Find 10 videos online showing objects moving from one place to another. Then, create a machine learning model that can accurately track the locations of these moving things in real-time.

Department of ECE
North South
University Dhaka,
Bangladesh

CSE 445.11
Project Group no: 07

Md Yeasin Arafat (2022320642), Ishfaque Islam (2022513 0 42), Murad Mustakim (2011650042),
Samor Deb Nath Konok (2014206642) Mahir Absar (2011685042)

Abstract—
In this reportIn this report, we present a machine learning-based system for real-time object tracking in football match videos. Our model integrates YOLO for object detection and ByteTrack for multi-object tracking, ensuring accurate and efficient tracking of players and the ball. We collected 10 football match videos and initiated the annotation process. So far, we have auto-annotated 1,000 frames and are actively working on manual annotation for improved accuracy. The system processes video frames, detects objects, and tracks their movement across sequences, handling challenges such as occlusions and fast motion. Our approach demonstrates high accuracy in object localization and tracking, with potential applications in sports analytics, player performance evaluation.

I. INTRODUCTION

Tracking moving objects in videos is a crucial task in computer vision, with applications in sports analytics, surveillance, and autonomous systems. Traditional tracking methods struggle with challenges such as occlusions, fast object movement, and varying lighting conditions. To overcome these limitations, we propose a deep learning-based approach that integrates YOLO for real-time object detection and ByteTrack for robust multi-object tracking. Our project focuses on tracking players and the ball in football match

videos to analyze movement patterns and interactions. We have collected 10 football match videos and initiated the annotation process, with 1,000 frames auto-annotated and ongoing manual annxotation efforts. most representative platforms for pre- owned bike listings in Bangladesh.

II. LITERATURE REVIEW

Real-time object tracking is crucial for applications in sports analytics, surveillance, and autonomous systems. YOLO is favored for its speed and efficiency in real-time tasks, while Faster R-CNN offers higher accuracy but at a slower speed. For tracking, SORT provides fast tracking using Kalman filtering but lacks robustness, while Deep SORT improves this by handling occlusions and identity switches. ByteTrack further refines tracking by associating both high and low-confidence detections, improving performance in complex environments.

Our project leverages YOLO for object detection and ByteTrack for multi-object tracking in football match videos. Despite its promising results, challenges like occlusions, motion blur, and varying frame rates remain, which we aim to address for improved accuracy and robustness.We aim to tackle these challenges for more accurate real-time tracking.

III. METHODOLOGY

Our system integrates YOLOv8 for object detection and ByteTrack for real-time object tracking. YOLOv8 detects objects in each video frame, providing bounding boxes and class labels, while ByteTrack associates these detections across frames, ensuring stable and accurate tracking.

To enhance tracking robustness, ByteTrack filters out false detections and maintains object identities even during occlusions. This makes it more reliable for fast-moving objects compared to traditional tracking methods.

For dataset preparation, we collected and annotated 10 football match videos, ensuring diverse player movements and ball trajectories. The model's performance is evaluated based on tracking accuracy, frame processing speed, and identity consistency. Our goal is to develop an efficient real-time tracking system for applications in sports analytics, automated player tracking, and performance evaluation.

IV. DATA SET

For this project, we've built a dataset using football match clips sourced from YouTube. These clips are typically 2 minutes long, and we selected a variety of matches to cover different scenarios. To create the dataset, we extracted frames from these clips at a rate of 5 frames per second , resulting in a total of 5,037 frames.To start the project process, we've been using Roboflow's auto-annotation tool. This allowed us to quickly label key objects, such as human and the ball in the frames. The auto-annotations provided bounding boxes and class labels, making it easier to get the ball rolling.

Right now, we're focusing on refining the dataset with manual annotations to improve its accuracy. manual adjustments ensure that the bounding boxes are more precise and the object labels are correctly assigned, especially in cases where the auto-annotations weren't perfect.

This dataset, once fully annotated, will be used to train and test our object detection and tracking models. The goal is to track moving objects across frames in real-time. By using clips from different matches, we're able to include a variety of scenarios, which will help build a more robust tracking system.

V. CONCLUSION

We are developing a real-time football match tracking system using YOLO for object detection and ByteTrack for tracking. Our dataset consists of 5,037 frames (5 fps) extracted from football videos, annotated using Roboflow to track players and the ball across diverse scenarios.

Currently, we are in the manual annotation phase, with the next steps focusing on training and evaluating our models based on accuracy and real-time performance.

For future improvements, we plan to explore additional models like Faster R-CNN, SSD, RetinaNet, EfficientDet, and tracking methods such as SORT, Deep SORT, and Kalman Filtering to enhance detection precision and tracking efficiency.

VI. REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 779-788, 2016, doi: 10.1109/CVPR.2016.91.

[2] Q. Wang and X. Huang, "ByteTrack: Multi-Object Tracking by Associating Every Detection Box," *arXiv preprint arXiv:2004.02795*, 2020. [Online]. Available: https://arxiv.org/abs/2004.02795.

[3] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, D. M. Dollar, and C. L. Zitnick, "Microsoft COCO: Common Objects in Context," *Proc. European Conf. Comput. Vis. (ECCV)*, 2014. [Online]. Available: https://arxiv.org/abs/1405.0312.