# Visual Reaction: Learning to Play Catch with Your Drone
# –Supplementary Material–

Kuo-Hao Zeng[1] Roozbeh Mottaghi[1,2] Luca Weihs[2] Ali Farhadi[1]
[1]Paul G. Allen School of Computer Science & Engineering, University of Washington
[2]PRIOR @ Allen Institute for AI

## 1. Complete list of objects

We use 20 objects for the experiments: *alarm clock*, *apple*, *basketball*, *book*, *bowl*, *bread*, *candle*, *cup*, *glass bottle*, *lettuce*, *mug*, *newspaper*, *salt shaker*, *soap bottle*, *statue*, *tissue box*, *toaster*, *toilet paper*, *vase* and *watering can*.

## 2. Results for the case that the camera is fixed

In Tab. 1, we provide the results for the case that the drone camera is fixed and does not rotate. In this experiment, we set horizon $H = 3$ and number of action sequences $N = 100,000$. The performance degrades for the case that the camera does not rotate, which is expected.

| | GT. | Est. | Fixed |
|---|---|---|---|
| **Ours, uniform AS** | 54.7 | 26.0 | 18.6 |
| **Ours, full** | 59.3 | 29.3 | 20.2 |

Table 1: **Camera orientation results.** We show the results for the scenario that the camera orientation does not change. **GT** corresponds to the case that we use the ground truth camera orientation at train/test time. The ground truth camera orientation is obtained via ground truth object's position and drone's position. **Est** denotes the case that we use the predicted object and drone positions to calculate to estimate the camera angle. **Fixed** denotes the case that the camera orientation is fixed.

## 3. More statistics of object properties

We show more statistics about our dataset in the Fig. 4, including the *mass*, *average acceleration* along the trajectories, *bounciness*, *drag*, and *angular drag*. Drag is the tendency of an object to slow down due to friction.

## 4. Details of the model architecture

Fig. 2 summarizes the details of the model architecture.

## 5. List of objects for the unseen categories experiment

We selected a subset of 5 objects as our held-out set such that they have different physical properties: *basketball*, *bowl*, *bread*, *candle*, *watering can*. We trained our model on the rest of the objects: *alarm clock*, *apple*, *book*, *cup*, *glass bottle*, *lettuce*, *mug*, *newspaper*, *salt shaker*, *soap bottle*, *statue*, *tissue box*, *toaster*, *toilet paper* and *vase*.

## 6. Visualization of forecasted trajectory

We visualize two examples of the forecasted trajectory in Figure 1. The examples demonstrate that the drone is able to plan according to the future positions of the objects predicted by the forecaster.

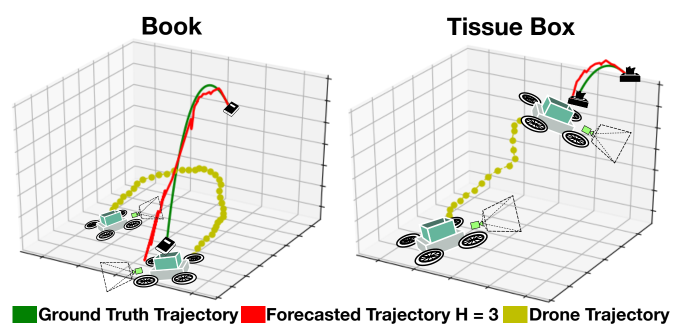


Figure 1: **Qualitative examples of the forecasting trajectory.** The green color, red color, and yellow color denote the ground truth object's trajectory, forecasted object's trajectory, and drone's moving trajectory, respectively.

## 7. Error of position, velocity and acceleration prediction.

The error of our method (L2 distance) for predicting position, velocity and acceleration are 0.644±0.389, 0.037±0.017, and 0.007±0.014, respectively. The error for the baseline CPP for example is 0.686±0.362, 0.148±0.033, and 0.007±0.013 for position, velocity and acceleration, respectively. This comparison is performed for $N = 100,000$.

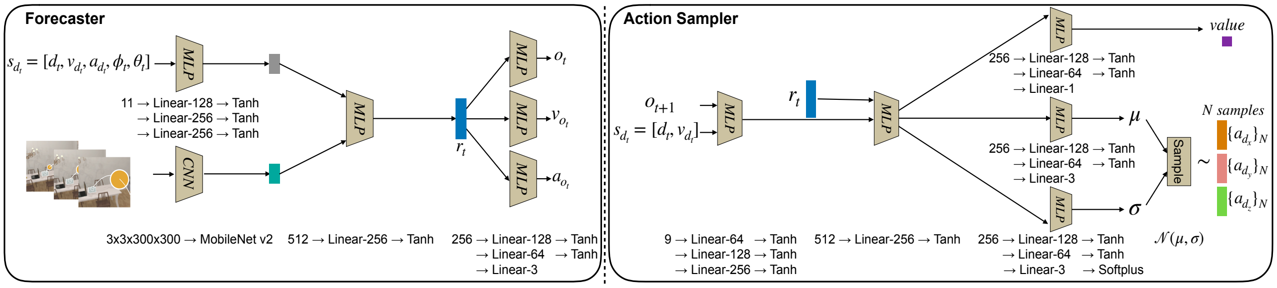


Figure 2: **Detailed architecture of the forecaster and action sampler.**

# 8. Different horizon length

Here, we show how the performance changes with varying the horizon length $H$ (Fig. 3). We observe a performance decrease for horizons longer than 3. The reason is that the learned forecaster has a small error and the error for each time-step accumulates. Thus, training an effective model with longer horizons is challenging and we leave it for future research.

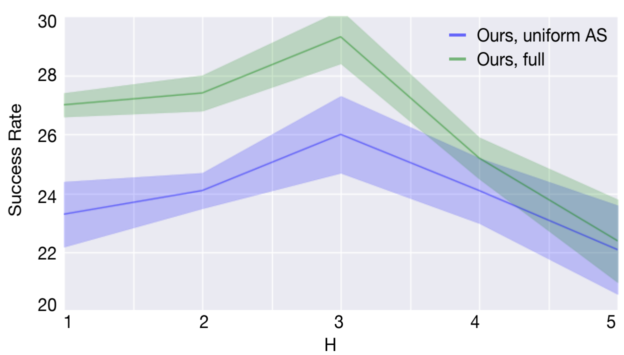


Figure 3: **Result for different horizon lengths.** We show how the performance changes by varying $H$.

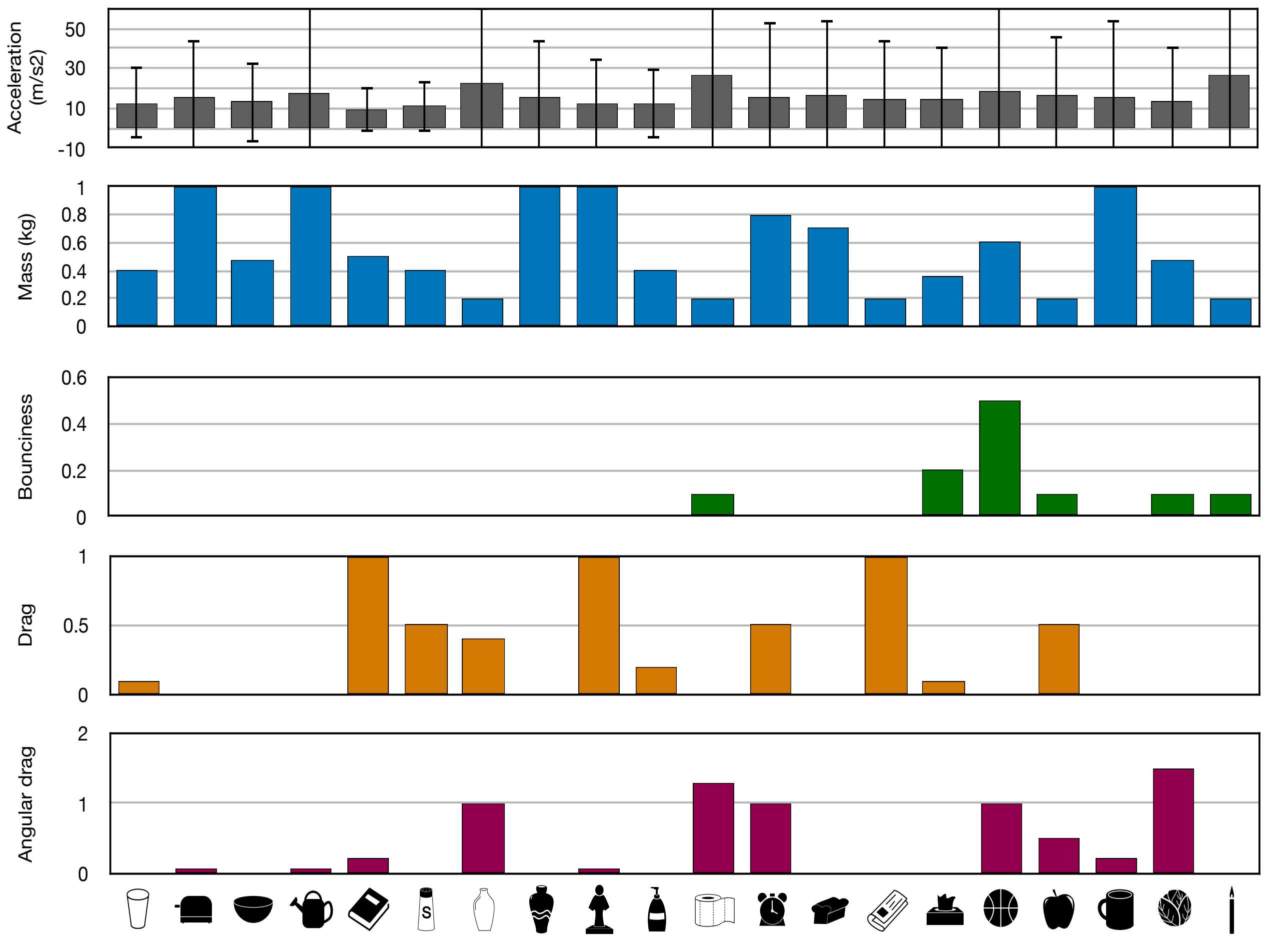


Figure 4: **More dataset statistics.** We provide more statistics for the 20 types of objects in our dataset. We illustrate the mass, average acceleration along the trajectories, bounciness, drag, and angular drag.