SPORTS VIDEOS PLAYER TRACKING MODEL REPORT

DS5216: ARTIFICIAL INTELLIGENCE

Presented By

M.S.M.SARIJ (DTS2434)

Introduction

THIS ASSIGNMENT FOCUSES ON BUILDING A COMPUTER VISION SYSTEM CAPABLE OF DETECTING AND TRACKING PLAYERS IN SHORT SPORTS VIDEOS. THE PROJECT USES A YOLO-BASED DETECTION FRAMEWORK TO IDENTIFY PLAYERS AND APPLIES AN ADDITIONAL KEYPOINT DETECTION MODEL TO ESTIMATE HUMAN POSE INFORMATION.

A DATASET OF FIVE SPORTS VIDEOS (RUGBY, BASKETBALL, CRICKET, VOLLEYBALL, FOOTBALL) WAS PROCESSED. EACH VIDEO MEETS THE ASSIGNMENT CRITERIA OF 5–10 SECONDS DURATION.

DATASET DIRECTORY:
https://github.com/mSarij/Player_Detection/tree/mSarij-patch-1/input%20videos

Methodology

1 Player Detection Model (YOLOv8n)

- **YOLOv8n pretrained model was used for fast and efficient player detection.**
- **Only the “person” class was detected.**
- **For each frame:**
  - **Bounding boxes were extracted**
  - **Detection confidence recorded**
  - **Players counted**

Github Repository link

2 Keypoint Detection Model (YOLOv8n-Pose)

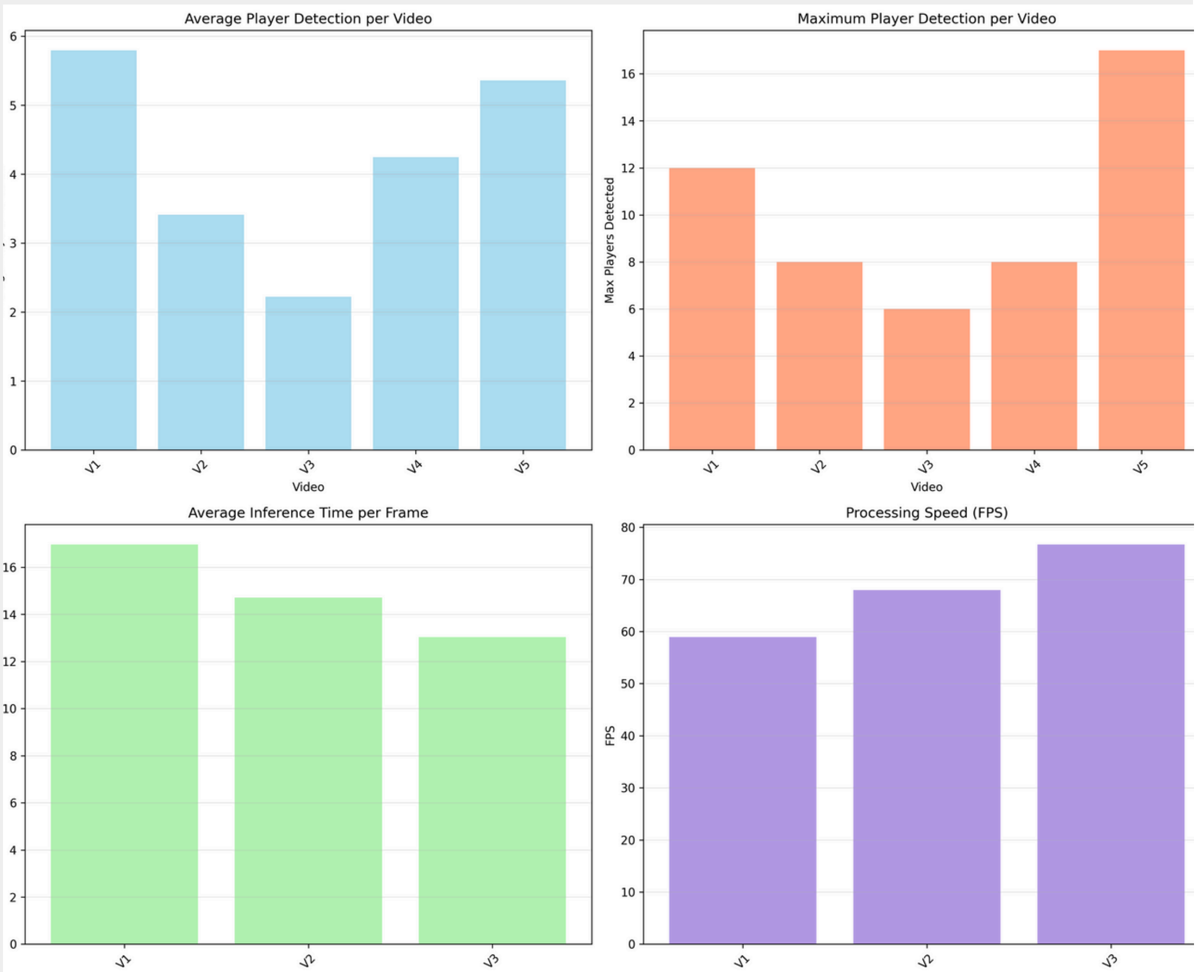
- The pose model detects 17 COCO keypoints per person.
- Skeletons were drawn on each frame.
- Average persons with keypoints were computed per video.

3 Inference & Performance Measurement

For each video:

- Average inference time per frame (ms)
- Processing speed (FPS)
- Average detections per frame
- Maximum detections
- Keypoint estimation performance

4 Performance Visualization



Performance Comparison of Models

PLAYER DETECTION PERFORMANCE (YOLO)

Model: YOLOv8n
Device: cuda

Detection Results:

Rugby.mp4:
- Average Players per Frame: 5.80
- Maximum Players Detected: 12
- Minimum Players Detected: 1
- Total Frames Processed: 225

Basketball.mp4:
- Average Players per Frame: 3.41
- Maximum Players Detected: 8
- Minimum Players Detected: 0
- Total Frames Processed: 240

Cricket.mp4:
- Average Players per Frame: 2.22
- Maximum Players Detected: 6
- Minimum Players Detected: 0
- Total Frames Processed: 125

Volleyball.mp4:
- Average Players per Frame: 4.24
- Maximum Players Detected: 8
- Minimum Players Detected: 1
- Total Frames Processed: 180

Football.mp4:
- Average Players per Frame: 5.36
- Maximum Players Detected: 17
- Minimum Players Detected: 0
- Total Frames Processed: 225

PERFORMANCE METRICS

Rugby.mp4:
- Average Inference Time: 16.96 ms
- Processing FPS: 58.96
- Average Confidence: 0.581
- Average Detections: 10.42

Basketball.mp4:
- Average Inference Time: 14.71 ms
- Processing FPS: 67.97
- Average Confidence: 0.506
- Average Detections: 7.50

Cricket.mp4:
- Average Inference Time: 13.04 ms
- Processing FPS: 76.69
- Average Confidence: 0.480
- Average Detections: 5.96

KEYPOINT DETECTION

Model: YOLOv8n-pose

Results:

Rugby.mp4:
- Average Persons with Keypoints: 6.78
- Total Frames Processed: 225

Basketball.mp4:
- Average Persons with Keypoints: 3.18
- Total Frames Processed: 240

Cricket.mp4:
- Average Persons with Keypoints: 3.12
- Total Frames Processed: 125

Volleyball.mp4:
- Average Persons with Keypoints: 3.64
- Total Frames Processed: 180

Football.mp4:
- Average Persons with Keypoints: 5.07
- Total Frames Processed: 225

Discussion

1 Model Performance

- YOLOv8n provides real-time performance with speeds between 59–76 FPS.
- Detection accuracy depends heavily on:
  - Video resolution
  - Lighting
  - Camera angle
  - Player distance
- Keypoint detection works best when players occupy larger pixel areas.

2 Limitations

- Missed detections in low-resolution or wide-field videos.
- Small, distant players produce fewer keypoints.
- Some false positives occur (referees, bystanders).
- No temporal tracking—each frame processed independently.
- Basketball poses challenges due to speed and occlusion.

“This project used the pretrained YOLOv8n and YOLOv8n-pose models without additional training. Therefore, loss curves such as box_loss, cls_loss, or keypoint_loss are not generated. If fine-tuning were performed, the loss curves would show convergence behaviour for model optimization.”

3 Possible Improvements

To enhance performance:

Detection

- Fine-tune YOLO on sports-specific training data
- Use YOLOv8m or YOLOv8l for higher accuracy

Tracking

- Integrate DeepSORT or ByteTrack
- Implement player re-identification (ReID)

Pose Estimation

- Use HRNet or OpenPose for higher accuracy

Extended Analytics

- Player speed estimation
- Heatmap generation
- Team movement analysis

Screenshots of Outputs

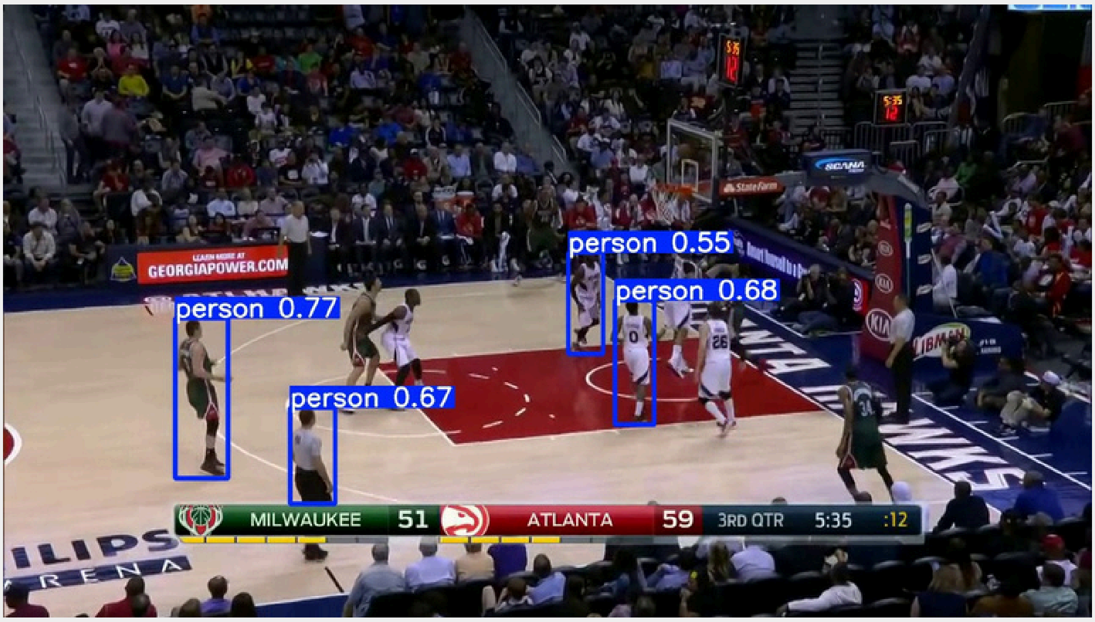
Player Detection

Detected Basketball players

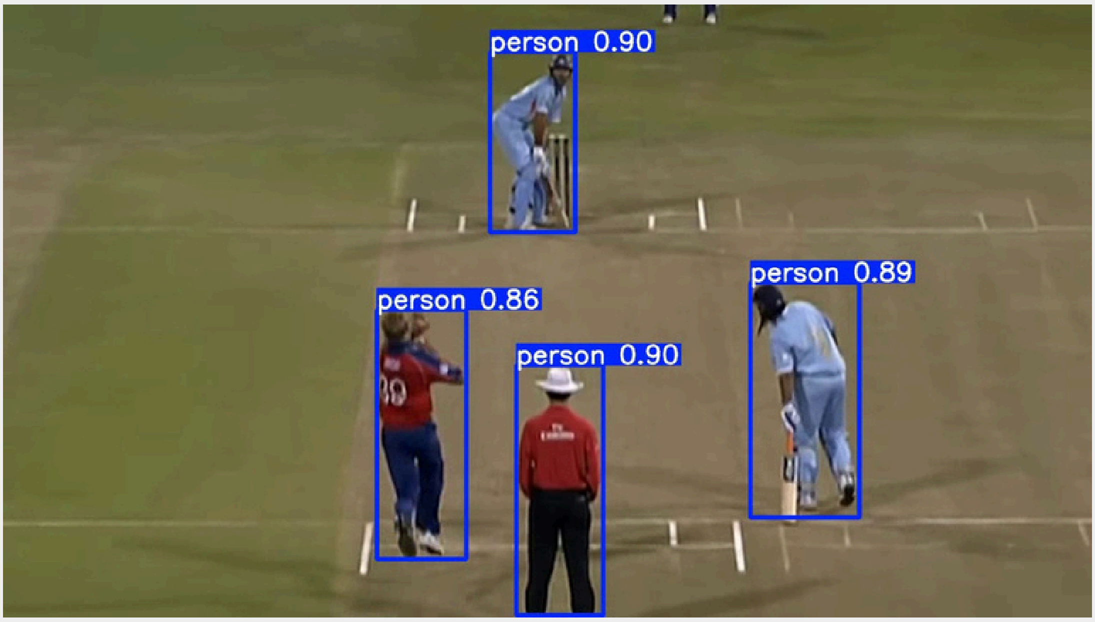
Detected Basketball players

Screenshots of Outputs

Player Detection



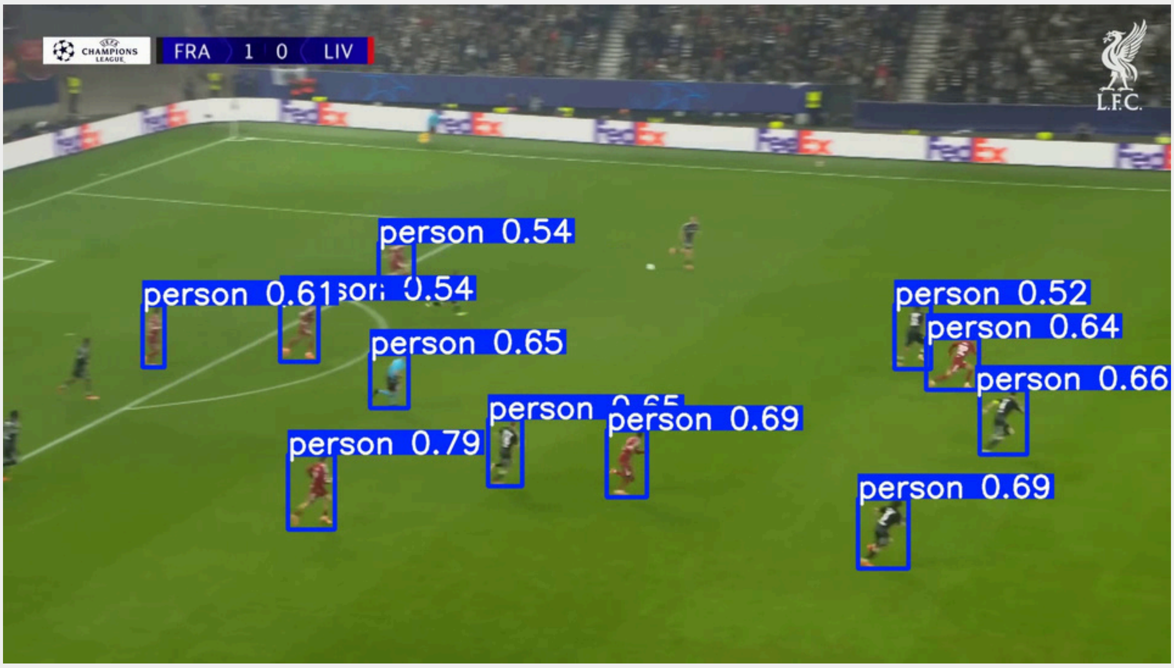
Detected Basketball players



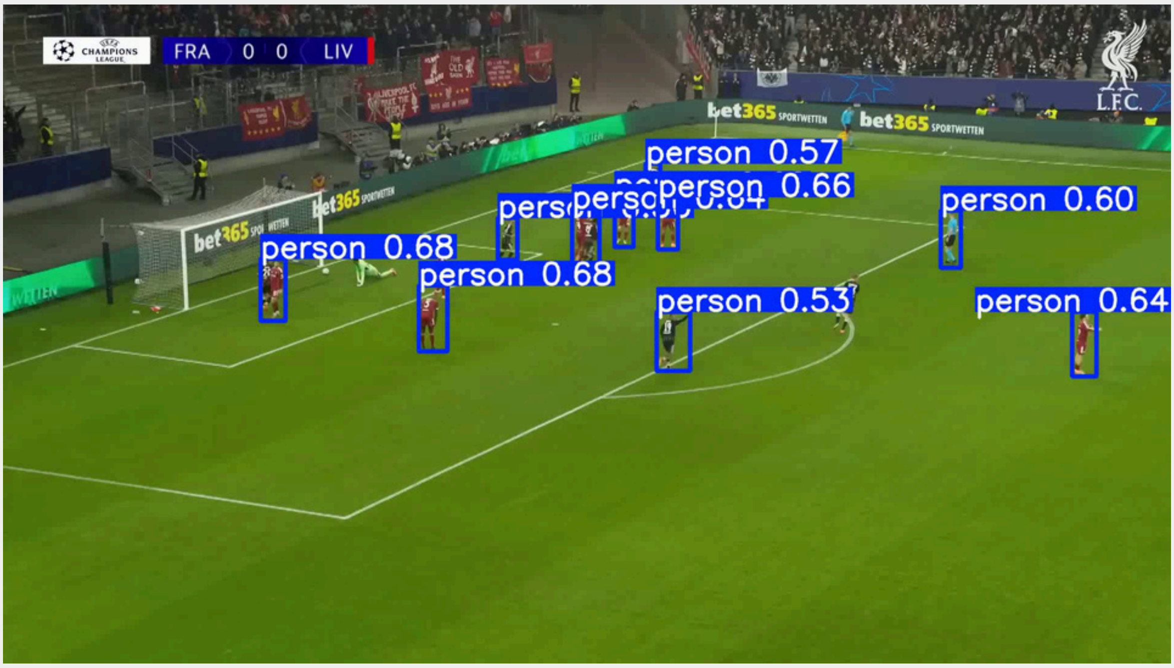
Detected Cricket players

Screenshots of Outputs

Player Detection



Detected Football players

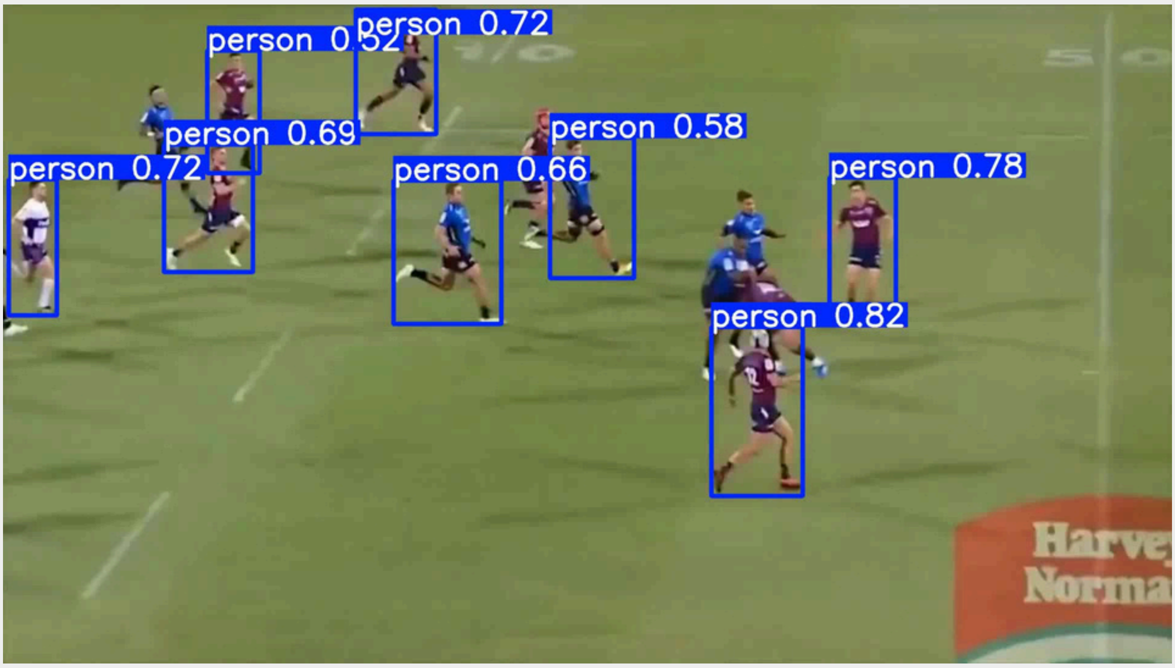


Detected Football players

Screenshots of Outputs

Player Detection

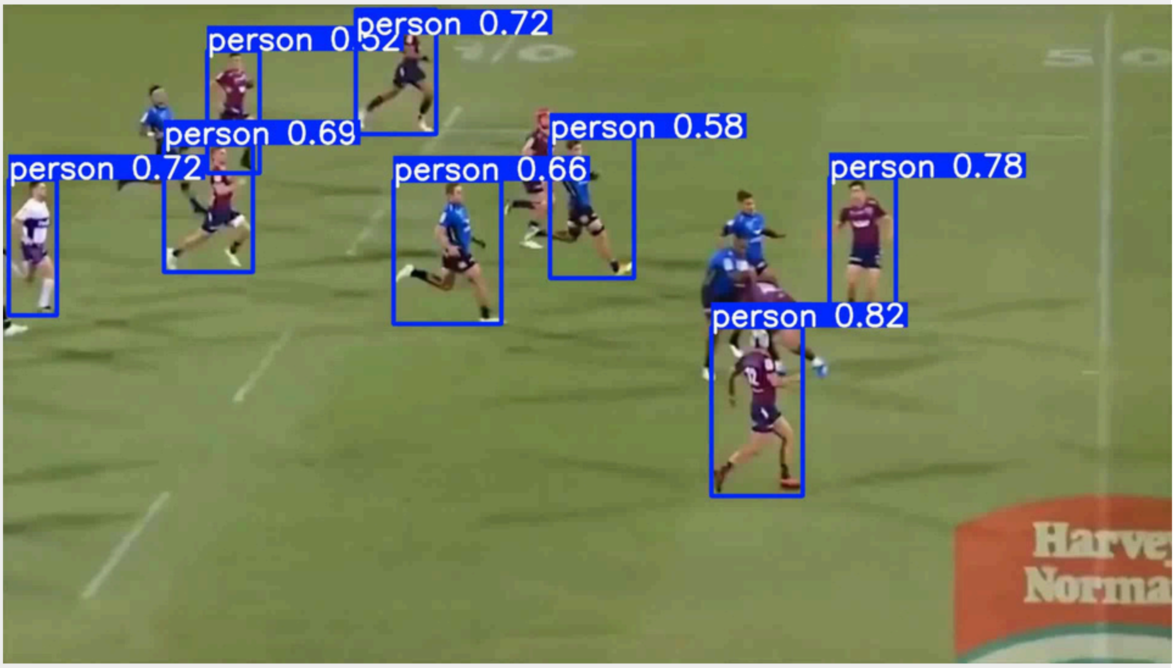
Detected Football players



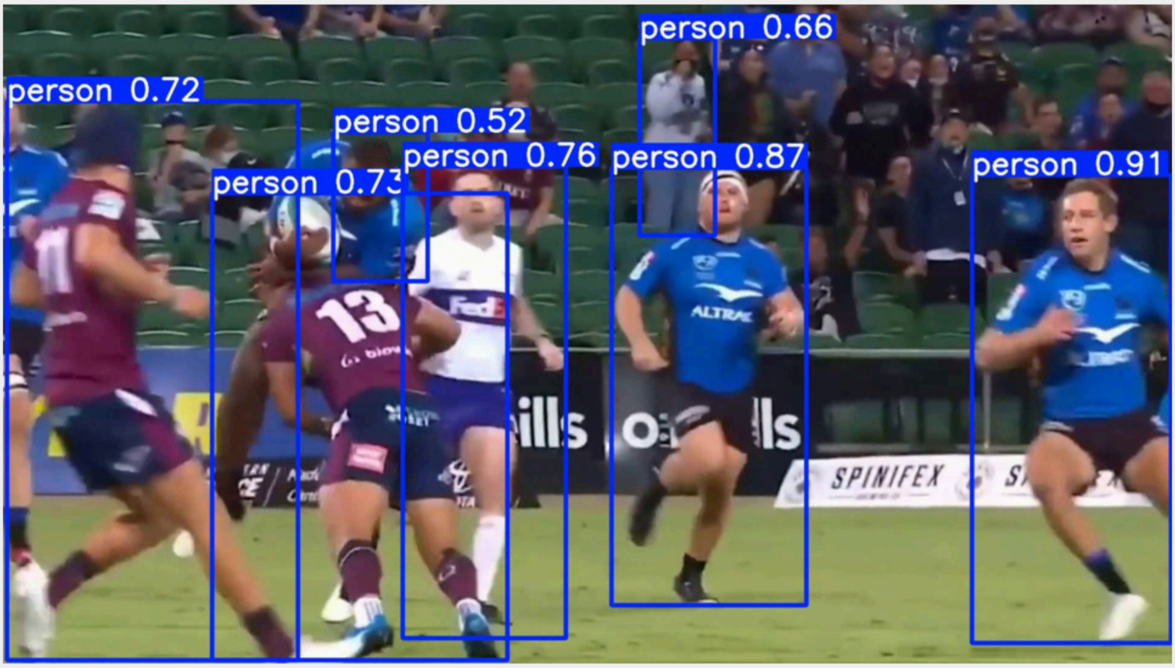
Detected Rugby players

Screenshots of Outputs

Player Detection



Detected Rugby players



Detected Rugby players

Screenshots of Outputs

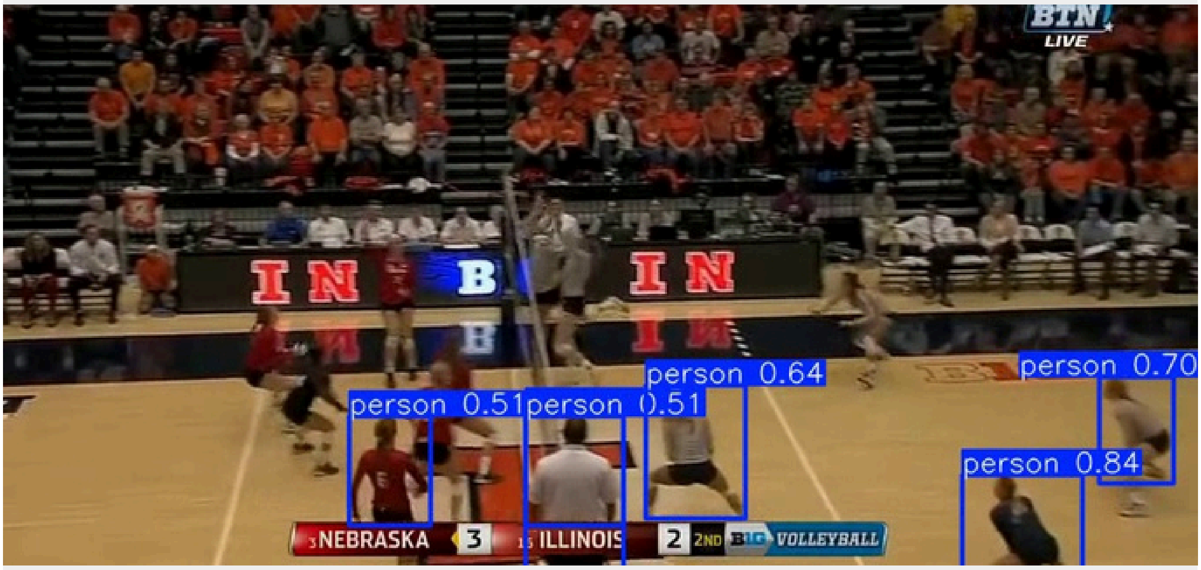
Player Detection

Detected Rugby players

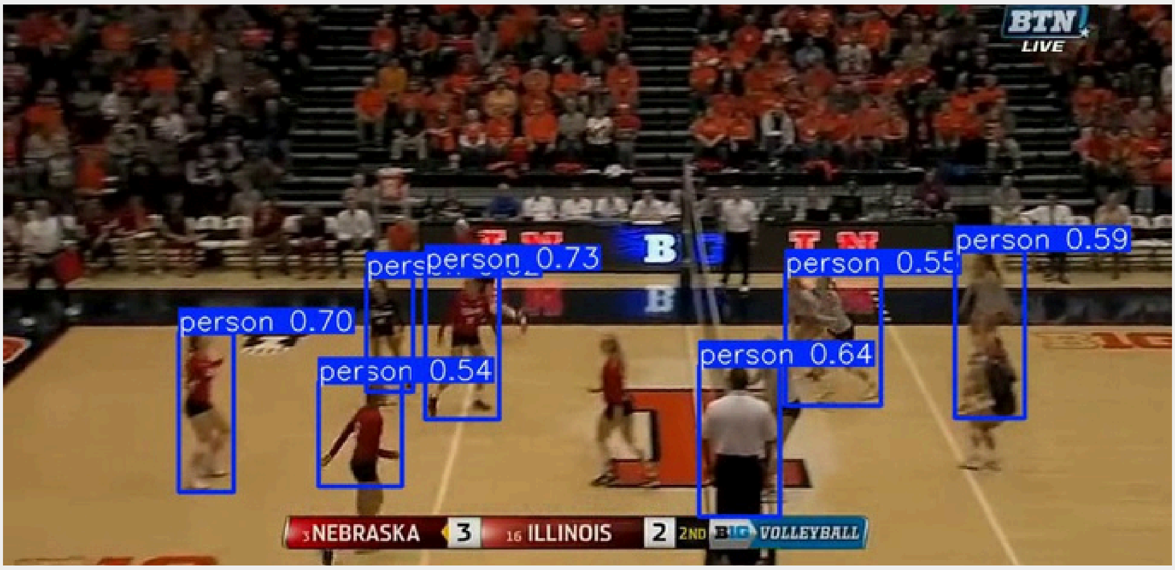
Detected Volleyball players

Screenshots of Outputs

Player Detection



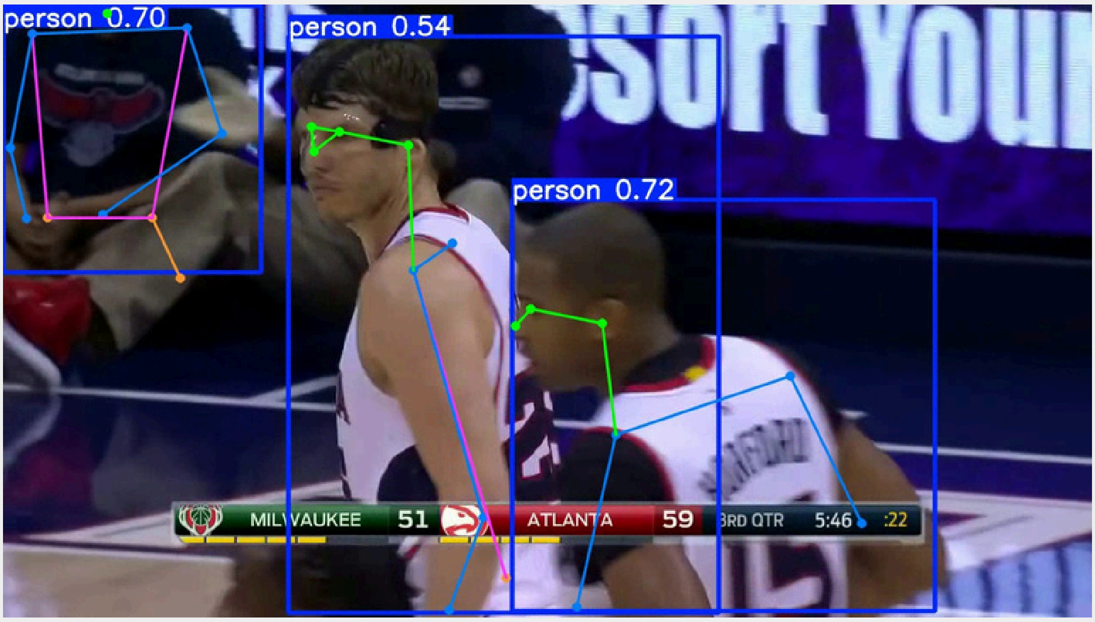
Detected Volleyball players



Detected Volleyball players

Screenshots of Outputs

Key point Detection

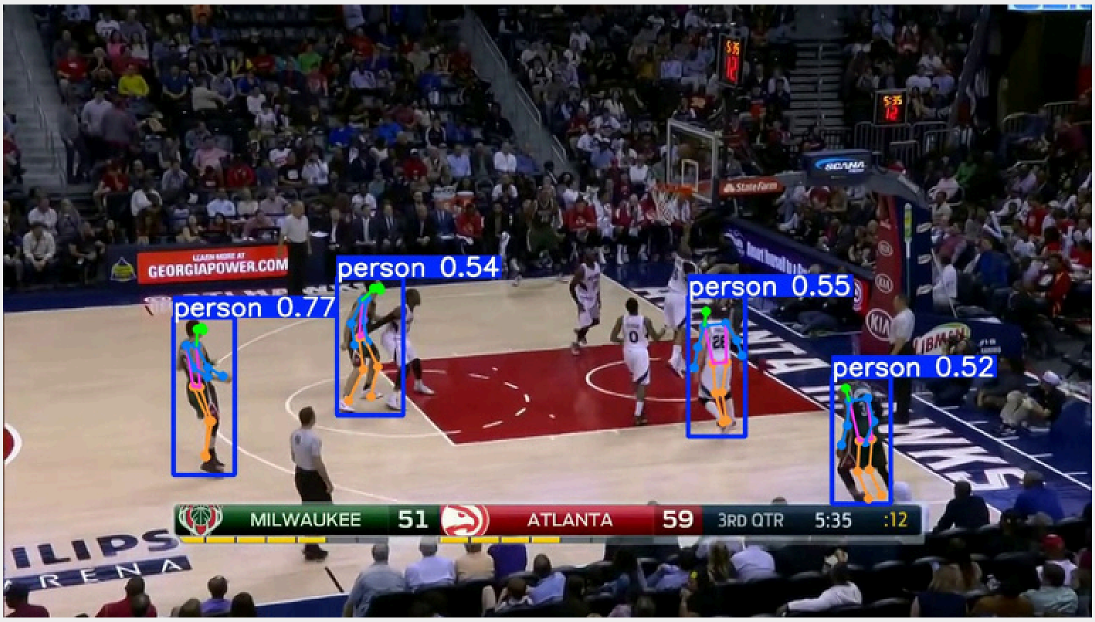


Key point Detection from detected Basketball players

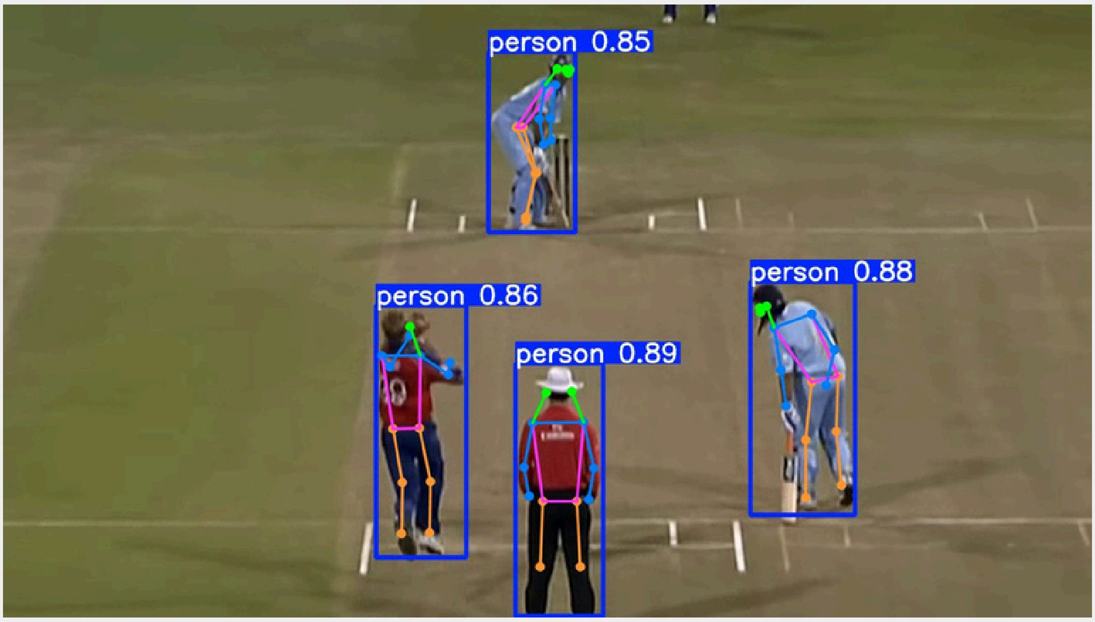
Key point Detection from detected Basketball players

Screenshots of Outputs

Key point Detection



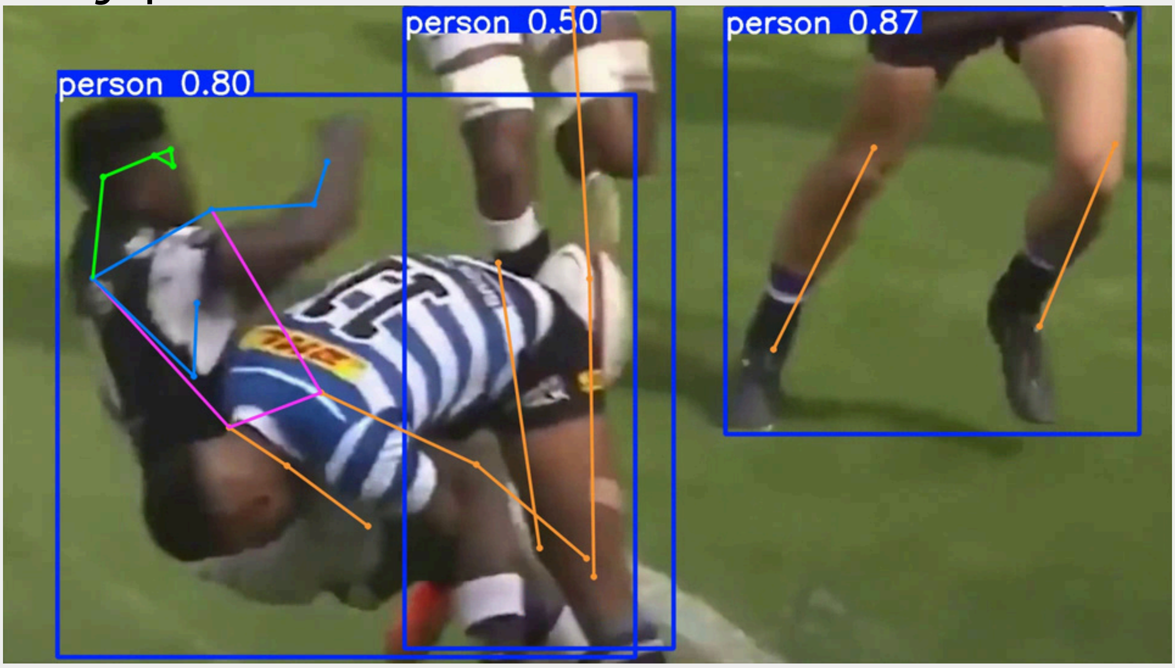
Key point Detection from detected Basketball players



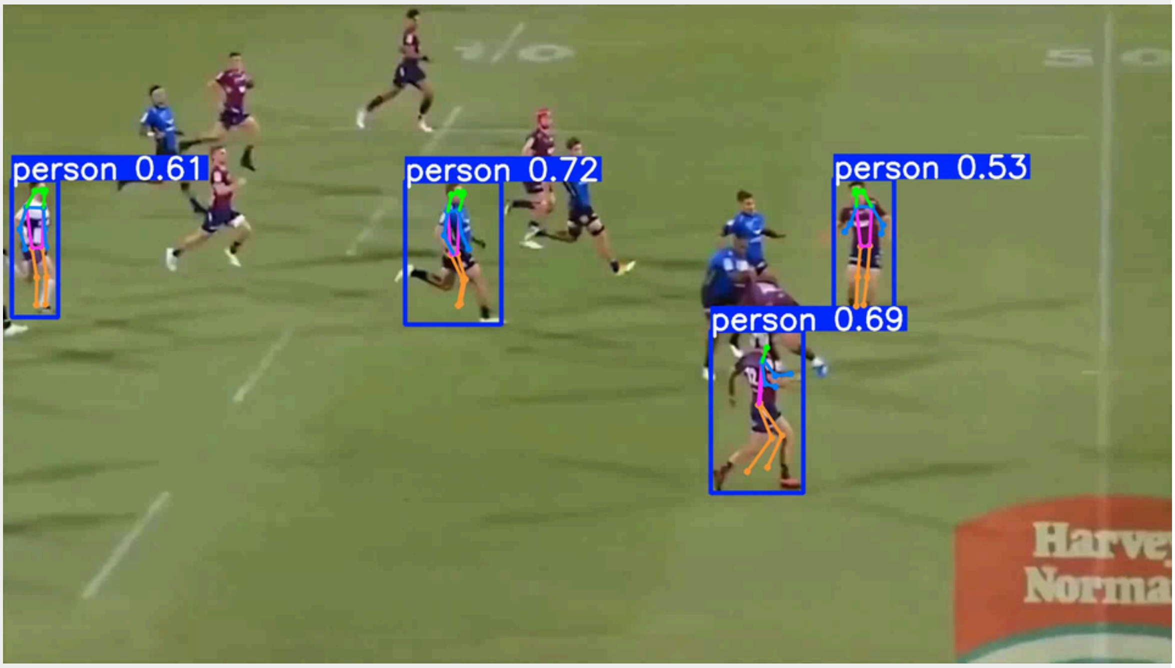
Key point Detection from detected Cricket players

Screenshots of Outputs

Key point Detection



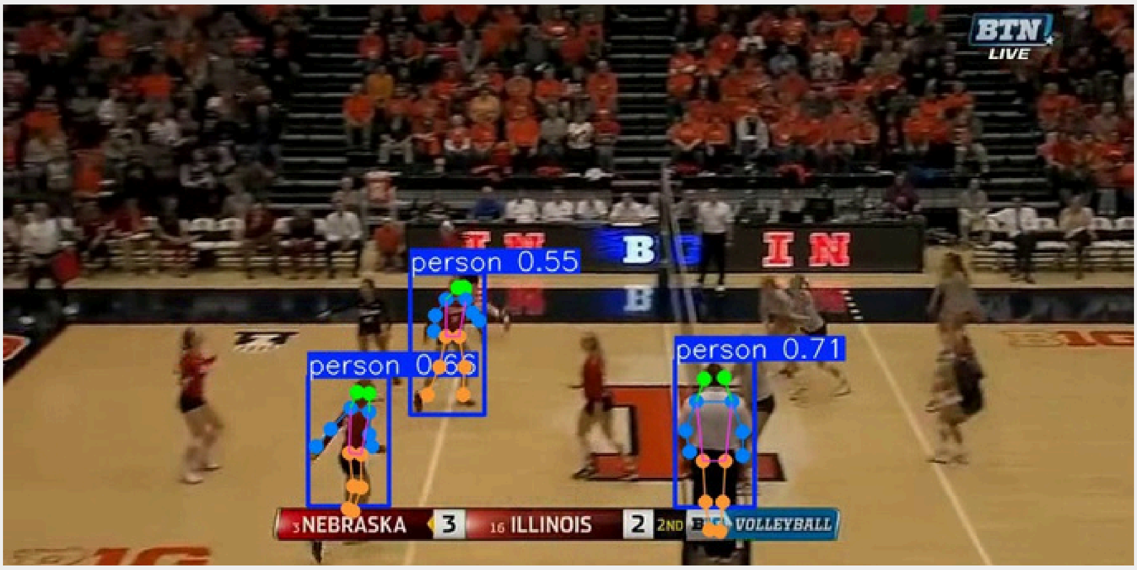
Key point Detection from detected Rugby players



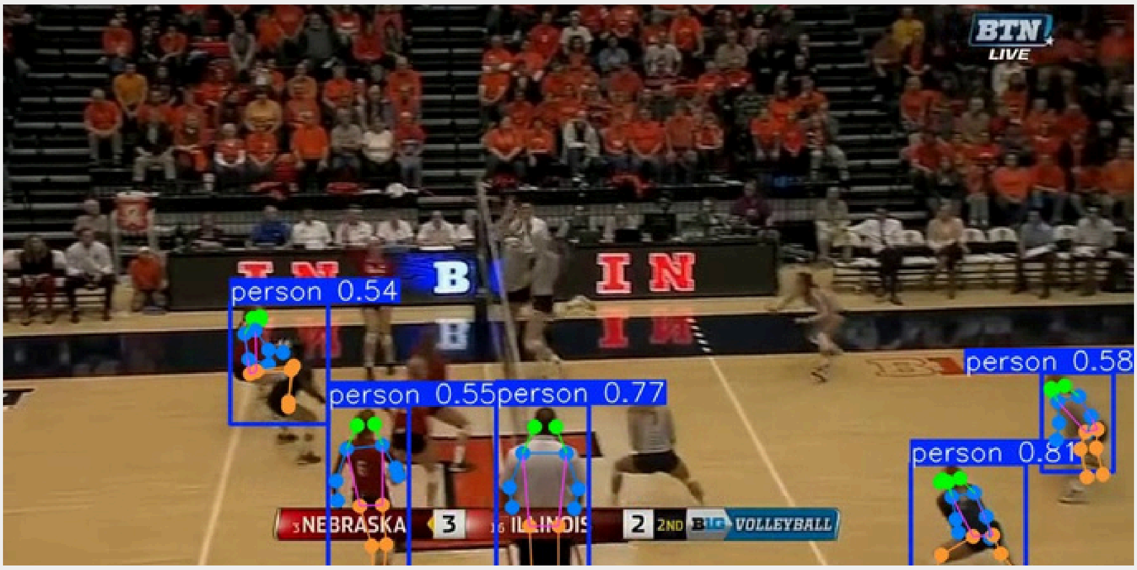
Key point Detection from detected Rugby players

Screenshots of Outputs

Key point Detection



Key point Detection from detected Volleyball players



Key point Detection from detected Volleyball players

Conclusion

This model successfully demonstrates the implementation of a complete YOLO-based sports player detection and analysis pipeline. Player detection, pose estimation, and performance benchmarking were performed across five different sports videos. The results show strong real-time performance and highlight important considerations when applying computer vision to sports analytics.
The system provides a solid baseline for future work, including tracking, classification, and advanced player analytics.

- YOLOv8n performs robust player detection across all sports.
- YOLOv8n-Pose works well when players occupy sufficient pixel area.
- Both models operate in real-time, making them suitable for live sports analysis.
- Performance drops mainly due to:
    - Distant camera views
    - Occlusions
    - Motion blur
    - Lower resolution sports videos