

# Soft, Compliant Tactile Robotic Manipulators

by

Sandra Q. Liu

Submitted to the Department of Mechanical Engineering  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MECHANICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2024

© 2024 Sandra Q. Liu. This work is licensed under a [CC BY-NC-ND 4.0](#) license.

The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Authored by: Sandra Q. Liu  
Department of Mechanical Engineering  
January 15, 2024

Certified by: Edward H. Adelson  
John and Dorothy Wilson Professor of Vision Science, Thesis Supervisor

Accepted by: Nicolas Hadjiconstantinou  
Professor of Mechanical Engineering  
Chairman, Department Committee on Graduate Theses



# Soft, Compliant Tactile Robotic Manipulators

by

Sandra Q. Liu

Submitted to the Department of Mechanical Engineering  
on January 15, 2024 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MECHANICAL ENGINEERING

## ABSTRACT

When we look to the future of soft robotics and manipulation, we begin to look towards sensory-rich and compliant grasping mechanisms. Not only do we want to capitalize on the significant advantages in safety and adaptability that soft robots have, we also want to incorporate high-resolution tactile sensors, which will allow soft robots to perform more tasks. One such system is the GelSight sensor, which is low-cost, effective, and high-resolution. However, the integration of these camera-based sensors into compliant manipulators is difficult due to the rigidity of the sensor backing. This thesis explores the design of multiple different compliant high-resolution tactile manipulators, along with some examples of their real-world uses. The first such design incorporates a simple camera-based tactile sensor into an exoskeleton-covered soft robot with vision-based proprioception. A later design integrates full camera-based tactile sensing capabilities into a flexible Fin Ray structure. Finally, the designs culminate in a novel soft-rigid human-inspired robotic hand with continuous tactile sensing which is capable of grasping heavier objects and safely interacting with humans. The incorporation of high-resolution tactile sensors into soft, compliant robots brings us closer to developing new manipulators that could someday match or exceed the capabilities of human hands.

Thesis supervisor: Edward H. Adelson

Title: John and Dorothy Wilson Professor of Vision Science



# Acknowledgments

I first would like to thank my amazing thesis committee, who has been so encouraging and inspirational to me.

Ted, of course, has been my advisor for all of my years at MIT and it has been an absolute wonder to work with him. Not only have I been greatly inspired by his basement tinkering, I am awed that even after all these years he is still excited about robotic hands and loves to prototype. He has given me a lot of encouragement to pursue my ideas over these few years and has pushed me to be more confident in myself. Before every talk I give, potentially to my labmates' chagrin, he will give me a pep talk telling me that my work is worth sharing, that I am an expert in my field, and that I should be proud of what I have done. It has really given me confidence and has pushed me to be willing to share my work with others and to be proud of the work I have done.

Similarly, Alberto has been very kind to me. As a new student starting out in a non-MechE lab, I was very intimidated by the process. No one in my lab was MechE and I felt very lost at the beginning with all the MechE requirements, but Alberto helped me out and immediately agreed to be my academic advisor. He gave me useful and insightful tips when I was starting out and has really helped welcome me to both the MechE and SidPac community.

Kait, to me, is an amazing powerhouse of a professor. From the beginning, I have been inspired by her drive, confidence, and just generally enthusiastic aura. She has been so down-to-earth and is always giving me wonderful and concise tips during my committee meetings, that have really helped inspire me as a person and a researcher.

I would also like to give a quick shout-out to my previous lab mates, which include a splattering of visiting students, post-docs, graduate students, and undergraduates. In particular, I want to give shoutouts to Yu, Filipe, and Radhen, who have really helped mentor me in various ways in my first few years here. I would also like to thank Satoshi, who has filled my first year memories with funny, random savage comments, and Achu who continuously insists that I am crazy and will tell people stories about his time together with us, including ones that I had actually forgotten. I would also like to thank Peiyu for being a bit of a rascal but fun nonetheless. Most of all, I want to say thank you to both Shawn and Leo.

Shawn is a year ahead of me and has always been so kind and so wonderful. He always has a smile on his face and his laughter is ridiculously contagious. He is also the person who introduced me to Lady M cakes and has a penchant for celebrating my achievements with said cakes. In the past year, I have really missed him randomly bursting into the song of his current obsession and scurrying around the lab.

Leo was the first UROP that I mentored. He is driven, capable, and extremely talented. I am very grateful to have had such a good experience with a UROP (which is something I have heard does not always happen), and I am extremely grateful that he is mischievous and plays along with jokes. If you ask him about decorative peppers, I am sure he will tell you.

Next, I would love to thank what I am calling the people in the robotics community who are not currently at MIT. Wenzhen, Rohit, and Jim, thank you guys so much for giving me the confidence to pursue what I am passionate about. You all have given me such wonderful opportunities to grow as a researcher and to share my research with other institutions. I am also grateful to have met so many people in the tactile and soft robotics communities at conferences and through collaborations. I am especially grateful to the soft roboticists, who have welcomed me into the community, especially after COVID took away a huge chunk of the networking that I had hoped to have in my first 3 years of graduate school. Thank you guys so much for everything and thank you for being so kind and encouraging. It has really meant a lot to me and has given me a lot of confidence to pursue the research I am passionate about.

And that of course leads me to my current lab members, who I would really like to take the time to thank. I am so grateful for Laurence and Mike, even though they have recently joined the lab, for their insights and absolutely hilarious comments. I also want to thank Maria for always having great tips for presentations and poster designs. I also would like to thank Elmer, our lab administrative assistant, who is always helping us with orders and reimbursements.

But the three current labmates I want to shout out the most are the ones who have somehow partaken in many of our lab shenanigans.

From his weird relationship with vegetables to his penchant of randomly forgetting important details to his excitement to take random trips, Alan has been an absolute hoot to be around. He is so funny and so talented, and I always go to him when I have any questions about programming, the panda robot, or machine learning. I am grateful for his larger than life personality and his extremely kind and generous nature. I think lab life would have certainly been less exciting without him around and scowling at the rest of us for being “creepy” and “suspicious.”

Yuxiang is essentially the meme of the lab who is suspiciously into songs about chickens. He is always somehow at the center of whatever weird thing is going on and if he has any sort of catchphrase, somehow everyone else in the lab copies him. This is probably why almost everyone in the lab mutters “useless” and “trash” when things are not going our way. But beyond that, he has been a fun workout buddy, an amazing collaborator, and the savage friend that keeps people laughing.

Last but not least among current labmates is Megha. I am not sure how to even begin to start talking about how much my life has improved after she joined our lab. She is so sassy, so fun, and so supportive. I am really going to miss not being in the same workspace as her and being able to gossip about everything and nothing, having too many inside jokes that sometimes we even cannot keep track of, and having a staunch supporter who is always gunning for my success. I am so grateful to have her in my life and to have someone who laughs so viciously at my jokes even when they are definitely not that funny.

Next, I want to thank my friends on the GSC AC and I especially want to thank Ruoxuan,

who is always planning fun events even now, and Henry who loves crab legs and souvenirs as much as I do.

I also want to thank my friends from the SP community, who have supported me so much throughout the years and taught me what it is like to be in a flash mob, to plan parties for way too many people and how to make strange symbols with my hands that definitely do not look like pineapples. I am so grateful for this community and in particular, I am thankful that it introduced me to Arko and Mo, who are almost always involved with various prank cakes. Arko was actually one of the first people I met during the MechE pub crawl and we have somehow been friends ever since. Mo is the person I go to when I want crazy ideas done because we do not know how to stop each other. As Mo says, whenever we come up with a dumb but stupid plan, I am half serious and he is half joking.

Another community I would like to thank is the MIT MechE and CSAIL one, which has been a huge part of my life. I would like to first thank Tyler, who really helped ease me into my first year at MIT and made sure that I was always sane and healthy even though we were literally dying while taking classes and were actually terrible at working on homework together. I am also thankful to Steven who is an adroit and fun storyteller and listener.

Another group of people I want to thank are the ones I became friends with during the pandemic, like Stephan, Simo, Kristan, Alan, and Kaitlyn. In particular, Alan and Kaitlyn are some of the best things that came out of quals. Alan took the three exact same quals as me and as a result they grouped us together when they were scheduling the quals. Having someone who knew exactly what I had just gone through was so helpful and grounding and I also enjoy the various times Alan and I have stopped by each other's labs to just talk.

Laura has also been a joy to hang around - she is so fun to talk to and so supportive. I have really enjoyed getting meals with her, cooking, drinking tea, and talking about books. I always have fun playing board games with her, even though I am almost always destroyed by her. Emily is a friend from Caltech who has continued to be an amazing friend at MIT. I adore her so much and I am always glad to hang out with her and eat desserts or dumplings and talk about slime reincarnation stories.

And of course, I would not be able to talk about being an MIT MechE graduate student without mentioning or talking about Rabab. Rabab has been a part of my graduate school life since the beginning. She is determined, smart, talented, and one of the warmest people I have ever met. I am always inspired by her kindness and I am so happy that she has been such a big part of my MIT life. We have done so much together, from picking random grapes at midnight, going to see water fire, going with Emily to see Salem, inadvertently talking about moon cycles, making styrofoam snow, doing yoga nights, conquering boda borg with Jihad, Gene, and Emily, the list could go on forever. I am so grateful that I got to meet her, even though we had trouble getting to hang out until 2.151.

I also want to thank all of my Caltech friends and friends from high school and before for all of their love, help, and support. I think being surrounded by everyone there really helped me grow and learn to be more confident. I am so grateful for all the fun memories, including the ones that I have pictures of but have no idea what was going on at the time - probably due to sleep deprivation.

I especially want to shout out the following people: Karim, who is my Avalon brother, which means that we betrayed each other on a weekly basis, Caleb, who is one of my favorite frosh from when I was a senior even though he hammocked my door shut, Patrick, Anne

Davis, and Tammer, who were very kind to me in my frosh year and taught me to stand up for myself and my feelings, Miguel, who is so fun and always down to try new food with me or sucker punch me when my football team is losing, David, who was my stand partner through most of college and who tricked me into taking an intense music history class but has great taste in mangas, Gyub, who is one of the cutest people I know and who has cheered me up greatly whenever I feel down, and Starry and Melissa, who both love gossiping as much as I do and always somehow ended up doing sort of stupid things with me probably because we were sleep deprived from talking too much. All of them have been such amazing friends to me and have taught me so much.

And last but not least, I would like to thank Jerry, who has been by my side for all of these years through Caltech and MIT, who has always supported me even when I felt like I did not deserve to be supported, who cuts watermelon like a pro, and is always encouraging me to challenge myself and to grow as a person.

My final shout-out goes to my family, who is absolutely precious to me. They have played a large role in shaping me to be who I am today. My grandparents have watched me grow up and helped raise me. They encouraged my creativity and taught me many skills and learning strategies.

My siblings, Alissa and Albert, are so fun to be around, whether it is playing game cube video games, throwing sticky weed at each other, or attempting to prank call each other or hiding behind bushes to spook each other. They are some of my biggest supporters and I am always excited to go home to hang out with them and crush them at Super Smash Bros.

My dad and mom have taught me a lot about overcoming adversity. They have always encouraged me to pursue my dreams and because of them, I was able to grow up without the thought that what I was doing was not okay, so much so that I somehow pigheadedly did not realize that in many of the activities I was in, I was the only female participant. They have taught me kindness, perseverance, and have given me so much support. In particular, my mom has been my biggest supporter throughout my entire life. She has always given me wonderful advice, listened to me, and comforted me when I was upset. Knowing that she is proud of me has allowed me to grow so much and I know that who I am today is because of my mom. I am so grateful to her beyond words and I am always inspired by her and her willingness to go the extra mile.



# Contents

<b>Title page</b>	<b>1</b>
<b>Abstract</b>	<b>3</b>
<b>Acknowledgments</b>	<b>5</b>
<b>List of Figures</b>	<b>11</b>
<b>List of Tables</b>	<b>17</b>
<b>1 Introduction</b>	<b>19</b>
<b>2 Background and Related Work</b>	<b>21</b>
2.1 Robotic Grippers . . . . .	21
2.1.1 Soft Robotic Grippers . . . . .	21
2.2 Tactile Sensors in Soft Robotic Grippers . . . . .	22
2.3 Camera-based Tactile Sensors . . . . .	22
2.3.1 GelSight-inspired Tactile Sensors . . . . .	23
<b>3 GelFlex</b>	<b>26</b>
3.1 Methods . . . . .	26
3.1.1 Hardware . . . . .	26
3.1.2 Software . . . . .	29
3.1.3 Experimental Setup . . . . .	33
3.2 Results . . . . .	35
3.2.1 Proprioception . . . . .	35
3.2.2 Object and Size Classification . . . . .	35
3.3 Discussion . . . . .	35
<b>4 GelSight Fin Ray</b>	<b>38</b>
4.1 Methods . . . . .	39
4.1.1 Sensor Design . . . . .	39
4.1.2 Software . . . . .	42
4.2 Experiment . . . . .	44
4.2.1 Experimental Setup . . . . .	44
4.2.2 Results . . . . .	45

4.3	Discussion . . . . .	47
<b>5</b>	<b>GelSight Baby Fin Ray</b>	<b>49</b>
5.1	Methods . . . . .	49
5.1.1	Hardware . . . . .	50
5.1.2	Software . . . . .	55
5.2	Analysis . . . . .	55
5.2.1	Simulation . . . . .	55
5.2.2	Fluorescent Silicone Paint . . . . .	57
5.3	Experiment . . . . .	60
5.3.1	Nut Classification . . . . .	60
5.3.2	Results . . . . .	61
5.4	Discussion . . . . .	61
<b>6</b>	<b>GelSight EndoFlex</b>	<b>63</b>
6.1	Methods . . . . .	65
6.1.1	Hardware . . . . .	65
6.1.2	Software . . . . .	67
6.2	Experiment . . . . .	69
6.2.1	Results . . . . .	70
6.3	Discussion . . . . .	71
<b>7</b>	<b>Conclusion and Discussion</b>	<b>73</b>
<b>8</b>	<b>Addendum: Previous GelSight Fin Ray Designs</b>	<b>75</b>
8.1	Optically Clear Fin Ray . . . . .	75
8.2	Early Hollow Fin Ray Prototypes . . . . .	75
<b>A</b>	<b>Tables</b>	<b>80</b>
<b>B</b>	<b>Figures</b>	<b>82</b>
	<b>References</b>	<b>86</b>

# List of Figures

2.1	A simplistic display of the aluminum cornflakes that are typically used in GelSight-inspired sensors. Flatter pieces will “perfectly” reflect light while more globular and rounder pieces will scatter the light. As a result, a combination of these pieces will allow the surface to capture more details in the tactile area and simultaneously allow smoother color gradients. . . . .	23
2.2	A simple visualization of a GelSight sensor pad as viewed by the camera. The red and green lights come from the same axes but from opposite directions, while the blue light comes from the bottom. If a spherical contact is pressed into the tactile surface, which is portrayed in grey, there should be a blending of color gradients which will create a rainbow circular contact patch. . . . .	24
2.3	Given a camera looking at a gel with markers on its surface, the camera can track the motion of the markers that are made as a spherical object is dragged along the surface of the gel. . . . .	24
3.1	<b>Left:</b> The soft finger robotic gripper, which is made of 2 exoskeleton-covered fingers, is grasping a metal mesh cup. Each finger is embedded with 2 intraphalangeal cameras. <b>Right:</b> The corresponding image from one of the cameras. Yellow dots on the side are designed for soft robot proprioception and the silver layer in the center is designed for tactile sensing. It can observe the mesh cup criss-cross patterns [13]. . . . .	27
3.2	The manufacturing process for the GelFlex finger gripper [13]. . . . .	28
3.3	Absolute joint angle naming convention for proprioception of the GelFlex finger [13]. . . . .	30
3.4	<b>Top:</b> Architecture of the proprioception neural network where $Ck$ represents a Convolution-BatchNorm-ReLU layer with $k$ filters. <b>Bottom:</b> Actual GelFlex finger with the intraphalangeal camera view. This image then gets fed into the CNN and the output joint angles are input into a RViz simulation of the finger [13]. . . . .	31
3.5	<b>First Row:</b> From left to right: a golf ball and a key. Both of these objects were brought into contact with the finger. <b>Second Row:</b> Their corresponding tactile imprints as viewed from the center intraphalangeal camera [13]. For more tactile images, including the difference images, see Appendix Fig. B.2. . . . .	32
3.6	The architecture of the MLP with Neural Incorporator used for size classification of the object set [13]. . . . .	33

3.7	Experimental setup for collecting data and testing. The soft robot gripper is attached to the end of the UR5 arm to keep it stationary. Four OptiTrack cameras track the position of the markers adhered to the finger. Testing is performed on the eight objects in the object set [13]. . . . .	34
3.8	A visual comparison between the actual finger and the simulated results as shown by RViz. The intraphalangeal camera view of the GelFlex is input into the proprioception neural net shown in Fig. 3.4 and the resulting predicted joint angles are output into RViz [13]. . . . .	36
4.1	In a), the GelSight Fin Ray gripper is holding a ball glass mason jar with its tactile sensing region along the indented letters "MASON" while b) denotes the 90° counterclockwise-rotated raw tactile image data where the top of the image generally denotes the tip of the finger and c) is its corresponding uncalibrated tactile reconstruction, where you can see most of the "S" letter and parts of the "O" and "A" letters as well [14]. Another example of a smaller object being grasped is shown in the Appendix Fig. B.4. . . . .	39
4.2	Manufacturing process flow of the GelSight Fin Ray finger, which includes finger design, tactile sensing pad manufacturing, and illumination [14]. . . .	40
4.3	Algorithm flow for obtaining the reference point matrix and live reconstruction of the raw tactile image. The reference point matrix is precalculated using a video taken of the finger without any tactile interference. The matrix is sparse and contains only one nonempty element per column. Subsequent raw tactile images are then compared to the reference point matrix via the dot centers, a difference image is taken, and an uncalibrated reconstruction image is calculated. Pictured in the raw tactile image is the head of a screw that was pressed into the sensing region at a slight angle [14]. . . . .	43
4.4	Objects used for the experimental setup. They include a mini screwdriver, plastic fruits, a Rubik's cube, a wine glass, a mason jar, and a squishy acrylic paint tube [14]. . . . .	44
4.5	Tactile sensing image reconstruction. The left most images show the objects (M4 heated insert, M2.5 screw) that were placed on the tactile sensing images, the middle images represent the raw tactile data, and the right most images display the reconstruction based on the raw data [14]. . . . .	46
4.6	Marker tracking with the GelSight Fin Ray. The yellow arrows denote the tracked position from the reference markers in the found reference image to the markers in the current image. In both images, the fingers are gripping a smooth acrylic cylinder. The left image (a) denotes shear along the cylinder from an external force while the right image (b) displays torsional external force being applied on the cylinder [14]. . . . .	47

4.7	Successful implementation of wine glass reorientation. A wine glass is handed to the gripper at an angle. Using the raw tactile image, the orientation of the wine glass stem is found through the reconstruction algorithm. Using this angle, the UR5 arm reorients the gripper so that it is holding the wine glass right side up. The arm then sets the wine glass down until the shear force from the contact goes above a certain threshold and causes the gripper to release the wine glass on the table without it tipping over. Note that the shear force detected by the marker tracking algorithm is along the major direction of the stem, indicating that the wine glass is upright [14]. . . . .	48
5.1	The Baby Fin Ray grabbing a walnut (left) and the corresponding cropped raw image of the mirror that displays the tri-colored tactile sensing region and the indentation of the walnut shell (right) [15]. . . . .	50
5.2	A side-by-side comparison of the original GelSight Fin Ray and the Baby Fin Ray [15]. . . . .	51
5.3	An exploded view of the Baby Fin Ray assembly. Not pictured are the camera, LEDs, and the diffuser [15]. . . . .	52
5.4	The 2D representation of a 100° field-of-view camera viewing the tactile sensing region through a mirror and its corresponding virtual camera. As the ray tracing shows, the camera is able to see the entire tactile sensing region while sitting at the base of the finger, where its rigidity will not impede with the deformable Fin Ray structure [15]. . . . .	54
5.5	FEM simulation results of the improved design and the “original” design, which is a smaller version of the previous GelSight Fin Ray for the sake of comparison. The indentations happen at the middle finger pad (lower location) and the fingertip a) Displacement fields of cylinder indentation; b) Displacement fields of cuboid indentation; c) Force-displacement curves of the improved and “original” fingers with cylinder indenter; d) Force-displacement curves of the improved and “original” fingers with cuboid indenter [15]. . . . .	56
5.6	Fluorescent silicone paint dogbone samples used for tensile testing on the Instron meeting. From left to right, we have our pink Silc Pig, pink synthesized paint, green Silc Pig, and green synthesized paint samples [15]. . . . .	57
5.7	Stress-strain curves of the fluorescent silicone paint, where the acrylic paints represent the synthesized silicone adhesive based paints. The graphs have been cut off at the point of breakage of the dogbone tensile testing pieces [15].	59
5.8	The architecture used for training our nut classification task. We take the raw image from our tactile sensor, unwarped the mirror image, before feeding it into a Resnet-50 architecture with four classes: Almond, Brazil Nut, Pecan, and Walnut [15]. . . . .	60

6.1	<b>Top</b> A CAD model of our GelSight EndoFlex gripper with some of the parts labeled. <b>Bottom</b> The GelSight EndoFlex is securely grasping a Rubik’s cube and the corresponding processed difference images of four of the six sensing regions are displayed. Of note is that the bottom two sensor images are showing continuous sensing along the side and corner of the cube, while the top two sensor images are showing one image each from the other two fingers [16]. . . . .	64
6.2	A close-up view of an EndoFlex finger with an exploded view. Each finger operates independently with one degree of freedom and can be quickly replaced if damaged [16]. A diagram of the continuous sensing region is shown in the Appendix Fig. B.5. . . . .	65
6.3	The manufacturing process for the EndoFlex sensor including assembly of electronics and casting of silicone [16]. . . . .	66
6.4	A close up image of the reticulated wrinkle surface of the GelSight EndoFlex sensor. The width of one of the wrinkles is approximately 0.4 mm wide and was only created when we first sprayed the paint on the mold surface before casting silicone inside [16]. . . . .	68
6.5	From left to right, we have raw sensor images of a 3.75 mm ball bearing array and a M2 screw, followed by their difference images from a reference image (no tactile contact), and the corresponding uncalibrated depth image [16]. . . . .	68
6.6	Neural net architecture for our single grasp classification. Once the object has been grasped, the six images are stitched together in a 2x3 array, thrown into our Resnet architecture and classified into a toy cup, an orange, or a Rubik’s cube [16]. . . . .	69
8.1	Optically clear Fin Ray structure with an embedded camera inside. On the right, is the corresponding image from the embedded camera with the markers on the side. . . . .	76
8.2	Optically clear fin ray structure deforming with a finger sliding along the side. Because of the silicone and the thin flexible plastic, the Fin Ray finger deflected backwards when an object was pressed against it. . . . .	76
8.3	The Onyx-based Fin Ray. The struts were thinned down and made narrower to help the structure deform, but the resulting structure still had signs of plastic deformation after a single grasp, which is shown in the curved struts and “tactile” section. . . . .	77
8.4	The SLA printed hollow Fin Ray. a) shows the 3D printed Fin Ray right off of the printer and b) shows the resulting post-processed Fin Ray structure. . . . .	78
8.5	The multimaterial Fin Ray structure with a PET backing and TPU struts (a). (b) displays the manufacturing process: we start with the 3D printed structures taken off of the printers with the support material structure still attached. After removing the support material, we note the close up of the “hair” leftovers from the support structure (b2), and then we clamp the two pieces together with epoxy. . . . .	79

B.1	Close up of a GelFlex finger with OptiTrack markers on it. Each link was represented by three separate OptiTrack markers to create a rigid body. . . .	82
B.2	Tactile images from the GelFlex. The first row shows the objects that are pressed into the finger, while the second row shows the raw image taken from the camera and the third row shows the corresponding difference image. From left to right, we have a cardboard box edge, painter’s tape, a finger, and a screw head. . . . .	83
B.3	Tactile neural network architecture. It takes a difference image of the GelFlex grasping an object as an input and outputs the label "Box" or "Cylinder," based on what the finger is in contact with. C5x5 represents a convolution with kernel size 5x5 while MP2x2 represents Max Pooling with a 2x2 kernel size. . . . .	83
B.4	The GelSight Fin Ray grasping a plastic strawberry from the YCB dataset [67]. In the bottom right corner, there is the raw image on the left and the corresponding reconstruction image on the right. . . . .	84
B.5	A side view of the GelSight EndoFlex with the camera field of views shown. When the finger is bent, we have continuous tactile sensing along the front and sides of each finger. . . . .	85





# List of Tables

5.1	Fluorescent Paint Comparisons [15] . . . . .	58
A.1	RMSE values between 82.55, 83.55, and 87.70 mm diameter dot pose data and "Testing" curves (all values in mm). The RMSE values are calculated between the dot locations estimated from the image (of these three differently sized objects) and the three testing diameters (which are all close in value to the actual diameter of the test object). . . . .	80
A.2	Confusion matrix for size classification of all 80 "bar stock" trials. There were 40 trials each for cylinders and squares. All values are percentages, unless otherwise noted. . . . .	80
A.3	Confusion matrix for size classification of the square "bar stock." All values are percentages, unless otherwise noted. The only sizes that were misclassified were the two larger sized boxes. This could have been caused by an unstable grip on the boxes by the soft robotic gripper. . . . .	81
A.4	Confusion matrix for size classification of the cylinder "bar stock." All values are percentages, unless otherwise noted. The size classification neural net appeared to perform better on cylinders than boxes. . . . .	81



# Chapter 1

## Introduction

Robotic grippers are extremely capable in industrial applications but have quite a few limitations. Many of them can only perform one task, cost exorbitant amounts, or require a very specific tool to be able to use. An ideal generalizable gripper can grasp many different objects and be used at-home or to assist the elderly as they age with dignity. The field of robotic hands is still a long way from this goal, but is getting closer through incremental improvements.

One of the ways to do this is to utilize compliance in soft robots. Soft robots are able to grasp a multitude of objects, just as humans can. They can pick up delicate objects, such as fruit, small objects such as pens, and larger objects such as mugs [1]. Leveraging this ability will help robots get a step closer to achieving generalizability in the wild and be useful for many different scenarios.

There are also severe limitations for robots that are soft and squishy. They lack feedback that could make them viable for control as well as sensing that could allow them to perceive the world around them. Many existing robots use external vision sensors that allow for this feedback. For example, with external cameras that act analogously to human eyes, a robotic hand is able to manipulate pegs into holes [2], interact with tangled cables [3], and identify the location of ripe fruit in orchards [4]. In fact, recent advances in computer vision have allowed robots to perform even more complicated tasks in the real world, including seeing and segmenting glass objects [5]. These examples are only a small subset of the tasks that robotic hands can now perform thanks to the utilization of external sensors.

However, all of these tasks can be augmented with tactile sensors. For the manipulation of a peg into a hole, although vision and a fine control system show promising results, having the addition of force sensing allows the control system to make mistakes and be able to account for more precise manipulation [6]. Tangled cables are visible when they are not grasped, but once they are grasped, they are obscured by the gripper and thus, information about the cable is hard to discern [7]. Similarly, even with a wrist-mounted camera, it is sometimes hard for the camera to look through clutter or distinguish minute changes in objects. Finally, although cameras can see the location of ripe fruit, tactile sensors can determine through touch if the fruit is actually ripe [8], and can also provide the necessary feedback to prevent damage to softer, fragile fruit such as raspberries or tomatoes [9].

In summary, we want compliance that can be provided by soft robotics and we want tactile sensing incorporated into them. Unfortunately, existing tactile sensors are either

rigid with high-resolution or are compliant with lower resolution [10], [11]. We want both high-resolution and compliance, because having both will allow robotic hands to perform a larger variety of everyday tasks, similar to what human hands are able to do. Human hands are capable of grabbing many objects, from gripping fragile, deformable fruits to grasping heavier tools. Furthermore, our skin is capable of detecting many nuanced textures over its entire surface, which allows us to classify objects with a single grasp or determine material, hardness, or if an object is slipping out of our hands.

To give robotic hands analogous sensing abilities, we turn to camera-based tactile sensors, which can help to mitigate the drawbacks of external vision cameras. Placing internal cameras into robotic hands has the potential to negate the necessity of external vision since all the sensing is contained within the hand itself. Moreover, cameras are able to capture extremely detailed information over a large viewing area, making them useful for tactile sensors. In particular, we are inspired by the GelSight sensor, which is high-resolution, effective, and low-cost [12].

The GelSight sensor is characterized by a semi-reflective silicone membrane, which is illuminated by colored LEDs, and viewed by a camera internally. When an object is pressed into the semi-reflective membrane, its silhouette and shape are illuminated by the different LEDs. The resulting image can then be processed using computer vision techniques to determine the properties of the object being pressed into the surface. If markers are added, the sensor can also track surface forces applied to the sensor via the object in contact.

Although GelSight sensors are useful for a variety of contact tasks, they are inherently “rigid” for ease of computer of vision processing and to house the electronics. As such, this thesis aims to address the seemingly disparate fields of rigid high-resolution visuotactile sensors and soft, compliant robotics.

We first design an exoskeleton-covered soft silicone finger that uses its backing to incorporate LEDs and cameras to determine the usefulness of incorporating camera-based tactile sensing and proprioception in soft fingers [13]. Next, we determine hardware and software methods to incorporate a GelSight-inspired sensor capable of 3D reconstruction and shear force tracking with a soft, compliant Fin Ray, which is capable of grasping many different objects. We demonstrate its usefulness in performing a simple manipulation task not easily done with pure external vision methods [14]. Then, we show that the design can be made more robust and compact with other design techniques, which allow us to potentially incorporate such rigid high-resolution tactile sensors in even more soft robots [15]. Finally, we show that all of these techniques can be used in a human-inspired hand design, which has continuous tactile sensing along its entire surface and can grasp heavy objects as well as be safe and compliant [16]. Its large sensing range also enables it to classify an object with a single grasp.

Overall, this thesis shows strategies for incorporating rigid camera-based tactile sensors into soft robots, combining the compliance and high-resolution tactile sensing necessary for robotic hands to progress even further towards human-level hands.

# Chapter 2

## Background and Related Work

### 2.1 Robotic Grippers

Robotic hands, or more specifically, prosthetic hands made from metal, have been in existence since around 75 AD when a physician chose to make a metal hand for a wounded soldier. In modern times, on Amazon, a consumer has a large selection of “picker uppers” targeted towards the elderly who wish to pick up trash from the floor without bending over or squatting. Analogously, some people choose to use a parallel jaw gripper equivalent, chopsticks, to eat their food without getting their own hands dirty.

Different types of grippers exist, and similarly, there has been extensive research to develop various types of prosthetics and robotic grippers. Currently, many industrial grippers utilize a parallel-jaw paradigm or have some form of suction. Research has also been done on more anthropomorphic grippers, with the intent to mimic human-hand capability in robotics. Although many robotic hands and grippers used to be designed with rigid materials, there has been increasing interest in the development of soft and compliant systems [1].

In particular, the compliance in soft robotics allows them to be useful in many applications, such as home-care, agriculture, and search and rescue [17]–[19]. Even though soft robots are more complex to model and do not have the well-defined kinematics of their rigid counterparts, their softness allows them to interact with a wide range of fragile materials without needing precise control [1], [20]. As a result of these advantages, many researchers have worked to develop different types of soft, compliant grippers.

#### 2.1.1 Soft Robotic Grippers

Compliant grippers exist in many forms and with various types of actuation. There are grippers that are passive and comply to forces imparted on them [21], soft fingers that are actuated pneumatically [20], [22], thermal-responsive materials such as shape memory alloys [23], and tendon-driven soft bodies [24], [25]. All of these grippers and robotic hands take advantage of such compliant and soft materials to better envelop an object that they are grasping, create a higher surface area contact, and thus, have a better grip on the held object.

The two types of actuations this thesis focuses on are tendon and passive actuation. Both of them are fast reacting, need less total power consumption, and require less onboard equipment compared to many other actuation methods for soft robots. Pneumatic and fluid

actuators, for instance, must have compressors, valves, and other apparatuses, which quickly scale up in size and become too bulky. Meanwhile, other actuation strategies like shape memory alloys require tens of minutes to fully actuate and are harder to control.

Tendon actuation generally operates on the same time scale as the motor it is controlled by, takes up a smaller amount of area, and is relatively simple to implement in a robotic finger design. Passive compliance, such as the one in the Fin Ray design, requires only a parallel jaw gripper and no other components aside from the soft structure itself. Both of these actuation methods also integrate well with soft, compliant robots and help robotic grippers to grasp objects with more stability and robustness.

However, their compliance and softness make them more difficult to interface with standard, rigid sensors. Many soft robots do not have any rigid structure to which such sensors can subsist. As a result, such “rigid” sensors can predominantly be used on rigid robots and other softer, more compliant sensors have been developed for soft robots.

## 2.2 Tactile Sensors in Soft Robotic Grippers

Although many different soft material sensors exist [26], one of the most commonly used tactile sensors for soft robotic grippers are strain sensors. Strain sensors are flexible, can provide some force feedback, and are relatively cheap to obtain [27]. Other soft robots utilize optical waveguide-based sensors, which work by detecting changes in the light filtering into an optical sensor and corresponding these changes to the finger position and force imparted [28]. Still others incorporate resistive or capacitive sensors. Resistive sensors are able to convert gel or soft material displacements into a corresponding electric signal while capacitive sensors can detect changes in electric fields due to tactile contact.

Although all of these types of sensors are quite useful and provide accurate force information, including magnitude and direction, these sensors are unable to display extremely high-resolution accuracy and are unable to tell distinct surface patterns or more finely detailed features.

For some dexterous manipulation tasks, being able to tell the 3D force vector is enough. Xu and Lin et al. and Khandate et al have performed complicated object reorientation tasks using a matrix of force sensors [29], [30]. However, for object classification or exploration, using the sensors to capture more detailed orientation can be helpful. To be able to capture both this detailed tactile surface information and surface forces while interacting with objects, we turn to camera-based tactile sensors.

## 2.3 Camera-based Tactile Sensors

The philosophy behind camera-based tactile sensors is as follows: (1) take a soft, sometimes elastomeric surface, (2) illuminate the surface with some lighting, and (3) capture the surface structure using a camera sensor. Then, using the images captured by a camera, the sensor will be able to discern different properties of the object that is in contact with the elastomeric surface. These properties range from material hardness [8], surface force information [31], [32], and material properties [33]. In general, these sensors are cheap to implement and

effective at sensing high-resolution details in its surface. The simplicity of the camera-based tactile sensor design idea also allows such sensors to exist in various forms and with different lighting configurations [34]–[36].

One variant of such a camera-based tactile sensor is the GelSight sensor [37], which has greatly inspired a family of sensors in recent years.

### 2.3.1 GelSight-inspired Tactile Sensors

Like other camera-based tactile sensors, the GelSight sensor also utilizes a camera, soft elastomer, and an illumination scheme. Two key components of a GelSight sensor are its tactile skin, which is composed of gray paint and aluminum flakes, and its RGB (red, green, blue) color scheme. Depending on whether the aluminum flakes are more globular or flat, the flakes will display different light reflective properties (Fig. 2.1). More flat pieces will exhibit “perfect” reflectance while globular pieces will scatter the light. In general, the aluminum flakes that are used are of the “cornflake” variety, whose shape is inspired by an eponymous American cereal brand. Cornflakes have a relatively even distribution of both flat and globular pieces due to their even, yet crinkled shapes. In aluminum form and mixed into silicone paint, they are able to help the sensor capture intricate tactile details and create smooth color gradients.

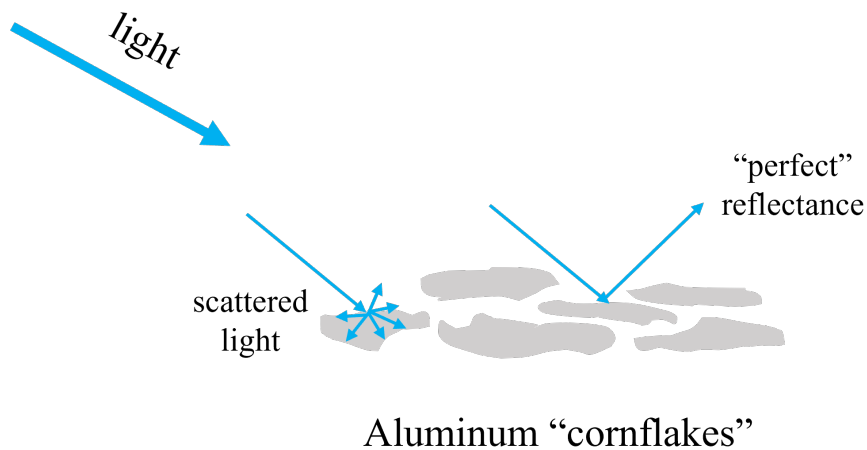


Figure 2.1: A simplistic display of the aluminum cornflakes that are typically used in GelSight-inspired sensors. Flatter pieces will “perfectly” reflect light while more globular and rounder pieces will scatter the light. As a result, a combination of these pieces will allow the surface to capture more details in the tactile area and simultaneously allow smoother color gradients.

The tri-color scheme is used to take advantage of the RGB color channels in a single image. By having these three colors illuminate the elastomer surface from different directions, we can get information on the gradient distribution in at least two directions using only a single image (Fig. 2.2).

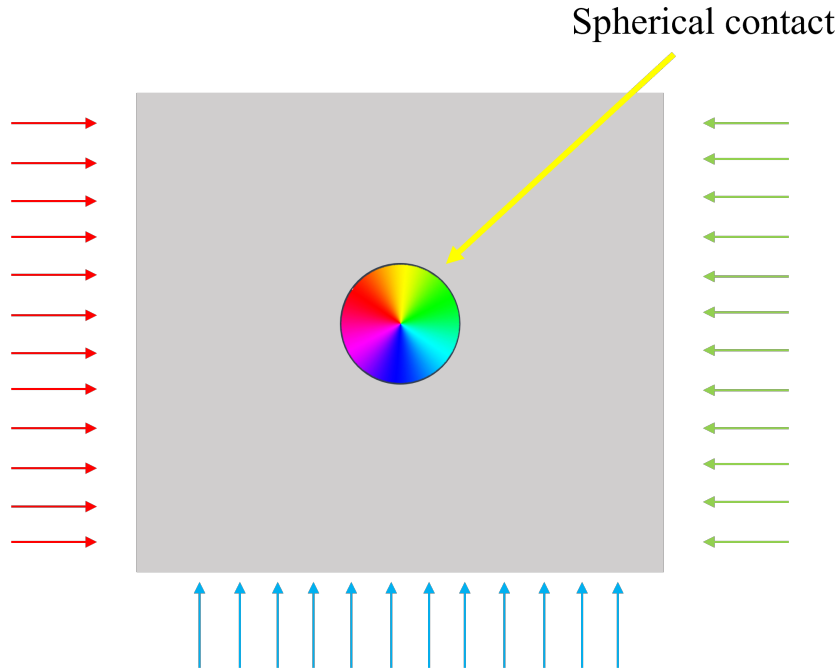


Figure 2.2: A simple visualization of a GelSight sensor pad as viewed by the camera. The red and green lights come from the same axes but from opposite directions, while the blue light comes from the bottom. If a spherical contact is pressed into the tactile surface, which is portrayed in grey, there should be a blending of color gradients which will create a rainbow circular contact patch.

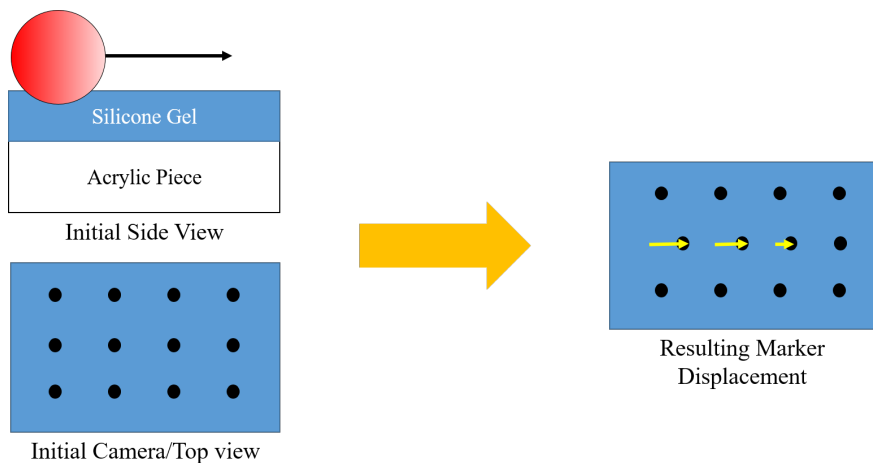


Figure 2.3: Given a camera looking at a gel with markers on its surface, the camera can track the motion of the markers that are made as a spherical object is dragged along the surface of the gel.

Furthermore, if we add markers to the surface of the elastomeric gel, we can track forces on the surface of the finger such as incipient slip or torsional shear (Fig. 2.3).

Many such GelSight-inspired sensors exist, from round, 360 degree ones to curved ones



to wedge-shaped ones [31], [38], [39]. However, all of these sensors, apart from the soft, elastomeric gel that encases them, are made from rigid structures. These rigid structures make them difficult to integrate with soft or compliant robots without compromising the softness of the robot. Nonetheless, to give soft robots high-resolution tactile sensing so they can have some feedback from the world, we need to be able to develop novel designs that will allow us to integrate the two seemingly disparate ideas of using a soft, compliant robot and a rigid sensor. Subsequent chapters will describe various solutions to this integration challenge.

# Chapter 3

## GelFlex<sup>\*</sup>

To gauge the usefulness of incorporating a camera, we design a soft finger structure that can hold a camera sensor. Previous work [40] allowed us to determine that it was necessary for our design to have features for the camera to use as a datum to help “track” proprioception. However, these features should not overtly obstruct the tactile sensing surface and cause occlusions on the surface. As such, we develop an exoskeleton-covered soft finger. The soft finger is composed of a solid rubber block with features on its side for the camera to track, while the exoskeleton is able to house the cameras and LEDs.

We then developed a novel robotic gripper composed of two of these exoskeleton-covered soft fingers, each with 2 cameras embedded inside. The robotic gripper can employ the embedded vision sensors for simultaneous proprioception and tactile sensing and have an accurate method for estimating the shape of the finger. Furthermore, the gripper can perform grasping tasks using the tendon-driven system and successfully differentiate between different shapes and sizes of bar stock, which is a classification problem useful for industrial and manufacturing robotic applications.

A picture of the gripper holding a mesh cup and the corresponding tactile image from one of the intraphalangeal cameras is shown in Fig. 3.1.

## 3.1 Methods

### 3.1.1 Hardware

The manufacturing of the robotic gripper includes casting processing, painting processing, finger assembly, and gripper assembly, and the overall process flow for manufacturing of the soft gripper is shown in Fig. 3.2.

**Casting processing** We first make the finger mold using a 3D printer (Onyx One 3D Printer, Markforged, Inc.) {C1}. The finger surfaces must be smooth since both the tactile and shape information are acquired via cameras observing these surfaces from inside. To improve the smoothness, we paint the inside of the 3D-printed finger mold with ultraviolet (UV) resin (200g Crystal Clear Ultraviolet Curing Epoxy Resin, LIMINO) {C2}. After exposing the

---

<sup>\*</sup>Most of this chapter is copied from the paper Exoskeleton-covered soft finger with vision-based proprioception and tactile sensing [13].

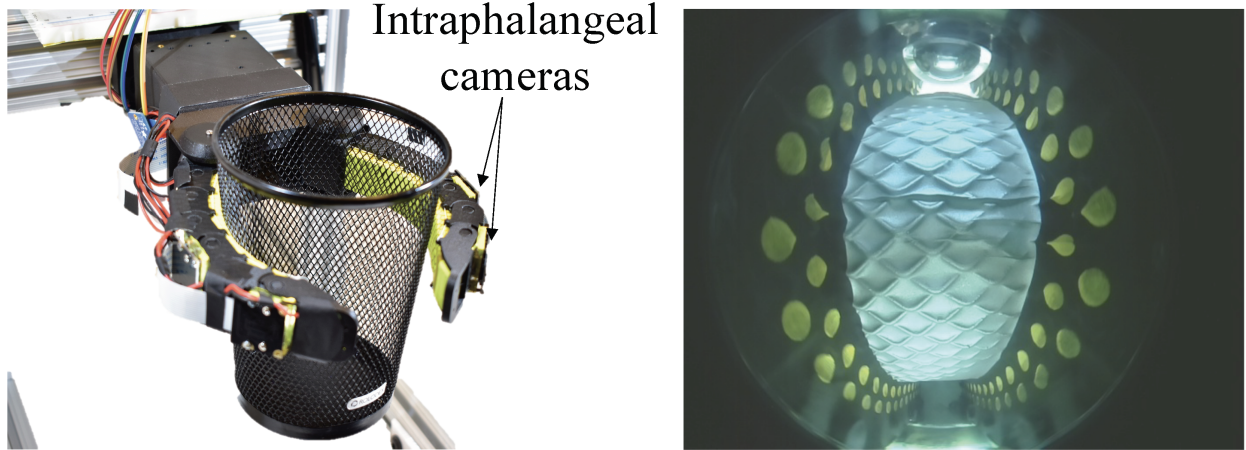


Figure 3.1: **Left:** The soft finger robotic gripper, which is made of 2 exoskeleton-covered fingers, is grasping a metal mesh cup. Each finger is embedded with 2 intraphalangeal cameras. **Right:** The corresponding image from one of the cameras. Yellow dots on the side are designed for soft robot proprioception and the silver layer in the center is designed for tactile sensing. It can observe the mesh cup criss-cross patterns [13].

resin to UV lights for 1 hour, the top of the mold becomes smooth. However, the UV resin contains chemicals which prevent the curing of silicone. To counter this, we coat the UV resin with a layer of an inhibit agent (Inhibit X, Smooth-On, Inc.), and then spray a layer of mold release (Ease Release<sup>TM</sup> 200, Smooth-On, Inc.) over it {C3}. We then embed clear acrylic hemisphere lenses (3/4 inch Hemispheres, California Quality Plastics) in the finger to allow for the integration of fish eye camera lenses {C4}. These hemispheres provide wide view angles for the camera and hide the camera within the finger. Since we use cameras for both proprioception and tactile sensing, the soft finger body must be optically transparent. To give our finger this property, we made it out of platinum-catalyzed translucent silicone (XP-565, Silicones, Inc.) {C5}. Prior to adding the silicone, we apply primer (DOWSIL<sup>TM</sup> PR-1200 RTV Prime Coat, DOW) on the outside surface of the lens to bond it to the silicone. This procedure prevents the cured silicone from easily being ripped off during larger deformations of the finger. To improve the softness and resilience of the silicone finger, we add dilute agent (LC1550 Phenyl Trimethicone, Lotioncrafter). The ratio of the components of silicon XP-565 A and B, and phenyl trimethicone is 1 to 15 to 3. After uniformly mixing all the components in a plastic cup, we vacuum it (90062-A 3 CFM Vacuum Pump, Mastercool, Inc.) to get rid of bubbles. Finally, we pour the mixture into the 3D printed mold.

**Painting processing** We first de-mold the silicone from the 3D printed mold and clean any residue on the silicone using soap {P1}. We then use different colors of silicone inks (Print-On Silicone Ink, Raw Material Suppliers), which are scratch-resistant and durable, to create dot features for the camera to track. Silicone solvent (NOVOCS Gloss, Reynolds Advanced Materials) is added to the silicone ink to dilute the ink. The ratio we use is 1 to 10 to 30 (silicone ink catalyst to silicone ink to silicone solvent). We apply black ink on both sides of the silicone finger and cure it in an oven at 200°F for 30 minutes {P2}. Then, we use a laser cutter (PLS6.75, Universal Laser Systems, Inc.) to engrave circular patterns on both sides of

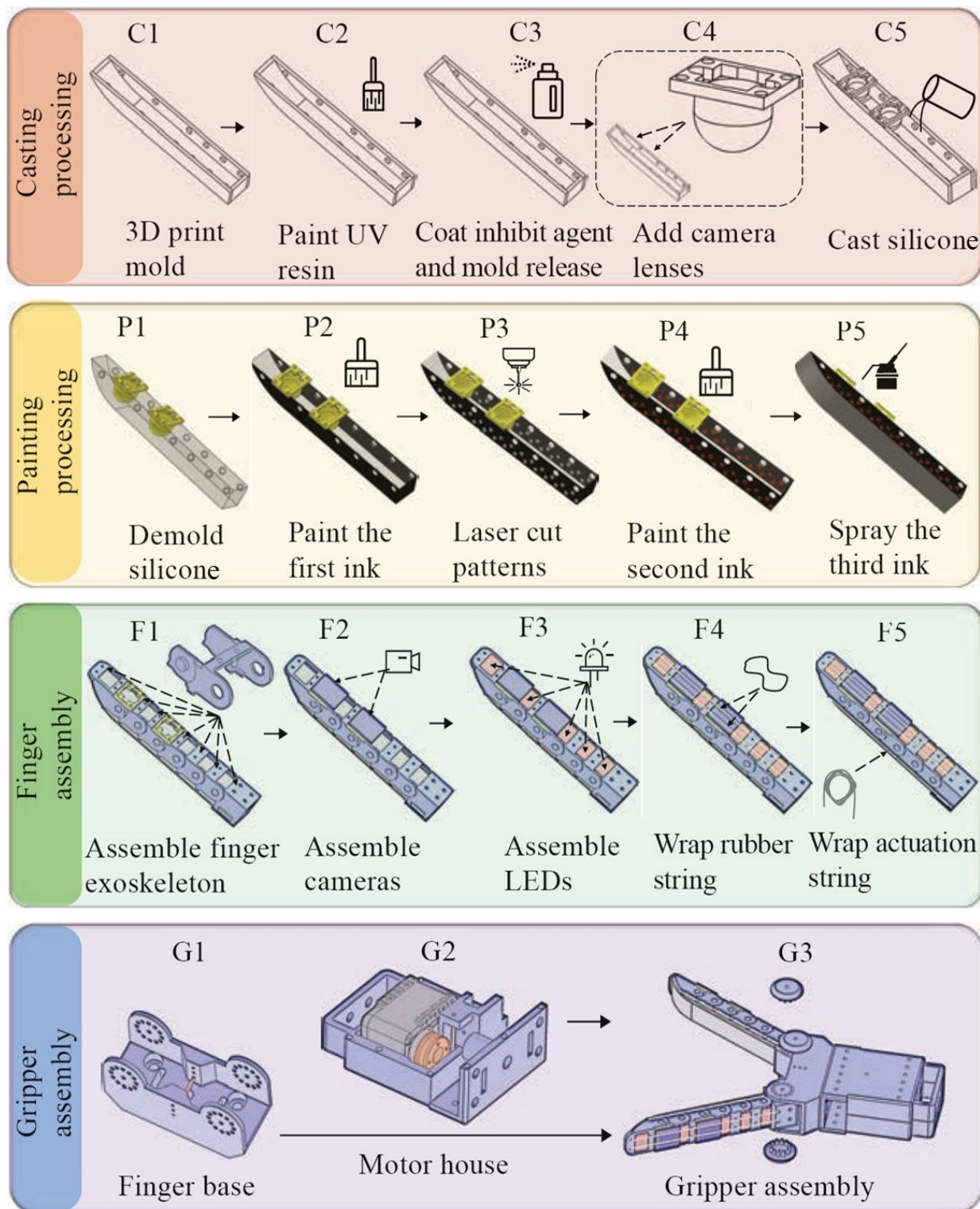


Figure 3.2: The manufacturing process for the GelFlex finger gripper [13].

the finger and remove the black paint in those areas {P3}. Next, we paint yellow silicone ink on both sides of the finger, creating yellow dots with a black background {P4}. These yellow

dots allow the camera to more easily capture finger deformation information. The next step is to spray a reflective ink at the bottom surface of the finger for tactile sensing {P5}. We opted to use semi-specular painting material (silver silicone ink) so that we could observe a higher-level of intricate details. Afterwards, we use an ultrasonic processor (VCX 750, Sonics & Materials, Inc.) to enhance the dissolution of the silicone ink into the NOVACS solvent. We use an airbrush (ZENY Pro 1/5 HP Airbrush Air Compressor Kit, ZENY) to paint the silver silicone ink on the bottom surface of the finger to ensure a smooth and uniform paint finish. Then we place the painted finger in an oven for 2 hours at 200°F before the finger is ready to use. Note that the painting order of the silicone ink can vary as long as one can create the desired features.

**Finger assembly** We design a highly underactuated exoskeleton that is comprised of 7 segments and 6 joints to actuate the soft finger. The exoskeleton is then integrated with the soft silicone trunk. We first make the finger exoskeleton with the 3D printer and assemble the finger segments, before pressing the soft silicone trunk into the exoskeleton {F1}. This design allows the finger to easily adapt to grasping various objects. The exoskeleton is designed with round pegs on the side to fit into the round slots on the side of the soft trunk. This prevents the soft trunk from easily detaching from the exoskeleton. Then, we add two cameras to the camera lenses that were fixed onto the silicone finger during the casting processing {F2}. Next, we add white LEDs, which provide internal illumination for the camera {F3}, and restrain them with rubber bands attached to the top of the exoskeleton {F4}. Finally, we add actuation strings through holes on each segment of the exoskeleton {F5}. One end of the string is tied to the last exoskeleton segment and the other end of the string is attached to a motor disk. This yields a highly underactuated exoskeleton-covered soft finger with a tendon-driven system.

Note, here the soft silicone trunk with its surface painting as well as the associated cameras and LEDs are referred to the *GelFlex*.

**Gripper assembly** The last step of the hardware process is to assemble the gripper. The gripper includes assembling 2 GelFlex fingers, a finger base {G1}, and a housing {G2} for a servo motor (DYNAMIXEL XM430-W210-T, ROBOTIS) and disk. The embedded cameras are controlled using Raspberry Pi boards. The orientation of the finger relative to the finger base is adjustable {G3}, allowing us to use it for a variety of tasks.

### 3.1.2 Software

To classify bar stock using tactile data and proprioception of the finger, a combination of three neural nets were used. One was for proprioception of the GelFlex finger, another was to use tactile sensing data to determine whether the grasped object was a cylinder or a box, and the final one combined the proprioception of the finger and tactile knowledge to classify the size of the object being grasped.

#### Proprioception

Proprioception of the finger was simplified to a 2D learning problem due to the mechanical constraints of the exoskeleton, and the assumption that the finger would not flex sideways

(z-direction) and would only bend in response to the tendon-driven system (x-, y-directions). Thus, only the 2D positions of each individual joint were considered for the neural network.

The first stationary joint was taken as the world reference frame origin and the absolute angle for each joint was calculated relative to the world reference frame. Each position of the finger is thereby defined by six numbers (Fig. 3.3).

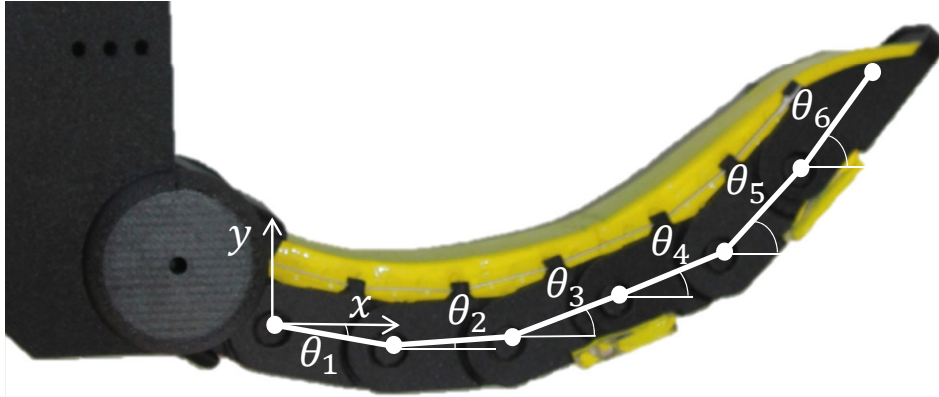


Figure 3.3: Absolute joint angle naming convention for proprioception of the GelFlex finger [13].

For the neural net, approximately 3,500 sets of images were collected (3,500 images each from the two embedded Raspberry Pi cameras) with their corresponding finger position absolute angles. The collected images include ones where the gripper is grasping various objects and no objects. These different images allow the neural net to be trained with many different parameters and possible finger configurations, and enable the neural net to ignore tactile information in images. The neural network used 80% of the images for training and 20% of the images for testing. Since the inputs of the neural net are 2D images, a fully convolutional neural network (CNN) [41] was implemented to process all of the data.

The neural net architecture for a single image is a sequential model of 4 convolutional layers with batch normalization and ReLU, which is followed by a single convolutional layer with sigmoid activation. The sigmoid activation then maps the extracted feature to the joint angle vector. This architecture is shown in Fig. 3.4.

To determine the effectiveness of utilizing one camera image versus using two camera images, two different neural networks were trained for proprioception. One used only the camera in the middle of the finger, while the other processed both of the camera images. The only difference between the two architectures was that for the network trained on two images, the two images were concatenated together, doubling the input channel number.

Because of the high variance in the training output angles, the training angles were normalized using their mean and standard deviation. They were appropriately re-scaled during the testing procedures. Zero mean Gaussian noise with small variance ( $1e-3$ ) was also added to these ground truth training angles to augment the data and simulate real-world perturbations that may not have been captured in the testing data. Similarly, to deal with image noise and potential unseen changes in lighting parameters, the input images were augmented with random changes in contrast and saturation.

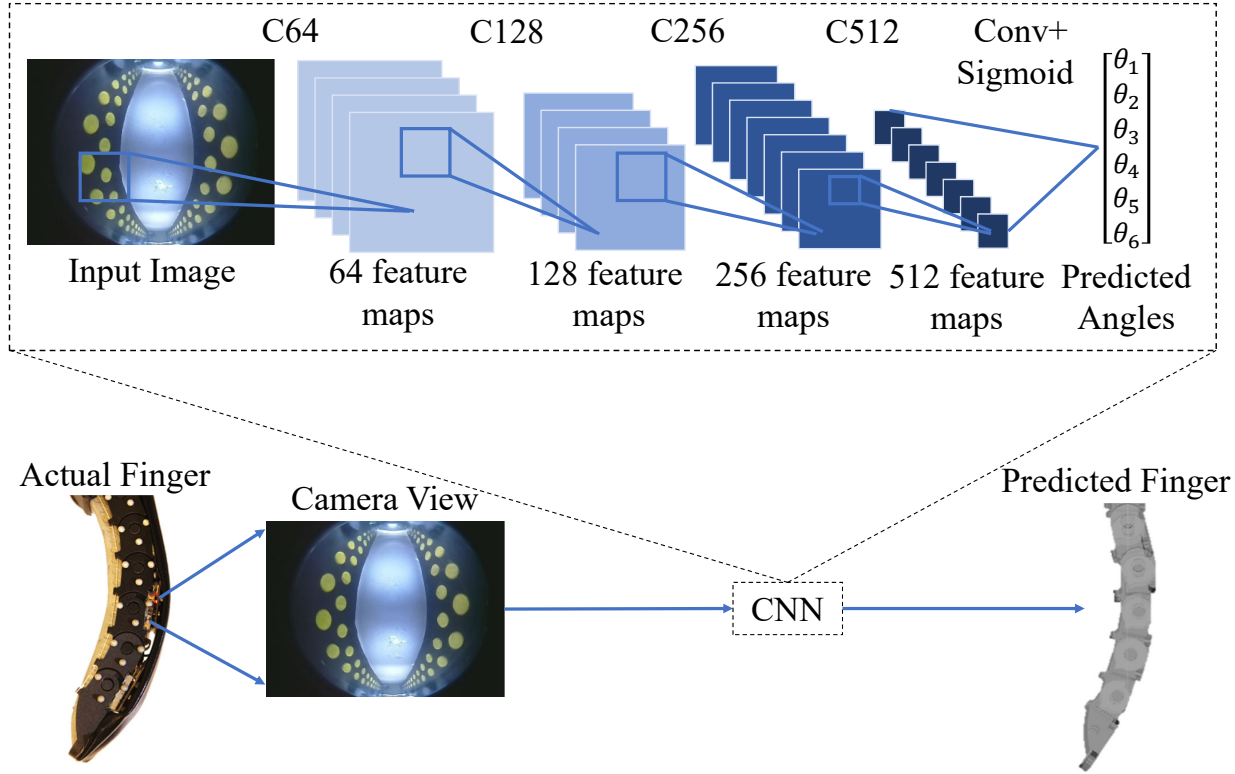


Figure 3.4: **Top:** Architecture of the proprioception neural network where  $Ck$  represents a Convolution-BatchNorm-ReLU layer with  $k$  filters. **Bottom:** Actual GelFlex finger with the intraphalangeal camera view. This image then gets fed into the CNN and the output joint angles are input into a RViz simulation of the finger [13].

### Tactile Sensing

Since all grasping contact is done near the middle of the finger, only the center camera of the GelFlex finger is utilized to obtain tactile information. The GelFlex finger works similarly to the way the GelSight sensor works [42] and is also capable of viewing highly detailed and intricate textures like the dimples in a golf ball or the ridges and notches in a key (Fig. 4.5). There were also observable differences between images the camera captured when the finger grasped sharp corners or edges (i.e. boxes) and when it grasped smoother curved surfaces (i.e. cylinders). This observation indicates that a neural network would be an effective way to classify between different types of bar stock (i.e. square versus round).

The neural network takes a cropped difference image as an input. The input is cropped to disregard the yellow dot pattern, which could interfere with the tactile classification neural network. Then, to ensure that the emphasis in the input data was on the tactile imprint, the difference image of the tactile contact image and a calibration image was calculated. The calibration image used was that of an image without tactile imprints at the starting non-actuated position of the GelFlex.

The network architecture is based on the LeNet-4 model [43] and is composed of a sequential model with two convolutional layers (kernel size  $5 \times 5$ ) that implement ReLU and max

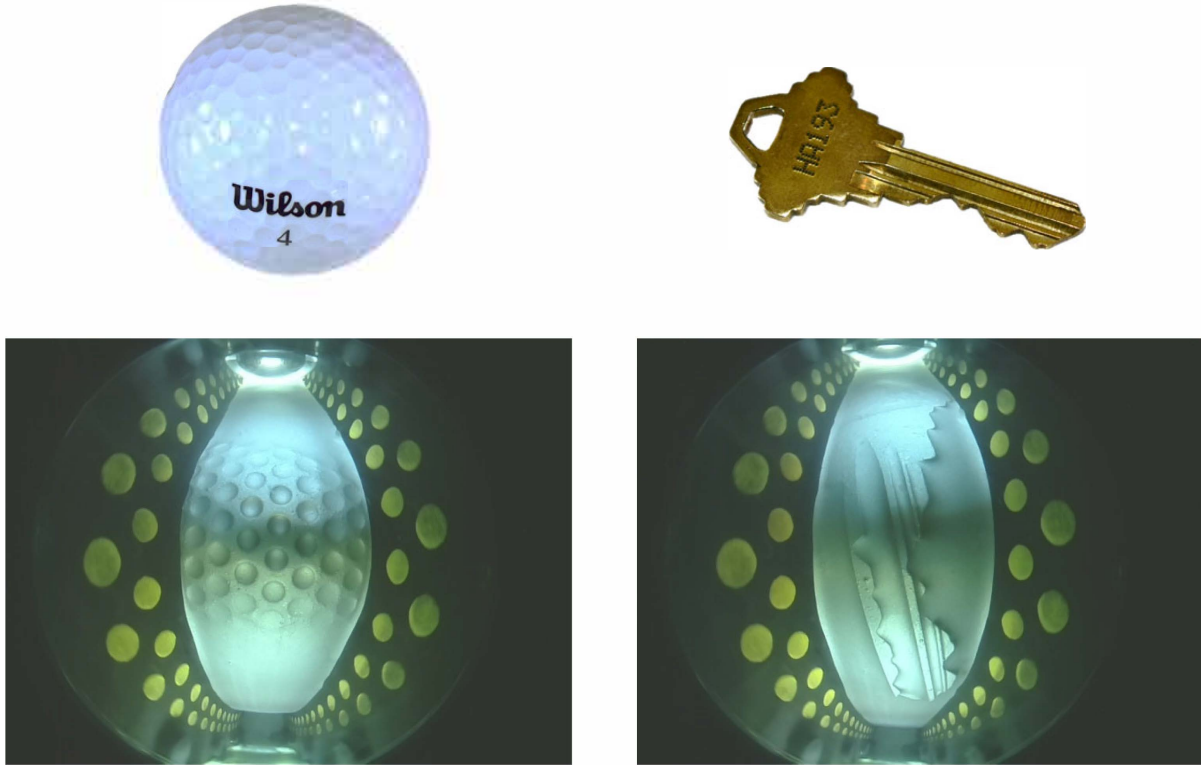


Figure 3.5: **First Row:** From left to right: a golf ball and a key. Both of these objects were brought into contact with the finger. **Second Row:** Their corresponding tactile imprints as viewed from the center intraphalangeal camera [13]. For more tactile images, including the difference images, see Appendix Fig. B.2.

pooling (kernel size 2x2), which is then followed by two fully connected layers. Subsequently, the last two layers use ReLU and softmax activation, respectively (see Appendix Fig. B.3).

A total of 850 pictures were collected for training the tactile CNN. These images were of the finger contacting different box and cylindrical objects at various orientations and contact points (425 pictures each of boxes and cylindrical objects). Of the images, 75% were used for training and the remaining 25% were retained for testing.

### Size Estimation

The classes used for the size estimation neural network were 4.25, 4.5, 4.75, and 5.0 inches. These classes were chosen because they were the shape sizes that the soft robotic gripper could comfortably and successfully grasp. A classification task was chosen as opposed to a linear regression task since bar stock are manufactured with specific dimensional increments.

In total, 800 predicted angles and shape labels were collected (200 for each class, 400 boxes and 400 cylinders); 90% were used for the training set while the remaining 10% were used for the testing set.

Three different network architecture models were considered for the incorporation of the class label (box versus cylinder) into the size estimation prediction method: multiple layer



perceptron (MLP), two-path MLP, and MLP with neural incorporator.

The MLP model provided a baseline comparison for the other models, since it is a standard deep, artificial neural network. For the size estimation task, the MLP model takes the concatenation of predicted angles and the class label (box/cylinder) as inputs. This model performed relatively well, with 88% accuracy on the testing set. However, the MLP model performed significantly worse compared to the 93% testing set accuracy of the two-path MLP, which separately processed the predicted angles and class labels before fusing them via channel-wise concatenation in the embedding space.

The neural incorporator model performed the best, with 100% accuracy on the testing set. The model is inspired by [44], and uses the features learned from the class labels, denoted as  $\beta$  and  $\gamma$ , to normalize features that the model learns from the predicted angles (Fig. 3.6). As such, the neural incorporator model was used for the GelFlex finger.

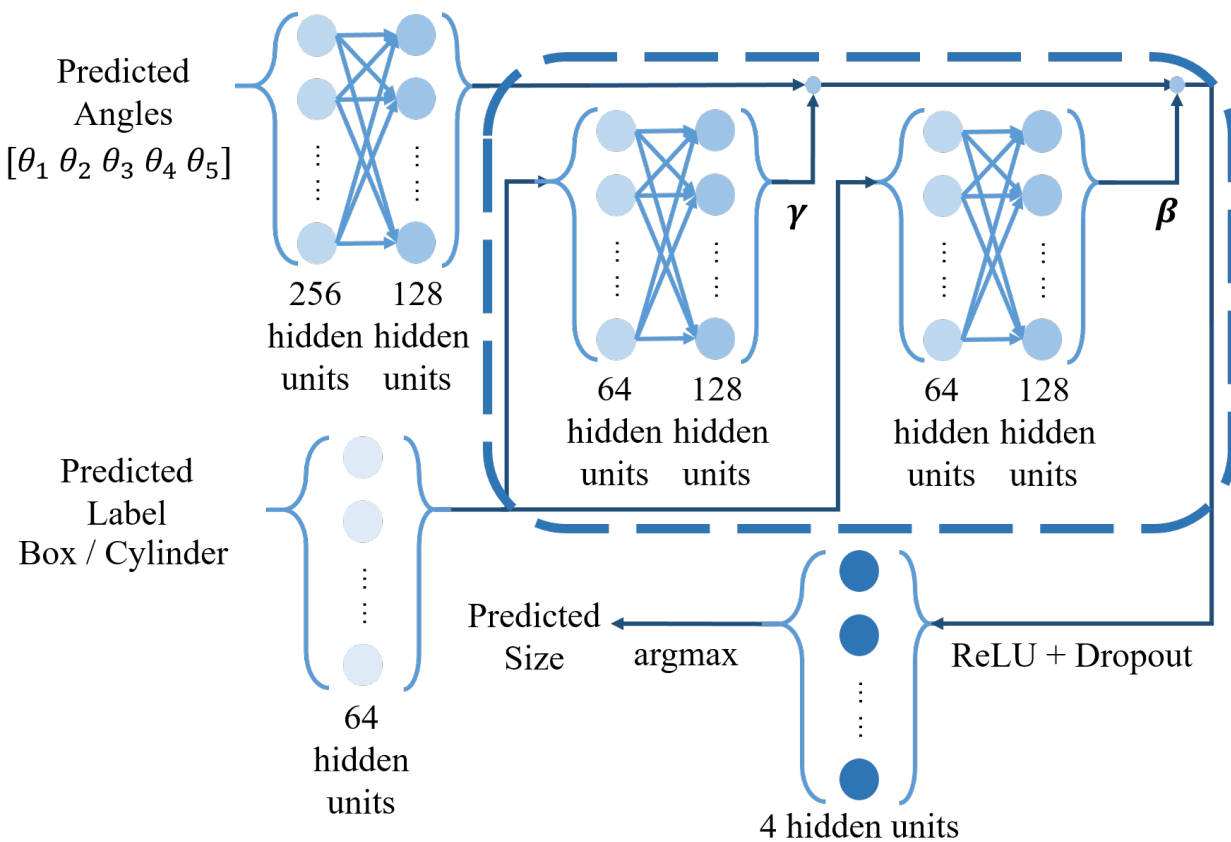


Figure 3.6: The architecture of the MLP with Neural Incorporator used for size classification of the object set [13].

### 3.1.3 Experimental Setup

The OptiTrack (OptiTrack, NaturalPoint Inc.) system was utilized to track the individual positions of the links on the finger. To define the positions of each rotation joint, each finger joint and the tip of the finger had three 3 mm OptiTrack face markers adhered onto their

surface. This defines a rigid body in the OptiTrack software and allows the positions to be streamed via *vrpn\_client\_ros*, which is an open-source ROS package [45].

For the bar stock identification task, the object set consisted of four cylinders and four rectangular prisms with square profiles that were created out of laser cut acrylic. These objects were chosen because they represent the two most common types of bar stock profiles. The cylinder diameters and square widths were 4.25, 4.5, 4.75, and 5.0 inches. These sizes were chosen since they were the smallest objects the gripper could securely grasp. A 0.25 inch difference (approximately 6.35 mm difference) was chosen to demonstrate the ability of the three neural nets to aggregately differentiate between typical differences in bar stock sizes. The entire setup is shown in Fig. 3.7. A close up picture of the GelFlex finger with OptiTrack markers is also shown in Fig. B.1

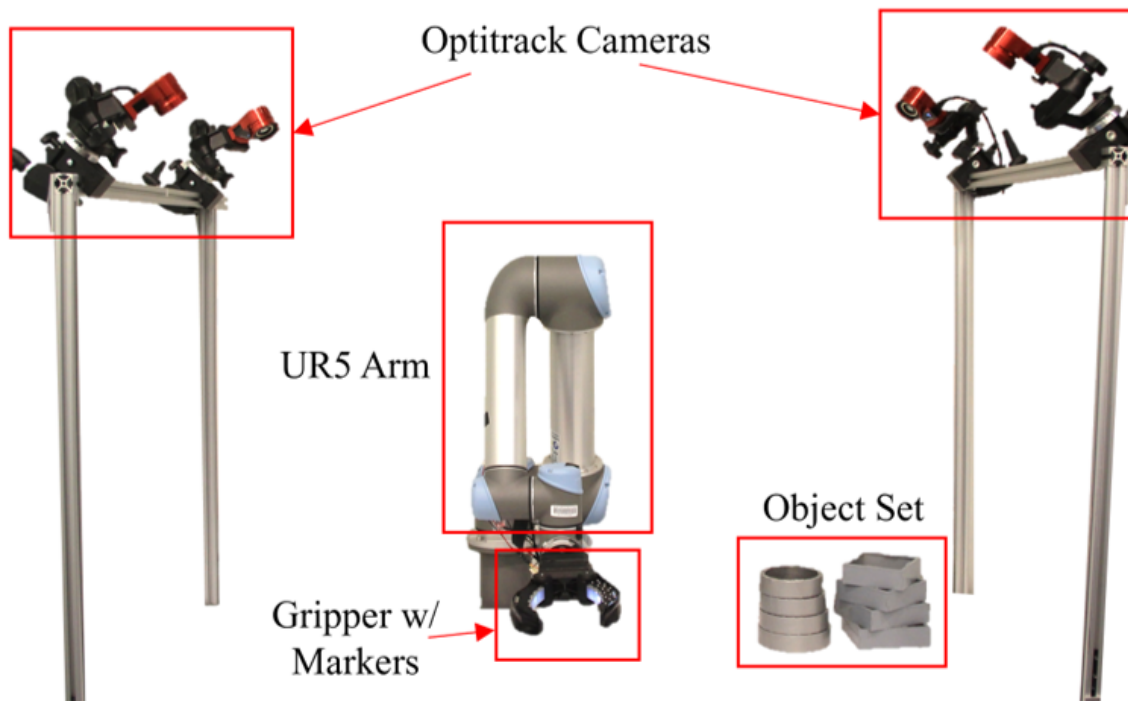


Figure 3.7: Experimental setup for collecting data and testing. The soft robot gripper is attached to the end of the UR5 arm to keep it stationary. Four OptiTrack cameras track the position of the markers adhered to the finger. Testing is performed on the eight objects in the object set [13].

Successful grasp of the object was determined by tracking the position of the motor. Once the position of the motor remained constant after the gripper attempted to grab an object, the algorithm implements the neural nets to determine the size and type of the object being grasped. Since the gripper configuration is symmetrical, the algorithm only obtains reading from one of the gripper fingers. The experiment is performed on a single TITAN X (PASCAL) GPU.

## 3.2 Results

### 3.2.1 Proprioception

Both neural nets (single image input, double image input) used for the GelFlex finger proprioception had extremely accurate results on the testing set, with over 99% accuracy where all six joint angles were within  $1^\circ$ . Live testing with the neural nets on the finger also produced accurate results when the estimated angle values were directly compared with the angles calculated using the OptiTrack positional data. The single input neural net average sum of absolute errors in the six angles was  $1.29^\circ$  while not grasping any objects and  $2.22^\circ$  while grasping objects in the object set. These errors for the double image input neural net were  $1.16^\circ$  and  $4.10^\circ$ , respectively.

To calculate the position errors ("accumulative error"), the ground-truth angles and neural network estimated angles were applied to back calculate the respective position values of the last joint in the gripper. For the one image input neural net, the average accumulative error was 0.070 mm while not grasping any objects and 0.77 mm while grasping objects in the object set. For the double image input neural net, these accumulative errors were 0.40 mm and 1.43 mm, respectively. The results predicted by the single image input neural net were also visually compared with the actual finger using Rviz (Fig. 3.8) [46]. The maximum accumulative error while gripping an object is small (1.71 mm distance,  $5.21^\circ$  over six total links).

In general, live testing results showed that the single image input neural network outperformed the double image input neural network with respect to accuracy and latency. On average, the inference times for the single image input and double image input neural nets are 7.1 and 7.3 ms, respectively.

### 3.2.2 Object and Size Classification

For live testing of the bar stock classification task, each of the eight objects in the object set were tested 10 times with the gripper to calculate the accuracy of the neural network. All box and cylinder shapes were classified correctly using the tactile information. Out of 80 total trials, only three objects were classified as the wrong size (Tables A.2, A.3, and A.4 in Appendix). The incorrect trials occurred only when the gripper was grasping the two largest boxes.

## 3.3 Discussion

In general, the proprioception neural net and tactile classification neural nets performed extremely well. Proprioception of the finger was highly accurate, robust, and capable of responding rapidly. The accumulative errors were small and were generally all sub-millimeter. These results were expected since the internal views of the finger showed visible changes for different finger configurations when they picked up different objects. The camera was in general able to differentiate between small deformations of the bend in the fingers.

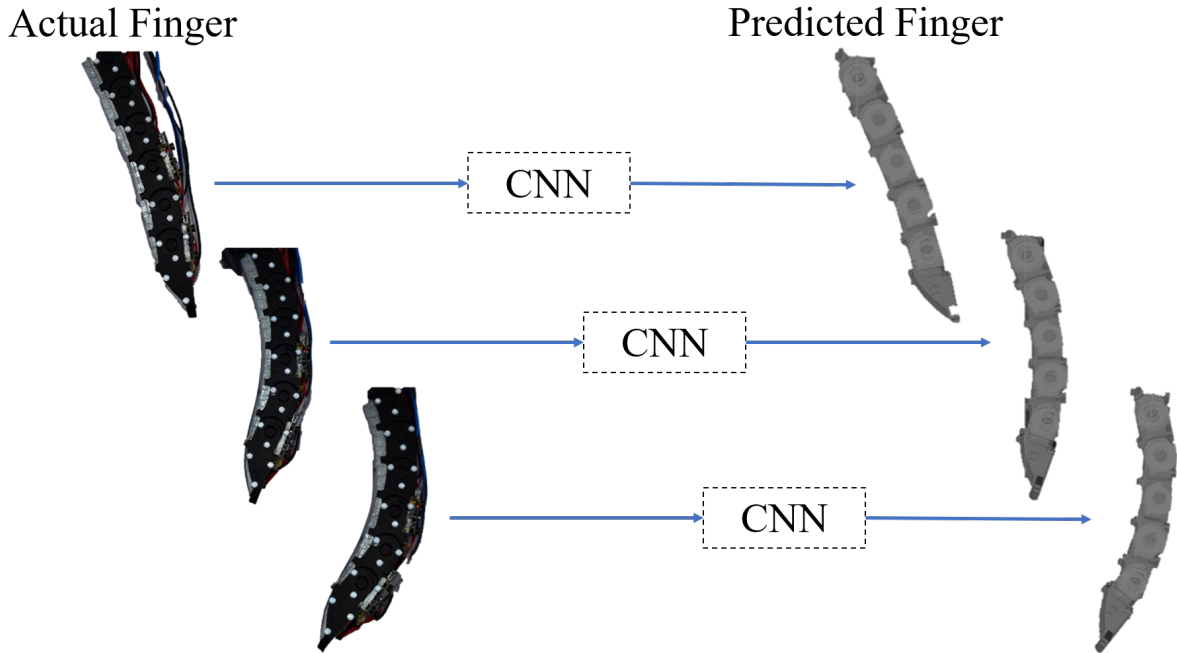


Figure 3.8: A visual comparison between the actual finger and the simulated results as shown by RViz. The intraphalangeal camera view of the GelFlex is input into the proprioception neural net shown in Fig. 3.4 and the resulting predicted joint angles are output into RViz [13].

One of the most surprising results was that the neural network using a single image input tended to outperform the neural network that used two images, which had more holistic information of the internal finger shape. Although the center camera could view most of the finger length, it was unable to view the tip of the finger. With the second camera, it was possible to directly look at the changes in the tip of the finger, which observed the most bending and the largest joint angle ranges. However, the mechanical constraints of the finger could have allowed the single image input neural network to generalize valid joint positions of the end of the finger relative to the other joints in the finger. Another possibility is that the double image input neural network could provide better results if the architecture was changed.

Nonetheless, the maximum and average accumulative error of the finger while grasping are small. This result is significantly better than human hand proprioception where the subject group average of the absolute distance between perceived and actual fingertip position was  $8.0 \text{ cm} \pm 1.0 \text{ cm}$  [47].

On the tactile data side, the GelFlex cameras were able to capture and perceive highly intricate textures and details. As a result, the tactile neural net was able to perform extremely well for the relatively simple human task of differentiating between square and cylindrical objects. This neural net, incorporated with the proprioception neural network, enabled the gripper to accomplish an important task in manufacturing and industry applications: bar stock classification.

The size and shape profile bar stock classification task is not one that can be accomplished without both proprioception and tactile sensing data. Although tactile sensing data can differentiate between features such as edges and rounded surfaces, it is unable to differentiate between different curvatures of cylinders or similarly placed edges of different squares. On the other hand, proprioception of the finger while gripping some cylinders may be confused with proprioception of the finger while it is gripping similarly shaped boxes. The results for the GelFlex finger show that the neural nets were able to very accurately determine the profile and size of the "bar stock" in the object set. The only time when size classification failed were with the two largest boxes, which was potentially caused by a non-firm grip on those objects. A non-firm grasp could have caused the proprioception neural network to output slightly incorrect estimates of the finger joint angles, which would cause the size classification neural network to incorrectly classify the box as a slightly smaller size.

One limitation is that the proprioception of the finger is only in 2D and is described by the shape of the exoskeleton as opposed to the infinite degrees of freedom that describe the entire flexible silicone body. In other words, it is impossible for the neural net to estimate more complex configurations (i.e. twisting) which cannot easily be measured by classic strain/force sensors but can be easily observed with an embedded finger camera.

Furthermore, the GelFlex finger is incapable of obtaining more meaningful tactile information via 3D reconstruction of the tactile data similar to what GelSight sensors can do [42] and it also can only grab larger items. In fact, its inability to grasp many different items eliminates the advantage of using a soft and compliant robotic finger, which is supposed to be better at grabbing a larger variety of objects. Although we showed that using a camera can give us viable proprioceptive and tactile sensing information, we were unable to utilize the camera well in a soft, compliant gripper.

# Chapter 4

## GelSight Fin Ray<sup>\*</sup>

One of the gripper designs we considered was the 3D printable Fin Ray Effect gripper, which has the ability to gently comply to an object it comes in contact with via the fin-like structure of the fingers [48]. The idea was originally inspired by the study of fish fin bone structure and its passive, deformable geometry [49]. This geometry allows Fin Ray fingers to conform to a surface, increasing the grasping contact surface area to allow a more secure grasp [50], [51]. These types of fingers, unlike other soft robotic fingers, are relatively simple to manufacture and easy to modify to incorporate various materials or control schemes [21]. Furthermore, their design does not require actuation to securely grasp objects, unlike many other rigid and soft robotic grippers, which allows Fin Ray fingers to potentially be used in cleaner or limited energy environments [52]. These qualities are important for performing manipulation in an unstructured home space, and as a result, a lot of research has been done to optimize the Fin Ray finger design.

To streamline the Fin Ray struts and its design to more optimally grasp a large variety of objects, both Shan et al. and Deng et al. developed mathematical models and simulations to assess and optimize its grasping capabilities [53], [54]. Crooks et al. worked on adapting the original Fin Ray design to ensure a preferred bending direction and to increase the grip force of the fingers [21] while Elgeneidy et al. worked on improving the design of a NinjaFlex Fin Ray finger through experimental design testing methods [55]. Another solution to improve grasp quality of a Fin Ray gripper involved adding electroadhesion pads to the tactile surface of the fingers [56].

Despite the progress made towards optimizing the Fin Ray finger design, not much work has been done to sensorize it. Although Yang et al. proposed a Fin Ray finger that incorporates an embedded force sensor for detecting the finger state and to measure contact force [57], it does not provide the higher-detailed tactile sensing necessary for more complex manipulation tasks. The addition of a high-resolution tactile sensor like GelSight could benefit the Fin Ray finger and allow it to perform manipulation tasks that require such sensing.

To our knowledge, this was the first instance of a soft robotic gripper that has incorporated the full functionality of a GelSight sensor, allowing tactile reconstruction, object orientation detection, and marker tracking. The design introduced in this paper also allows a GelSight sensor to be elongated and flexible, which captures a larger tactile surface area

---

<sup>\*</sup>Most of this chapter is copied from the paper GelSight Fin Ray: Incorporating Tactile Sensing into a Soft Compliant Robotic Gripper [14].

and enables compliant grippers to utilize GelSight sensor technology.

This chapter presents the following contributions:

- A novel design of a fully-functioning flexible, elongated tactile sensing surface with an embedded vision sensor for tactile reconstruction, orientation estimation, and slip detection;
- A novel design of a sensorized GelSight Fin Ray inspired gripper with tactile sensing (Fig. 4.1);
- An algorithm that allows the gripper to be handed a thin object, determine its orientation, and gently set it down upright using only tactile sensing.

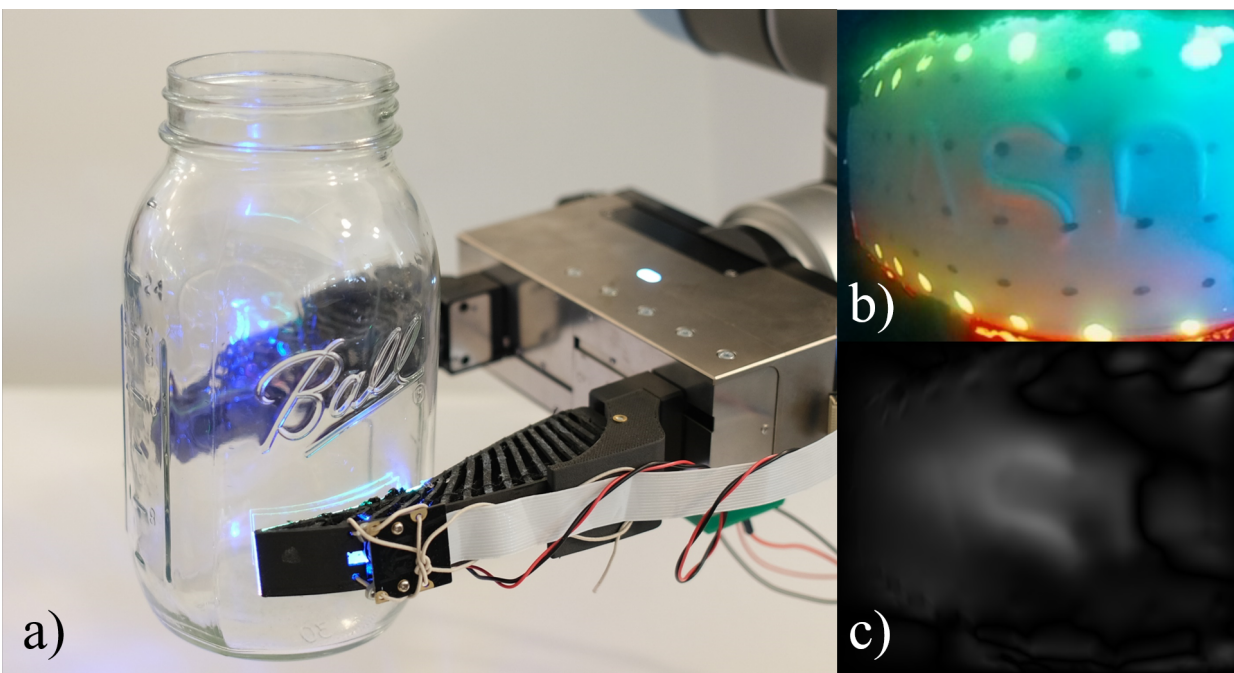


Figure 4.1: In a), the GelSight Fin Ray gripper is holding a ball glass mason jar with its tactile sensing region along the indented letters “MASON” while b) denotes the  $90^\circ$  counterclockwise-rotated raw tactile image data where the top of the image generally denotes the tip of the finger and c) is its corresponding uncalibrated tactile reconstruction, where you can see most of the “S” letter and parts of the “O” and “A” letters as well [14]. Another example of a smaller object being grasped is shown in the Appendix Fig. B.4.

## 4.1 Methods

### 4.1.1 Sensor Design

Overall, the sensor design involves three major components: finger design, tactile sensing pad manufacturing, and illumination. The resulting finger was attached to a parallel gripper

configuration. Fig. 6.3 shows the entire manufacturing process of the finger. We also briefly consider the fabrication of a completely optically clear Fin Ray; that design idea quickly became infeasible due to its lack of rigidity. The optically clear Fin Ray and other manufacturing processes we tried for the Fin Ray are discussed more in Chapter 8.

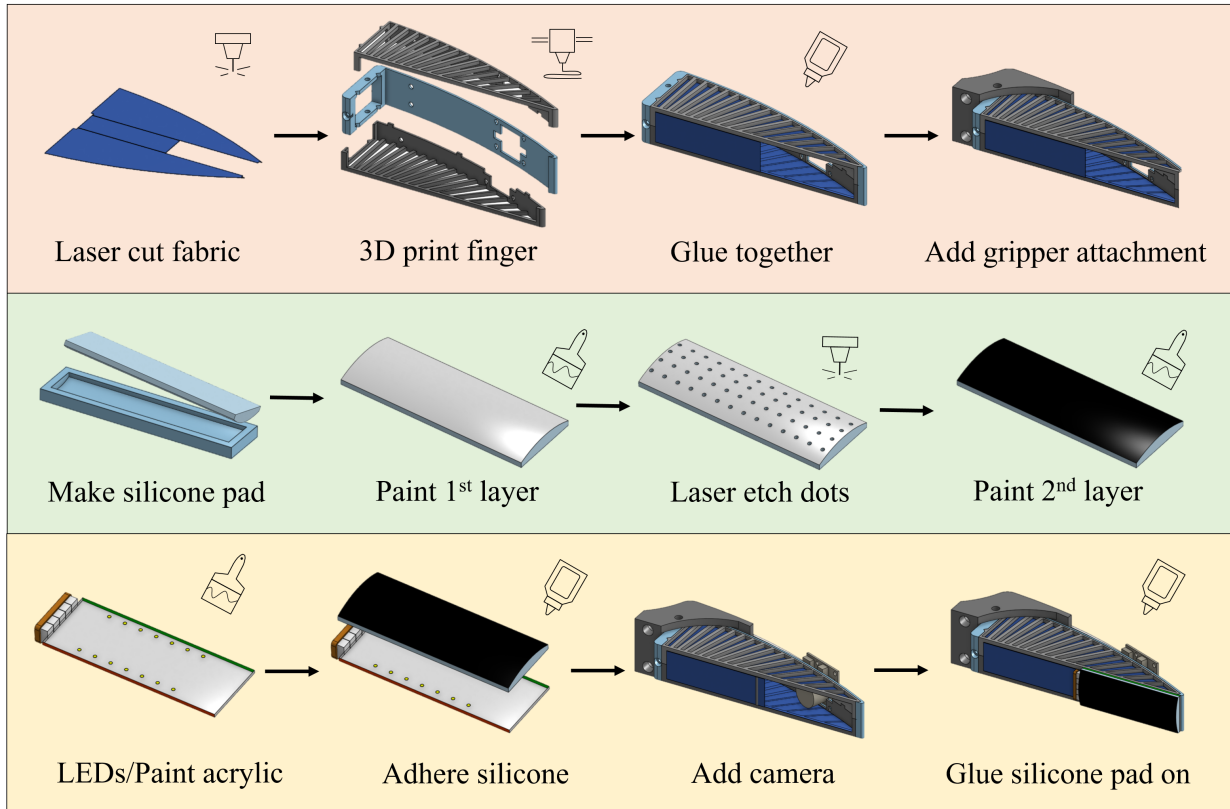


Figure 4.2: Manufacturing process flow of the GelSight Fin Ray finger, which includes finger design, tactile sensing pad manufacturing, and illumination [14].

### Finger Design

The struts of the finger were designed based on the design proposed by Elgeneidy et al. [58]. To create a finger that could house and allow usage of a vision sensor while retaining the structural features of a Fin Ray, the middle part of the finger was hollowed out. This removal of the middle support material caused the finger to become too flexible, so a harder material (1.75 mm TPU, Overture, printed on Prusa MK3S) was used for the finger struts while an even more rigid material (Onyx, Markforged) was attached to the back of the finger as support. To circumvent the unwieldy removal of TPU support material from the main finger print, the struts portion of the finger was separated into two parts along its length. Doing so allowed the parts to be printed with the outer section against the build plate, enabling them to be printed with no support material and eliminating extensive post-manufacturing processes while also shortening the printing processes.

Since the sensor utilizes a camera, it was important that any ambient lighting be obscured so that it did not interfere with the tactile sensor. However, any material added to the finger also had to be stretchy so that it would not interfere with the compliant characteristics of



the finger. To solve this issue, a laser cut piece of black, stretchy cloth (spandex/nylon) was adhered to the inside of the struts and across the bottom half of the open part of the finger before the three 3D printed parts were pieced together. The bottom half was covered since most of the manipulation and grasp tasks are performed near the tip of the finger, where the finger is more compliant. The dark cloth obstructed outside environmental lighting while its elasticity allowed the finger to function as it would without the addition of the cloth. Afterwards, the two halves of the finger were glued onto the rigid base with cyanoacrylate glue.

### **Tactile Sensing Pad Manufacturing**

The tactile sensing silicone pad was cast as a rectangular-profile piece with dimensions of 2.5 mm by 18 mm, with a 25 mm curved profile along one of the 18 mm surfaces to serve as the tactile sensing surface. A curved profile was chosen to help guide the side illumination and evenly distribute the coloration along the sensor center. The mold was 3D printed using the Markforged Onyx material. To ensure that the tactile curved surface was smooth, a piece of plastic mylar sheet with 6 mil (0.15 mm) thickness was superglued onto the inner curved mold surface. A platinum catalyst translucent silicone (XP-565, Silicones, Inc.) was combined with a plasticizer (LC1550 Phenyl Trimethicone, Lotioncrafter) in a ratio of 11 to 3 (1 part XP-565 activator to 10 parts base to 3 parts plasticizer) to be used for the silicone pad. The mixture was used to ensure a clear, transparent sensing region for the camera sensor and a plasticizer was added to increase the softness and robustness of the silicone pad. After degassing the silicone mixture and pouring it into a pre-prepared mold, the mold was left to cure for 24 hours at room temperature.

Once the silicone pad was cured, a thin layer of silicone paint was brushed on the curved side of the silicone with a foam brush. This paint layer formed a semi-specular tactile sensing surface. To create this painted surface, which would allow the sensor to more easily pick up small details and dimmer illumination in the tactile sensing surface, a mixture of 1 part silicone ink catalyst to 10 parts gray silicone ink base (Raw Materials Inc.) to 2.5 parts 4  $\mu\text{m}$  Aluminum cornflakes (Schlenk) to 30 parts NOVOCS Gloss (Smooth-on Inc) was used. After allowing the painted surface to fully cure, 1 mm circular dots spaced 4 mm apart were laser engraved (Epilog Laser Cutter System) on the painted surface, effectively removing the painted layer. A layer of black silicone paint (1 part catalyst to 10 parts black paint to 30 parts NOVOCS Gloss) was then sprayed on using an airbrush, creating patterned black dots on the silver-gray tactile surface. The dots allow for the use of marker tracking to determine motion or slippage of a grasped object.

One of the manufactured silicone pads was then glued to a piece of 0.5 mm thick acrylic to provide the silicone with a rigid backing and prevent the silicone from folding in on itself when it comes in contact with an object. A mixture of 1 part Silpoxy (Smooth-on Inc) to 10 parts NOVOCS gloss was used so that the silicone could bond to the acrylic and prevent delamination of materials over repeated use.

The thinness of the acrylic is to ensure that the finger has overall compliance and can still bend when it comes into contact with an object. Acrylic was also chosen as the rigid base of the silicone pad for its ease of manufacturing and its reasonably close index match with silicone, which helps to mitigate light interference of the sensing surface caused by refraction. This acrylic piece was then attached to the 3D printed finger.

## **Illumination**

To illuminate the finger, the sides of the acrylic were respectively covered with a layer of red and green fluorescent acrylic paint (Liquitex BASICS ACRYLIC) and a line of blue LEDs (450 nm) was adhered to one end of the acrylic piece to illuminate the silicone pad and cause the paints to fluoresce. A short strip of TPU filament was superglued to its back and along the empty middle section of the finger to hold the blue LEDs in place. Together with the fluorescent paints, the blue LEDs helped to provide a tri-color sensing region, similar to how a GelSight wedge and digger fingers are illuminated [31], [59]. The presence of all three colors help the sensor to perform depth reconstruction of objects the gripper comes into contact with.

The camera is set opposite the tactile surface near the tip at the back of the finger so that it can view the most compliant parts of the gripper as well as where the majority of the tactile interactions will occur. To mitigate the saturation of blue LEDs viewed by the camera, a yellow filter is placed atop the camera sensor. The filter helps make the less vivid red and green colors visible by the camera.

Because illumination of the finger as viewed by the camera changes as the finger flexes, separate yellow fluorescent dots (Liquitex BASICS ACRYLIC) were painted on both sides of the acrylic surface to provide a reference for the shape of the finger. The color was deliberately chosen so that it could easily be visible in the dim illumination and so as not to interfere with the tactile sensing. These dots are separate from the black silicone dots used for marker tracking, which depict the shear information of the silicone surface.

## **Gripper Assembly**

The fully assembled finger was then attached to the parallel WSG 50-110 (Weiss) gripper using an Onyx 3D printed part designed to fit as a replacement for the traditional Weiss gripper fingers.

### **4.1.2 Software**

Similar to a GelSight sensor, the Fin Ray finger is able to perform tactile sensing, measure the orientation of an object it comes into contact with, and track markers for slip and twist detection [42].

All images were captured using a Raspberry Pi 160° field of view (FOV) camera, before being cropped slightly so that only the tactile sensing portion of the image was shown. A wide FOV camera was used so that it can sense along the entire length of the sensing region.

#### **Tactile Sensing/Reference Image**

The main challenge with incorporating tactile sensing into the Fin Ray finger is that the finger deforms when any object comes in contact with it. As a result, the camera sees both the proprioceptive change when the finger bends, and the minute changes in the silicone pad corresponding to the tactile surface interactions. Tracking only the fine-detailed tactile information requires knowledge of the proprioceptive state of the finger or the bending positions of the acrylic. As such, yellow fluorescent dots were painted on the sides of the acrylic so that the motion of the finger could be tracked and used as a reference. This reference was then used to isolate the small-scale tactile information from the GelSight sensor.

A reference video was created with the finger where a human hand flexed the sides of the finger which are not visible by the camera sensor. As such, the camera was able to capture the different illumination schemes without any tactile interference. Using the video, the dots were thresholded out via HSV color segmentation, opening and closing operations were performed on the thresholded image, and the individual dot centers were extracted from each frame of the reference videos using a Python OpenCV contour finding algorithm and converted into a Numpy array for better latency [60], [61]. Subsequent tactile images were then compared to reference images using the dot segmentation method to find the image that most closely matched the current image. The general algorithm flow is shown in Fig. 4.3.

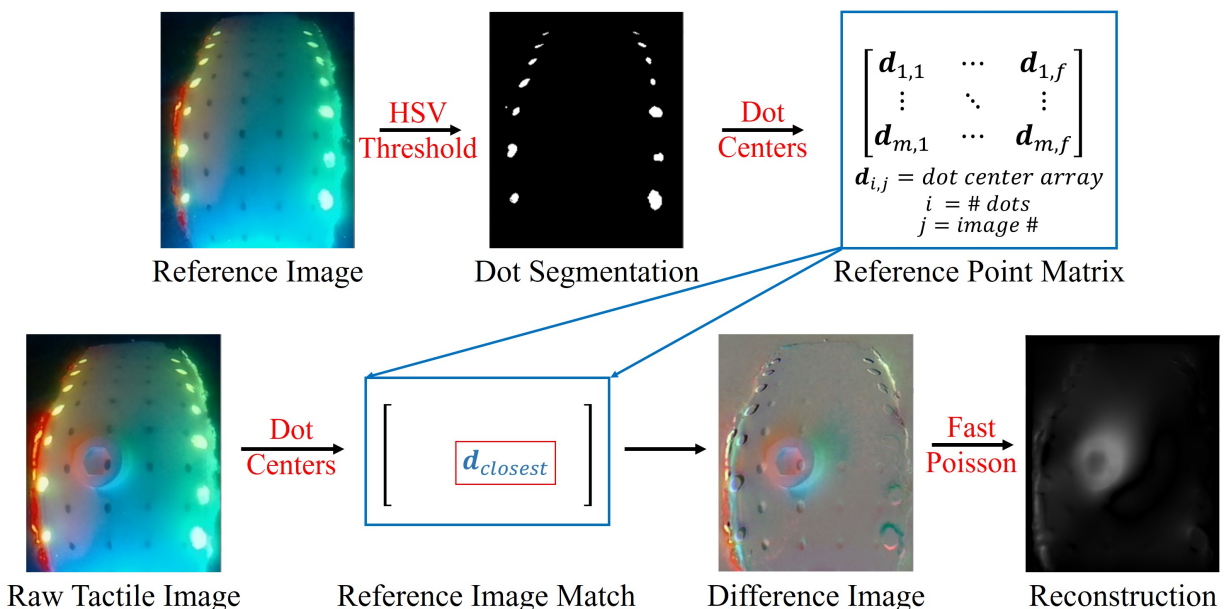


Figure 4.3: Algorithm flow for obtaining the reference point matrix and live reconstruction of the raw tactile image. The reference point matrix is precalculated using a video taken of the finger without any tactile interference. The matrix is sparse and contains only one nonempty element per column. Subsequent raw tactile images are then compared to the reference point matrix via the dot centers, a difference image is taken, and an uncalibrated reconstruction image is calculated. Pictured in the raw tactile image is the head of a screw that was pressed into the sensing region at a slight angle [14].

### Object Orientation

After matching the tactile image with a reference image, difference images were taken and then Poisson reconstruction was used to calculate uncalibrated depth reconstruction images [62]. The contour finding algorithm was then used on thresholded versions of the depth images and principal component analysis was used to find the orientation of the object that was being grasped.

### Marker Tracking

Marker tracking was performed using the difference image between the luminosity channel

of the LAB color space and its corresponding median-blurred image. This method allowed the markers to be segmented out of each image. Arrows were then drawn from the reference image marker to its closest marker in the actual image. If the distance between them was significant, the arrow length was scaled by a factor of three to emphasize the shear and twist motion of the object in contact.

## 4.2 Experiment

### 4.2.1 Experimental Setup

Software was run via Python and ROS [45] and incorporated object orientation estimation and marker tracking of the tactile surface. The camera was run on a Raspberry Pi board and images were streamed to the computer using `mjpg_streamer`. The fingers were attached to the Weiss gripper, which was mounted on a UR5 arm. Although both of the fingers in the parallel gripper were Fin Rays, only one of them was sensorized.

To test the compliance and gripping ability of the fingers, many test objects were handed to the gripper to be grasped. These objects included a mini screwdriver, a plastic strawberry, a plastic lemon, a plastic orange, a Rubik’s cube, a wine glass, a ball glass mason jar, and a deformable, squishy acrylic paint tube (Fig 4.4).



Figure 4.4: Objects used for the experimental setup. They include a mini screwdriver, plastic fruits, a Rubik’s cube, a wine glass, a mason jar, and a squishy acrylic paint tube [14].

To test the orientation estimation and marker tracking, the gripper was programmed to be handed a wine glass stem, measure its orientation based on tactile results, and rotate the gripper so that the wine glass would be oriented upright. Afterwards, the gripper would set the wine glass down until the total magnitude, or shear force, detected by the marker tracking algorithm exceeded a threshold, indicating that the glass had touched the table. A

wine glass was chosen because it is difficult to segment transparent objects using vision, and thus, tactile sensing greatly simplifies the reorientation and placement problem.

## 4.2.2 Results

### Grasping

The Fin Ray fingers were able to comply to and grasp all of the test objects, which included a variety of different sizes and shapes. For curved objects, the Fin Ray fingers were able to wrap around the curvature of the grasped objects while for the more linear objects, such as the screwdriver and the Rubik’s cube, the fingers acted more like a rigid, parallel gripper. For the squishy paint tube, the Fin Ray fingers were able to comply around the object without breaking or collapsing the paint tube. Although the grips for heavy and asymmetrically weighted objects were subject to slippage, the grasp became sturdier once the force the gripper could apply was increased.

This difficulty in grasping particular objects may be partially due to the slight inadvertent twisting motion that the fingers sometimes exhibited due to the hollowing out of the structure inside, which most likely decreased the gripping strength of the fingers. Despite the addition of the acrylic mount, which did limit the twisting motion, there was still a minute amount of twist in the finger itself. Furthermore, the acrylic piece could have also created challenges in grasping since it may have limited the flexibility of the finger and required more force to ensure greater compliance of the finger to the object it was grasping. There was a tradeoff between the twisting motion of the structure and its compliant flexibility. However, this tradeoff did not result in hindering any of the compliance or grasping abilities of the Fin Ray finger for the objects in the testing set.

In summary, the GelSight Fin Ray gripper was able to grasp a multitude of objects, showcasing that the fingers still retain their compliance properties, which makes them useful for a large variety of grasping tasks that potentially need robust, universal grasping abilities combined with tactile sensing.

### Tactile Sensing

The tactile surface of the sensor is able to detect minute details, such as some of the threads of a M2.5 screw, and the indents on the outside of a M4 heated insert. All of the reference images were able to help provide good reconstruction images, despite the slightly nonuniform fluorescing quality of the red and green paints. The raw tactile images and reconstruction images are shown in Fig. 4.5.

Generally, the dot center matching algorithm was quite robust, but it seemed to struggle more when the GelSight Fin Ray fingers wrapped around heavier objects such as the large mason jar. The extra weight caused the fingers to twist a bit and would shift the raw tactile image down by a nontrivial amount, causing some noise in the uncalibrated reconstruction image since it was difficult to include twisting motion in the reference images without interfering with the tactile sensing region. Regardless, the marker tracking was able to sense the shear force motion due to twisting by gravity and was able to sense twist and shearing motions along the center of the tactile surface (Fig. 4.6).

Due to the proximity of the black dots to the fluorescent dots, there were cases where some of the black dots could not be perfectly tracked. This phenomena also occurred for

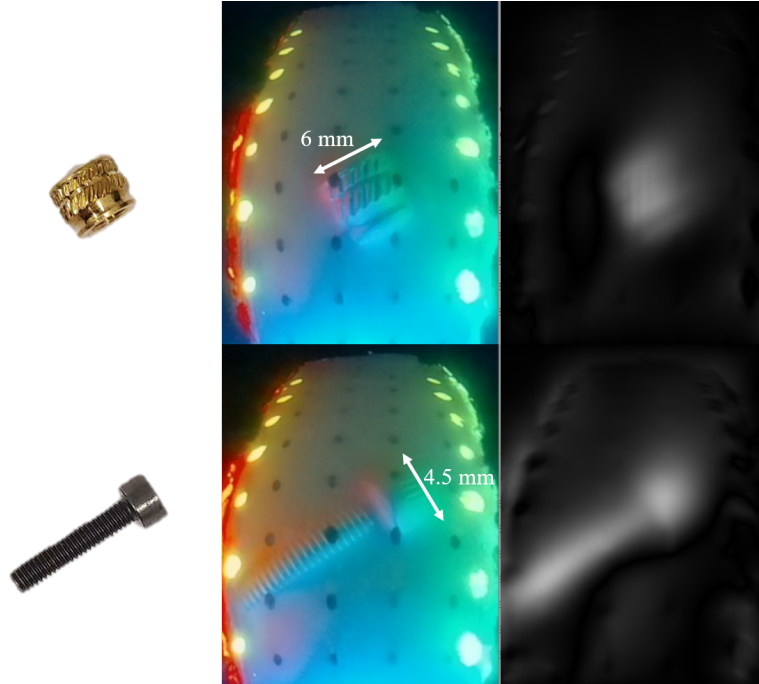


Figure 4.5: Tactile sensing image reconstruction. The left most images show the objects (M4 heated insert, M2.5 screw) that were placed on the tactile sensing images, the middle images represent the raw tactile data, and the right most images display the reconstruction based on the raw data [14].

the dots in some of the regions at the top of the raw tactile image, which are furthest from the illumination of the blue LEDs. Having black dots also made the algorithm track a few black spots at the outer edges of the tactile sensing, which served to add some noise into the system as well. Nevertheless, all of this noise did not affect the marker tracking overall, and can be potentially optimized for future design iterations of the GelSight Fin Ray.

### Wine Glass Reorientation

Without tactile sensing, the task of reorienting and setting down a wine glass without tipping it over or crashing it into the table becomes more complex due to its transparency, which makes it harder to see. With the GelSight Fin Ray, however, wine glass reorientation was generally successful (Fig. 4.7). Out of the 10 trials that were performed, the algorithm was able to succeed in 7 of them; the GelSight Fin Ray finger was able to allow the gripper to successfully reorient and set the wine glass down without it tipping over. In the three cases where the algorithm failed, the grasp was not secure enough and the wine glass slightly slipped out of the Fin Ray grasp, causing the reconstruction of the image to fail, which in turn caused failure of the reorientation portion of the wine glass. However, in those cases, the gripper was still able to detect when the wine glass came in contact with the table and stop itself from crashing the wine glass into the table further and potentially breaking it.

To conclude, the marker tracking portion of the software, despite some noise in the system, was more robust than the reconstruction algorithm. This robustness most likely occurs because the marker tracking algorithm does not require firm contact of the silicone pad sensing surface with its object. Bare minimum contact is needed as long as the top part of the silicone pad can be visibly displaced by external forces. Since the silicone pad is soft,

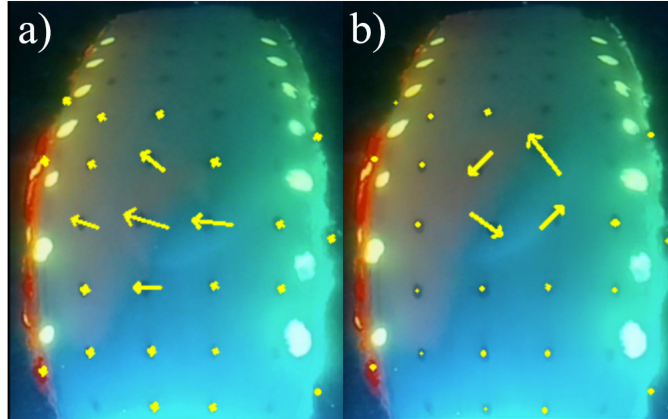


Figure 4.6: Marker tracking with the GelSight Fin Ray. The yellow arrows denote the tracked position from the reference markers in the found reference image to the markers in the current image. In both images, the fingers are gripping a smooth acrylic cylinder. The left image (a) denotes shear along the cylinder from an external force while the right image (b) displays torsional external force being applied on the cylinder [14].

the top portion of the pad was easily deformable, causing the marker tracking portion of the tactile sensing to be inherently more robust. The future addition of a neural network could also generally increase the accuracy and results of tactile reconstruction and marker tracking.

### 4.3 Discussion

The incorporation of tactile sensing has the potential to allow soft compliant robots to mimic or outperform human capabilities to reconstruct the world around them through tactile manipulation. Such abilities may ultimately allow soft robots to improve their performance in medical fields (i.e. surgical robots or prosthetics), during human-robot interactions, or on rescue missions in unknown situations. In particular, Fin Ray fingers allow manipulation to be performed in limited energy environments due to their simplicity in actuation, enabling them to be of great usage in many different environments around the world. However, many existing sensors that provide very intricate tactile information are rigid, compact, and not suitable for utilization in a soft robot.

This chapter introduced a novel flexible, elongated GelSight sensor that can be used in conjunction with a Fin Ray inspired finger, allowing it to perform reconstruction, object orientation estimation, and marker tracking. The combination of the two technologies allow the Fin Ray finger to perform a greater variety of manipulation tasks, including the successful reorientation and placement of a transparent wine glass without any catastrophic failures.

A few of the limitations are inherent to the twisting motion of the hollowed out finger, which is traded off with the rigidity of the acrylic at the back of the sensing region. This limitation caused the Fin Ray finger to require a nontrivial amount of force to comply to more curved objects. Additionally, if the object being grasped was too heavy or if the force was not enough to grab a smaller object to impart a large imprint on the silicone pad, the

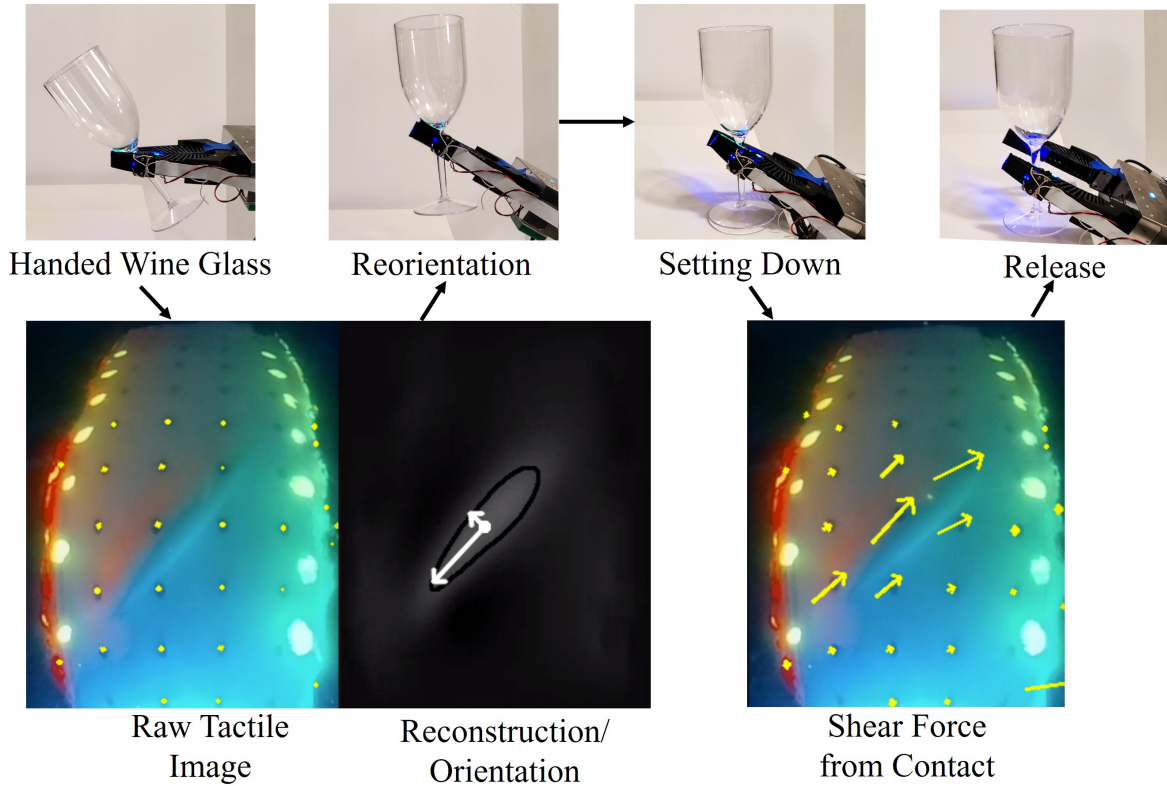


Figure 4.7: Successful implementation of wine glass reorientation. A wine glass is handed to the gripper at an angle. Using the raw tactile image, the orientation of the wine glass stem is found through the reconstruction algorithm. Using this angle, the UR5 arm reorients the gripper so that it is holding the wine glass right side up. The arm then sets the wine glass down until the shear force from the contact goes above a certain threshold and causes the gripper to release the wine glass on the table without it tipping over. Note that the shear force detected by the marker tracking algorithm is along the major direction of the stem, indicating that the wine glass is upright [14].

gripper would occasionally struggle with successful implementation of reconstruction and object orientation estimation. Finally, the addition of the camera, semi-flexible acrylic piece and LEDs to the Fin Ray created some robustness and compliant issues. Both the camera and acrylic piece added rigidity to the Fin Ray system, causing it to have less compliance when grasping smaller objects. Furthermore, there were problems with gripping placement, where if an object was grasped closer to the base of the Fin Ray finger, it could accidentally displace the LEDs. For long-term manipulation tasks or learning paradigms, the robustness and compliance of the system needed to be addressed.

These issues can be fixed with future optimization of the design of the hollowed out finger, which we will discuss in the next chapter.



# Chapter 5

## GelSight Baby Fin Ray\*

Overall, we still want the safety and compliance given to us by soft robotic gripper designs combined with the high-resolution tactile sensing that camera-based sensors can provide us. The development of the original GelSight Fin Ray [14] has helped us bring these seemingly disparate characteristics together. Even so, there are existing issues with its design such as its slight loss of compliance due to sensor integration. To further solve this integration problem, we present the following contributions:

- A novel design of a soft, compliant and robust GelSight sensor in the Baby Fin Ray (Fig. 5.1);
- Synthesis and analysis of fluorescent pigment that has potential use in other soft robotic integration of camera-based tactile sensors;
- Successful performance of nut classification (i.e. identifying textures) using the Baby Fin Ray.

### 5.1 Methods

The improvements to the original GelSight Fin Ray encompass mainly hardware updates and result in the development of the Baby Fin Ray. A side-by-side comparison of the two fingers are also shown in Fig. 5.2, and the exploded version of the Baby Fin Ray is shown in Fig. 5.3.

These advancements were made with the goal to improve the design elements of the sensorized Fin Ray itself, while also showing more of its potential capabilities for tasks that require rich tactile sensing. Another important aspect is the improvement of fabrication techniques that could be utilized in other types of soft, compliant camera-based tactile sensors.

---

\*Most of this chapter is copied from the paper GelSight Baby Fin Ray: A Compact, Compliant, Flexible Finger with High-Resolution Tactile Sensing [15].



Figure 5.1: The Baby Fin Ray grabbing a walnut (left) and the corresponding cropped raw image of the mirror that displays the tri-colored tactile sensing region and the indentation of the walnut shell (right) [15].

### 5.1.1 Hardware

Much of the hardware design for the Baby Fin Ray is inspired by the GelSight Fin Ray paper [14]. Like the original design, we create a hollow inner structure inside of the finger to allow unobstructed viewing of the tactile sensing surface. We utilize a 3D printed 1.75 mm TPU 95A material (Sainsmart) for the Fin Ray struts with a more rigid backing (Onyx material, Markforged) to prevent unwanted twisting motions in the finger that would be induced by hollowing out the structure.

For the silicone gel pad, we use a mixture of 1 to 10 to 3 parts of XP-565 Parts A and B (Silicones Inc.) and a plasticizer (LC1550 Phenyl Trimethicone, Lotioncrafter), respectively. The XP-565 silicone is used because of its translucency. However, because the silicone itself is rigid, the plasticizer is added to soften the material and make it more suitable as a tactile sensor.

Our mold is 3D printed out of Onyx material and a thin piece of 6 mil (0.15 mm) mylar is adhered to the bottom curved portion of the mold with cyanoacrylate glue. Doing so provides a smooth surface for the tactile sensing portion of the gel pad. The mold has a 2.5 mm by 18 mm rectangular cross sectional area with a 25 mm radius curved profile along one of the longer sides to ensure even side illumination of light. To ensure levelness of the sensing pad, a flat piece of translucent acrylic sheet is placed on top of the mold, leaving an area in

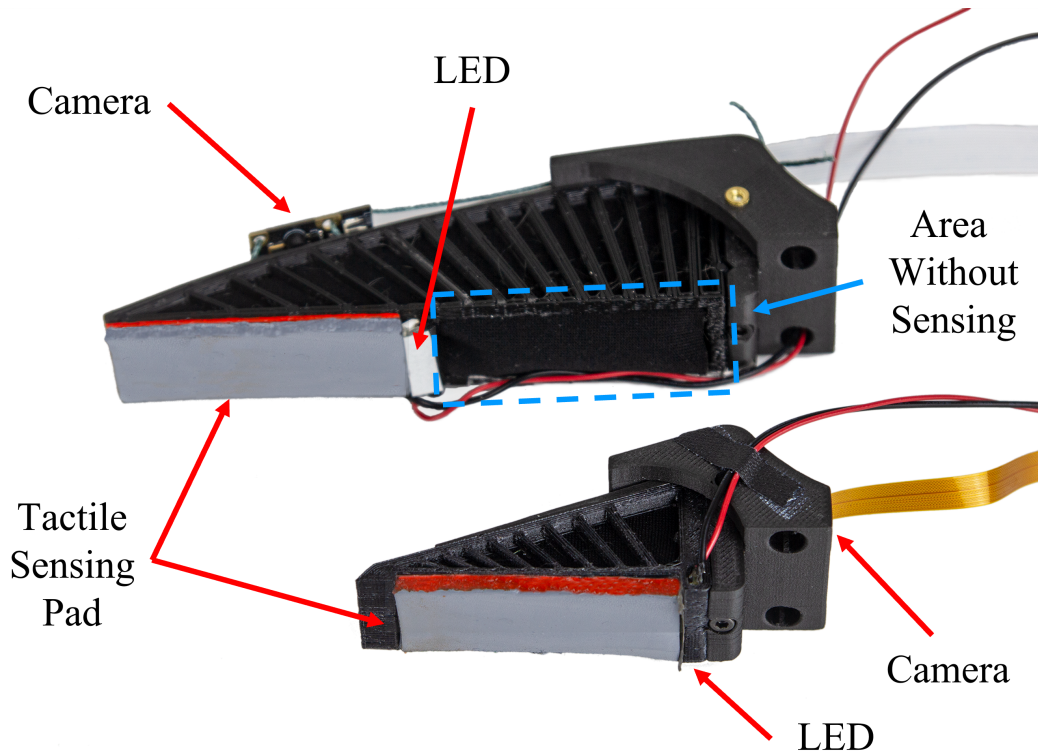


Figure 5.2: A side-by-side comparison of the original GelSight Fin Ray and the Baby Fin Ray [15].

which the silicone could be poured. The whole ensemble is secured with rubber bands and tilted at a slight angle to allow air pockets to rise to the top of the pouring pocket. Once the bubbles escape out of the mold, the mold assembly is placed inside of an oven and cured for 5 hours at 52°C (125°F).

Afterwards, we synthesize our sensing membrane paint, which covers the curved surface of the silicone gel pad. This paint, like the one for the original GelSight Fin Ray, is composed of 1 part silicone ink catalyst to 10 parts gray silicone ink base to 2.5 parts 4  $\mu\text{m}$  aluminum cornflakes to 30 parts NOVOCS Gloss (Raw Materials Inc., Schlenk, Smooth-on Inc). This is brushed only on top of the silicone gel pad, so that fluorescent paint can be applied to the sides of the silicone gel and so that the camera can see through the gel volume to the sensing surface.

However, the Baby Fin Ray differs from its predecessor in many aspects. Specifically, there were a few problems with the design integration of the GelSight sensor into the Fin Ray in the original design. Despite its ability to comply to different object shapes while performing high-resolution tactile sensing, which made it able to do simple tactile manipulation tasks, the hardware was not very robust. The sensing region was limited to only half of the sensor despite the use of a large, wide angle fisheye lens camera, and the finger was unable to comply as much as a non-sensitized Fin Ray would be able to. These issues were addressed by our updated hardware design. The main problems we chose to fix included the following:

- Design for Usability
- Camera Placement

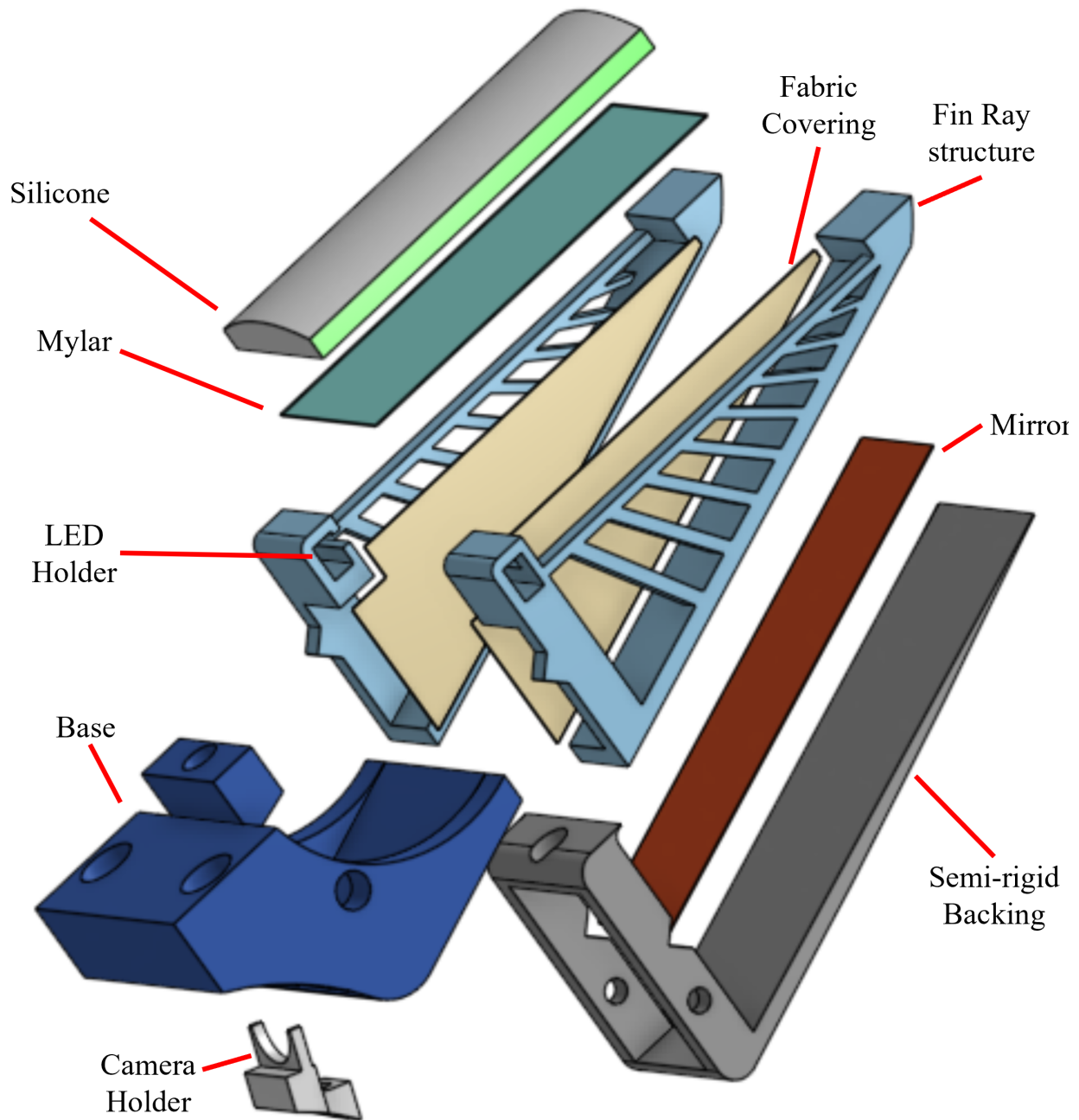


Figure 5.3: An exploded view of the Baby Fin Ray assembly. Not pictured are the camera, LEDs, and the diffuser [15].

- Illumination & Paint

Solving these problems also gives us the future potential to incorporate some or all of these methods in the development of other soft sensors or robots integrated with high-resolution camera-based sensors.

**Design for Usability** Despite the ability of the original GelSight Fin Ray to perform touch-based manipulation tasks, such as wine glass reorientation, it is not able to robustly perform

many different types of tasks. Firstly, the LEDs are attached to the deformable acrylic piece that houses the silicone gel pad, meaning that if the Fin Ray grasps a large object or grasps it incorrectly, the LEDs can potentially be dislodged over time. Another issue is that the silicone gel pad is extended to the tip of the sensor, meaning that if we wanted to use the Fin Ray to grab an item out of clutter, the gel pad could potentially be dislodged. Finally, the long length of the Fin Ray (110 mm) makes it impossible to be viewed by a single camera.

As such, we halve the length of the Fin Ray design so that it is coincidentally around the size of a human finger. This length reduction serves two purposes: (1) the entire front of the sensor can be visible by a single camera; (2) one single illumination source is needed for the entire sensing region.

We also place the blue LEDs (Chanzon 2835 SMD LEDs) at the base of the finger, protectively encased inside the flexible structure. Doing so further allows us to place a piece of VViViD Air-Tint Dark Air-Release Vinyl Wrap Film over the LEDs as a diffuser, which helps to prevent bright spots in the sensing region that were prevalent in the original design.

The silicone gel pad is recessed into the Fin Ray structure so that there is less of a non-smooth interface between the tip of the Fin Ray and the silicone gel pad. We also sharpen the Fin Ray tip to somewhat emulate a fingernail. These changes serve to make the Baby Fin Ray more useful for a different variety of tasks, such as digging through a cluttered, obstructed environment to grab an object.

**Camera placement** One of the most prevalent issues with the GelSight Fin Ray is the large obtrusive camera on the backbone. Although this camera provides high-resolution wide-angle imaging, it also prevents the back of the Fin Ray from flexing. This obstruction makes it difficult for the overall structure to then comply to objects.

A way to resolve this issue is to use a smaller camera that then corresponds with a smaller Fin Ray structure. However, this solution does not completely solve the problem of introducing unnecessary rigidity in the back of the structure, so we turn to the usage of flexible mirrors with inspiration from GelSlim and the GelSight Wedge [31], [32].

In particular, we attach a 0.2 mm PET sheet to the inner back surface of the Fin Ray structure and place a camera at the base of the structure. The camera is angled in a way such that it can unobtrusively see the entire tactile sensing surface using the reflective surface of the PET sheet. Additionally, because the PET sheet is so thin, it can conform with the flexibility of the back structure, and we can see in Fig. 5.4, using CAD software, that we only need to utilize a 100° field-of-view camera to see the entire tactile sensing region.

As such, the camera that we use for our design is a Raspberry Pi Zero Spy Camera with a 120° field of view. This camera is small enough to fit within the base, and it also provides a suitable viewing range for our tactile sensing region. Then, a yellow filter (Rosco E-Colour 765 Sunlight Yellow) is placed under the lens of the camera to help filter out the blue light, which could overpower the rest of the colors in the sensing region.

**Illumination & Paint** Another issue we had to resolve was the slightly rigid acrylic piece in the original design, which serves to house the paint and provide a semi-rigid, deformable tactile sensing surface. The acrylic piece allows the acrylic fluorescent paint to deform with and illuminate the Fin Ray structure without the paint delaminating from silicone. This delamination occurs because the acrylic paint and silicone do not bond well together, necessitating the addition of an acrylic substrate. However, this addition increases the

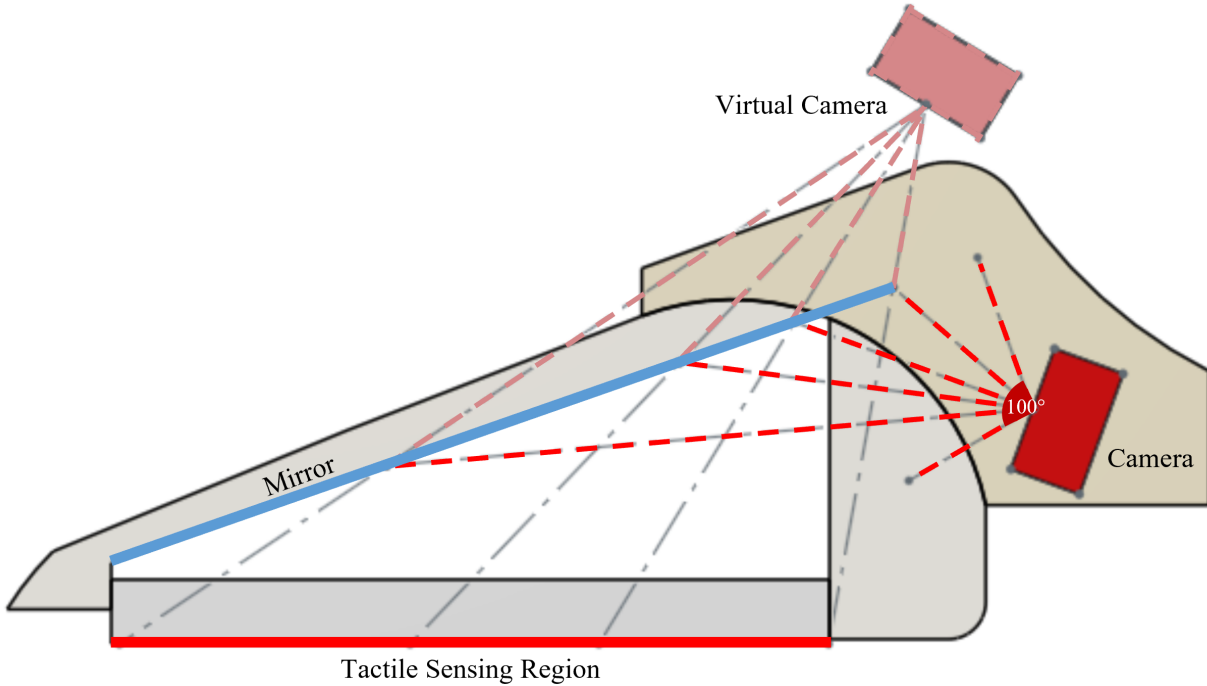


Figure 5.4: The 2D representation of a  $100^\circ$  field-of-view camera viewing the tactile sensing region through a mirror and its corresponding virtual camera. As the ray tracing shows, the camera is able to see the entire tactile sensing region while sitting at the base of the finger, where its rigidity will not impede with the deformable Fin Ray structure [15].

rigidity of the finger.

We choose to completely eliminate the need for such an acrylic piece for this improved design. Not only does this improve upon the flexibility, it can also provide potential design strategies for non-LED illumination systems to other soft camera-based sensors.

To be able to create a flexible, fluorescent paint that can bond to silicone, we use two different types of paint: a commercial fluorescent silicone pigment and a silicone paint of our own design. We choose to focus on adding as much pigment as possible to the fluorescent silicone mixtures so that we can have a more vivid tri-color sensing region. As a result, we do not optimize for the elongation, tensile strength, or other potentially relevant properties of the silicone paint.

For the commercial fluorescent silicone pigment, we use Smooth On's Silc Pig Electric Green and Pink pigments. Following the instructions provided, we add only 3% of the total silicone system weight to our mixture of XP-565 and NOVOCS Gloss (11 to 3), before we stir it, degas it, and paint it onto the long sides of the silicone gel pad in multiple layers. The NOVOCS Gloss is added to the mixture to help thin it out to make it easier to paint onto the silicone piece. The silicone gel pad is then placed into the oven to cure for 4 hours at  $52^\circ\text{C}$  ( $125^\circ\text{F}$ ).

In an attempt to create a customizable flexible paint, which could work with multiple types of acrylic paint, we also synthesize a paint using a silicone adhesive base. In a small

container, we measure out 0.5 grams of acrylic paint (Liquitex Basics Acrylic Paint Fluorescent Green and Red) and 1.5 grams of A564 medical adhesive silicone (Factor II, Inc). Because silicone adhesive cures faster with moisture and acrylic paint is water-based, we make sure that the silicone adhesive and paint are not in contact with one another as we add both of them to our container.

Once both have been added, we vigorously stir until the mixture is mostly homogeneous and immediately add 1.0 gram of NOVOCS Gloss to thin out the mixture while constantly stirring. After the clumps of acrylic paint and silicone adhesive have dissolved into the NOVOCS Gloss, we slowly add 3.0 more grams of NOVOCS gloss in 1.0 gram increments, stirring after each addition. Doing so allows the paint to slowly thin out until it can be more easily applied to the sides of the silicone gel pad. To avoid large clumps in the paint, the mixture is strained over a 190 Micron filter (TCP Global) and then multiple layers of the paint are applied onto the silicone gel pad sides. The mixture is left to cure at room temperature for 30 minutes. We also note that this exact same procedure can be done with Silpoxy (Smooth On), and can potentially be performed with other silicone adhesives. However, we choose the A564 for its clearer coloring as opposed to other silicone adhesives that have more of a hazy quality, which we believe could limit the paint fluorescence.

### 5.1.2 Software

The camera is connected to a Raspberry Pi and video is streamed using mjpg-streamer. For utilizing the data for tasks, we use a warp perspective function on the visible mirror in the image to convert it into a rectangular image with the correct ratio of the actual flexible mirror. This accounts for slight differences in manufacturing that might lead to changes in the "negative" space around the flexible mirror and tactile sensing region.

## 5.2 Analysis

### 5.2.1 Simulation

To improve the compliance of the original GelSight Fin Ray and the robustness of hardware when doing manipulation, we make a more compact design by removing the acrylic piece and by changing the camera location. In order to prove that the new design has better compliance, we did a series of FEM simulations using ABAQUS software. We used continuum shell elements for thin sheets like the Mylar and flexible mirror, while we used eight-node linear brick elements for the main structures of the Fin Ray, e.g. silicone, semi-rigid backing, base, and the Fin Ray structure. Because of the large displacement, a non-linear analysis step was adopted to simulate the complex jamming behavior between the Fin Ray ribs.

As shown in Fig. 5.5, we make a comparison between the Baby Fin Ray and the "original" Gelsight Fin Ray. The "original" design is a smaller and comparable version of the previous original GelSight Fin Ray with the rigid acrylic piece inside the gel pad and the camera on the backbone. A cylinder indenter (shown in Fig. 5.5 a) and a cuboid indenter (shown in Fig. 5.5 b) are pressed into two fingers at two different locations, middle finger pad (lower location) and fingertip (higher location). Fig. 5.5 c and d show the Force-displacement curves of the

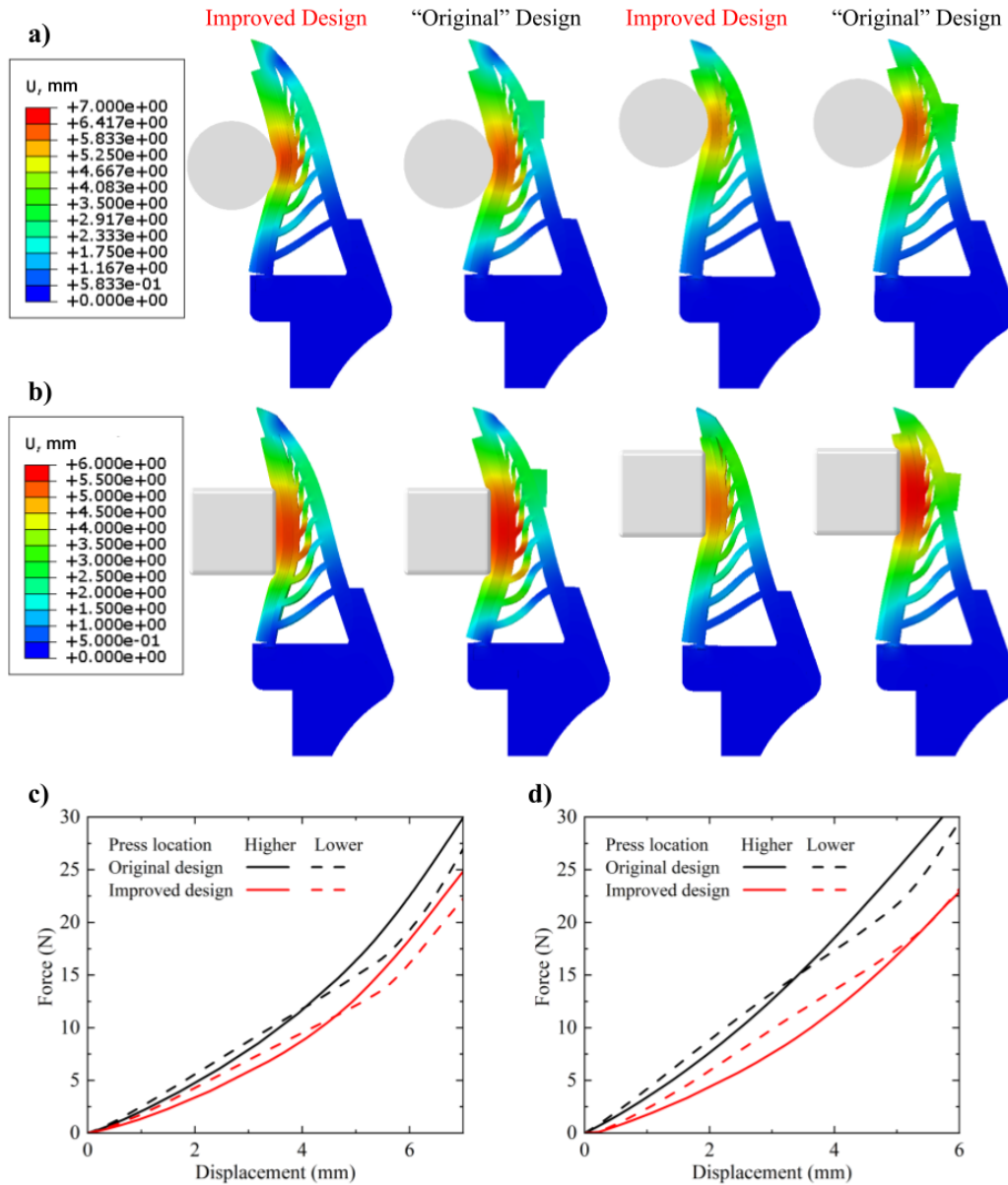


Figure 5.5: FEM simulation results of the improved design and the “original” design, which is a smaller version of the previous GelSight Fin Ray for the sake of comparison. The indentations happen at the middle finger pad (lower location) and the fingertip a) Displacement fields of cylinder indentation; b) Displacement fields of cuboid indentation; c) Force-displacement curves of the improved and “original” fingers with cylinder indenter; d) Force-displacement curves of the improved and “original” fingers with cuboid indenter [15].

two fingers pressed by the cylinder indenter and the cuboid indenter respectively. Depending on the shape of the indenter, indenter location, and depth, the Fin Ray fingers can display very different stiffness or compliance, but the improved design in general needs less pushing force to get the same amount of indentation depth. Overall, the Baby Fin Ray is 18-32% more compliant than the original design.





Figure 5.6: Fluorescent silicone paint dogbone samples used for tensile testing on the Instron meeting. From left to right, we have our pink Silc Pig, pink synthesized paint, green Silc Pig, and green synthesized paint samples [15].

### 5.2.2 Fluorescent Silicone Paint

To analyze our synthesized fluorescent materials and compare them with one another and to the original acrylic paint on an acrylic piece, we performed tensile testing and created our own metric to analyze the illumination and 2D localization through reconstruction of each lighting scheme.

**Tensile Testing** Tensile testing was performed on the two types of synthesized silicone-based fluorescent paint. We chose not to perform this test on the dried acrylic paint since acrylic paint is brittle once dried.

Test samples were prepared in a dogbone mold based on the ASTM D412 standard. The molds were printed with Onyx material on the Markforged printer. A laser cut 0.15 mm Mylar piece was placed on the bottom of the mold and an acrylic piece was clamped on the top of the mold to enforce an even thickness across the entire dogbone specimen. After preparing the samples mixtures, the mixtures were poured into the molds and left to cure per the paint preparation instructions. Some examples of our samples are shown in Fig. 5.6.

After curing, all of the paint samples experienced some shrinkage. Although the dimensions of the dogbone profile of each piece was close to the others (sub-millimeter differences), the thicknesses varied a bit more (1 mm differences). As such, the dimensions were re-measured after curing and right before the dogbones were put into the tensile testing machine. Testing was performed on an Instron machine and samples were stretched until they broke. Results were then normalized to the thicknesses of the samples.

**Illumination/2D Localization** To test the integrity of our illumination schemes using different paints, we compared the reconstruction outputs (2D localization) of each paint illumination pattern against our “base case”: the acrylic paint on a deformable acrylic piece.

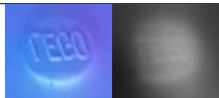

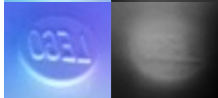
We created three different tactile sensing samples and placed them on a Baby Fin Ray. Three different objects, a 5 mm by 10 mm rectangular block, a 4.75 mm diameter ball bearing, and a Lego block piece were then lightly pressed into the three different sensing

regions. Corresponding reference images, images where no tactile imprint was given to the sensing region, were used to obtain a difference image that only showed the tactile imprint of our objects. This difference image was then used to obtain a reconstruction image, which we used as a 2D localization metric for comparing the position of the pressed object with the image reconstruction to determine the paint fidelity with respect to a GelSight sensor.

This comparison was made by first manually segmenting the ball bearing and rectangular block. We then used a distance error metric for the center of the ball segmentation and utilized the Dice coefficient metric to compare the similarities of the rectangular segmentation image with the 2D localization results [63]. A spherical object and rectangular object were chosen to best represent the lighting illumination changes where both a smooth and a discrete object were pressed into the sensing region. In particular, we chose a spherical object to see if the lighting illumination was uniform enough for our reconstruction to determine the appropriate placement of the circle center. The Lego block piece was chosen to determine, in a qualitative way, how finely detailed the reconstruction image could be. However, for many manipulation tasks, it is unnecessary to have extremely high-resolution sensing that would allow us to see the tiny words on a Lego block piece, which are approximately 0.2 to 0.3 mm in width.

## Paint Analysis Results

Table 5.1: Fluorescent Paint Comparisons [15]

Name	Lego Pic. & Reconstruction	Circle Seg. Error	Rect. Dice Score	Elongation @ Break	UTS
Acrylic on Acrylic		3.6 pixels	72.3%	N/A	N/A
Acrylic + Silicone Adhesive		1.4 pixels	85.0%	1,212%	1.37 MPa
Silc Pig Pigment		3.6 pixels	75.0%	267%	0.23 MPa

As shown in Fig. 5.7, we see that both of the synthesized (acrylic) paints outperformed the Silc Pig based fluorescent paints in percent elongation at breakage and their ultimate tensile strengths (UTS). Whereas the green and red acrylic paints gave us a percent elongation of 1,348% and 1,212%, and a UTS of 1.37 MPa and 1.65 MPa, respectively, the corresponding values for the Silc Pig paints were 267%, 269%, 0.23 MPa, and 0.26 MPa.

We believe that these differences in values between the two types of paint could in part be due to the extensive amount of Silc Pig pigment we put into the silicone mixture. Although the pigment is synthesized to work well with and be miscible with silicone, the pigment has the potential to disrupt crosslink formation of the silicone, which would cause the silicone material to have lower percent elongation and ultimate tensile strength values.

This phenomena could have also occurred with the addition of acrylic paint to the silicone adhesive. However, we believe that because silicone adhesive is designed to adhere to many different types of surfaces, disruption of crosslink formation would not have had as severe of an effect on the integrity of the synthesized paint.

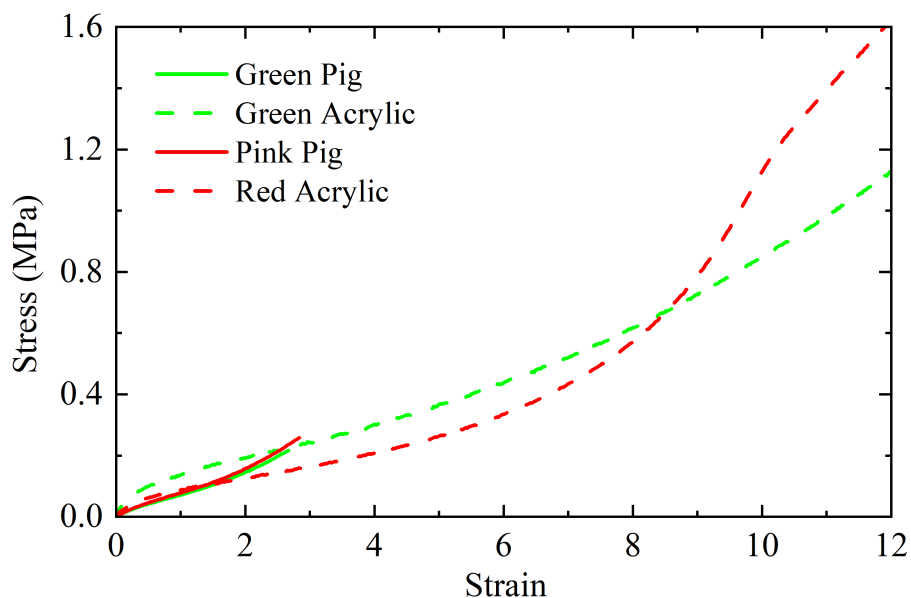


Figure 5.7: Stress-strain curves of the fluorescent silicone paint, where the acrylic paints represent the synthesized silicone adhesive based paints. The graphs have been cut off at the point of breakage of the dogbone tensile testing pieces [15].

Another interesting effect we noticed was that the acrylic paints had different strain-stress curve behaviors. Although both curves behaved in similar ways at the beginning of the tensile test, the red acrylic paint began to exhibit a nonlinear behavior before breakage. We believe this behavior can be attributed to the material differences of the red and green acrylic paints from Liquitex and how miscible they are with the NOVOCSS Gloss and silicone adhesive mixture. The Silc Pig mixtures, on the other hand, behaved similarly most likely because these pigments were manufactured to perform in this way.

With regards to illumination schemes, we found that although the acrylic paint on an acrylic sheet visually performed equivalently with the other two at discerning finer details, it was not better than the other two paints at 2D localization. In fact, the silicone adhesive acrylic paint mixture slightly outperformed all of the paints, but not significantly so. The reconstruction algorithm gave a 1.4 pixel error for the circle centers and a 85.0% Dice score for the rectangular reconstruction segmentation image, compared to 3.6 pixel errors for both the Silc Pig and acrylic on acrylic and a 75.0% and 72.3% similarity coefficient, respectively. A comprehensive list of our results is shown in Table 5.1.

Overall, our metrics show that there is not a significant difference between the different types of illumination schemes. This metric could be different if we were to include the reconstruction normal gradients; however, we found that including normal gradients was not useful for many applications and that 2D localization was an acceptable alternative.

We also note that for the acrylic on acrylic illumination scheme, there was more of a uniform lighting scheme throughout the entire sensor. Specifically, the tip of the sensor for both the Silc Pig paint and the synthesized paint had a slightly dimmer illumination at the tip compared to the base of the sensor. The nonuniform illumination, which did not affect our experiment results, is most likely due to the lower index of refraction of silicone compared

to acrylic. As a result, more of the blue light will have a higher chance of refracting out of the silicone material instead of experiencing total internal reflection within the material, as it did with the blue LEDs pressed against an acrylic piece.

## 5.3 Experiment

### 5.3.1 Nut Classification

To leverage the new, improved hardware modifications, we chose to utilize the Baby Fin Ray in an object “sorting” task, using nuts that were still in their shells. This task shows how the new sensor can classify objects based on their shape and texture.

For our task, we chose to use almonds, Brazil nuts, pecans, and walnuts, which all have semi-distinct shell textures and are of various shapes and sizes. We collected 500 tactile images each of the four different nuts, pressed against different parts of the sensor in various orientations and with flipped lighting configurations. The grasps were guided by hand. Next, the images were processed via an image unwarping process and the dataset was split into 80% for training and 20% for validation, with data augmentation applied on the training dataset. We used a Resnet-50 architecture, shown in Fig. 5.8, for training on the unwarped images with four total classes representing the different types of nuts [64]. For our neural net, we used the stochastic gradient descent optimizer with a learning rate of 1e-3, and a learning rate scheduler with a step size of 7 and a gamma of 0.1.

We also attempted to train the same neural network on raw images instead of on our warped images.

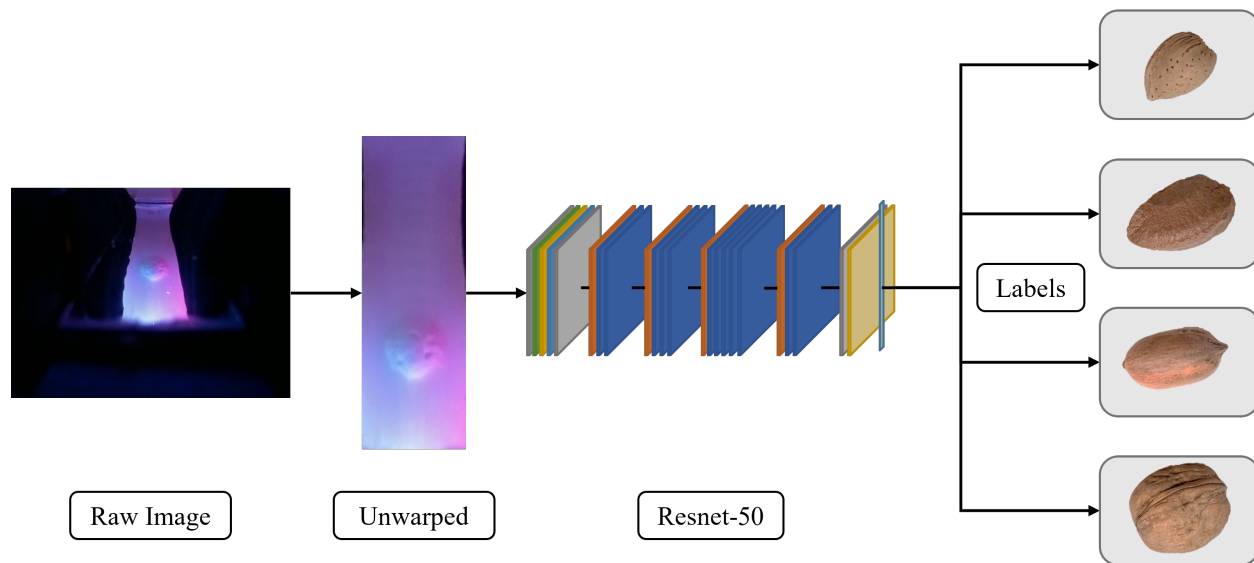


Figure 5.8: The architecture used for training our nut classification task. We take the raw image from our tactile sensor, unwarped the mirror image, before feeding it into a Resnet-50 architecture with four classes: Almond, Brazil Nut, Pecan, and Walnut [15].

### 5.3.2 Results

Without unwarping, our neural network was able to achieve 99.5% accuracy, while with unwarping, our neural net achieved a 95.8% accuracy on the validation set. We believe that this slight discrepancy is due to the amount of “negative” space in our unprocessed images. Although it was not immediately visible to our eyes, there are potential discrepancies in the black covering of the Fin Ray when different sized objects are being grabbed. As such, the “negative” space might lead to higher accuracy for classifying the nut shells over this dataset, but could potentially have a lower accuracy on sensors that might be manufactured slightly differently.

Overall, our unwarped neural net had 98.1%, 96.8%, 96.9%, and 91.3% accuracy on classifying almonds, Brazil nuts, pecans, and walnuts, respectively. We also found that the Baby Fin Ray was able to both conform well to these differently sized nuts and was robust enough to withstand multiple digging attempts through our bowl of nuts. Not only did the synthesized paint stay bonded to the silicone, the design changes also allowed us to have more compliance and gave us more repeatable lighting results in between multiple grasps. The tapered end of our structure also means that the Baby Fin Ray could be used for digging tasks. These changes overall make the Baby Fin Ray more usable in manipulation tasks than its predecessor.

## 5.4 Discussion

The ability to leverage both the compliance and flexibility of soft robots with the intricate sensing afforded by camera-based tactile sensing allows us to perform more manipulation tasks. In this paper, we show an updated version of the GelSight Fin Ray, which not only encompasses both abilities, but has been improved so that it can potentially perform robustly in manipulation tasks.

We increased its robustness by making hardware changes to the GelSight Fin Ray, including an update to the camera placement and a revamp of the illumination design for better compliance. Additionally, we specifically designed the finger for digging through clutter tasks. Our analysis of the different fluorescent paint illumination schemes also shows us that although there is not too much difference between the different schemes, our synthesized silicone adhesive and acrylic paint mixture showed the highest elongation at breakage and ultimate tensile strength. These properties show promise for the synthesis of other types of soft grippers with proprioception or tactile sensing, since there are many soft grippers that either expand or are constantly experiencing some form of deformation.

Finally, we showcase that these new design implementations and sensing capabilities allow us to repeatedly grab nuts with the conformability of the Baby Fin Ray, and then use the camera-based sensor to very accurately classify between the various distinct textures. These abilities opens up avenues for us to, in the future, perform complicated digging tasks through cluttered household environments to retrieve specific objects. We also keep in mind that, in the future, adding markers like the force tracking ones in the original GelSight Fin Ray, would allow us to perform incipient slip analysis and also allow us to potentially perform proprioception.

Overall, we show strategies to integrate more intricate, highly-detailed sensing into soft, compliant robots, and to do so without jeopardizing the original flexible and conformable capabilities of the soft robot. In the following chapter, we implement some of these manufacturing ideas into the development of the GelSight EndoFlex, which is based in anthropomorphism and designed from the ground up to include continuous tactile frontal and side sensing.

# Chapter 6

## GelSight EndoFlex<sup>\*</sup>

To incorporate continuous tactile sensing into a soft finger and hand design, we built a hand from the ground up, incorporating cameras and LEDs into an endoskeleton, which also serves to improve the strength of the finger. Concurrently, we wanted frontal and side sensing that could be provided by the cameras.

The ability to identify an object using a single grasp is important and requires “complete” sensing along the grasping surfaces of a finger. Even though many current finger-inspired sensors can perform object recognition well with high-resolution finger tip sensors or with low-resolution larger tactile sensors, they either require that the object be in full contact with the finger tips or multiple regrasps to classify the object within the hand [65], [66]. Furthermore, they do not have the compliance afforded by soft robotics, which can greatly improve secure grasping abilities of the gripper or make them safer for interaction with the world around them.

In other words, soft robotic manipulators could greatly benefit from having structural compliance and rigidity, along with high-resolution sensing of tactile sensors. To this end, we present the following contributions:

- A novel design of a continuous high-resolution tactile sensor along a doubly curved surface;
- An endoskeleton finger design for a human-inspired gripper that incorporates tactile sensing (Fig. 6.1);
- A neural net that can utilize the tactile images from just a single grasp to classify objects.

---

<sup>\*</sup>Most of this chapter is copied from the paper GelSight EndoFlex: A Soft Endoskeleton Hand with Continuous High-Resolution Tactile Sensing [16].

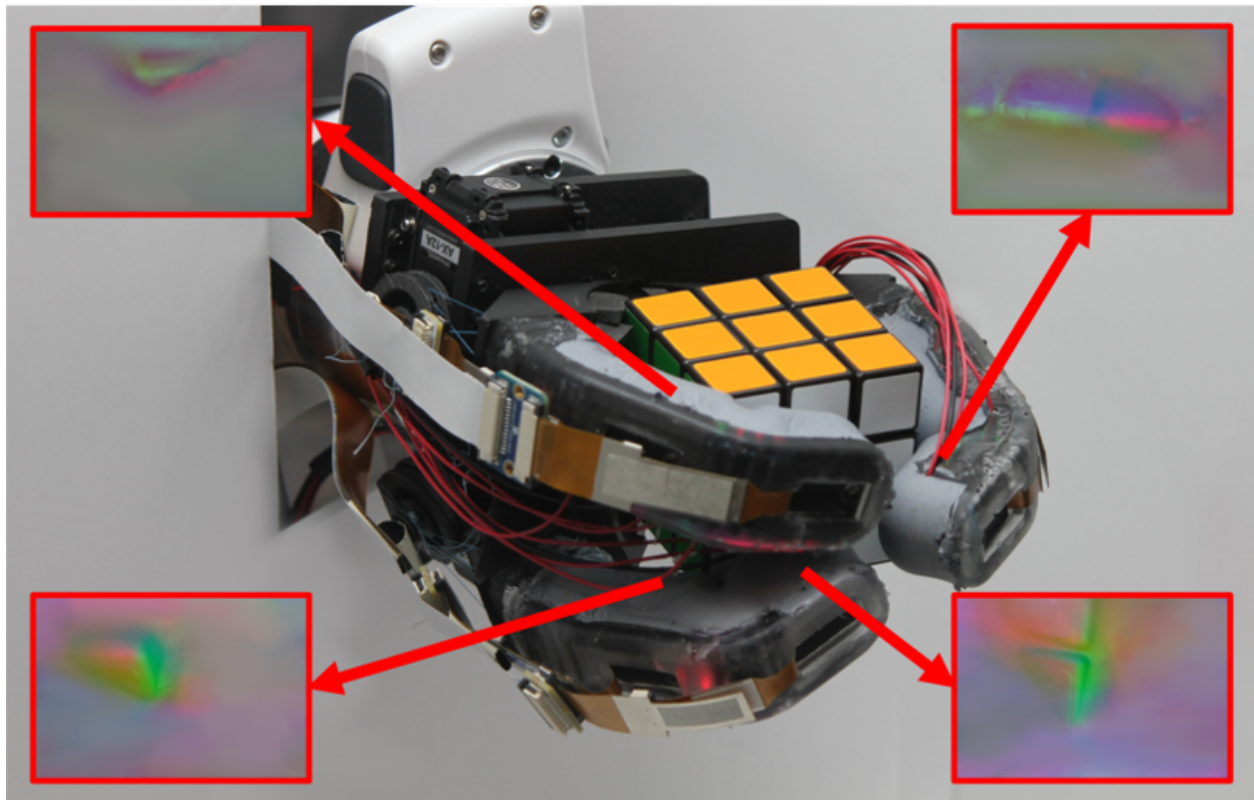
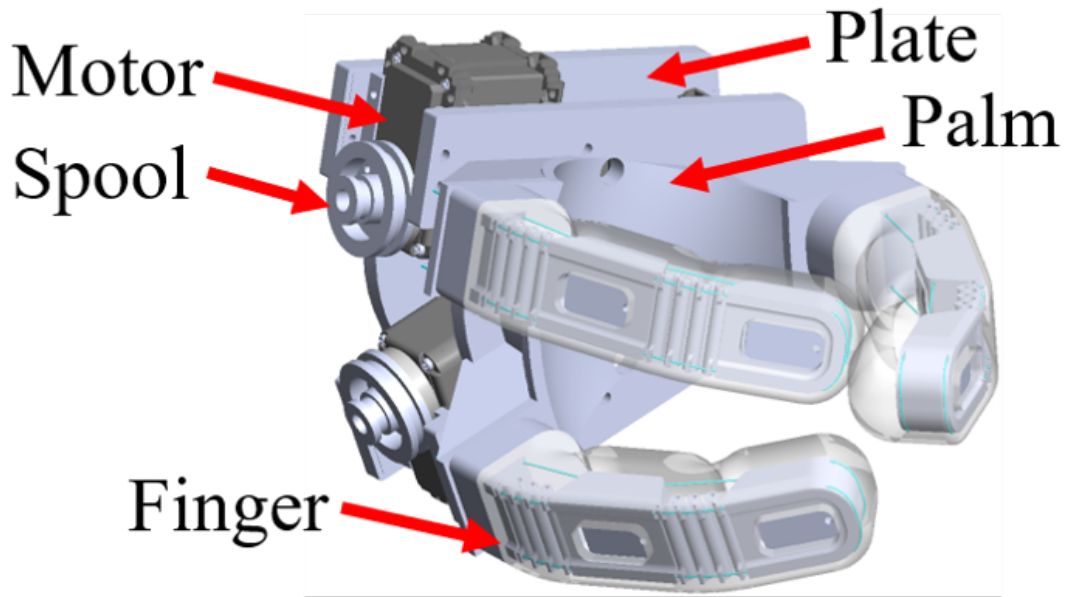


Figure 6.1: **Top** A CAD model of our GelSight EndoFlex gripper with some of the parts labeled. **Bottom** The GelSight EndoFlex is securely grasping a Rubik's cube and the corresponding processed difference images of four of the six sensing regions are displayed. Of note is that the bottom two sensor images are showing continuous sensing along the side and corner of the cube, while the top two sensor images are showing one image each from the other two fingers [16].



## 6.1 Methods

### 6.1.1 Hardware

The EndoFlex sensor is composed of an endoskeleton encased in silicone with two embedded cameras for continuous sensing (Fig. 6.2). Each endoskeleton was designed to be one continuous piece with a pair of rigid segments and flexures to form joints. This design minimizes the number of parts required to fabricate one finger when compared to traditional rigid fingers. The flexure design was chosen for its high compliance and low deformation of individual elements to reduce silicone delamination. We 3D printed the endoskeleton using an Onyx One printer with Markforged Onyx plastic for its combination of high strength and relatively low tensile modulus when compared to other extruded plastics. This combination of properties allowed minimal force loss during actuation.

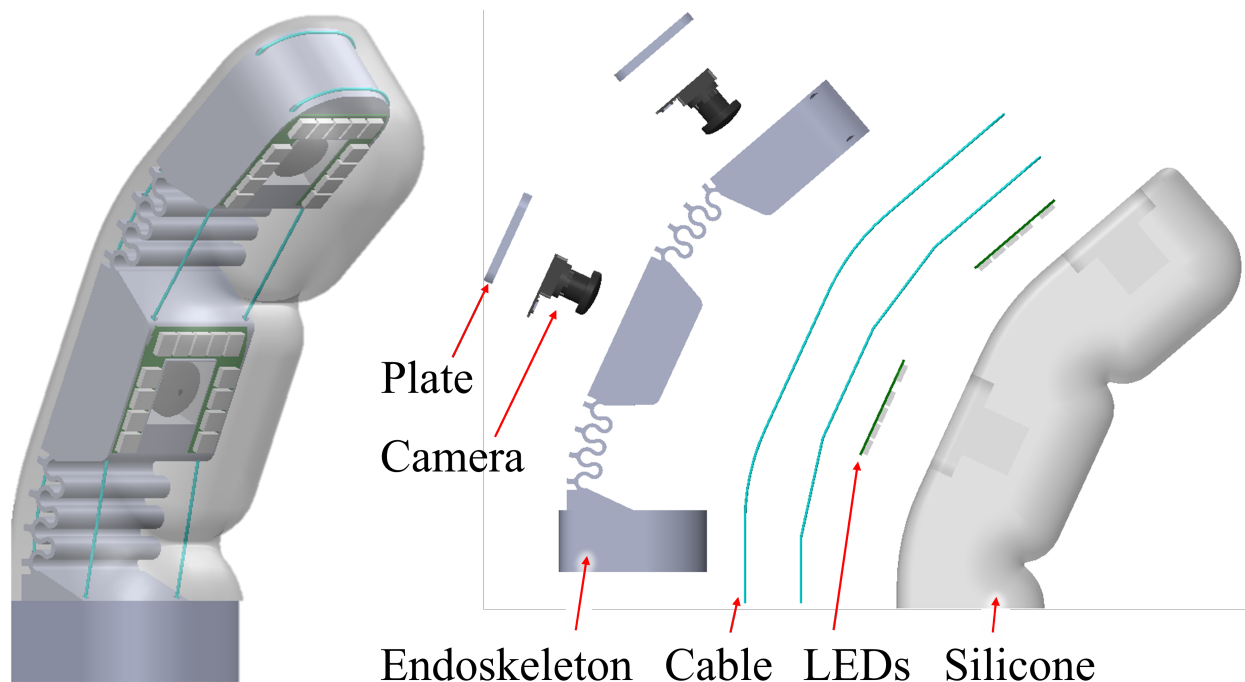


Figure 6.2: A close-up view of an EndoFlex finger with an exploded view. Each finger operates independently with one degree of freedom and can be quickly replaced if damaged [16]. A diagram of the continuous sensing region is shown in the Appendix Fig. B.5.

A camera was mounted into each endoskeleton segment to prevent any shifting during actuation. Three sets of red, green, and blue LEDs were mounted with cyanoacrylate adhesive onto the rigid segment of the endoskeleton. They were spaced 90 degrees apart to create a colored light gradient for the GelSight algorithm. Finally, the endoskeleton was threaded with Piscifun Onyx Braided Fishing Line soaked in Smooth-On Universal Mold Release to reduce friction when cast in silicone. We chose to use cable-driven actuation to reduce potential camera-view obstructions and also so that we could more easily integrate the camera into the finger skeleton.

A rigid three finger palm was designed with temporary fasteners to allow for fast replacement of damaged fingers or for future iterations. Fingers were positioned in a ‘Y’ pattern with two fingers and an opposing thumb. The pair of fingers was spaced thirty degrees apart to distribute grasping force without creating collisions. The palm was designed to have a “cupped” shape with a polyurethane foam layer to add grasping ability. A separate rigid plate was designed as a mount for the hand to sit on the Panda robotic arm. Three Dynamixel AX-12A servos were mounted between the plate and the palm and served as the actuation method for the fingers through double axle spools. The double axle design allowed for actuated contraction and extension of each finger. The palm, plate and spools were all printed with Markforged Onyx plastic using a Markforged Onyx printer.

As part of our finger manufacturing process, which is fully shown in Fig. 6.3, a two-part mold is designed for casting silicone to create the optically clear medium for the GelSight sensor. The mold was designed to hold the endoskeleton during the casting process which removed the need for fasteners or adhesives to hold the silicone layer. The mold had high curvature to create a rounded finger, much like a human finger. One major benefit of the curved surface was the high reflection of lights within the silicone which aided in sensing by removing shadows of pressed objects. The mold design removed any air gap between the camera lens and cast silicone to minimize the refraction of light. The mold was produced using a Formlabs 2 SLA printer for its high resolution. To achieve the optical clarity required to use GelSight, the mold was incrementally sanded with sandpaper starting at 800 grit and eventually reaching 2000 grit. No additional polishing or coating step was added as sanding with 2000 grit sandpaper produced a clear surface finish in the tactile sensing region.

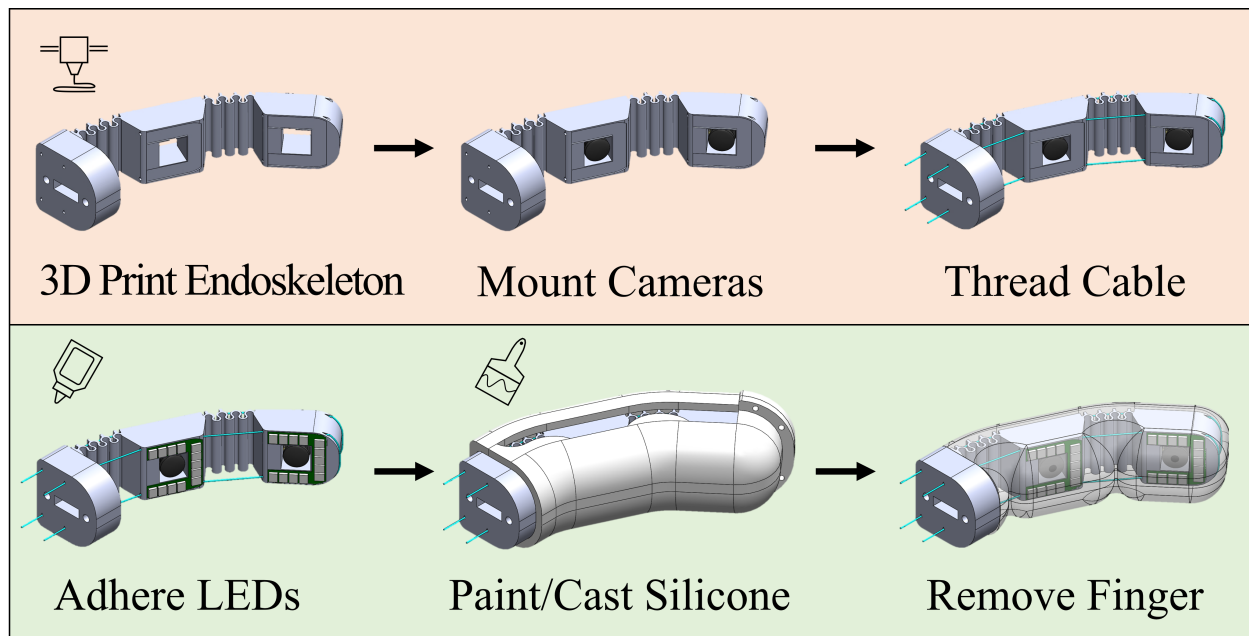


Figure 6.3: The manufacturing process for the EndoFlex sensor including assembly of electronics and casting of silicone [16].

To allow the silicone to compress when the tendons pulled the endoskeleton finger to a closed grasp position, we chose to synthesize a softer silicone for the finger. As a result,

we used a ratio of 1 to 15 to 5 parts of XP-565 parts A and B, and plasticizer (Phenyl Trimethicone, Lotioncrafter). Decreasing the ratio of part A to B for the XP-565 is equivalent to adding less catalyst, which increases the softness of the silicone, while the addition of the plasticizer also causes the resulting cured silicone to have a softer texture.

Before pouring the silicone mixture into the mold, we used a paint brush to paint a thin layer of Inhibit-X (Smooth-On Inc). After waiting a few minutes for it to dry, we sprayed a layer of Ease Release 200 (Smooth-On Inc) on the mold. To create the sensing surface, we combined 5 parts 4  $\mu\text{m}$  Aluminum cornflakes (Schlenck) with a mixture of 22 parts silicone ink catalyst and gray silicone ink base (Raw Materials Inc.) and 60 parts NOVOCSS Gloss (Smooth-On Inc), and mixed it for a minute using an ultrasonicator. This mixture was then sprayed into the inside of the top mold with an airbrush and left to dry for at least 10 minutes before we fit the threaded endoskeleton inside of the mold and screwed the mold halves together. Remaining holes and the lips of the mold were covered in a thin layer of brushed-on silicone adhesive (Devcon), which created a seal for the mold and prevented any silicone leakage outside of the mold that could be caused by mold warping or other printing imperfections.

Once the main body silicone mixture had been degassed, we slowly poured the mixture into the prepared mold. The entire mold assembly was placed on top of a vibrating plate for 10 minutes to get rid of any bubbles in the camera-viewable areas. These bubbles may have been induced by the silicone pouring over the flexures, electronics, and other 3D printed parts inside of the mold. Some of the bubbles were retained along the side of the sensor surface, which is not viewable by the camera and did not negatively affect the sensor integrity.

Finally, the mold was placed inside of an oven at 125°F (52°C) for 12 to 15 hours. This temperature was chosen to prevent any of the electronics or inner structures from reaching their glass temperatures and causing delamination of the parts from the silicone. Once the finger was removed from the mold, the gray sensing membrane surface was no longer smooth and instead had a reticulated wrinkled texture (Fig. 6.4). This phenomena only occurred when we sprayed the paint on the mold first and did not occur if we chose to cure the finger first without the paint in the mold and spray the paint on the finger surface afterwards.

The modular fingers were then placed on our palm plate to create our completed gripper. We also note that this configuration can be changed to enable different types of grasps, although we chose an enveloping grasp to maximize the amount of sensing the gripper could obtain from grasping an object in its palm.

### 6.1.2 Software

Each finger was equipped with two Raspberry Pi Zero spy cameras with a 160° field of view, for a total of six cameras. All of the cameras were able to view a curved segment of the finger, which was illuminated by tri-directional LEDs. The finger segment images were individually streamed using the mjpg-streamer package and can be processed using OpenCV and fast poisson solver [60], [62] to get difference images and uncalibrated reconstruction images, as shown in Fig. 6.5.

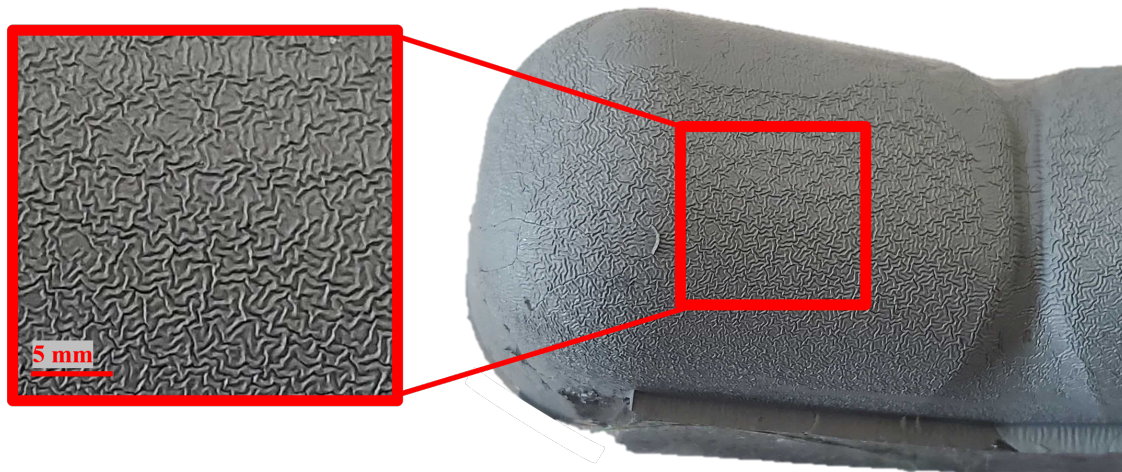


Figure 6.4: A close up image of the reticulated wrinkle surface of the GelSight EndoFlex sensor. The width of one of the wrinkles is approximately 0.4 mm wide and was only created when we first sprayed the paint on the mold surface before casting silicone inside [16].

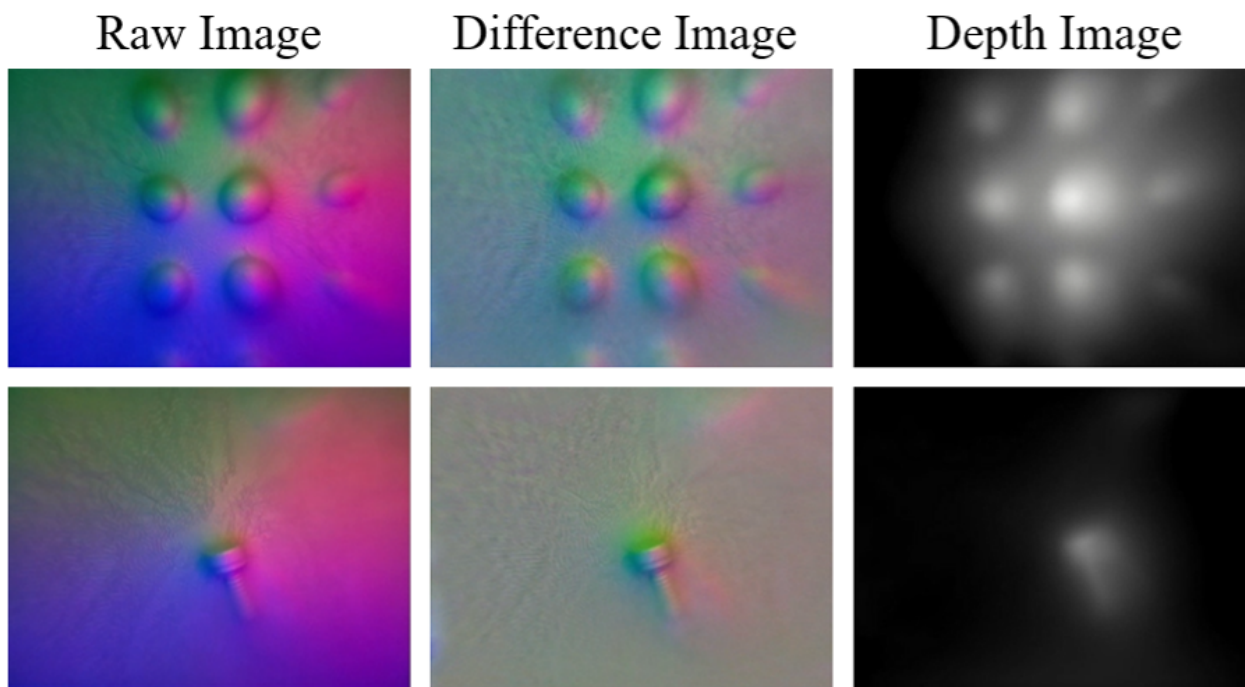


Figure 6.5: From left to right, we have raw sensor images of a 3.75 mm ball bearing array and a M2 screw, followed by their difference images from a reference image (no tactile contact), and the corresponding uncalibrated depth image [16].

## 6.2 Experiment

To show the usefulness of having continuous sensing, we collected single grasps of various objects and performed a classification task based on the entire finger sensing region. Previous works show that object classification using finger tip sensing or low-resolution palm sensing is accurate, but only when the objects were in contact with the fingertips or multiple touches have been performed [65], [66].

Our grasping object set included three distinct objects from the YCB dataset: the Rubik’s cube, one of the toy stacking cups, and a plastic orange [67]. These three objects are shown in Fig. 6.6. For each object, we collected approximately 500 different grasps using all six of the cameras inside the fingers to obtain a holistic, “full-hand” tactile view of the entire object. To capture many different grasps, we had assistants manually reorient each object randomly such that it could still be feasibly grasped with the gripper, which allowed different parts of the sensor images to capture different features of the object that was being grabbed. We also attempted grasps utilizing a couple of the fingers instead of all of the fingers in the cases that the third finger did not have a solid contact with the object in its hand.

For each set of six images we captured, we stitched them together into a 2 by 3 array and used them as inputs for a Resnet-50 neural net architecture with the three outputs as the objects we used for our grasping data set [64]. We chose to use stochastic gradient descent as our optimizer, with a learning rate of 1e-3 and a learning rate scheduler with a step size of 7 and a gamma set to 0.1. We also implemented data augmentation on the entire set of images to deal with potential inconsistent lighting or random noise output of the images, and to account for eventual wear and tear in the silicone over time. We split our data into training and validation sets in a 80% to 20% ratio. The complete neural net architecture is shown in Fig. 6.6.

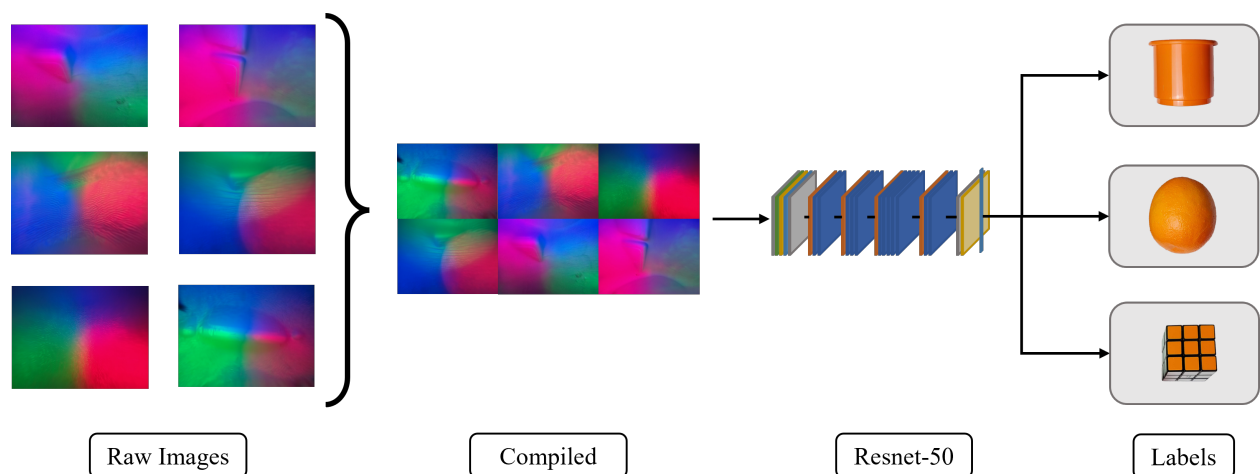


Figure 6.6: Neural net architecture for our single grasp classification. Once the object has been grasped, the six images are stitched together in a 2x3 array, thrown into our Resnet architecture and classified into a toy cup, an orange, or a Rubik’s cube [16].

## 6.2.1 Results

**Grasping** The GelSight EndoFlex was able to easily and very securely grasp all of the objects in our object set. In particular, the polyethylene foam layer on the palm provided a compliant, deformable surface that the grasped objects could be pressed against. The hand was also able to grasp empty water bottles without crushing them, as well as heavier objects, like a drill with a battery, without dropping them. As expected, the compliance of the soft gel allowed us to grasp more fragile objects, while the rigid endoskeleton allowed the fingers to withstand the force and weight of a heavier object.

Each finger was also able to bend to around  $60^\circ$  at each flexure point using the Dynamixel motors. Because the silicone was quite soft and because we added human finger-inspired grooves along the flexures, when the fingers bent, the silicone was able to more easily compress around the sides. However, the silicone still obstructed some of the bending angle, and as a result, the endoskeleton finger was unable to bend to its full  $90^\circ$  range that it would have been able to otherwise. Furthermore, deepening the grooves to facilitate bending would have limited the sensing area and ultimately interfered with the continuous sensing. Nonetheless, this limitation in motion did not severely limit the hand's ability to grasp objects because the deformable silicone surface over the endoskeleton finger helped to accommodate any loss of motion with its compliance and softness.

Casting the mold while the finger was in a slightly bent position helped to prevent creases in the surface of the silicone when the finger bent. Doing so also prevented silicone creasing when the finger was straightened out since the sensing surface was pulled in tension. Unfortunately, over time, pulling the silicone finger in tension caused parts of the silicone in the base of the finger to slightly tear. We believe that this problem could potentially be mitigated by using a softer silicone with higher elongation.

Finally, given a different arrangement of fingers or with a finger that could behave more thumb-like with an added degree of freedom, we believe that these fingers have the potential to grasp an even larger variety of objects.

**Tactile Sensing** As designed, the finger was able to continuously sense along the entire length of the finger when it was in a "closed" position. The fingers were also able to sense along the sides as well.

Overall, the finger was able to provide extremely high resolution sensing and the raw sensor images were able to capture details that previous GelSight sensors could sense, but with additional sensing coverage due to its rounded shape and the wider camera field of view. However, the wider camera field of view and the curved shape caused some distortion in the sensing image, which is most apparent on the sides of the image frame.

Additionally, some of the sensing surfaces appeared to have distinct rings of lights around the different color channels instead of the blending we would have expected from using a Lambertian paint on the surface of the silicone gel. We believe that this phenomena could have been caused by slight delamination of the silicone from the LEDs. The addition of the air interface will cause the light to refract from the air to the silicone face and potentially cause these rings of light to form and prevent even blending of the light within the silicone. In particular, we noticed that when objects were pressed against these sensing surfaces, the light circles began to dissipate. Nonetheless, this did not affect the sensor resolution and the

distinct features of the objects were still distinguishable as the tactile sensor had extremely high-resolution.

Finally, we noticed that the wrinkles, which were manufactured on some of the finger sensing surfaces, were helpful in preventing tears in the silicone membrane. Unlike the smoother sensing surfaces, it seems like the wrinkles helped to mitigate the high stress points caused by sharp corners poking into the sensor surface. The surfaces with wrinkles also felt like they had less friction than the smoother surfaces. Although the wrinkled surface made surface reconstruction difficult because the wrinkled texture appeared in difference images, they did not seem to negatively affect our object classification. The effect could also have been mitigated since we noticed that if enough pressure was put on the sensor surface, the wrinkles would smooth out slightly, which would not affect object classification results.

**Object Classification** Our object classification model was able to obtain 94.1% accuracy on our validation set. In live testing, which consisted of our robotic hand grasping the 3 objects ten times each, we were able to correctly classify 80% of the objects. The orange was able to be recognized 9 out of 10 times, while the classifier slightly struggled with distinguishing between the Rubik’s cube and the toy cup (80% and 70% accuracy, respectively). We believe that the discrepancy in the validation set results and the live testing results could be due to slight tears that developed over the course of the data collection and testing. Regardless, the hand was able to use only a single grasp to recognize the identity of an object.

As we expected, the orange, which had the most distinguishable tactile features, was the easiest for our model to recognize. Not only was the orange covered in a unique bumpy skin texture, it also had a distinctive stem portion. On the other hand, unless the fingers directly pressed against a corner of the Rubik’s cube or along multiple smaller cubes, it was hard to visually distinguish some of its edges from the edges at the bottom and top of the toy cup.

We believe that this confusion between the Rubik’s cube and the toy cup could be mitigated by adding a palm, which could also provide additional sensing. The added sensing from a larger area on the palm could have helped capture more tactile details that may have been missed by the fingers. Regardless, the object classification using continuous sensing along the multi-fingered hand was fairly robust and able to perform well on our object set. Specifically, it could be useful for grabbing objects in the dark or in an occluded environment where external vision would not be useful or could not be used. Analogously, humans can also capture distinguishing tactile features with an initial grasp, before performing additional exploration or using parameters such as friction, weight, and thermal conductivity.

## 6.3 Discussion

In this chapter, we present the novel design of a continuous high-resolution tactile sensor incorporated into a finger, which was then integrated into a human-like hand. The hand was then able to use these large sensing areas to be able to accurately classify objects using a single grasp. The ability to identify an object with a single grasp is akin to the way we as humans are able to grab an object with some priors and without external vision and determine almost immediately what we are holding.

Although recent research has focused a lot on large area low-resolution tactile sensors or high-resolution fingertip sensors for dexterous manipulation, not much research has been done on high-resolution sensing across the majority of a finger’s surface. Having this added sensing allows us to perform many useful classification tasks, and doing so in a soft, compliant gripper allows us to also safely and securely interact with objects and the surrounding environment. Sensors similar to the GelSight EndoFlex have the ability to be used for home-care robots or for human-robot interaction, where compliance and sensing are key to success.

Future work on this gripper involves adding a thumb-like joint, as well as full fingertip sensing, which can greatly improve the usability of the gripper for sensing and dexterous manipulation tasks. We can also continue to draw inspiration from GelSight sensors and add markers which could help track slip and shear or torsional forces along the surfaces of the finger. Overall, our novel endoskeleton finger design begins to solve the problem of designing human-inspired soft-rigid robotic hands with high-resolution sensing that are capable of performing more and more complicated tasks.

Also of note is that although this gripper is inspired by the human hand, it is unclear if this is the best form factor for a soft robotic gripper compared to grippers such as the Fin Ray effect inspired one. Furthermore, another interesting question about the resolution of tactile sensing is worth pursuing: do we need so much tactile acuity in our sensors? This uncertainty is another avenue worth pursuing as the soft robotics and tactile sensing community both continue to develop.



# Chapter 7

## Conclusion and Discussion

Over the course of this thesis, four different soft, compliant fingers were designed to incorporate high-resolution tactile sensing that is befitting of camera-based tactile sensors. In particular, we started out by showing that integrating a camera into a soft finger exoskeleton allowed us to use a single sensor to get both proprioception and exteroception through tactile sensing. From there, we developed new techniques to integrate such rigid GelSight-inspired sensors into both a Fin Ray gripper, which is better suited for grasping objects, and human-inspired gripper. More importantly, we showed that these soft fingers, similar to more rigid fingers, could do 3D reconstruction, classify objects, and track shear and torsional forces. These novel sensors even allowed us to perform different tasks such as wine glass manipulation, texture classification, and reaching into clutter.

Although the two fields of rigid-structure high-resolution tactile sensing and soft, compliant robots seem vastly different, this thesis begins to address this discrepancy and bridge this gap. Previously, many robotic grippers either had low-resolution force sensing with soft, compliant materials or high-resolution tactile sensing over a small, rigid area like a robotic fingertip. High-resolution tactile sensing can be useful in achieving more intricate manipulation tasks.

One important idea to consider is that there is still no rubric for designing soft, compliant hands, nor is there a philosophy for what type of tactile resolution sensing is actually necessary for many tasks. Despite this thesis attempting to address this gap, there is still no conclusive solution.

Even as I write this thesis, there are still countless of labs around the world working on various styles of hands: from anthropomorphic to octopus tentacles to a small array of delta robots that work in tandem to manipulate an object.

On the tactile sensing side, many researchers have grown to see its usefulness, but there is still a lot of reliance on utilizing external vision systems or wrist-mounted cameras. In many cases, such as a robotic hand digging through a dark, cluttered space, or even interacting with an object that has already been grasped and is occluded, having tactile sensing becomes vital to dexterous manipulation.

The idea of high-resolution tactile sensing is that one can always downsample or downsize the amount of information that is received. As such, being able to include even a camera into these soft, compliant systems without lowering their softness or compliance is important, and this work demonstrates the various ways we can approach this integration problem. From

the development of a flexible, strong fluorescent silicone paint to an integration of electronics into an exo- or endoskeleton, this thesis shows that the incorporation of this “rigid” high-resolution tactile sensor is possible, and provides utility in the ongoing pursuit of perceptive robotic hands.

## Chapter 8

# Addendum: Previous GelSight Fin Ray Designs

### 8.1 Optically Clear Fin Ray

One of the earliest considerations for the Fin Ray design was to create an optically clear Fin Ray structure. Having the structure be optically clear would allow for a camera to look through it while still maintaining the integrity of its fins.

To develop such a structure, we used 1 mm thick mylar as the base of the fins in the Fin Ray and filled the existing structure with a 1-to-1 mixture of Ecoflex Gel (Smooth On Inc) and XP-565 (Silicones Inc). Both of these materials have somewhat similar indexes of refraction, which we hoped would make it easier for a camera to see through the internal structure. We also used a 1-to-1 mixture of the two silicones because we wanted the resulting cured silicone to have a shore hardness between the gel and the XP-565.

A laser engraved piece of acrylic was used to rigidly hold the mylar pieces in place, before the silicone mixture was poured in between the struts. The base or bottom mylar piece had a circular hole laser cut out of it to hold an 18 mm diameter glass hemisphere for a camera housing. Once the mixture was poured in, after abiding by the manufacturing directions provided by the company, another symmetrical laser engraved piece of acrylic was placed over the mylar struts so that they would not flex and ruin the curing silicone. The resulting optically clear fin ray and the corresponding camera image is shown in Fig. 8.1.

Although the image inside the optically clear Fin Ray succeeded visually, with minimal internal reflection, the structure was too compliant. As a result, even though the structure could comply to an object pressed against it, the fin ray also flexed backwards, making it barely able to grasp any object (Fig. 8.2).

As a result, we reconsidered using a totally optically clear Fin Ray, and instead considered hollowing out the Fin Ray structures instead.

### 8.2 Early Hollow Fin Ray Prototypes

We considered three different hollow Fin Ray prototypes before settling on the final design shown in Chapter 4. The first prototype we attempted to use was one made of Onyx material

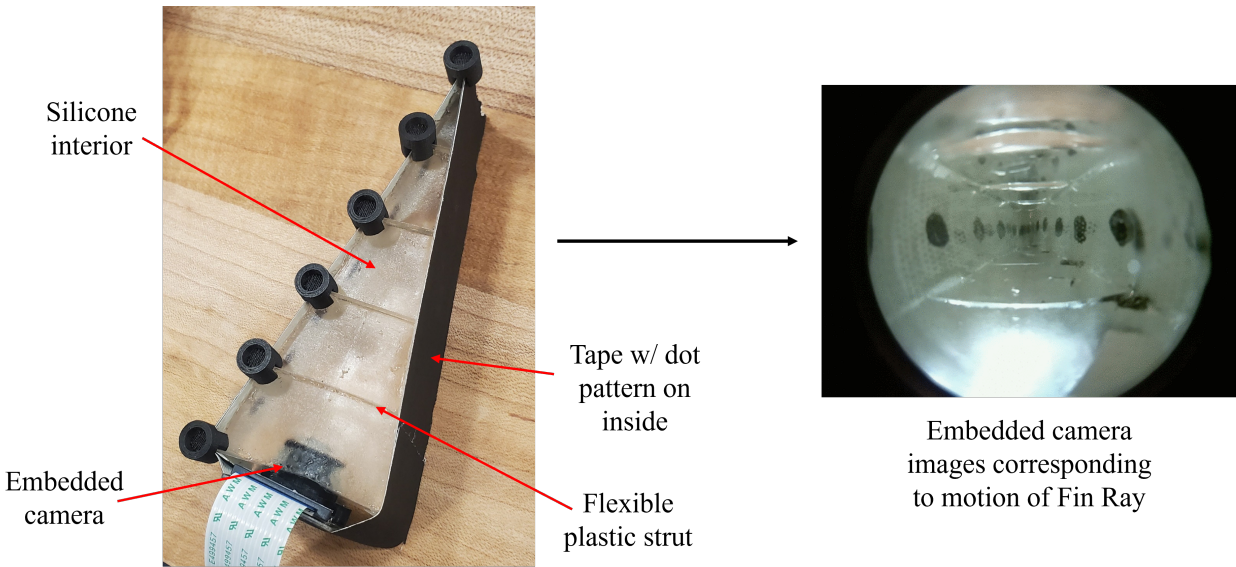


Figure 8.1: Optically clear Fin Ray structure with an embedded camera inside. On the right, is the corresponding image from the embedded camera with the markers on the side.

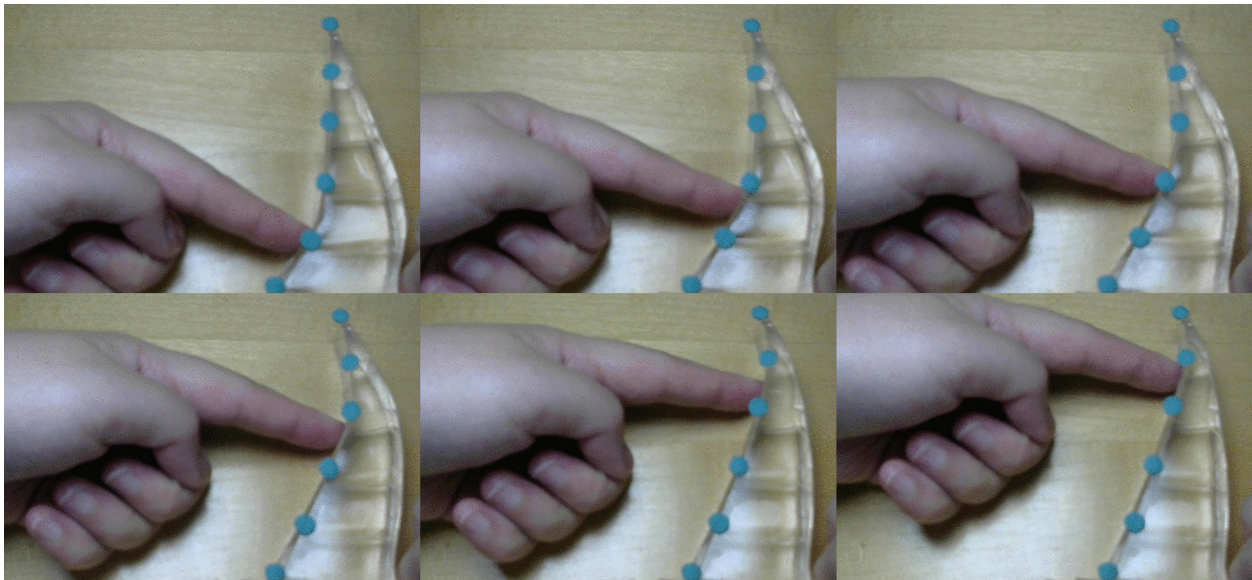


Figure 8.2: Optically clear fin ray structure deforming with a finger sliding along the side. Because of the silicone and the thin flexible plastic, the Fin Ray finger deflected backwards when an object was pressed against it.

from Markforged. Since the Onyx material has a higher elastic modulus than TPU and other rubbers, we thinned out the structure and decreased the fin widths. The part was printed in a single piece and the support material was easily removed. Despite there being no unwanted twisting in the structure, the Onyx plastically deformed after grasping smaller objects and quickly became infeasible to use (Fig. 8.3).

Another consideration was to SLA print the hollow structure using Flexible 80A Resin



Figure 8.3: The Onyx-based Fin Ray. The struts were thinned down and made narrower to help the structure deform, but the resulting structure still had signs of plastic deformation after a single grasp, which is shown in the curved struts and “tactile” section.

(Formlabs), which is shown in Fig. 8.4. The support material was not as easy to remove as the MarkForged Onyx material, but it was still doable within 10 minutes. However, because the material was so compliant, the material twisted and compressed too much for it to be useful for manipulation. We considered printing a semi-rigid backing to attach to the structure, but found that printing using a FDM printer was a lot faster.

The third hollow Fin Ray prototype was the FDM-printed multi-material Fin Ray (Fig. 8.5). The struts of the Fin Ray are made of flexible TPU 95A, while the backing is made of 3D printed PET. This structure manufacturing technique and final structural compliance had the most advantages of the other techniques we tried. Notably, it was relatively fast to print, did not have much twisting due to the rigid backing, and had the compliance needed to grasp different items. On the other hand, the support material and post-processing was very difficult. Not only did it require precision, it left a “hairy” facade afterwards and looked terrible. As a result, we moved to printing two halves of the flexible material portion of the Fin Ray separately, negating the need for any support material.

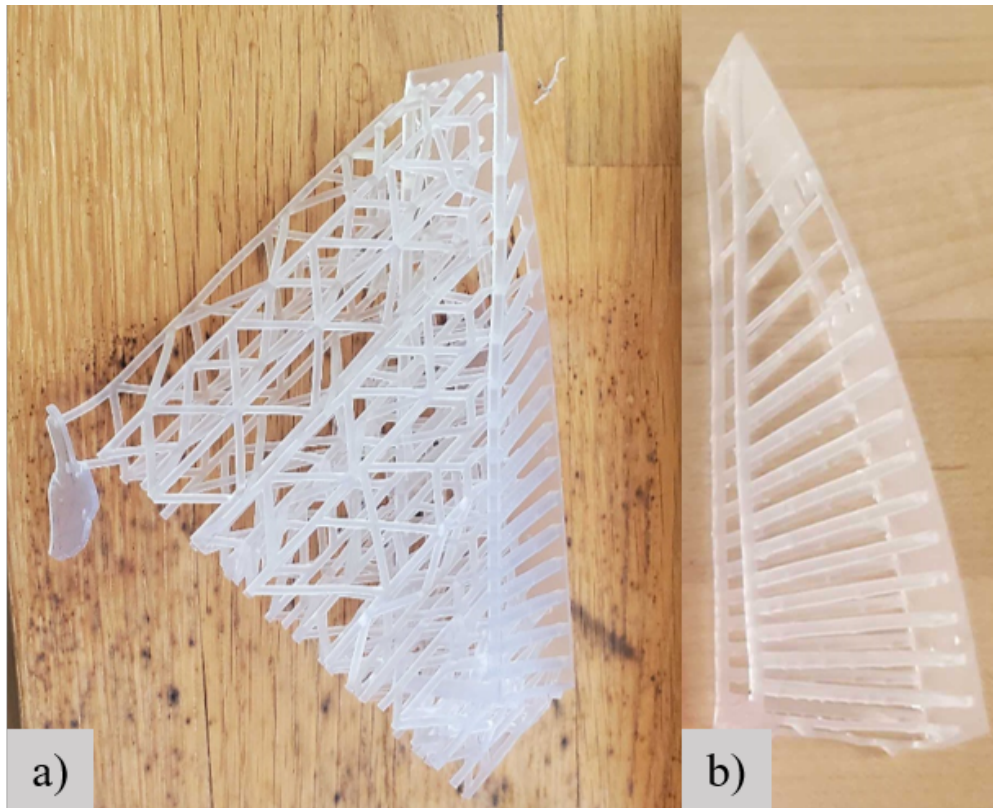


Figure 8.4: The SLA printed hollow Fin Ray. a) shows the 3D printed Fin Ray right off of the printer and b) shows the resulting post-processed Fin Ray structure.

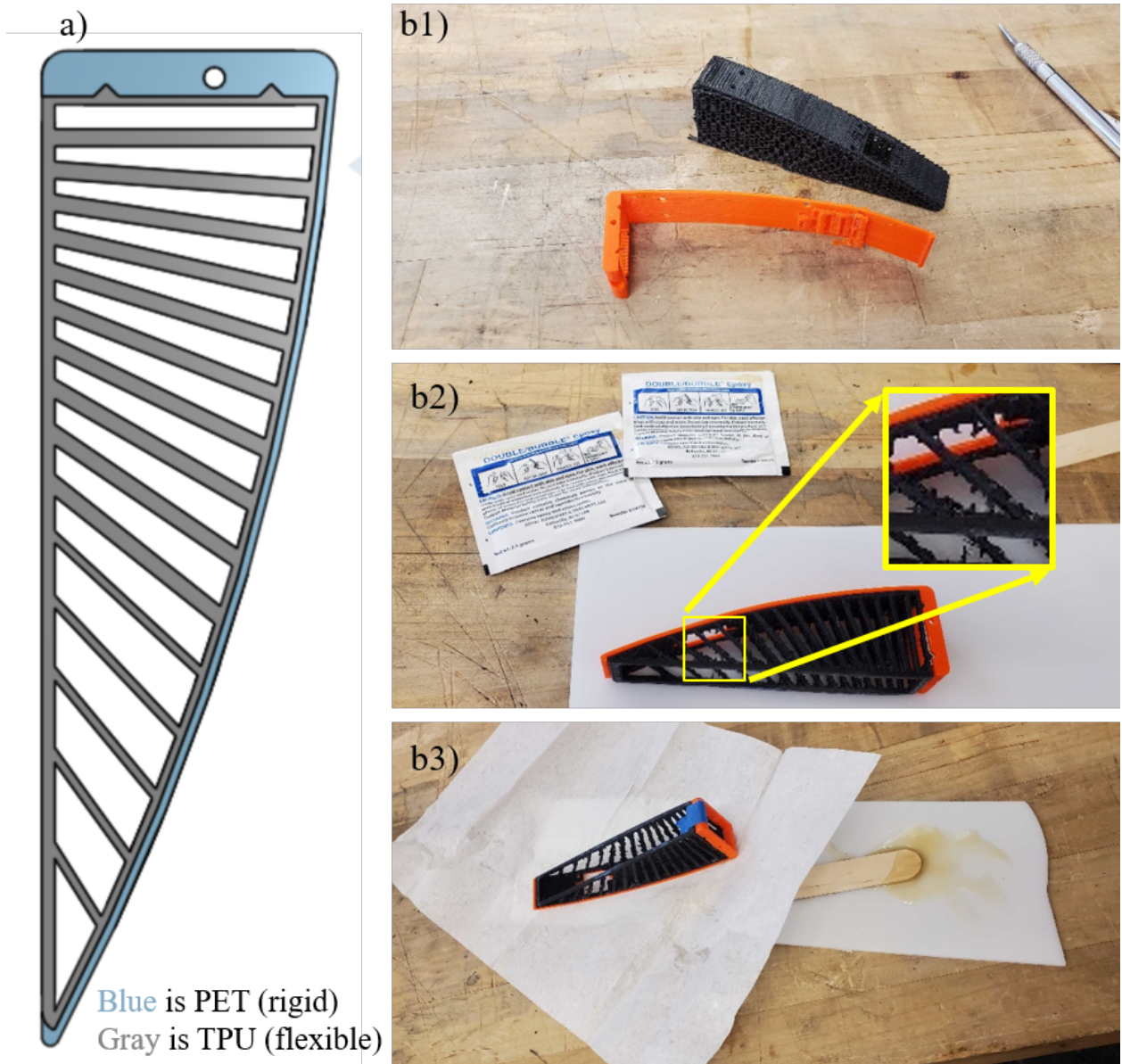


Figure 8.5: The multimaterial Fin Ray structure with a PET backing and TPU struts (a). (b) displays the manufacturing process: we start with the 3D printed structures taken off of the printers with the support material structure still attached. After removing the support material, we note the close up of the “hair” leftovers from the support structure (b2), and then we clamp the two pieces together with epoxy.

# Appendix A

## Tables

Table A.1: RMSE values between 82.55, 83.55, and 87.70 mm diameter dot pose data and "Testing" curves (all values in mm). The RMSE values are calculated between the dot locations estimated from the image (of these three differently sized objects) and the three testing diameters (which are all close in value to the actual diameter of the test object).

		Testing		
		82.55 mm	83.55 mm	87.70 mm
Actual	82.55 mm	0.60	0.65	0.86
	83.55 mm	0.44	0.49	0.69
	87.70 mm	0.27	0.30	0.47

Table A.2: Confusion matrix for size classification of all 80 "bar stock" trials. There were 40 trials each for cylinders and squares. All values are percentages, unless otherwise noted.

		Predicted			
		4.25 in	4.5 in	4.75 in	5.0 in
Actual	4.25 in	<b>100</b>	0	0	0
	4.5 in	0	<b>100</b>	0	0
	4.75 in	0	0	<b>90</b>	10
	4.75 in	0	0	5	<b>95</b>



Table A.3: Confusion matrix for size classification of the square "bar stock." All values are percentages, unless otherwise noted. The only sizes that were misclassified were the two larger sized boxes. This could have been caused by an unstable grip on the boxes by the soft robotic gripper.

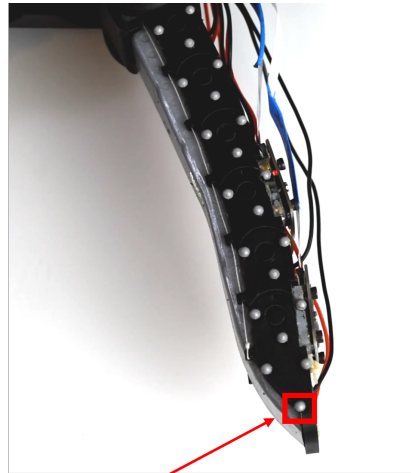
		Predicted			
		4.25 in	4.5 in	4.75 in	5.0 in
Actual	4.25 in	<b>100</b>	0	0	0
	4.5 in	0	<b>100</b>	0	0
	4.75 in	0	0	<b>80</b>	20
	4.75 in	0	0	10	<b>90</b>

Table A.4: Confusion matrix for size classification of the cylinder "bar stock." All values are percentages, unless otherwise noted. The size classification neural net appeared to perform better on cylinders than boxes.

		Predicted			
		4.25 in	4.5 in	4.75 in	5.0 in
Actual	4.25 in	<b>100</b>	0	0	0
	4.5 in	0	<b>100</b>	0	0
	4.75 in	0	0	<b>100</b>	0
	4.75 in	0	0	0	<b>100</b>

# Appendix B

## Figures



OptiTrack Marker

Figure B.1: Close up of a GelFlex finger with OptiTrack markers on it. Each link was represented by three separate OptiTrack markers to create a rigid body.

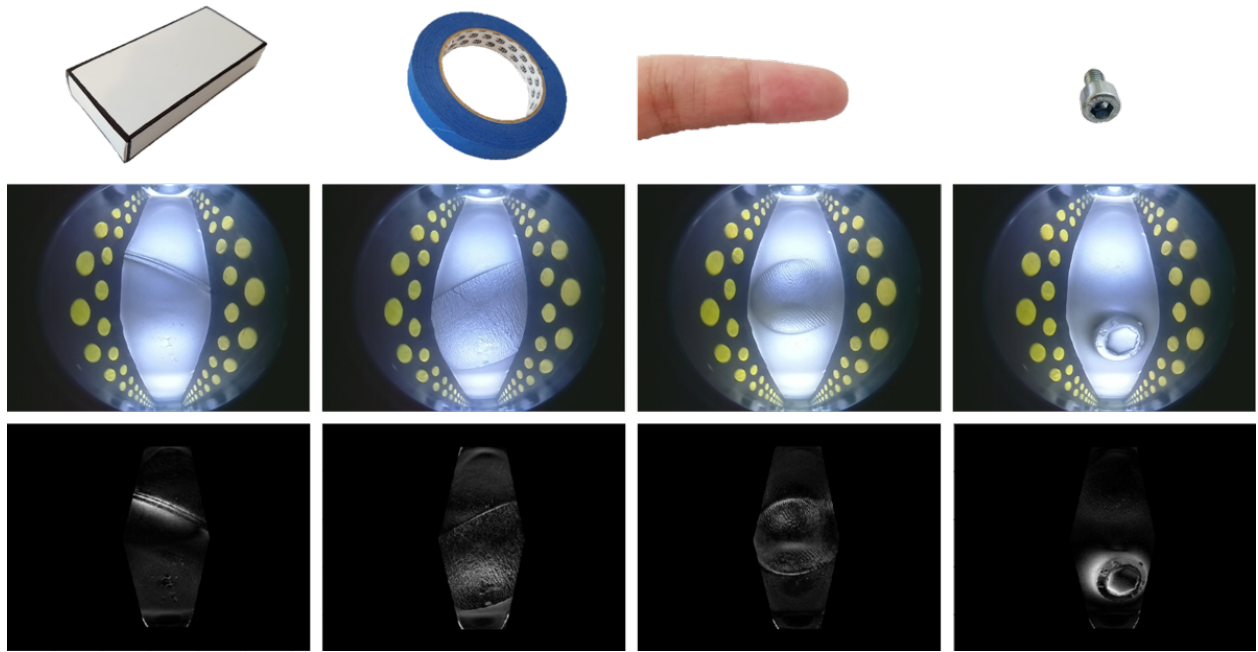


Figure B.2: Tactile images from the GelFlex. The first row shows the objects that are pressed into the finger, while the second row shows the raw image taken from the camera and the third row shows the corresponding difference image. From left to right, we have a cardboard box edge, painter's tape, a finger, and a screw head.

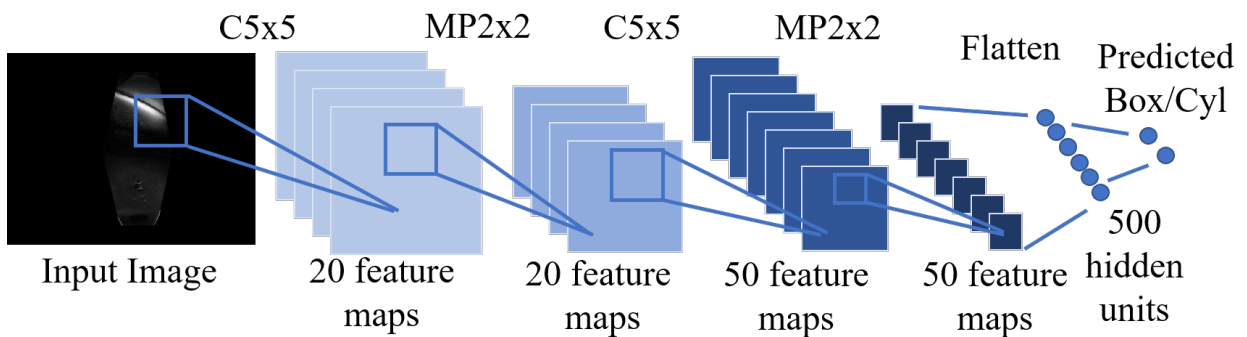


Figure B.3: Tactile neural network architecture. It takes a difference image of the GelFlex grasping an object as an input and outputs the label "Box" or "Cylinder," based on what the finger is in contact with. C5x5 represents a convolution with kernel size 5x5 while MP2x2 represents Max Pooling with a 2x2 kernel size.

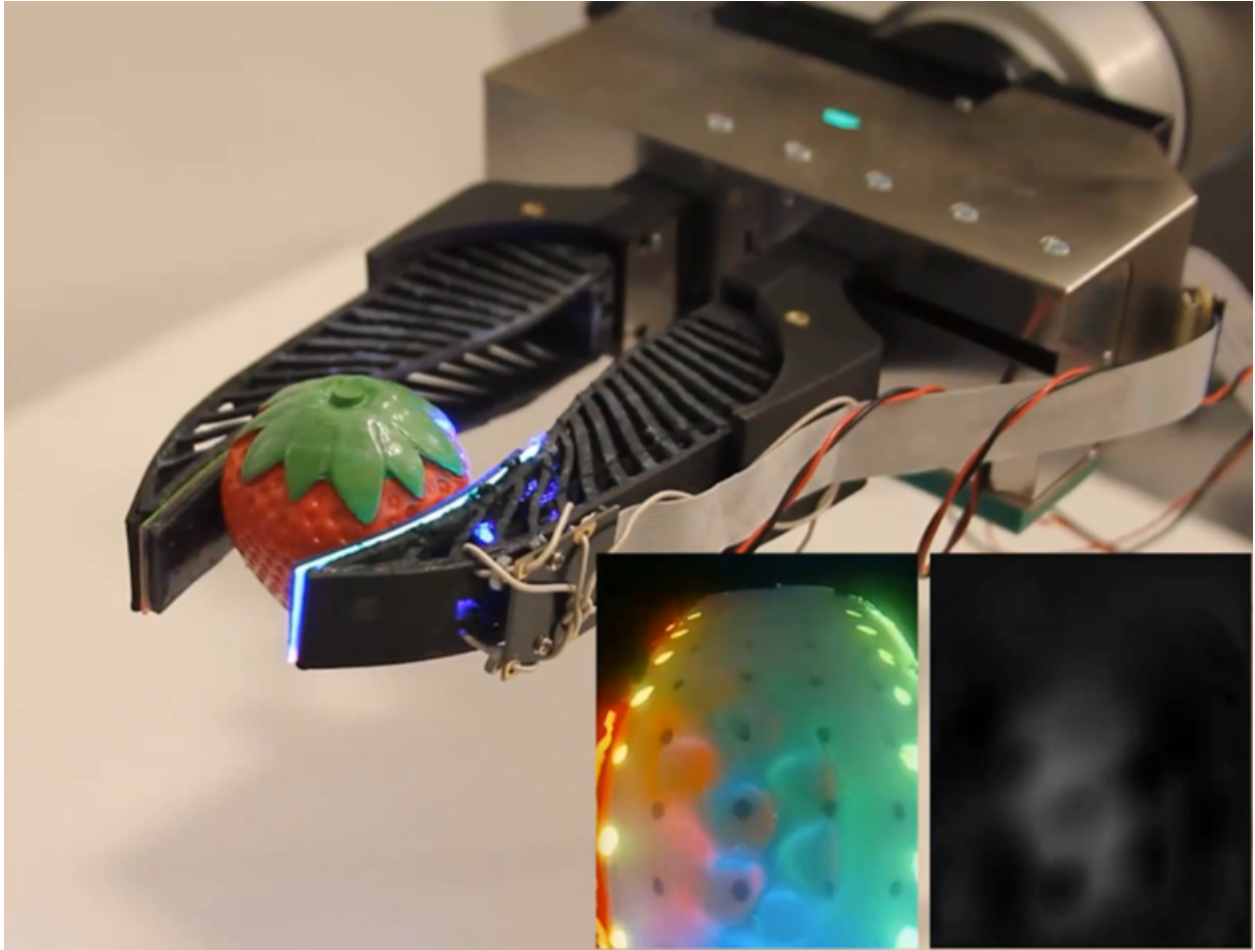


Figure B.4: The GelSight Fin Ray grasping a plastic strawberry from the YCB dataset [67]. In the bottom right corner, there is the raw image on the left and the corresponding reconstruction image on the right.

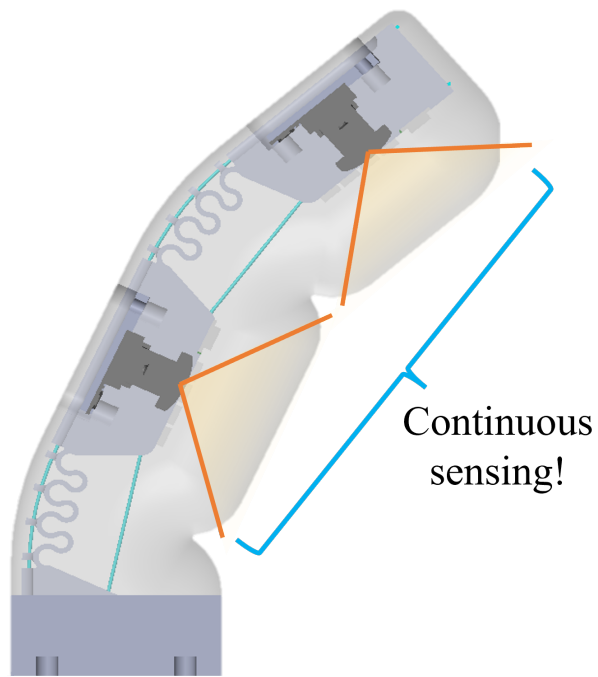


Figure B.5: A side view of the GelSight EndoFlex with the camera field of views shown. When the finger is bent, we have continuous tactile sensing along the front and sides of each finger.

# References

- [1] H. Wang, M. Totaro, and L. Beccai, “Toward perceptive soft robots: Progress and challenges,” *Advanced Science*, vol. 5, no. 9, p. 1 800 541, 2018.
- [2] Y. Huang, X. Zhang, X. Chen, and J. Ota, “Vision-guided peg-in-hole assembly by baxter robot,” *Advances in Mechanical Engineering*, vol. 9, no. 12, p. 1 687 814 017 748 078, 2017. DOI: [10.1177/1687814017748078](https://doi.org/10.1177/1687814017748078).
- [3] S. Lin, X. Jiang, and Y. Liu, “Cable manipulation with partially occluded vision feedback,” in *2022 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2022, pp. 1245–1250. DOI: [10.1109/ROBIO55434.2022.10011815](https://doi.org/10.1109/ROBIO55434.2022.10011815).
- [4] Y. Tang, M. Chen, C. Wang, L. Luo, J. Li, G. Lian, and X. Zou, “Recognition and localization methods for vision-based fruit picking robots: A review,” in *Frontiers in plant science*, vol. 11, 2020. DOI: [10.3389/fpls.2020.00510](https://doi.org/10.3389/fpls.2020.00510).
- [5] E. Xie, W. Wang, W. Wang, M. Ding, C. Shen, and P. Luo, “Segmenting transparent objects in the wild,” *arXiv preprint arXiv:2003.13948*, 2020.
- [6] S. Kim and A. Rodriguez, “Active extrinsic contact sensing: Application to general peg-in-hole insertion,” in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 10 241–10 247. DOI: [10.1109/ICRA46639.2022.9812017](https://doi.org/10.1109/ICRA46639.2022.9812017).
- [7] Y. She\*, S. Wang\*, S. Dong\*, N. Sunil, A. Rodriguez, and E. H. Adelson, “Cable manipulation with a tactile-reactive gripper,” *CoRR*, vol. abs/1910.02860, 2019.
- [8] W. Yuan, C. Zhu, A. Owens, M. Srinivasan, and E. Adelson, “Shape-independent hardness estimation using deep learning and a gelsight tactile sensor,” May 2017, pp. 951–958. DOI: [10.1109/ICRA.2017.7989116](https://doi.org/10.1109/ICRA.2017.7989116).
- [9] K. Junge and J. Hughes, “Soft sensorized physical twin for harvesting raspberries,” in *2022 IEEE 5th International Conference on Soft Robotics (RoboSoft)*, 2022, pp. 601–606. DOI: [10.1109/RoboSoft54090.2022.9762135](https://doi.org/10.1109/RoboSoft54090.2022.9762135).
- [10] M. Lambeta, G. Kammerer, D. Jayaraman, *et al.*, “Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation,” *IEEE Robotics and Automation Letters*, vol. PP, pp. 1–1, Feb. 2020. DOI: [10.1109/LRA.2020.2977257](https://doi.org/10.1109/LRA.2020.2977257).
- [11] S. Puhmann, J. Harris, and O. Brock, “IRBO hand 3/i: A platform for soft dexterous manipulation,” *IEEE Transactions on Robotics*, vol. 38, no. 6, pp. 3434–3449, Dec. 2022. DOI: [10.1109/tro.2022.3156806](https://doi.org/10.1109/tro.2022.3156806). [Online]. Available: <https://doi.org/10.1109%2Ftro.2022.3156806>.

- [12] W. Yuan, S. Dong, and E. H. Adelson, “Gelsight: High-resolution robot tactile sensors for estimating geometry and force,” *Sensors*, vol. 17, no. 12, p. 2762, 2017.
- [13] Y. She\*, S. Q. Liu\*, P. Yu\*, and E. H. Adelson, “Exoskeleton-covered soft finger with vision-based proprioception and tactile sensing,” in *Proc. of The International Conference in Robotics and Automation (ICRA)*, 2020.
- [14] 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)*, 2022, pp. 925–931. DOI: [10.1109/RoboSoft54090.2022.9762175](https://doi.org/10.1109/RoboSoft54090.2022.9762175).
- [15] 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)*, Mar. 2023.
- [16] 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.
- [17] L. B. François Schmitt Olivier Piccin and B. Bayle, “Soft robots manufacturing: A review,” *Front Robot AI*, vol. 5, p. 84, 2018. DOI: [10.3389/frobt.2018.00084](https://doi.org/10.3389/frobt.2018.00084).
- [18] G. Chowdhary, M. Gazzola, G. Krishnan, C. Soman, and S. Lovell, “Soft robotics as an enabling technology for agroforestry practice and research,” *Sustainability*, vol. 11, p. 6751, Nov. 2019. DOI: [10.3390/su11236751](https://doi.org/10.3390/su11236751).
- [19] E. W. Hawkes, L. H. Blumenschein, J. D. Greer, and A. M. Okamura, “A soft robot that navigates its environment through growth,” *Science Robotics*, vol. 2, no. 8, eaan3028, 2017. DOI: [10.1126/scirobotics.aan3028](https://doi.org/10.1126/scirobotics.aan3028).
- [20] K. Becker, C. Teeple, N. Charles, Y. Jung, D. Baum, J. C. Weaver, L. Mahadevan, and R. Wood, “Active entanglement enables stochastic, topological grasping,” *Proceedings of the National Academy of Sciences*, vol. 119, no. 42, e2209819119, 2022. DOI: [10.1073/pnas.2209819119](https://doi.org/10.1073/pnas.2209819119).
- [21] W. Crooks, G. Vukasin, M. O’Sullivan, W. Messner, and C. Rogers, “Fin ray effect inspired soft robotic gripper: From the robosoft grand challenge toward optimization,” *Frontiers in Robotics and AI*, vol. 3, Nov. 2016. DOI: [10.3389/frobt.2016.00070](https://doi.org/10.3389/frobt.2016.00070).
- [22] Y. Hao, Z. Gong, Z. Xie, S. Guan, X. Yang, Z. Ren, T. Wang, and L. Wen, “Universal soft pneumatic robotic gripper with variable effective length,” in *2016 35th Chinese Control Conference (CCC)*, 2016, pp. 6109–6114. DOI: [10.1109/ChiCC.2016.7554316](https://doi.org/10.1109/ChiCC.2016.7554316).
- [23] Z. Ren, M. Zarepoor, X. Huang, A. P. Sabelhaus, and C. Majidi, “Shape memory alloy (sma) actuator with embedded liquid metal curvature sensor for closed-loop control,” *Frontiers in Robotics and AI*, vol. 8, 2021, ISSN: 2296-9144. DOI: [10.3389/frobt.2021.599650](https://doi.org/10.3389/frobt.2021.599650). [Online]. Available: <https://www.frontiersin.org/articles/10.3389/frobt.2021.599650>.

---

\*denotes authors with equal contribution

- [24] Y. Kim and Y. Cha, “Soft pneumatic gripper with a tendon-driven soft origami pump,” *Frontiers in Bioengineering and Biotechnology*, vol. 8, 2020, ISSN: 2296-4185. DOI: [10.3389/fbioe.2020.00461](https://doi.org/10.3389/fbioe.2020.00461). [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fbioe.2020.00461>.
- [25] P. Rao, Q. Peyron, S. Lilge, and J. Burgner-Kahrs, “How to model tendon-driven continuum robots and benchmark modelling performance,” *Frontiers in Robotics and AI*, vol. 7, 2021, ISSN: 2296-9144. DOI: [10.3389/frobt.2020.630245](https://doi.org/10.3389/frobt.2020.630245). [Online]. Available: <https://www.frontiersin.org/articles/10.3389/frobt.2020.630245>.
- [26] Y. Hao, Z. Shixin, B. Fang, F. Sun, H. Liu, and H. Li, “A review of smart materials for the boost of soft actuators, soft sensors, and robotics applications,” *Chinese Journal of Mechanical Engineering*, vol. 35, p. 37, Dec. 2022. DOI: [10.1186/s10033-022-00707-2](https://doi.org/10.1186/s10033-022-00707-2).
- [27] H. Yang and W. Wu, “A review: Machine learning for strain sensor-integrated soft robots,” *Frontiers in Electronic Materials*, vol. 2, 2022, ISSN: 2673-9895. DOI: [10.3389/femat.2022.1000781](https://doi.org/10.3389/femat.2022.1000781). [Online]. Available: <https://www.frontiersin.org/articles/10.3389/femat.2022.1000781>.
- [28] K. C. Galloway, Y. Chen, E. Templeton, B. Rife, I. S. Godage, and E. J. Barth, “Fiber optic shape sensing for soft robotics,” *Soft Robotics*, vol. 6, no. 5, pp. 671–684, 2019, PMID: 31241408. DOI: [10.1089/soro.2018.0131](https://doi.org/10.1089/soro.2018.0131).
- [29] J. Xu\*, H. Lin\*, S. Song, and M. Ciocarlie, “Tandem3d: Active tactile exploration for 3d object recognition,” in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 10 401–10 407. DOI: [10.1109/ICRA48891.2023.10161091](https://doi.org/10.1109/ICRA48891.2023.10161091).
- [30] G. Khandate, M. Haas-Heger, and M. Ciocarlie, “On the feasibility of learning finger-gaiting in-hand manipulation with intrinsic sensing,” in *2022 International Conference on Robotics and Automation (ICRA)*, IEEE Press, 2022, pp. 2752–2758. DOI: [10.1109/ICRA46639.2022.9812212](https://doi.org/10.1109/ICRA46639.2022.9812212).
- [31] S. Wang, Y. She, B. Romero, and E. Adelson, “Gelsight wedge: Measuring high-resolution 3d contact geometry with a compact robot finger,” Jun. 2021.
- [32] E. Donlon, S. Dong, M. Liu, J. Li, E. Adelson, and A. Rodriguez, “Gelslim: A high-resolution, compact, robust, and calibrated tactile-sensing finger,” in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1927–1934.
- [33] W. Yuan, Y. Mo, S. Wang, and E. Adelson, “Active clothing material perception using tactile sensing and deep learning,” Nov. 2017.
- [34] N. Kuppuswamy, A. Alspach, A. Uttamchandani, S. Creasey, T. Ikeda, and R. Tedrake, “Soft-bubble grippers for robust and perceptive manipulation,” in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 9917–9924. DOI: [10.1109/IROS45743.2020.9341534](https://doi.org/10.1109/IROS45743.2020.9341534).



- [35] C.-Y. Lin, Z.-X. Lin, S. Wang, and H. Xu, “Dtact: A vision-based tactile sensor that measures high-resolution 3d geometry directly from darkness,” *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 10 359–10 366, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:252568209>.
- [36] A. Chaudhury and C. Atkeson, “Controlled illumination for perception and manipulation of lambertian objects,” Apr. 2023.
- [37] M. K. Johnson and E. H. Adelson, “Retrographic sensing for the measurement of surface texture and shape,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1070–1077. DOI: [10.1109/CVPR.2009.5206534](https://doi.org/10.1109/CVPR.2009.5206534).
- [38] M. H. Tippur and E. H. Adelson, “GelSight360: An omnidirectional camera-based tactile sensor for dexterous robotic manipulation,” in *2023 IEEE Conference of Soft Robotics (RoboSoft)*, Apr. 2023.
- [39] J. Zhao and E. H. Adelson, “GelSight Svelte: A human finger-shaped single-camera tactile robot finger with large sensing coverage and proprioceptive sensing,” in *2023 International Conference on Intelligent Robots and Systems (IROS 2023)*, 2023.
- [40] S. Q. Liu, “Vision-based proprioception of a soft robotic finger with tactile sensing,” 2020.
- [41] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 3431–3440. DOI: [10.1109/CVPR.2015.7298965](https://doi.org/10.1109/CVPR.2015.7298965).
- [42] S. Dong, W. Yuan, and E. H. Adelson, “Improved gelsight tactile sensor for measuring geometry and slip,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2017, pp. 137–144.
- [43] Y. LeCun, L. Jackel, L. Bottou, *et al.*, “Comparison of learning algorithms for handwritten digit recognition,” in *INTERNATIONAL CONFERENCE ON ARTIFICIAL NEURAL NETWORKS*, 1995, pp. 53–60.
- [44] T. Park, M.-Y. Liu, T.-C. Wang, and J.-Y. Zhu, “Semantic image synthesis with spatially-adaptive normalization,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019.
- [45] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Ng, “Ros: An open-source robot operating system,” vol. 3, Jan. 2009.
- [46] H. Kam, S.-H. Lee, T. Park, and C.-H. Kim, “Rviz: A toolkit for real domain data visualization,” *Telecommunication Systems*, vol. 60, pp. 1–9, Oct. 2015. DOI: [10.1007/s11235-015-0034-5](https://doi.org/10.1007/s11235-015-0034-5).
- [47] C. T. Fuentes and A. J. Bastian, “Where is your arm? variations in proprioception across space and tasks,” *Journal of Neurophysiology*, vol. 103, no. 1, pp. 164–171, 2010.
- [48] FESTO, “Multichoicegripper,” 2014. [Online]. Available: [https://www.festo.com/net/SupportPortal/Files/333986/Festo%5C\\_%20MultiChoiceGripper%5C\\_en.pdf](https://www.festo.com/net/SupportPortal/Files/333986/Festo%5C_%20MultiChoiceGripper%5C_en.pdf).

- [49] J. L. Tangorra, G. V. Lauder, I. W. Hunter, R. Mittal, P. G. A. Madden, and M. Bozkurttas, “The effect of fin ray flexural rigidity on the propulsive forces generated by a biorobotic fish pectoral fin,” *Journal of Experimental Biology*, vol. 213, no. 23, pp. 4043–4054, Dec. 2010, ISSN: 0022-0949. DOI: [10.1242/jeb.048017](https://doi.org/10.1242/jeb.048017).
- [50] J. Shintake, V. Cacucciolo, D. Floreano, and H. Shea, “Soft robotic grippers,” *Advanced Materials*, vol. 30, no. 29, p. 1707035, 2018. DOI: <https://doi.org/10.1002/adma.201707035>.
- [51] M. H. Ali, A. Zhanabayev, S. Khamzhin, and K. Mussin, “Biologically inspired gripper based on the fin ray effect,” in *2019 5th International Conference on Control, Automation and Robotics (ICCAR)*, 2019, pp. 865–869. DOI: [10.1109/ICCAR.2019.8813388](https://doi.org/10.1109/ICCAR.2019.8813388).
- [52] W. Crooks, S. Rozen-Levy, B. Trimmer, C. Rogers, and W. Messner, “Passive gripper inspired by manduca sexta and the fin ray<sup>®</sup> effect,” *International Journal of Advanced Robotic Systems*, vol. 14, Jul. 2017. DOI: [10.1177/1729881417721155](https://doi.org/10.1177/1729881417721155).
- [53] X. Shan and L. Birglen, “Modeling and analysis of soft robotic fingers using the fin ray effect,” *The International Journal of Robotics Research*, vol. 39, Apr. 2020. DOI: [10.1177/0278364920913926](https://doi.org/10.1177/0278364920913926).
- [54] Z. Deng and M. Li, “Learning optimal fin-ray finger design for soft grasping,” *Frontiers in Robotics and AI*, vol. 7, Feb. 2021. DOI: [10.3389/frobt.2020.590076](https://doi.org/10.3389/frobt.2020.590076).
- [55] K. Elgeneidy, P. Lightbody, S. Pearson, and G. Neumann, “Characterising 3d-printed soft fin ray robotic fingers with layer jamming capability for delicate grasping,” in *2019 2nd IEEE International Conference on Soft Robotics (RoboSoft)*, 2019, pp. 143–148. DOI: [10.1109/ROBOSOFT.2019.8722715](https://doi.org/10.1109/ROBOSOFT.2019.8722715).
- [56] R. Chen, R. Song, Z. Zhang, L. Bai, F. Liu, P. Jiang, D. Sindersonberger, G. J. Monkman, and J. Guo, “Bio-inspired shape-adaptive soft robotic grippers augmented with electroadhesion functionality,” *Soft robotics*, 2019.
- [57] 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*, vol. 12, no. 10, 2021, ISSN: 2072-666X. DOI: [10.3390/mi12101141](https://doi.org/10.3390/mi12101141).
- [58] K. Elgeneidy, A. Fansa, I. Hussain, and K. Goher, “Structural optimization of adaptive soft fin ray fingers with variable stiffening capability,” in *2020 3rd IEEE International Conference on Soft Robotics (RoboSoft)*, 2020, pp. 779–784. DOI: [10.1109/RoboSoft48309.2020.9115969](https://doi.org/10.1109/RoboSoft48309.2020.9115969).
- [59] R. Patel, R. Ouyang, B. Romero, and E. Adelson, “Digger finger: Gelsight tactile sensor for object identification inside granular media,” in *Experimental Robotics*, B. Siciliano, C. Laschi, and O. Khatib, Eds., Cham: Springer International Publishing, 2021, pp. 105–115, ISBN: 978-3-030-71151-1.
- [60] G. Bradski, “The OpenCV Library,” *Dr. Dobb’s Journal of Software Tools*, 2000.
- [61] C. R. Harris, K. J. Millman, S. J. van der Walt, *et al.*, “Array programming with NumPy,” *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. DOI: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2).

- [62] J. Doerner, *Fast poisson reconstruction in python*, <https://gist.github.com/jackdoerner/b9b5e62a4c3893c76e4c>, 2014.
- [63] K. Zou, S. Warfield, A. Bharatha, C. Tempany, M. Kaus, S. Haker, W. Wells, F. Jolesz, and R. Kikinis, “Statistical validation of image segmentation quality based on a spatial overlap index,” *Academic radiology*, vol. 11, pp. 178–89, Feb. 2004. DOI: [10.1016/S1076-6332\(03\)00671-8](https://doi.org/10.1016/S1076-6332(03)00671-8).
- [64] K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, 2015. DOI: [10.48550/ARXIV.1512.03385](https://doi.org/10.48550/ARXIV.1512.03385). [Online]. Available: <https://arxiv.org/abs/1512.03385>.
- [65] W. Bottcher, P. Machado, N. Lama, and T. M. McGinnity, “Object recognition for robotics from tactile time series data utilising different neural network architectures,” 2021. DOI: [10.48550/ARXIV.2109.04573](https://doi.org/10.48550/ARXIV.2109.04573). [Online]. Available: <https://arxiv.org/abs/2109.04573>.
- [66] S. Pohtongkam and J. Srinonchat, “Tactile object recognition for humanoid robots using new designed piezoresistive tactile sensor and dcnn,” *Sensors*, vol. 21, no. 18, 2021, ISSN: 1424-8220. DOI: [10.3390/s21186024](https://doi.org/10.3390/s21186024). [Online]. Available: <https://www.mdpi.com/1424-8220/21/18/6024>.
- [67] B. Çalli, A. Walsman, A. Singh, S. S. Srinivasa, P. Abbeel, and A. M. Dollar, “Benchmarking in manipulation research: The YCB object and model set and benchmarking protocols,” *CoRR*, vol. abs/1502.03143, 2015. eprint: [1502.03143](https://arxiv.org/abs/1502.03143).