Pool2Ocean: Synthetic Data Generation for Underwater Object Detection Using CycleGAN

Christopher Morse

Faculty Mentor: Junaed Sattar, Ph.D.

University of Minnesota – Twin Cities

Introduction

- In recent years, Autonomous Underwater Vehicles (AUVs) have been in the forefront of oceanic research and exploration.
- Some AUVs depend on visual input to move, interact with the environment, and communicate.
- Necessary to ensure that these AUVs are properly trained on underwater imagery so they can make accurate generalizations about their visual environment.
- Collecting underwater imagery is often time consuming, expensive, and hazardous for both the humans and robots involved.

Motivation and Objective

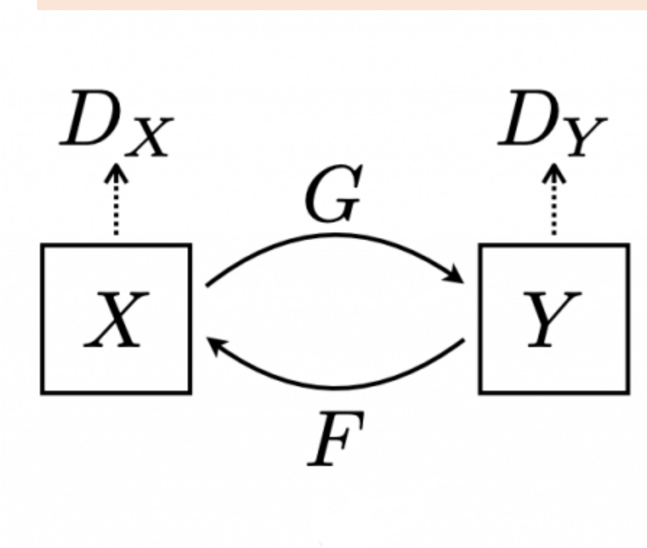
- Desire to explore a more accessible method of data augmentation, using deep learning for synthetic data generation, to improve underwater robot-to-robot detection.
- Conversion of images of a robot operating in a pool to images that mimic real ocean images.
- These resulting 'fake' ocean images can be used to expand an existing dataset.

Hypothesis

- Augmenting a dataset of real ocean images with synthetic ocean images will lead to better performance from RetinaNet for robot detections, as determined by the resulting precision, recall, and average precision (AP) scores during evaluation.

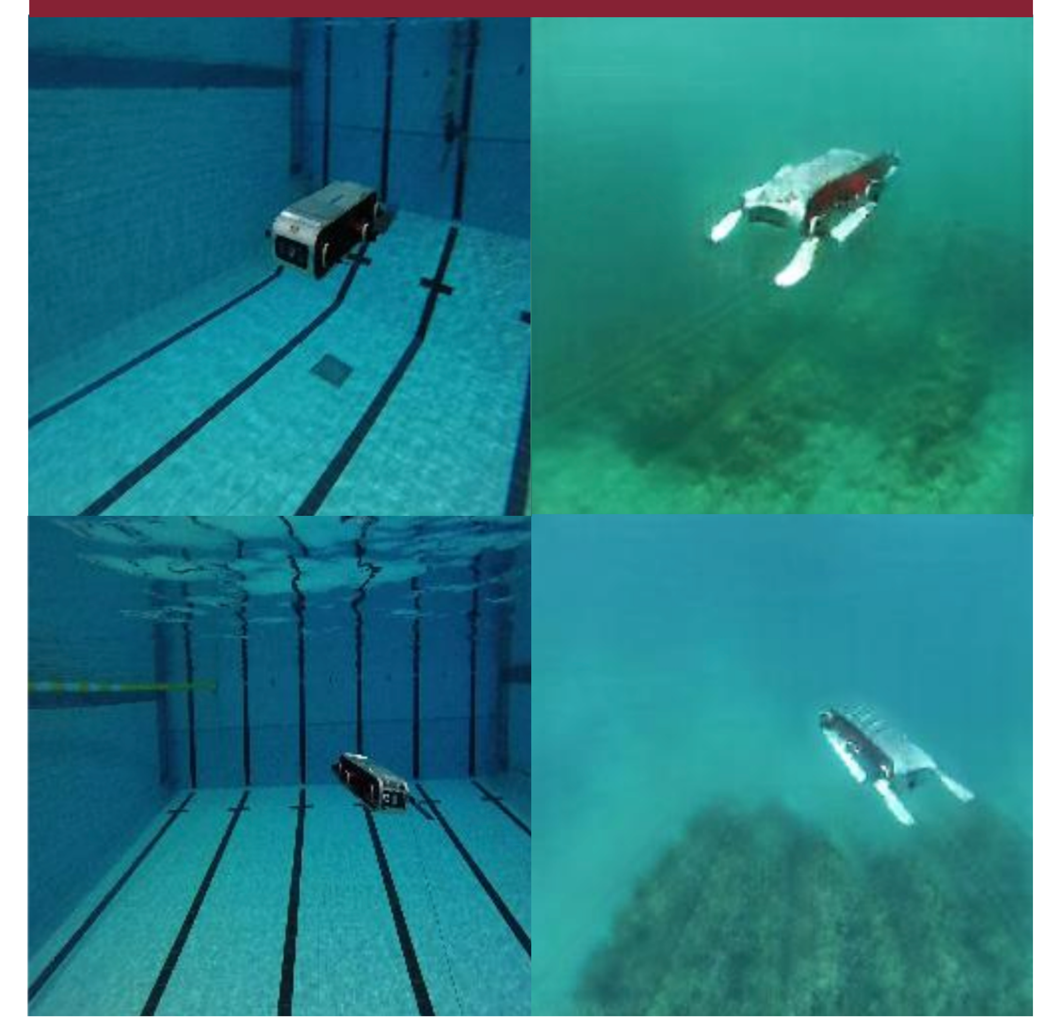
CycleGAN Data Generation

- Datasets were compiled from videos of an Aqua robot ("Minnebot") in pools and oceans collected from the Interactive Robotics and Vision (IRV) Laboratory.
- CycleGAN was used to translate images from the source domain (pool) to the target domain (ocean).



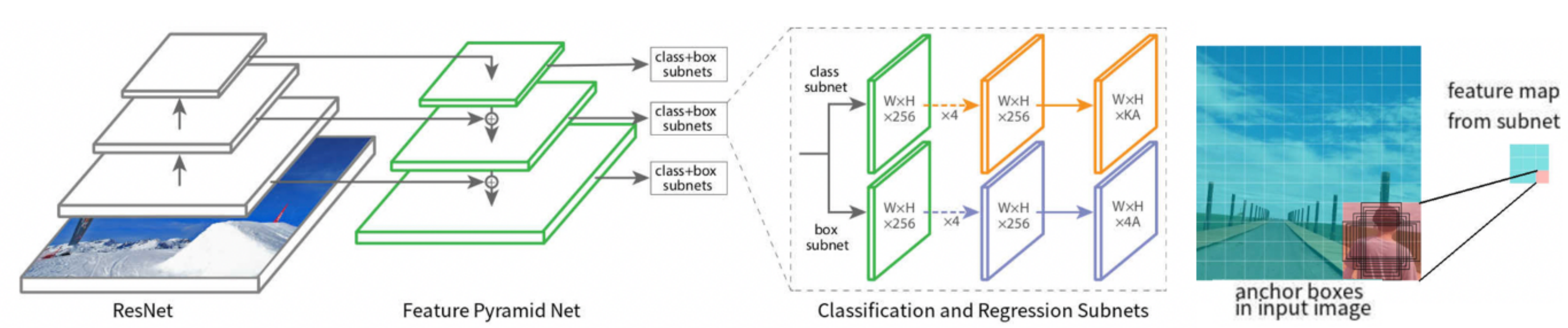
- Generator and discriminator networks operate on the source (X) and target (Y) domains.
- Translations from mappings G and F are used as input for discriminators D_Y and D_X, respectively.

Original Image **Generated Image**



RetinaNet Object Detection

- Three RetinaNet detection models were trained for 75 epochs: the first on real ocean images, the second on fake ocean images, and the third on all images.



Results

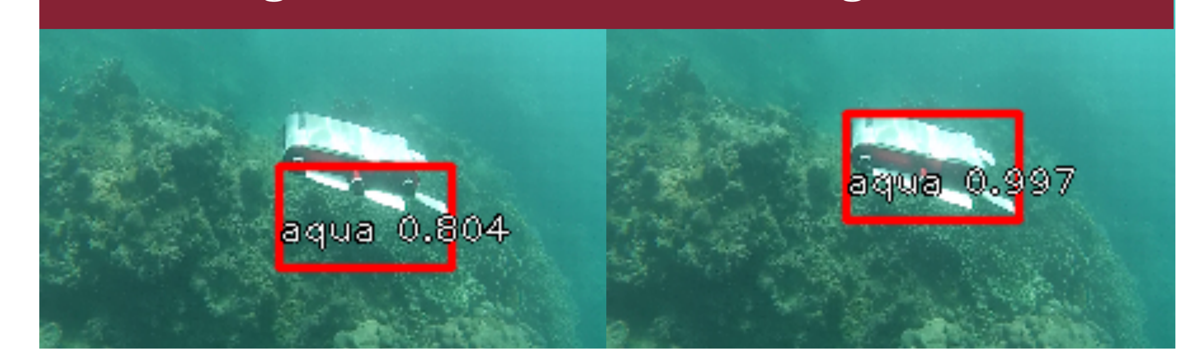
- An independent testing set of 102 real ocean images was used for evaluation.
- Each model was evaluated on the testing set to produce precision and recall scores, as well as the average precision (AP) across various detection thresholds:

	Real Ocean Images	Fake Ocean Images	All Images
Precision	0.6211	0.4897	0.9059
Recall	0.4836	0.2956	0.6481
AP	0.5273	0.2774	0.9573

Discussion

- These results suggest that synthetic data generation, for the purpose of data augmentation, can be an effective method to improve a RetinaNet's model's ability to detect underwater robots.
- Performance when a model is exclusively trained on fake images is poor, likely due to translation imperfections and subtle artifacts (e.g. pool lines).
- Drastic improvement in performance when the context of the real ocean domain is included in the combined dataset.

Pre-Augmentation **Post-Augmentation**



Contributions

- Safe, fast, non-invasive, inexpensive, and effective method for dataset building.
- Demonstrates the potential for unpaired image-to-image translation to augment existing datasets with synthetic data to improve robot-to-robot detections.

Acknowledgements

This research was supported by the University of Minnesota Undergraduate Research Opportunities Program (UROP).

Special thanks to the Interactive Robotics and Vision (IRV) Laboratory, found at: http://irvlab.cs.umn.edu/.