CS 671 - Deep Learning and its Applications
Project Presentation on Cricket Commentary Generation using Deep Learning
Group-05

Vishal Anand(B16040), Aman Jain(B16044), Yash Agrawal(B16120)

April 17, 2019

Problem Statement

Automatic generation of text-based ball by ball commentary of live cricket match by applying deep learning models.

- Teams playing, striker and non-striker batsman, bowler, shots played eg. cover drive and lofted, 1,2, four, six or out.

- A full fledged human-like commentary for live matches.

Motivation

- Cricket being a popular game in Indian subcontinent, but not much of practical research is done in this domain.
- Online cricket websites and apps such as Cricbuzz, provides live text-based commentary of cricket matches which is manually written.
- Cricket videos are accompanied by detailed commentaries available online.

Challenges

- Getting the full match video, trimming and splitting the videos into frames.
- Detection of batsman, non-striker and bowler's names.

- Classification of different types of cricket shots played by the batsman and runs obtained on that shot.
- Generation of text commentary by creating a sentence from the key words such as name of batsman, bowler, etc obtained in above steps.

Dataset

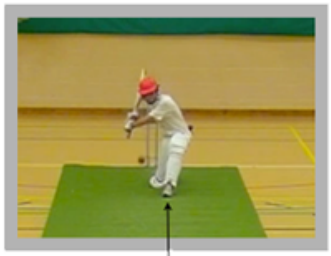
- Full match length cricket videos on websites like Youtube, etc.
- Making the video into frames and then splitting it into set of frames for each ball.
- Manually making the bounding boxes and labelling it as the player's name for batsman, bowler and non-striker.
- A list of vocabulary/ dictionary corresponding to each keywords like: out: great bowling by *bowlerName*, 6: Into the air and fantastic shot by *batsmanName* , etc.

Proposed Methodology

- Three sub tasks:
  1. Player Detection
  2. Shot detection
  3. Commentary generation

If attempting a 'slog sweep' take large backswing

Spot ball is full for the sweep and take stride forward



Make sure front foot is far enough away from the ball to give enough room to swing through

Strike under ball to get it into the air

Proposed Methodology

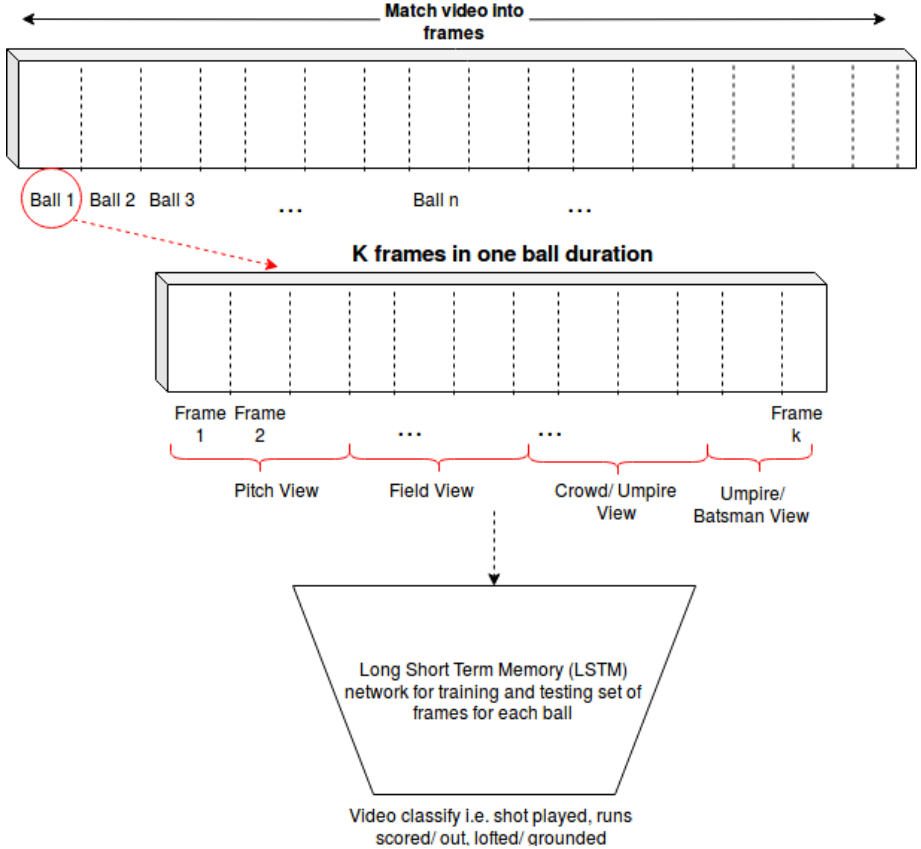
- **Player Detection:** Faster RCNN

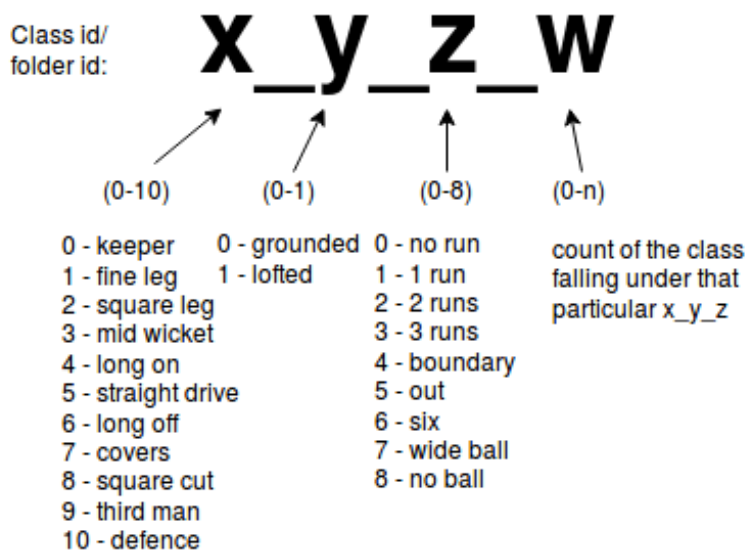
Image	Class	x1	y1	x2	y2
1.jpg	3	23	55	65	122
1.jpg	23	133	45	157	87
1.jpg	12	77	34	101	98
2.jpg	44	33	78	121	156
...					
...					

Class	Player
1	Virat Kohli
2	Rohit Sharma
3	Lasit Malinga
...	
...	

Proposed Methodology

- **Shot Detection:** RNN -> LSTM





Proposed Methodology

- **Commentary Generation:** Randomized algorithm
  We got certain keywords i.e. batsman, bowler, non striker's name, lofted/ grounded shot played, name of the shot played and the result at the end of the ball (1s, 2s, out, 6).
  Have separate vocabulary/ dictionary corresponding to each keywords like: out: great bowling by *bowlerName*, 6: Into the air and fantastic shot by *batsmanName*, etc.

Methodology Explored

- First sub task in our proposed methodology includes detection of batsmen, bowler and non-striker. Since there are more than one object that needs to be detected, object detection techniques such as RCNN, Fast RCNN, Faster RCNN can be used.
  Following articles explain about different object detection techniques :
  1. https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e
  2. https://towardsdatascience.com/object-detection-using-deep-learning-approaches-an-end-to-end-theoretical-perspective-4ca27eee8a9a?fbclid=IwAR0AzYf56EdElovqOPJcIw84ZdgWZIQnt8jAYpMcd2qfruO9RinFZxc7UyI
  3. https://www.analyticsvidhya.com/blog/2018/11/implementation-faster-r-cnn-python-object-detection/

Methodology Explored

- Next sub-task involves predicting the type of shot played by the batsman, final result(runs obtained on that shot) on a particular ball.
- Each ball is a small video of approx 1 min consisting of different frames. Each of these small videos needs to be classified into different categories based on the type of shot, runs scored. Since each ball is a sequence of frames, where one frame is dependent on the previous frames, video classification techniques such as Recurrent Neural Network is used.
- Following articles explain how recurrent neural networks can be used for video classification :
  - https://towardsdatascience.com/recurrent-neural-networks-and-lstm-4b601dd822a5
  - https://blog.coast.ai/continuous-video-classification-with-tensorflow-inception-and-recurrent-nets-250ba9ff6b85?fbclid=IwAR2oDcj-W-grq1MkE3dgfKOr7eazm5Zqpwnfn08qG09cS$_n$6*KCNi*5*nyn* − *Y*

Results and Discussions

- The RCNN technique used for object detection of batsman, bowler, non-striker requires lot of time for training and is slow because convolution neural networks need to run multiple times for multiple region proposals.

- Some of the previous work done related to our problem have obtained good results which can be useful for different sub-tasks in our project.

    1. **Cricket shot detection from videos** : Used saliency and optical flow to bring out static and dynamic cues and on Deep Convolutional Neural Networks(DCNN) and SVM. Accuracy: 83% for right handed and 65% for left handed batsman. Link to paper
    2. **Fine grain annotation of cricket videos** : Given a Cricket Video and corresponding textual Commentary, temporally align/ map them with the video shots. Obtained an accuracy of around 80%. Link to paper

Conclusion

- Using Faster RCNN as the object detection technique, model will tell us the batman, bowler and non-striker by training it for multiple such frames from complete video of cricket match.
- Rest of the information such as the type of shot played by the batsman, on which side of the ground the ball went and the runs scored on that ball will be obtained from the LSTM model which will classify each ball into a particular class.
- The final commentary will be generated using these keywords obtained from the above two tasks.

Future Work

- Currently the model will be trained only for specific players and specific team. Later, it can be trained for data of all the team and all the players.
- The current model does not provide any information about the fielder to which the ball goes. This can also be included in future.
- Improving the accuracy of the model is one of the main goal to focus on in future with more datasets.
- Improving the sentences formed from the keywords obtained from the deep learning models so that a whole new sentence is formed each time.