IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Fast Moving Object Detection Using Trajectory Based Estimation

Nandini N
Associate Professor
Department of Computer Science and Engineering,
Dr. Ambedkar Institute of Technology,
Bengaluru-560056, India

Chandana N
Department of Computer Science and Engineering,
Dr. Ambedkar Institute of Technology,
Bengaluru-560056, India

Ashwas Ajay
Department of Computer Science and Engineering,
Dr. Ambedkar Institute of Technology,
Bengaluru-560056, India

Bhargav Urs S
Department of Computer Science and Engineering,
Dr. Ambedkar Institute of Technology,
Bengaluru-560056, India

Deeksha H T
Department of Computer Science and Engineering,
Dr. Ambedkar Institute of Technology,
Bengaluru-560056, India

Abstract : The goal of this research is to create a reliable trajectory-based estimating system for fast-moving object detection. Utilizing the YOLOv8 model to find, track, and estimate the trajectory of a cricket ball in low-resolution MP4 footage taken with a smartphone is the specific goal. The output consists of a reconstructed MP4 movie that follows the ball precisely and shows its trajectory path as a series of dots in each frame. The dataset utilized is made up of 18,000 1920x1080 resolution frames. Each frame is annotated by utilizing the "makesense.ai" website to build a bounding box around the cricket ball and name it as a "ball". The YOLOv8 model has a mean average precision (mAP) of 99.1%, precision of 98.5%, and recall of 96.9% after being trained for 114 epochs with a batch size of 16. Additionally, a user-friendly interface makes it simple for users to choose and upload movies for prediction. With a focus on determining the trajectory of cricket balls, this study seeks to offer an efficient method for the precise identification and tracking of swiftly moving objects.

***IndexTerms* - Deep Learning, YOLOv8, Cricket ball tracking, tiny object detection, Sports AI, Trajectory estimation.**

I. INTRODUCTION

A prominent area of research now is the extraction of information from films, particularly in the context of image processing and deep learning. Videos contain a huge number of visual sensor records that can provide useful information. Videos are essential for post-game analyses and tactical analysis in applications like sports analysis and athlete training. High-resolution and high-frame-rate video systems, which are frequently utilised in professional sports for data collecting or referee aid, are, however, generally pricey and unavailable to consumers or amateurs. Therefore, to enable large-scale sports data collecting, it is imperative to design a low-cost method for data acquisition from broadcast footage.

One of the most valuable pieces of information for game analysis is ball trajectory data. However, in sports like cricket, tennis, badminton, and baseball, tracking the ball presents unique challenges. Images of the ball are small and distorted due to its small size and its fast movements, which can reach several hundred km per hour. We are building a project that precisely positions cricket balls in films taken by consumer electronics like cellphones using a deep learning network called YOLOv8 to address this issue. By learning the trajectory information of the balls, the model is able to detect occluded balls and overcome problems like residual and hazy images.

In the current system, scientists and engineers have created innovations like Hawkeye [1], which helps umpires by using several high-resolution and high-frame-rate cameras to aid in decision-making. These solutions, however, require a substantial hardware investment and are primarily made for professional cricket. There aren't many object-tracking methods that are explicitly designed to help umpires track cricket balls. The aim of this project is the construction of a system that continuously monitors the ball, tracks its trajectory, and gives precise placement. The goal of the suggested approach is to assist umpires in making accurate decisions at a low cost. It particularly targets local teams and up-and-coming aspiring amateur players who want to grow and become professionals.

NEED OF THE STUDY.

The study presented in this paper is driven by the need to address several critical aspects related to ball tracking in sports videos. These aspects include:

- **Accessibility and Affordability:** High-resolution and high-frame-rate video systems used in professional sports for data collection or referee assistance are often expensive and not accessible to individuals or amateur teams. There is a need to develop a low-cost solution that allows for data acquisition from broadcast footage, enabling wider participation and massive sports data collection.
- **Challenging Ball Tracking:** Tracking the ball in fast-paced sports such as cricket, tennis, badminton, and baseball is particularly challenging due to the small size of the ball and its high-speed movements. These factors result in small and distorted images, making it difficult to accurately track the ball using conventional methods.
- **Post-Game Analysis and Tactical Insights:** Videos play a crucial role in post-game analysis and tactical insights for sports teams and athletes. Ball trajectory data is one of the fundamental pieces of information required for game analysis, providing valuable insights for improving performance, strategic decision-making, and training.
- **Umpiring Assistance:** Ball tracking technologies can assist umpires in making accurate decisions during matches. Existing solutions, such as Hawk-eye, primarily cater to professional cricket and involve substantial hardware investments, limiting their accessibility to non-professional teams and local matches. There is a need for affordable ball tracking systems that can provide accurate information to support umpires in decision-making.

Overall, studying pregnancy assistance is vital for improving maternal and child health, filling information gaps, enhancing accessibility to care, empowering pregnant women, promoting preventive care, and supporting healthcare professionals. It allows us to develop innovative solutions that leverage technology to provide valuable support and guidance during this important stage of life.

II. LITERATURE REVIEW

[1] The Hawk-Eye system, which is owned by Sony, was created by Paul Hawkins in the United Kingdom. It has been widely utilized in professional tournaments to determine the trajectories of balls and aid referees in resolving contentious decisions by providing 3D visual representations. However, the system requires expensive high-end cameras and dedicated operators placed strategically at specific positions and angles. As a result, the cost associated with implementing the Hawk-Eye system is prohibitive for non-professional teams.

[2] Udit Arora, Sohit Verma, Sarthak Sahni, and Tushar Sharma proposed a method for detecting the cricket ball using Support Vector Machine (SVM) and Histogram of Oriented Gradient (HOG). The authors collect positive and negative data samples to train SVM models specifically for HOG objects such as the ball and batsman. The video is divided into multiple image frames with a fixed time interval. The frame difference technique is utilized to identify regions that are likely to contain the ball. Each window frame is converted to grayscale, and HOG features are extracted from these frames. These features are then inputted into the SVM model to detect the ball in the areas that exhibit differences when compared to subsequent frames using the frame subtraction method.

[3] M. Archanaa and M. Kalaisevi Geetha proposed a method for ball detection in tennis videos. They employ frame differencing to identify potential ball candidates, which are then verified based on size and shape parameters. In cases where multiple detections occur, the technique selects the region closest to the expected position. To address the issue of double detections, the model applies logical AND operations between the image difference result and the background image region within the expected intensity range for the ball.

[4] Arjun Nelikanti, G. Karuna, and G. Venkata Rami Reddy focused on detecting object regions in video sequences under various occlusion conditions. The paper utilizes the Kalman filter to detect objects following different trajectories within a fixed time window. The proposed approach demonstrates successful detection of thrown objects in diverse scenarios, including both occlusion and non-occlusion conditions, and with different background environments.

III. RESEARCH METHODOLOGY

3.1 Population and Sample

In the context ball tracking, the population in this case would be all the potential cricket training centers or academies where cricket footage can be recorded for ball tracking analysis. The population represents the entire set of such establishments.
The sample refers to the specific cricket training center called "Jawahars Cricket Academy" where we collected our data. This training center is a subset of the population mentioned above.

To further elaborate, the data collection process involved recording approximately 6 hours of low-resolution MP4 footage at the "Jawahars Cricket Academy." This sample represents a specific cricket training center from which we obtained our data. The dataset was then cleaned and processed, resulting in 18,000 frames with a resolution of 1920x1080 pixels.
It's important to note that while the sample represents a single cricket training center, the collected data and subsequent analysis can provide insights and generalizability beyond the specific training center. The findings and models developed using this sample can potentially be applied to other similar cricket training centers and even broader contexts involving ball tracking in cricket.

3.2 Data and Sources of Data

To As the researcher collecting the data for our project, we visited the "Jawahars Cricket Academy," which served as our cricket training centre and data source. To gather our dataset, we recorded several hours of low-resolution MP4 footage using a smartphone and tripod stand. We ensured that the footage captured various scenarios and movements relevant to ball tracking in cricket.

After the data collection phase, we carefully cleaned and processed the footage. From the initial 6 hours of recorded videos, we extracted and isolated individual frames for analysis. The resulting dataset consisted of 18,000 frames, each with a resolution of 1920x1080 pixels. To annotate the frames, we employed the "makesense.ai" website, which facilitated the process of labeling and creating bounding boxes around the cricket ball. We manually reviewed each frame, precisely drawing bounding boxes around the ball and labeling it as a "ball." This annotation process ensured that our dataset was accurately labeled and ready for training and evaluation of our deep learning model.

By utilizing the "Jawahars Cricket Academy" as our data source and employing smartphone-recorded footage, we aimed to simulate real-world conditions and challenges faced in ball tracking during cricket matches. This approach provided us with a diverse and representative dataset for training our model and addressing the objectives of our project effectively.

3.3 Theoretical framework

The theoretical framework for the project can be outlined as follows:

- **Computer Vision:** The project is grounded in the principles of computer vision, a field of study focused on enabling computers to gain a high-level understanding of visual data. Computer vision techniques are utilised to analyse and extract meaningful information from the low-resolution MP4 videos recorded at the cricket training centre.
- **Object Detection and Tracking:** Object detection and tracking are core components of the project's theoretical framework. These techniques involve the identification and localisation of objects within images or videos, as well as the subsequent tracking of their movements across frames. The YOLOv8 model, a state-of-the-art deep learning architecture for object detection, is employed to accurately detect and track the cricket ball in recorded videos.
- **Deep Learning:** Deep learning serves as the underlying theoretical framework for the project. Deep learning algorithms, such as YOLOv8, leverage artificial neural networks with multiple layers to learn complex patterns and features directly from the data. By training the YOLOv8 model on the annotated dataset, the project harnesses the power of deep learning to achieve robust and accurate detection and tracking of the cricket ball.
- **Data Annotation:** Data annotation is a crucial step in the project's theoretical framework. Each frame in the dataset is manually annotated by drawing bounding boxes around the cricket ball and labelling it as a "ball." This annotated dataset is then used to train the YOLOv8 model, enabling it to learn the visual characteristics of the cricket ball and accurately detect and track it in the videos.
- **Interactive User Interface:** The project incorporates the development of an interactive web-based user interface using Python's Streamlit framework. This interface allows users to conveniently upload their videos for prediction and visualise the results. The user-friendly interface enhances the accessibility and usability of the system, providing a seamless experience for users to interact with the YOLOv8 model and observe the detected and tracked cricket balls in the uploaded videos.
- **Video Analysis and Sports Applications:** The theoretical framework also recognises the broader context of video analysis and its applications in sports. Post-game analysis, tactical insights, and umpiring assistance are identified as key motivations for accurate ball tracking. By leveraging computer vision techniques and the YOLOv8 model, the project aims to provide valuable information for performance improvement, strategic decision-making, and training in cricket. The use of YOLOv8 enables the accurate estimation of the ball's trajectory, facilitating comprehensive analysis and support for players, coaches, and umpires.

Overall, the project's theoretical framework combines principles from computer vision, deep learning, data annotation, and interactive user interface development. By leveraging computer vision techniques and YOLOv8, the project aims to develop a robust system for fast-moving object detection, specifically targeting accurate ball tracking and trajectory estimation in cricket videos.

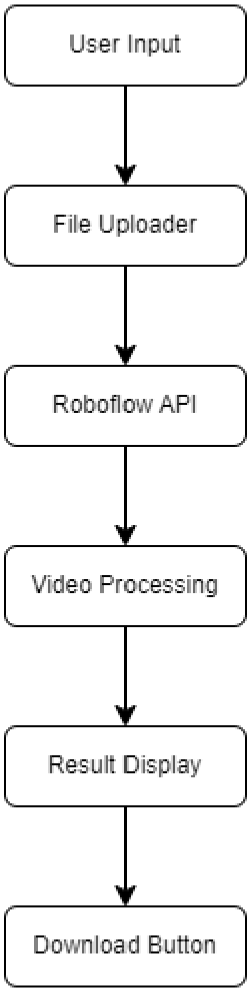

Fig 2.1 Work Plan of the System

Figure 2.1 represents the system's work plan, describing the overall functioning of our model. The system begins with an upload page where users can submit a video for training. They have the option to either drag and drop the video file or browse their system contents to upload it. The maximum file size allowed for upload is 200MB, and the supported file types are limited to mp4 and mpeg4. The dataset given to Roboflow includes YOLO labels along with the associated images. YOLOv8, developed by Ultralytics, is considered the latest version of the YOLO (You Only Look Once) architecture and is known for its state-of-the-art performance. Once the video is uploaded, it undergoes training on Roboflow, specifically for detecting the ball and its trajectory. The video is divided into frames, which are then sent to the Roboflow project API. In response, the API provides location coordinates of the ball within each frame. These coordinates are used to mark a circle on the corresponding frames. Finally, the original video is reconstructed based on its original frames per second (fps) value.

3.4 Statistically descriptive:

The properties of the gathered data can be summed up and described using descriptive statistics. Measures like mean, median, standard deviation, and range can shed light on how variables like ball trajectory, speed, or distance traveled are distributed and vary.

3.4.1 Hypothesis testing:

To evaluate the importance of links or differences within the data, hypothesis testing can be used. To find out if there are appreciable variations in ball trajectory between various players or training sessions, for instance, statistical tests like t-tests or chi-square tests might be performed.

3.4.2 Regression Analysis:

Regression analysis can be used to examine how different variables are related. For instance, a regression model could be created to investigate how variables like ball speed, pitch conditions, and player technique affect ball trajectory. In order to evaluate the combined effect of multiple variables on the trajectory of the ball, multiple regression can be utilized.

3.4.3 Time series analysis

Time series analysis can be used to look for patterns and trends in data on ball trajectories over time. To find seasonality, trends, and other temporal patterns in the ball trajectory, methods like autoregressive integrated moving average (ARIMA) or seasonal decomposition of time series (STL) can be utilized. (RNNs) can be used to identify intricate correlations and forecast future ball trajectories with accuracy.

3.4.4 Machine learning techniques:

Machine learning techniques can be used to analyze and forecast ball trajectory based on past data using a variety of machine learning methods. Recurrent neural networks (RNNs), decision trees, and random forests are some examples of models that can be used to capture complicated relationships and predict future ball trajectories with accuracy.

3.4.5 Spatial analysis:

Examining the spatial distribution and patterns of ball trajectories can be done using spatial analysis techniques. Ball trajectory data can be used to find clusters or spatial dependencies using tools like spatial autocorrelation analysis and spatial clustering algorithms.

The links, patterns, and trends found in the gathered data can be understood using these statistical methods and econometric models. They can help in determining the elements affecting ball trajectory, identifying performance variations, and formulating predictions or suggestions for enhancing training or games tactics. The kind of data being used, the goals of the research, and the exact hypothesis being evaluated will all influence the tools and models that are specifically chosen.

IV. RESULTS AND DISCUSSION

The results of the project include:

- **Ball Detection and Tracking:** The implementation of the YOLOv8 model for ball detection and tracking is expected to yield accurate and reliable results. The model should be able to successfully identify and localize the cricket ball in each frame of the low-resolution MP4 videos.

- **Trajectory Estimation:** By continuously tracking the ball across frames, the system should be able to estimate the trajectory of the cricket ball. The trajectory estimation results can be visualized by displaying the path of the ball through a series of dots or markers overlaid on the video frames.

- **User Interface and Video Upload:** The interactive web-based user interface developed using Python's Streamlit framework allows users to conveniently upload their videos for prediction. The interface should provide a user-friendly experience and enable users to visualize the detected and tracked cricket ball in the uploaded videos.

- **Accuracy and Performance Metrics:** The YOLOv8 model's performance can be evaluated using various metrics such as mean average precision (mAP), precision, and recall. The reported mAP of 99.1%, precision of 98.5%, recall of 96.9% and F1 Score of 29.07 indicate impressive accuracy in ball detection and tracking.

- **Practical Applications:** The accurate detection, tracking, and trajectory estimation of the cricket ball have practical applications in sports analysis, post-game analysis, and umpiring assistance. The results obtained from the system can provide valuable insights for improving performance, strategic decision-making, and training in cricket.

The performance evaluation of the above model is presented using a confusion matrix, depicting the classification results for the detected objects and their ground truth labels.

	Actual Positive	Actual Negative
Predicted Positive	969	Unknown
Predicted Negative	31	Unknown

Fig.4.1. First page

Figure 4.1 portrays the first page of the trajectory detection system, presenting a visual representation of the interface users encounter during their initial interaction with the detection model.The image displays various elements such as an upload option, a limit on video size, acceptable video formats, and a browse button for selecting and uploading the desired video.This depiction aims to illustrate the user-facing aspect of the detection model, specifically highlighting the components related to uploading videos.

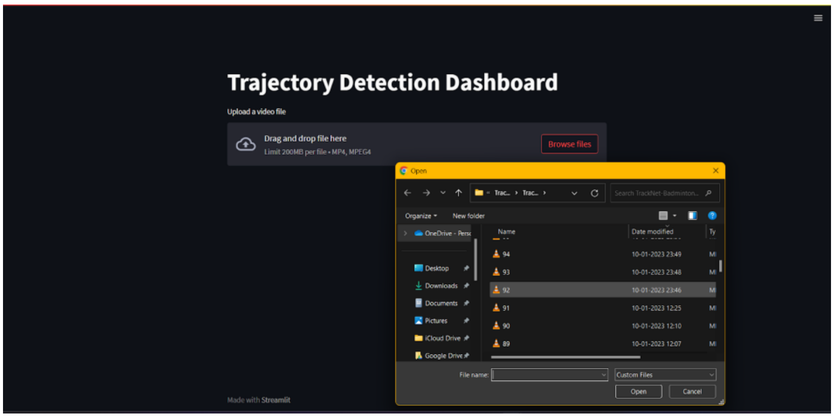

Fig.4.2. Upload video

Figure 4.2 When the user interacts with the "browse" button, they are presented with a prompt to select a file for uploading. The visual representation in the provided image illustrates this step, which serves as the initial stage of a processing model.This user upload portion allows individuals to choose a file that will be used for further processing.

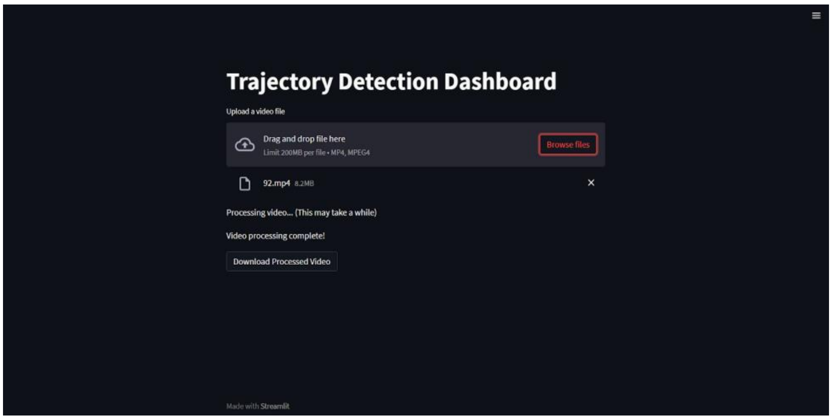

Fig.4.3.Video Processing

Figure 4.3 visually represents the video processing step within the system, offering an illustration of the processing phase along with the displayed outcome. The image highlights the processing stage and provides a visual representation of the result that is generated as a part of this step. Additionally, the depicted interface allows users to download the obtained result, giving them the option to save and access it for further use.

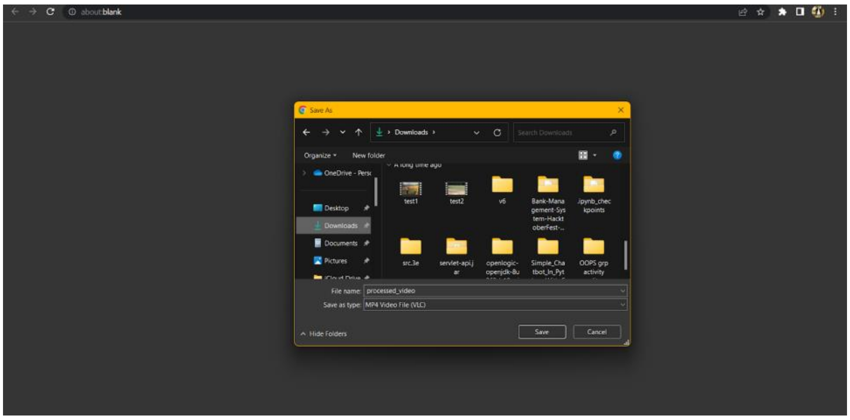

Fig.4.4. Downloading the processed video

Figure 4.4 visually represents the post-download action that occurs after the user clicks the download button. The image showcases the interface where the user can specify the desired location for saving the video file they have downloaded. This depiction emphasizes the user's ability to control the location where the downloaded video will be stored, enabling them to choose a convenient location for viewing.

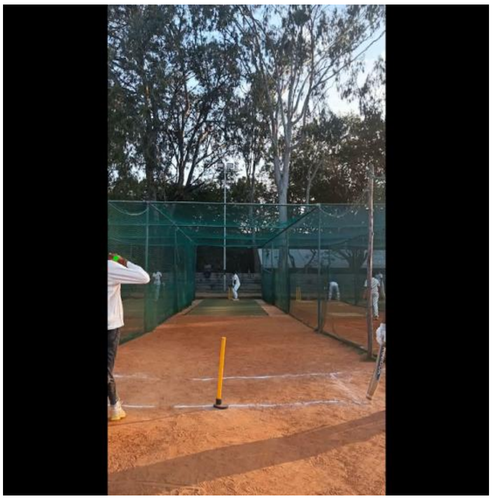

Fig.4.5. Ball detection

The result video provides visual information about the ball and its trajectory. Figure 4.5 specifically displays a single frame from the result video, showcasing the ball and its corresponding trajectory. This image serves as an example, giving us a glimpse of the captured frame that highlights the ball's presence and the path it follows.

Fig.4.6.Trajectory and trajectory pattern

Figure 4.6 provides a visual representation of the ball's trajectory. By observing successive frames in the video, additional details of the ball's trajectory become more apparent. This image serves as an example, illustrating how the trajectory evolves and becomes more discernible as more frames are observed.

Fig.4.7. Trajectory pattern

Figure 4.7 displays the concluding frames of the result video, showcasing the trajectory of the ball. By examining these final frames, one can gain a better understanding of the game by having appropriate knowledge of the ball's position. This image serves as a visual aid, offering valuable insights into the ball's path and enabling a deeper comprehension of the game.

REFERENCES

[1] Hawk-eye, https://en.wikipedia.org/wiki/Hawk-Eye.

[2] Cricket umpire assistance and ball tracking system using a single smartphone camera Udit Arora1, Sohit Verma1, Sarthak Sahni1,Tushar Sharma1, Division of Computer Engineering, Netaji Subhas Institute of Technology, New Delhi 110078, India,10 Nov 2017.

[3] M Archana and M. K. Geetha, “Object detection and tracking based on trajectory in broadcast tennis video,” Procedia Computer Science, vol. 58, pp. 225–232, 2015.

[4] FMODetect: Robust Detection of Fast Moving Objects, Denys Rozumnyi1,Jiri Matas, Filip Sroubek, Marc Pollefeys, Martin R. Os, arXiv: 2012.0821 6v2, 17 Aug 2021.

[5] “Moving Object Detection and Tracking under Different Occlusion Conditions in Video Sequences” Arjun Nelikanti, G. Karuna, G. Venkata Rami Reddy, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.