Detection and Classification of Fishes

Saha, Rudra rsaha3@asu.edu Khandelwal, Prakhar pkhande1@asu.edu

I. Introduction

The ability to extract information about a very specific object type from an image which occupies only a small patch of it can be utilised to embolden the efficacy of relevant downstream tasks which rely on it. An effective model to gauge the image and identify such relevant portions of it can be built using Attention neural network architecture. Our goal for the project is to show the effectiveness of such a paradigm. Working with the data provided by The Nature Conservancy hosted on kaggle, the task is to develop algorithms to automatically detect and classify species of tunas, sharks and more that fishing boats catch in order to accelerate their video review process.

The video frames present in the dataset shows fishes present in an environment riddled with noise where running a simple object recognition module on top of it will result in sub-par accuracy as shown by the benchmark results. Instead of delving into techniques to reduce noise, our approach is to tend to the patch with the object of attention in it and classify the object by only accounting for that patch, where objects here are the various kinds of fishes.

II. Project Timeline

Timeline		
Task	Assigned to	Tentative Deadline
Literature Survey	Prakhar, Rudra	24th Feb
Creation of Data Set	Prakhar, Rudra	4th March
Attention Neural Network	Prakhar, Rudra	12th March
CNN Training	Prakhar, Rudra	19th March
Testing and Validation	Prakhar, Rudra	5th April

References

[1] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhutdinov, Richard Zemel and Yoshua Bengio, Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. (ICML 2015)

[2] Sergey Zagoruyko and Nikos Komodakis, Paying more attention to attention: Improving the performance of CNN via Attention Transfer. (ICLR 2017)

[3] Volodymyr Mnih, Nicolas Heess, Alex Graves and Koray Kavukcuoglu, Recurrent Models of Visual Attention. (NIPS 2014)

[4] J Long, E Shelhamer and T Darrell, Fully convolutional networks for semantic segmentation. (IEEE 2015)

[5] Satrajit Acharya and R Vimala Devi, Image retrieval based on visual attention model. (Elsevier 2012)

[6] I Sutskever, O Vinyals and QV Lem Sequence to Sequence learning with neural networks. (NIPS 2014)

[7] Karol Gregor, Ivo Danihelka, Alex Graves, Danilo Jimenez Rezende and Daan Wierstra, DRAW: A Recurrent Neural Network For Image Generation. (JMLR 2015)

[8] C Liu, J Mao, F Sha and A Yuille, Attention Correctness in Neural Image Captioning. (AAAI 2017)