Beyond Success Rate: Empirical Evidence That Compliance Dominates in Deformable-Object Grasping

Adedayo Akinade*, Keenan Granland †, Jürgen Leitner* and Michael Burke*
*Department of Electrical and Computer Systems Engineering
Monash University, Australia
Email: adedayo.akinade@monash.edu
†Department of Mechanical and Aerospace Engineering
Monash University, Australia

***Abstract*—Vision-based grasp synthesis algorithms trained on rigid objects are often deployed on deformable objects without adaptation, yet their behaviour under that distribution shift is not well characterised. We study two off-the-shelf algorithms, GG-CNN and a ResNet-based grasp synthesis algorithm, on six 3D-printed EGAD geometries fabricated in Varioshore TPU at two gyroid infill densities (5% and 20%), giving a controlled stiffness manipulation. Grasps are executed on a Franka Emika Panda and scored by combining task success with the density-aware Chamfer distance between point clouds of the object before and during grasping. Both algorithms drop substantially in grasp score on the softer infill variant of every object tested (relative reductions of 14%–25% per object), while differing from each other by less than 0.04 in every condition. For successful grasps, material compliance can exceed the algorithm-choice effect by an order of magnitude, suggesting that uncontrolled compliance is a substantial confound in deformable-object grasping benchmarks. The study also demonstrates that simple, accessible 3D printing with infill-controlled compliance is sufficient to produce deformable objects with characterised mechanical properties suitable for grasp benchmarking.**

I. Introduction

Vision-based grasp synthesis is mature for rigid objects [8, 9, 4], where the natural outcome variable is binary success: an object is either grasped and picked up or it is not, and if successful, its geometry is unchanged. Many real-world objects, however, deform under contact [11, 7]. A grasp on a deformable object can complete the task while permanently distorting or damaging it, but binary success rates do not distinguish such outcomes from gentle successful grasps. The density-aware Chamfer distance (DCD) [12] between point clouds captured before and during grasping has recently been applied to soft-grasp evaluation [6] as a complement to binary success.

This raises an empirical question: *how do off-the-shelf vision-based grasp synthesis algorithms actually perform on volumetric deformable objects?* Existing benchmarks for these methods test them on rigid objects; deformable-object studies have largely focused on developing specialised methods rather than characterising standard methods under direct deployment. This workshop paper reports the core empirical finding of a controlled study on this question.

Many grasp synthesis algorithms learn from depth or RGB-D observations of rigid objects [9, 8, 5], with evaluation typically on rigid object sets such as YCB [2] or EGAD [10] under success-rate protocols [1]. For deformable manipulation, Huang et al. [7] evaluate predefined candidate grasps on volumetric deformable objects in simulation; Greenland et al. [6] demonstrated DCD as a visual deformation measure for soft grasping. Neither characterises how off-the-shelf rigid-grasping methods behave when applied directly to physical deformable objects under controlled material compliance, which is the gap this work addresses.

II. Methodology

A. Test Objects and Compliance Manipulation

Seven EGAD geometries [10] (A0, B1, C2, D3, E4, F5, G6) are fabricated in Varioshore TPU on a Bambu Lab P1P with gyroid infill at two densities (5% and 20%). All other print parameters are held constant. Compression testing on a calibrated Lloyd LRX-Plus load frame confirms that the infill manipulation produces a stiffness contrast of $1.3\times$–$3.3\times$ in compression modulus across six of the selected objects, and the bioyield force scales accordingly, with the 20% variants withstanding substantially larger loads before transitioning to nonlinear deformation. The seventh EGAD geometry (G6) was excluded because its modulus was constant across all infill densities and patterns tested, so the intended manipulation could not be verified mechanically. The full material-property measurements are reported in Table I. These measurements demonstrate that simple infill-density control in a consumer-grade 3D printer is sufficient to produce a deformable object set with quantitatively characterised mechanical properties suitable for grasp benchmarking.

B. Grasp Deformation Scoring

A deformation grasp score is defined on successful grasps only; failure characterisation is the subject of established rigid-object benchmarks [1] and is out of scope. For each successful trial, the score is:

$$s = 1 - \frac{d_{\mathrm{DCD}}}{2}, \tag{1}$$

where $d_{\mathrm{DCD}} \in [0,1]$ is the density-aware Chamfer distance [12] between the pre-grasp point cloud $\mathcal{P}_1$ and the during-grasp point cloud $\mathcal{P}_2$ of the object, computed as

$$\begin{aligned} d_{\mathrm{DCD}}(\mathcal{P}_1, \mathcal{P}_2) = \frac{1}{2}\Bigg[& \frac{1}{|\mathcal{P}_1|} \sum_{x \in \mathcal{P}_1} \left(1 - \tfrac{1}{n_{\hat{y}}} e^{-\alpha \|x - \hat{y}\|_2}\right) \\ & + \frac{1}{|\mathcal{P}_2|} \sum_{y \in \mathcal{P}_2} \left(1 - \tfrac{1}{n_{\hat{x}}} e^{-\alpha \|y - \hat{x}\|_2}\right) \Bigg], \end{aligned} \tag{2}$$

where $\hat{y} = \arg\min_{y \in \mathcal{P}_2} \|x - y\|_2$ and $\hat{x} = \arg\min_{x \in \mathcal{P}_1} \|y - x\|_2$ are nearest-neighbour correspondences, $n_{\hat{y}}$ and $n_{\hat{x}}$ count the number of times a point is referenced as a nearest neighbour, and $\alpha = 50$ is a density-sensitivity parameter following [12, 6]. The grasp score in (1) is bounded in $[0.5, 1]$: a gentle successful grasp approaches 1, a highly deforming grasp leans toward 0.5. Per trial, d_{DCD} is the mean of the lift-checkpoint and stability-checkpoint values. The during-grasp point cloud is isolated from the gripper using MobileSAM [13] on a colour image obtained from a wrist-mounted RGBD camera.

C. Execution and Algorithms

Each object instance (and stiffness level) is presented in six randomly selected poses. We evaluate two algorithms: GG-CNN, a lightweight fully convolutional baseline [9], and a ResNet-based grasp synthesis algorithm [3]. Both are used with the publicly released models without fine-tuning on deformable objects, reflecting the typical deployment scenario. Trials in which the object is dropped are not included in the reported grasp scores; the drop counts are themselves informative and are reported alongside the main results. Specifically, GG-CNN experienced 4 drops and the ResNet-based algorithm experienced 3 drops across the experimental sessions, giving the two algorithms comparable failure rates over the same condition set.

TABLE I

MEASURED MECHANICAL PROPERTIES OF THE PRINTED TEST OBJECTS AT GYROID INFILL. VALUES ARE MEAN ACROSS FOUR REPLICATES. E: COMPRESSION MODULUS; F_y: BIOYIELD FORCE.

Object	Infill (%)	E (MPa)	F_y (N)
A0	5	0.29	204.90
A0	20	0.52	339.65
B1	5	0.09	102.74
B1	20	0.15	144.86
C2	5	0.20	131.97
C2	20	0.33	182.31
D3	5	0.10	78.24
D3	20	0.36	211.33
E4	5	0.29	178.24
E4	20	0.47	258.11
F5	5	0.06	46.16
F5	20	0.10	57.28
G6	5	0.03	26.59
G6	20	0.03	27.11

Fig. 1. Experimental setup. The Franka Emika Panda robot is fitted with a parallel-jaw gripper and two wrist-mounted Intel RealSense cameras: a D435 for grasp synthesis from a wider view and a D405 for short-range deformation observation. The inset shows the six 3D-printed EGAD test objects (A0, B1, C2, D3, E4, F5, G6) fabricated in Varioshore TPU; each object is printed at two gyroid infill densities (5% and 20%).

Experiments are conducted on a Franka Emika Panda 7-DOF arm fitted with a parallel-jaw gripper (FIgure 1. Two Intel RealSense cameras are wrist-mounted: a D435 for grasp synthesis from a wider scene view, and a D405 for short-range deformation observation. Both stream RGB and depth at 640×480 at 30 Hz.

III. RESULTS

A. Algorithm Performance

Overall, the two algorithms produce closely matched scores. GG-CNN reaches 0.830 ± 0.104; the ResNet-based algorithm reaches 0.834 ± 0.097. Per-object, the two algorithms differ by less than 0.04 on every object. Welch's t-tests comparing the two algorithms within each infill condition find no statistically significant difference between them: at 5% infill, $t(70.0) = -1.07$, $p = 0.286$, Cohen's $d = -0.25$; at 20% infill, $t(66.1) = 0.45$, $p = 0.655$, $d = 0.11$. The effect sizes are negligible to small, supporting the descriptive observation that on this object set the choice between the two methods does not produce a measurable difference in grasp performance.

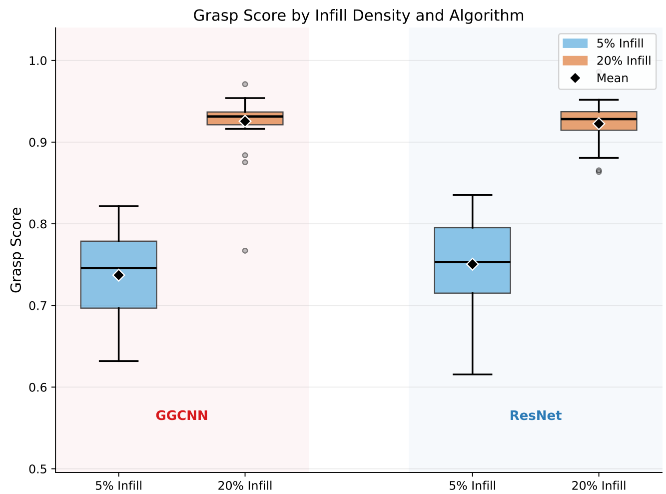

Fig. 2. Distribution of grasp scores by infill density and algorithm. Both algorithms exhibit a large drop on the softer 5% infill variant; the two algorithms behave almost identically within each infill condition.

B. Effect of Compliance

The dominant pattern in the data is a large, consistent drop in grasp score on the softer 5% infill variant relative to the 20% variant (Figure 2). For GG-CNN the mean grasp score falls from 0.926 at 20% infill to 0.738 at 5%, a relative reduction of 20.3%; for the ResNet regressor the corresponding reduction is $0.923 \rightarrow 0.748$, or 19.0%. Standard deviation is also visibly larger on the 5% infill variant ($\sigma = 0.053$) than on the 20% variant ($\sigma \approx 0.03$), so softer objects produce both lower-mean and more-variable performance. Welch's t-tests confirm this pattern statistically: within GG-CNN, $t(57.2) = -18.15$, $p < 0.001$, Cohen's $d = -4.28$ (large); within ResNet, $t(50.2) = -17.84$, $p < 0.001$, $d = -4.20$ (large). The effect sizes are an order of magnitude above the conventional threshold for a large effect ($|d| > 0.8$), indicating that the infill manipulation produces a much larger shift in grasp performance than algorithm-to-algorithm variation within either condition.

C. Consistency Across Objects

The infill effect appears consistently across every object tested. For every object, the 20% infill mean exceeds the 5% infill mean by between 0.13 (B1, ResNet) and 0.23 (C2 and D3, GG-CNN). No object reverses the direction of the effect, and no algorithm is consistently better than the other across objects. The within-object infill effect (0.13–0.23) is roughly an order of magnitude larger than the algorithm-to-algorithm difference (< 0.04 per object).

IV. Discussion

Studies that evaluate grasp synthesis algorithms on deformable objects without controlling for material compliance are likely to be measuring compliance effects as well as algorithm effects, with the two confounded. In our data the compliance effect on a single object exceeds the algorithm-choice effect by an order of magnitude; a study that drew samples across compliance conditions without recording or controlling them would attribute that variation to algorithm performance.

The near-identical behaviour of GG-CNN and the ResNet-based regressor is itself worth recording. Both methods were trained on rigid-object datasets and their architectural differences are substantial, yet on this object set their grasp-score distributions are not statistically distinguishable in either infill condition (both $p > 0.28$), with effect sizes that are negligible-to-small ($|d| = 0.25$ at 5% infill; $|d| = 0.11$ at 20% infill; Section III). One reading is that both algorithms are limited by features the rigid-object training distribution does not provide; another is that the grasp score we report is insensitive to algorithm-level differences that would appear under other metrics. Evaluating methods trained or fine-tuned on deformable data, and reporting additional performance dimensions beyond DCD, are natural next steps.

The empirical analysis here is conditional on grasp success: trials in which the algorithm's selected grasp fails to lift the object are excluded. The paper therefore characterises *how* grasps deform objects when they succeed, not how often they succeed. The deformation measurement is also purely geometric and does not estimate internal stress or proximity to material yield; integrating a stress-based component into the scoring formulation would meaningfully extend the methodology and is a direction we believe worth pursuing.

V. Conclusion

We presented a method for empirical characterisation of off-the-shelf vision-based grasp synthesis algorithms. Our results show that accessible 3D printing techniques can be used to generate controllable deformation characteristics for grasp benchmarking. We used this approach to characterise two vision-based grasp synthesis algorithms on six 3D-printed EGAD geometries with controlled, characterised material compliance. The two algorithms behave comparably to each other on every object; both show a large and consistent drop in grasp score on the softer infill variant. The compliance effect exceeds the algorithm-to-algorithm difference by roughly an order of magnitude, which suggests that uncontrolled material compliance is a substantial confound in deformable-object grasping benchmarks.

References

[1] Yasemin Bekiroglu, Naresh Marturi, Máximo A. Roa, Komlan Jean Maxime Adjigble, Tommaso Pardi, Cindy Grimm, Ravi Balasubramanian, Kaiyu Hang, and Rustam Stolkin. Benchmarking protocol for grasp planning algorithms. *IEEE Robotics and Automation Letters*, 5 (2):315–322, 2020. doi: 10.1109/LRA.2019.2956411.

[2] Berk Calli, Aaron Walsman, Arjun Singh, Siddhartha Srinivasa, Pieter Abbeel, and Aaron M. Dollar. Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set. *IEEE Robotics & Automa-*

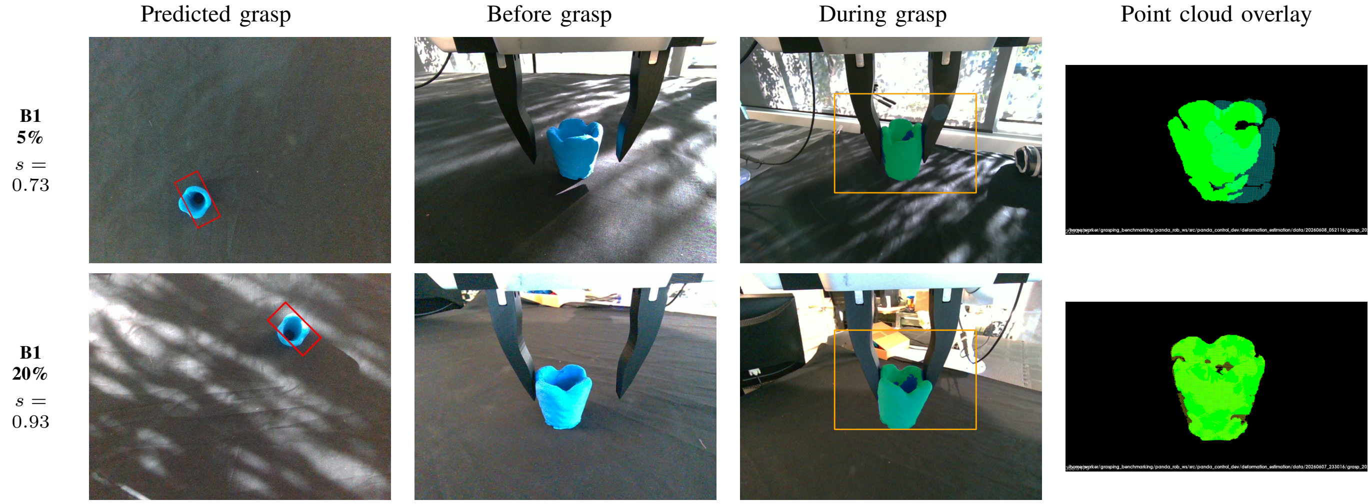

Fig. 3. Grasp progression for the B1 EGAD geometry at the two infill densities. Each row shows, from left to right, the predicted grasp pose overlaid on the pre-grasp RGB image, the pre-grasp scene, the object held in the gripper during lift, and the point cloud overlay in which the pre-grasp cloud (green) is rendered alongside the during-grasp cloud (orange) after SAM-based segmentation. The trial's grasp score s is reported at the left of each row. On the softer 5% infill variant (top row), the during-grasp point cloud departs visibly from the pre-grasp shape, corresponding to a lower grasp score. On the stiffer 20% variant (bottom row), the two clouds align closely and the grasp score remains high.

tion Magazine, 22(3):36–52, 2015. doi: 10.1109/MRA.2015.2448951.

[3] F. Chu, R. Xu, and P. A. Vela. Real-world multiobject, multigrasp detection. In *IEEE Robotics and Automation Letters*, volume 3, pages 3355–3362, Oct 2018. doi: 10.1109/LRA.2018.2852777.

[4] Hao-Shu Fang, Chenxi Wang, Minghao Gou, and Cewu Lu. Graspnet-1billion: A large-scale benchmark for general object grasping. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11441–11450, 2020. doi: 10.1109/CVPR42600.2020.01146.

[5] Haoshu Fang, Chenxi Wang, Hongjie Fang, Minghao Gou, Jirong Liu, Hengxu Yan, Wenhai Liu, Yichen Xie, and Cewu Lu. Anygrasp: Robust and efficient grasp perception in spatial and temporal domains. *IEEE Transactions on Robotics*, 39:3929–3945, 2023. URL https://api.semanticscholar.org/CorpusID:254823166.

[6] Benjamin G. Greenland, Josh Pinskier, Xing Wang, Daniel Nguyen, Ge Shi, Tirthankar Bandyopadhyay, Jen Jen Chung, and David Howard. SoGraB: A visual method for soft grasping benchmarking and evaluation. In *2025 IEEE 8th International Conference on Soft Robotics (RoboSoft)*, pages 1–6, 2025. doi: 10.1109/RoboSoft63089.2025.11020918.

[7] Isabella Huang, Yashraj Narang, Clemens Eppner, Balakumar Sundaralingam, Miles Macklin, Ruzena Bajcsy, Tucker Hermans, and Dieter Fox. Defgraspsim: Physics-based simulation of grasp outcomes for 3d deformable objects. *IEEE Robotics and Automation Letters*, 7(3):6274–6281, July 2022. ISSN 2377-3766. doi: 10.1109/LRA.2022.3158725.

[8] Jeffrey Mahler, Jacky Liang, Sherdil Niyaz, Michael Laskey, Richard Doan, Xinyu Liu, Juan Aparicio Ojea, and Ken 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.

[9] Douglas Morrison, Peter Corke, and Jürgen Leitner. Closing the loop for robotic grasping: a real-time, generative grasp synthesis approach. In Hadas Kress-Gazit, Siddhartha S. Srinivasa, Tom Howard, and Nikolay Atanasov, editors, *Robotics*, Robotics: Science and Systems, United States of America, 2018. The MIT Press. ISBN 9780992374747. doi: 10.15607/RSS.2018.XIV.021.

[10] Douglas Morrison, Peter Corke, and Jürgen Leitner. Egad! an evolved grasping analysis dataset for diversity and reproducibility in robotic manipulation. *IEEE Robotics and Automation Letters*, 5(3):4368–4375, July 2020. ISSN 2377-3766. doi: 10.1109/LRA.2020.2992195.

[11] Jose Sanchez, Juan-Antonio Corrales, Belhassen-Chedli Bouzgarrou, and Youcef Mezouar. Robotic manipulation and sensing of deformable objects in domestic and industrial applications: a survey. *The International Journal of Robotics Research*, 37(7):688–716, 2018. doi: 10.1177/0278364918779698. URL https://doi.org/10.1177/0278364918779698.

[12] Tong Wu, Liang Pan, Junzhe Zhang, Tai Wang, Ziwei Liu, and Dahua Lin. Density-aware chamfer distance as a comprehensive metric for point cloud completion. In *In Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

[13] Chaoning Zhang, Dongshen Han, Yu Qiao, Jung Uk Kim, Sung-Ho Bae, Seungkyu Lee, and Choong Seon Hong. Faster segment anything: Towards lightweight sam for mobile applications. *arXiv preprint arXiv:2306.14289*, 2023.