Automated Fish Counting Using Sonar

Issue 1: March 2025

mahobley@caltech.edu

In This Issue…

Introduction

Welcome to the first of what we intend to be a quarterly newsletter. These newsletters will summarize what we are currently working on, our results and our goals. It will be distributed via email, they will also be available on our website.

The Project Goal

The goal of this project is to automatically count returning salmonids using data from ARIS and DIDSON sonar cameras. We hope that this will lessen the burden of those who are currently manually counting and annotating salmon, freeing them to pursue other projects that help the species.

The Team

We are a nationwide team mostly based out of three academic institutions: Caltech, MIT, and UMAss Amherst.

The Professors: Pietro Perona (Caltech), Sara Beery (MIT), Grant Van Horn (UMass Amherst), and Georgia Gkioxari (Caltech).

The Researchers: Michael Hobley (Caltech), Suzanne Stathatos (Caltech), Justin Kay (MIT), Timm Haucke (MIT), Kai Van Brunt (MIT), Sevan Brodjian (Caltech), and Madison Van Horn (Caltech).

We are also helped by Erik Young, a volunteer and coordinator for the project. We are also grateful for the help of all of our collaborators, which include fishery agencies at the State, Federal, and international levels, and environmental nonprofits.

Current Results

As you probably know, the basic approach is to use existing data to "train" an algorithm that can automatically recognize, track, and count fish. Two questions arise regarding the training video being used. First, what is the result when appplying an algorithm to a river when the training data is restricted to that same river and location? Second, can an algorithm be developed that is sufficiently accurate if the training data is developed from other rivers? The latter case would be preferable, as it would minimize the process of retraining every time the camera location is changed.

Based on the data we have gathered to date, We are currently able to achieve approximately an accuracy level (based on comparisons to manual counts) of:

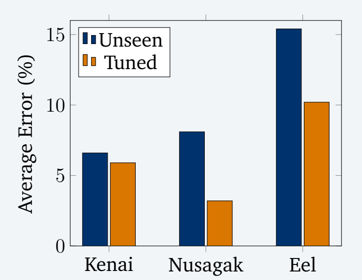
- 6.4% when we use an algorithm that has been trained based on data from the same river.
- 10% when we use use an algorithm from rivers for selected rivers we have not previously seen.

However, it should be noted that some rivers are more difficult than others and so our performance decreases in these settings.

The difficulty of counting a river is based on many factors, for example:

- Water clarity
- Occluded areas due to rocks or topography
- Differing fish size, velocity, density, and swimming behaviour
- Focal length; distance of the fish from the camera
- Sonar echoes

Example Results for Three Rivers

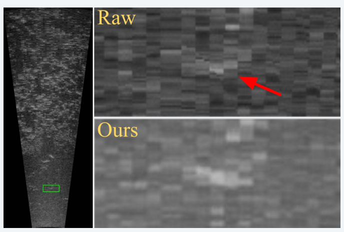


The results improve when we are able to fine-tune a model on a given river. However, this requires annotated data from that location. (Blue) results on a given river when trained on only data from other rivers, (Orange) results on a given river when trained on other rivers and that river.

Research Update

Suzanne: with assistance from the team, submitted a paper for publication at a top conference on a new denoising method, SAVeD. SAVeD focuses on removing background noise in rivers while exaggerating fish via their motion to make their downstream detection and tracking easier. SAVeD is a (relatively) lightweight preprocessing step that has the ability to make use of all of the data that we have, annotated or not. We use Elwha as our test river site, as it is one of the rivers our methods find most difficult. The detection, and subsequent tracking and counting, performance of SAVeD denoised frames is markedly better than the three-channel (raw, background-subtracted, and frame-to-frame-absolute difference) frames that we have previously used. We find that SAVeD boosts detection performance by 5.1%. Tracking performance improves by 26.7% (from 37.4 to 47.4), and count error decreases by 38.1% (from 54.8 to 33.9). Suzanne will search for ways to improve the computational efficiency of these methods and continue her effort to remove components from the machine learning pipeline that require labels, so that we can more quickly integrate more of the data that we have been given into a model.

An example fish with and without SAVeD



SAVeD clarifies fish appearance. Here, the fish becomes a clear and contiguous object. This makes detection easier, which in turn improves tracking and counting performance.

Kai: Inspired by the prevalence of echogram usage by the community of Aris and Didson users, Kai developed a lightweight method which utilises only this spatio-temporal representation. The benefit of finding fish from echograms, rather than from the raw videos, is that an echogram compresses several hundred video frames into a single image allowing for very efficient analysis.

Kai's method was also able to utilise unannotated data to improve the accuracy of their method without the need for further laborious annotations.

Using only the echogram, which is a few hundred times more space efficient than the raw video frames, Kai was able to achieve a 23%-26% error on data gathered from Kenai.

Justin's manuscript on adapting models to new deployment locations was recently accepted for publication in the journal Transactions on Machine Learning Research. The method we introduced, ALDI, automatically adapts to new environments and can improve performance by upwards of 20% at new rivers without requiring human annotation effort. Justin and team also recently submitted a follow up paper focused on label-efficient verification of model performance that can be used to select the best candidate model for deployment with as few as ten labelled images.

Engineering Development

In parallel with the scientific research efforts, Madison is developing a robust and efficient codebase. The design requirements of this engineering stage are based off insights gained from our stakeholders in our meetings over the last year. This codebase will operate to allow stakeholders to utilise the methods developed by our researchers locally, on their own data. This code base will be released open-source, free, and available for all.

An initial mock up is displayed below. This would allow for visual inspection of the results. The results would also be compiled into a csv or excel file to allow for easy analysis.