

目标识别与力估计使用 GelSight Baby Fin Ray

Sandra Q. Liu

Department of Mechanical Engineering
Massachusetts Institute of Technology United States
sqliu@mit.edu

Yuxiang Ma

Department of Mechanical Engineering
Massachusetts Institute of Technology United States
yxma@csail.mit.edu

Edward H. Adelson

Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology United States
adelson@csail.mit.edu

Abstract: 近期在软体机器人手和触觉传感方面的进展使得两者能够在机器学习的帮助下执行越来越多的复杂任务。特别是，我们在先前的工作 [1] 中介绍了 GelSight Baby Fin Ray，它将相机与柔软、合规的 Fin Ray 结构相结合。基于相机的触觉感知赋予了 GelSight Baby Fin Ray 捕捉丰富的接触信息（如力、物体几何形状和纹理）的能力。此外，我们的先前工作表明，GelSight Baby Fin Ray 能够穿过混乱环境，并对壳内坚果进行分类。为进一步考察 GelSight Baby Fin Ray 的潜力，我们利用学习来区分壳内坚果的纹理并执行力和位置估计。我们在流行的神经网络结构中实现了消融研究，包括 ResNet50、GoogLeNet 以及 3 层和 5 层层卷积神经网络 (CNN) 结构。我们得出结论，机器学习是一种有前景的技术，可以从高分辨率触觉图像中提取有用信息，并赋予软体机器人更好地理解 and 与环境互动的能力。

Keywords: 触觉传感器，鱼鳍状结构，接触感应，物体感知

1 介绍

仿鳍手指的设计灵感来自于鱼鳍的变形，其中鳍会根据施加在其结构上的力弯曲和适应 [2]。因此，仿鳍可以应用于不同的机器人任务，并且由于其结构简单，易于打印和原型制作。尽管许多研究集中在机械行为优化上 [3, 4]，但将传感器集成到仿鳍中的工作却很少见。如果具有触觉反馈，仿鳍将会受益，一些作品专注于添加力传感器或使用嵌入式相机和机器学习来进行力估计 [5, 6]。

最近，已经有一些工作致力于将传统刚性的基于摄像头的传感器整合到各种软手指设计中 [7, 8, 9, 10]。具体来说，基于摄像头的传感器可以提供更高分辨率的细节，分辨物体几何形

状，并测量接触力 [11]。原始的 GelSight Fin Ray 能够利用其顺应结构的被动适应性和由摄像头传感器提供的高分辨率触觉感知能力 [12]。这些功能使其能够抓取各种物体、跟踪表面力并执行触觉变形的三维重建。进一步的机械改进带来了 GelSight Baby Fin Ray，它比其前身更加顺应性、坚固和紧凑 [1]。



图 1: 左边，GelSight Baby Fin Ray 抓住了一个核桃；右边是相应的原始触觉图像 [1]。

我们展示了进一步利用 Fin Ray 结构中基于相机的触觉传感器的工作。在这项工作中，我们实现了多种机器学习算法来分类壳内坚果的纹理，并进行力估计和沿触觉传感器长度的位置确定。

2 相关工作

机器学习技术已被用于多个刚性手指传感器，与基于摄像头的触觉传感器相结合。郭等人和黄等人能够利用机器学习方法补充这些传感器，以估算容器内固体颗粒和液体的特性 [13, 14]。其他论文使用各种学习方法提取有关表面力和触觉变形三维几何重建的信息 [15, 16, 17]。还有一些人进行了纹理分类 [7]。另一方面，只有少数论文对软体机器人进行本体感觉研究，即状态估计或力预测，并没有真正专注于识别纹理 [18, 19, 9]。这项工作实现了单一摄像头传感器集成到一个柔软、顺从的机器手指中，利用纹理和力估计来进行物体分类。

3 硬件

我们利用了我们在 GelSight Baby Fin Ray[1] 上之前工作中的硬件设计。该设计的表示如图 2 所示。

4 实验与结果

4.1 坚果分类

GelSight Baby Fin Ray 论文最初执行的任务之一是带壳坚果分类。我们针对在 [1] 中使用的 ResNet50 架构 [20] 进行了消融研究，并将结果与 GoogLeNet[21]、一个 3 层和一个 5 层的 CNN[22]、使用径向基函数 (RBF) 和多项式函数的 support vector machine (SVM)[23] 以及 k-nearest neighbors (KNN) 分类算法 [24] 进行了比较。对于 ResNet、CNN 和 GoogLeNet 架

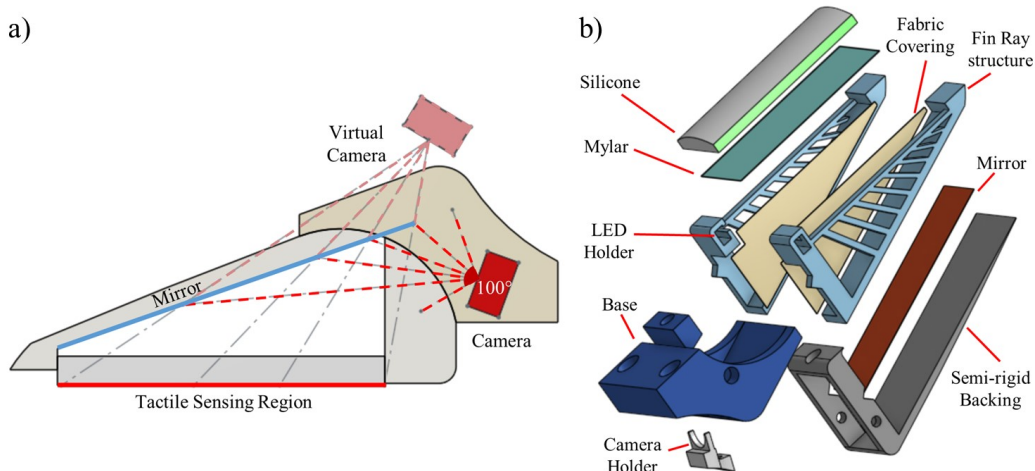


图 2: a) GelSight Baby Fin Ray 的二维表示，展示了相机和镜子能够观察到整个触觉感应区域。b) 不含相机、LED 和扩散器 [1] 的装配分解视图。

构，我们使用未经变形且经过数据增强后的图像作为输入，从而有效地消除了图像中的“负”空间。对于 KNN 和 SVM 算法，我们使用图像中的原始像素创建一个特征向量作为输入。我们通过 `mjpg_streamer` 从树莓派相机流式传输的图像中收集数据。坚果以各种角度在整个触觉表面被压下。每种类型的坚果共捕捉 500 张图像。数据被分为 80% 的训练集和 20% 的验证集。

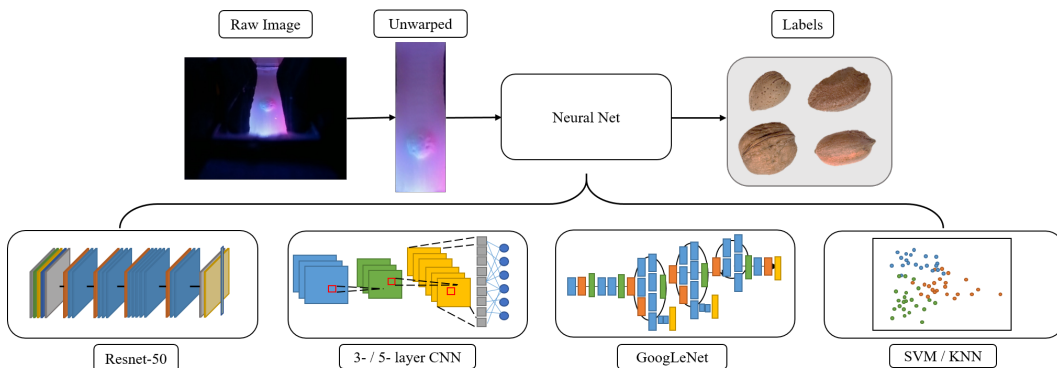


图 3: 为了分类我们的四个类别，杏仁、巴西坚果、山核桃和核桃，我们采用原始图像，进行校正并使用多种神经网络架构（包括 ResNet50、GoogLeNet、3 层和 5 层层的 CNN、SVM 和支持向量机）对其进行训练。

我们的结果如表 1 所示。总体而言，更深的网络在数据集上的表现更好。我们相信这些结果可能是因为更深层次的架构能够识别出更多细微和复杂的细节，而这些细节可能会被较浅层的网络或使用 SVM 和 KNN 方法所忽略。尽管如此，所有的算法和神经网络都表现出色，表明坚果壳的高分辨率触觉细节很容易分辨。

表 1: 坚果分类的准确性

Network	Overall	Almonds	Brazil Nuts	Pecans	Walnuts
3-layer CNN	73.0%	73.6 %	84.0 %	70.8 %	64.0 %
5-layer CNN	87.5%	82.1 %	87.2 %	91.7 %	89.4 %
ResNet	95.6%	98.1%	96.8 %	96.9 %	91.3 %
GoogLeNet	97.5%	96.2 %	98.9%	96.9 %	98.1%
SVM (poly)	86.3%	83.0 %	86.0 %	79.0 %	97.0 %
SVM (rbf)	79.0%	81.0 %	77.0 %	68.0 %	90.0 %
KNN	73.3%	71.0 %	78.0 %	72.0 %	72.0 %

4.2 力估计

我们还利用 Gelsight Baby Fin Ray 建立了基于学习的力估计方法。一个圆柱体或长方体压头和一个 Gelsight Baby Fin Ray 安装在 Franka Panda Hand 上，如图 4 所示。我们编程机器人手将压头推入 Gelsight Baby Fin Ray，同时使用 Nano 17 Force/Torque (ATI) 传感器记录压入力。同时，通过 mjpg_streamer 收集触觉图像。接触点沿 Fin Ray 长度分布在 10 mm 到 50 mm 之间，法向力的范围为 0 到 25 N。每个压头类型捕获总共 60000 张图像。该网络预测法向力和接触位置。

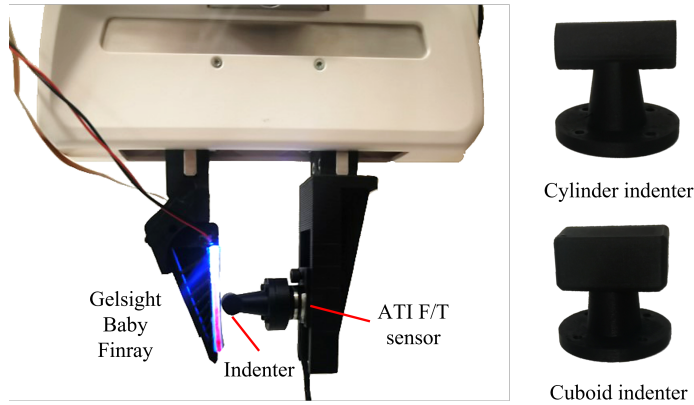


图 4: 使用圆柱形压头和立方体压头的力测量实验装置。

表 2: 接触位置和法向力估计的误差估算

Network	Contact Position Error (mm)	Normal Force Error (N)
3-layer CNN	2.38	1.24
5-layer CNN	5.09	2.21
ResNet50	2.15	1.21
GoogLeNet	2.02	1.18

我们使用 ResNet50、GoogLeNet 以及 3 层和 5 层的 CNN 进行了消融研究。在回归任务中，5 层 CNN 往往具有较大的梯度，导致准确率最低。与分类任务类似，在接触位置估计和法

向力估计方面，GoogLeNet 的表现优于所有其他架构。原因可能是 GoogLeNet 的 inception 模块能够有效地分析不同大小的接触印记。

5 结论

基于相机的触觉传感器结合学习方法可用于同时确定纹理识别、力估计和触觉位移的位置。总体而言，这些能力使我们能够尽量减少手指所需电子设备的数量，并设计策略将这些传感器集成到类似鱼鳍状的顺应性结构中。总之，这些软触觉传感器在未来许多操作任务中有潜在的应用。

Acknowledgments

丰田研究院、海军研究办公室以及 SINTEF BIFROST (RCN313870) 项目提供了资金支持这项工作。

参考文献

- [1] S. Q. Liu, Y. Ma, and E. H. Adelson. Gelsight Baby Fin Ray: A compact, compliant, flexible finger with high-resolution tactile sensing. In 2023 IEEE International Conference on Soft Robotics (RoboSoft), 03 2023.
- [2] O. Pfaff, S. Simeonov, I. Cirovic, P. Stano, et al. Application of fin ray effect approach for production process automation. *Annals of DAAAM & Proceedings*, 22(1):1247–1249, 2011.
- [3] X. Shan and L. Birglen. Modeling and analysis of soft robotic fingers using the fin ray effect. *The International Journal of Robotics Research*, 39(14):1686–1705, 2020.
- [4] Z. Deng and M. Li. Learning optimal fin-ray finger design for soft grasping. *Frontiers in Robotics and AI*, page 161, 2021.
- [5] Y. Yang, K. Jin, H. Zhu, G. Song, H. Lu, and L. Kang. A 3d-printed fin ray effect inspired soft robotic gripper with force feedback. *Micromachines*, 12(10):1141, 2021.
- [6] W. Xu, H. Zhang, H. Yuan, and B. Liang. A compliant adaptive gripper and its intrinsic force sensing method. *IEEE Transactions on Robotics*, 37(5):1584–1603, 2021.
- [7] A. Amini, J. I. Lipton, and D. Rus. Uncertainty aware texture classification and mapping using soft tactile sensors. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 4249–4256. IEEE, 2020.
- [8] Y. She, S. Q. Liu, P. Yu, and E. Adelson. Exoskeleton-covered soft finger with vision-based proprioception and tactile sensing. In 2020 IEEE International Conference on Robotics and Automation (ICRA), pages 10075–10081. IEEE, 2020.
- [9] O. Faris, R. Muthusamy, F. Renda, I. Hussain, D. Gan, L. Seneviratne, and Y. Zweiri. Proprioception and exteroception of a soft robotic finger using neuromorphic vision-based sensing. *Soft Robotics*, 10(3):467–481, 2023.
- [10] S. Q. Liu*, L. Z. Yañez*, and E. H. Adelson. Gelsight EndoFlex: A soft endoskeleton hand with continuous high-resolution tactile sensing. 2023 IEEE International Conference on Soft Robotics (RoboSoft), 2023.
- [11] W. Yuan, S. Dong, and E. H. Adelson. Gelsight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors*, 17(12):2762, 2017.

- [12] S. Q. Liu and E. H. Adelson. Gelsight Fin Ray: Incorporating Tactile Sensing into a Soft Compliant Robotic Gripper. In 2022 IEEE 5th International Conference on Soft Robotics (RoboSoft), pages 925–931, 2022. doi:10.1109/RoboSoft54090.2022.9762175.
- [13] X. Guo, H.-J. Huang, and W. Yuan. Estimating properties of solid particles inside container using touch sensing, 2023.
- [14] H.-J. Huang, X. Guo, and W. Yuan. Understanding dynamic tactile sensing for liquid property estimation, 2022.
- [15] C. Lin, Z. Lin, S. Wang, and H. Xu. Dtact: A vision-based tactile sensor that measures high-resolution 3d geometry directly from darkness, 2022.
- [16] S. Wang, Y. She, B. Romero, and E. Adelson. Gelsight wedge: Measuring high-resolution 3d contact geometry with a compact robot finger. 06 2021.
- [17] I. H. Taylor, S. Dong, and A. Rodriguez. Gelslim 3.0: High-resolution measurement of shape, force and slip in a compact tactile-sensing finger. In 2022 International Conference on Robotics and Automation (ICRA), pages 10781–10787, 2022. doi:10.1109/ICRA46639.2022.9811832.
- [18] R. Wang, S. Wang, S. Du, E. Xiao, W. Yuan, and C. Feng. Real-time soft body 3d proprioception via deep vision-based sensing. IEEE Robotics and Automation Letters, 5(2):3382–3389, 2020.
- [19] R. B. Scharff, G. Fang, Y. Tian, J. Wu, J. M. Geraedts, and C. C. Wang. Sensing and reconstruction of 3-d deformation on pneumatic soft robots. IEEE/ASME Transactions on Mechatronics, 26(4):1877–1885, 2021.
- [20] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition, 2015. URL <https://arxiv.org/abs/1512.03385>.
- [21] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1–9, 2015. doi:10.1109/CVPR.2015.7298594.
- [22] Y. LeCun and Y. Bengio. Convolutional Networks for Images, Speech, and Time Series, page 255 – 258. MIT Press, Cambridge, MA, USA, 1998. ISBN 0262511029.
- [23] M. Hearst, S. Dumais, E. Osuna, J. Platt, and B. Scholkopf. Support vector machines. IEEE Intelligent Systems and their Applications, 13(4):18–28, 1998. doi:10.1109/5254.708428.
- [24] T. Cover and P. Hart. Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1):21–27, 1967. doi:10.1109/TIT.1967.1053964.