A Robust Grasping Policy Based on Derivative-Free Optimization and Grasp-Quality Neural Networks Trained with Synthetic Point Clouds and Grasps from Dex-Net 2.0

Jeffrey Mahler[1], Sherdil Niyaz[1], Ken Goldberg[1,2]

***Abstract*— Recent research demonstrates that grasping policies based on Grasp Quality Convolutional Neural Networks (GQ-CNNs) that are trained to classify grasp success from massive datasets of synthetic 2.5D point clouds, grasps, and analytic grasp metrics can be used to reliably plan grasps on a physical robot with no additional training on real data. Previous experiments suggest that GQ-CNN-based policies tend to predict few false positives on real data but often fail to find a robust grasp for objects that can only be successfully grasped in a small set of precise gripper configurations, such as when a finger must be placed in a small opening. Thus, we explore a robust grasping policy that uses the cross entropy method (CEM), a derivative free optimization technique, to optimize for the most robust grasp predicted by the GQ-CNN on a given input depth image. In 100 grasping trials across 40 novel household objects, we find that the CEM-based grasping policy has 99% precision (only one false positive out of 69 grasps classified as successful) and increases the overall grasp success rate from 80% to 94%. This result, along with visualizations of the probability of success predicted by GQ-CNNs over the continuous input space of grasps, suggest that GQ-CNNs learn a detailed representation of grasp quality and the collision space between the gripper and object. We also explore finetuning the fully connected layers of the GQ-CNN on real data and find that performance does not improve, suggesting that more sophisticated transfer learning methods may be necessary to further improve performance.**

I. Introduction

Reliable robotic grasping across a wide variety of objects and environments is challenging due to imprecision in sensing, actuation, and models, which leads to uncertainty about properties such as object shape, pose, and mass. Recent results [11], [14] demonstrate that training deep neural networks on large, empirically collected datasets of grasping outcomes can generalize across many objects, consistent with results in computer vision based on deep learning from datasets containing millions of examples [9], [17]. However, this generalization ability requires months of data collection time on a physical system and performance appears to plateau around 80%. On the other hand, analytic methods based on physical models, such as robust grasp planning and caging, can be scaled to generate millions of examples using Cloud Computing [13] but may be subject to modeling errors, as they assume precise knowledge of physics and known uncertainty distributions.

[1] Dept. of Electrical Engineering and Computer Science;
[2] Dept. of Industrial Operations and Engineering Research;
University of California, Berkeley, USA {jmahler, sniyaz, goldberg}@berkeley.edu

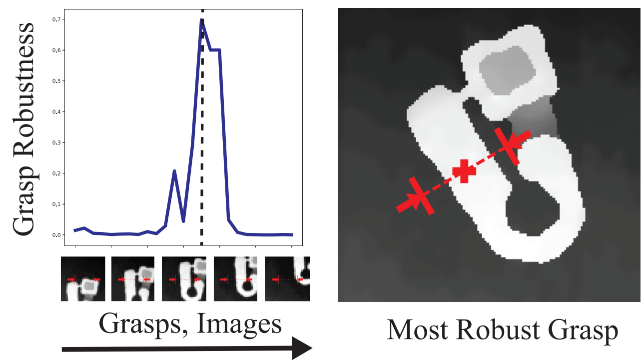

Fig. 1: (Left) Grasp robustness predicted by a Grasp Quality Convolutional Neural Network (GQ-CNN) trained with Dex-Net 2.0 over the space of depth images and grasps for a single point cloud collected with a Primesense Carmine. As the center of the gripper moves from the top to the bottom of the image the GQ-CNN prediction stays near zero and spikes on the most robust grasp (right), for which the gripper fits into a small opening on the object surface. This suggests that the GQ-CNN has learned a detailed representation of the collision space between the object and gripper. Furthermore, the sharp spike suggests that it may be difficult to plan robust grasps by randomly sampling grasps in image space. We consider planning the most robust grasp using the cross-entropy method on the GQ-CNN response.

Our recent research [12] suggests that Grasp Quality Convolutional Neural Networks (GQ-CNNs) trained on Dex-Net 2.0, a dataset of 6.7 million synthetic point clouds, grasps, and analytic grasp metrics generated from 3D models, can be used to reliably plan grasps on a physical robot without any fine-tuning on real point clouds. However, our grasp planning was based on a heuristic random sampling procedure to generate candidates from a continuous set of possible gripper configurations (a planar pose and height about a table). This resulted in non-deterministic behavior and often failed on objects that required precision to grasp, such as placing a gripper jaw in a small opening.

In this work we explore a robust grasping policy that optimizes the GQ-CNN output over a set of continuous grasp candidates using the cross entropy method, a derivative-free optimization technique. To evaluate the optimization we also develop robustness maps, visualizations of the function learned by the GQ-CNN in image space for a fixed grasp axis and height above the table. Experiments with an ABB YuMi suggest the CEM-based robust grasping policy achieves a 99% precision score and 94% success rate in 100 trials on a dataset of 40 novel household objects without any fine-tuning on physical grasping outcomes. We also explore finetuning the fully connected layers of the GQ-CNN on batches of 100 physical datapionts and find that finetuning decreases performance for all values of the hyperparameters we tried.

This suggests a need for either improved models of sensors and grasping, more sophisticated transfer learning methods, or larger datasets of physical outcomes in order to further improve performance.

II. Related Work

Grasp planning considers finding a gripper configuration that maximizes a success (or quality) metric subject to kinematic and collision constraints. There are three primary methods for grasp planning: *analytic* methods [16], which consider performance according to physical models given a perfect known world state [15], *empirical* (or data-driven) methods [2], which learn a model to plan grasps on sensory data usuing human labels [1] or physical trials [11], [14], and *hybrid* methods, which learn a grasping policy from large datasets generated using automated methods such as dynamic simulation [6] or robust grasp analysis [12]. We consider improving the performance of hybrid grasp planners for grasping novel objects on a physical system with derivative free optimization.

Analytic Methods. Analytic robust grasp planning (RGP) methods the expectation value of analytic metrics (such as the ability to resist external wrenches) under uncertainty in perception and actuation. A common method to deploy analytic methods is to precompute a database of grasps for known 3D objects and the quality of each such as GraspIt! [4], and to index this database from sensory data [3], [8]. Recent research suggests that RGP can be accelerated using adaptive sampling over massive datasets of 3D objects and robust grasps such as the Dexterity Network (Dex-Net) 1.0 [13].

Empirical Methods. Research on empirical methods has primarily focused on associating human labels with graspable regions in RGB-D images [10] or point clouds [5]. Another line of research on empirical grasp planning has attempted to optimize success in physical trials directly by collecting large datasets of grasps over several months on a physical robot [11], [14]. In contrast, we consider learning a GQ-CNN model for grasp planning on real point clouds from massive datasets of simulated point clouds and grasps.

Hybrid Methods. To reduce data collection time, recent research has considered learning a model for grasp planning on simulated point clouds using either human labels [7], dynamic simulation [6], or analytic grasp metrics [12]. In our recent work, we generated Dex-Net 2.0, a dataset containing 6.7 million point clouds, grasps, and analytic metics, and used it to learn a GQ-CNN model that reliably classifies grasps as robust on real point clouds without being trained on physical outcomes. While the GQ-CNN had very few false positives on the physical robot, there were several objects for which our grasp planner was not able to find a robust grasp. Thus, in this work we consider derivative-free optimization over candidate grasps and fine-tuning the learned grasping policy on physical grasping outcomes.

III. Problem Statement

We consider the problem of finding a robust grasping policy π_θ parameterized by a set of neural network weights θ that plans a parallel-jaw grasp $\mathbf{u}$ from a depth image $\mathbf{y}$ of a singulated object lying on a planar worksurface to . We parameterize grasps as the planar pose and height of a gripper above the worksurface and assume that the depth sensor is overhead, has known intrinsics, and is approximately orthogonal to the worksurface.

A. Model

We model the image $\mathbf{y}$ as a partial, noisy observation of a true world state $\mathbf{x}$ specifying the camera pose, calibration parameters, object geometry, object pose, and material properties. Let $S(\mathbf{u}, \mathbf{x}) \in \{0, 1\}$ be a binary-valued grasp success metric such as whether or not a grasp $\mathbf{u}$ would lift the object in state $\mathbf{x}$. We define the *robustness* of a grasp [12] as the probability of success under uncertainty about the possible state of the world given the partial observaion: $Q(\mathbf{u}, \mathbf{y}) = \mathbb{E}\left[S \mid \mathbf{u}, \mathbf{y}\right]$.

B. Objective

Our goal is to learn policy parameters θ that maximize the success rate of planned grasps over a distribution on point clouds that can be generated from a set of possible objects $\mathcal{D}$:

$$\theta^* = \underset{\theta \in \Theta}{\operatorname{argmax}}\ \mathbb{E}_{p(\mathbf{y}|\mathcal{D})}\left[Q(\pi_\theta, \mathbf{y})\right]. \qquad \text{(III.1)}$$

IV. Methodology

We propose to approach the objective of equation III.1 by restricting ourselves to a greedy policy $\pi_\theta(\mathbf{y}) = \operatorname{argmax}_{\mathbf{u}\in\mathcal{C}} Q_\theta(\mathbf{u}, \mathbf{y})$ with respect to an estimated robustness function $Q_\theta(\mathbf{u}, \mathbf{y})$, where $\mathcal{C}$ specifies constraints on candidate grasps such as kinematic feasibility. Our approach to grasp planning consists of two parts: (1) learning a robustness function from a massive dataset of synthetic point clouds, grasps, and robust grasp metrics generated from thousands of 3D object models and (2) optimizing for the most robust grasp at runtime using derivative-free optimization.

A. Learning a Robustness Function

We learn a robustness function $Q_{\hat{\theta}}$ represented as a Grasp Quality Convolutional Neural Network (GQ-CNN) from a dataset $(S_i, \mathbf{u}_i, \mathbf{x}_i, \mathbf{y}_i)_{i=1}^N$ sampled i.i.d from a generative model of the grasping process based on physics models of grasp contact and image formation by minimizing the cross-entropy loss for classification:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{i=1}^{N} \mathcal{L}(S_i, Q_\theta(\mathbf{y}_i)).$$

In particular, we use a GQ-CNN trained from the Dex-Net 2.0 dataset of 6.7 million point clouds, grasps, and metrics. The GQ-CNN architecture automatically transforms the pixels of the original image to align the grasp center with the image center and the grasp axis with the middle row of pixels before applying convolutional layers to reduce the dimensionality of input data for learning. See [12] for details. Example data from the Dex-Net 2.0 dataset is available at `http://bit.ly/2pkP1TM`.

B. Evaluating the Robust Grasping Policy

One challenge with planning grasps on physical robot is selecting the greedy grasp with respect to the GQ-CNN predictions. This can be difficult when an object can only be grasped in a small set of precise configurations, such as the example in Fig. 1. Thus in this work we use the cross entropy method (CEM) [11], [18], a form of derivative-free optimization, to optimize for the most robust grasp by iteratively resampling grasps from a learned distribution over robust grasps and updating the distribution. Our method, illustrated in Algorithm 1, models the distribution on promising grasps using a Gaussian Mixture Model (GMM) and seeds the initial set of grasps with antipodal point pairs. The algorithm takes as input the number of CEM iterations m, the number of initial grasps to sample n, the number of grasps to resample from the model c, the number of GMM mixure components k, a friction coefficient μ, and the GQ-CNN Q_θ, and returns an estimate of the most robust grasp $\mathbf{u}$. The qualitative performance of our method on several examples from our experiments is illustrated in Fig. 2.

Input: Num rounds m, Num inital samples n, Num CEM samples c, Num GMM mixture k, Friction coef μ, Elite percentage γ, Robustness function Q_θ
Result: $\mathbf{u}$, most robust grasp
$\mathcal{U} \leftarrow$uniform set of n antipodal grasps;
for i = *1, ..., m* **do**
  $\mathcal{E} \leftarrow$top $\gamma-$percentile of grasps ranked by Q_θ;
  $M \leftarrow$GMM fit to $\mathcal{E}$ with k mixtures;
  $G \leftarrow c$ iid samples from M;
end
return $\underset{\mathbf{u}\in\mathcal{U}}{\text{argmax}}\ Q_\theta(\mathbf{u}, \mathbf{y})$;

Algorithm 1: Robust Grasping Policy using the Cross Entropy Method on a Learned GQ-CNN

V. Experiments

We evaluated the physical performance of our robust grasping policy on an ABB YuMi to benchmark the performance of grasping novel objects presented in isolation. Our experimental procedure was identical to that of [12] except that on each trial we uniformly sample an object from the test set rather than evaluating each object for a fixed number of iterations. Our primary performance metric was precision, or the percentage of executed grasps classified as robust that succeeded on the real robot. We also measured the success rate, or percentage of all grasps that were successful, to penalize the robot for not finding a robust grasp for an object. All experiments ran on a Desktop running Ubuntu 14.04 with a 2.7 GHz Intel Core i5-6400 Quad-Core CPU and an NVIDIA GeForce 980, and we used an NVIDIA GeForce GTX 1080 for training the GQ-CNN.

A. Hyperparameter Optimization

We set the hyperparameters of our algorithm based on grasping performance on the eight adversarial training objects from [12]. The parameters of the CEM-based grasping policy were $m = 5$, $n = 200$, $c = 50$, $k = 5$, and $\gamma = 0.25$. This policy achieved 93% success on the adversarial dataset.

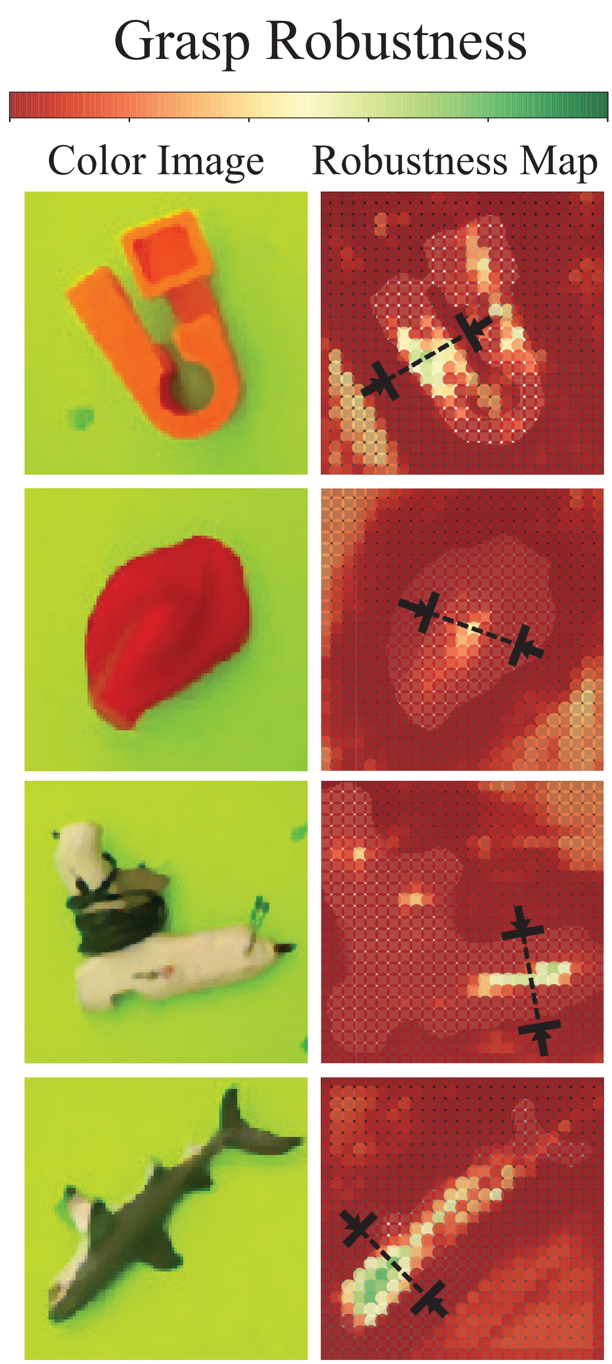

Fig. 2: Example input color images and maps of the grasp robust estimated by the GQ-CNN over grasp centers for a constant grasp axis angle in image space and height above the table, with the grasp planned by our CEM-based robust grasping policy shown in black. CEM is able to find precise robust grasping locations encoded by the GQ-CNN that are very close to the global maximum for the given grasp axis and height. The GQ-CNN also appears to assign non-zero robustness to several grasps that completely miss the object. This is likely because no such grasps are in the training set, and future work could augment the training dataset to avoid these grasps.

B. Performance on Novel Objects

We evaluated performance on a test set of 40 novel household objects augmented from the 10 object test set of [12] which contains objects of various sizes, shapes, and material properties. The dataset is illustrated in Fig. 3. Objects

Fig. 3: The test set of 40 household objects using for evaluating the generalization performance of our robust grasping policy. The dataset contains both fully rigid objects (e.g. bolt), partially rigid objects (e.g. spray bottle), and deformable objects (e.g. wires) with sizes between $3cm$ and $30cm$.

in the dataset were retricted to weight less than $0.25kg$ (the payload of the YuMi) and to have at least one grasp that can be reached with a gripper width of $5cm$ for each stable resting pose on the table.

On this dataset our GQ-CNN achieved a precision of 98.6% (68 true positives out of 69 grasps classified as robust), a success rate of 94%, and planned grasps in an average of $2.46sec$ (about $3\times$ slower than previous work). The single false positive was on a large serving spoon, where the robot attempted to grasp the bowl of the spoon and failed. We are interested in discussing possible improvements to the runtime of the CEM-based robust grasping policy evaluation such as reducing the dimensionality of input data or using alternative optimization methods.

VI. Discussion: Updating the Robustness Function with Physical Data

We also ran experiments on finetuning the GQ-CNN robustness function using a batch method inspired by Q-learning [19]. First, we used the initial GQ-CNN trained on synthetic data to collect a set of physical outcomes on the adversarial training set of [12] using epsilon greedy with respect to our robust CEM-based grasping policy. Then we finetuned the fully connected layers of the GQ-CNN for a small number of epochs on only the collected data.

Interestingly, we found that finetuning on real data led to increased precision but decreased the success rate on all values of hyperparameters tested ($\epsilon = 0, \epsilon = 0.4, \epsilon = 1.0$ and a batch size of 100). The original policy had a precision of 91% and a success rate of 93%. The best finetuned policy used $\epsilon = 0.4$ and had a precision of 94% and a success rate of 83%. We are interested in discussing alternative transfer learning methods such as active learning and domain confusion for improving the performance of the policy on real data to near 100% success.

References

[1] R. Balasubramanian, L. Xu, P. D. Brook, J. R. Smith, and Y. Matsuoka, "Physical human interactive guidance: Identifying grasping principles from human-planned grasps," *Robotics, IEEE Transactions on*, vol. 28, no. 4, pp. 899–910, 2012.

[2] J. Bohg, A. Morales, T. Asfour, and D. Kragic, "Data-driven grasp synthesisa survey," *IEEE Transactions on Robotics*, vol. 30, no. 2, pp. 289–309, 2014.

[3] C. Goldfeder and P. K. Allen, "Data-driven grasping," *Autonomous Robots*, vol. 31, no. 1, pp. 1–20, 2011.

[4] C. Goldfeder, M. Ciocarlie, H. Dang, and P. K. Allen, "The columbia grasp database," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 1710–1716.

[5] A. Herzog, P. Pastor, M. Kalakrishnan, L. Righetti, J. Bohg, T. Asfour, and S. Schaal, "Learning of grasp selection based on shape-templates," *Autonomous Robots*, vol. 36, no. 1-2, pp. 51–65, 2014.

[6] E. Johns, S. Leutenegger, and A. J. Davison, "Deep learning a grasp function for grasping under gripper pose uncertainty," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 4461–4468.

[7] D. Kappler, J. Bohg, and S. Schaal, "Leveraging big data for grasp planning," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2015.

[8] B. Kehoe, A. Matsukawa, S. Candido, J. Kuffner, and K. Goldberg, "Cloud-based robot grasping with the google object recognition engine," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 4263–4270.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[10] I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," *The International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 705–724, 2015.

[11] S. Levine, P. Pastor, A. Krizhevsky, and D. Quillen, "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *arXiv preprint arXiv:1603.02199*, 2016.

[12] J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg, "Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics," *arXiv preprint arXiv:1703.09312*, 2017.

[13] J. Mahler, F. T. Pokorny, B. Hou, M. Roderick, M. Laskey, M. Aubry, K. Kohlhoff, T. Kröger, J. Kuffner, and K. Goldberg, "Dex-net 1.0: A cloud-based network of 3d objects for robust grasp planning using a multi-armed bandit model with correlated rewards," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*. IEEE, 2016.

[14] L. Pinto and A. Gupta, "Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2016.

[15] F. T. Pokorny and D. Kragic, "Classical grasp quality evaluation: New algorithms and theory," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 3493–3500.

[16] D. Prattichizzo and J. C. Trinkle, "Grasping," in *Springer handbook of robotics*. Springer, 2008, pp. 671–700.

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.

[18] R. Y. Rubinstein, A. Ridder, and R. Vaisman, *Fast sequential Monte Carlo methods for counting and optimization*. John Wiley & Sons, 2013.

[19] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press Cambridge, 1998, vol. 1, no. 1.