

# Gate Detection for Autonomous MAV Race

## Overview

This thesis investigates the development of a vision based navigation system to be applied in autonomous drone racing, in particular the race at IROS 2018 in Madrid. The race court consists of several gates that need to be passed one after another. The image is recorded from a monocular camera located at the front of the drone. An example for the race court can be seen on the right.

Goal of this research is to detect one or multiple gates in an image and potentially do a pose estimation. The output can be used by the drone's navigation system to efficiently fly through the race court. The preliminary research questions are summarized in **Research Questions** the approach is shown in **Approach**.

## Approach

1. Generate data to evaluate different models
2. Evaluate different common object detectors on generated data
3. Record 'real' data for evaluation
4. Customize and implement model on the MAV, optimize with respect to efficiency and detection performance

## Data Generation

At this point no labeled data is available. Training samples are generated by a graphic engine that loads the 3D model of the gate and replaces the background with a random image. In later stages training data will be collected from a manufactured race court.

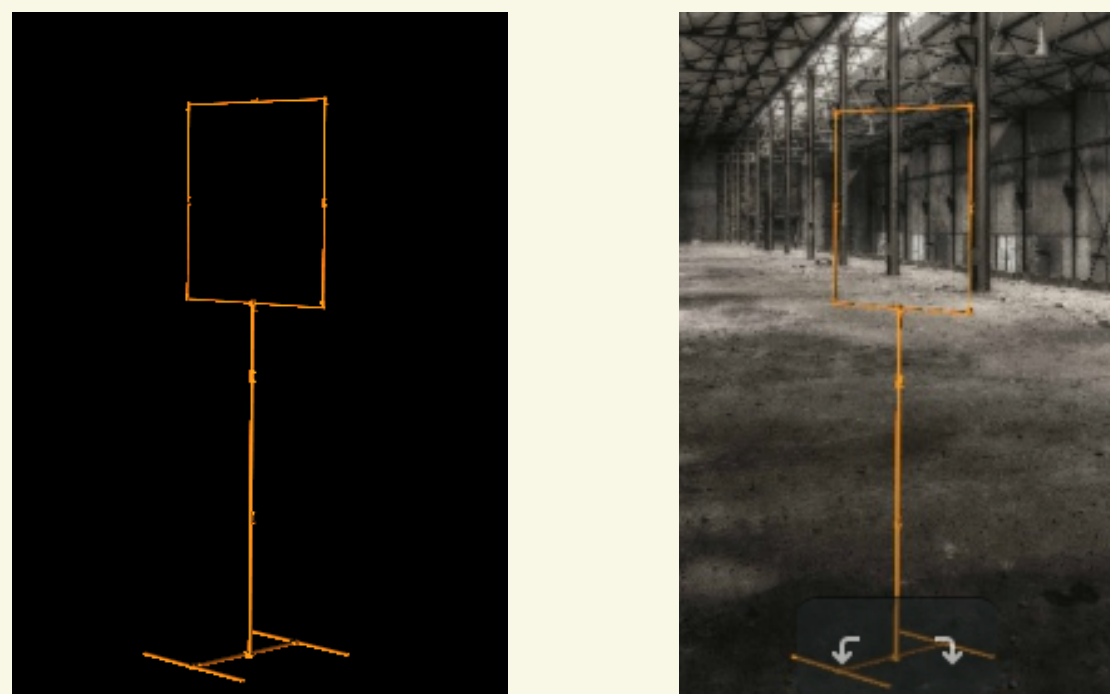


Figure: Example for Generated Sample

## Research Questions

1. How can we detect objects that consist mainly of background, namely the gates in an autonomous drone race?
2. How can we deal with the limited amount of available training data?
3. How can we deal with the limited computational resources on the MAV?



Figure: Race Court at IROS2011

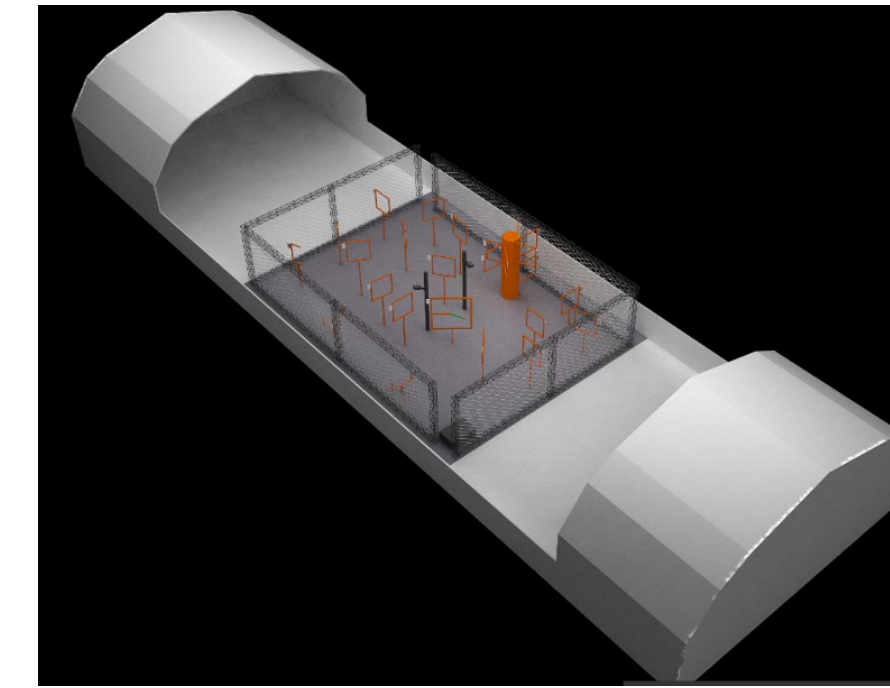
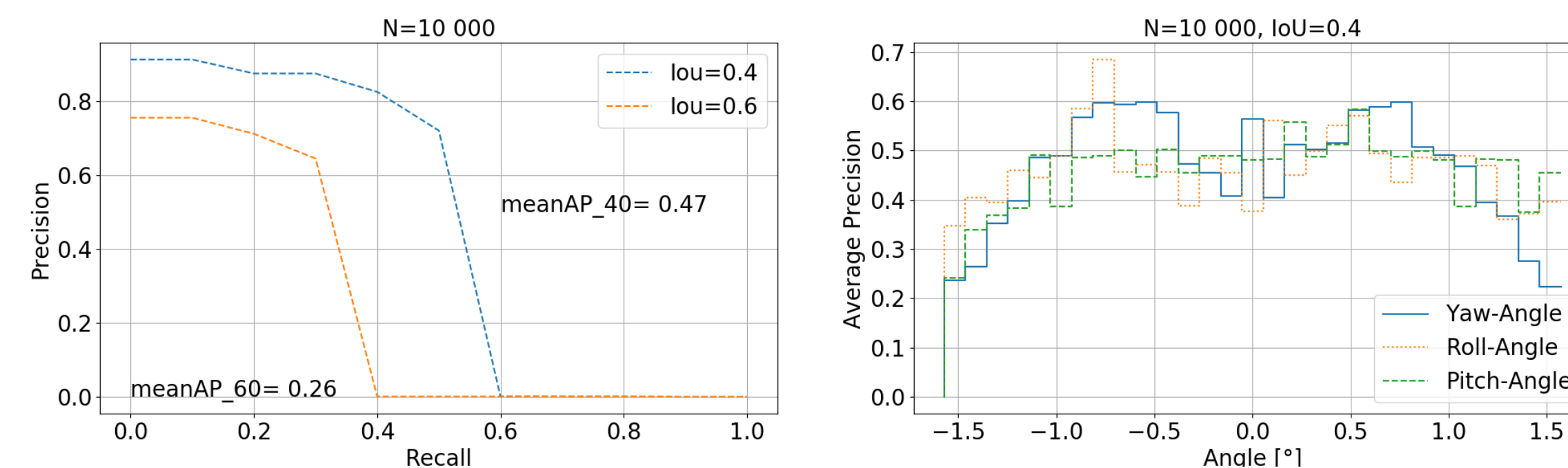
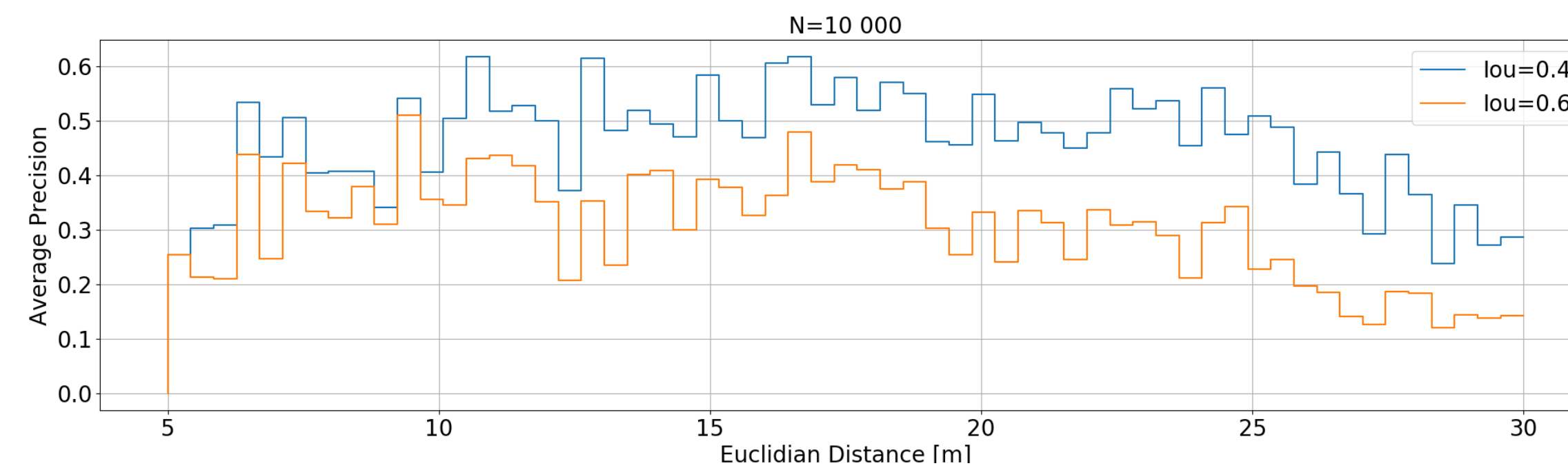


Figure: 3D Modell of Race Court



## Results

The results presented are evaluated on a generated dataset with 10 000 images. A detection is noted as correct if the intersection-over-union (IOU) between the predicted and the true bounding box exceeds a threshold. The average precision corresponds to the average precision over all recall levels. To evaluate the effects of the camera position results are grouped in bins of 0.5m / 6°. For each bin the average precision is calculated.

## Object Detectors

Object detection is a well-studied field in computer vision. Several benchmarks like the Pascal VOC challenge evaluate methods in terms of detection performance. For this thesis speed plays an important role. Methods with high detection performance that take several seconds to process on image can't be used on the MAV. To this end one model has been evaluated namely the "Yolo" - object detector [1]. Potential further methods are "SSD" [2] or DPMs [3]. An example prediction for Yolo and SSD can be seen in Figure 4.

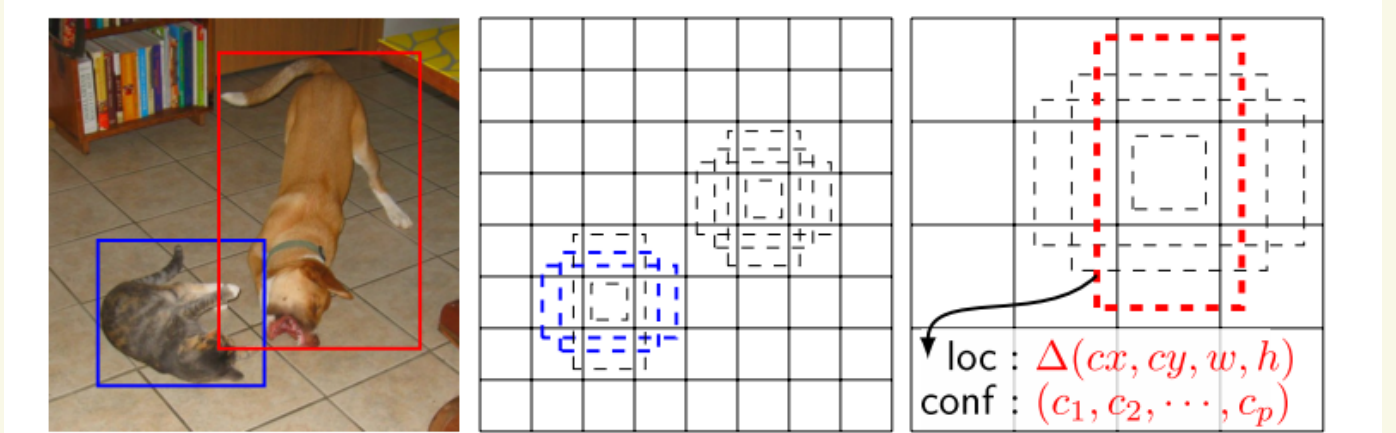


Figure: Example SSD [2]

## Conclusion

The overall meanAP of 47 % with an IoU of 40% can't be considered as decent result yet. However, this also contains results at positions that are out of the expected flight range. The results across precisions are quite inconsistent, one can see a slight deterioration in performance at higher and lower distances. Increasing the IoU threshold to 0.6 leads to heavy deterioration in performance. Hence, the bounding boxes are not that accurate yet. As this is crucial for the MAV navigation, improvement is required.

## References

1. Redmon et al. YOLO9000: Better, Faster, Stronger. *Data Mining with Decision Trees*, 7(3):352350, 2016.
2. Wei et al. Liu. SSD: Single Shot MultiBox Detector.
3. Pedro Felzenszwalb, David McAllester, and Deva Ramanan. A discriminatively trained, multiscale, deformable part model. *In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.