r-Convolution Neural Network and faster r-Convolution Neural Network [18], [7], for the most part start with the area of interest (proposition age) to get a lot of up-and-comer locales, around then utilize a completely related layer toward the conclusion to classifying goals.

Recent approaches declared above concentrate on recognition of events in football game from the perception of classifiers, models, and feature extraction for extracting simple events. Such methods lack in meaningful representation of in between events. Introducing semantic explanation and essential information in these approaches is quite problematic. Thus, this urges us to begin from the essential structure squares and recreate a framework that grants abusing the semantic information around event, which can be used to distinguish the high-level and intermediate type of events.

As far as we could possibly know, however there are frameworks that consequently recognize essential evidences, correspondingly the position and movements of the player, there are no programmed pointers for semantically composite event, such as scoring on a corner kick or penalty kick.

3. Proposed Work

3.1 Video Events

An exact ontological description of the event is an open fact. The idea of this section is to define our event detection in precise manner. There are so many events were conducting but the video events is the most interesting one. For example, a football game is video event. In this "Pass the ball" is one of the events in which the distance among player1 and the ball rises and the distance among the player2 and the ball fall.

3.1.1 Low-Level Events

A low-level event type is defined as follows:

$$SE\ \left(I, simpleEventType, t, (role_1, oType_i), \ldots., (role_n, oType_n\,)\right) \qquad (1)$$

Where, *I* is the attribute which represent the event, simpleEventT*ype* is the type of event, e.g. "*throwing the ball*", and *t* is the duration of time within which the event take place, $role_1$..., $role_n$ ($n = 1, n_{max}$) are the task that dissimilar objects play in type of event. For example, "throwing the ball" is a subject who throws the ball is first task and second task is thrown ball; at last $oType_i$ is the legal type of object that can play the task $role_i$, e.g., it is only balls can be thrown and only players who can throw. Examine, the entire definition of the type of event "*throwing the ball*" is

$$(I, Throwing\_the\_ball, t, \left(throwing_{Player}, player\right), (throwing\_object, ball)) \qquad (2)$$

A particular instance of a low-level type of event defined in (1) the following tuple:

$$(I, seType, t, (role_1, O_1), \ldots., (role_n, O_n) \tag{3}$$

Where, *I* is the attribute which represent the event, O_1 and O_n are objects identifier observed in the frame related with the duration of time *t*, correspondingly.

$$(12, Throwing\_the\_ball, t, (throwing\_Player, obj02), throwed\_Object, obj01)) \tag{4}$$

The event of "Throwing_the_ball" defines a low-level type of event "throwing the ball" that occurred at time duration t, anywhere the obj01 is thrown by obj02. Also, obj01 and obj02 are two objects in the structure were detected corresponding to duration of time t, of type player and ball in that order.

3.1.2 High-Level Events

High-level events are built for the purpose of more accuracy. High-level events are detected by the simple events. We can use temporal and logical operators to detect high-level event. The hierarchy events are definite by low-level event to high-level event. In the following we have two types of complex event one is logical event and another one is temporal event.

3.1.3 Logical events

A logical event originates from the function of logical operators such as NOT, OR, AND to a set of events which might be complex or simple.

$$LCE = (I, complexEventType, t, L =< e_1\ op\ e_2\ op \ldots..\ ope_n >) \tag{5}$$

Where, *I* is the attribute which represent the event, *complexEventType* is the high-level type of event (such as "The goal is valid only if there is no foul"), *t* is the duration of time in which the high-level type of event exists, *L* is the lower-level set of events e_1. . . . . . e_n joined by op logical operators (*i.e. NOT, AND, OR*).

3.1.4 Temporal events

A temporal event stems from the application of temporal operation than as follows:

$$TCE(I, complexEventType, t, L =< e_1\ THEN\ e_2 \ldots.THEN\ e_n >) \tag{6}$$

Where, *I* is the attribute which represent the event, *complexEventType* is the high-level type of event (such as “player 1 passes the ball to player 2"), *t* is the duration of time, *L* is the arrangement of simple events, e_1, e_n that should follow in the order. For precedent, e_1, e_2, e_3, e_4 may be, correspondingly, “player1 possesses the ball", “player1 kicks the ball", “the ball approaches player 2", “player2 gets in possession of the ball".

4. Event Types

The mainly important factors in analyzing a soccer game is the capacity to identify the goal, kicks, passes, offside, cards, ball possession and many other events in a video. Till date the event recognition in Soccer field is done by using immovable cameras to capture the movements of players and the ball. By using this method, the accuracy of movement tracking in the soccer field has enhanced to greater levels but this method is too costly. The video clip that is used can also be accessed from Internet. Here we are explaining a tiny event which is the most complex event in soccer game which are in continuous frames, for instance the ball possession and the player stroking the ball found on the distance among the bounding boxes and the rules which is a mix of logical and temporal operations is set for each and every event.

In this process, the event is defined by a set of rules and regulations, but this will not be completely flexible and on further up gradation of this event detection may have a possibility of detecting the events automatically by using deep learning techniques.

4.1 Ball possession Event

The two terms TBP and PBP are combined known as ball possession. TBP means team ball possession and PBP means player ball possession. These two possessions have same starting point but different ending point. When the ball is in play it determined the ball possession by the intervals of time. Player ball possession starts when a player starts to kick the ball at initial point and it ends when the ball halt in the field.

PBP can be said as "An event takes place when the space between the ball and player is not beyond the threshold value and the ball is close to that player".

$$(I, PlayerBallPossession, t+\bar{k}, (PossPlayer, p_i), (PossObject, b)) \leftarrow$$

$$player(p_i), ball(b), D(p_i, b, t) < T_h,$$

$$\forall j \neq i, player(p_j), D(p_j, b, t) > D(p_i, b, t)^\wedge$$

$$\forall k = 1 \ldots . \bar{k}, D(p_i, b, t+k) \approx 0$$

The PBP takes place at the time interval $t+\bar{k}$ and when the distance D between the ball b and the player p_i at time t(p_i, b, t) is below the threshold T_h and D(p_j, b, t) is the distance among any one player p_j and the ball b is more than D(p_i, b, t). Also, the communication between the players and the ball is not always the same; it will be less after an action up to certain number of continuous frames. The threshold value T_h, is an experimental calculation of the physical interaction among the player and the ball.

4.2 Kicking the ball Event

In a video clip of soccer which is a continuous arrangement of frames, the type ofevent regarding to the Identifying the ball after kicking the ball, and the distance among the player and the ball vary continuously. At the beginning, the distance among the player and the ball is less and then later the distance among the player and the ball is more for certain number of frames and after certain period of time the interaction between the player and the ball can no longer happen and so we can define the type of event "kicking the ball" as follows:

$$(I, KickingTheBall, t+\bar{k}, (KickingPlayer, p_i), (KickedObject, b)) \leftarrow$$

$$player(p_i), ball(b), D(p_i, b, t) < T_h \;^\wedge$$

$$\forall k = 0 \dots \bar{k} - 1, D(p_i, b, t + k) < D(p_i, b, t + k + 1)$$

The above expression can hold as long as the distance among the player and the ball raises after the player interact with the ball. In soccer game there are many types of kicks. They are Cornor kick, Penalty kick, Goal kick, Free Kick, and so on.

4.3 Restrictions of the proposed work

When we evaluating a game, there can be many fouls happening throughout the match but we are considering only the major fouls. In certain cases, the contact between the player and the ball may not be there but the player runs nearing the ball to reach the ball. The interaction among the player and the ball happens throughout the match. In other cases, when the player and the ball are near but the ball is behind the player i.e. the player is standing showing his backside to the ball. For a superior understanding among the different types of kicks, we can identify it by the velocity at which the ball travels after the interaction among the ball and the player. After dribbling a ball, it will be slower when compared to the ball after a kick/shoot. We too taking into account of the threshold level of every players by considering about each and every players profile linked to their usual communication with the ball.

5. Methodology

The block diagram of the proposed architecture shows the method of work. The data consist of approximately three minutes long video, which consists of about 6200 frames. Using SSD players and the soccer ball are detected from each single frame. After that unwanted parts of the frames in the input were filtered out because they are not defined about events. In filtering process they keep the data around the bounding boxes using logical and temporal operators as a result we detected the fouls.

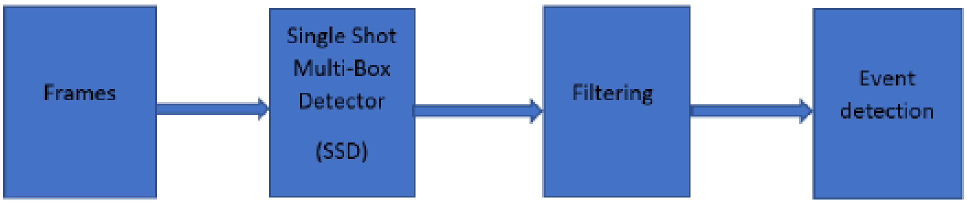


Fig.1 Block diagram of the proposed architecture

FRAMES: Frames are nothing but it is a sequence of events running as video. Each and every frame consists of bounding box set. It represents the dimension and position of the object. To detect the object first the sequence of frames were given as input to the SSD and coordinates of the object were known by the bounding boxes with confident score.

SSD: SSD stands for "**Single Shot Multi-Box Detector".** In early application one method is common for object detection that is the system provides a part of region which re-samples the pixels and bounding box provided with features followed by classifier. It detects the event but it is very expansive for low frame rate in computation process. There are another method in detection of object is High-speed single shot multi-box detector. It contains two main tasks one is classifier and another one is region proposal. The main plan behind single shot multi-box detector is s a narrow convolution filter is applied to meet the bounding boxes.

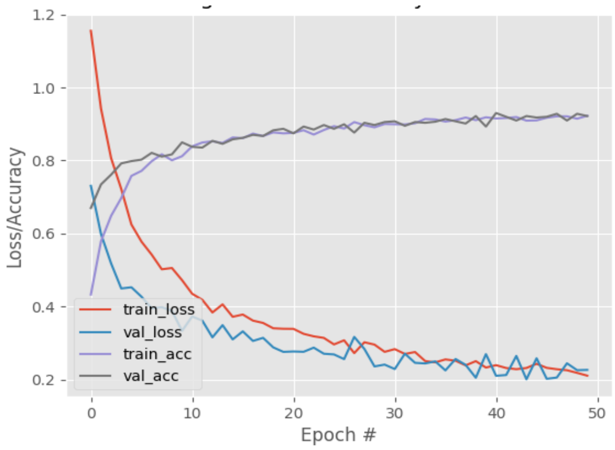


Fig.2 Performance accuracy plots

SSD requires an image that should be supplied as input and the information about the ground for each object class while training the frames. For training the images, first the models of a real soccer match video are required. Vatic Animation tool allows annotating of players and ball inside each frame drawing a bounding box around the players. The outcome of this method is a group of pictures with coordinates of bounding boxes stored in Pascal VOC layout. The performance accuracy plot is shown in fig.2 and the comparison of existing algorithm is shown in fig.3. Table 1 shows the manually annoted numbers of an object, and it is used for the training and test of single shot multi-box detector.

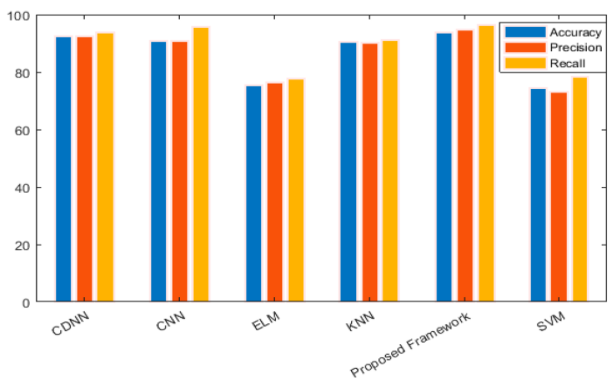


Fig.3 Comparisons with existing algorithms

Table 1. Manually Annotated Object to Train and Test the SSD

	Player	Ball	Flag	Player Movements
Training	1235	529	125	15001
Test	867	318	76	11440
Result	2102	847	201	26441

Irish Interdisciplinary Journal of Science & Research (IIJSR)
Vol.4, Iss.3, Pages 66-78, July-September 2020

In Table 1 the training set has been used to form a SSD representation. In the below Table 2 the test set of average precision value is given. For instance from the soccer match the input and the output appearance to SSD is made known in Figure 4 and Figure 5, correspondingly.

Filtration: Filtration can be done by the particular threshold value of the object detected from the single shot multi-box detector. For example, in a single frame multiple players were detected but we required only a particular threshold values which defines about the events other than that we can filter out.

Table 2. Average Precision of the System

	FP	FN
TP	2	2
TN	189	150

Detection of Event: In a real-time application, we have some detection regarding video events. In that case there was sequence of frames running over a time period. We cannot study about all the frames running behind the video clips and it is highly complex. This methodology is going to have a solution for problem. The system can detect the events by this process and there are two main steps have been done. One is the detection and another one is a collection of simple low-level events. By keeping this as a base we can also detected the highly-complex events. The system also detects the type of event.

We can also feed the new upcoming events with unique event type identifier. The events can be monitored by two levels, high-level events and simple low-level events. The simple low-level events recognize the simple event and high-level events recognize the complex events.

The detection of event by the system revives bounding boxes with confident score as input initially it process the input to identify which type of events based on the rules and it stores the event in the memory. The input bounding boxes nothing but it defines about the object coordinates. After the event detection we use temporal and logical operators to know the event is high-level event. There are so many programming languages are there for event detection but we have taken python language because it is general purpose syntax.

Fig.4 Original frame

Irish Interdisciplinary Journal of Science & Research (IIJSR)
Vol.4, Iss.3, Pages 66-78, July-September 2020

Fig.5 Detection of object by SSD with confidence score

6. Experimental Results

The proposed systems are used to detect the low-level complex occurrence in the real football video such as kicking the ball and ball possession. The work has been tested by 3 minutes video consists of approximately 6.2k frames. While experiment that process we know that comparing to other techniques convolution neural network detects few more number of events in a small dataset and in future we will use larger dataset to increase the number of events. Table 3 shows the results of detecting events. In a table for kicking the ball event totally 2o events were considered, 13 out of 15 events were successfully detected but 3 events were missed because when the player kick the ball it reaches another player less than 3 frames. For ball possession event we have taken 10 events, 9 events were detected correctly. In missing cases the 2 players are very close to one another. Event detection happen when the players were met in an appropriate number of repeated frames. For instance, if the distance among the player and the ball is less than 3 frames we go for event type as ball possession.

Table 3. Results for Detection of Event

Detected Events	**Total**	**Missed Detection**	**Accuracy**
Ball Possession	10	9	92%
Kicking the ball	15	13	81%
Pass the ball	14	11	84%
Shot on goal	8	7	93%

While keeping this as a base we also detected the highly complex events like pass the ball and shot on the goal. As a result, we merge the simple low-level events and high-level events using temporal and logical operators. To classify the highly composite event, we have considered different cases at various abstraction levels. Pass the ball events are detected when simple low-level events occur and it can use for referring to the players of the same team. For example, consider player 1 and player 2 of the same team. The player 1 pass the ball to the player 2 as referring both players are same team this detection had been done. This can be achieved by the jersey color of the players. When we obtained a successive sequence of frame, event detection starts, once the Player1 owns the ball is recognized it calculate the distance among the ball and the player1 if it is low for a small number of frames after that distance among the ball and the player1 raises in number of succeeding frames it checks for simple low-level Kicking the ball events. When the distance between the ball and the player2 decreases it is highly complex "Pass the ball events".

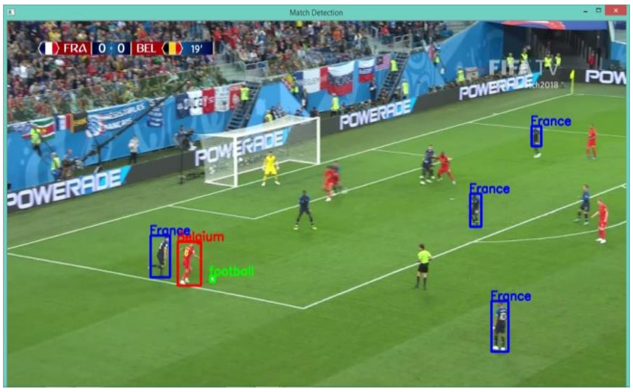
$$(23, Pass, t + \bar{k}, (passingPlayer, p_1), (receiviedPlayer, p_2), (passedObject, ball)) \tag{7}$$

Where, 23 is the variable, Pass is the type of event, $t + \tilde{k}$ is duration of time in which the event exists. Received player is the task execute by player p2 on the ball. Passing player is the task execute by player p1 on the ball.

In "shot on goal" event type consider three entities goal post, player and ball. The highly complex "Shot on goal" occurs, when the player kick the ball, the distance among the player and the ball increases at the same time the distance among the goal post and the ball loss the particular threshold.

$$(20, ShotOnGoal, +\bar{k}, (KickingPlayer, p), \left(\begin{matrix} KickedObject \\ t + \bar{k} \end{matrix}, ball\right), (GoalPost, G)) \tag{8}$$

Where, 20 is the variable, "short on the goal" is the type of event, Kicking player is the task execute by player P , *GoalPost* is the task of object *G*, while the ball motion towards goalpost.



6(a)

6(b)

Fig.6(a) & (b) Final Results

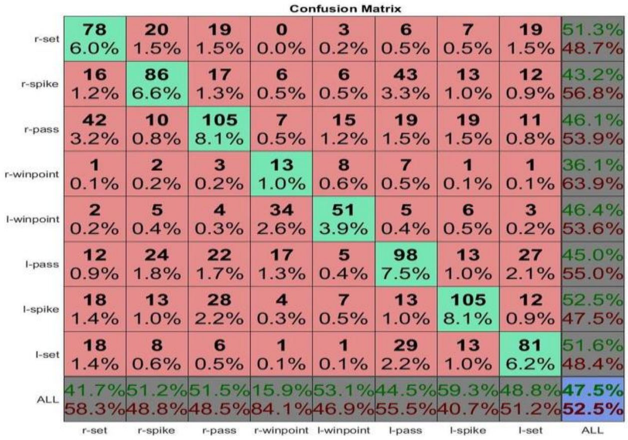


Fig.7 Confusion matrix

Table 4. Performance Measure

Parameters	SVM	RNN	CNN
Accuracy	79.4%	84.5%	87.6%
Sensitivity	92.1%	81.7%	76.8%
Specificity	77.6%	80.7%	86.2%

7. Conclusion

There are numerous events present in the football games and the proposed work tends to classify the four events of the football games. The implemented work attains the accuracy of 87.63% with the reduction in the sensitivity and the specificity range of 72.5 and 86.2%. The confusion matrix of the proposed work gives the classification accuracy during the training phase. For each mass category, the bounding boxes related with a confidence score taken as input for event detection. The low-level complex events are successfully detected by the system such as: "Ball possession" and "Kicking the ball ". The final results indicate the effectiveness and validity of the proposed technique. The future work implies the consideration of more number of events within the classification process.

Declarations

Source of Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing Interests Statement

The authors declare no competing financial, professional and personal interests.

Consent for publication

We declare that we consented for the publication of this research work.

Code availability

The programming code that we have used for this research is available and authors are willing to share when it is required.

References

[1] Carcel Folch, E. (2011). Rich internet application for semi-automatic annotation of semantic shots on keyframes.

[2] Hern, E., Beltr, Z. Z., & Zannatha, J. M. I. (2011, November). Soccer ball speed estimation using optical flow for humanoid soccer player. In 2011 IEEE Electronics, Robotics and Automotive Mechanics Conference (pp. 178-183). IEEE.

[3] Mackenzie, R., & Cushion, C. (2013). Performance analysis in professional soccer: Player and coach perspectives. In Performance analysis of sport IX (pp. 49-57). Routledge.

[4] Kjellström, H., Kragić, D., & Black, M. J. (2010, June). Tracking people interacting with objects. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 747-754). IEEE.

[5] Netcraft, L. T. D. (2009). Netcraft web server survey. URL http://news.netcraft.com

[6] Priya, G. L., & Domnic, S. (2011, June). Video cut detection using Hilbert transform and GLCM. In 2011 International Conference on Recent Trends in Information Technology (ICRTIT) (pp. 749-753). IEEE.

[7] de Sousa, S. F., Araújo, A. D. A., & Menotti, D. (2011, January). An overview of automatic event detection in soccer matches. In 2011 IEEE Workshop on Applications of Computer Vision (WACV) (pp. 31-38). IEEE.

[8] Molchanova, V. S., Vidishcheva, E. V., & Stojsavljević, R. (2013). Designing Republic of Abkhazia Tourism Site Web 2.0 Technology-Based as a Key Competitive Advantage of the Recreational Industry. European Journal of Economic Studies, (4), 215-220.

[9] Zhu, G., Xu, C., Huang, Q., Rui, Y., Jiang, S., Gao, W., & Yao, H. (2008). Event tactic analysis based on broadcast sports video. IEEE Transactions on Multimedia, 11(1), 49-67.

[10] Xu, P., Xie, L., Chang, S. F., Divakaran, A., Vetro, A., & Sun, H. (2001, August). Algorithms And System For Segmentation And Structure Analysis In Soccer Video. In ICME (Vol. 1, pp. 928-931).

[11] Xie, L., Xu, P., Chang, S. F., Divakaran, A., & Sun, H. (2004). Structure analysis of soccer video with domain knowledge and hidden Markov models. Pattern Recognition Letters, 25(7), 767-775.

[12] Wan, K., Yan, X., Yu, X., & Xu, C. (2003, November). Real-time goal-mouth detection in MPEG soccer video. In Proceedings of the eleventh ACM international conference on Multimedia (pp. 311-314).

[13] Shimawaki, T., Sakiyama, T., Miura, J., & Shirai, Y. (2006, August). Estimation of ball route under overlapping with players and lines in soccer video image sequence. In 18th International Conference on Pattern Recognition (ICPR'06) (Vol. 1, pp. 359-362). IEEE.

[14] Nunez, J. R., Facon, J., & de Souza Brito, A. (2008, June). Soccer Video Segmentation: referee and player detection. In 2008 15th International Conference on Systems, Signals and Image Processing (pp. 279-282). IEEE.

[15] Liang, D., Liu, Y., Huang, Q., & Gao, W. (2005, November). A scheme for ball detection and tracking in broadcast soccer video. In Pacific-Rim Conference on Multimedia (pp. 864-875). Springer, Berlin, Heidelberg.

[16] Leo, M., Mosca, N., Spagnolo, P., Mazzeo, P. L., & Distante, A. (2008, September). Real-time multi-view event detection in soccer games. In 2008 Second ACM/IEEE International Conference on Distributed Smart Cameras (pp. 1-10). IEEE.

[17] Cho, J. U., Jin, S. H., Dai Pham, X., Jeon, J. W., Byun, J. E., & Kang, H. (2006, October). A real-time object tracking system using a particle filter. In 2006 IEEE/RSJ international conference on intelligent robots and systems (pp. 2822-2827). IEEE.

[18] Kataoka, H., Hashimoto, K., & Aoki, Y. (2011, September). Player position estimation by monocular camera for soccer video analysis. In SICE Annual Conference 2011 (pp. 1985-1990). IEEE.

[19] von Hoyningen-Huene, B. (2011). Real-time Tracking of Player Identities in Team Sports (Doctoral dissertation, Technische Universität München).

[20] Hirano, S., & Tsumoto, S. (2004, September). Finding interesting pass patterns from soccer game records. In European Conference on Principles of Data Mining and Knowledge Discovery (pp. 209-218). Springer, Berlin, Heidelberg.