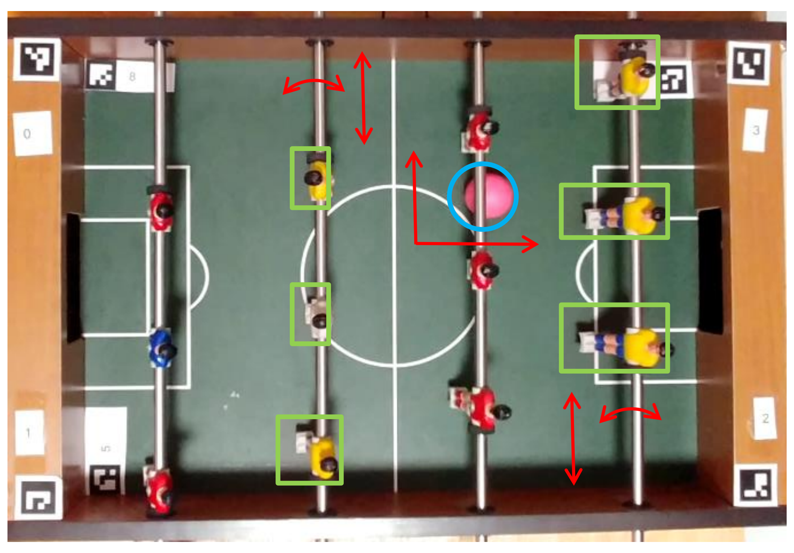
Foosball robot object detection

Joseph Lundy
Department of Electrical and Computer Engineering, University of Toronto

Motivation

For a foosball robot, need to track the following in real-time:
- ball position
- foosman position and rotation

Challenges
- Foosball frequently occluded by foosmen and rods
- Placing sensors to measure rod angle is costly and results in undesired changes to the foosball table

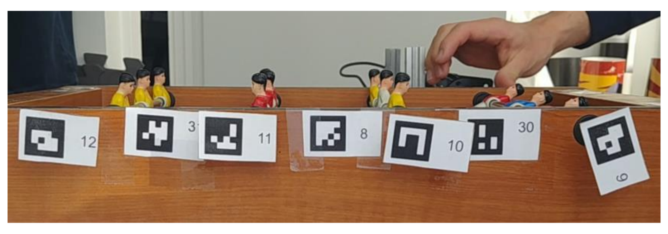


Idea
- Use overhead camera to detect and track the ball and foosmen

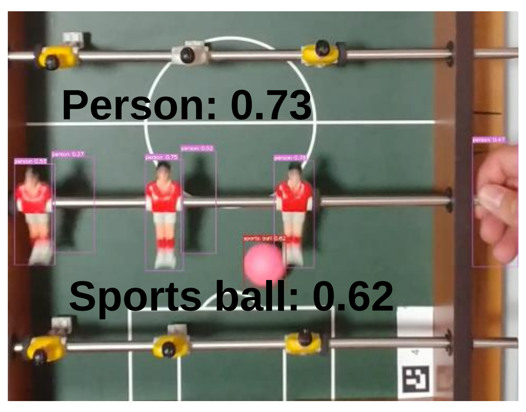
Previous approach [1]
- Used color based tracking for ball tracking
- Imprecise, sensitive to different lighting conditions

Related Work

Highly reliable fiducial markers [2]
- Used aruco markers for ground truth angle estimation
- Want to avoid modifying table with markers on the foosmen



Yolov7 [3]
- Very fast object detector
- Not trained for this application but suitable candidate for transfer learning



Segment Anything [4]
- Very powerful segmentation tool
- Can't be used real-time
- Used SAM2 to compute bounding boxes for training data
- Output very sensitive to initialization

Legacy tracking algorithms

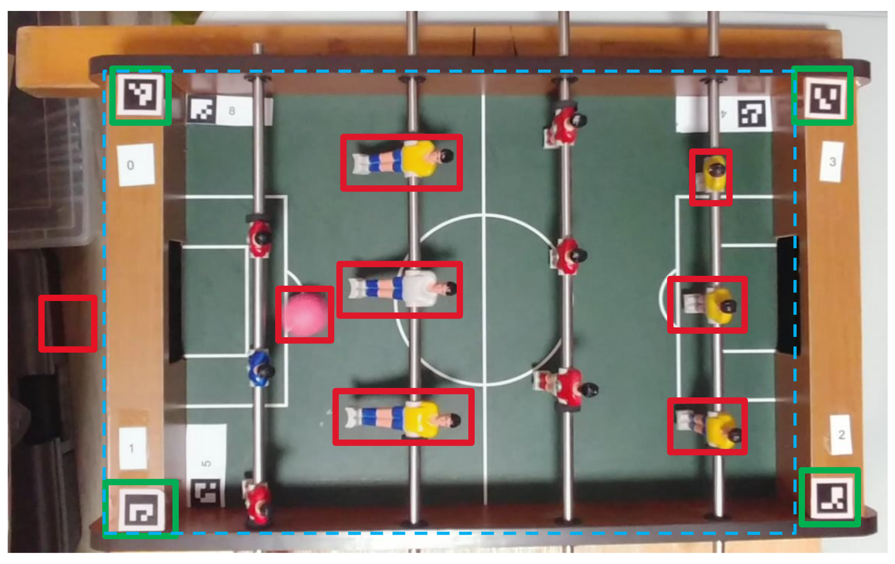
Name	Performance	FPS
CSRT	Very good - Tracks ball until very last frames	44
Boosting	Mediocre – loses ball after occlusion; tracks foosman poorly	62
KCF	Very poor - Loses ball and foosman very quickly	69
MIL	Mediocre – loses ball immediately; can track foosman	26
TLD	Good – can track ball throughout, very noisy detection; performs worse on foosman	29
Median Flow	Poor – loses ball immediately; can track foosman but detection box grows	80
MOSSE	Very poor – loses ball immediately; loses foosman after rotation	80

References

[1] Joseph Lundy, "Foosball Robot 2.0," Joseph Lundy, Apr. 21, 2022. https://joelundy.wordpress.com/foosball-robot-2-0/
[2] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," Pattern Recognition, vol. 47, no. 6, pp. 2280–2292, 2014
[3 N. Ravi et al., "SAM 2: Segment Anything in Images and Videos," arXiv preprint arXiv:2408.00714, 2024
[4] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. 2022

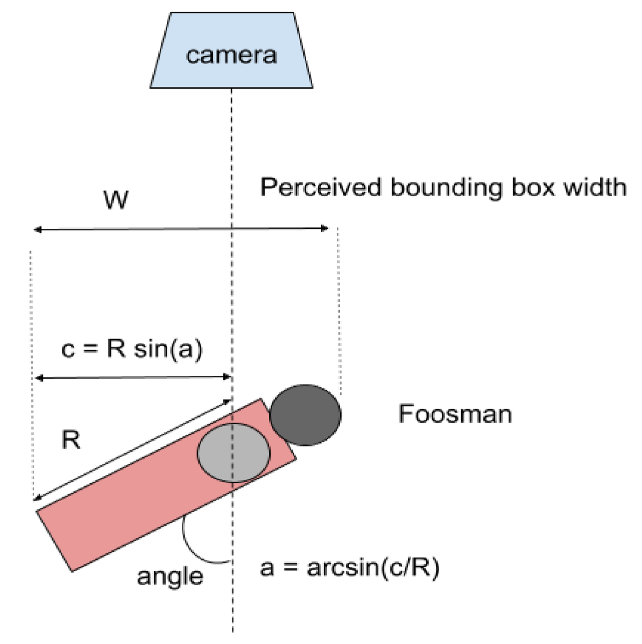
New Technique

General Approach
- Train YOLOv7 CNN detector to detect ball and foosmen
- Use linear regression on bounding boxes to compute foosman / rod rotation angle
- Use aruco markers for ground truth angle estimation data



Proposed algorithm for tracking
1. Detect aruco markers to narrow down area of interest
2. Run retrained yolov7-tiny model to get bounding boxes
3. Ignore reported objects outside area of interest
4. Use Kalman filter and Hungarian algorithm for object tracking
5. Estimate measured foosman angle from bounding box
6. Apply Kalman filter on angle estimations

Proposed approach for angle estimation
1. Train Yolov7 CNN model to detect not just foosman location but classify it based on its rotation quadrant
2. Use arcsine method on bounding box
3. Compute 'full' rotation angle with knowledge of quadrant



Experimental Results

Retrained Yolov7 Detection

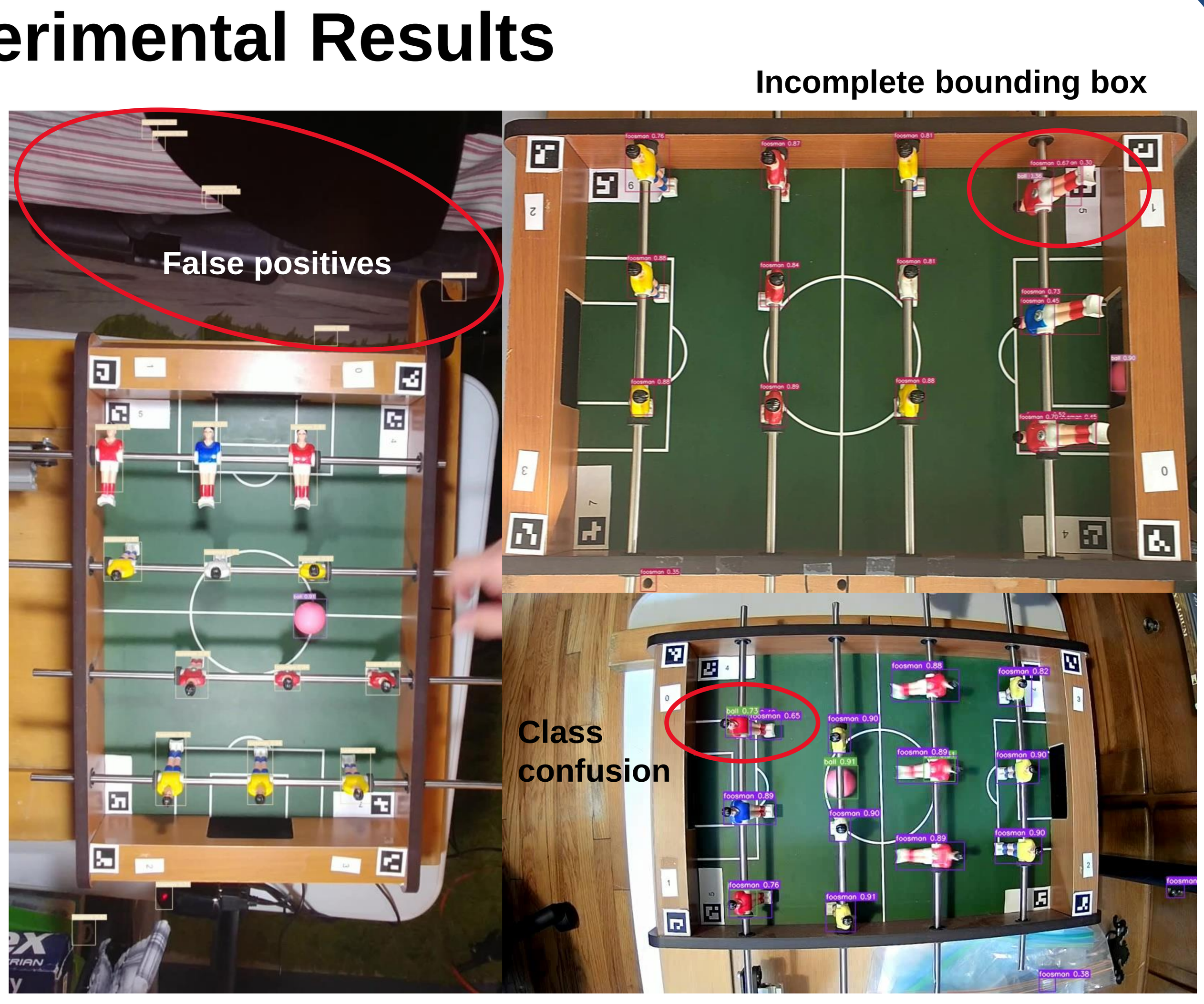
	Train	Test	Val
Precision	0.996	0.999	0.988
Recall	0.991	0.989	0.968

Detection speed

	Inference	NMS	Total FPS
Typical vid.	3 ms	0.7 ms	270
Worst vid.	3.3 ms	2.5 ms	172

Camera view, scene generalization

Desc.	All	Ball	Foos
View rotated 90	P 0.976 R 0.929	P 0.979 R 1	P 0.973 R 0.858
Webcam – warped	P 0.827 R 0.938	P 0.655 R 0.957	P 1 R 0.92

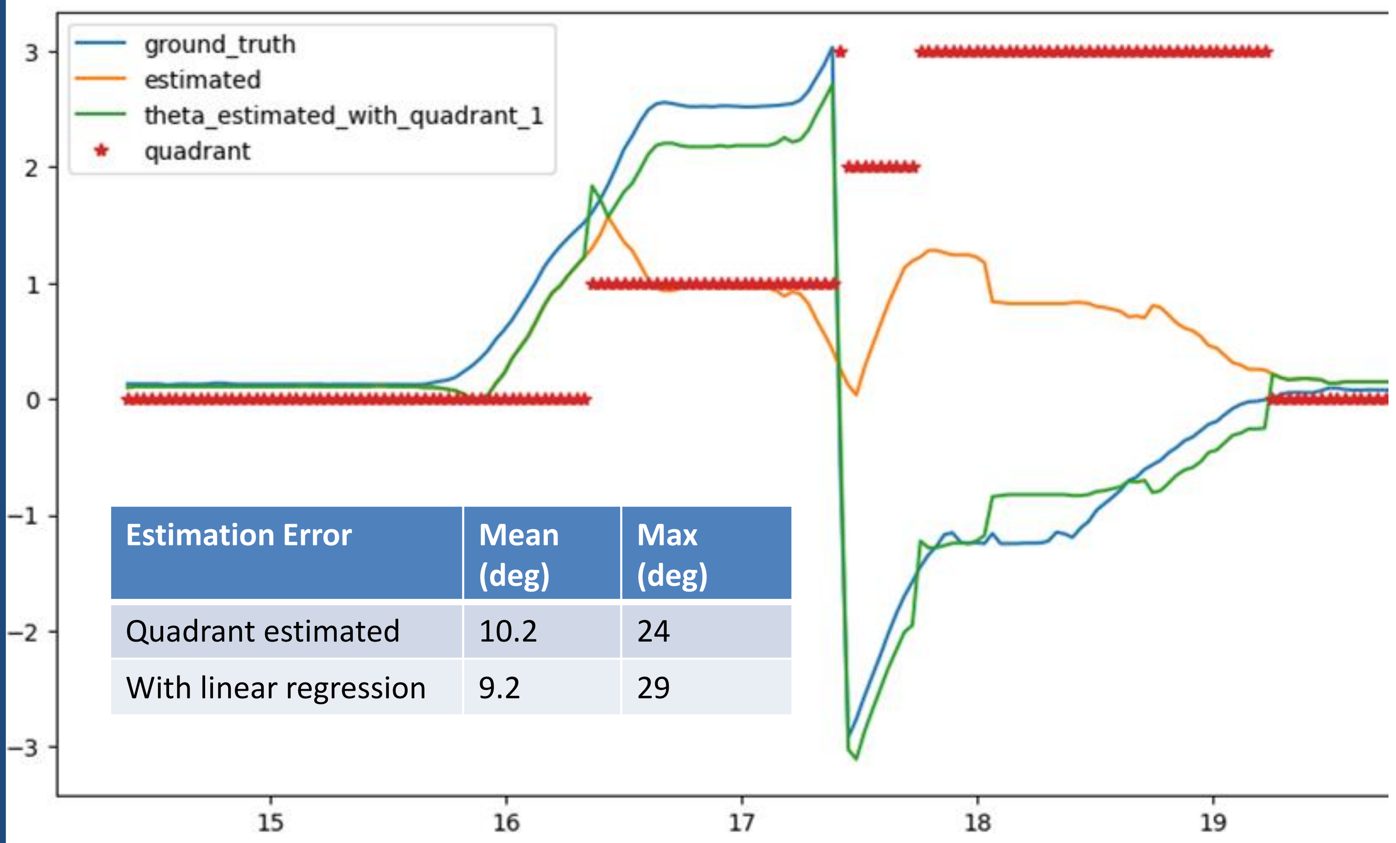


Ball tracking generalization

Color	Performance	Ball Prec.	Ball Recall
Pink	Excellent detection and classification; Handles motion blur and occlusions very well	0.991	1
Orange	Robust detection but moderate confidence scores (0.46 – 0.7)	1	1
L. Orange	Fair detection; ball found ~70% of time; low confidence (0.2 – 0.4); frequent class confusion	0.666	0.4
Green	Poor detection; ball found ~ 50%; bounding box is less accurate; class almost always wrong	0.998	0.5
Blue	Robust detection; frequent class confusion; sometimes misses when close to foosman	0.778	0.7
White	Robust detection; however ball always confused as foosman	0	0
All colors	(all 60 images)	0.692	0.6

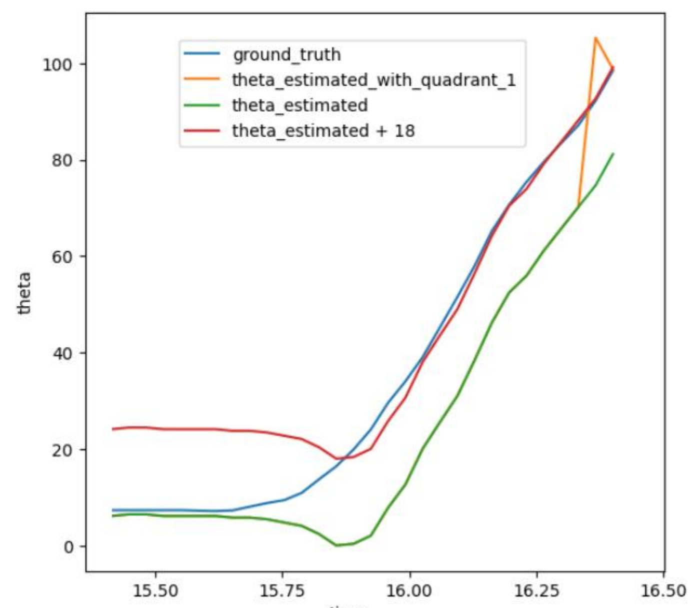
Angle Estimation Study

Full 360 estimation if quadrant is known

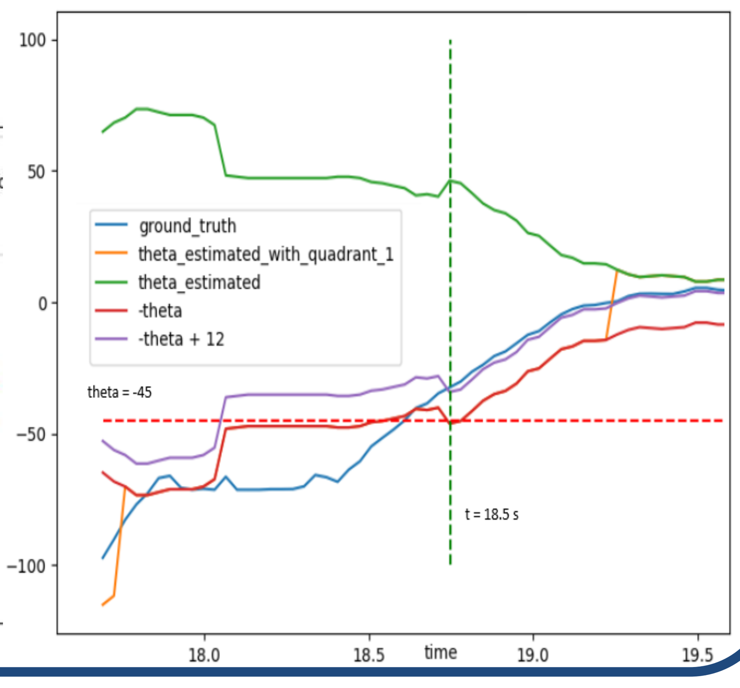


Estimation Error	Mean (deg)	Max (deg)
Quadrant estimated	10.2	24
With linear regression	9.2	29

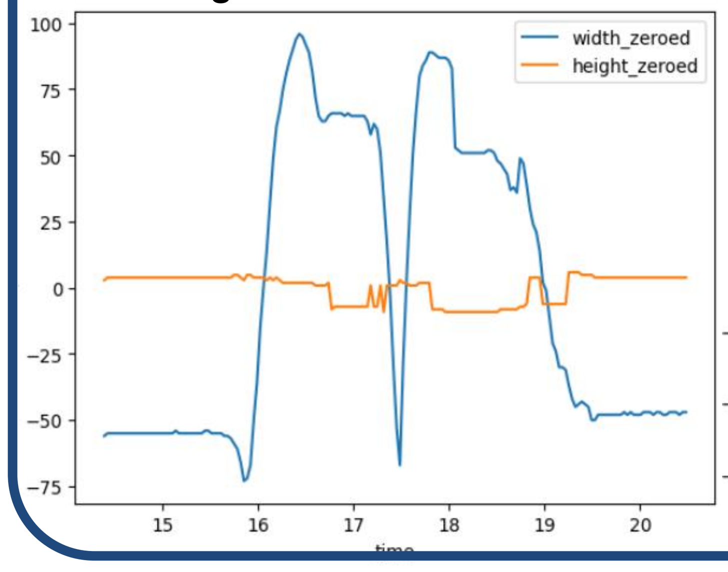
Alignment error and non-linearities: Positive angles



Alignment error and non-linearities: Negative angles



Bounding box width trend



Bounding box center trend