COMS 4733
Problem Set #4

Mutian Tong - mt3566@columbia.edu

April 28, 2023

Problem 2

(a)

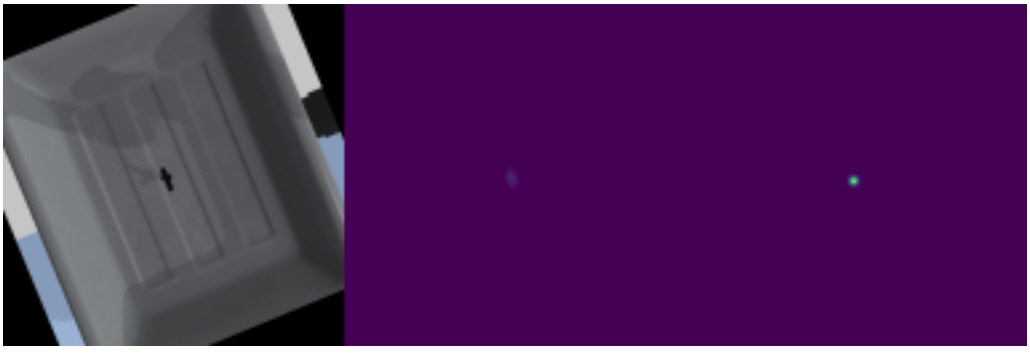
Using a Gaussian scoremap can introduce more smoothness to the data distribution. Different from one-hot pixel image, Gaussian model can represents continuous values in the scoremap, and thus makes the model easier to learn and can also provde better gradient in the backpropagation process in training part.

(b)

Self.aug-pipeline applies a random translation and rotation at 70% of times. For the times it indeed applies the pipeline, it first randomly translates the image both horizontally and vertically to a degree of 0.2 times their original position (either increase or decrease the coordinates). And also randomly rotates the image and keypoints to a angles in the range between 22.5/2 to -22.5/2

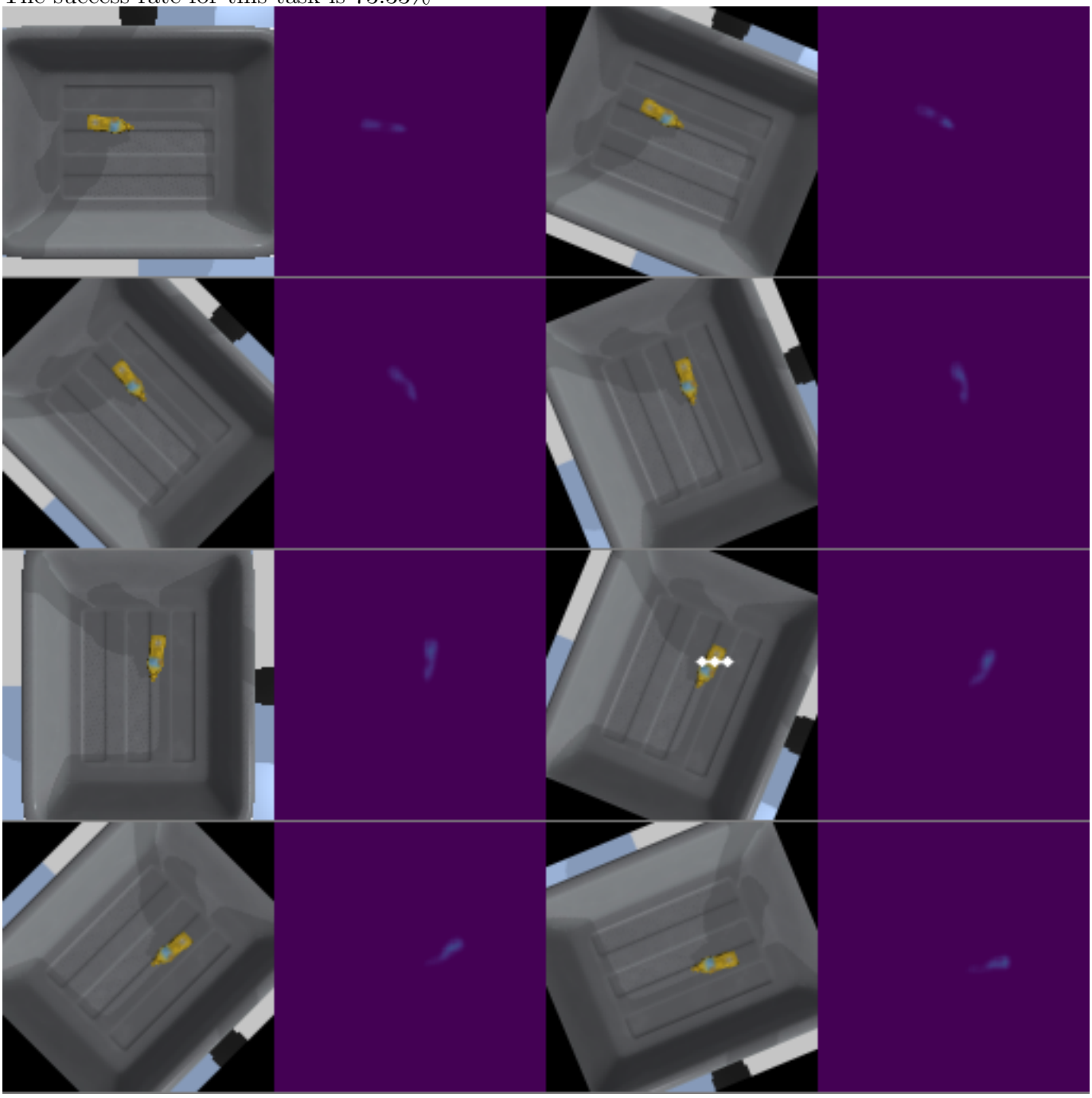
(d)

The training error is 0.0013, test error is 0.0011



(f)

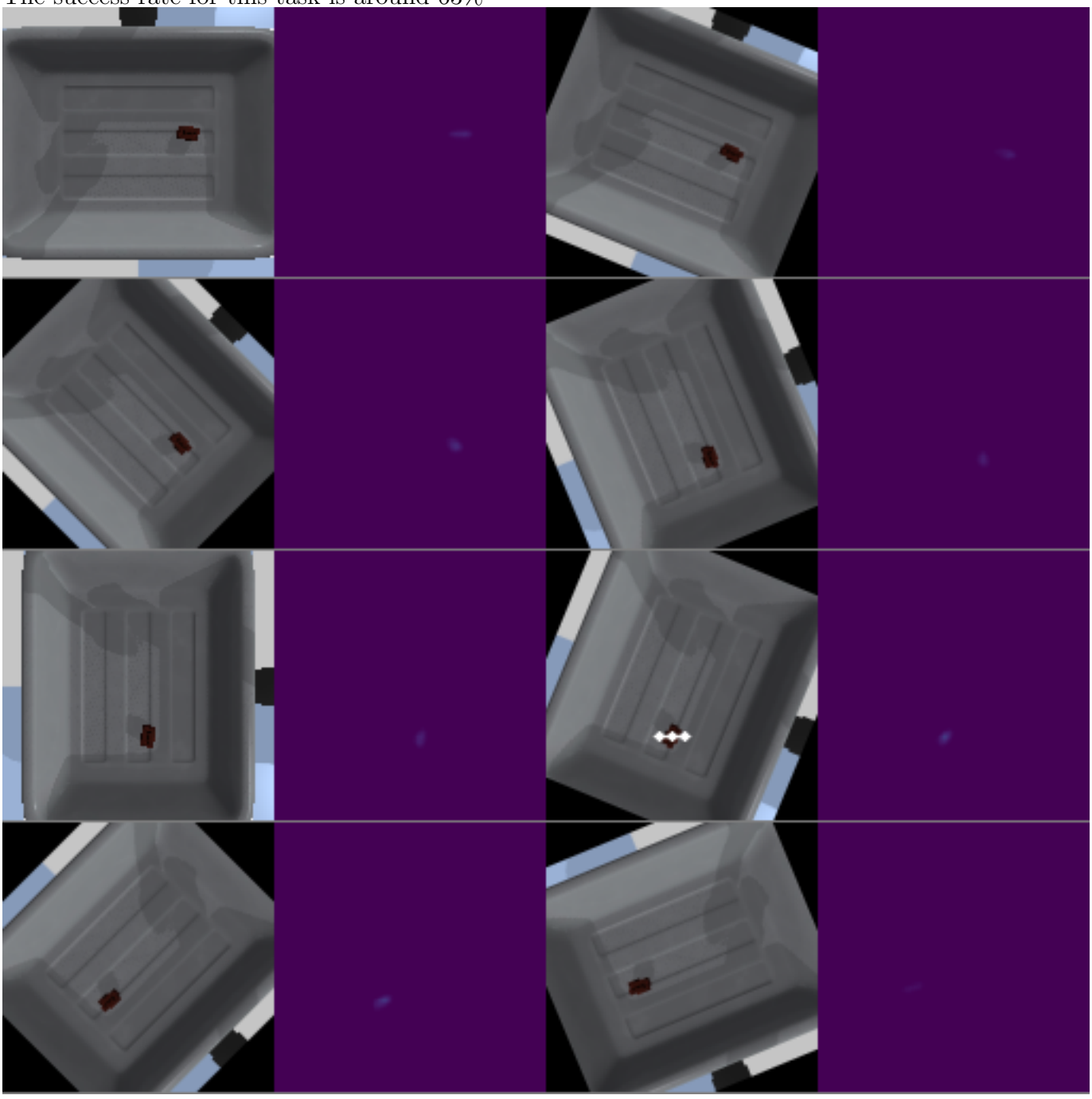
The success rate for this task is 73.33%



link for the video: Clickhere

(g)

The success rate for this task is around 63%



The corresponding map displays a slightly stronger shift to the true position of the object compared with the correspondence in the training sets' object
Link to the video: clickhere

(h)

There are 6 objects left in the end, with 9 objects moved.
Link to the video clickhere

(i)

Firstly, we have employed the data augmentation techniques to generate new samples from the given dataset and efficiently increases variety of training examples. Despite, we have also chosen the Gaussian scoremaap instead of 0-1 hot map, which makes the model more generalizable.

Problem 3

(a)

Link to the video: clickhere

(b)

link to the video: ClickHere

(c)

Link to the video:ClickHere

(d)

the performance of evaluation on the training group holds almostly the same due to the fact that it is very rare to get the same grasp for different trail. However, both the success rate on the testing objects and mixed objects increase quite a bit. The accuracy rate on the testing objects increases from around 63 to 85 (20% higher) while the model now can pick 13 objects from 15 in the empty bin task compared with 9 before.

Problem 4

(a)The test loss is around 0.02, while the training loss is around 0.01

(b)

The success rate for this part is around 6.67%

Link to the video: Clickhere

(c)

Link to the video: Clickhere
14 objects are left in the end.
This method performs pretty bad due to the fact that it still performs regression in a 3-dimension actions vector, while our previous affordance map contains much more information and learning signal and it is easier for it to learn the complex relation between variables and result. On the contrary, a single action vector as the output of the action regression model may not be able to account for all the variabilities in the data.