IBC2022

AUTOMATED CONTENT ENRICHMENT IN SPORTS BROADCAST

M. Izadi, G. Sarafraz, A. Behnamnia, M. Hosseinzadeh, S. Satish, M. Abdollahpour, E. Goodarzi, H. Rabiee, A. Izmailov

TVCONAL, Singapore https://tvconal.com/
info@tvconal.com

ABSTRACT

Data will continue to play a major role in sports from both performances as well as consumer engagement perspectives. An increasing amount of digital sports content is generated and made available through television broadcasts or streaming over the internet. Such content is an empowering source of data in sports. However, the content is not used to its full potential due to the tedious amount of work required to tag, curate, and extract value. To date, most tactical analyses are performed by reviewing match videos manually. In an elite broadcast production setting, partial automation is currently possible to only retrieve the finest and the most contentious moment of a match. Metadata tags are either generated using the pitch of audio signals, or generic classifications of themes or are purely generated by human loggers manually. In this paper, we address the problem of automated tagging in a domain-specific software system that is built on tracking technologies and action recognition using computer vision and deep learning models. This is beyond creating general tags for highlight generation. We examine the sport of cricket as a case study and present the practical impacts of this technology in a match broadcast and other sectors of the sports industry such as scouting and high-performance training.

INTRODUCTION

Detailed accurate data collection will become more important and valuable as the sport evolves in the future. Increased levels of data that teams and fans are interested in require a larger and more skilled team to collect and log for such uses.

The majority of sporting leagues dynamically require large logging teams, with sometimes as many as twenty individuals logging a live match content which is processed in a control room for broadcast purposes such as replay and for use by teams coaching and analytics departments. The introduction of automation across the entire workflow of data collection, processing, curation, and maintenance will significantly assist the industry in overcoming the challenges of collecting and logging data from live sporting events.

This paper discusses an AI application that impacts the entire workflow of collecting and logging data from live sporting events. The introduced technology provides reliable and accurate data to broadcasters and teams. For cricket alone, it generates over fifty tags for each ball bowled in a match. Each tag is associated with millions of data points collected in every frame of a ball delivery content. The rich data is then consolidated using fifty parameters in a team profile and seventy parameters in the player's profile to extract insights.

Essentially our AI-powered system discretizes the video to game-events and non-game event segments and further breaks down each game-event segment into a single unit of analysis (shot, play, strike, etc.) in that game. We examine the sport of cricket as a case study and present an architecture that allows automatic recognition of deliveries. Cricket is the second most-watched sport after soccer with over a billion fans. Cricket matches are long in duration compared to other sports. An accurate recognition model that detects deliveries and distinguishes their semantic taxonomy allows searching the content and retrieving any event types they encode. This rich information, combined with the quantitative nature of cricket matches where many metrics are measured for every delivery, provides an unprecedented opportunity to create innovations, investigate game tactics, and create customized highlights and other sources of fan engagement.

The system is compatible with both archived and live sports content. We use the asynchronous and multi-process implementation to perform in almost real-time. In building the underlying models, we have considered a wide variety of sports production types and have trained these models with a large amount of content. This ensures the models are generalizable to any conditions or production setups. We investigate the utilization of natural language processing on top of the automated tags to generate automated commentary scripts in English. This can further enhance the sports production in the absence of professional English commentary, which is the case for lower-tier broadcast and streaming productions.

BACKGROUND

With increasing demand to consume sports content on Over The Top, OTT, platforms and the technological advances in data streaming [1], these platforms will also need to generate consumable sports data and content from a large body of archives or from live streams almost immediately after the live broadcast. Untagged or minimally tagged piles of content can quickly become a logistical nightmare when it comes to retrieving the right content for a specific consumption such as innovative highlights, player promotions, or performance improvement and scouting.

Content logging in sports broadcast

Live data collection and logging are demanding and challenging tasks in sports matches. In a multilevel prospective study [3], many sports media outlets confirmed that they lack the time and personnel to cover every sport match manually, especially concerning pre-match reports and editorials that need data collection from multiple matches and teams. These outlets see automation as a helpful tool that complements their work to cover more matches and focus on better quality content. Computer vision is at the forefront of applications development in sports analytics with many commercial products available for gaining insights into coaching, planning against opponents, and more appealing visualizations for sports viewers [4]. Some of these applications are focused on generating metadata tags in sports broadcasts. Currently available commercial solutions for automated logging in sports are focused on identifying exciting moments of the match for the purpose of creating highlights. These solutions are either based on selecting clips with high audio intensity as candidates for highlights or are based on using Optical Character Recognition (OCR) to identify events such as wickets and boundaries from the scorecard part of the video [2]. Either way, the tags are not comprehensive enough to support all types of events and are not semantically specific to all events type in a single sport.

To develop models for automated logging in more specific details, many manually annotated tags on every specific detection are required. The lack of such datasets introduces limitations in developing AI models that are robust in automatically logging each and every event in a video regardless of their excitement level.

Related Work

Video analytics in sports technology is progressing rapidly in the area of tracking [5,6]. Recent advances in such technologies have gained a great amount of success in the detection of balls and athletes in both indoor and outdoor sports. Many previous works focused on detecting actions on sports broadcasts for various sports. In Football, classified game events are detected by combining visual features of players in each video frame and taking into account the feature representation on previous frames. Their model classifies each frame into three classes; Pass, Dribble, and Shoot. First, the ball and each player in the frame are detected using an object detector. Then visual features of the ball and players are aggregated using a Long Short-Time Memory (LSTM) sequence model. Finally, another LSTM sequence model processes the frame features conditioned on previous frame features. In Baseball, [6] performed action classification, pitch type classification, and pitch speed regression on custom-based data set from the MLB-YouTube dataset containing densely annotated frames of broadcast baseball videos suitable for fine-grained activity recognition. The dataset includes two parts: segmented videos, short clips of actions, and continuous videos, which are one to two minutes long videos. In cricket, [2] proposed a model that summarizes highlights of cricket matches based on recognizing and clipping the important events. The algorithm first splits the video into video shots. It sequentially scans the keyframes in the video shot to detect an action. Once an action is detected, the algorithm goes to previous frames to find the starting action frame. They rely on various features such as audio cues for celebration detection and the presence of the scoreboard for replay detection.

Our work is substantially different from the existing research in the literature in several ways. First and foremost our models do not depend on the audio cues and in fact in the absence of any spectators or commentary they detect the actions and special moments of the game. Secondly, our algorithms are built for more than making highlights; they detect every single delivery and tag them with a large number of tags. This is the only way the technology can lift the requirement for engaging a large number of loggers. In the next section, we present details of our methodology.

METHODS

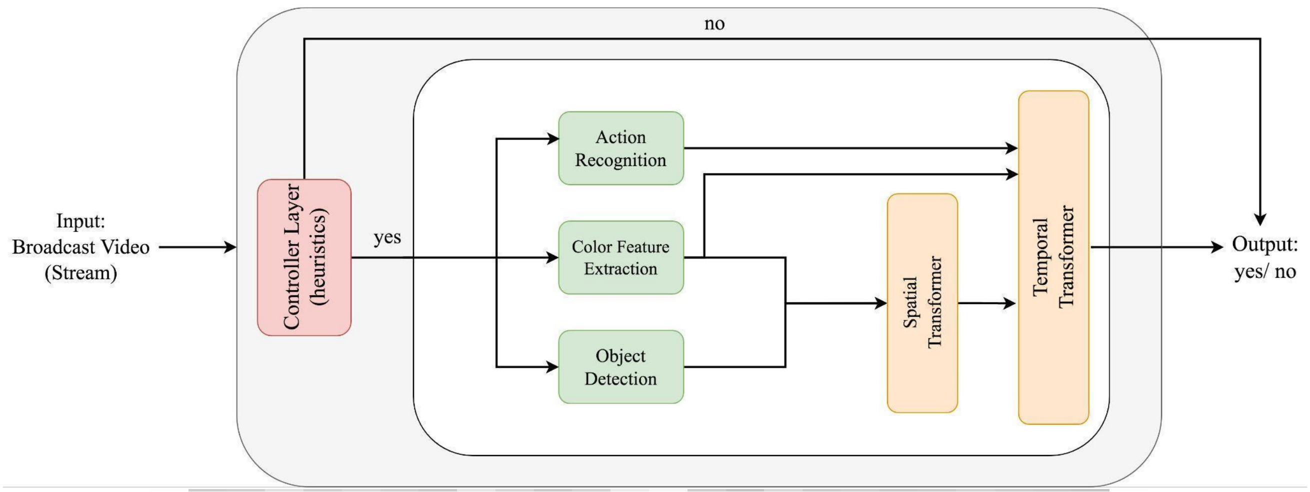
Our general approach to sports logging is to divide all the events that happen during the broadcast of a match and contained in the broadcast content into two broad categories of game events, and non-game events. Then we further categorize the game events into different types of playing actions. Similarly, the non-game events are categorized as more refined events. Game events are in fact modular sport-specific events that are called play, shot, delivery, kick, stroke, etc. depending on the domain. We call these events a unit of analysis and try to provide as many descriptive tags as possible for every unit. We use customized methods in each sport to detect game events from non-game events.

Game-Events Detection

Generally, each unit of analysis needs to be distinguished from the rest of the content. This step is the most important part of tagging. Ball detection and tracking, player detection, and tracking are the components of game-event detection. Additionally, sports gear such as bat, helmet, pad, racket, etc. is needed to be detected correctly so that some playing actions and some playing roles can be described heuristically through the detection of these gears. In this context, the line marks on the pitch, court, field, and ground should be detected or determined as well.

Non-Game-Events Detection

The second category of events that we need to distinguish and further analyse are largely in common among sports. This includes the crowds or spectators, celebrations, interviews with players and coaches, emotional moments, view of the stadium surroundings, fans, graphics, sponsors, and advertisements. We use a combination of Convolutional Neural networks (CNNs), computer vision techniques, and customized heuristics to detect such events.



Our approach to detect a unit of analysis event from match videos relies on a carefully designed combination of spatial and temporal features. For each frame in a video buffer, our model extracts spatial features and processes these features with a temporal model to predict if the buffer contains a unit of analysis or not. In order to model spatial and temporal relations between features, we use the attention mechanism proposed in [7]. Before processing temporal or spatial features, we first check to see if the frame is a field scene. We use the hue histogram of the frame to detect whether it is showing a sports field or not. This may improve prediction accuracy as many completely unrelated scenes are discarded. It also reduces computational cost significantly because the model stops processing the shot when detected as a non-field scene. We use an online shot change detection model to obtain event intervals in the video stream and decide if each of these chunks corresponds to the desired event or not. Figure 1 shows the proposed game event detection pipeline.

Spatial Features

If the frame is in the field, the view of the frame is passed to an object detection model. Detected objects are grouped and aggregated to correct the model's prediction for complex objects, especially the ball. Also, some colour features such as hue are extracted from the frame. Moreover, we use a state-of-the-art action detection model to extract action features from the video frame. We compute flux tensors from the detected objects to detect movements in the scene. The actions, objects, and shot changes are different in different sports and are detected based on pre-trained models. The object features are processed and aggregated through a spatial transformer.

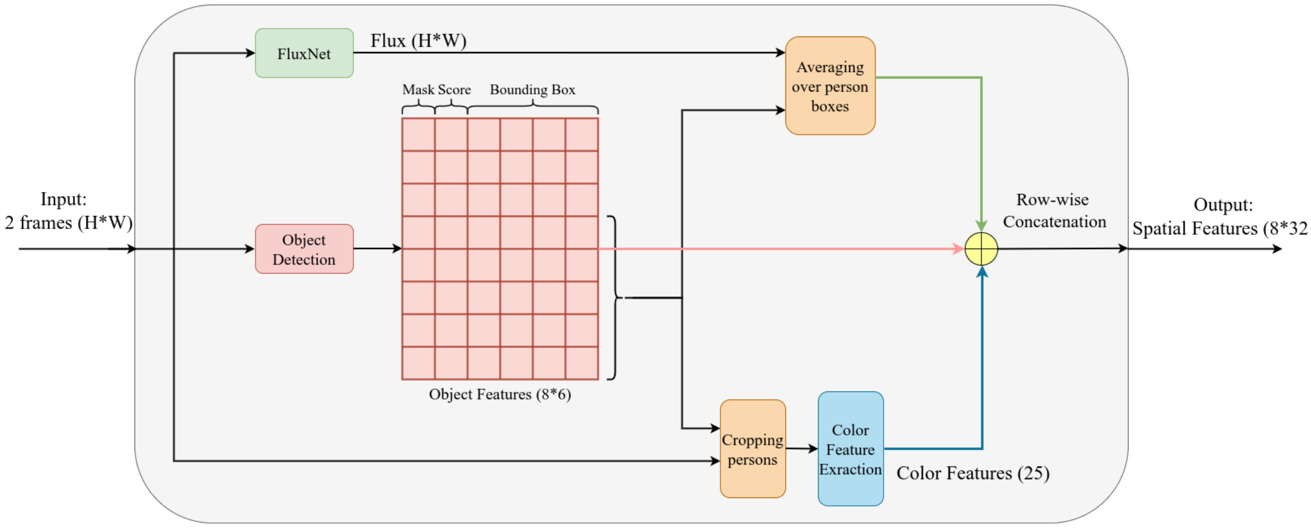


Figure 2 - Spatial feature extraction module

Figure 2 presents the spatial feature extraction component of our general modelling pipeline.

Temporal Features

After spatial features are extracted, the spatially processed object features, action features, and colour features are passed to a temporal transformer to model the relation of different frames through time and classify the event. If a shot change is detected, the algorithm will reset the temporal features in the memory.

CRICKET CASE STUDY

In this paper we focus on the professional sport of cricket. Therefore, the unit of analysis is a delivery and the game events are related to bowling a ball and hitting with a bat, and what happens after until the ball stops. Table-1 presents a list of tags that are associated with a delivery or relate to the sport of cricket. Our system is intended to generate these tags automatically and close to real-time in order to replace a human logger. The information related to the state of the match, innings, over, ball, striker, the non-striker, bowler, teams, and venue are extracted from the match information and score-card. In detecting most of the non-game events we followed the approach in [2] on using video shots with the assumption that the pixels of successive video frames differ vastly at the beginning and at the end of a video shot. This assumption mostly holds in multi-camera production.

Game events	Non-game events
Full delivery (Legal, Wide and No Ball)	Pitch Report
Batting Action (Horizontal Bat Shots [Late Cut, Square Cut, Pull Shot, Hook Shot and	Toss

Sweep], Vertical Bat Shots [Leg Glance, Cover/ Straight/ On/ Off Drive and Defense] and Others [Reverse Sweep, Switch Hit, Scoop Shot, Helicopter Shot and Slog])	
Bowling Action (Medium Pace, Fast, Slow, Off-break, Leg-break and Orthodox)	Team Representation
Fielding Action (catch [high, mid, slip] and drop catch)	Opening/ Closing Ceremony
Scoring runs 0,1,...,6	Interviews
Dismissal	Prize Distribution
Replay	Advertisement
Decision Review System demonstration	Sponsors
	Graphics
	Celebration
	Miscellaneous

Table 1 – List of tags associated with event-driven metadata generation

Delivery detection is the heart of automated logs, where we detect the starting time of a delivery event. Once a delivery event is confirmed then the possible bowling, batting, and fielding events are identified in consecutive frames from the start of a unit of a delivery. For instance a batting action on average is not more than 1.5 second from the beginning of a delivery. This is translated to the number frames to analyse based on the fps property of the captured content.

Based on the proposed pipeline for game event detection, our model for cricket delivery detection consists of field view detection, action recognition, object detection, colour feature extraction, and motion estimation. After filtering and determining the delivery starts, the first 36 frames of each delivery were considered bowling, and the second 36 frames with six overlapped frames were considered batting sub-clip. These extracted clips are used to train the action recognition module to further detect the type of batting action, bowling action, or fielding action. In this section, we elaborate on each of these models in detail.

Field View Detection (Controller heuristics)

To detect scenes that contain a cricket field, we use the peak of the hue histogram of the video frames. We collect peak values for hue histograms of random samples from delivery frames and use them as a reference set. Any frame with all peaks outside this set is considered a non-field scene.

Action Recognition

In the designed pipeline, to estimate the action performed in each of the 12-frame sub-clips online, we used MoViNet, an efficient 3D convolutional network proposed in [8]. The MoviNet model is used in the stream and causal mode. In this process, which is online and real-time, each 12-frame sub-clip is assigned to one of the three specified categories: batting, bowling, and others. At the final step, the action recognition module will return the feature map of the last layer, which is a vector with a dimension of 2048. This vector will be one of the input features of the main transformer. To train this module, the weights are initialized using pre-trained MoviNet; next, using 12-frame sub-clips from pre-processed 36-framed videos collected in batting, bowling, and other samples, the model is re-trained. To improve the final accuracy in the stream mode, a 36-frame-long buffer is used, and for each step, model optimization is done when the buffer is emptied.

To further distinguish between different types of an action such as batting action, we develop heuristics that are built on top of the object detection and action recognition to describe each type of an action. For instance, once a batting action is detected, in order to classify batting types, we describe three levels of batting action depending on the position of the bat at the time of contact with the ball. A horizontal batting position at time of contact for example, corresponds to Late Cut, Square Cut, Pull Shot, Hook Shot, Sweep type of batting.

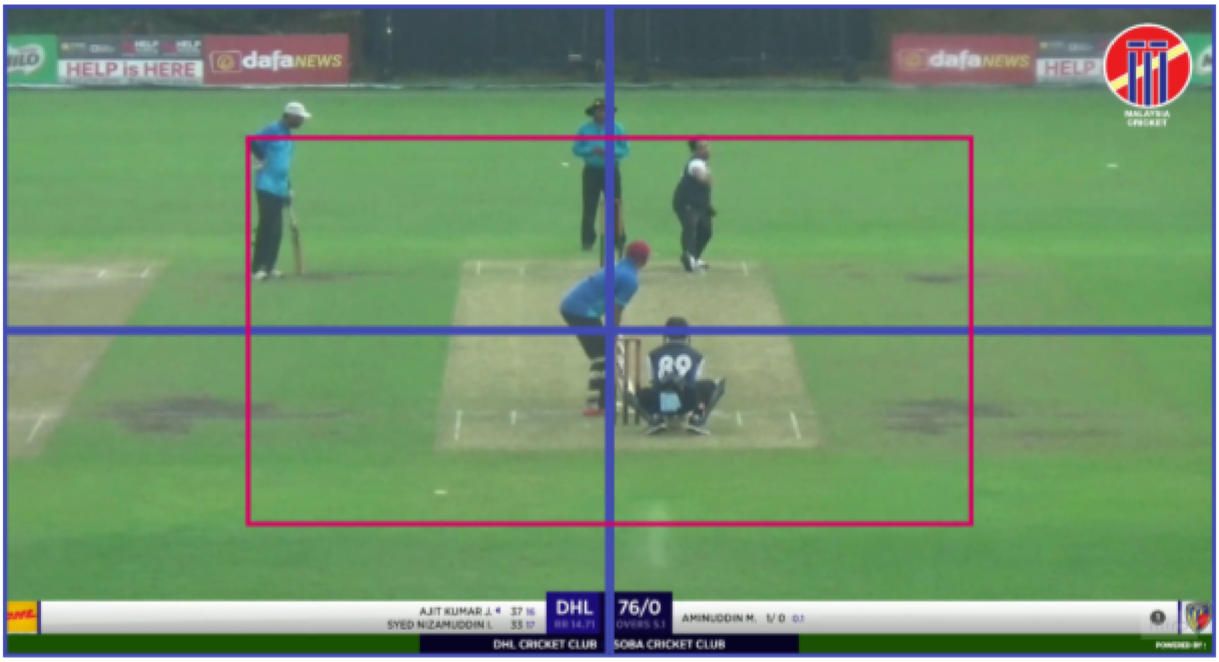
Object Detection and Feature Extraction:

To further understand the delivery scenes, we use an object detection model to detect the desired objects in each frame and combine the detection result with the action recognition result and colour features. As the ball, bats, and persons are the fundamental entities present in almost all delivery intervals, we train an object detector, YOLOX-s [9], on these three classes. The detection score for each of the detections is compared with the specific threshold of that class. The detections passing the threshold are modelled as a feature vector corresponding to eight instances - one ball, two bats, and five persons. For each instance, the detection score and bounding box coordinates are stored. Also, a binary value as the detection mask is added to determine whether the detection is valid; when the number of detected instances for a class is fewer than the predicted number stated above, the valid detections are flagged with mask one. The remaining ones are filled with flag zero. At specific time steps when the features need to be fed to the transformer to detect the delivery, the object features for the frames inside the window are aggregated. The best feature vector representing the whole window is selected in the aggregation process. The persons within the frame with the maximum number of valid persons are chosen to represent the five persons for the window. Also, the centres of the bounding boxes for balls are averaged and represented as the estimated ball for the whole window. We note that the ball is the most challenging class to detect due to its high speed and small size of the ball. Hence, the model might fail to detect the ball in many frames or give false predictions as balls. Average over ball locations in consecutive frames provide a better estimation for frames where the ball is not detected, but it also diminishes the effect of false-positive ball predictions. As a result of the aggregation process, a feature vector of size 8*6 is constructed for the window and combined with other features to be used in the transformer.

Colour Features

We extract different colour features from video frames to model various aspects of the scene. We consider five different boxes from the image, four of them being the four quarters of the image and one larger box at the centre of the image. We extract hue histogram, LBP histogram, and edge direction histogram from each box and concatenate them as a final colour feature. This way, we consider the whole image information for our features and focus more on the scene's centre, where the most important activities happen. Figure 3 illustrates the boxes that colour features were extracted from.

We also extract colour features from the bounding box of each detected person in the scene (the ones detected by the object detection model). We extract a hue histogram from the bounding box of each person to represent the colour of their clothes. These features distinguish players of opposing teams and the umpire from each other.



Motion Estimation

To detect moving objects from the video, we compute the flux tensor in each detected person’s bounding box. Flux tensor is a low-cost alternative to optical flow for motion estimation. It is calculated by taking an average over the derivative of the image intensity gradient with respect to time in a small neighbourhood over the input pixel in the image. Persons with higher flux tensor values move faster.

Figure 4 shows different types of batting actions that are generally logged manually.

Figure 4 – Sample of batting action variations

Data

For the main task of delivery detection, we prepared, annotated, and used more than 20,000 delivery clips from broadcast feeds of different matches and tournaments for the main delivery classifier. An equal number of non-delivery clips are also extracted from the match videos. The non-delivery clips included footage that seemed similar to delivery, where players were running or throwing the ball at each other. These two data classes are used to train the main module for delivery detection.

Depending on the type of production, a delivery clip can look very different. For example a delivery in an elite broadcast where many cameras are included in the match coverage versus a streaming production with one fixed camera at one end-filed looks different to a human viewer and therefore should be similarly detected by the system. To have a more scalable and adaptable system, we need a generalized model that is trained with diverse samples from different types of production. Our data consist of footage from three types of cricket match coverage: multi-camera where more than three cameras are used to show a delivery event; a 3-camera production where there are two fixed cameras at two ends of the pitch and one for tracking the ball; and single-camera streaming that is used in clubs or academies.

To train the YOLOX-s model for the object detection task, we use a dataset of 7500+ images from real cricket matches with annotations for balls, bats, shoes, and persons. Furthermore, to achieve better detection results we used synthetic data to increase the volume of the dataset. We used an algorithmic approach to augment difficult and rare cases in frames. For some objects such as bats, we augment extra bats in each image being occluded by a person. As the bat is often occluded due to the batsman's posture while batting the ball, this augmentation helps the model better detect the bats in the batsman's hand. The colour of the cricket ball is different in different formats and that can impact the detection accuracy. Therefore, we use the content related to the Test matches as well as the Limited Overs format in our training data.

Results

We have evaluated the performance of our system with respect to different performance metrics in a large number of experimental setups. The space is limited to present the complete list of results. Table 2 shows the evaluation results in an elite production is superior in detection of actions versus the productions with only 3 cameras. This partially due to the fact that there is a more established pattern of camera changes to show the player actions with higher number of cameras and it is partially due to have better quality footage in this type of production. Over all the results are very encouraging with respect to both sensitivity and specificity. The timeliness is suitable for live tagging as indicated in Table 2. We observed that the detection of some of the non-game events are still not as accurate as the case for specific game events. We plan to further develop customised models for these events. But in general, the game-events are the ones that are subject of interest to most users.

Table 3 presents the results of our delivery detection that is the central part to automated tagging of many game events. The production quality is important in the performance of our pipeline. We saw the best performance when the feeds of runout cameras are used in delivery detection. These cameras are usually installed in all four corners of the pitch, specifically looking at the bowler and the batsman from left and right side. We have used slightly different models and customised heuristics in consuming the feeds from the runouts as the pattern of actions are substantially different than the broadcast feed.

Production type	Module	Accuracy	Recall	Memory	inference run-time
Multi-cam (>3) Broadcast Feed	Batting Action Recognition	95%	93.6%	2.5 GB	76 fps
Multi-cam (>3) Broadcast Feed	Bowling Action Recognition	97%	94.9%		
Multi-cam (>3) Broadcast Feed	Replay Detection	92.8%	86.7%		
Multi-cam (>3) Broadcast Feed	Average Non-game Events	88.9%	85.2%		
3-cam Broadcast Feed	Overall Action Recognition	91.3%	96.1%		
Single-Camera Runout	Object Detection	Ball: 97.2%	83.4%	1 GB	90 fps
		Person: 99.94%	99.66%		
		Bat: 96.1%	97.4%		

Table 2 – Performance Evaluation of Sample Modules in Different Production Types

Production Type	Accuracy (%)	Recall (%)
Multicam (>3) Broadcast Feed	95.4	96.1
3-cam Broadcast Feed	90.6	93.5
Single-Camera Runout	99.2	96.6

Table 3 – Delivery Detection Performance in Different Production Types

APPLICATION IMPACT

Automation, the ease of use, time and cost-saving have great impacts on different segments of the sports data ecosystem: broadcasters, publishers, teams, leagues, and sports governing bodies.

The live coverage of sports competitions has become the centre of the digital strategy to generate more traffic and engage sports fans. Our detailed tags organized as game-events and non-games events allows the generation of more innovative and more engaging highlights.

It is exciting that with our automated tagging we are also well-positioned to redefine the commentary and the way to cover competitions for the viewers in the absence of human commentators. This is mostly the case in match streaming. Innovations with the commentary text and audio complement and substantially improves the feel of match production. Another area we have already started to explore is the narrative reformulations in covering a match by commentators. We have been able to create a variety of innovative highlights on specific players or on certain playing actions as well as on exciting moments of the match for our cricket partners. We have used our automated tags to generate commentary for the highlights with audio integration of chatbots trained with sports casters voice. In another use case we have integrated the tags generated from our technology to create automated posts for pushing to social networks. Together with the metadata results, the bot-generated tweets with match statistics seem to be engaging for many cricket fans to follow. It is noteworthy to mention that similar work started in the area of sports commentary [10], and in sports journalism [11]. The efficiency and the profitability of this technology reinforce the idea that this application impact serves to satisfy users' need for the latest news in audio, video, and on all types of devices.

The timeliness of providing a team analyst with the content and insights is vitally important. The metadata tags can be linked to the statistics and data points collected throughout the match to create a comprehensive scouting resource. Scouting presents one of the crucial parts of team strategy. According to scouting reports, specific team tactics are organized based on the research on the opponent. Information that would require days or weeks of research into statistics and review of video content can be provided within seconds as our technology enriches the content with many search tags.

We have used this application in a few innovative ways to serve the teams and their analysts with collecting valuable information from searching hours of content. A few biomechanical KPIs can be collected with our technology when we have access to multiple camera feeds specifically setup for high performance analysis. This includes the wrist profile and foot profile of the player while delivering an action and detecting patterns in a player facing an opponent to confirm coaching hypothesis. However, from the practical perspective, one can trade off the accuracy for the quality of the input data feeds and apply the automated tagging in usecases for which high quality data feeds are not cost-efficient or not feasible to have. For instance, in the context of training nets, a cell phone on a tripod can capture the players' actions in a training session instead of runout broadcast cameras. In some preliminary work we have assessed this usecase. While the video can be sliced into an array of clips corresponding playing actions and tagged with specific action type, since each player repeats actions many times during the training missing a few actions will not hurt and the accuracy can be compromised to some levels.

We used cricket as a case study in this paper, however, our approach is applicable to other sports as well. We are currently extending this application to racket sports including table tennis and badminton. Given the fact that the relative size of the playing surface to the players size in an image is much smaller in racket sports in comparison with cricket, there is less requirement for the number of cameras and hardware to track the players and the ball everywhere. Additionally, the accuracy of events detetion is almost hundred percent. This means, there is no need even for a human scorers and that can significantly simplify the competitions workflow.

CONCLUSIONS

Automation holds considerable future potential in enabling accelerated workflows in broadcast and media, and it is an empowering tool for players and content owners. In this paper we carefully selected a set of computer vision and deep learning models to explore automation in the particular application of sports tagging. This application makes it possible to scale up content utilization in fan engagement, as well as player development, and open avenues to diversify revenue streams for content owners.

REFERENCES

1. Aaron Chadha, Eirina Bourtsoulatze, Ilya Fadeev, Vasileios Giotsas, Sergio Grce, and Yiannis Andreopoulos, Deep-learning based precoding techniques for next-generation video compression. Proceedings of the IBC technical papers, 2019.
2. Shukla, P., Sadana, H., Bansal, A., Verma, D., Elmadjian, C., Raman, B., & Turk, M., Automatic cricket highlight generation using event-driven and excitement-based features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 1800-1808) 2018.
3. Jessica Kunert, Automation in sports reporting: Strategies of data providers, software providers, and media outlets, Media and Communication, 2020.
4. Graham Thomas, Rikke Gade, Thomas B. Moeslund, Peter Carr Computer vision for sports: Current applications and research topics, BBC, White Paper, April 2017.
5. Tsunoda, T., Komori, Y., Matsugu, M., & Harada, T., Football action recognition using hierarchical lstm. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 99-107), 2017.

6. Piergiovanni, A. J., & Ryoo, M. S., Fine-grained activity recognition in baseball videos. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 1740-1748), 2018.
7. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. Attention is all you need. *Advances in neural information processing systems*, *30*, 2017.
8. Kondratyuk, D., Yuan, L., Li, Y., Zhang, L., Tan, M., Brown, M., & Gong, B., MoviNets: Mobile video networks for efficient video recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 16020-16030), 2021.
9. Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. Yolox: Exceeding Yolo series in 2021. *arXiv preprint arXiv:2107.08430*, 2021.
10. Greg Lee and Vadim Bulitko and Elliot A., Automated Story Selection for Color Commentary in Sports, IEEE Transactions on Computational Intelligence and AI in Games, 2014.
11. Carlson, M. The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. Digital Journalism, 3(3), 416–431, 2015.

ACKNOWLEDGEMENTS

The authors would like to thank NEP-Group as the broadcast technology partner and multiple cricket boards and teams including the Pakistan Cricket Board, Malaysian Cricket Associations, and Titans Cricket for partnering with us and for their content and data used in this research.