An Underwater Behavior Recognition System for Marine Life

Favour Nerrise[1], H. Seckin Demir[2], Sule Ozev[2], Jennifer Blain Christen[2]
[1]University of Maryland, College Park, [2]SenSIP Center, School of ECEE, Arizona State University
Tempe, AZ 85287, USA
sensip@asu.edu

Abstract— **Bycatch, unintended capture of marine species, is a prominent issue that affects sea animals like sea turtles and damages the habitat. Several cyber physical systems have been implemented to reduce bycatch, with varied success rates. Smart Nets is an object recognition system that detects sea turtles and uses LED illumination levels as stimuli to warn turtles and big sea animals of potential danger. We aim to build on Smart Nets by developing a behavior recognition system that identifies sea turtle response behavior to stimuli, specifically U-turns and reversals. This would enable automated optimized stimuli to achieve effective results in reducing the bycatch incidents.**

***Keywords*—sea turtles, behavior recognition, U-turns, reversals, Smart Nets, object detection, activity recognition**

I. Introduction

Behavior recognition of marine life is a largely under-investigated area in marine research. Generally, objection detection and tracking of sea animals requires high energy consumption, expensive equipment for sensing, and long-term capture. This becomes very difficult to maintain over time, and there is a pertinent to reduce the energy needed and consumed for autonomous marine life detection. To date, there have been little to no developments to building a behavior recognition system for sea animals.

Smart Nets is a machine learning enabled cyber physical system, to explore the efficacy of deterrent devices that trigger species-specific sensory cues to optimize bycatch reduction technology while avoiding the enormous effort and time required for traditional large-scale studies [1]. The system addresses a major challenge of underwater imaging systems, which often demand severe consumption constraints due to the limited resources. Accordingly, the system implements a recognition model that adaptively adjusts system parameters like the camera frame rate and LED illumination level, based on the environmental factors to optimize the energy consumption while ensuring a high recognition accuracy. However, there remains a great need for automating the animal behavior analysis for system performance evaluation and stimuli selection process.

In this paper, we propose a behavior recognition system for marine life, specifically u-turns and reversals. Reversal usually occurs when the sea animal gets untangled from the net; this involves no change in orientation. U-turn behaviors often occur when the animal sees a danger and changes its orientation 180 degrees in a short time span. The testing dataset contains 40 video files with 270 clips of manually identified sea turtle behaviors by marine scientists. The average duration of U-turn behavior in U-turn clips (*n=141*) was 5.5 seconds, and the average reversal behavior clips (*n=129*) was approximately 7.8 seconds.

This project is sponsored by NSF Award 1659871.

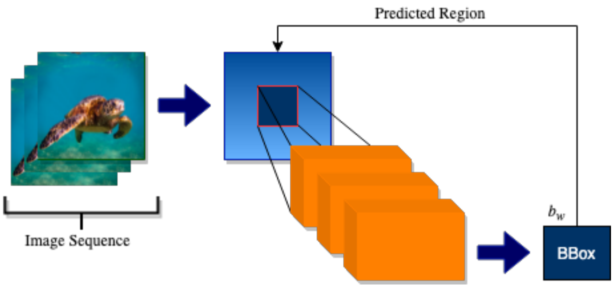

Fig. 1. Block diagram of behavior recognition system

The input features for the model will be the estimated orientation and depth of the sea animal, identified through monocular 3D object detection. Previously, depth estimations have been generated using a monocular image single-frame depth network like SfMLLearner [2]. Our approach uses YOLO v4 tiny [3], a convolutional neural network, to predict and retrieve bounding box coordinates, applied to an extended Kalman filter [4]. The filter computes the distance between the sensor and the sea turtle's real-time location to estimate depth in a 3D-space, which is measured for accuracy against ground truth labels of manually estimated sea turtle depth. Currently, the object detection module performs highly with a mAP@0.50 of 0.8567.

Once the features are extracted, we will feed this into a rule-based classification model to identify two behaviors: U-turn and reversal. This rule-based classification approach could be further improved to adapt to a neural network-based model [5].

References

[1] H. S. Demir, J. B. Christen, and S. Ozev, "Energy-Efficient Image Recognition System for Marine Life," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 39, no. 11, pp. 3458–3466, Nov. 2020, doi: 10.1109/TCAD.2020.3012745.

[2] T. Zhou, M. Brown, N. Snavely, and D. Lowe, "Unsupervised Learning of Depth and Ego-Motion from Video," 2017. [Online]. Available: https://arxiv.org/abs/1704.07813

[3] Z. Jiang, L. Zhao, S. Li, and Y. Jia, "Real-time object detection method based on improved YOLOv4-tiny," *arXiv:2011.04244 [cs]*, Dec. 2020, Accessed: Jun. 25, 2021. [Online]. Available: http://arxiv.org/abs/2011.04244

[4] A. Mauri *et al.*, "Deep Learning for Real-Time 3D Multi-Object Detection, Localisation, and Tracking: Application to Smart Mobility," *Sensors (Basel)*, vol. 20, no. 2, Jan. 2020, doi: 10.3390/s20020532.

[5] A. Wedel, U. Franke, J. Klappstein, T. Brox, and D. Cremers, "Realtime Depth Estimation and Obstacle Detection from Monocular Video," in *Pattern Recognition*, vol. 4174, K. Franke, K.-R. Müller, B. Nickolay, and R. Schäfer, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 475–484. doi: 10.1007/11861898_48.