Microsoft

# AI for Earth Grantee Profile

## Duke Marine Robotics and Remote Sensing Lab
## Monitoring whales with drones and machine learning


## Summary

Climate change is disrupting the pristine ecosystem around the western Antarctic peninsula, a globally significant center of biodiversity based around the presence of krill that provide sustenance for many other species. Many of the world's whales spend their summers here as their primary feeding grounds. By monitoring the size and health of the whales, it's possible to gain insights on the abundance of krill and the ecosystem as a whole. Satellites and drones now enable vast amounts of image and video data on the whales to be collected—more than could ever be efficiently processed by people. The Duke University Mobile Robotics and Remote Sensing Lab is developing machine learning models on Microsoft Azure that can manage this massive data and quickly provide the statistics needed to help further research and environmental protection efforts. Making these models available as APIs on Azure will also enable other researchers to improve their work.


## Monitoring climate change in the Antarctic with machine learning

Life teems on Earth even in some of the seemingly most inhospitable climes. The western Antarctic peninsula is one such hotspot, the biodiversity epicenter of Antarctica. There, many different creatures from crabs to penguins up to many of the world's whales go to sustain themselves, subsisting on the millions of krill living in the icy waters.

> As the ice declines, so will krill populations—and all the other animals that depend upon that.

However, even this pristine environment, untouched by agricultural or industrial pollution, is still threatened by climate change. The ice is melting, and it appears to be happening faster here in this vibrant ecosystem. Overall rising sea temperatures and changes in oceanographic currents are contributing to this decline, and as the ice declines, so will the krill population—and all the other animal populations that depend upon that. As one of the first major ecosystems to be disrupted by climate change, and a pristine one at that, it's important to understand the changes happening here, to better understand global climate change.

It’s difficult to study the krill directly, as they are masses of millions of tiny creatures. But it’s easier to monitor the great whales that spend the Antarctic summer feeding to gain the body mass they’ll need to survive the rest of the year. Since the whales rely on an abundant population of krill, healthy whales indicate a healthy ecosystem. The whales’ health can be monitored in part by measuring their sizes to compare how they’ve grown in a season versus previous years, and that measurement can be done with photographs. With modern high-resolution imaging and video available from satellites as well as drone-based cameras, we have good tools to gather lots of information about the whales. Processing all that information is another matter—it’s now possible to collect more data in the three months of the Antarctic summer than a team of human researchers can review and analyze in the remaining nine months of the year.

### Measuring whales with machine learning

Patrick Gray, a marine scientist with Duke University’s Marine Robotics and Remote Sensing Lab (MaRRS, led by Dr. David Johnston), is a member of one such research team monitoring whales in Antarctica. In collaboration with fellow Duke PhD student KC Bierlich, Gray is enlisting the aid of artificial intelligence through a grant from Microsoft to solve this data analysis problem. For a human, the process of marking up whale images to calculate their sizes could take five to six minutes per whale, or four to five work days to do a couple thousand images. But the MaRRS lab is training machine learning models in Microsoft Azure to identify whale species in the imagery and calculate the sizes from the length and girth visible. A trained neural network can do this work in a couple seconds per image—or a couple thousand images in an hour. That speed will enable them to process and analyze the massive amounts of data they bring back from Antarctica in a timely fashion, gaining useful insights to help direct further research and environmental preservation.

> “Our team is using machine learning to monitor the size and health of whales to gain insights on the abundance of krill and the ecosystem as a whole.”
> —Patrick Gray

For this project, the team is using Azure Data Science Virtual Machines with ArcGIS Pro 2, Python, and enterprise R for processing RGB, multispectral, and hyperspectral imagery in the cloud. Azure machine learning resources train deep convolutional neural networks and recurrent neural networks on multiple-GPU compute clusters (with training datasets up to 10 TB and inference datasets of 100+ TB). Training and inference jobs may use up to 100 instances concurrently and take up to two days to complete. The team also employs traditional image classification workflows available in ArcGIS Pro (both pixel and object-based), in comparison to new, high-efficiency neural network workflows.

### Surveying whales with satellite imagery

In related work, the team can apply these machine learning models to very high-resolution imagery from DigitalGlobe's WorldView satellites to find, identify, and measure whales in different areas of the world. WorldView imagery is produced for a variety of different clients who are typically interested in shipping lanes and areas within ten miles of coastlines, not the middle of oceans, so the inclusion of whales is incidental. However, it's also produced regularly, frequently, and in massive amounts—a human analyst would need years to scan through what's produced in a month, but a neural network can do it automatically, providing a valuable trove of data. And in addition to helping monitor the whales' health during the rest of the year, it also can help define their habitats to better create marine protected areas. Potentially it could even help with near-real time dynamic traffic management to reduce conflicts between shipping and whales and help protect the endangered species.

In an article published through the *British Ecological Society* in June 2019, Gray and his collaborators documented the results of their efforts. The automated photogrammetry method compared very favorably with conventional techniques, with 90 percent of the automated length measurements being within 5 percent of the manual measurements—accurate enough to establish size classes of whales automatically. Further, the model was 98 percent accurate at predicting humpback, minke, or blue whale species in the sample set. These results show the potential for these automated methods to help scientists quickly proceed from the raw data analysis to the vital questions of their research.

### Moving forward

Gray believes the process of automating the species ID, measurements, and behavior analysis will be transformative for the whole field of ecology. Managing and processing all the data available from new input streams such as camera traps, drones, and satellites is a big task for most teams. Currently, hundreds of whale researchers across the globe are using different manual analysis processes, and thus getting different results from the data they can process. Gray has turned the team's machine learning model into an API (available on GitHub) that runs in the Microsoft Cloud in Azure Container instances, where it is available to all these different research groups. Researchers can upload their images and metadata, and get back species IDs and photogrammetric measurements. Researchers could also provide collections of images for other species and have the API trained on the new datasets, extending its capabilities for many more researchers.

# About Patrick Gray

Patrick is a PhD student, marine scientist, and computer scientist with Dr. David Johnston at Duke's Marine Robotics and Remote Sensing Lab. He is interested in combining artificial intelligence and remote sensing for ocean science and conservation, as well as for exploring and understanding other ocean worlds within our solar system. Before his graduate work at Duke, Patrick worked with Dr. Pete Girguis at Harvard University as a research technician, at WayPaver Foundation as Chief Technology Officer, and at Moon Express as a software engineer.

# About the Duke University Robotics and Remote Sensing Lab

The Duke Marine Robotics and Remote Sensing Lab was established in 2016 with the aim to advance robotic applications, platforms, and sensors in the service of marine science and conservation. Supported by an infrastructure grant from the National Science Foundation, the lab has established state-of-the-art engineering and teaching spaces, and is active in five areas of science and practice, including conducting research, supporting external research programs, sensor and platform testing, and executing both formal and informal educational and outreach programs. The lab operates and maintains 18 different unmanned aircraft system (drone) platforms (both fixed wing and multirotor) which can carry a variety of RGB, multispectral and thermal payloads for a broad suite of remote sensing and data collection tasks. At present, the lab has projects focused across disciplines in marine sciences, ranging from coastal geomorphology to animal behavior, and is funded through grants from federal and state government agencies as well as private sources.

# Resources

### Websites

Duke Marine Robotics and Remote Sensing Lab (MaRRS)
Cetacean Photogrammetry API on GitHub
Measuring Cetaceans from Remotely Sensed Data on GitHub—documents the code needed to train and deploy this model, and makes the data open access

### Publications

Gray, P., et al. "Drones and convolutional neural networks facilitate automated and accurate cetacean species identification and photogrammetry." *British Ecological Society* (2019). https://doi.org/10.1111/2041-210X.13246.