Available online at www.sciencedirect.com
ScienceDirect
ICT Express 6 (2020) 170–174
www.elsevier.com/locate/icte

Hockey activity recognition using pre-trained deep learning model

Keerthana Rangasamy[a,*], Muhammad Amir As'ari[a,b], Nur Azmina Rahmad[a], Nurul Fathiah Ghazali[a]

[a] School of Biomedical Engineering and Health Sciences, Faculty of Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia
[b] Sport Innovation and Technology Center (SITC), Institute of Human Centered Engineering (IHCE), Universiti Teknologi Malaysia, Johor Bahru, Malaysia

Received 6 January 2020; received in revised form 15 April 2020; accepted 28 April 2020
Available online 21 May 2020

Abstract

Activity recognition in sports is often complex task resulting from the rapid dynamic interaction within players. In this paper, pre-trained VGG-16, deep learning based hockey activity recognition model has been proposed. Own hockey dataset consisting of four main activity includes free hit, goal, penalty corner and long corner was constructed as there are no existing field hockey datasets available. Experimental results indicate that the pre-trained deep learning model generates comparative results on this challenging dataset by tweaking the hyperparameters of this pre-trained model.

Keywords: Sport video analysis; Deep learning; Activity recognition

1. Introduction

Sports play a significant role in the field of entertainment in these times [1]. As a consequence, coaches are seeking numerous ways to boost the performance capabilities of the players. It is impossible to remember and insight all the movements and actions of the player at the end of the match by coach to employ that information to train their players in terms of improving possible mistakes. As a result, performance analyst also known as notational analyst takes the role by recording the entire event, collecting data such as identifying the activities of the player, player movement, time of a specific activity and presents those crucial findings to coaches [2]. Later, coaches will use those data to train their players and improve the performance level of the players. However, it is a heavy burden to the performance analyst to annotate each activity one by one manually in an effort to identify the activity being performed by the players. Hence, automated activity recognition system is proposed to automatically recognize the activity of the players on the hockey pitch. This research mainly focuses on hockey sport as only little research was conducted in hockey. Almost 80% of sport related research had focused on soccer, basketball, baseball and tennis [3].

Activity recognition as above-mentioned is one of the most fundamental tasks for the performance analyst to begin any of their sport analysis. Although there are several tools such as Dartfish, Sportcode to be used by the performance analyst, they still require to watch the entire sport video in order to label activity performed by players for further analysis nonetheless coaches and players needed the output of the analysis within less time [4]. So, this paper is focused on hockey activity recognition in the field of computer vision to ease coaches in evaluating the players performance. A model has been developed which uses hockey video images as an input to automatically recognize the four main activities of the hockey match. The activities are penalty corner, goal, long corner and free hit. Penalty corner, long corner and free hit will be given against the defending team for any defensive infringement. These three activities are considered important in hockey game as it has power to change score of the game. And goal on the other hand, the team with many goals is the winner of the game. Thus, analysing the movement pattern and the position of the player before and after these activities could play huge part in improving match performance.

In order to achieve the objective of this research deep learning approach was implemented. Using deep neural network one is able to learn and extract the features acquired directly from the inputs unlike traditional machine learning approach

* Corresponding author.
E-mail addresses: keerthana2@live.utm.my (K. Rangasamy), amir-asari@biomedical.utm.my (M.A. As'ari), nazmina4@live.utm.my (N.A. Rahmad), fathiah5@live.utm.my (N.F. Ghazali).

Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

https://doi.org/10.1016/j.icte.2020.04.013

Table 1
Traditional Method of Sport Analysis Model.

Article	Proposed Model	Sports Category
[6]	Proposed discriminative model in order to learn the hierarchical structure of sport video events.	Field Hockey
[7]	Design a hierarchical method for detection of play-break in ice hockey match.	Ice-Hockey
[8]	Introduce rule based algorithm by employing low-level features	Soccer
[9]	Develop a model using isophote's curvature and discriminative features to detect ball in soccer game	Soccer
[10]	Present a method for basketball highlight generation using excitement modelling	Basketball

where the features to be extracted from the inputs needed to be handcrafted. The proposed model utilizes pre-trained VGG-16 network which was fine tuned for classification of hockey activities (free hit, goal, long corner and penalty corner). A dataset of hockey images has been constructed from broadcasted hockey match found in YouTube since there is no publicly accessible benchmark hockey dataset for activity recognition. In this hockey dataset, activities of free hit, goal, long corner and penalty corner are labelled which will be used for training using the model.

The main contribution of this paper is to propose automated activity recognition model for broadcasted hockey match from own datasets consisting of four main hockey activities: free hit, goal, long corner, and penalty corner. The rest of this paper is organized as given below, in Section 2 are reviews on related work for sports related activity recognition. Section 3 focuses on methodology and Section 4 presents results and discussion and finally Section 5is conclusion.

2. Related work

2.1. Sport video analysis

With the huge amount of available free online datasets and the breakthrough of the Convolution Neural Network (CNN) in object detection and image classification in the field of computer vision has brought sport analysis as current topic of interest for many researchers [5]. Although previously, sport video analysis was done using the traditional machine learning method but with the current development of technology, has leads to new evolution for sport video analysis which achieving outperform results through deep learning approach. Sport analysis is a huge domain where each sports has its own unique characteristics. During the early stages, before the series of development in deep learning, experts design each feature to extract the desire features for a specific game. Due to its complexity and inflexibility, these handcrafted machine learning models were just limited to specific sports only such as soccer game, basketball, baseball and tennis game [3]. Table 1 shows the list of some handcrafted sport analysis models.

After the breakthrough of CNN in object detection and image classification, researchers started to implement and evolve

Table 2
Deep Learning Method of Sport Analysis Model.

Article	Proposed Model	Sports Category
[12]	Introduce end-to end soccer video scene and event classification using CNN based deep transfer learning approach	Soccer
[13]	Develop model using CNN and RNN to draw frame features and frames temporal relationship	Soccer
[14]	Addressed event recognition as action localization issues by using Segment-CNN	Soccer
[15]	Discover a model consisting of VGG-16 and LSTM that spot events in long soccer games	Soccer
[16]	Present puck possession events classification model using RNN	Ice Hockey
[17]	Proposed 2-stage deep temporal model using LSTM model	Volleyball

the traditional machine learning models using deep learning approach [11]. It starts from text classification, slowly develops to human activity recognition, simple action recognition till complex group activity recognition and sportvideo analysis. In the area of sport video analysis it can be divided into several categories such as trajectory and tracking, implementing spatial features, temporal features and also combination of spatial temporal features. However, this paper is just focused on utilization of spatial feature through CNN as discussed earlier. Table 2 shows some of the previous research on deep learning based sport video analysis.

3. Methodology

3.1. Dataset preparation

Since, there is no publicaly available hockey dataset, own hockey dataset was collected form International Hockey Federation (FIH) YouTube videos of (Hockey World Cup 2018). The dataset consisting of video frames of free hit, goal, long corner and penalty corner is collected from 12 broadcasted hockey matches as shown Fig. 1. The video resolution was 1280x720. The gathered videos are converted to video frames. The frame rate of 25 fps was used for video frame extraction and all the desired activities are manually annotated. From the extracted frames only total of 400 key hockey activity frames were collected in these datasets which is 100 per class to ensure the datasets are evenly distributed. These hockey datasets are challenging datasets as it includes video frames from various camera viewpoints, variation in the scale, appearance and position of the player.

3.2. Frame selection

As aforementioned, datasets of hockey activity frames were selected from the collected 12 YouTube videos. For each class only 100 frames were selected, so a total of 400 frames were used in this experiment. These RGB frames were resized to 224x224 at an input size of VGG-16 model. The datasets were normalized and saved as numpy array before passing them to the model. In this study 10-fold cross-validation was utilized.

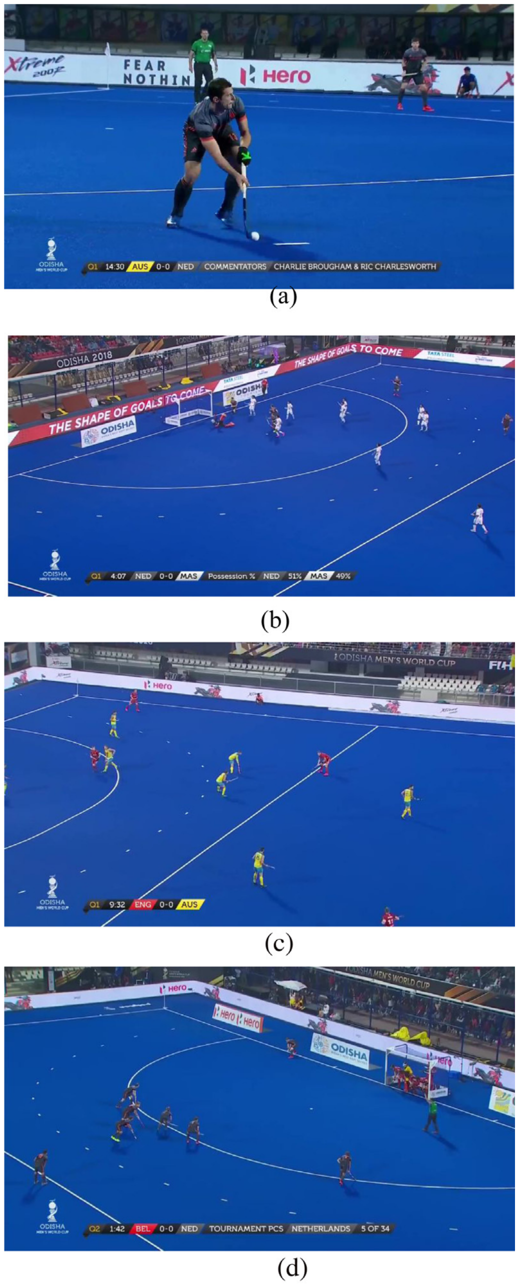

Fig. 1. Dataset of hockey (a) Free Hit, (b) Goal, (c) Long Corner and (d) Penalty Corner.

3.3. Frame-level activity recognition model

In this frame-level activity recognition system, VGG-16 model which has pre-trained by Imagenet was used as shown in Fig. 2. Using transfer leaning method for training these collected datasets is more efficient than training using CNN from scratch. The VGG-16 model parameters were fixed where 16 convolutional layers with 3-fully connected layers were followed by softmax layer. The last fully connected (FC) layer was removed and replaced with a new fully connected (FC) layer for this activity classification model. Considering the collected datasets are close to Imagenet dataset, it is not necessary to fine-tune whole layers.

In this classification model, video frames of hockey activities are the input of the model. The inputs were passed on to VGG-16 model which was fine tuned for this hockey activity recognition task. The output was obtained from the highest probability of the softmax layer. This model extracts frame level features from the entire frame on the image. It learns patterns from the entire frame and extracts from low-level features to high-level features as it passes from the 1st input layer till the end of the layer of this VGG-16 model through consecutive convolutional and pooling process. The training was conducted using 10-fold cross-validation. And the overall accuracy was measured using categorical crossentropy.

This activity recognition model was implemented using Keras (version 2.3.1) with Tensorflow backend. The training process was repeated three times by using different number of epoch 100, 200 and 300 to study one of the hyperparameter which is epoch. The batch size was fixed to the default value of 32. Adam optimization with learning rate of 0.0001 was used for optimization of the model and the values of beta 1 and beta 2 were fixed to default values as well which are 0.9 and 0.999 respectively. Adam is a prevalent algorithm in the field of deep learning, as it attains excellent results fast. The models were trained with a Nvidia GeForce GTX 1050 Ti GPU. The framework of proposed model is illustrated in Fig. 3.

4. Results and discussion

As mentioned earlier, the proposed model recognizes hockey activities which are free hit, goal, long corner, and penalty corner. Datasets were manually created and tagged as a consequence of lack of open and high quality hockey video dataset. Stratified 10-fold cross-validation was utilized in this hockey

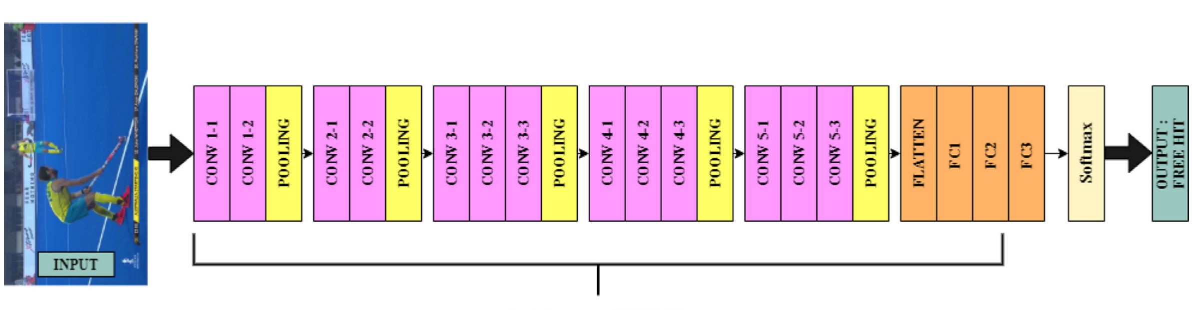

Fig. 2. Architecture of Proposed VGG-16 Model.

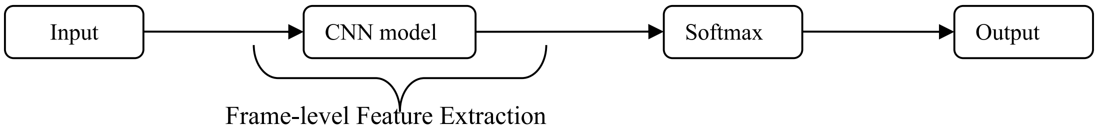

Fig. 3. Framework of Proposed Classification Model.

Table 3
Accuracy Matrix of proposed Model.

No. of epochs	Class	Precision	Recall	F1-Score	Accuracy
100	Free Hit	1.00	0.80	0.89	0.90
	Goal	0.93	0.93	0.93	
	Long Corner	0.82	0.90	0.86	
	Penalty Corner	0.86	1.00	0.92	
200	Free Hit	1.00	0.80	0.89	0.95
	Goal	1.00	1.00	1.00	
	Long Corner	0.83	1.00	0.91	
	Penalty Corner	1.00	1.00	1.00	
300	Free Hit	1.00	0.90	0.95	**0.98**
	Goal	1.00	1.00	1.00	
	Long Corner	0.91	1.00	0.95	
	Penalty Corner	1.00	1.00	1.00	

activity model. The term stratifies here means the splitting of data for each class is equal. In this model, only single hyperparameter was tuned which is number of epochs. This research was repeated three times with same datasets and model architecture but with distinct number of epochs which were 100, 200 and 300. The precision, recall, and F1 score as well as confusion matrix for this research for each epoch were presented in Table 3 and Fig. 4 respectively.

Based on Table 3, we tabulated the precision, recall and F1 score of the model for each number of epoch. The precision examines how accurate a model is, recall examines how many of the actual positives were correctly predicted while F1 score measures the balance between precision and recall. Since, in the studies both precision and recall play vital part for accuracy measurement, we consider F1 score for evaluation of the model. Based on the F1 score of the model, epochs with 300 have the highest score. It has the highest accuracy as well which is 98.0%. The model was slightly confusing between free hit and long corner as both of the activities shares mostly similar visual pattern in terms of position and appearance of the player. Fig. 4 displays confusion matrix that is used to evaluate the performance of the model in terms of simple visualization.

5. Conclusion

In this work, a deep learning based transfer learning model, VGG-16 has been proposed on activity recognition in field hockey. Four main activity recognition which are free hit, goal, long corner and penalty corner are identified from the collected hockey dataset through this pre-trained model. The highest accuracy achieved by the model is 98.0% in hockey activity recognition. This finding is promising and future work should emphasize on incorporating more types of hockey activities and also including spatial features with temporal features by

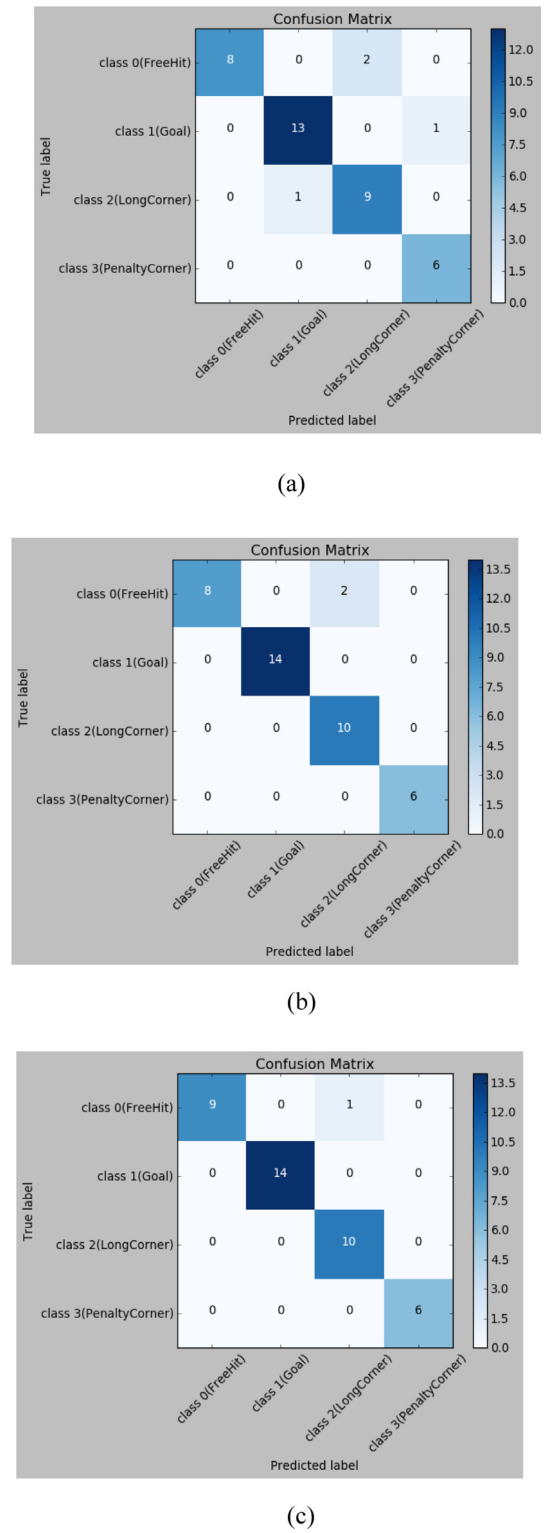

Fig. 4. Confusion matrix proposed method for different no of epoch (a) 100, (b) 200 and (c) 300.

incorporating LSTM model as well, as the proposed pre-trained VGG-16 model has just focused on spatial features on frame-level hockey activity recognition.

CRediT authorship contribution statement

Keerthana Rangasamy: Conceptualization, Data curation, Writing original draft, Methodology, Investigation. **Muhammad Amir As'ari:** Supervision, Project administration, Writing - review & editing. **Nur Azmina Rahmad:** Writing - review & editing. **Nurul Fathiah Ghazali:** Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to express their appreciation to Universiti Teknologi Malaysia (UTM) for endowing this research and the Ministry of Higher Education (MOHE), Malaysia for supporting this research work under Zamalah Scholarship and Research Grant No. Q.J130000.2651.16J20.

References

[1] M.A. Russo, A. Filonenko, K.H. Jo, Sports classification in sequential frames using CNN and RNN, in: 2018 Int. Conf. Inf. Commun. Technol. Robot. ICT-ROBOT 2018, 2018, pp. 1–3.

[2] M. Stein, et al., Bring it to the pitch: Combining video and movement data to enhance team sport analysis, IEEE Trans. Vis. Comput. Graphics 24 (1) (2018) 13–22.

[3] H.C. Shih, A survey of content-aware video analysis for sports, IEEE Trans. Circuits Syst. Video Technol. 28 (5) (2018) 1212–1231.

[4] D. Gu, Analysis of tactical information collection in sports competition based on the intelligent prompt automatic completion algorithm 35, 2018, pp. 2927–2936.

[5] G. Yao, T. Lei, J. Zhong, A review of convolutional-neural-network-based action recognition, Pattern Recognit. Lett. 118 (2019) 14–22.

[6] T. Lan, L. Sigal, G. Mori, Social roles in hierarchical models for human activity recognition, in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2012, pp. 1354–1361.

[7] M.A. Carbonneau, A.J. Raymond, E. Granger, G. Gagnon, Real-time visual play-break detection in sport events using a context descriptor, in: Proc. - IEEE Int. Symp. Circuits Syst., Vol. 2015-July, 2015, pp. 2808–2811.

[8] D.W. Tjondronegoro, Y.P.P. Chen, Knowledge-discounted event detection in sports video, IEEE Trans. Syst. Man Cybern. A 40 (5) (2010) 1009–1024.

[9] P.L. Mazzeo, P. Spagnolo, M. Leo, T. De Marco, C. Distante, Ball detection in soccer images using isophote's curvature and discriminative features, Pattern Anal. Appl. 19 (3) (2016) 709–718.

[10] G.G. Lee, H.K. Kim, W.Y. Kim, Highlight generation for basketball video using probabilistic excitement, in: Proc. - 2009 IEEE Int. Conf. Multimed. Expo, ICME 2009, 2009, pp. 318–321.

[11] A. Khan, A. Sohail, U. Zahoora, A.S. Qureshi, A survey of the recent architectures of deep convolutional neural networks, 2019, pp. 1–62.

[12] Y. Hong, C. Ling, Z. Ye, End-to-end soccer video scene and event classification with deep transfer learning, in: 2018 Int. Conf. Intell. Syst. Comput. Vision, ISCV 2018, Vol. 2018-May, 2018, pp. 1–4.

[13] H. Jiang, Y. Lu, J. Xue, Automatic soccer video event detection based on a deep neural network combined CNN and RNN, in: Proc. - 2016 IEEE 28th Int. Conf. Tools with Artif. Intell., ICTAI 2016, 2017, pp. 490–494.

[14] T. Liu, et al., Soccer video event detection using 3D convolutional networks and shot boundary detection via deep feature distance, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), in: LNCS, vol. 10635, 2017, pp. 440–449.

[15] J. Yu, A. Lei, Y. Hu, Soccer video event detection based on deep learning, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), in: LNCS, vol. 11296, 2019, pp. 377–389.

[16] M.R. Tora, J.J. Little, Classification of puck possession events in ice hockey, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW, 2017, pp. 147–154.

[17] M. G. Ibrahim M.S., S. Muralidharan, Z. Deng, A. Vahdat, A hierarchical deep temporal model for group activity recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1971–1980.