

# Identifying “Something Interesting” in a Marine Underwater Image

Author: Jennifer Crosby

Final Project Report

**Dataset:** Video from NOAA Okeanos underwater autonomous ROV expeditions to the Marianas Trench

One of the main oceanographic projects today is to discover and protect marine life and vegetation, done while mapping deep water locations through video through unmanned expeditions to areas that are otherwise impossible to see. During and shortly after each expedition, hundreds of expert scientists are available to view this video and analyze and classify what is being seen. However, they allot little time outside of the expedition window to watch the videos. At the current time, these experts must watch the entire videos in order to determine if anything captured is of use to their research. These videos are often more than four hours long, and require massive time expenditures from these experts, for sometimes 10 minutes of value. Our problem consists of helping in this project by building and training a neural network to shorten the amount of time the experts need to spend watching the video by reducing the video watched to only potentially interesting images. Ultimately, the goal is to classify what is seen and parse that video to smaller videos that would be sent to the respective scientists who have interest.

NOAA’s (National Oceanic and Atmospheric Administration) expeditions cost millions of dollars each, which makes it important that the information is utilized. Currently, however, they have too much data for people to effectively analyze and much has not been even viewed. This results in much wasted money and time for both NOAA and the experts. Our contribution will be the first step in enabling faster and more useful data analysis in the future

There was no shared work due to my partner dropping the class, so this whole project ended up being an individual project.

**Results.** Using VGG16 Convolution Neural Network I was able to obtain initially 78% accuracy on 400 images and running 3 epochs on 1500 images the accuracy was 60%. Running 30 epochs will have an even better accuracy.

## Summary and Conclusions

VGG16 produced better results than my initial model based off of my Exam 2 code. Adding additional layers would improve the model.

I don't know how to calculate this, as all ideas and research came from the internet and class code.

## References

Ullah, J. Ahmad, K. Muhammad, M. Sajjad and S. W. Baik, "Action Recognition in Video Sequences using Deep Bi-Directional LSTM With CNN Features," in IEEE Access, vol. 6, pp. 1155-1166, 2018.

<https://towardsdatascience.com/detecting-animals-in-the-backyard-practical-application-of-deep-learning-c030d3263ba8>