6.857 Final Project: Milestone 2

Laura Jarin-Lipschitz, Sebastiani Aguirre Navarro and Rachel Holladay

October 11, 2017

1 Algorithms and Evaluation

(To be written more formally and filled in)

- Our goal is to predict grasp success given a grasp and image - We will use the dataset provided by big bird[1] that gives point clouds of objects - we will use software released by authors of [2] to generate grasps and we will implement the labeling scheme provided in the paper to create the data set - the data set is quite large so maybe we will only use a portion? (like maybe only a subset of the objects?)

- in the paper they implement four-layer CNN that is the same structure as LeNet [3]. "two convolutional/pooling layers followed by one inner product layer with a rectified linear unit at the output and one more inner product layer with a softmax on the output. " - Input is representation of object surfaces as seen by depth sensor and the grasp candidate. The output is prediction of whether or not the grasp candidate is a grasp. - Use stochastic gradient descent with specified learning rate - They explored the grasping representation, whether the model was trained with CAD models and level of information about object

- however lenet was originally developed for handwritten character recognition, where they define the MNIST data set on that paper. while this problem is still looking for patterns, it is a different application - therefore in contrast we want to change the CNN structure and learning design choices to see their effect on performance. - we will treat what is in the paper as the baseline and compare it against model alterations - More specifically we will change **RH: these edits to the CNN - need to fill some different things we could try**

- in the paper they use several metrics: they compare accuracy as a function of number of training iterations. **RH: We will also use this?** - they also run a series of real robot experiments. as mentioned in milestone 1, we will not be doing this - **RH: We will use what other metrics?**

2 Risk Management

(To be more formally written out later)

- We are provided the data set and grasping generation algorithm. We have to implement the labeling algorithm from the paper, which means there is a risk that we make a mistake in our labeling implementation. This would lead to a biased data set, which could break everything. we plan to stress test our labeling procedure and visually inspect (?) the labeling

- we hope to treat the network in the paper as the baseline and, by exploring other architectures, outperform the network. however there is a chance that we cannot. however even if we cant do better we hope to explore the effects of different architectures to better understand.

- we are going to be testing everything in simulation. This is due to time and resource limitations. However, robotics is inherently about a real robot interacting with the physical world. Therefore, there is a part (from the robotics side, not the ML side) that is incomplete

3 Division of Labor

RH: I'm really not sure what to say about this

References

[1] A. Singh, J. Sha, K. S. Narayan, T. Achim, and P. Abbeel, "Bigbird: A large-scale 3d database of object instances," in *ICRA*, pp. 509–516, IEEE, 2014.

[2] A. t. Pas, M. Gualtieri, K. Saenko, and R. Platt, "Grasp pose detection in point clouds," *arXiv preprint arXiv:1706.09911*, 2017.

[3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.