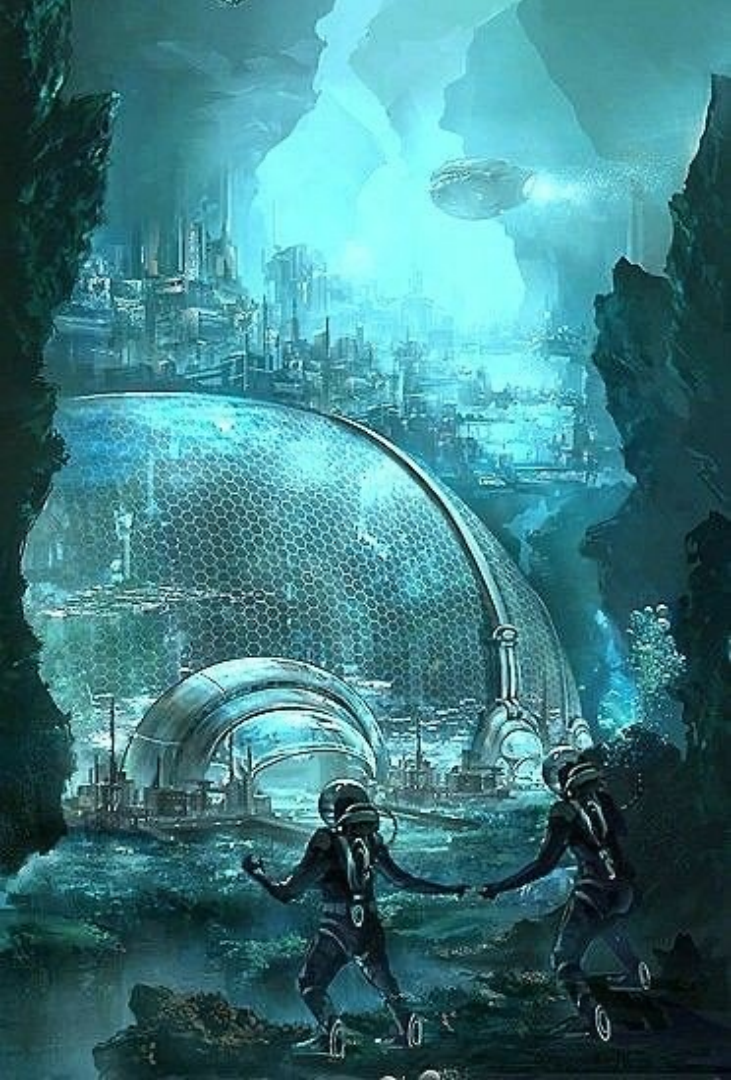
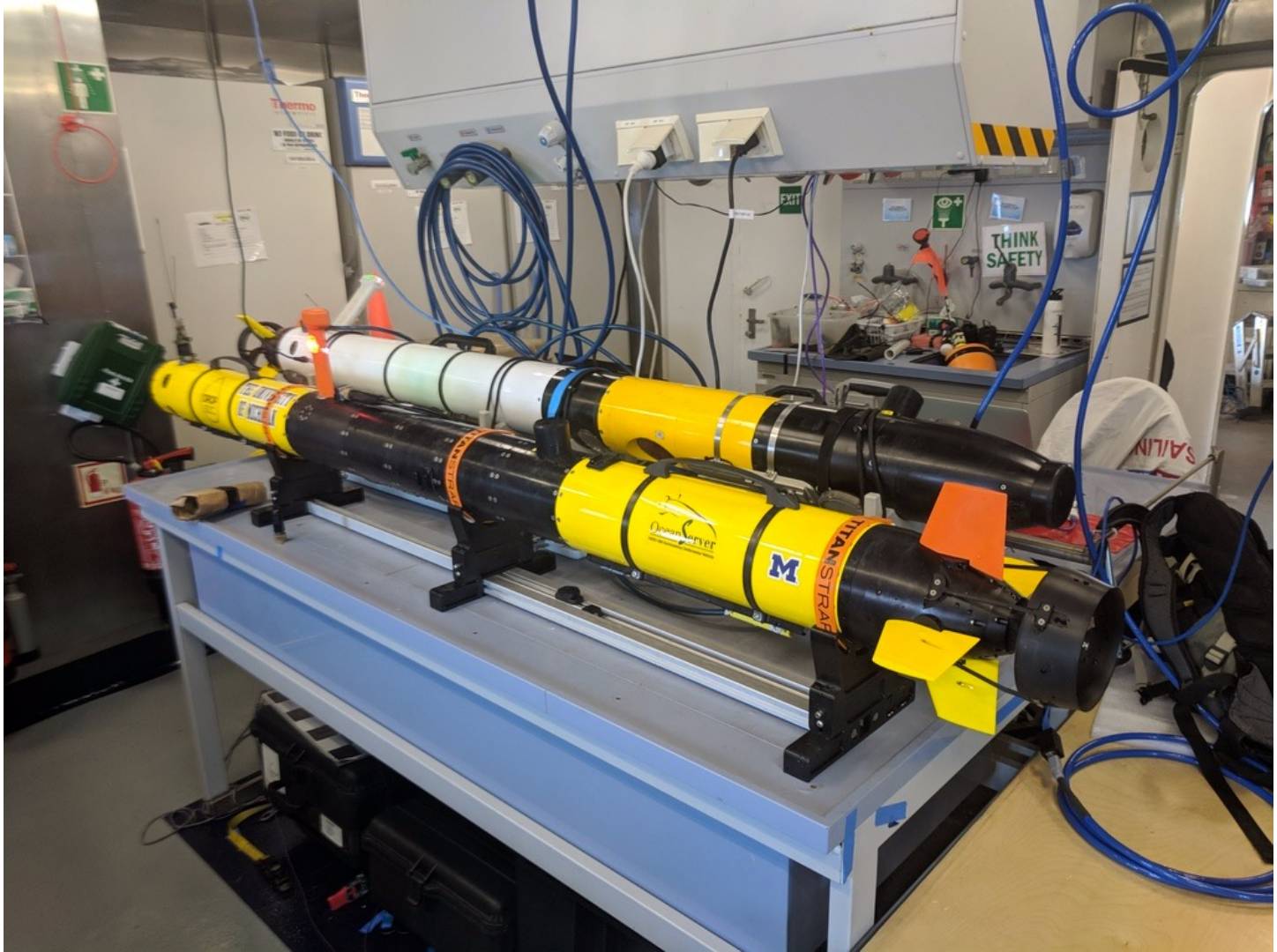
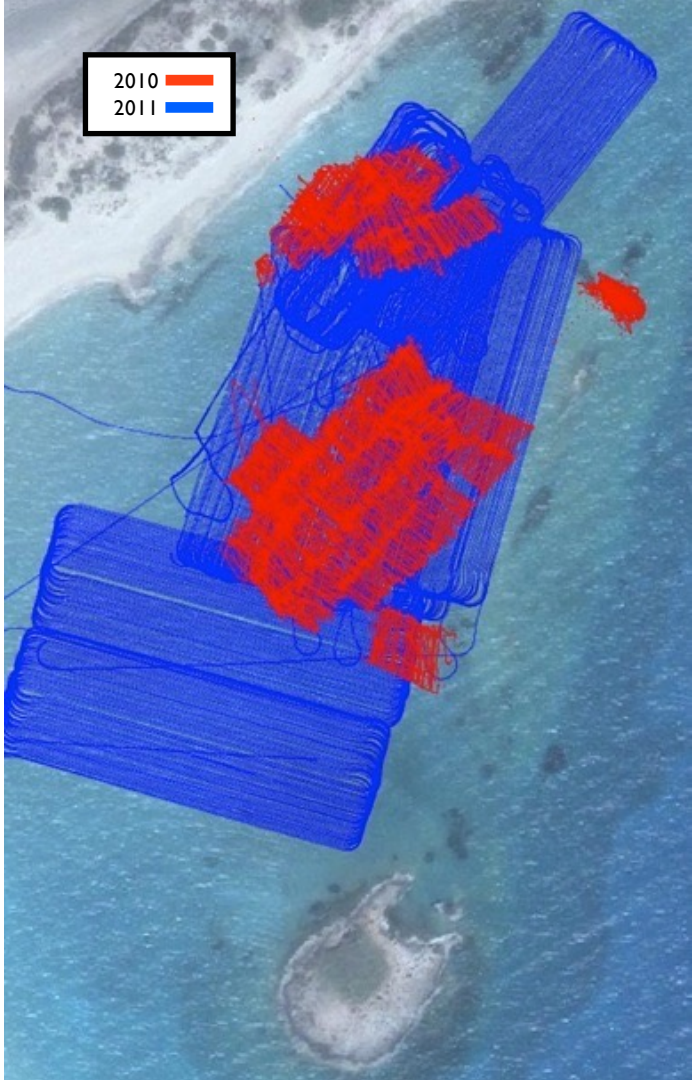
Lessons from the field:
Deep Learning and Machine Perception for field robots

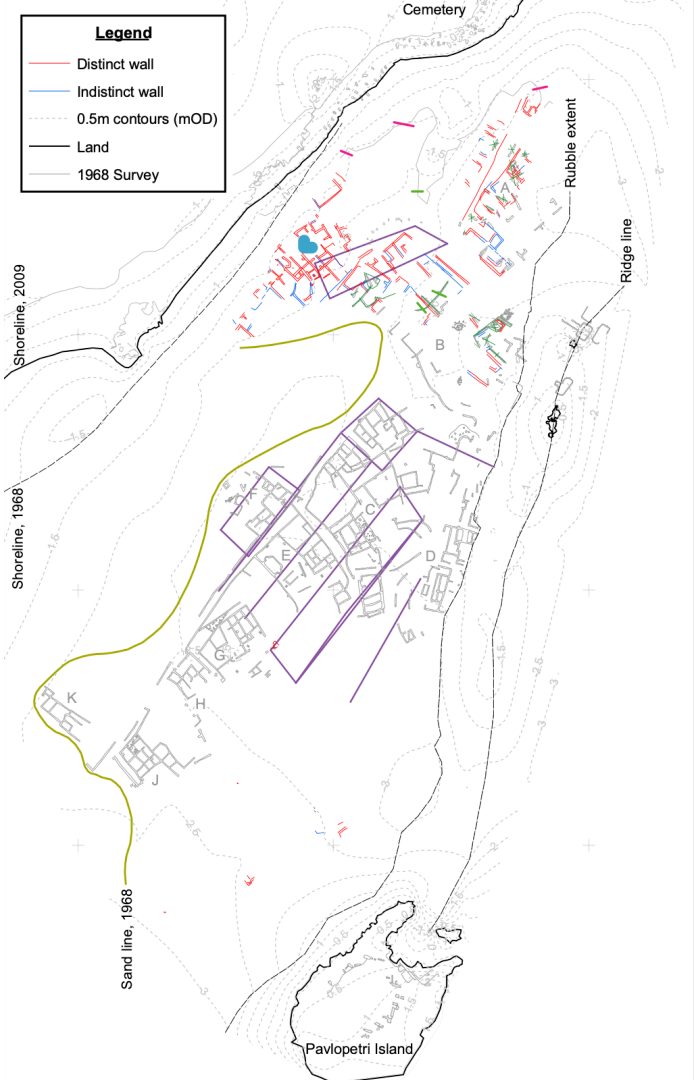
Matthew Johnson-Roberson
CMU
2022

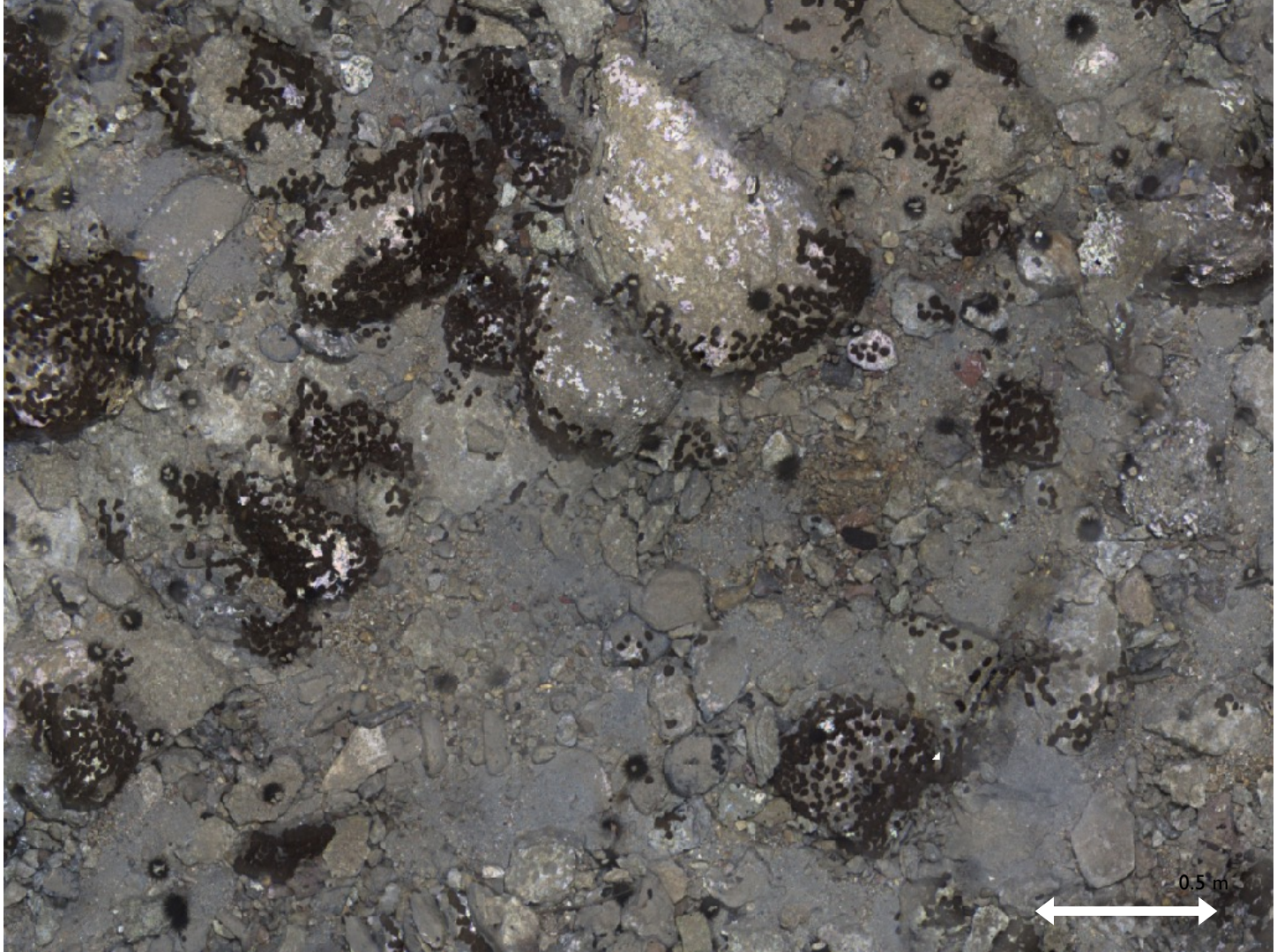


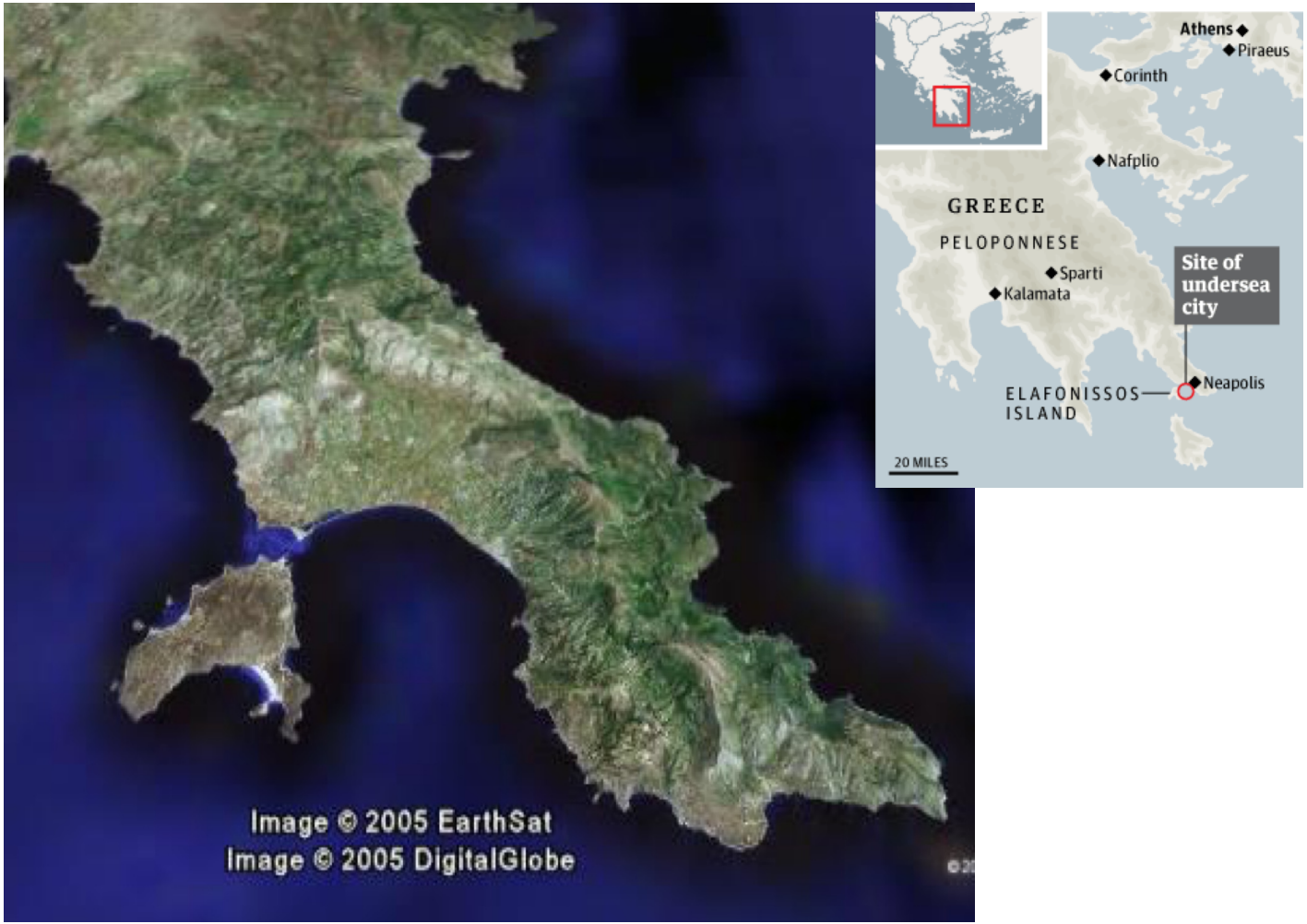
The Future


THINK
SAFETY
EXIT
Thermo
M

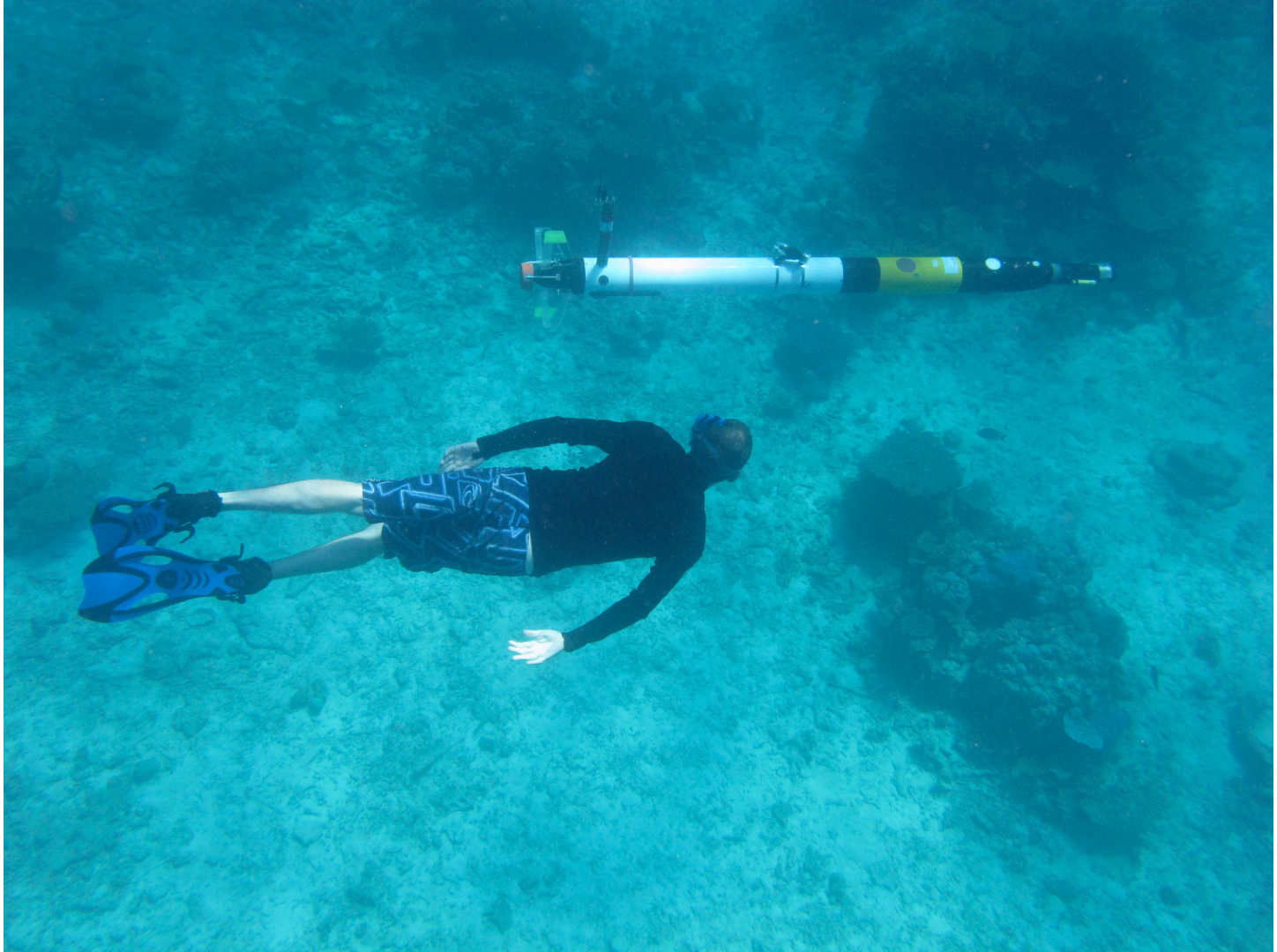

2010
2011

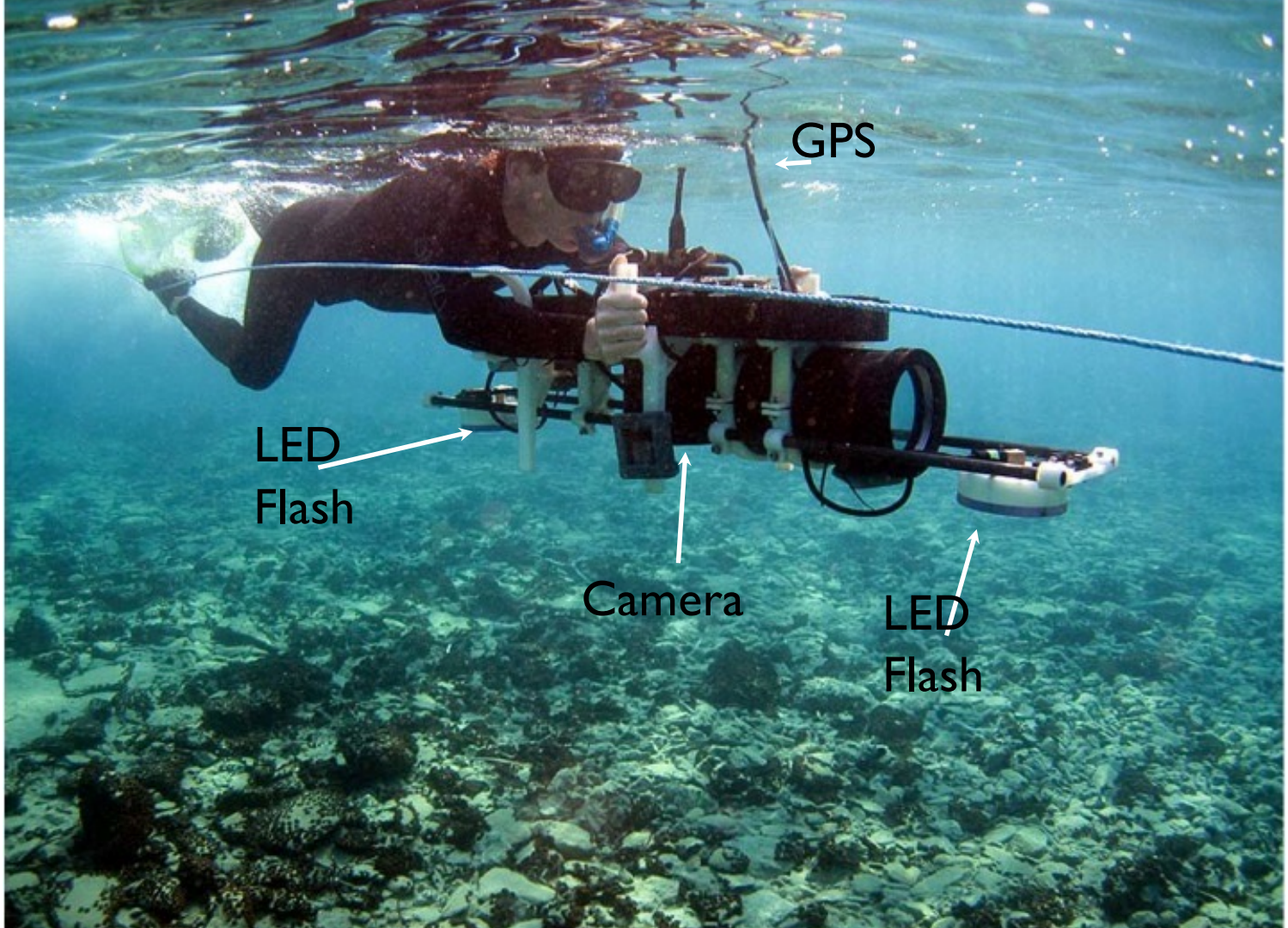

Legend
Distinct wall
Indistinct wall
0.5m contours (mOD)
Land
1968 Survey
Cemetery
Rubble extent
Ridge line
Shoreline, 2009
Shoreline, 1968
Sand line, 1968
Pavlopetri Island
A
B
C
D
E
F
G
H
J
K


0.5 m

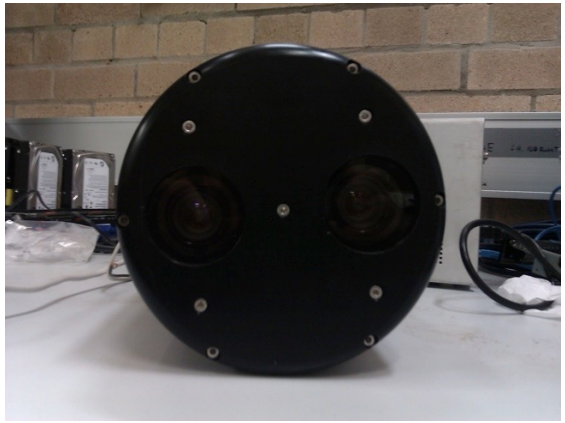

Athens
Piraeus
Corinth
Nafplio
GREECE
PELOPONNESE
Sparti
Kalamata
Site of undersea city
Neapolis
ELAFONISSOS ISLAND
20 MILES
Image © 2005 EarthSat
Image © 2005 DigitalGlobe

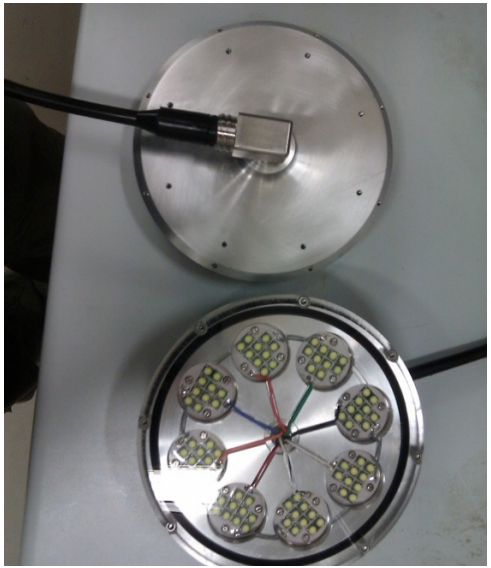
INSTITUTE FOR EXPLORATION
MYSTIC AQUARIUM
INSTITUTE FOR EXPLORATION
MYSTIC AQUARIUM
PREDATOR

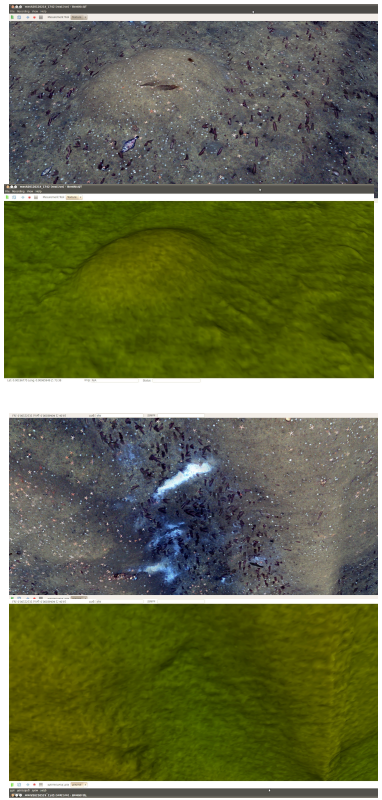


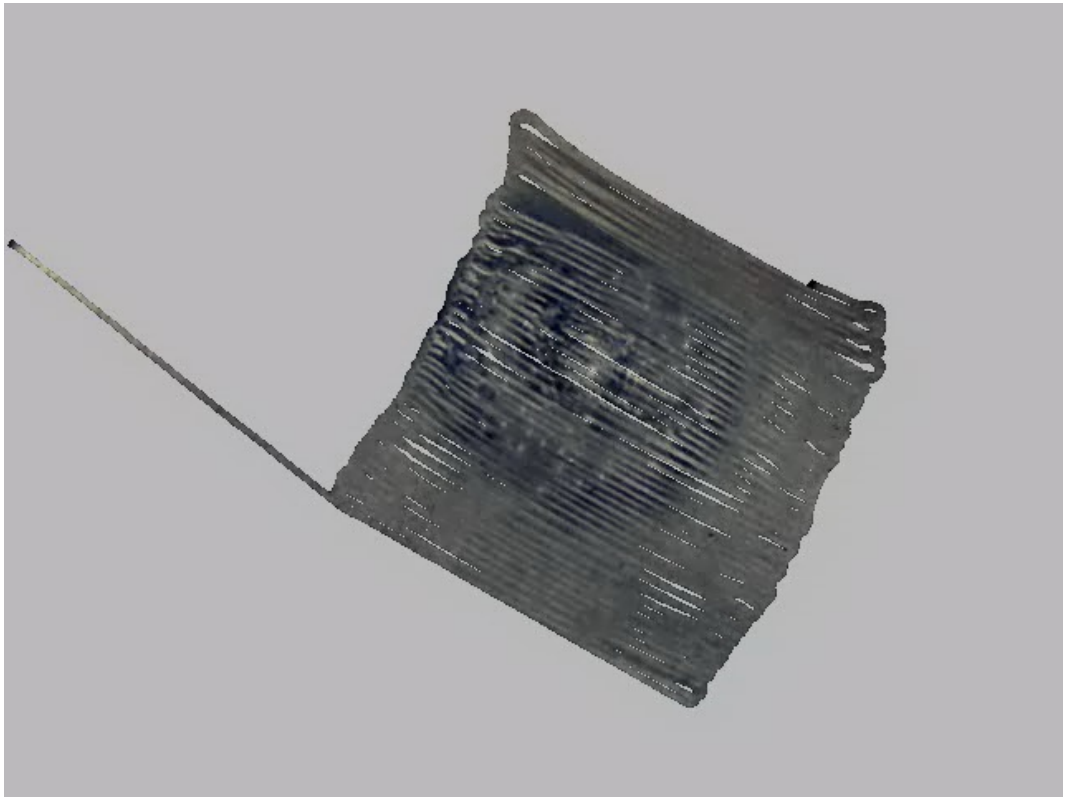

GPS
LED
Flash
Camera
LED
Flash

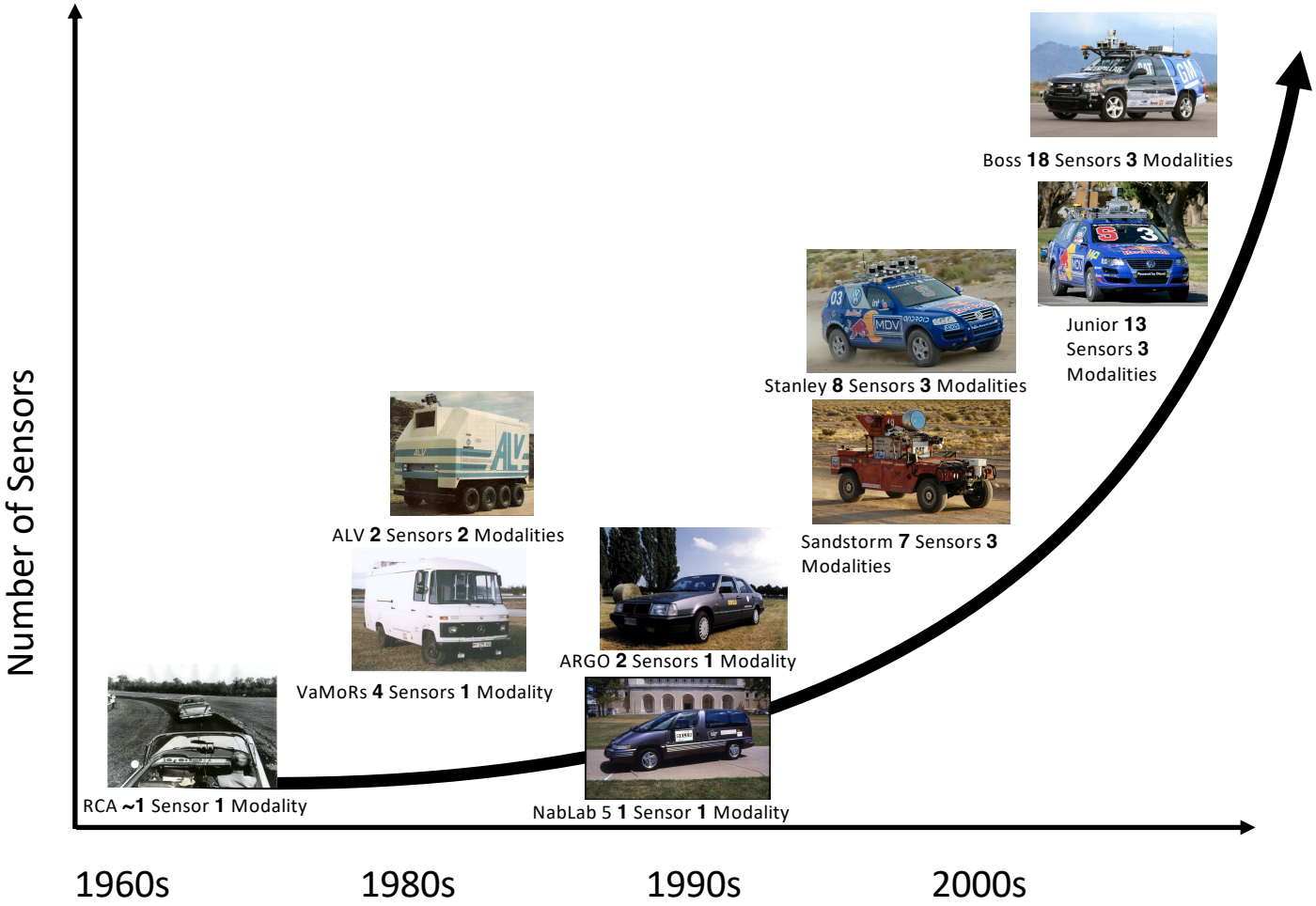
SENTRY
NSF
Woods Hole Oceanographic Institution



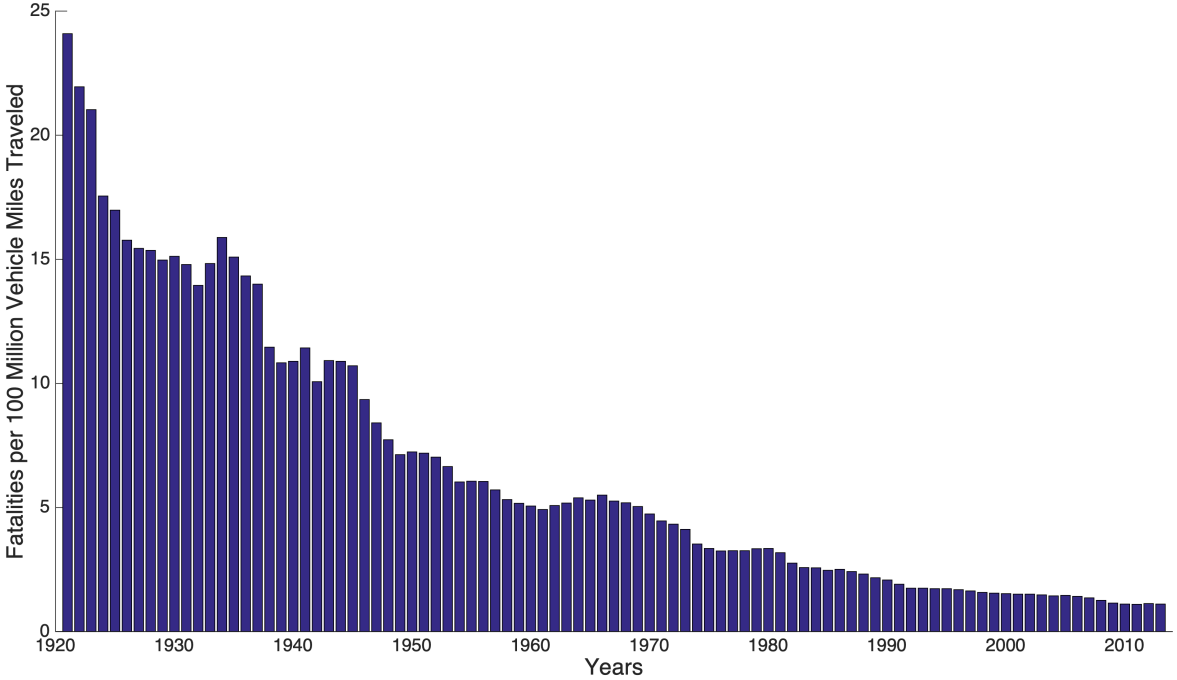


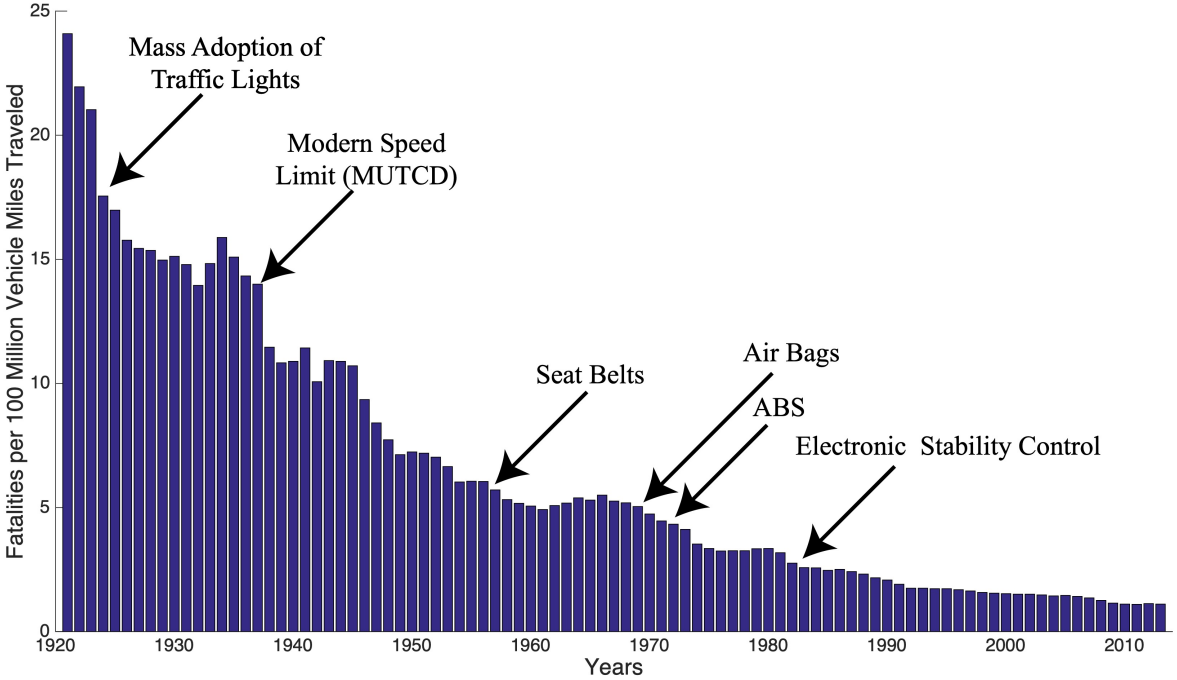





Number of Sensors
1960s
1980s
1990s
2000s
RCA ~1 Sensor 1 Modality
VaMoRs 4 Sensors 1 Modality
ALV 2 Sensors 2 Modalities
NabLab 5 1 Sensor 1 Modality
ARGO 2 Sensors 1 Modality
Sandstorm 7 Sensors 3 Modalities
Stanley 8 Sensors 3 Modalities
Junior 13 Sensors 3 Modalities
Boss 18 Sensors 3 Modalities

Why is this hard?


Fatalities per 100 Million Vehicle Miles Traveled
25
20
15
10
5
0
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
Years


Mass Adoption of
Traffic Lights
Modern Speed
Limit (MUTCD)
Seat Belts
Air Bags
ABS
Electronic Stability Control
Fatalities per 100 Million Vehicle Miles Traveled
Years
0
5
10
15
20
25
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010

How is learning relevant?

[redacted]

Kingfisher ASV

Seabotix ROV

BlueROV2 ROV

Bluefin HAUV

Iver AUV

REMUS 6000 AUV

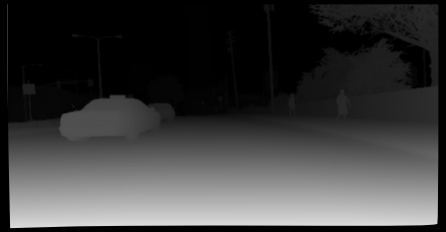
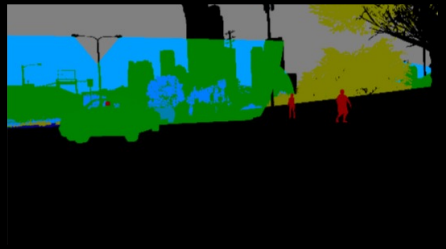
Ok Robots…. but how to escape labeling?

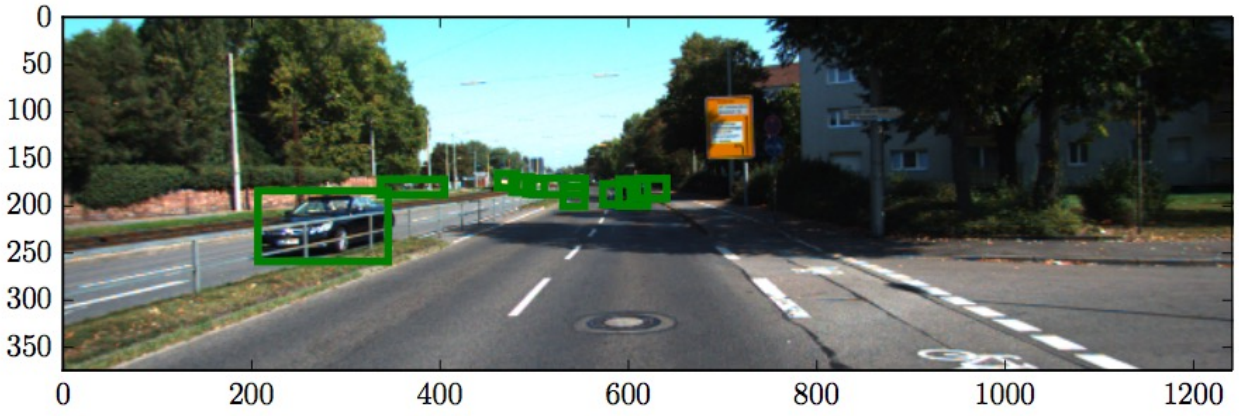
- Self-supervision/ Unsupervised Learning
- Physics
- Simulation

Can simulation
give us good
training data?

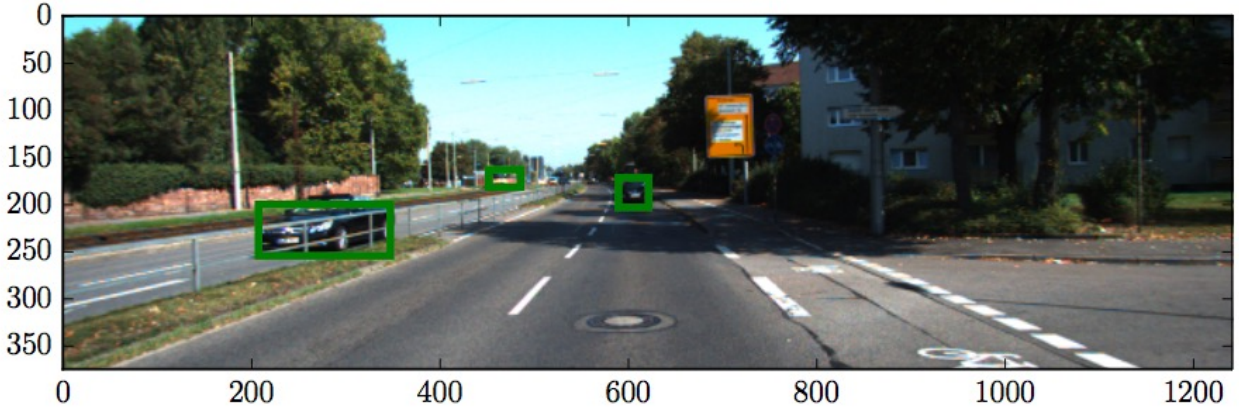
Discount
BEAUTY
STORE

Discount
BEAUTY
STORE

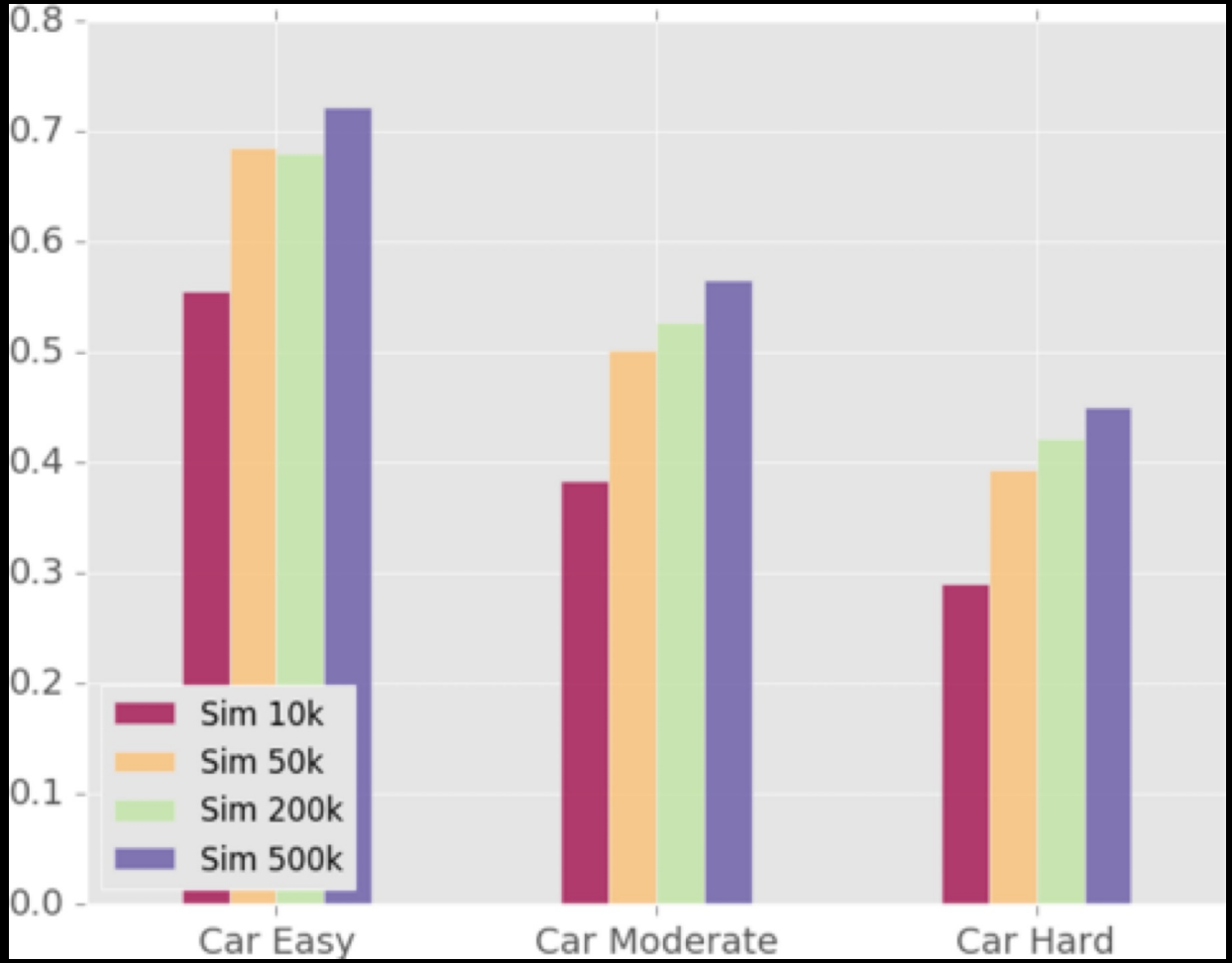



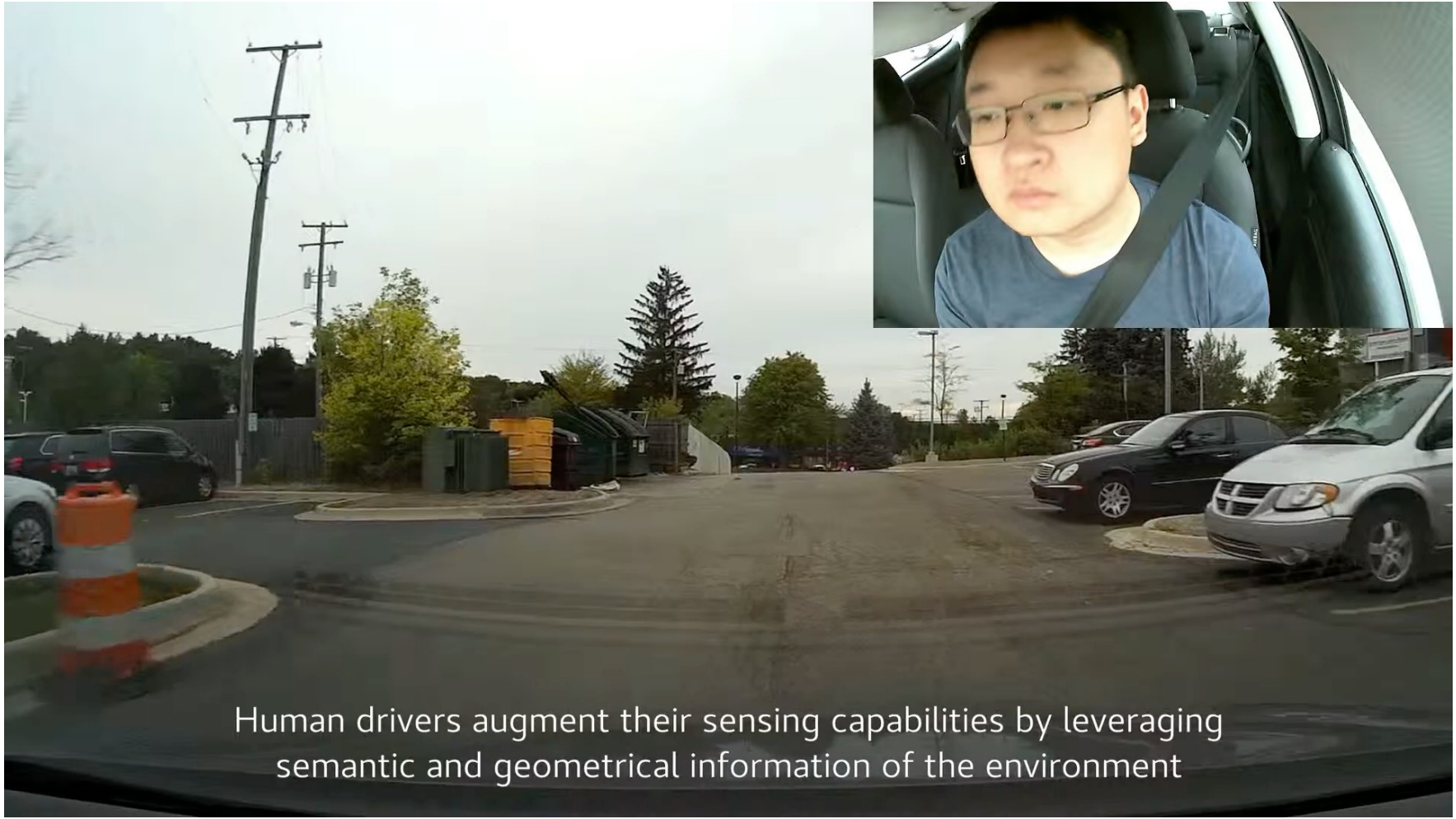


(a) Cityscapes



(b) 200k

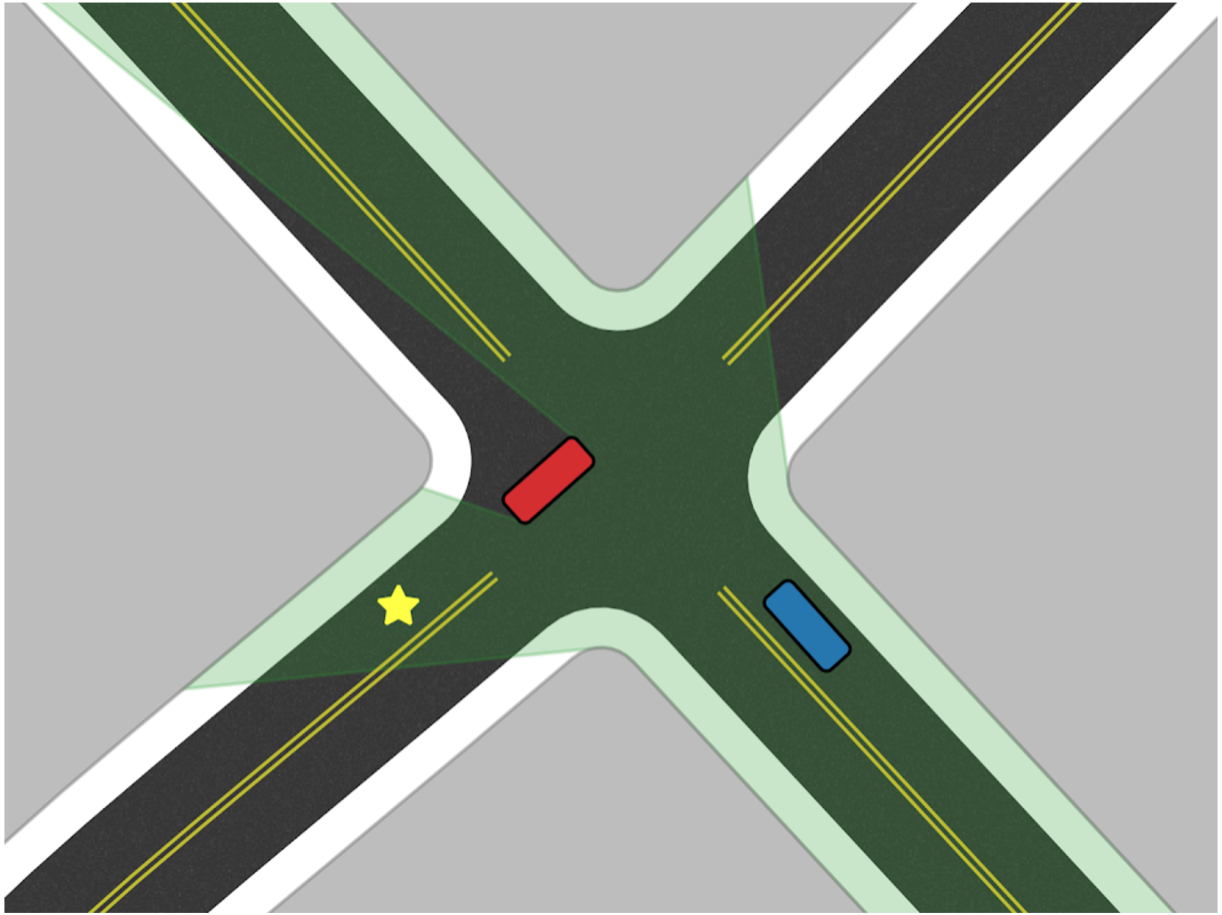

0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Sim 10k
Sim 50k
Sim 200k
Sim 500k
Car Easy
Car Moderate
Car Hard

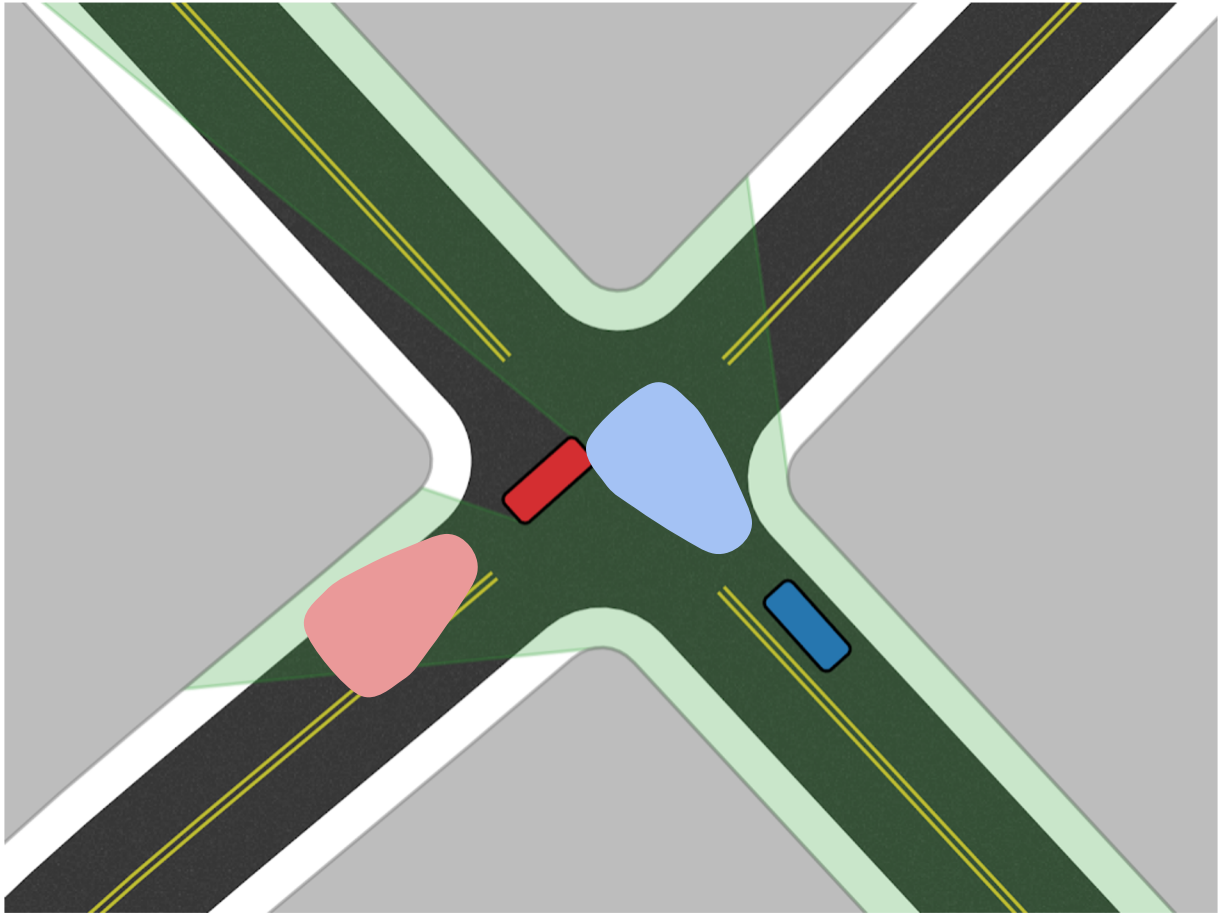

Human drivers augment their sensing capabilities by leveraging
semantic and geometrical information of the environment

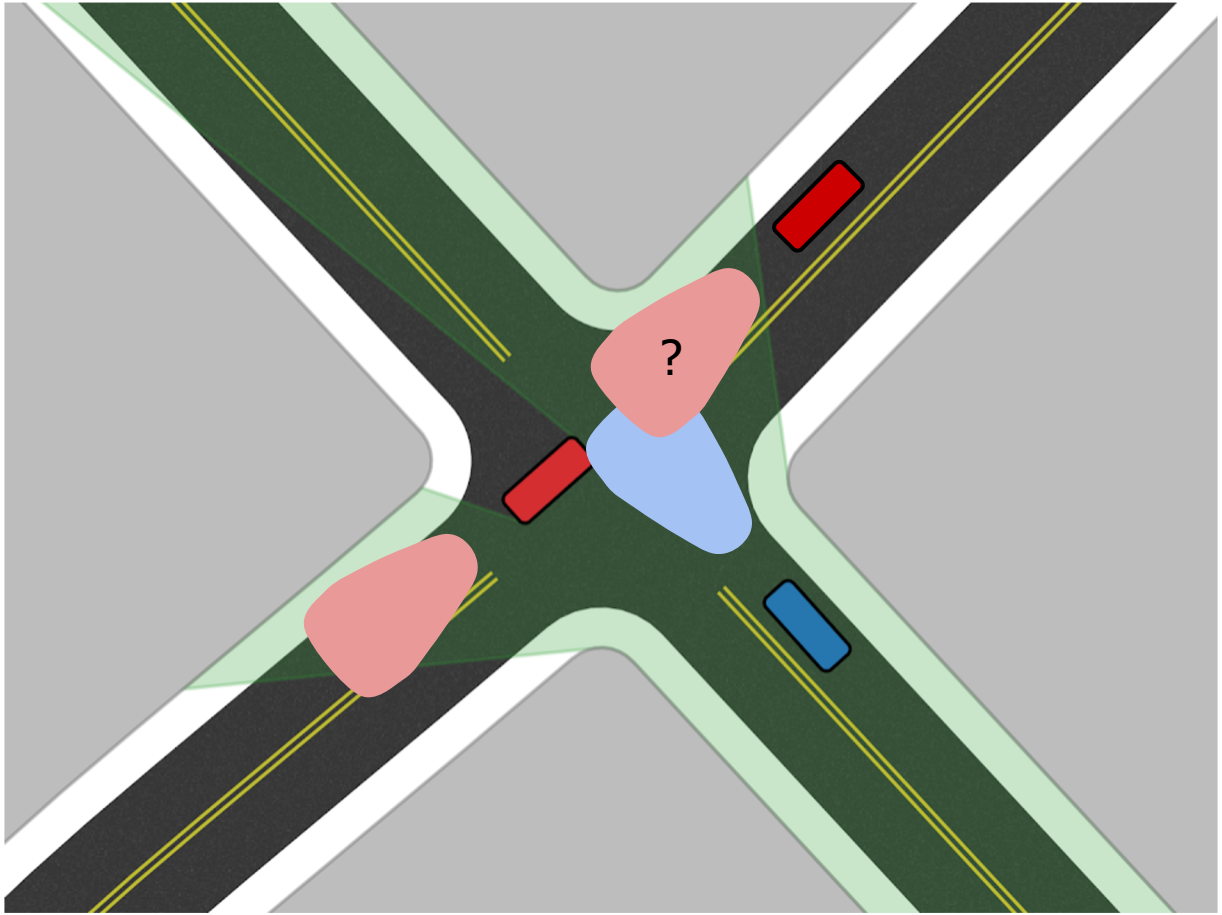
Risk Assessment and Planning with Bidirectional Reachability for Autonomous Driving

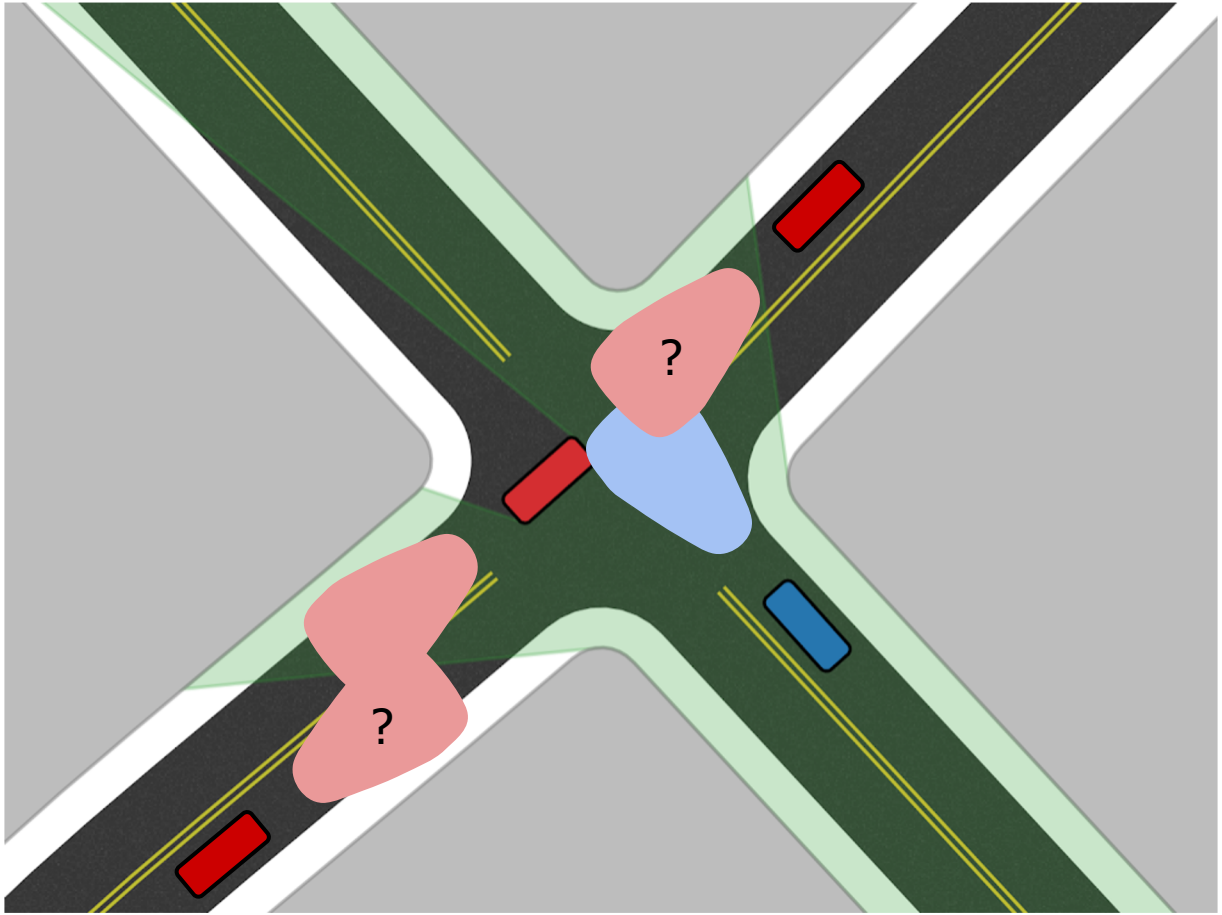
Ming-Yuan Yu
Ram Vasudevan
Matthew Johnson-Roberson

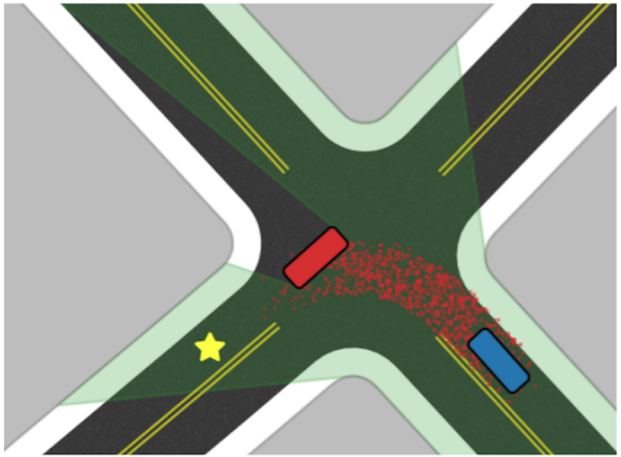
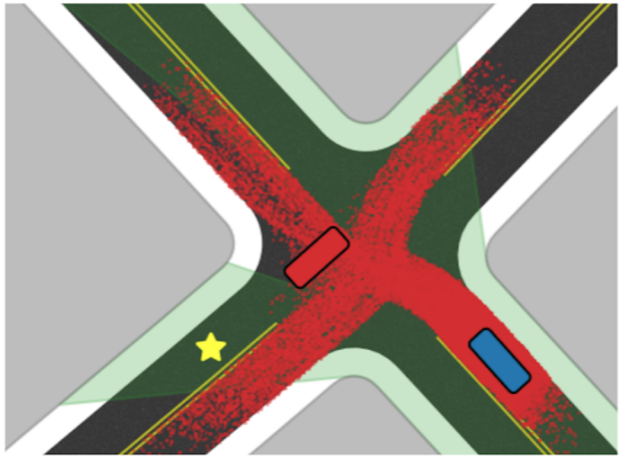
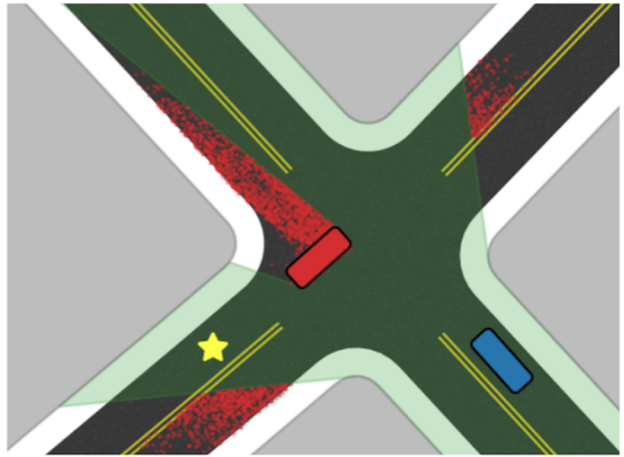
ICRA 2020

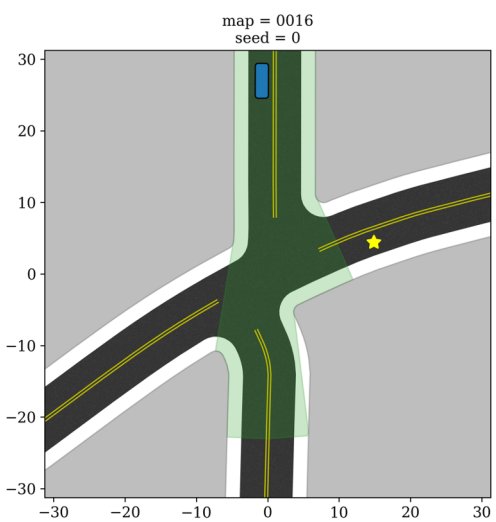


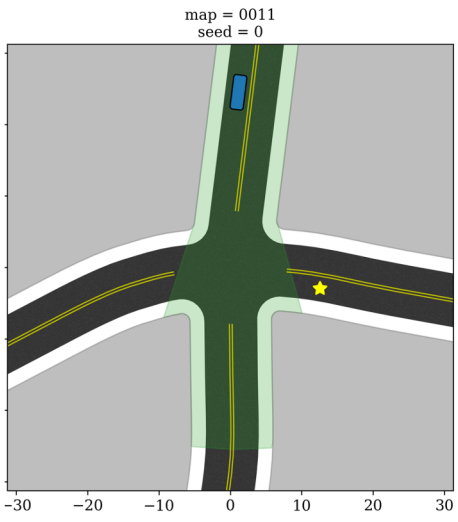


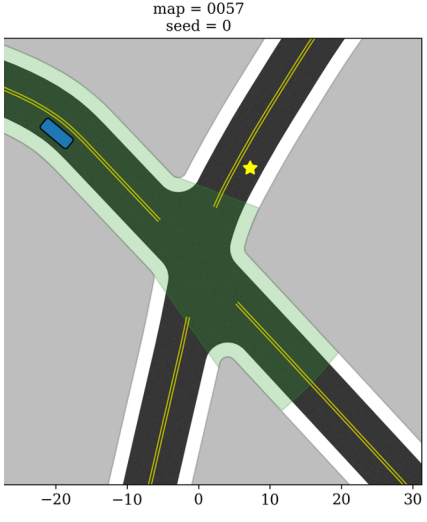


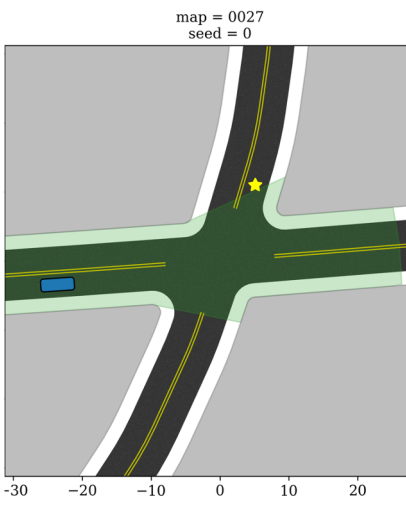


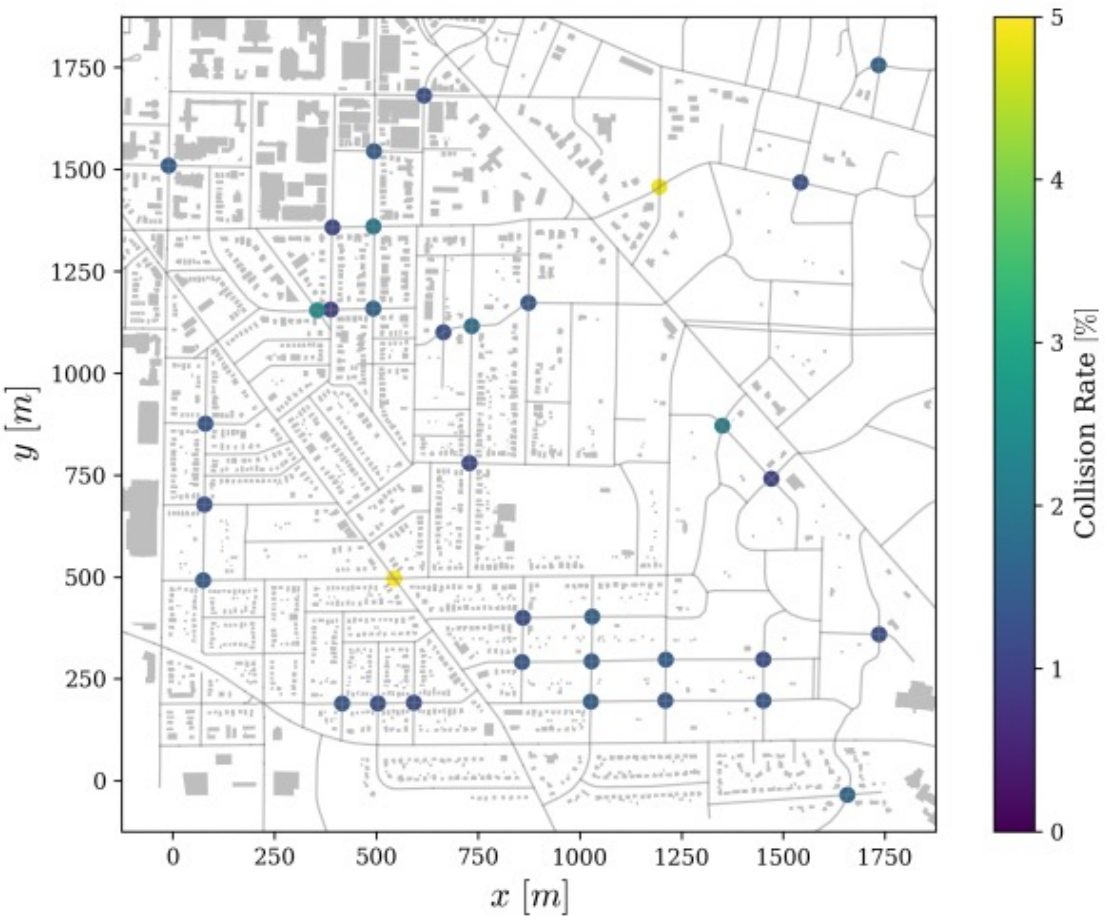





map = 0016
seed = 0
30
20
10
0
−10
−20
−30
−30
−20
−10
0
10
20
30


map = 0011
seed = 0
−30
−20
−10
0
10
20
30


map = 0057
seed = 0
−20
−10
0
10
20
30


map = 0027
seed = 0
−30
−20
−10
0
10
20


1750
1500
1250
1000
750
500
250
0
y [m]
0
250
500
750
1000
1250
1500
1750
x [m]
5
4
3
2
1
0
Collision Rate [%]

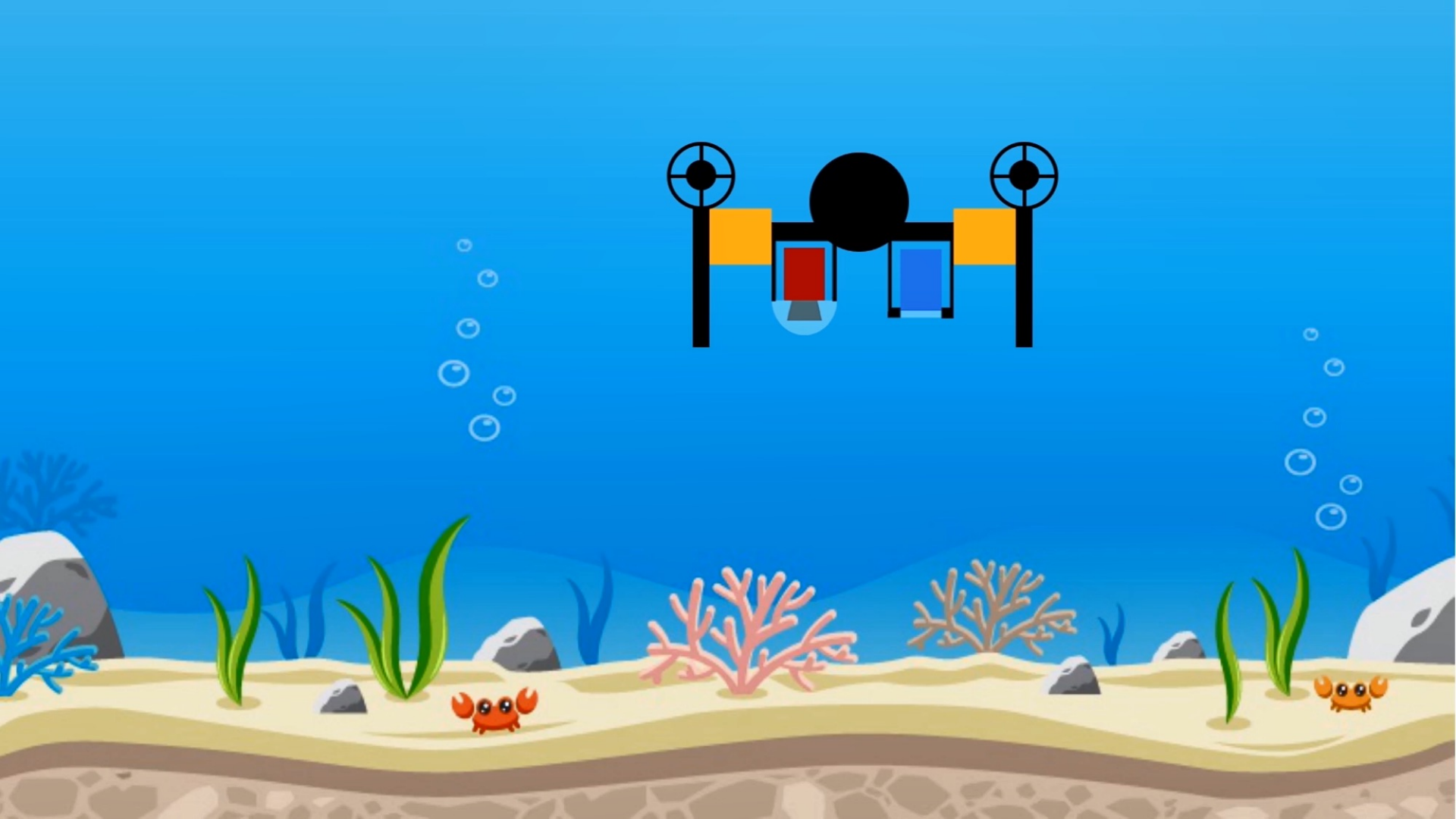
Underwater Image Physics

Kathrine Skinner

Now Assistant Professor
(University of Michigan

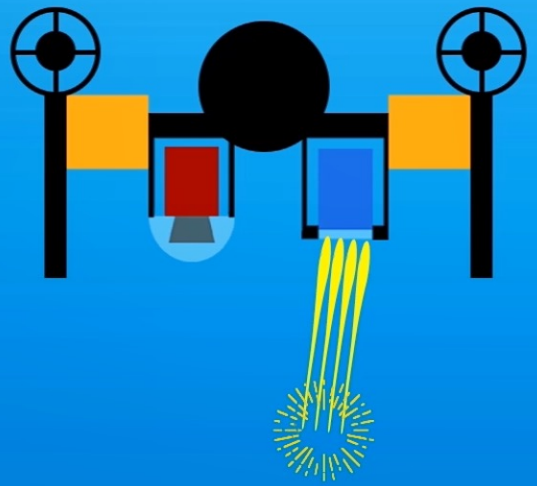
Jie Lie

Now Research Scientist
(Toyota Research Institute)



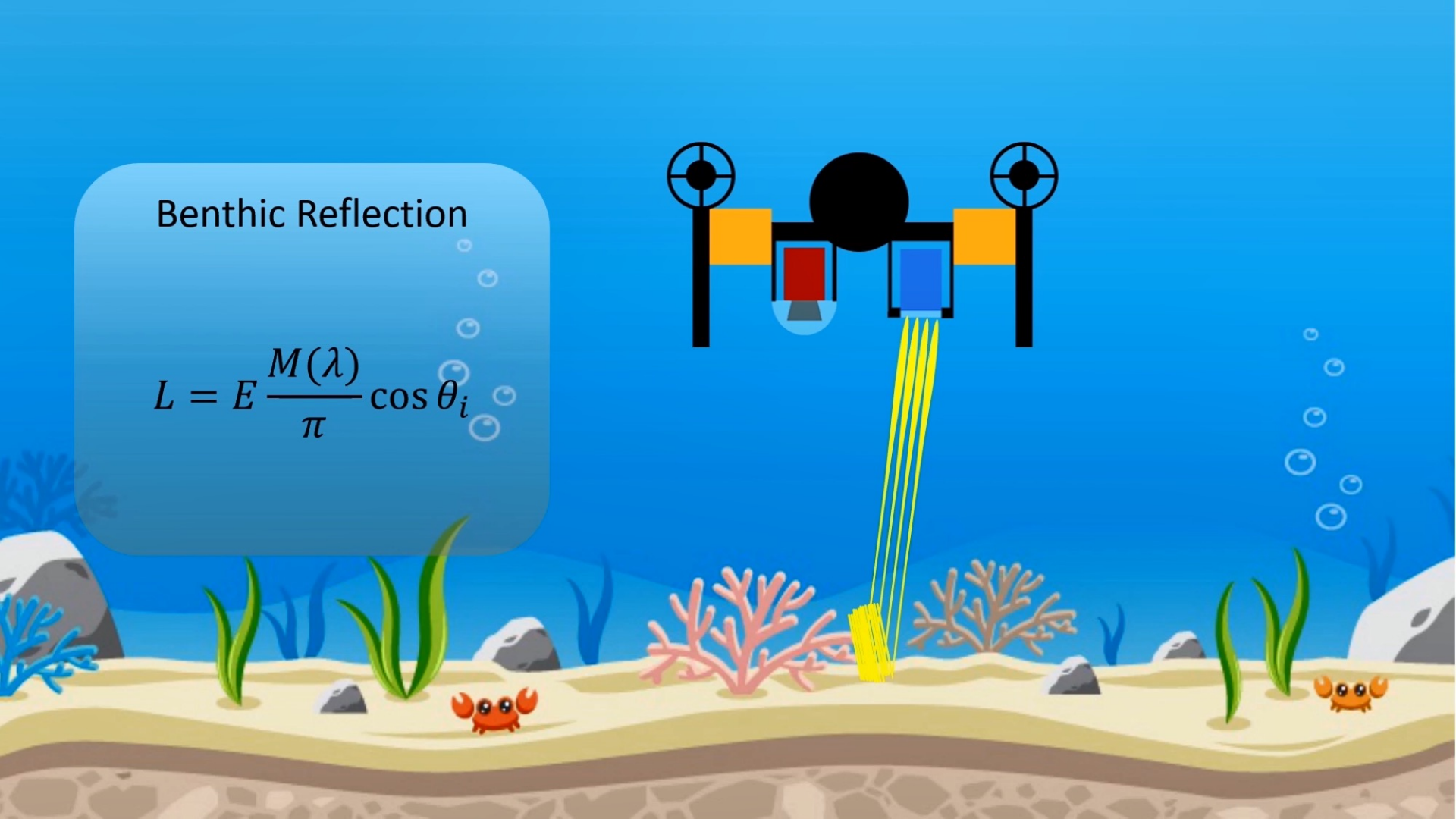
Water Column Attenuation

$$L = Re^{-b(\lambda)d}$$



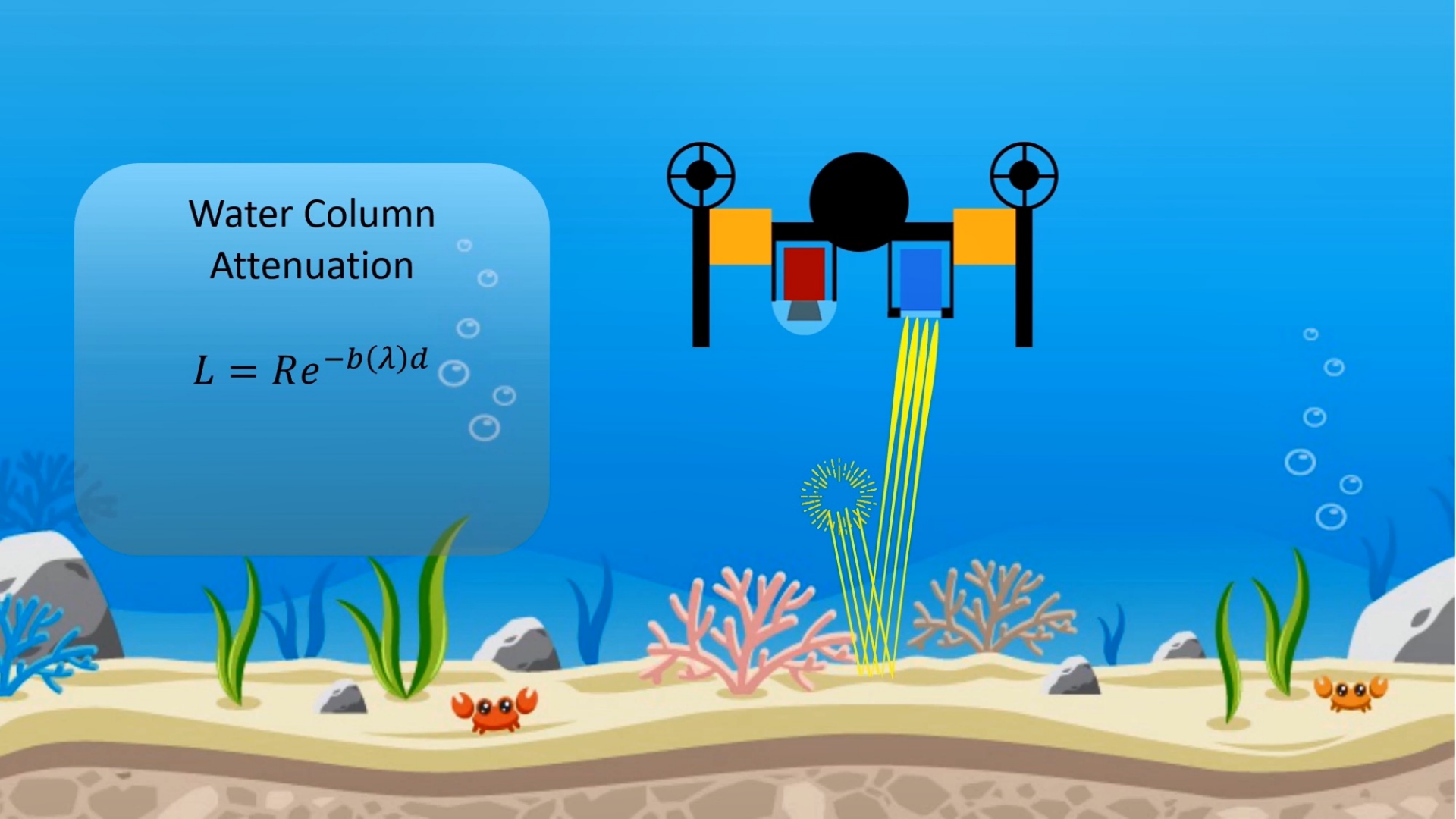
Benthic Reflection

$$L = E\frac{M(\lambda)}{\pi}\cos\theta_i$$



Water Column Attenuation

$$L = Re^{-b(\lambda)d}$$



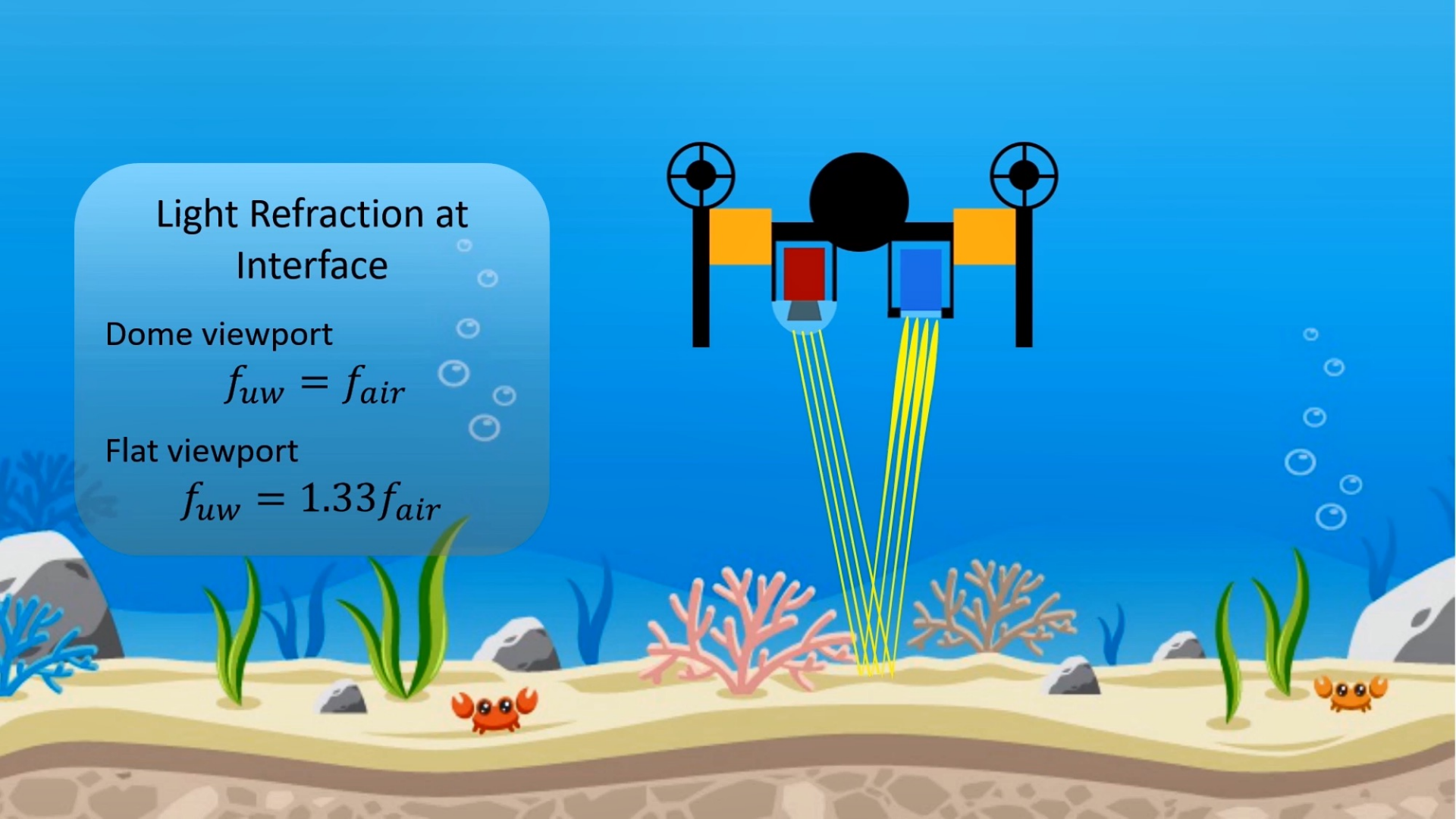
Light Refraction at Interface

Dome viewport

$$f_{uw} = f_{air}$$

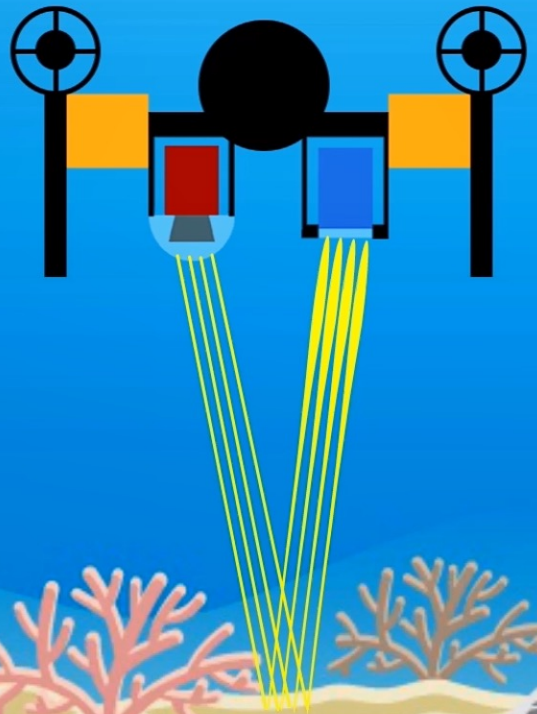
Flat viewport

$$f_{uw} = 1.33 f_{air}$$



Lensing Effects

$$E_I = L\frac{\pi}{4N^2}cos^4(\alpha)$$



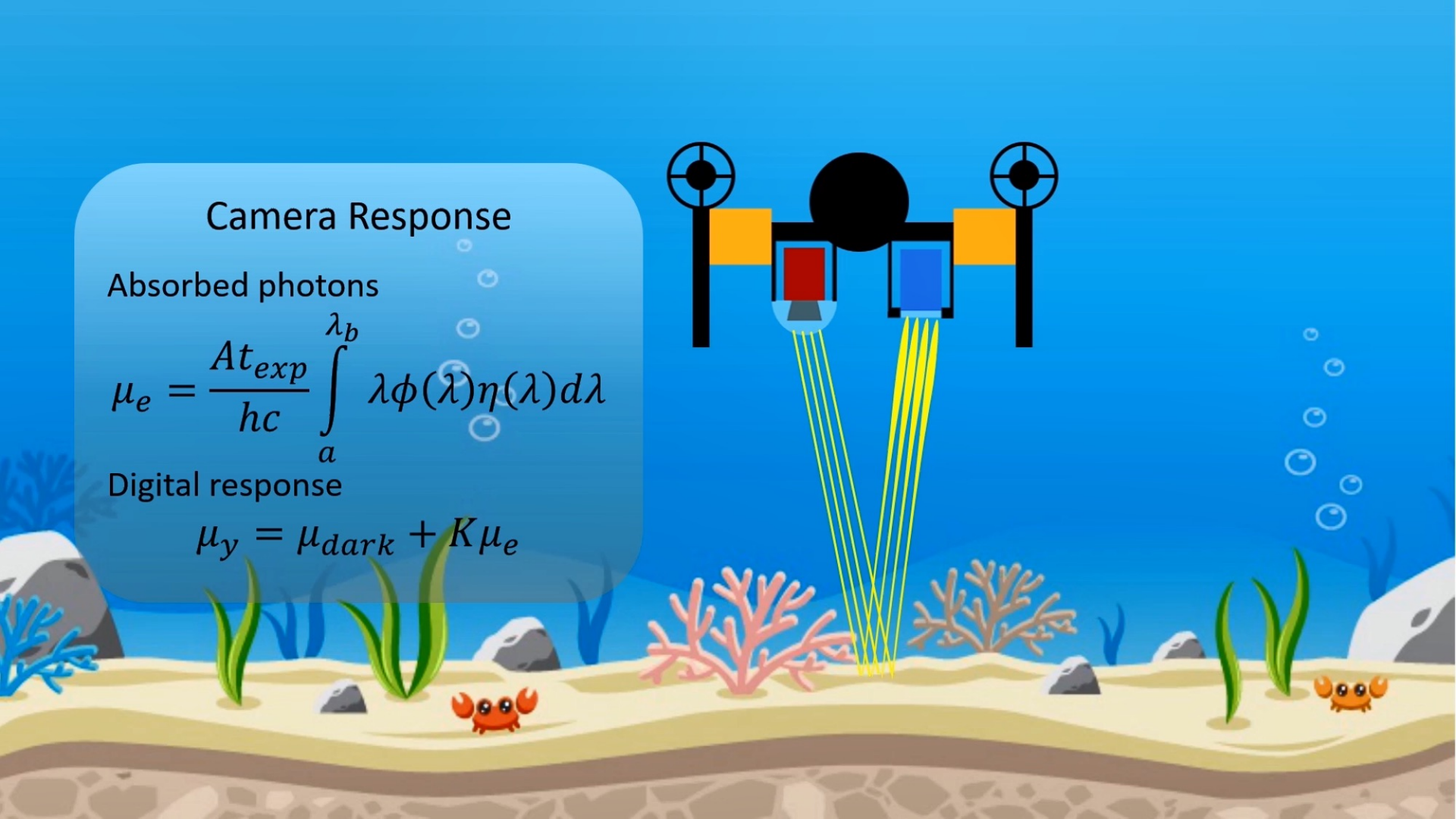
Camera Response

Absorbed photons

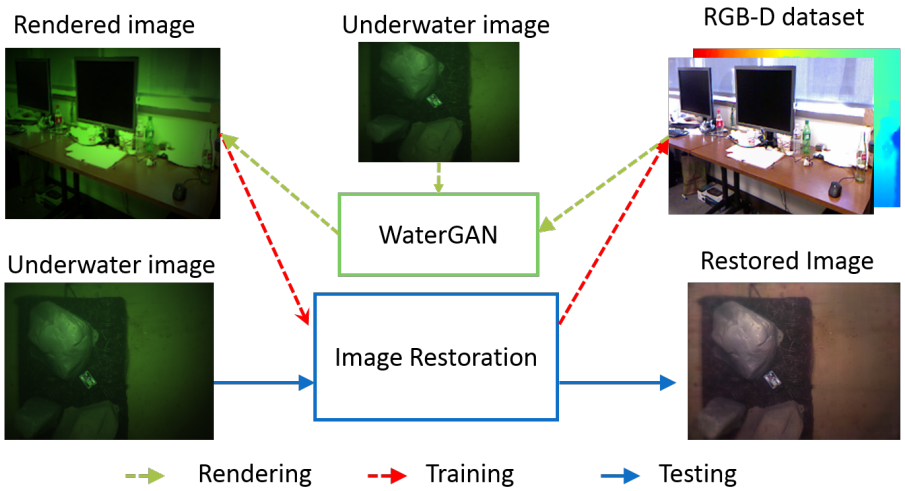
$$\mu_e = \frac{At_{exp}}{hc} \int_a^{\lambda_b} \lambda \phi(\lambda) \eta(\lambda) d\lambda$$

Digital response

$$\mu_y = \mu_{dark} + K\mu_e$$

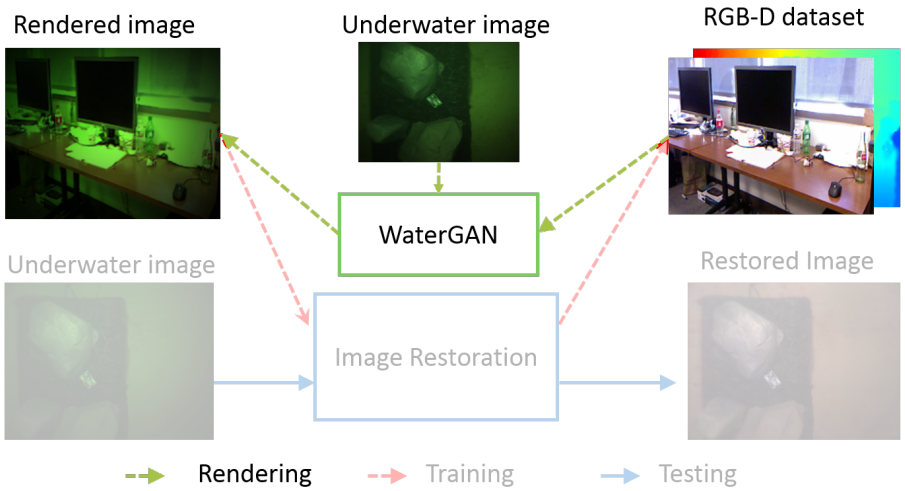


WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images



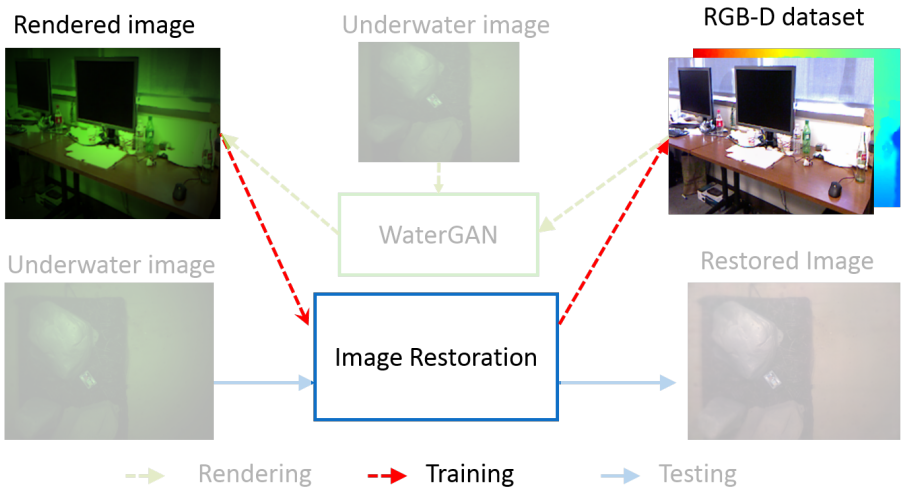
Jie Li*, Katherine A. Skinner*, Ryan Eustice and Matthew Johnson-Roberson, "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images." In IEEE RA-L, 2017. *The authors contributed equally to this work.

WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images



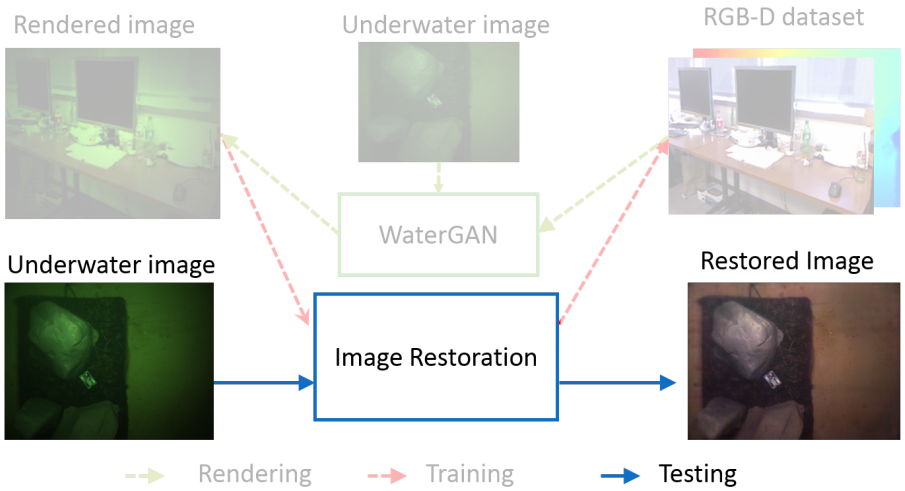
Jie Li*, Katherine A. Skinner*, Ryan Eustice and Matthew Johnson-Roberson, "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images." In IEEE RA-L, 2017. *The authors contributed equally to this work.

WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images



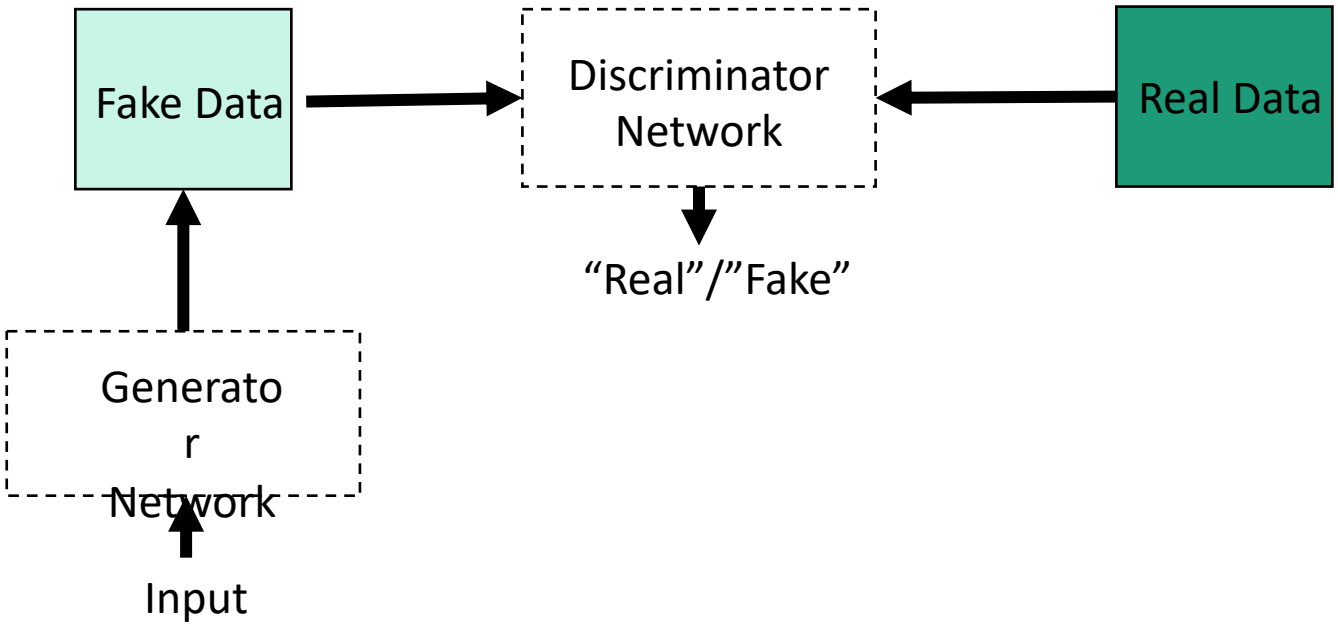
Jie Li*, Katherine A. Skinner*, Ryan Eustice and Matthew Johnson-Roberson, "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images." In IEEE RA-L, 2017. *The authors contributed equally to this work.

WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images



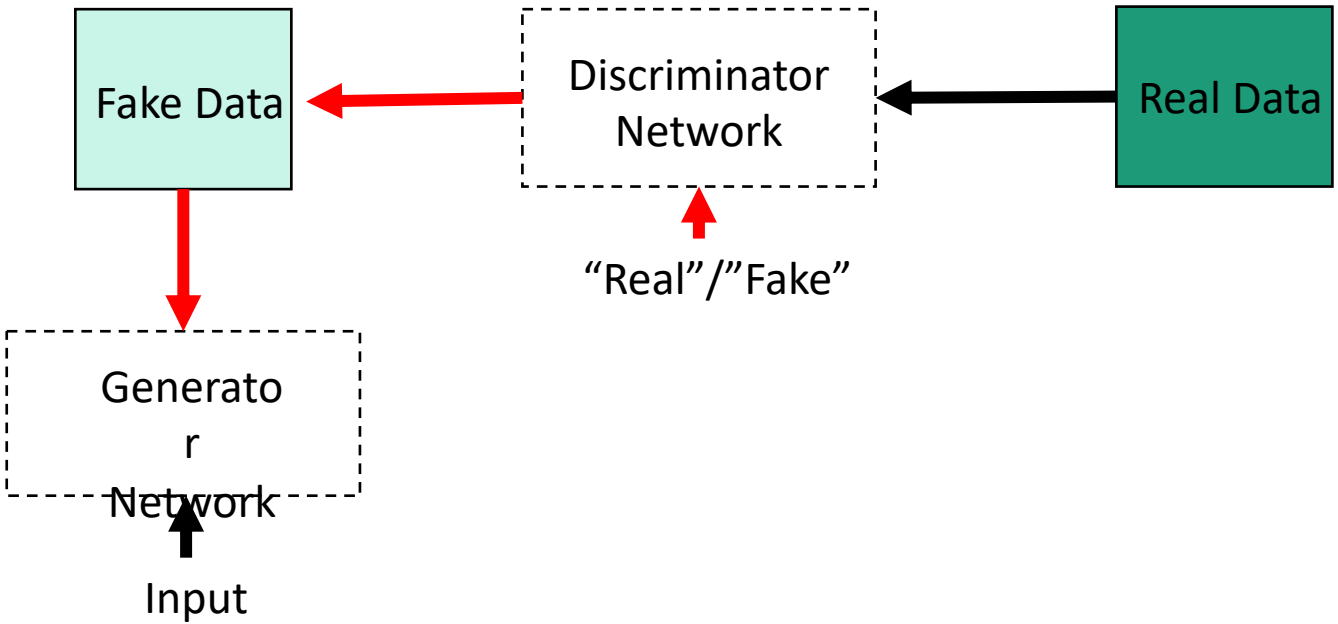
Jie Li*, Katherine A. Skinner*, Ryan Eustice and Matthew Johnson-Roberson, "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images." In IEEE RA-L, 2017. *The authors contributed equally to this work.

Generative Adversarial Networks (GANs)



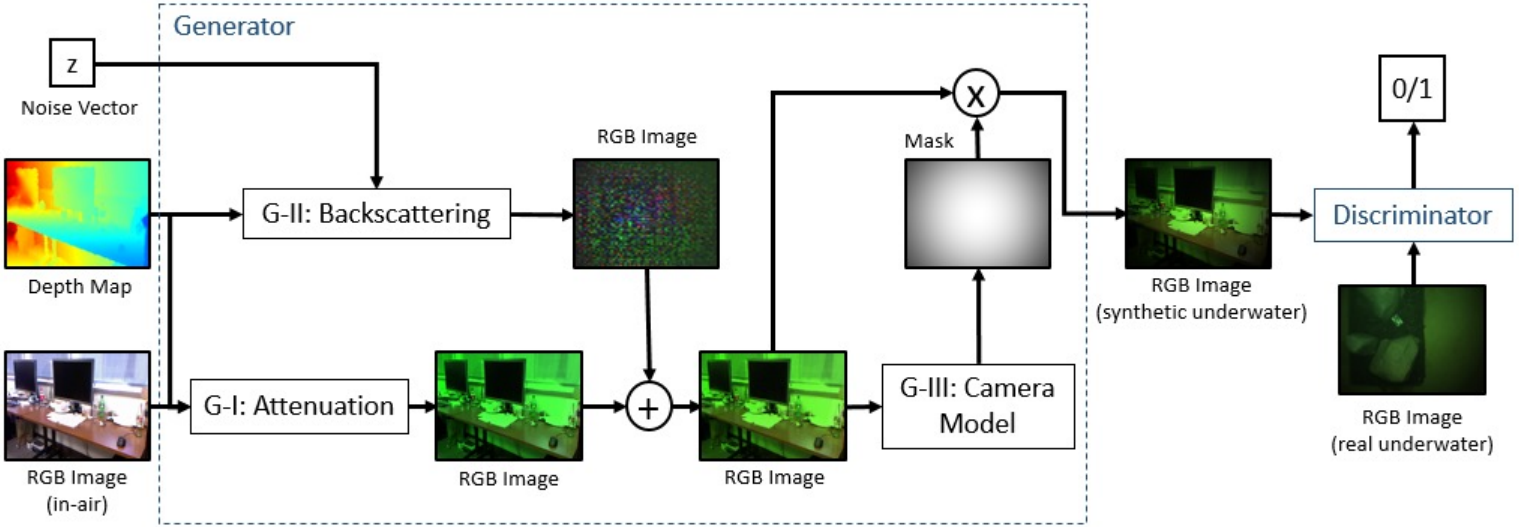
Reference: I. J. Goodfellow, et al. "Generative adversarial networks," *NIPS*, 2014.

Generative Adversarial Networks (GANs)

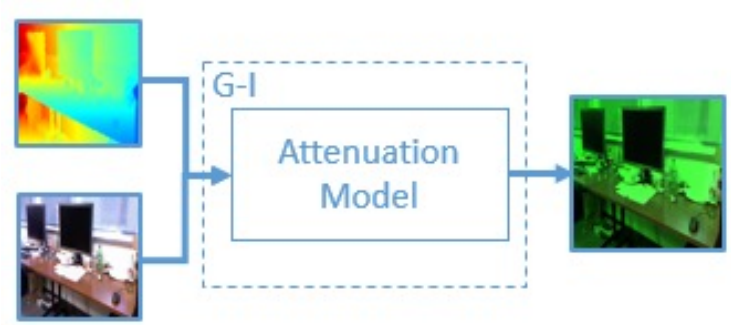


Reference: I. J. Goodfellow, et al. "Generative adversarial networks," *NIPS*, 2014.

Generating Realistic Underwater Images



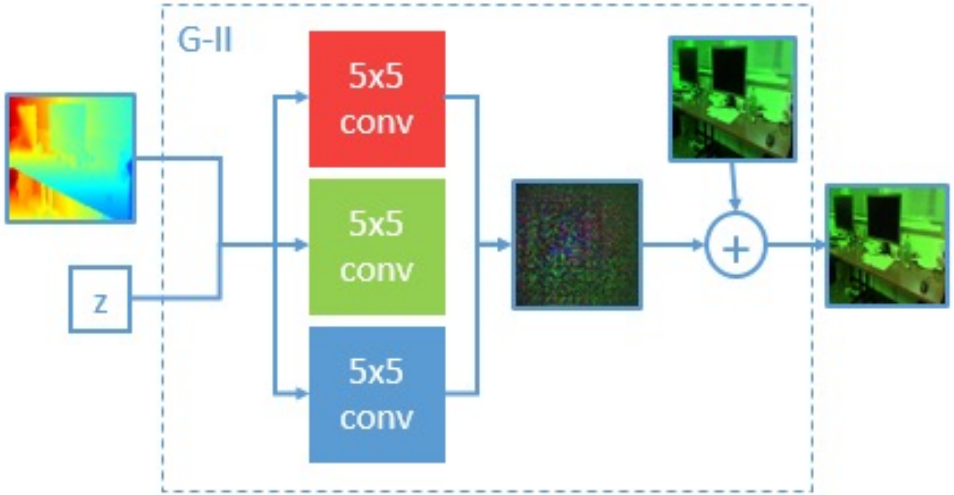
Stage G-I: Attenuation



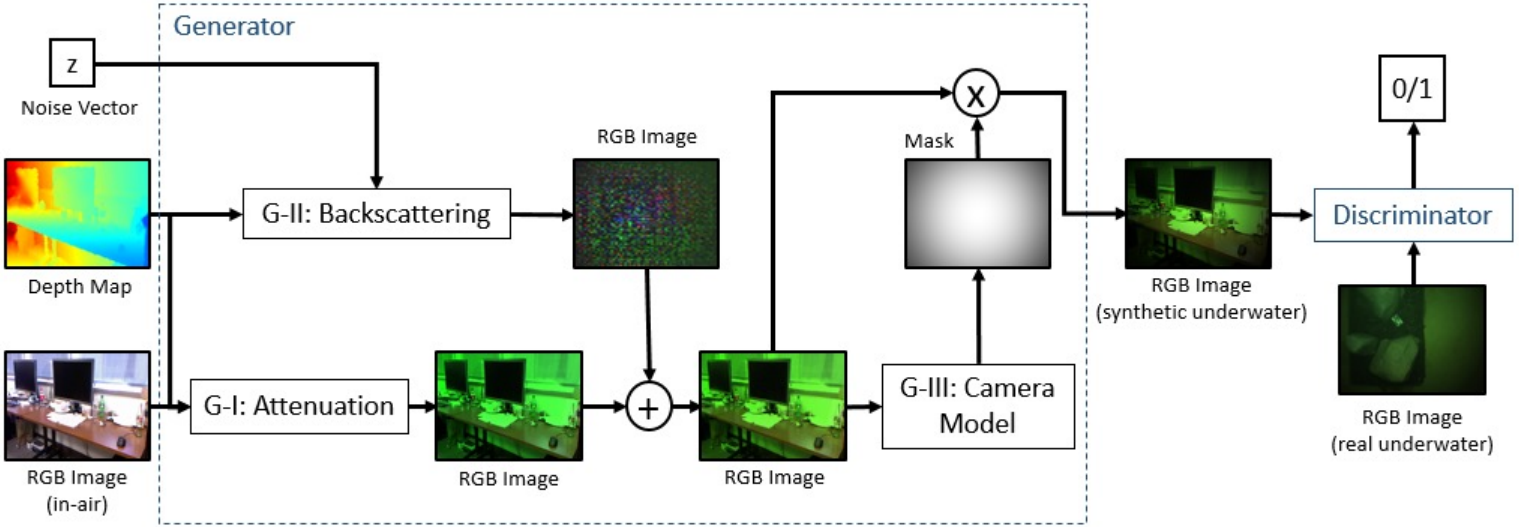
$$G_{1,C} = I_{air,C} e^{-\beta_C \Delta z}$$

Δz = Distance along line of sight
β_C = Effective wideband atten. coeff.
I = Image
C = Color channel

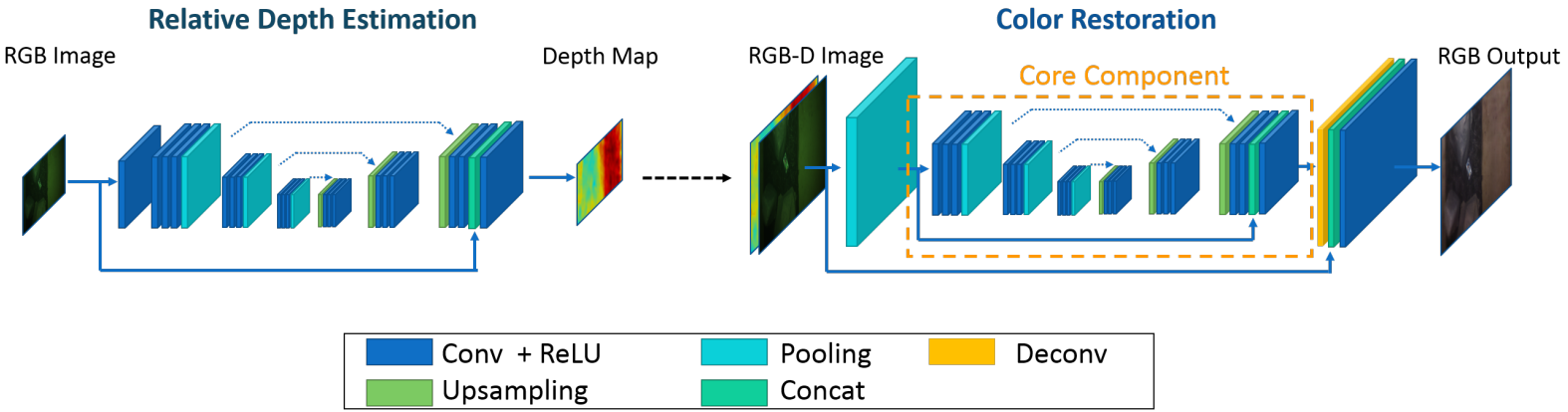
Stage G-II: Backscattering



Generating Realistic Underwater Images



Underwater Image Restoration Network

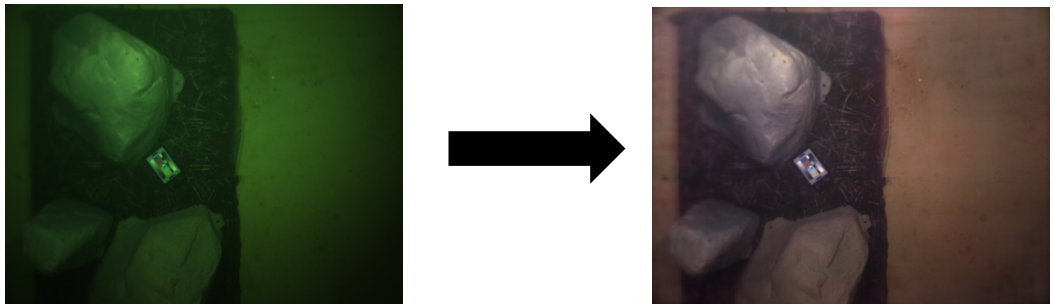


*Network developed by Jie Li.

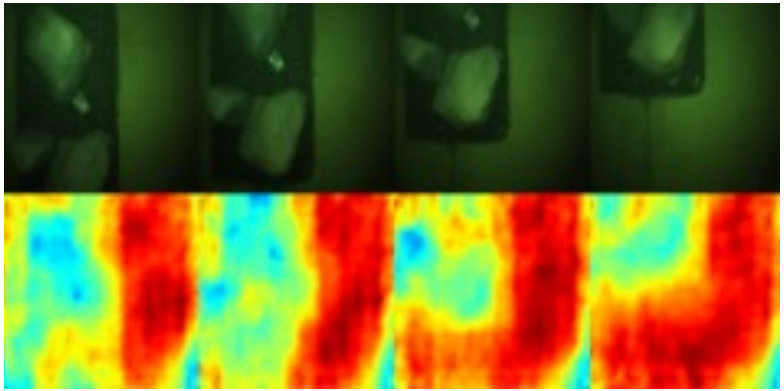
Results: Comparison to Other Methods

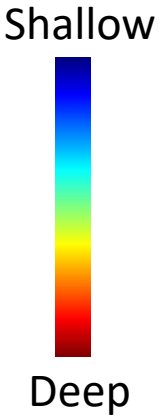
Raw | Histogram Equalization | Gray World | Modified Jaffe-Mcglamery | Shin et al. | WaterGAN

Unsupervised learning enabled by leveraging physics-based model of underwater image formation



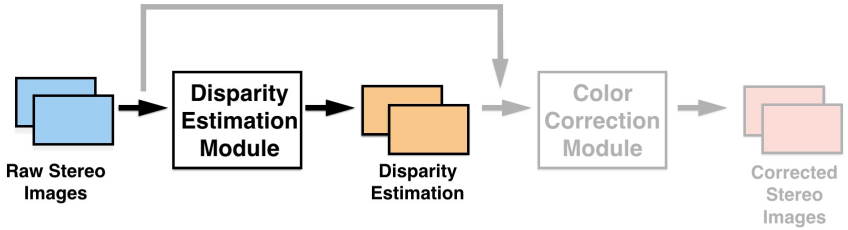
Depth Estimation

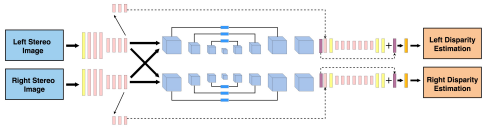




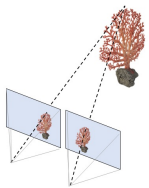
Results for WaterGAN depth estimation

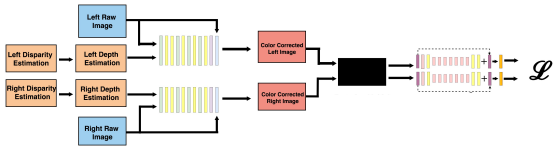
UWStereoNet





Geometric Constraints





Physics-based Models

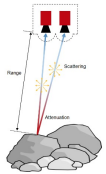
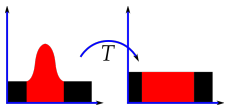
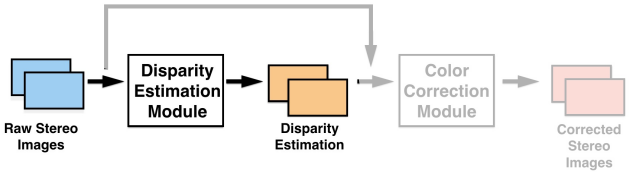


Image Processing



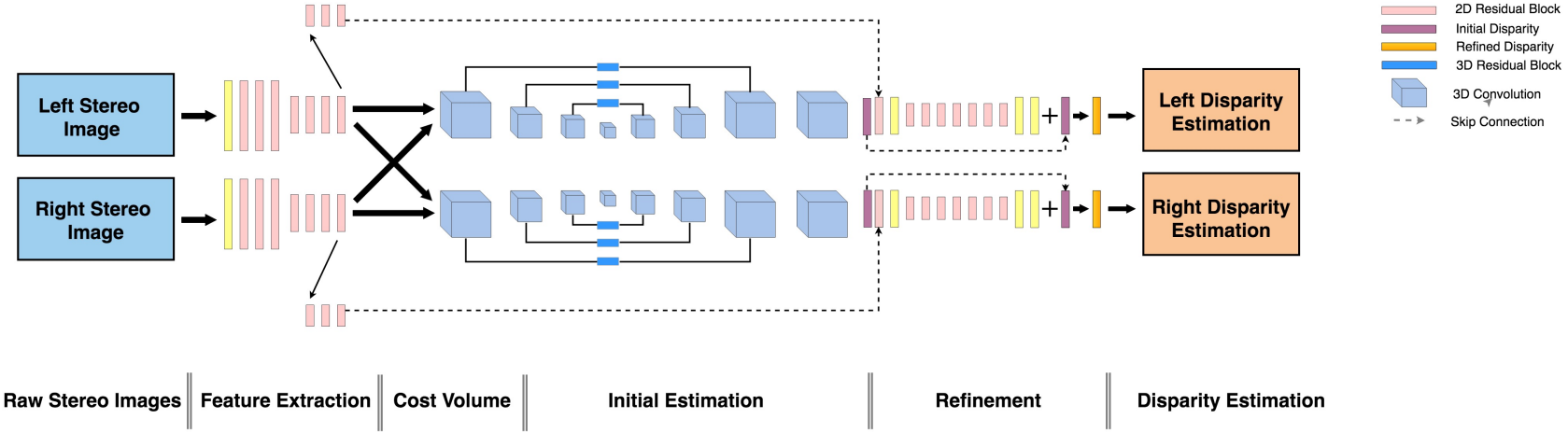
UWStereoNet



$$Loss = \alpha_1 L_{disp_init} + \alpha_2 L_{disp_ref}$$

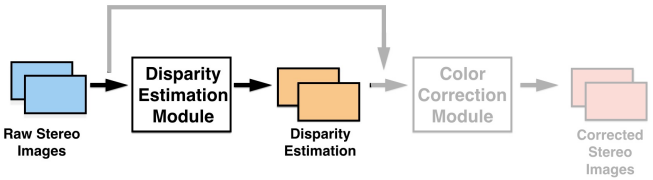
$$L_{disp_init} = \beta_1 L_{disp_warp} + \beta_2 L_{consist} + \beta_3 L_{reg}$$

$$L_{disp_ref} = \gamma_1 L_{disp_warp} + \gamma_2 L_{consist} + \gamma_3 L_{smooth}$$

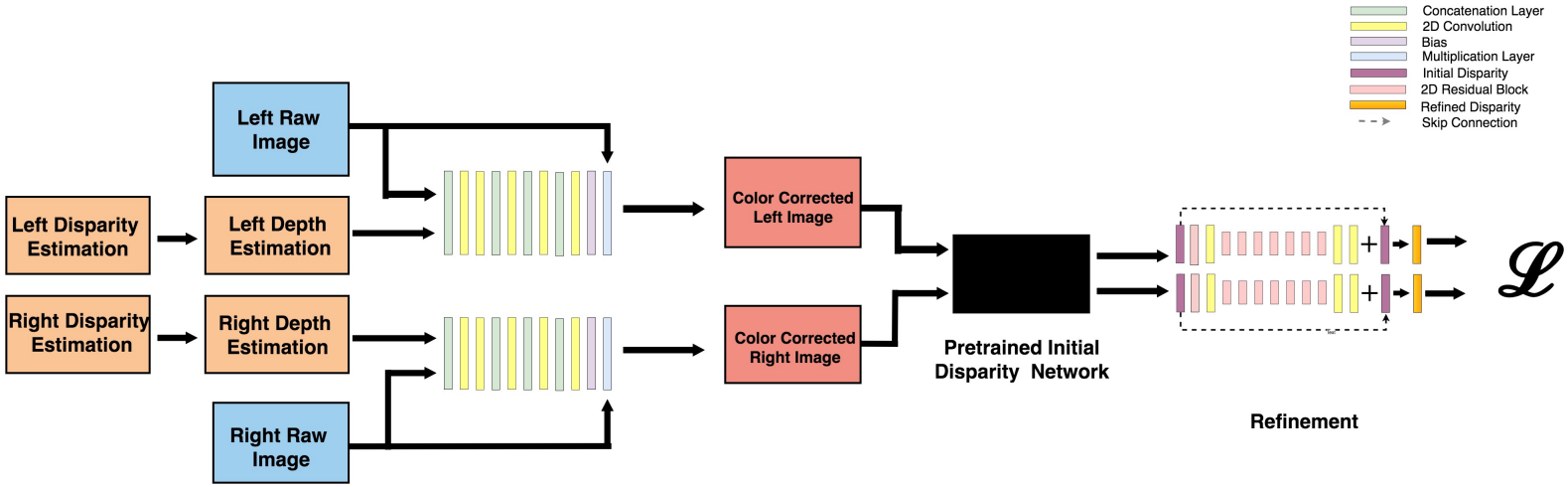


*Disparity network developed by Junming Zhang.

UWStereoNet

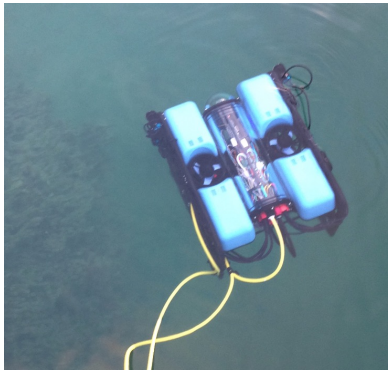


$$Loss = \theta_1 L_{gray} + \theta_2 L_{IQ} + \theta_3 L_{color_warp} + \theta_4 L_{color_cyc} + \theta_5 L_{disp_ref}$$

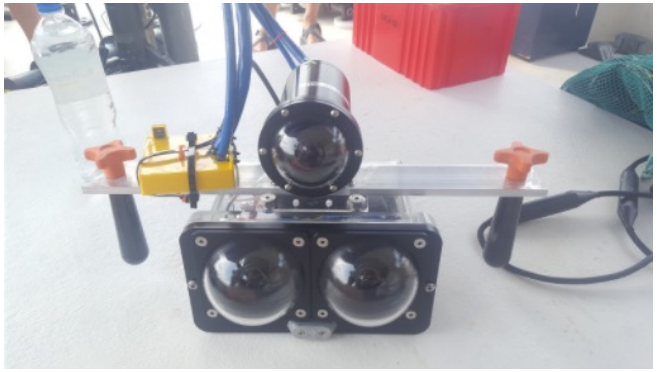


Experimental Setup

- Coral reef surveys near Bermuda Institute of Ocean Sciences (BIOS)
- Rock platform surveys near Hawaii Institute of Marine Biology (HIMB)

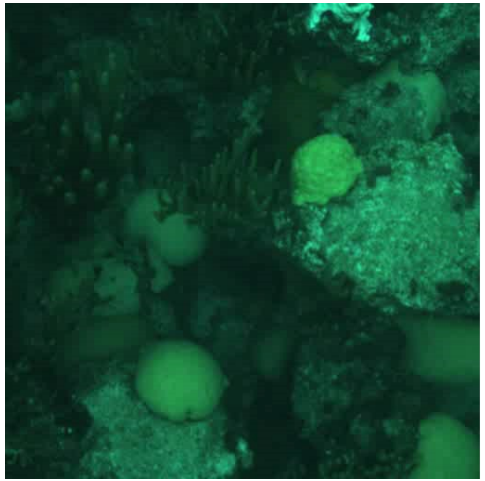


BlueROV survey at HIMB with ground truth structure and color

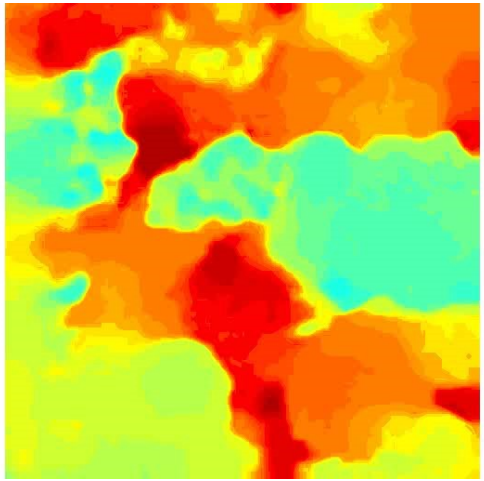


Diver rig survey near BIOS

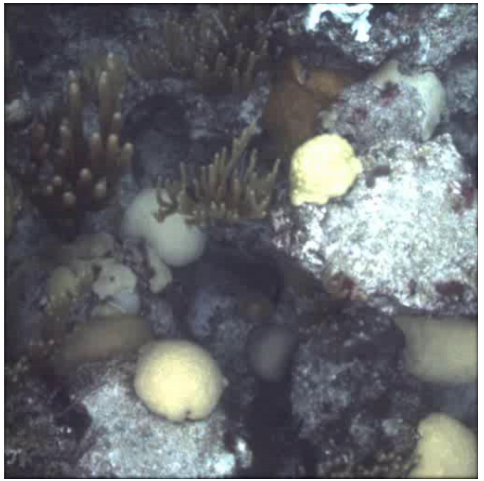
Results



Raw Image



Depth Map

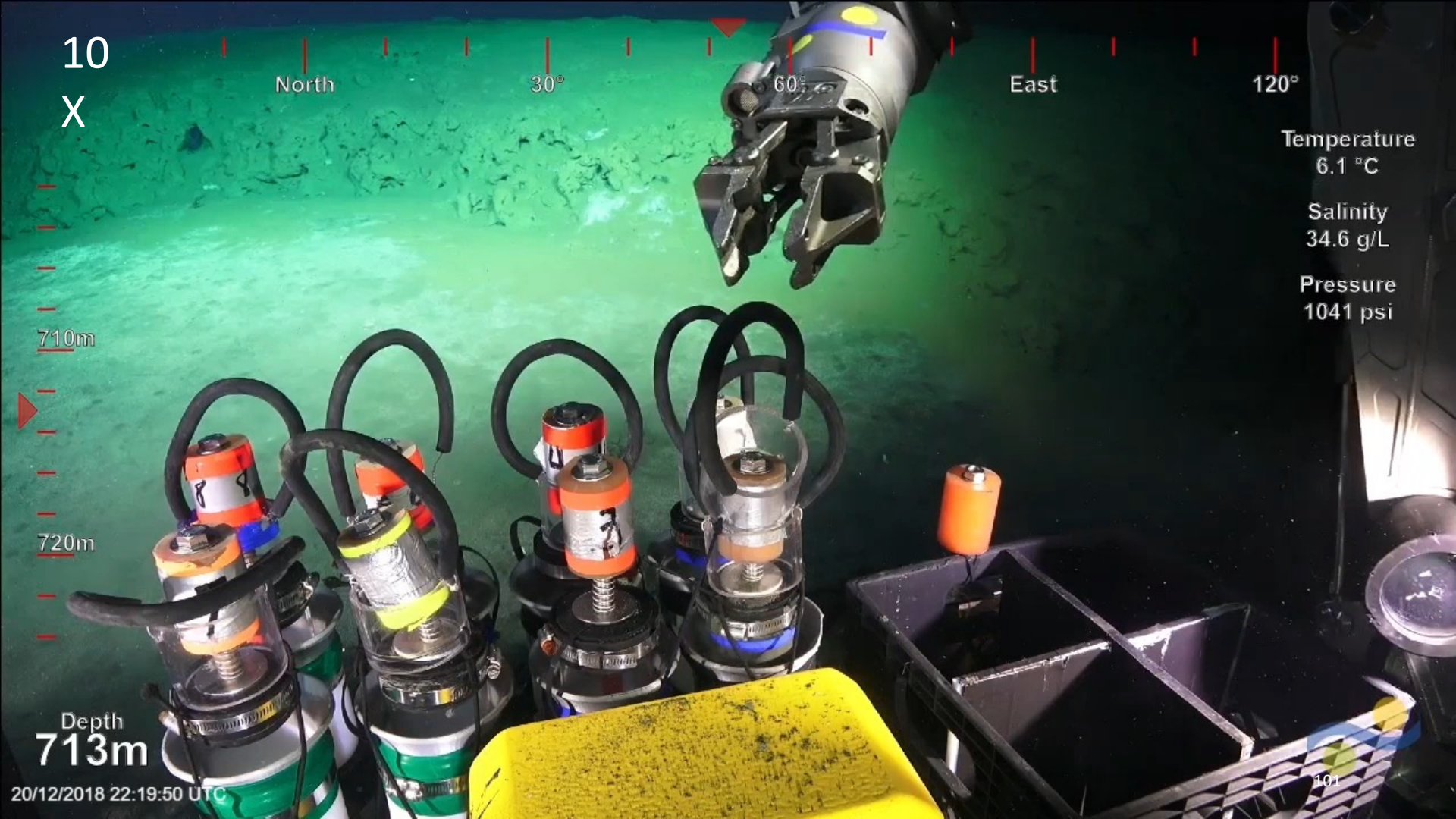


Corrected Color

Underwater Manipulation

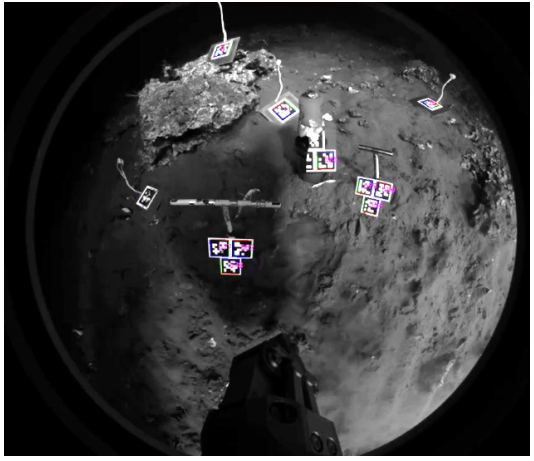
Gideon Billings

PhD Graduate
University of Michigan

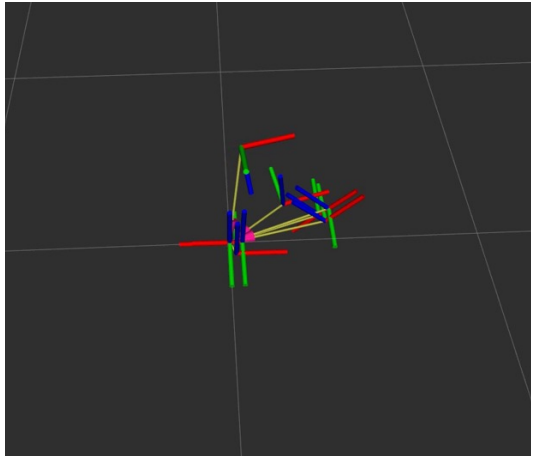

10
X
North
30°
60°
East
120°
Temperature
6.1 °C
Salinity
34.6 g/L
Pressure
1041 psi
710m
720m
Depth
713m
20/12/2018 22:19:50 UTC

UWHandles – Underwater Handle Pose Estimation Dataset

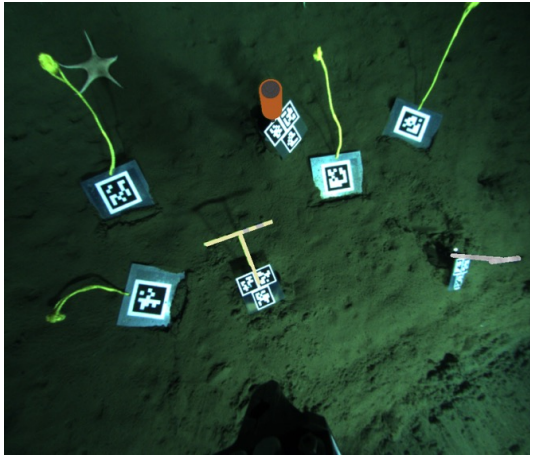
1. Detect Tags



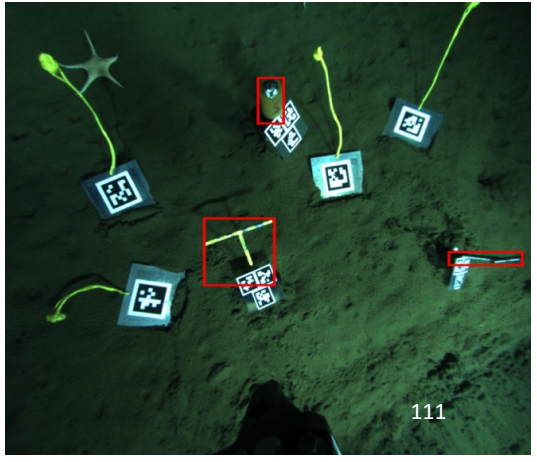
2. Tag SLAM



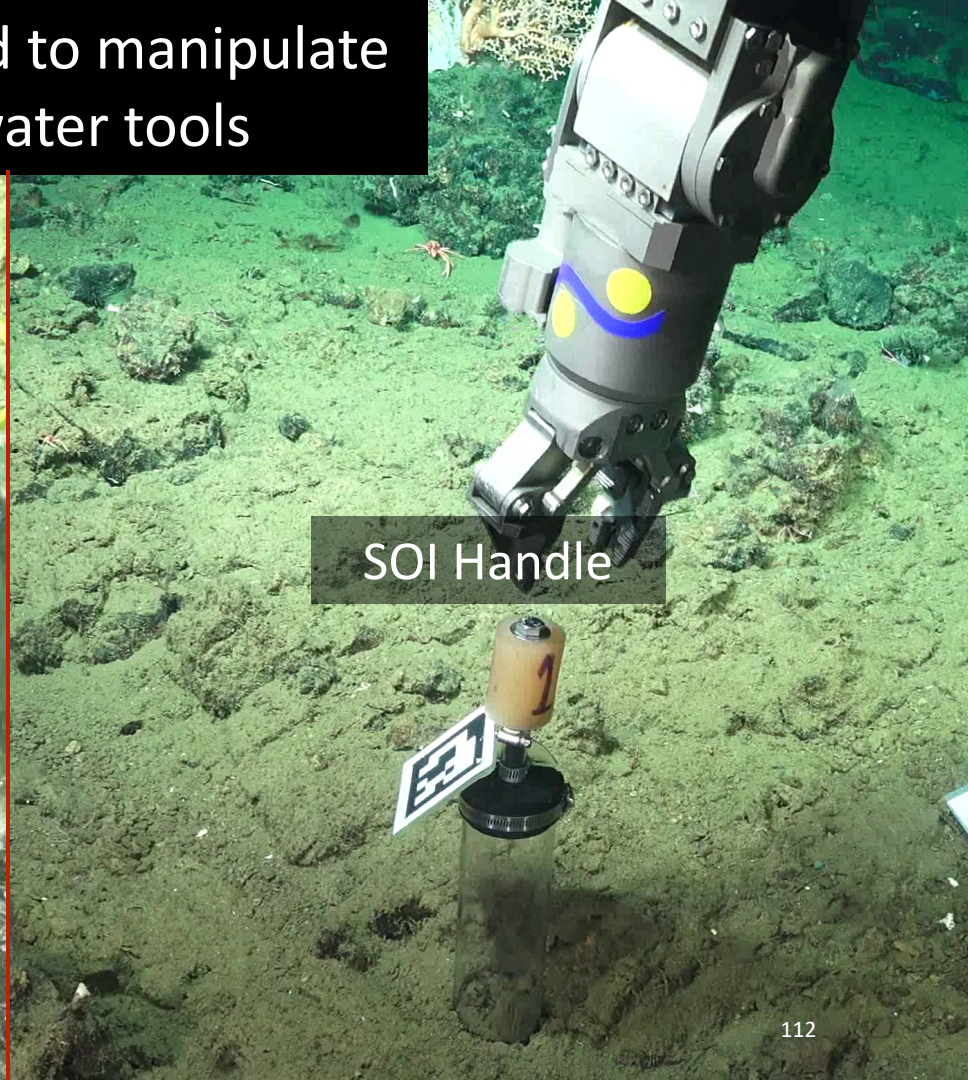
3. Fit 6D poses



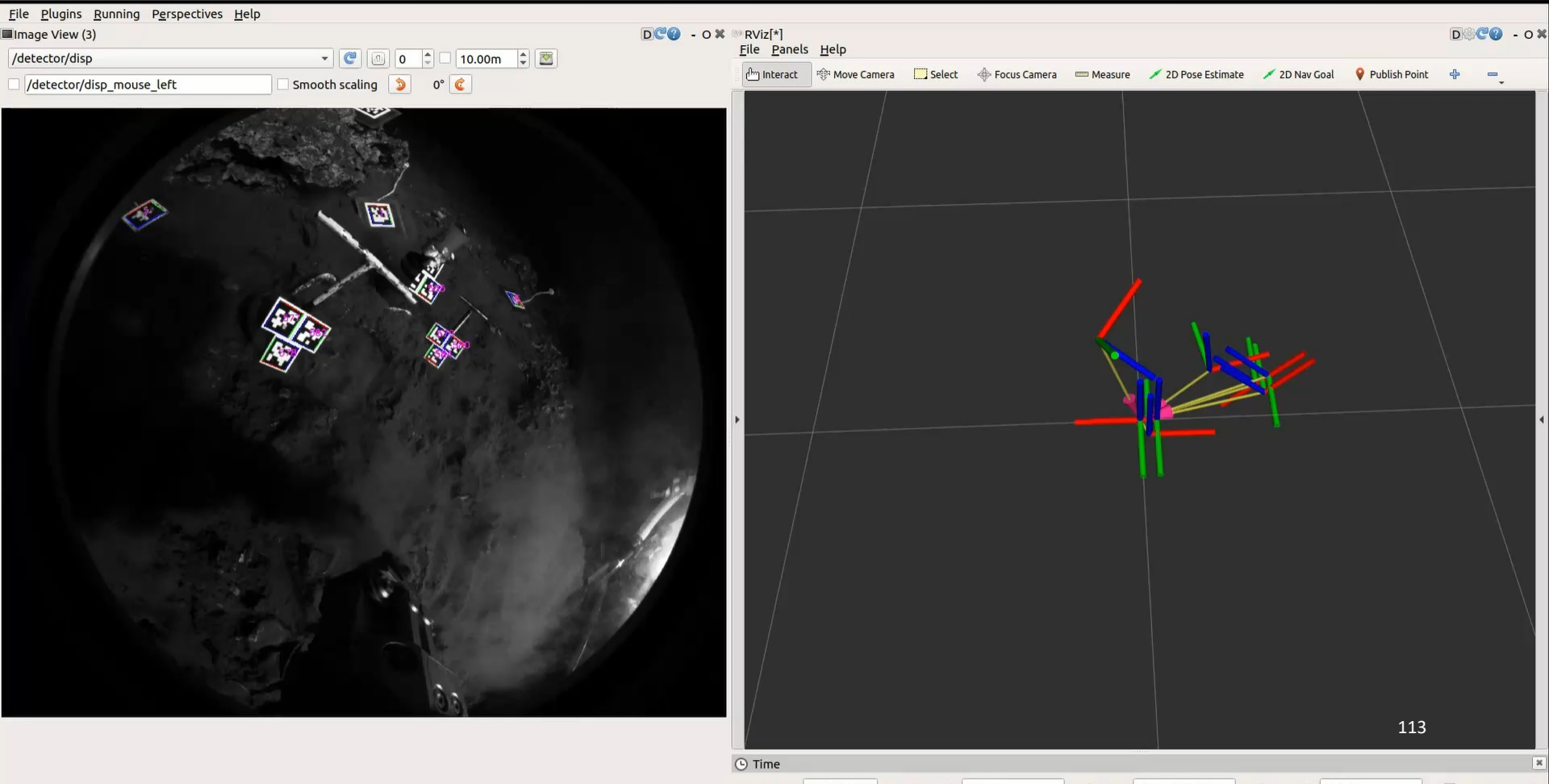
4. Annotate



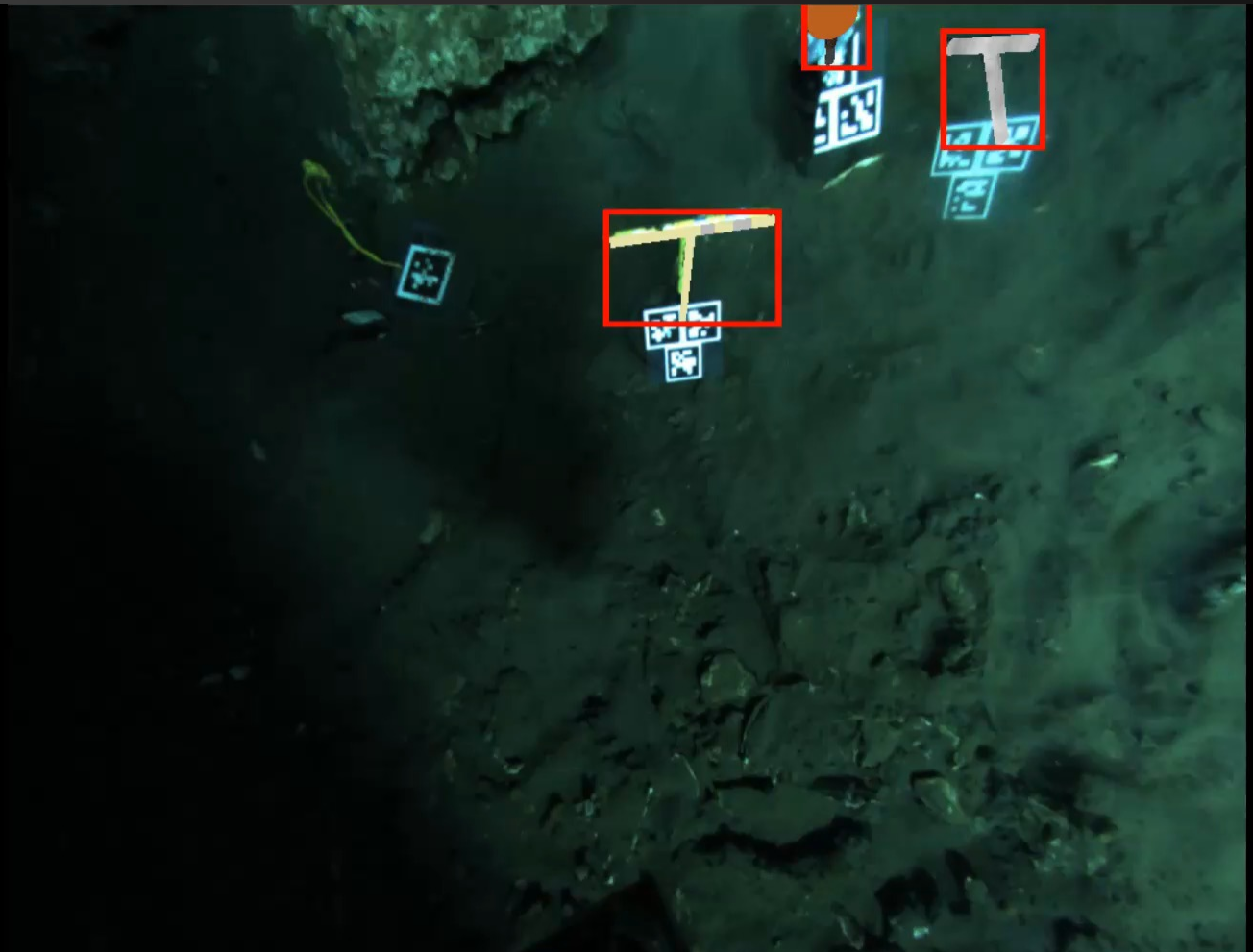
Handles used to manipulate underwater tools



ROS Based AprilTag Detection and SLAM



VisPose – Tool for Annotating Pose in Monocular Image Sequences

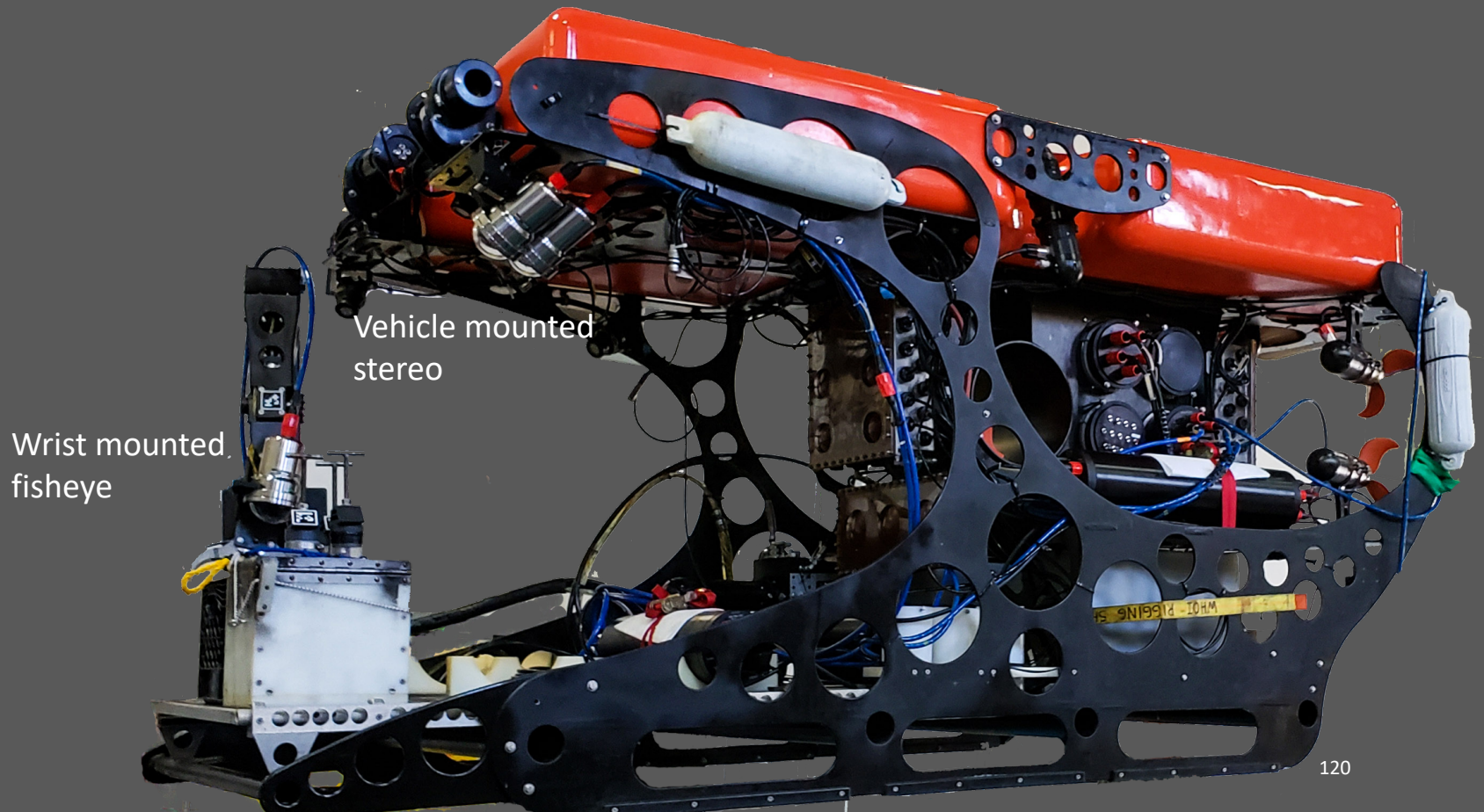


Automating ROV Manipulation

Gideon Billings

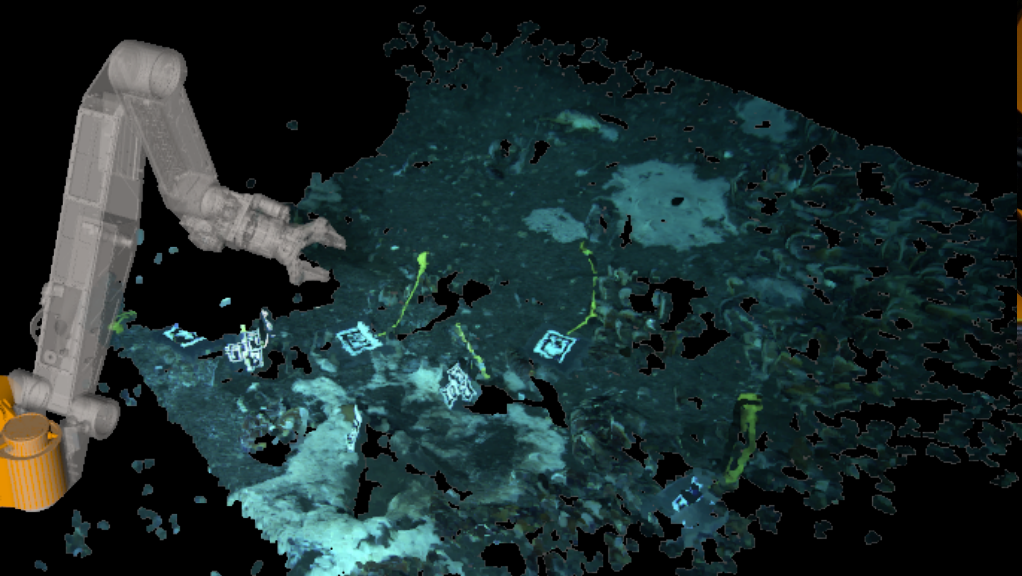
ATOM - Autonomous Testbed for Obstructed Manipulation

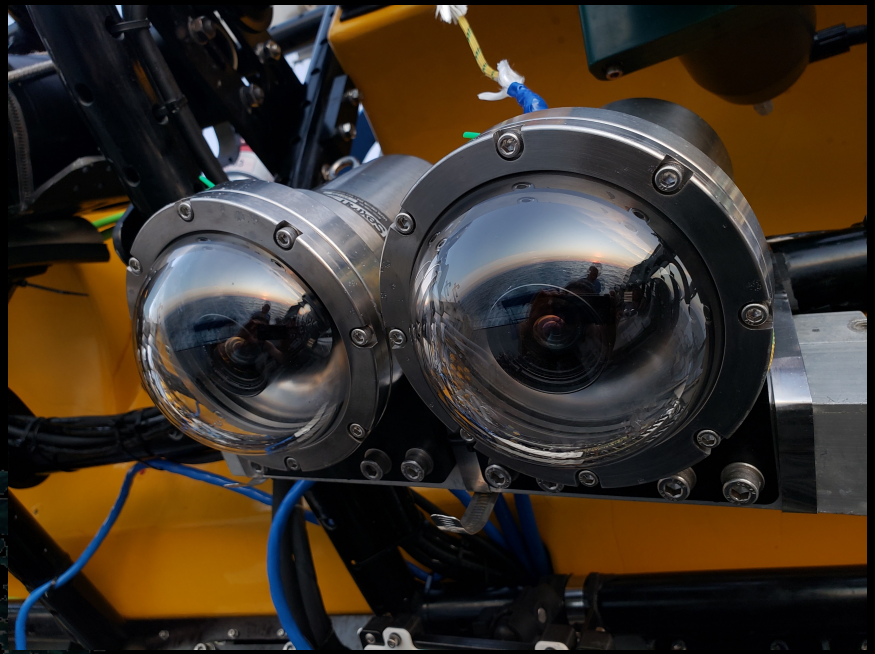
SMIRC - Stereo and Manipulator Imaging and Reconstruction Cameras



Stereo and Manipulator Imaging and Reconstruction Cameras

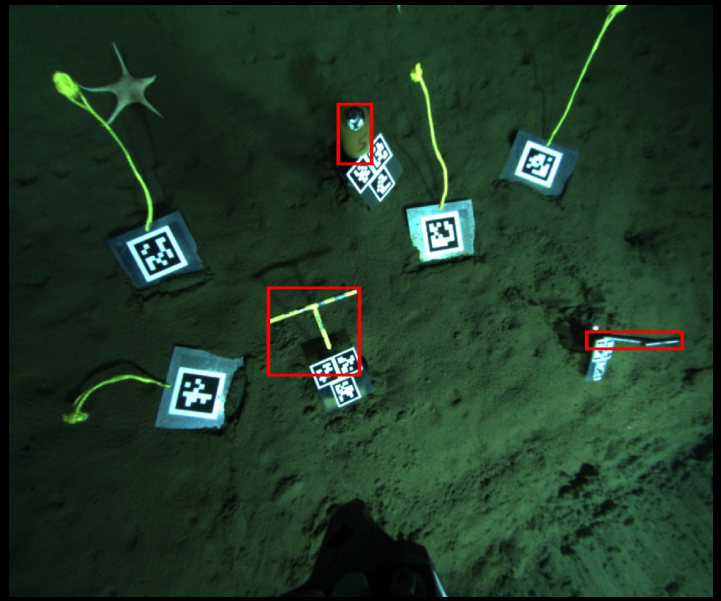
Stereo for 3D reconstruction and visual kinematic calibration



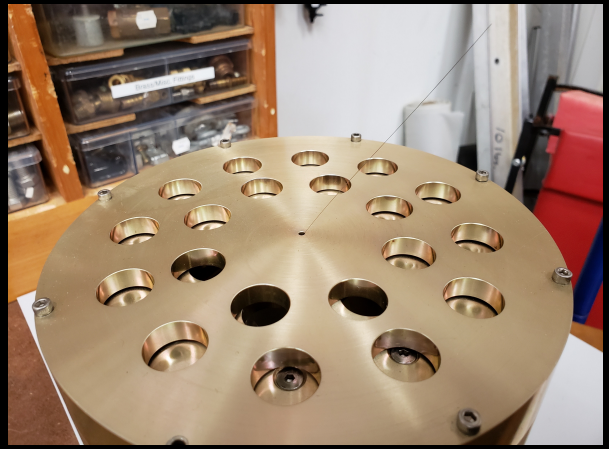


Stereo and Manipulator Imaging and Reconstruction Cameras

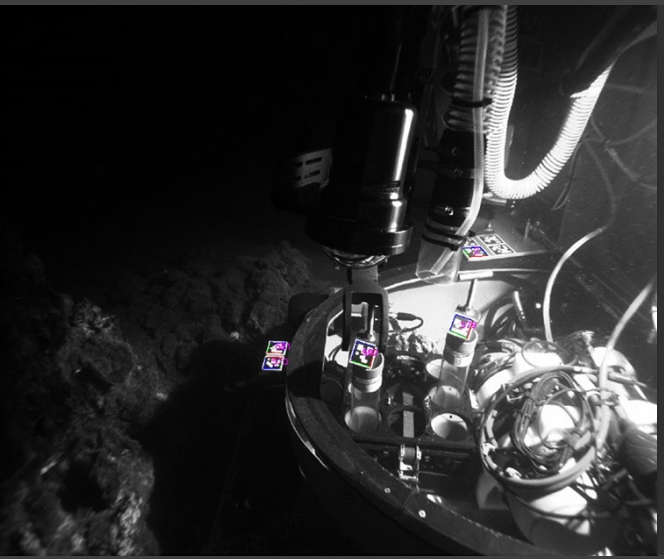
Fisheye for object detection and visual servoing

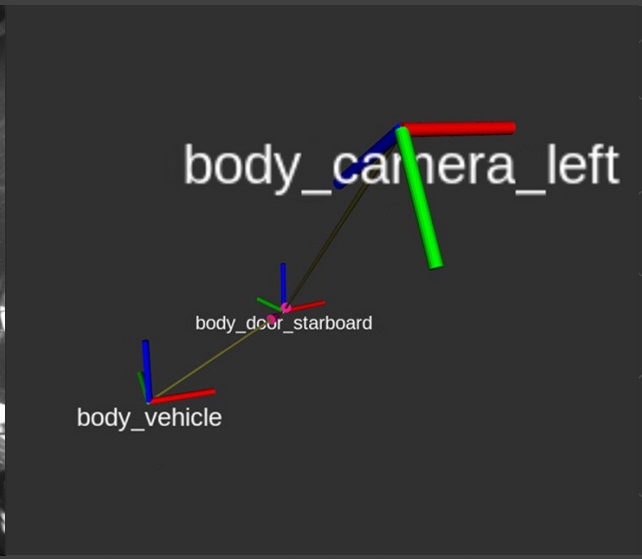


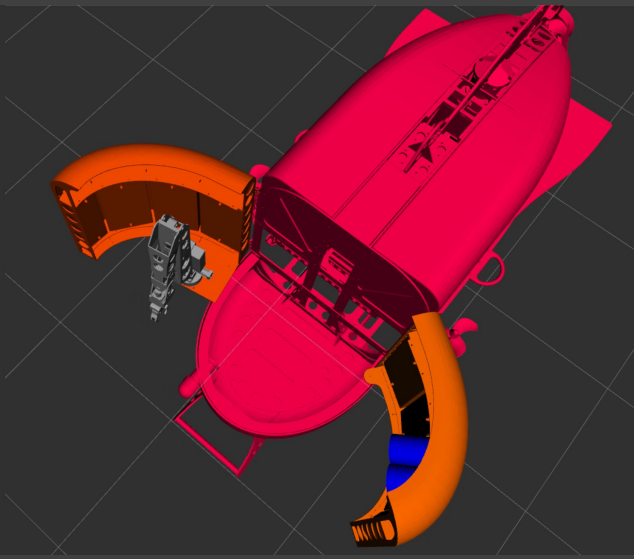
Nereid Under Ice Hybrid Tether ROV



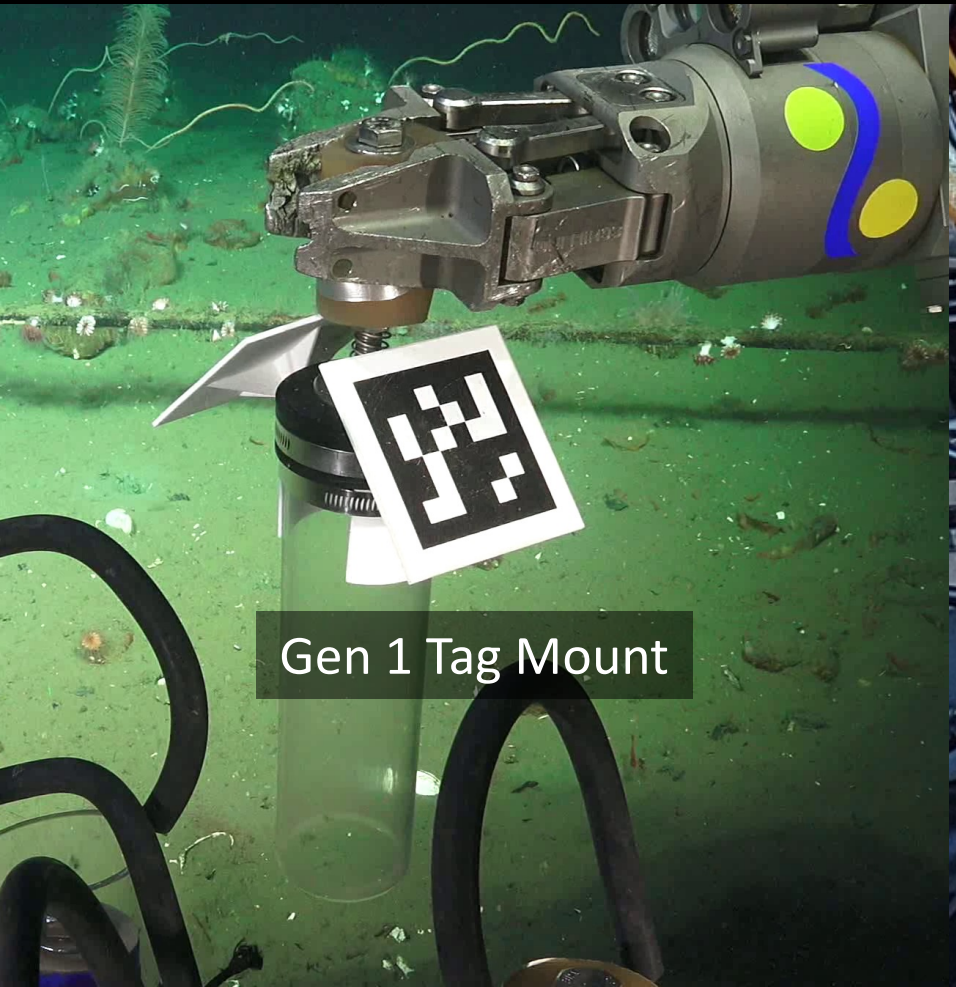
Modeling NUI Doors



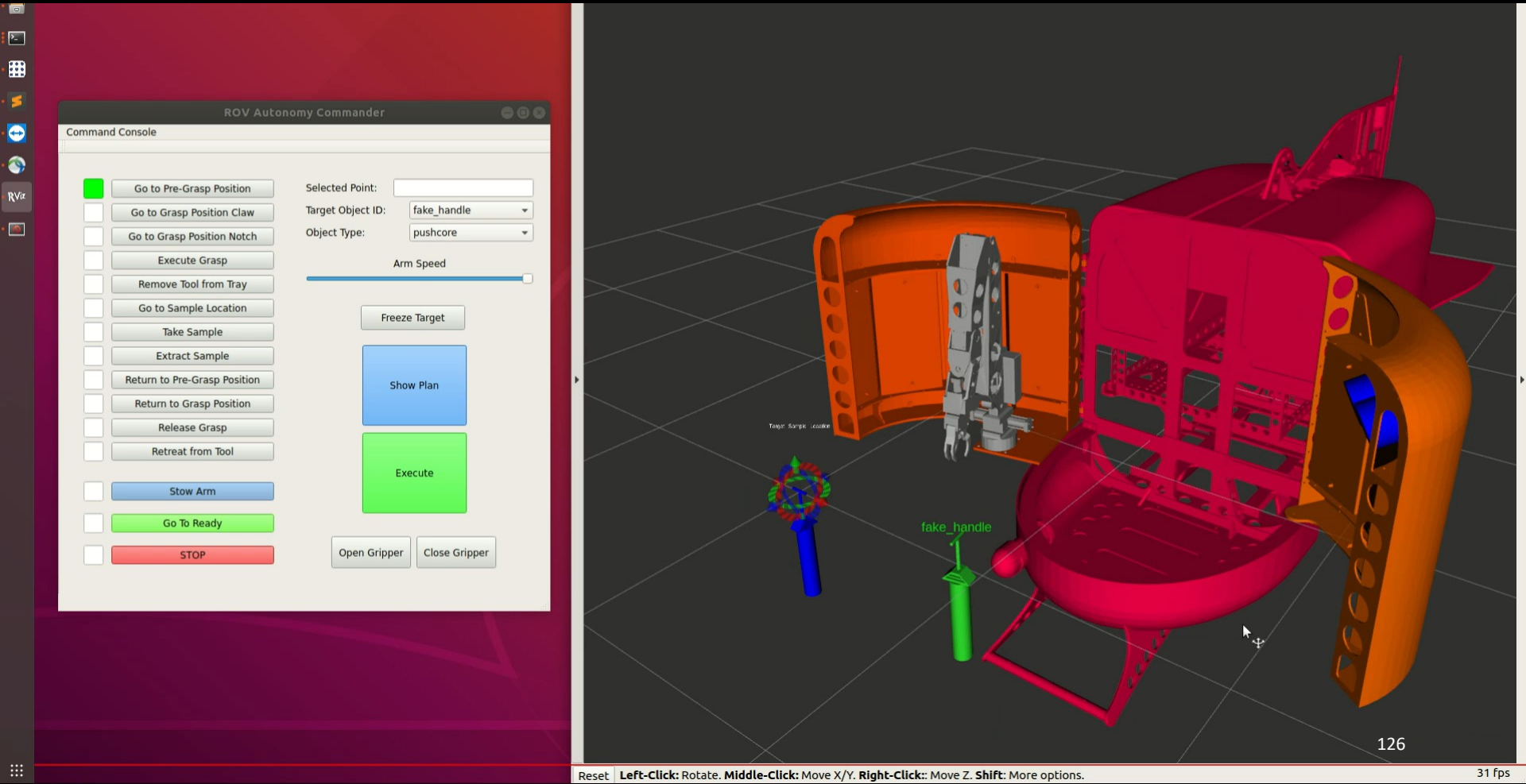




AprilTags for ground truth handle pose

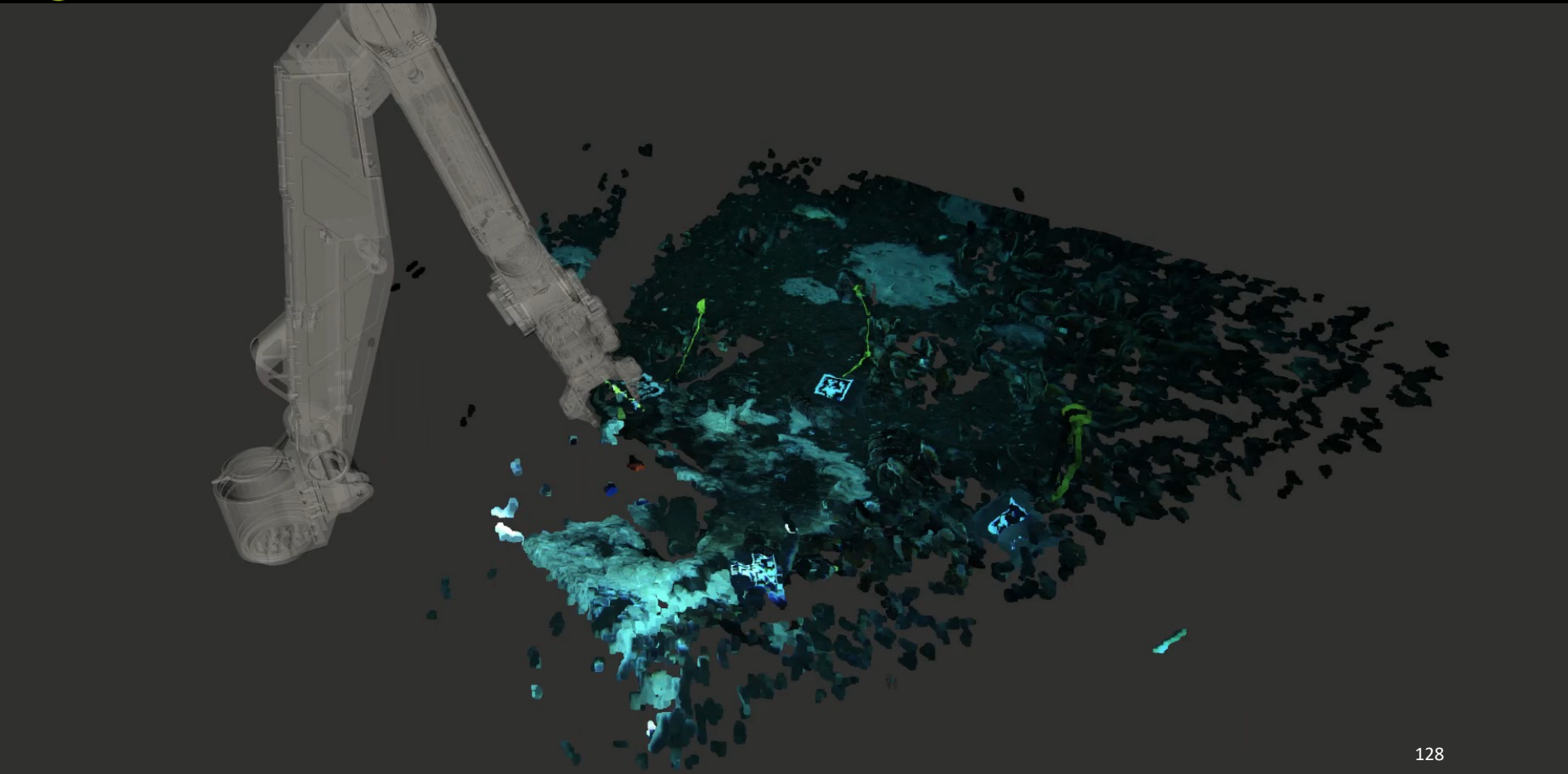


Pick and Place Pipeline

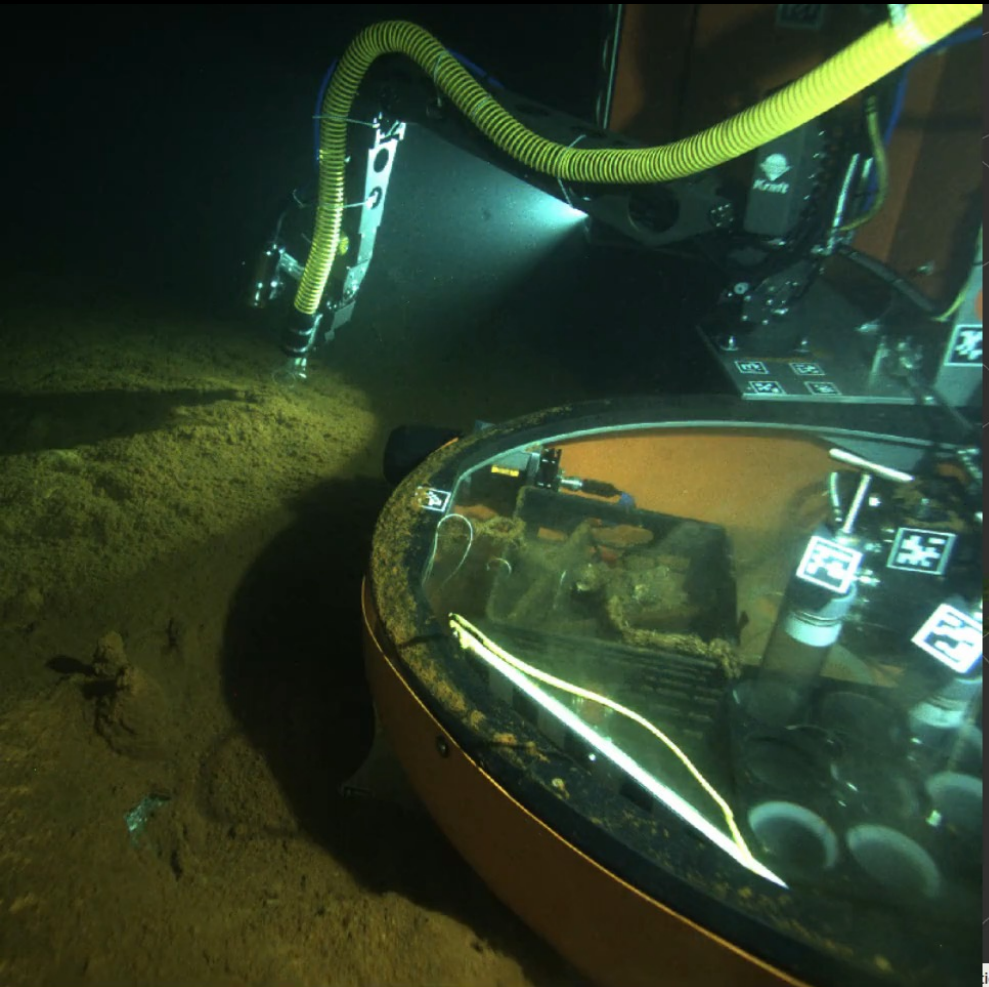


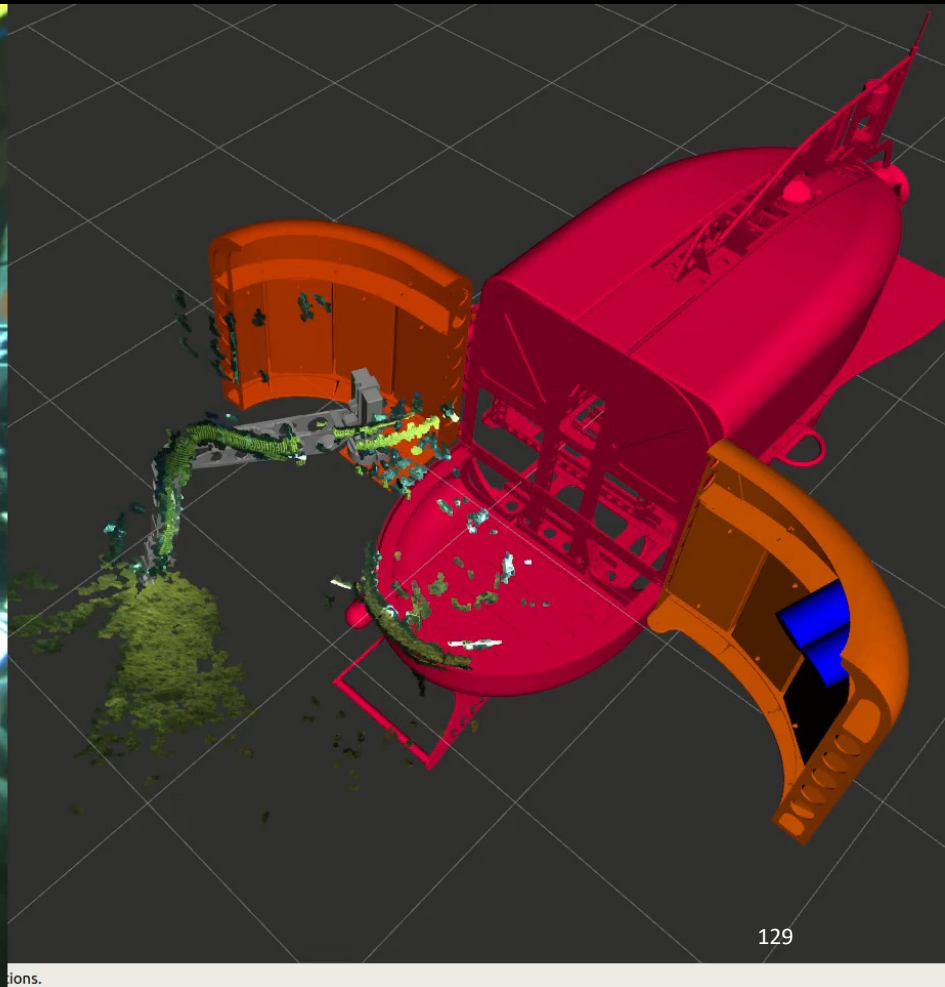
Field Demonstrations

SOI Sebastian ROV Realtime Kinematics & Stereo Point Cloud

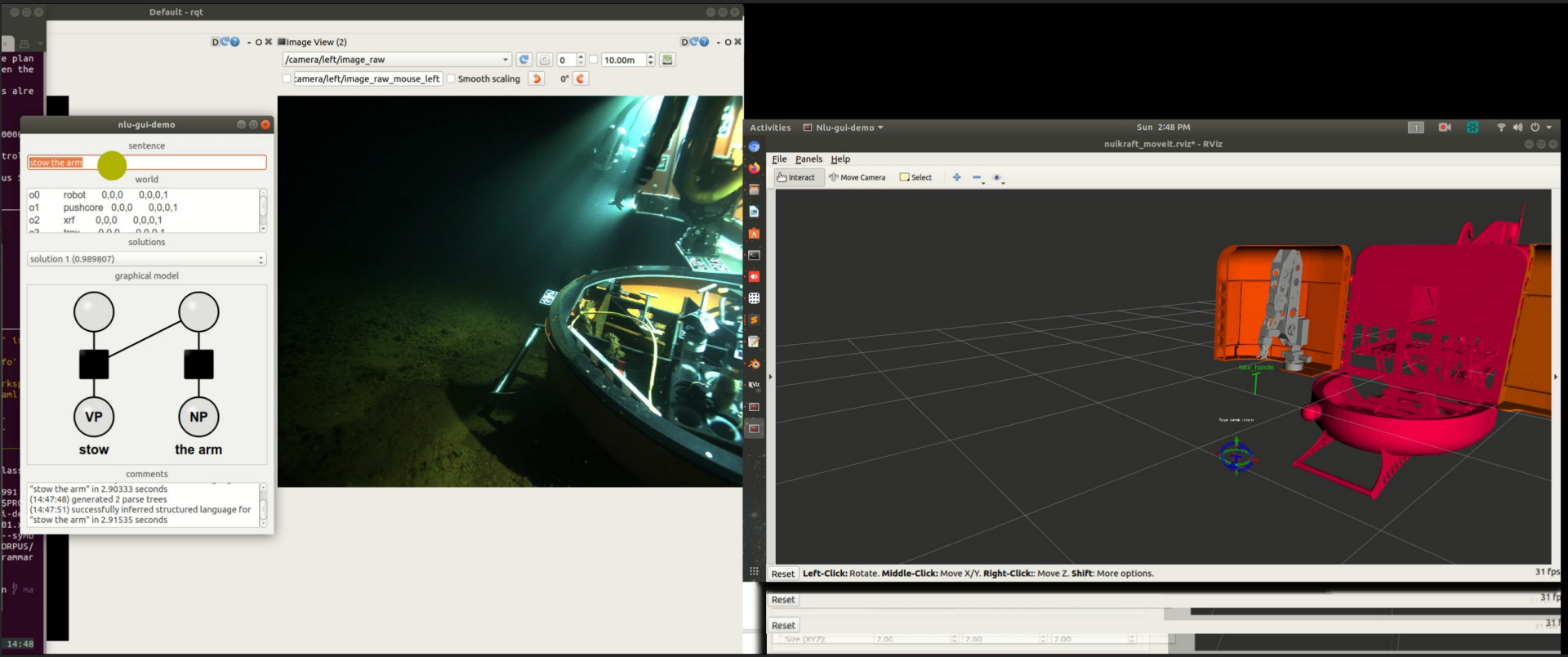


Planner Controlled Slurp Gun Sample





Natural Language Control



Thank you
for your attention