Automatic Cricket Commentary Generation: A Review

Ira Arora[*1], Anita Choudhary[*2]

*[*1]Student, Department of Computer Science Engineering, Dronacharya College of Engineering, Khentawas, Haryana, India*
*[*2]Assistant Professor, Department of Computer Science Engineering, Dronacharya College of Engineering, Khentawas, Haryana, India*

ABSTRACT
Cricket is the most popular sport in India and is played almost everywhere but not much of the practical research is done in this domain. Online cricket websites and apps such as Cricbuzz, provides live text-based commentary which is manually written. This motivated us for this project.

The recognition of human activities is one of the key problems in video understanding. Action recognition is challenging even for specific categories of videos, such as sports, that contain only a small set of actions. Interestingly, sports videos are accompanied by detailed commentaries available online, which could be used to perform action annotation in a weakly-supervised setting. For the specific case of Cricket videos, we address the challenge of temporal segmentation and annotation of actions with semantic descriptions

I. INTRODUCTION
Computer vision is an interdisciplinary field that has been gaining huge amounts of traction in the recent years (since CNN) and self-driving cars have taken center stage. Another integral part of computer vision is object detection. Object detection aids in pose estimation, vehicle detection, surveillance etc.

The difference between object detection algorithms and classification algorithms is that in detection algorithms, we try to draw a bounding box around the object of interest to locate it within the image. Also, you might not necessarily draw just one bounding box in an object detection case, there could be many bounding boxes representing different objects of interest within the image and you would not know how many beforehand.

The major reason why you cannot proceed with this problem by building a standard convolutional network followed by a fully connected layer is that, the length of the output layer is variable — not constant, this is because the number of occurrences of the objects of interest is not fixed. A naive approach to solve this problem would be to take different regions of interest from the image, and use a CNN to classify the presence of the object within that region.

The problem with this approach is that the objects of interest might have different spatial locations within the image and different aspect ratios. Hence, you would have to select a huge number of regions and this could computationally blow up. Therefore, algorithms like R-CNN, YOLO etc. have been developed to find these occurrences and find them fast.

II. METHODOLOGY
Literature Review and Techniques Used
The major reason why you cannot proceed with this problem by building a standard convolutional network followed by a fully connected layer is that, the length of the output layer is variable — not constant, this is because the number of occurrences of the objects of interest is not fixed. A naive approach to solve this problem would be to take different regions of interest from the image, and use a CNN to classify the presence of the object within that region. The problem with this approach is that the objects of interest might have different spatial locations within the image and different aspect ratios. Hence, you would have to select a huge number of regions and this could computationally blow up. Therefore, algorithms like R-CNN, YOLO etc have been developed to find these occurrences and find them fast.

Computer vision is an interdisciplinary field that has been gaining huge amounts of traction in the recent years(since CNN) and self-driving cars have taken center stage. Another integral part of computer vision is object detection. Object detection aids in pose estimation, vehicle detection, surveillance etc. The difference between object detection algorithms and classification algorithms is that in detection algorithms, we try to draw a bounding box around the object of interest to locate it within the image. Also, you might not necessarily draw just one bounding box in an object detection case, there could be many bounding boxes representing different objects of interest within the image and you would not know how many beforehand.

The major reason why you cannot proceed with this problem by building a standard convolutional network followed by a fully connected layer is that, the length of the output layer is variable — not constant, this is because the number of occurrences of the objects of interest is not fixed. A naive approach to solve this problem would be to take different regions of interest from the image, and use a CNN to classify the presence of the object within that region. The problem with this approach is that the objects of interest might have different spatial locations within the image and different aspect ratios. Hence, you would have to select a huge number of regions and this could computationally blow up. Therefore, algorithms like R-CNN, YOLO etc have been developed to find these occurrences and find them fast.

Proposed Methodology

Cricket is a complex game that has a lot of rules. Thus, a large range of events might be important depending on circumstantial factors.Cricket matches usually have a longer duration than most sports, with the shortest version of the game lasting for 3 hours and the longest lasting for 5 days.Event-driven features are used for extracting the four major events in a cricket match, .i.e., wickets, boundaries, sixes, and milestones, while the remaining important events are identified with the aid of excitement features. Event-driven strategies make use of OCR, playfield scenarios and replays, while the complimentary excitement-based strategy makes use of audio-based classifiers and replays to reinforce the relevance of an event in a cricket match.

Three sub tasks:
o   Player Detection
o   Player Identification
o   Commentary generation

Video shot segmentation

Video shot detection is an important aspect of sports video summarization. A complete cricket match can be broken down into a series of video shots, i.e., a set of frames that are part of the same camera take. Therefore, events in a cricket match can be considered as sequences of video shots.

Scoreboard Detection

The scoreboard refers to the ball-by-ball statistics generated on screen during a cricket match. Scoreboards can be used to predict important events, such as boundaries, sixes, or the total number of wickets that have fallen. They also can tell us about the start or end of an inning. Scoreboards can display constant as well as intermittent information about the match. For example, while the total number of runs and wickets are always displayed on the scoreboard, other stats, like the number of overs, speed of a delivery, and the current score of a batsmen, are displayed once in a while. Accurate detection of scoreboard is important to identify key events because it changes rapidly in a cricket game. Detection of events, such as boundaries, sixes and wickets, rely heavily on the scoreboard. Thus, scoreboard recognition must be performed during the entire course of the game, meaning that the Optical Character Recognition (OCR) engine must be very effective.

Detecting Playfield Scenarios

Once an important event is discovered, a question arises: How the video shots comprising an event should be clipped? The methodology designed to solve this problem relies on recognizing playfield scenarios and uses domain knowledge for labeling the key frame of a given video shot as the starting frame. Almost all events in a cricket match start with the bowler running up to bowl a delivery. In such cases, the camera focuses is either on the bowler or it incorporates the bowler,the striker, the non-striker, and the umpire, along with the pitch. Detecting these key frames would lead to capturing the entire video shot where the bowler runs and bowls a delivery. These video frames can then be used to mark the start of an event. The intuition is to develop a model that is able to distinguish key frames (starting frames) of a video shot that depict a bowler running out to bowl a delivery from other key frames that depict other aspects of the game.

Working
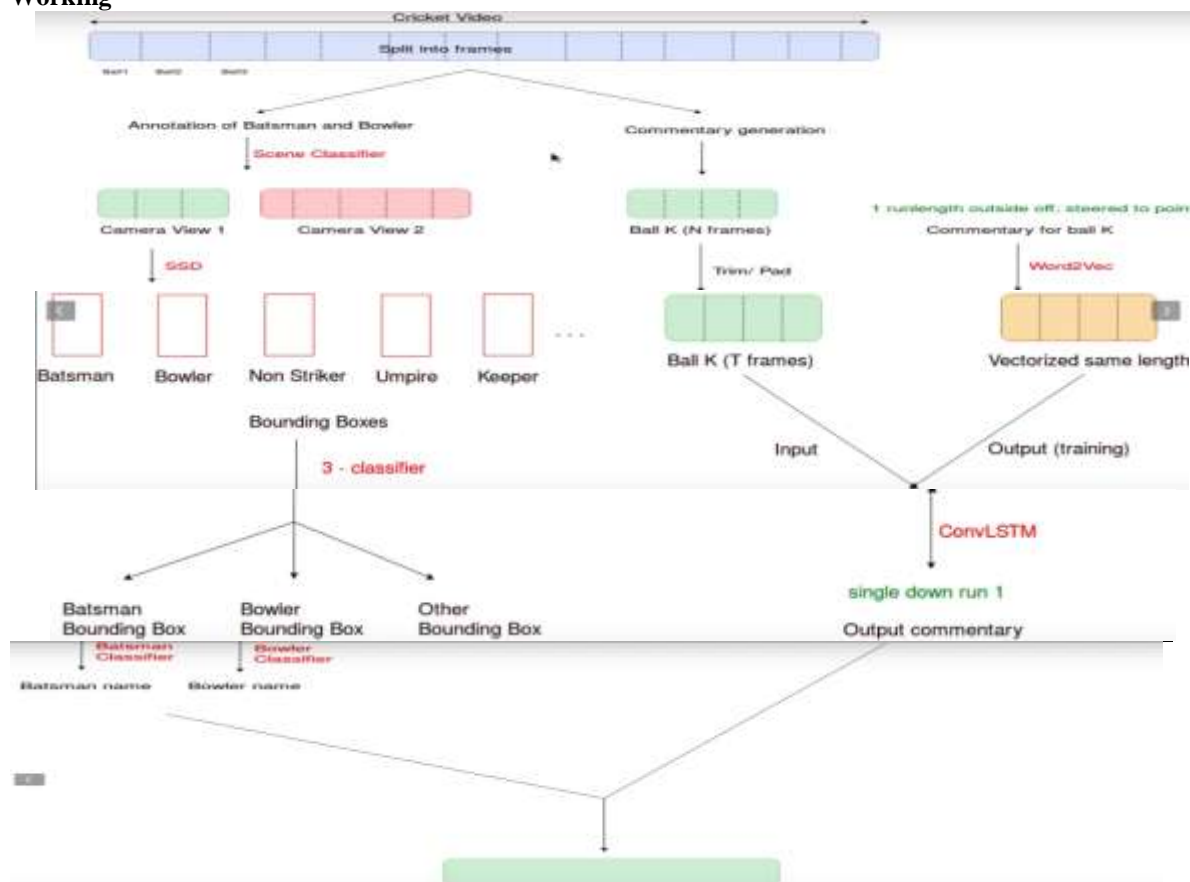
Flowchart showing the working

III. ANALYSIS

In this paper, we present a solution that enables rich semantic annotation of Cricket videos at a fine temporal scale. Our approach circumvents technical challenges in visual recognition by utilizing information from online text-commentaries. We obtain a high annotation accuracy, as evaluated over a large video collection.

- The R-CNN 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.
- Given a cricket video and corresponding text commentary, temporarily align/map them with the video shots. Obtained an accuracy of around 80%.
- Using SSD as the object detection technique, model will tell us the batsman and the bowler by training it for multiple such frames from complete video of cricket match.

- Rest of the information such as the type of the 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 LSTM model which will generate the commentary.
- The final commentary will be the embedding of the batsman and the bowler name with the previous generated commentary of the shots.

Document key expressions have empowered quick and accurate searching for a given archive from a huge text

SYSTEM ANALYSIS
PRELIMINARY INVESTIGATION

Things are expected to get even more critical since the company's growing numbers of clients and related requirements have been projected to demand a massive number of employees in the coming future from the past and the today's date. Such events and projections have forced a strong need for modification in the current way of handling activities. It is better to implement the latest of it rather than to go through the pain of

updating the system over and over again. Also the solution would be developed by in-house developers. Their time have to be managed with their other client dependent schedules.

FEASIBILITY STUDY

Depending on the results of the initial investigation, the survey is expanded to a more detailed feasibility study. Feasibility study is a test of system proposal according to its workability, impact on the organization, ability to meet user needs, and effective use of resources. The objective of the feasibility study is not to solve the problem but to acquire a sense of its scope . During the study, the problem definition is crystallized and aspects of the problem to be included in the system are determined. Consequently, costs and benefits are described with greater accuracy at this stage.

TECHNICAL FEASIBILITY

This involves financial considerations to accommodate technical enhancements. If the budget is a serious constraint, then the project is judged not feasible.

ECONOMICAL FEASIBILITY

With the help of banking application it will lead to decrease in cost of opening and maintaining offices which will be more than the cost of developing and maintaining the Application.

OPERATIONAL FEASIBILITY

This Application is very easy to operate as it is made user friendly. Main consideration is user friendly access to all the functionality of the Application.

IDENTIFICATION OF NEED

To identify the need for software we use the Principles of Requirement Engineering. Requirement engineering provides the appropriate mechanism for understanding what the customer wants, analyzing need, assessing feasibility, negotiating a reasonable solution, specifying the solution unambiguously, validating the specification and managing the requirement as they are transformed into an operational system. The requirement engineering process can be described in five distinct steps: -
Requirement elicitation.
Requirement analysis & negotiation.
Requirement specification.
System Modeling.
Requirement validation.
Requirement Management.

In other words we can say that requirement analysis is a software task that bridges the gap between system level requirement engineering and software design. Requirement analysis allows the software engineering to refine the software allocation and build models of the data, functional and behavioral domains that will be treated by software. Requirement analysis provides the software designer with a representation of information, function and behavior that can be translated into data, architectural, interface and component level design; finally the requirement specification provides the developer and the customer with the means to assess quality once software is built.

The most commonly used requirement technique is to conduct a meeting or interview. The first meeting between a software engineer and customer can be likened to the awkwardness of a first date between to adolescents. Neither person knows what to say or ask, Both are worried that what they do say will be misinterpreted, both are thinking about where it might lead (Both likely have radically different expectation here) Both want to get the think over with, but at the same time, both want it to be a success.

Here according to this principle the analyst starts by asking context-free-questions. That is a set of question that will lead to a basic understanding of the problem, the people who want a solution, the nature of solution that is desired, and the effectiveness of the first encounter itself. The first set of Context-free question focuses on the customer, the overall goals, and the benefits. For example, the analyst might ask: -
Who is behind the request for this work?
Who will use the solution?
What will be economic benefit of a successful solution?
Is there another source for the Solution that you need?

The next set of the questions enables the software engineer to gain a better Understanding of the problem and the customer to voice his or her perceptions about a solution: -
How would you characterize "good" output that would be generated by a successful solution?
What problem(s) will this solution address?
Can you show me (or describe) the environment in which the solution will be used?
Will special performance issues or constraints affect the way of the solution is approached?
Are you the right person to answer these questions?
Are your answers? Official"?

Preliminary Investigation

The first step in the system development life cycle is the preliminary investigation to determine the feasibility of the System. The purpose of the preliminary investigation is to evaluate project requests. It is neither a design study nor does it includes the collection of details top describe the business system in all respect. Here for the "PedisAlert", I have worked on the preliminary investigation that accomplished the following objectives:

Clarify and understand the project request.

Determine the size of the project.

Assess costs and benefits of alternative approaches.

Determine the technical and operational feasibility of alternative approaches.

Feasibility Study

Not everything imaginable is feasible, not even in software, as it may appear to outsiders. Feasibility is the determination of whether or not a project is worth doing. On the contrary, software feasibility has seven solid dimensions as below: -
1. Technical feasibility.
2. Operational feasibility.
3. Economic feasibility.

The process followed in making this destination is called a feasibility study. This type of study determines when a project can or should be taken. Once it has been determined that the project is feasible, and then analyst can go ahead and prepare the project specification, which finalizes project requirements. Generally, feasibility studies are undertaken within time constraints and normally written in feasibility report. The contents and recommendation of such a study will be used as a sound basis for deciding whether to proceed, postpone or cancel the project.

In other words we can say that a feasibility study is conducted to select the best system that meets performance requirement. This contains identification description, an evaluation of candidate system and the selection of best system for the job.

Project Planning

The planning, design, and installation of a system is termed a project and is directed by a project leader who uses the available resources to produce a new or better system for the organizations.

Thus, project planning for any company has the following four main steps:
1. Organizing the resources available for the project.
2. Scheduling the events of the project.
3. Evaluating the progress.
4. Establishing standards for the project.

Project Scheduling

In project management, a schedule consists of a list of a project's terminal elements with intended start and finish dates. Terminal elements are the lowest element in a schedule, which is not further subdivided. Those items are often estimated in terms of resource requirements, budget and duration, linked by dependencies and scheduled.

The chart techniques are the scheduling tools of the project planner. Even the simplest project should be framed as a chart so that progress can be measured. The Gantt chart is effective for simple projects, especially when the interrelationships among events are not too complex. Complicated scheduling usually requires a PERT chart.

A schedule must be flexible because unexpected events may occur that may affect the schedule for the development of the systems.

If a project leader cannot enforce a schedule, someone else should be leading the project. The following are variety of options that are open to the project leader:
• Increase the budget.
• Increase manpower in the form of overtime or additional people.
• Add equipment.
• Change priorities.
• Replace the individual responsible.

GANTT CHART OF THE ACTIVITY

A Gantt chart is a type of bar chart that illustrates a project schedule. Gantt charts illustrate the start and finish dates of the terminal elements and summary elements of a project.

Development Phase	120 Days						Duration (Day)
	0 to 15 Day	16 to 30 Day	31 to 45 Day	46 to 60 Day	61 to 75 Day	76 to 90 Day	
Requirement Gathering	▭						10
Analysis	▭						15
Design		▭					30
Coding			▭				25
Testing					▭		18
Implementation						▭	08
Documentation	▭						
Total Time (Days)							90

Terminal elements and summary elements comprise the work breakdown structure of the project. Some Gantt charts also show the dependency (i.e., precedence network) relationships between activities.
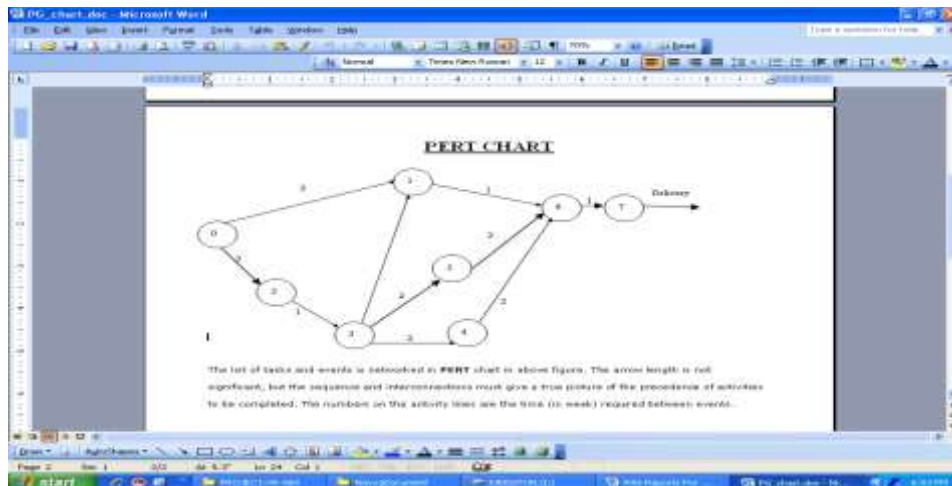
The numbers on the activity lines are the time (in week) required between events.

IV. CONCLUSION

In this paper, we present a solution that enables rich semantic annotation of Cricket videos at a fine temporal scale. Our approach circumvents technical challenges in visual recognition by utilizing information from online text-commentaries. We obtain a high annotation accuracy, as evaluated over a large video collection.

- The R-CNN 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.
- Given a cricket video and corresponding text commentary, temporarily align/map them with the video shots. Obtained an accuracy of around 80%.
- Using SSD as the object detection technique, model will tell us the batsman and the bowler by training it for multiple such frames from complete video of cricket match.
- Rest of the information such as the type of the 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 LSTM model which will generate the commentary.
- The final commentary will be the embedding of the batsman and the bowler name with the previous generated commentary of the shots.

V. ACKNOWLEDGEMENTS

REFERENCES

- https://www.google.com/url?sa=t&source=web&rct=j&url=https://cse.iitk.ac.in/users/cs365/2014/_submissions/sumedh/project/report.pdf&ved=2ahUKEwidl6-x5qjmAhUAzDgGHSApBagQFjADegQIBRAB&usg=AOvVaw2tHMlP6BB-sFjB7aMlXZXY&cshid=1575902561721
- https://www.google.com/url?sa=t&source=web&rct=j&url=https://pillow.readthedocs.io/en/3.1.x/reference/Image.html&ved=2ahUKEwi-tIvu5qjmAhUGwTgGHaaXAAUQFjAAegQIBBAC&usg=AOvVaw3s3-ER8v99-Z-1JlI086gn&cshid=1575902656350
- https://ieeexplore.ieee.org/abstract/document/6012139/
- https://ieeexplore.ieee.org/abstract/document/7101847/
- https://scholar.google.co.in/citations?user=pslH9f0AAAAJ&hl=en&oi=sra
- https://www.youtube.com/watch?v=DZOQS_a0dEo&feature=youtu.be
- https://youtu.be/Kwu1yIC-ssg