Project Title:

Fish monitoring system with stereo camera on edge-computing device

Short abstract:

This thesis proposes the study and development of a fish monitoring system on an edge-computing device, using a stereo camera integrated into a fishing net. Objectives include fish count, species classification, and size estimation. Deep learning models will be trained and adapted, leveraging the disparity information from the stereo images.

Description:

This master's thesis proposal focuses on the development of a fish monitoring system using a stereo camera system integrated with a fishing net. The main objectives of this research project include fish instance counting, species classification, and estimation of average size. This monitoring system intends to be embedded in an existing system that is in development. The proposed methodology centers around employing deep learning models from the YOLO family for fish detection and classification and deployed on an edge-computing device. Additionally, the stereo images captured by the camera system will be utilized to enhance the performance of these tasks. The proposal also explores alternative approaches utilizing stereo images and investigates potential methods for fish measurement.

The study will commence with implementing and fine-tuning the YOLOv7 model for fish detection and classification in stereo images. Since no publicly available dataset closely resembles the data obtained from the system (shown in Figure 1), the proposal also includes exploring effective methods to annotate or exploit the great number of available unlabelled data.

The main focus of the project will be leveraging the information provided by stereo images. The study will explore different approaches to utilize disparity information for validating and refining detected fish instances. *Can the detections in both stereo images validate each other or be used as a hard example mining process*? Are some of the first ideas that could arise, and will be focused on throughout the project. In addition, the depth cues present in stereo images can be used to estimate the size and distance of the fish.

Considering that the fish monitoring system will run on a Jetson Xavier NX device, the speed performance of the model will also be taken into account. Although having available video sequences in the dataset, the deployment of the algorithm on an edge-computing device leads to the discarding of the methods that work on video sequences. The project will focus on working on single stereo frames, making the monitoring system more robust for different scenarios. This could include a scenario in which the framerate is lowered to save the limited battery and be able to keep track of much larger trawls.

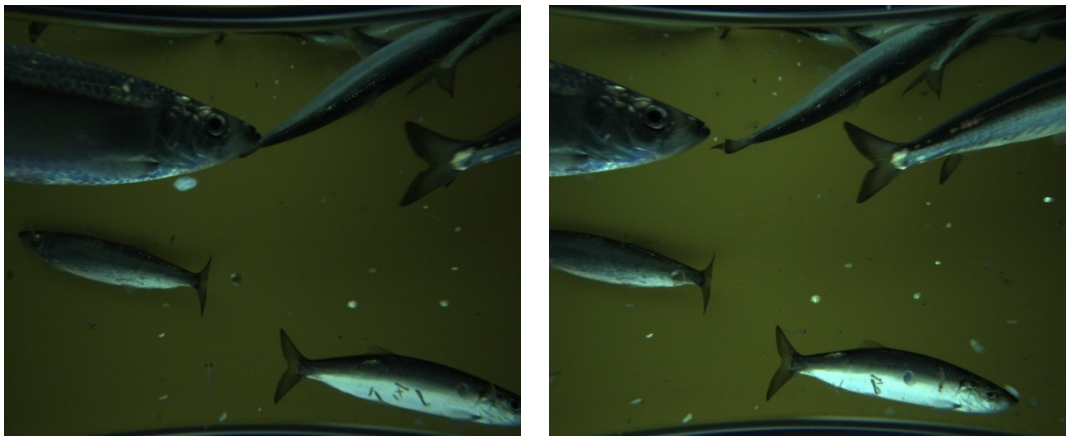

Figure 1. Stereo frame (Left and Right) from acquisition system.