Massachusetts Institute of Technology

Work on Efficient Computer Vision Methods in FIFA 23

6.506 Final Presentation

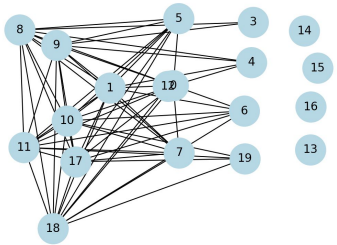
Ian Gatlin
Work on Efficient Computer Vision Methods in FIFA 23

May 14th, 2024
Demo Video: https://github.com/ijga/FIFACoachBot/blob/main/demo.mov
Github Repo: https://github.com/ijga/FIFACoachBot

Overview of Project

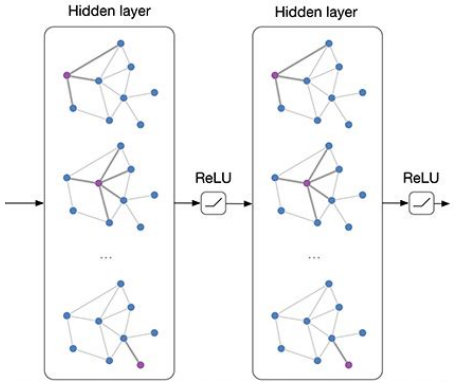
Goal

- Quickly create high quality graph representations of soccer matches
- Create a ML coaching model that gives instant coaching based on the state of the game



Challenges

- Finding small objects, specifically the ball
- Personal laptops have slower hardware compared to the state of the art systems



Three Interesting Problems

Training a Quality CV Model

- Used bounding box object detection
- 100 images with little variability

Applying CV Model to Footage

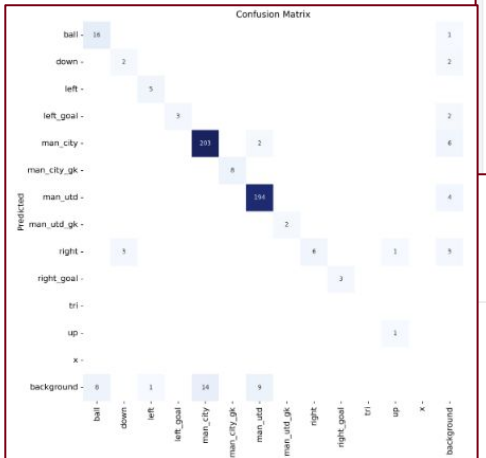
- Tiling frames and processing them sequentially and in parallel
- No downscaling of frames

Consistent Representations with Different Camera Angles

- A lot of image processing to determine angle from fixed camera
- Discuss other work

CV Model Generation with Yolov8 and Roboflow

- **Labeled 100 images**
  - Attempted to encode controller input
- **Tiled Images 2 rows x 3 columns** [1]
- **1440 Total Images**
- **200 Epochs**



Images
100
0 missing annotations
0 null examples

Annotations
1,966
19.7 per image (average)
Across 13 classes

Average Image Size
2.07 mp
from 2.07 mp
to 2.07 mp

Median Image Ratio
1920x1080
wide

Class Balance

Manage Classes | Rebalance Splits

all | train | valid | test

Train | Valid | Test

man_city	859	
man_utd	831	
ball	94	under represented
right	34	under represented
man_city_gk	33	under represented
left	27	under represented
right_goal	25	under represented
left_goal	16	under represented
man_utd_gk	16	under represented

Preprocessing

Auto-Orient: Applied

Tile: 2 rows x 3 columns

Augmentations

Outputs per training example: 3

Blur: Up to 0px

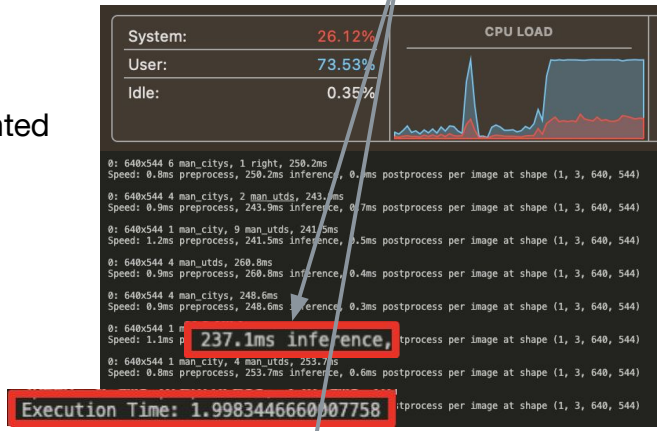
Noise: Up to 0.49% of pixels

Applying CV Model to Footage - Tiling [2]

Serial

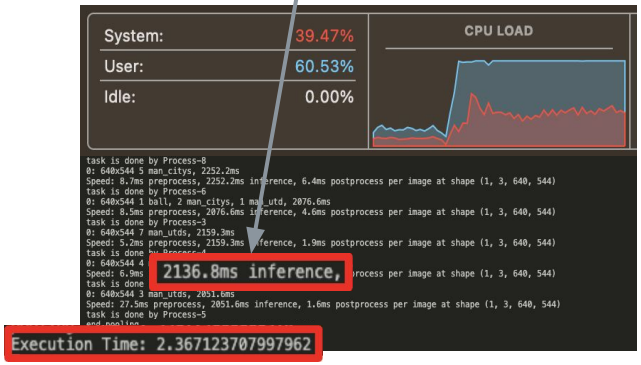
- Library implemented in parallel?

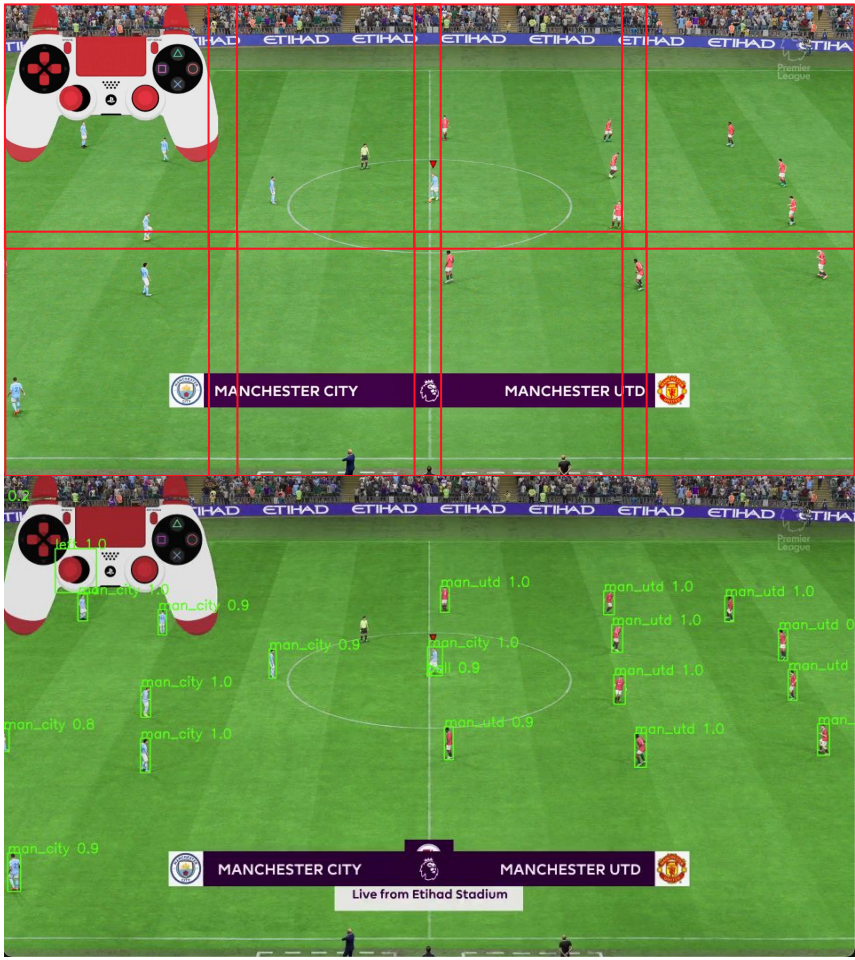
Wall-clock time to complete object detection on a tile



Parallel [3]

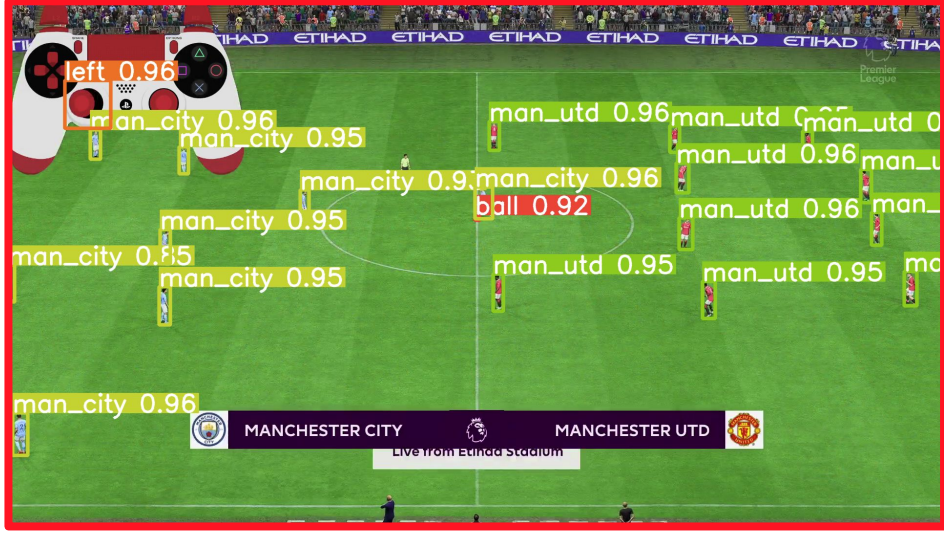
- Memory sharing problems (1MB)
- Process management overhead





Applying CV Model to Footage
- No downscaling

- Manually set image size to not downscale (1080p) when running object detection
- Honestly works the best
- Kinda boring though




```
video 1/1 (frame 1/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 11 man_utds, 1688.9ms
video 1/1 (frame 2/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 9 man_citys, 11 man_utds, 1579.6ms
video 1/1 (frame 3/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 9 man_citys, 11 man_utds, 1418.9ms
video 1/1 (frame 4/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 9 man_citys, 11 man_utds, 1355.9ms
video 1/1 (frame 5/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1415.7ms
video 1/1 (frame 6/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1384.3ms
video 1/1 (frame 7/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1359.9ms
video 1/1 (frame 8/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1368.2ms
video 1/1 (frame 9/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1409.3ms
video 1/1 (frame 10/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1359.9ms
video 1/1 (frame 11/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1355.3ms
video 1/1 (frame 12/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1422.0ms
video 1/1 (frame 13/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1434.0ms
video 1/1 (frame 14/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_utds, 1413.0ms
video 1/1 (frame 15/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_ut
video 1/1 (frame 16/52938) /Users/iangatlin/Documents/FIFACoach/506/videos/coop/game1.mov: 1088x1920 1 ball, 1 left, 8 man_citys, 10 man_ut 1364.5ms
```

Consistent Representations with Different Camera Angles

Homography transform to project players into a non-warped frame

- Improves consistency and quality

Time is dominated by object detection

- 60ms for dynamic homography
- 200-1500ms for object detection

Other Steps

- Graph creation
- Duplicate bounding box deletion

Real world applications

- One camera needed

Conclusions

- Very difficult to get graph representations in an quick enough time for instant coaching feedback on personal laptops
- Need to explore different computer vision models besides bounding box object detection with Yolov8
    - Yolov8 is very easy to work with
    - Image segmentation
    - Label more data to create better model
- Python multiprocessing is missing documentation for shared memory and has its limitations
    - Python is used a lot for computer vision
- Coach for end of game feedback
    - Still need performance since most frames will have no coaching insights

Thank you!

References:

[1] https://blog.roboflow.com/detect-small-objects/#how-to-effectively-detect-small-objects

[2] https://github.com/niconielsen32/tiling-window-detection/blob/main/tiling.py

[3] https://www.digitalocean.com/community/tutorials/python-multiprocessing-example