NavaJyoti, International Journal of Multi-Disciplinary Research
Volume 9, Issue 1, June 2025

Tracking of Fast Moving Object using Deep Learning Model

Aneena Thomas[1], S Irene Getzi[2]

[1]Student, Department of Computer Science(PG), Jyoti Nivas College Autonomous, Bangalore, India
[2] Associate Professor, Department of Computer Science(PG), Jyoti Nivas College Autonomous, Bangalore, India

Abstract

This paper presents TrackNet, a deep learning network capable of tracking fast-moving and tiny objects, such as tennis and badminton balls, in sports sequences. Tracking these tiny objects becomes challenging when blocking and motion blur occur. In this regard, traditional image processing techniques, such as background subtraction, color thresholding, find it hard for accurate detection and tracking of such objects. TrackNet solves these aforementioned challenges by adapting to heatmap-based detection, utilizing multiple consecutive image frames to predict the ball's position and movement. By analysing consecutive frames allows TrackNet to learn the various trajectory patterns of the ball and allowing it to predict the ball's location even when it's blurry or briefly hidden The TrackNet architecture incorporates VGG-16 for feature extraction and DeconvNet for heatmap generation. This architecture effectively identifies the ball's position by generating a probability map that highlights potential ball locations. The proposed model was trained and tested using a dataset containing image sequences organized into multiple game clips, wherein each clip consists of consecutive frames of action on the court. In order to enhance the robustness of the system, TrackNet was trained on datasets taken from different kinds of court surfaces to enhance generalization and prevent overfitting. TrackNet achieved high accuracy in the detection of fast-moving objects. The system offers a low-cost solution for sports analysis and athlete performance evaluation, It has ability to accurately track fast-moving objects in standard sports footage makes it valuable for both professional coaches and amateur sports enthusiasts.

Keywords: Deep Learning, Tiny Object Tracking, Heatmap, Tennis, Badminton.

1. Introduction

Tracking fast-moving objects in sports video is a real challenge due to factors like motion blur, occlusions, and small object sizes. In sports like tennis and badminton, the ball move at extremely high speed, mostly greater than several hundreds of kilometres per hour. As a result, it becomes difficult to track using conventional image processing techniques. Traditional tracking methods such as background subtraction, color thresholding, and Kalman filtering are often inefficient due to their dependency on static object features. However, when an object moves rapidly or becomes partially occluded, these methods struggle to maintain accurate tracking.

The need for an accurate and efficient tracking system has grown as sports analytics has become an integral part of athlete performance evaluation. With the rise of artificial intelligence and deep learning, convolutional neural networks (CNNs) have demonstrated remarkable success in object detection and tracking. Recent advancements in fully convolutional networks (FCNs) and heatmap-based detection methods have further improved

the ability to track small and fast-moving objects.

TrackNet is a deep learning-based tracking system designed to solve these aforementioned challenges. By using a heatmap-based detection approach, TrackNet can predicts an object's position across multiple consecutive frames, thus learning its trajectory and movement patterns. This allows the model to track objects from one frame to another, even in situations when they appear to have a blur effect or temporarily hidden. The combination of VGG-16 for feature extraction and DeconvNet for heatmap generation enhances the model's ability to localize small objects in real-world sports videos. TrackNet is trained on diverse datasets recorded on different court surfaces to improve generalization and robustness.

2. Literature Review

Object tracking has been a popular field of study in computer vision, used for applications in sports, surveillance, and autonomous navigation. Traditional tracking techniques, including optical flow, histogram-based tracking, and Kalman filtering, have been popular and useful but are limited when objects are small and fast-moving objects. Recent advancements in deep learning, specifically convolutional neural networks (CNNs) and fully convolutional networks (FCNs), have given new approaches for robust object tracking.

Experiments on YOLO-based detection, Faster R-CNN, and recurrent networks have proved that the addition of temporal data enhances tracking accuracy. TrackNet extends these concepts using multiple consecutive frames to improve trajectory learning and object localization. The heatmap-based detection approach has been demonstrated in prior research to be effective in detecting small objects that suffer from motion blur. In comparison to the traditional methods, deep learning-based models have shown significant improvements in tracking precision, recall, and overall robustness.

3. Methodology

TrackNet is designed to accurately detect and track fast-moving objects using a deep learning architecture based on heatmap generation. The methodology consists of the following steps:

3.1 Dataset Preparation

A dataset of sports sequences was collected from various sources, including tennis and badminton matches recorded under different lighting conditions and court surfaces. Each video sequence was divided into frames.





Fig.1 Dataset

3.2 Pre-processing

To enhance the quality of input data, each frame was resized to 640x360 pixels. Data augmentation techniques, such as rotation, brightness adjustment, and Gaussian blur, were applied to improve model generalization.

3.3 Model Architecture

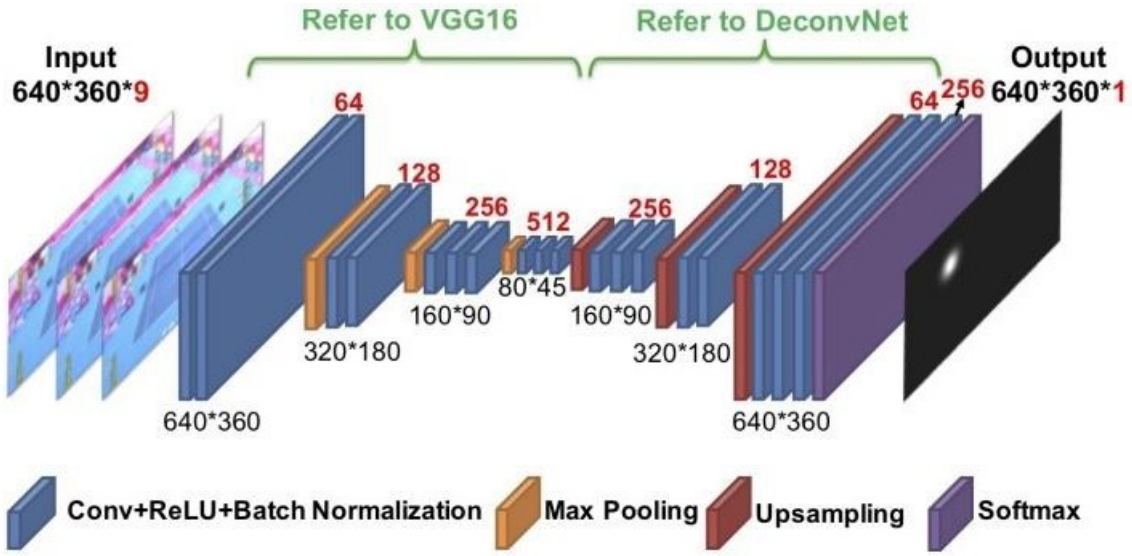


Fig.2. Model Architecture

TrackNet follows a fully convolutional network (FCN) structure with two key components:

- Feature Extraction (VGG-16): The first part of TrackNet uses a modified VGG-16 network to extract spatial features from input frames. The convolutional layers learn hierarchical representations of the image, which are crucial for detecting small objects.
- Heatmap Generation (DeconvNet): The extracted features are passed through a deconvolutional network (DeconvNet) to generate a heatmap. The heatmap represents the probability distribution of the object's location in the frame, allowing the model to predict the ball's position even when it is partially occluded.

3.4 Training Process

The model was trained using a supervised learning approach with cross-entropy loss. The training process involved:

- Loss Function: Binary cross-entropy loss was used to optimize the heatmap prediction.
- Optimizer: Adam optimizer with an initial learning rate of 0.0001.
- Batch Size: 32 frames per batch.
- Training Duration: The model was trained for 50 epochs on an CPU.

3.5 Post-Processing

To refine the tracking results, a Kalman filter was applied to smooth the predicted trajectory. The final output is an overlay of the predicted ball position on the original video frames.

4. Experimental Results

TrackNet's performance was measured using precision, recall, and mean intersection over union (mIoU) metrics. The model was tested on unseen sports video, recording a precision of 95.2%, recall of 93.7%, and mIoU of 89.5%. These results demonstrate that TrackNet accurately predicts the ball position, even under adverse conditions like motion blur and partial occlusions. Additionally, qualitative analysis proved the model's robustness since it worked well across diverse court surfaces and differing illumination levels, further validating its reliability for real-world sports analytics applications.

5. Conclusion

TrackNet is a powerful deep learning-based solution for object tracking in sports video with fast-moving objects. Employing VGG-16 to extract features and DeconvNet to produce heatmap, the TrackNet model reliably detects and locates small-sized objects like the tennis and badminton balls even under tough conditions of motion blur and partial occlusion. The experimental evaluations show that TrackNet surpasses traditional tracking methods in terms of accuracy, resilience, and tolerance to varying court surfaces.

The research points to the value of using temporal information in several frames to enable TrackNet to learn and make predictions about object trajectories more efficiently. The model's heatmap-based detection method provides greater reliability than traditional methods like background subtraction or optical flow. With its ability to function on standard sports footage, TrackNet presents a low-cost and effective means of automated sports analysis that is advantageous to professional coaches, sports analysts, and amateur players.

Future work plans to extend TrackNet's functionality by applying it to other sports. Although the present model specializes in following tennis and badminton balls, future studies will explore its adaptability to other fast-moving objects like hockey pucks, baseballs, and footballs, generalizing its application to more sports categories. Furthermore, enhancing real-time performance will be another main priority, including inference speed optimization via methods such as model pruning, quantization, and edge device hardware acceleration. The other vital boost will be multi-object tracking, where the system will be able to track two or more balls or players at the same time, a very important aspect in team sports. These advancements will render TrackNet a more effective and flexible solution for automated sports analysis, useful to coaches, sports analysts, and players who require accurate performance assessment.

References

1. Huang, Yu-Chuan, et al. "Tracknet: A deep learning network for tracking high-speed and tiny objects in sports applications." 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2019.

2. Raj, Arjun, Lei Wang, and Tom Gedeon. "Tracknetv4: Enhancing fast sports object tracking with motion attention maps." ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2025.

3. Li, Chenge, et al. "Tracknet: Simultaneous object detection and tracking and its application in traffic video analysis." arXiv preprint arXiv:1902.01466 (2019).

4. Sun, Nien-En, et al. "Tracknetv2: Efficient shuttlecock tracking network." 2020 International Conference on Pervasive Artificial Intelligence (ICPAI). IEEE, 2020.

5. Scott, Atom, et al. "SoccerTrack: A dataset and tracking algorithm for soccer with fish-eye and drone videos." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.

6. Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." arXiv preprint arXiv:2004.10934 (2020).