AFSC PROCESSED REPORT 2023-01

Alaska Fisheries Science Center
Fisheries Monitoring and Analysis Division

# Development of Electronic Monitoring (EM) Computer Vision Systems and Machine Learning Algorithms for Automated Catch Accounting in Alaska Fisheries

MARCH 2023



U.S. Department of Commerce | National Oceanic and Atmospheric Administration | National Marine Fisheries Service

AFSC Processed Report

This document should be cited as follows:

Lee Son, G. S., Romain, S.,  Rose, C. S., Moore, B. J., Magrane, K. A.,Packer, P. S., and Wallace, F. R.  2023. Development of electronic monitoring (EM) computer vision systems and machine learning algorithms for automated catch accounting in Alaska Fisheries. AFSC Processed Rep. 2023-01, 113 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

This document is available online at: https://repository.library.noaa.gov/

# Development of Electronic Monitoring (EM) Computer Vision Systems and Machine Learning Algorithms for Automated Catch Accounting in Alaska Fisheries

G. S. Lee Son[1], S. Romain[1], C. S. Rose[2], B. J. Moore[1], K. A. Magrane[1], P. S. Packer,[3] and F. R. Wallace[3]

[1] Pacific States Marine Fisheries Commission

7600 Sand Point Way NE

Seattle, WA 98115

[2] FishNext Research, LLC

Seattle, WA

[3] Fisheries Monitoring and Analysis Division

Alaska Fisheries Science Center

National Marine Fisheries Service

National Oceanic and Atmospheric Administration

7600 Sand Point Way NE

Seattle, WA 98115

March 2023

# ABSTRACT

Successful fisheries management is dependent upon the collection of data from fishing activities. Fishing data supports and improves stock assessments to ensure that catch limits are sustainable in the long term. Electronic monitoring (EM) has been shown to be an effective tool to meet fisheries monitoring objectives, particularly in compliance-based programs. However, while current systems allow for an alternate method for acquiring data, these collections require manual review and analysis to extract the meaningful catch accounting information. This can be a costly and time-consuming effort. Additionally, challenges in deploying EM camera systems arise due to complex hardware and software operational requirements, varied boat sizes, designs, and gear types, and the damage that can be done to electronics when exposed to harsh ocean environments. The EM Innovation (EMI) project, supported by the Fisheries Monitoring and Analysis Division (FMA) of the Alaska Fisheries Science Center (AFSC), aims to address these issues by researching and piloting cost-effective and durable machine learning and computer vision (CV) advancements for EM camera system deployments, with the goal of providing near-real time, automated, catch accounting and reporting.

EMI research consists of the development and deployment of camera systems for acquiring imagery and the development and integration of automated CV machine learning algorithms and applications. The purpose of these systems is to detect and identify catch events from fishing imagery and to classify those detection to species or larger taxonomic groups. Once these detections are made, further data analysis about the catch event can be obtained, such as length estimation and count information. Algorithms have different requirements based on the detection types and the fisheries environment involved. EMI identified multiple fisheries applications where CV can be of use. These include 1) automated species detection, identification, and length estimation of fish as it is caught at the rail of fixed gear (longline) vessels; 2) species identification of fish images collected in controlled environments; 3) detection, count, and length estimation of Pacific halibut (*Hippoglossus stenolepis*) bycatch; and 4) detection, count, and distinction of salmon from processing plant belts containing multiple species of fish. Additional algorithm functionality for the detection and monitoring of crew member activity on vessel decks is also presented.

Research is conducted through an iterative development lifecycle with four principal areas. These include the following: 1) the development and deployment of the camera systems for data collection; 2) the annotation and cataloging of the collected data; 3) the development and training of the algorithms based on the annotated data; and 4) the analysis and interpretation of the algorithm results. Each cycle has an associated output, these being: an EM camera system a set of algorithms or a combination of both. Results have been very promising and are presented in this Processed Report for each functional stream. By leveraging these latest developments in computer vision, cost-effective and timely extraction of scientific data from images will provide greater certainty for resource management and will support sustainable fishing practices.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

# 1. INTRODUCTION

Approximately 58% of U.S. seafood was landed in Alaska in 2018 (NMFS 2020), making the federally managed commercial fisheries off Alaska one of the largest in the United States. The North Pacific Observer Program (Observer Program), administered by the Alaska Fisheries Science Center's (AFSC) Fisheries Monitoring and Analysis Division (FMA), plays a vital role in the conservation and management of these Alaska groundfish and halibut fisheries. FMA monitors groundfish and halibut fishing activities in the Federal fisheries off Alaska and conducts research associated with sampling commercial fishery catches, estimation of catch and bycatch mortality, and analysis of fishery-dependent data. FMA is responsible for training, briefing, debriefing, and oversight of observers who collect catch data onboard fishing vessels and at onshore processing plants. Successful fisheries management is dependent upon the collection of data from these fishing activities. Fishing data, such as the number of fish that are caught, the fishing effort (the number of hours or days spent fishing), and bycatch information, support and improve stock assessments and ensure that catch limits are sustainable in the long term.

Data are traditionally collected by at-sea and shoreside observers, who collect data on what fishermen land and discard. More recently, FMA has invested in alternative data collecting methods for which the deployment of observers is difficult or restrictive. Electronic monitoring (EM) is one such data collection method. EM technologies are being used in several applications in the North Pacific and elsewhere. In the North Pacific, video technology is used for two purposes: 1) estimating catch and discard for fisheries management and 2) monitoring for compliance with regulations.

The Observer Program Annual Deployment Plan (ADP) dictates when and where to deploy observers based on a scientifically defensible deployment plan reviewed annually by the North Pacific Fishery Management Council. Since 2018, the deployment of EM systems has been included in the ADP as a monitoring option for longline gear, and in 2019 was extended to include EM as a monitoring option for pot gear.

While EM holds the potential to make data collection timelier, more accurate, and more cost-efficient, there are some drawbacks and limitations which need to be overcome for this potential to be realized. Some of the real-world practical challenges include complex hardware and software, varied boat sizes and designs, and the damage to electronics when exposed to saltwater and pounding waves. Additionally, the video collected requires human review to extract the data needed for quota management and stock assessments which is a costly and time consuming effort. The EM Innovation (EMI) project, under FMA, aims to address these issues by researching and piloting cost-effective and durable methods for collecting and automatically processing the data from EM system deployments. This automated processing can be achieved using computer vision (CV).

CV is a field of artificial intelligence (AI) that trains computers to interpret and understand the visual world. Using digital images from cameras and videos and deep learning models, machines can accurately identify and classify objects. For fishery EM, CV algorithms and models are trained using data collected from EM camera systems. Once trained these models can be used to detect fish and fishing events within EM video data. Further deep learning models can then be applied to make determinations based on these detections, such as identifying the fish to species or determining the length of the fish.

The accuracy and dependability of the results of CV algorithms and models is dependent on the training input provided. For the EMI project, specific EM training imagery was needed based on the automation required. For example, longline imagery is needed for training the longline detection algorithms, while multiple images of various fish species in various poses are needed for the identification models. Camera specific training imagery is also needed, such as stereo imagery for determining lengths and multispectral imagery for determining training features. EMI conducts the research, development, and deployment of both the camera systems needed for the collection of training data as well as the CV algorithms and models for fish detection, fish tracking from one frame to another (in some cases the tracking of multiple fish in the frame at the same time), fish species classification, and fish length estimation.

The EMI project aims to provide integrated CV EM camera systems and applications for a more cost-effective and timelier method for collecting and providing analyzed data for catch accounting purposes. Any realized cost savings could allow for increasing coverage among vessels for the same total costs. The advances made in this research have the potential to benefit similar EM programs that require CV, as the CV algorithms and models can be re-trained for new datasets based on training imagery collected from other camera systems.

This Processed Report aims to detail the research methods and results for camera systems and algorithm development. The work conducted by the EMI falls under Goal 3 of the NOAA Fisheries Strategic Plan (NOAA 2019): Improve organizational excellence and regulatory efficiency through institutionalizing the use of innovative technologies. Supporting the development, leveraging, and use of powerful technologies (e.g., artificial intelligence and machine learning), digital platforms and electronic monitoring, for conducting surveys, enhancing, and improving the accuracy of observing systems, and collecting and sharing data in cost-effective, transparent, and real-time approaches.

## 2. METHODS

The objective of EMI is to provide EM camera systems and computer vision algorithms to collect and analyze fisheries data for North Pacific commercial fisheries. The basic functions of the algorithms are to detect and identify catch events from fishing imagery and classify that detection to species or a larger taxonomic group. Further analysis tasks include collecting length estimations and counts. This automated collection and analysis technology can be applied to the hook and line, trawl, and pot fisheries. While the functions of the algorithms are fundamentally the same (detection, classification, tracking, and length estimation) they have different requirements based on the detection types and fisheries environment. For instance, detecting and analyzing a fish from an image captured as the fish is being hauled over the rail at sea, with variable lighting, mixed background, and object occlusion, has distinct factors to consider compared to analysis of a clear, top-down image of a fish acquired with consistent lighting and a stationary featureless background. As such, EMI has defined its research areas and work streams based on the requirements needed for each CV algorithm function as well as the types of camera systems and image data involved. These work streams are detailed in Table 1.

Table 1. -- EMI research streams based on functional algorithm requirements.

| Functional algorithm requirements | Purpose | Applicable fishery | EM camera system type | Research output |
|---|---|---|---|---|
| Fish detection, species identification and length estimation | Automating the detection, tracking, classification, retention, and length measurements of catch events at the rail on fixed gear vessels in an uncontrolled natural environment (mixed weather and lighting conditions) | Hook and Line (HAL) | Stereo CV camera system | EM camera system and automated algorithms capable of running in real-time (as imagery is collected) and running post-processed (once all imagery has been collected and is ready for analysis) |
| Fish species identification | Automating species identification in a controlled environment under optimal conditions (clarity, focus, lighting) | Trawl, Others as needed | Single camera system for data collection | Machine learning algorithms for post processing analysis |
| Halibut detection and length estimation | Automating the count and length estimation of halibut discards from on-deck sorting. | Trawl | Single camera enclosed chute system | Integrated EM camera chute for real-time analysis of halibut bycatch |
| Salmon detection on belts of processing plants | Automating the detection and count of salmon bycatch on processing plant belts for validation compliance | Trawl | Single or multiple camera system | Automated algorithms capable of running real-time and post processed |
| Crew detection and activity monitoring on deck | Automating the detection of crew and tracking activity of crew actions | HAL, Trawl, Others as needed | Single camera systems | Automated algorithms capable of running real-time and post processed |

Each stream follows the same research development cycle. This iterative cycle consists of areas: 1) the development and deployment of the camera systems for data collection; 2) the annotation and cataloging of the collected data; 3) the development and training of the algorithms based on the annotated data; and 4) the data analysis and interpretation of the algorithm results. Each cycle has an associated output, these being an EM camera system, a set of algorithms, or a combination of both. Figure 1 depicts the iterative research method.

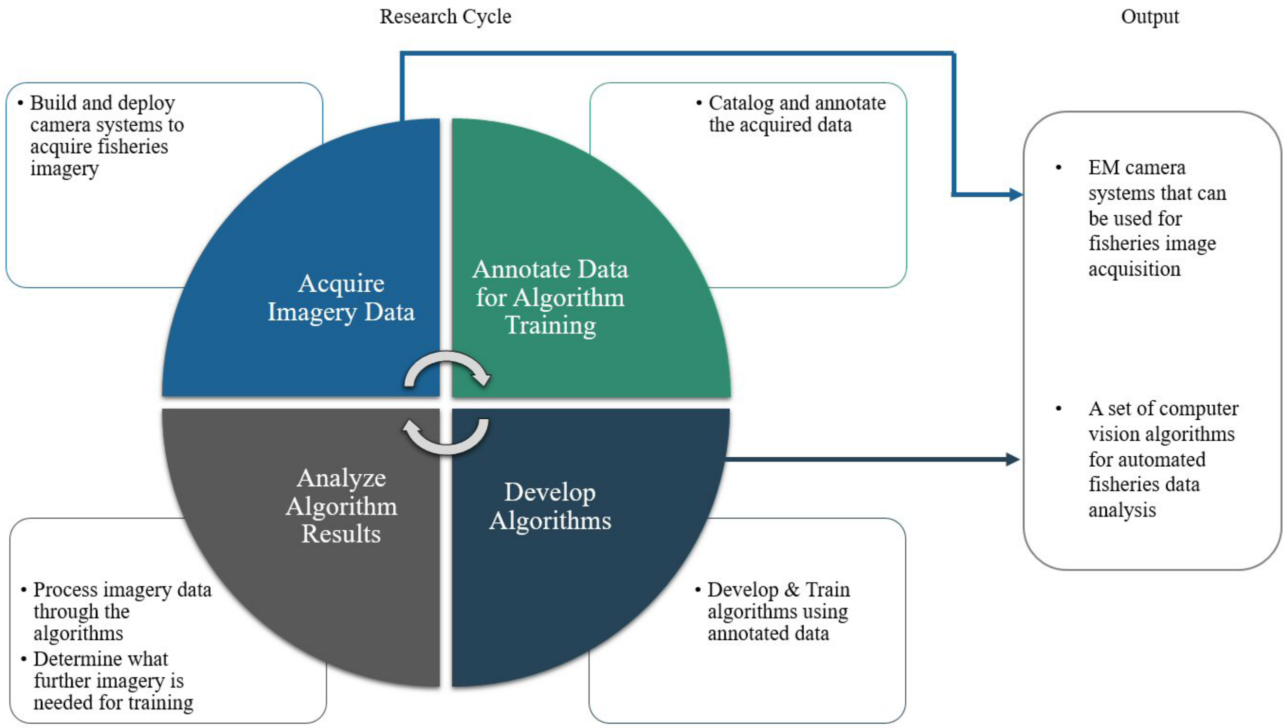



Figure 1. -- Development cycle and project output.

The EMI team consists of fishery biologists, application developers, electrical engineers, and data annotators. Development of automated EM systems capable of identification of species and size from collected imagery relies on interdependent camera systems (hardware and associated software) and image analysis CV algorithms, models, and applications. Hardware designs are uniquely adapted to collect imagery that meets the specific image quality requirements of the algorithms. During development and testing of the systems, hardware is continually adapted as the image analysis algorithms become more accurate and more robust. It is important to note that the images collected by camera systems are not data that can be used by fisheries managers as data; images must first be processed into numeric data representing catch composition and quantity to be used in estimation, stock assessment work, and other analytic activities. Currently, to perform this data transformation, manual review of the imagery is needed to extract required information. To extract and transform this data automatically the CV algorithms must be able to detect, track, and classify an object or behavior of interest, as well as estimate the size of the object of interest. Images from fishing vessels vary between gear type, target species and lighting conditions throughout collections. Hence, the hardware (cameras, sensors, and associated software) is developed to support the data needs of computer vision algorithms that provide the actual data used in fisheries management.

The computer vision algorithms are developed using imagery acquired through EMI systems deployed on volunteer commercial fishing vessels and scientific research surveys (IPHC, and NMFS Sablefish and BSAI/GOA Trawl). This imagery, in the form of photos and videos, is collected using camera systems and hardware sensors built and designed by the project for the purpose of image acquisition and subsequent analysis. Imagery is acquired, cataloged, and annotated by EMI staff, and then passed on to the research team at the University of Washington's Department of Electrical and Computer Engineering (UWECE) for algorithm development and deep learning model training. The results of the algorithm development are then jointly tested and analyzed, and refinements to the data requirements are specified. The acquisition camera systems are then updated and redeployed to collect the new image data and the iterative development cycle starts again.

Each work stream cycle is detailed in its own section below with focus on the four iterative development steps. Experiments were periodically conducted to apply the algorithms to areas beyond the scope of the work streams, such as applying the fish species identification algorithms on birds for classification of birds. These experiments are highlighted together with discussions on the steps needed for operational readiness.

# 3. SPECIES DETECTION, IDENTIFICATION AND LENGTH ESTIMATION - EMI RAIL

EM systems are currently deployed on small, fixed gear vessels with the purpose of collecting data for review (50 C.F.R. §679.51, 2021). These surveillance systems record video of catch events as the fish are being hauled over the rail as well as the deck to monitor crew activity. Videos are then manually reviewed frame by frame to extract meaningful catch accounting information. This is a labor- and cost-intensive task with teams of dedicated video reviewers. The time between the actual haul event and final analysis can be quite long as data drives are only collected once fishing trips are completed and transferred to the reviewing agency. Additionally, the time to complete a review of the data averages around an hour of review time for every hour of video data (NMFS 2019). Thus, it is apparent that this review process could significantly benefit from image processing automation.

The objective of EMI Rail is to develop automated video analyses to count, identify to species or species group, determine catch disposition, and measure fish as gear is retrieved during multi-species longline fisheries. These computer vision algorithms could be integrated into EM systems to achieve real-time analysis for catch accounting. While this goal is challenging due to environmental and technical limitations, such as power consumption and processing capabilities, it is the long-term goal of EMI Rail.

The current focus of EMI Rail is to integrate the algorithms into the human review process to supplement and aid the human review output as well as to alleviate the workload required for processing such vast amounts of data. Under laboratory environment testing, the current rail algorithms can successfully detect fish within a frame, identify fish to species, and determine if that fish had been previously identified from a prior frame (tracking). Fish length is estimated from stereo image data. Tracking also determines if the fish was kept or discarded. Further determining fate disposition, (discarded live, discarded dead) may be possible, although this needs to be investigated further. Figure 2 below summarizes the iterative research and development cycle for EMI Rail and anticipated output.

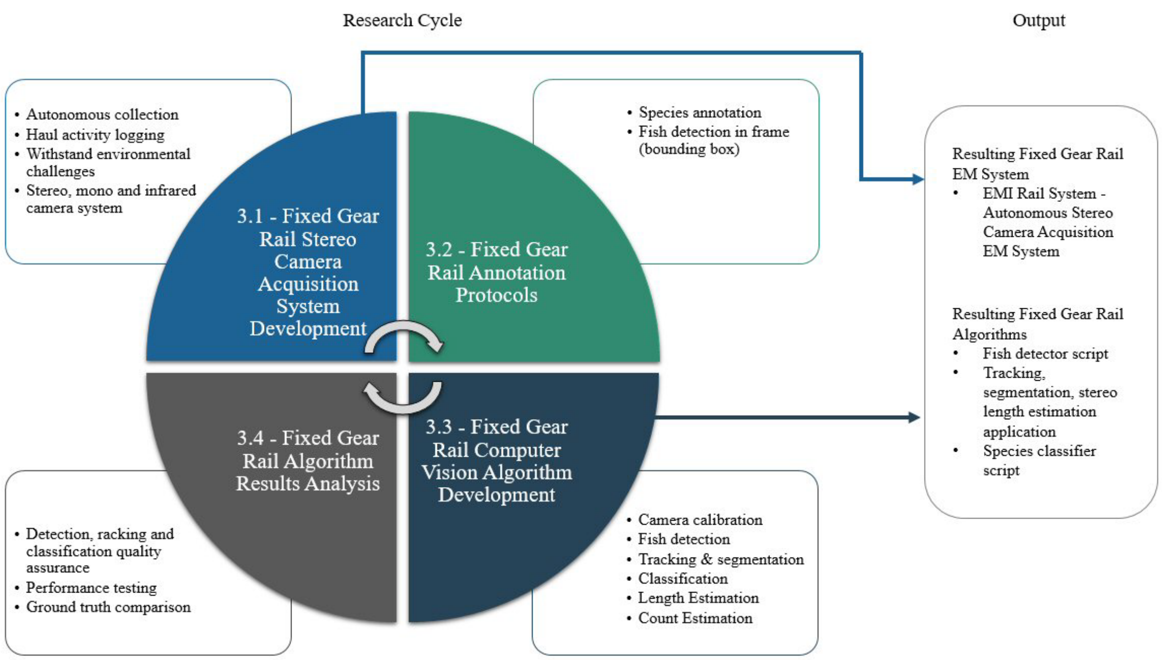



Figure 2. -- System and algorithm development cycle for fixed gear fisheries.

## 3.1 EMI Rail System - Stereo Camera Image Acquisition

To estimate fish length, stereo imagery is needed to project the shape of the fish in 3D to determine depth perception. Depth data allows the machine learning algorithms to discriminate objects or behaviors of interest with precision unavailable from two-dimensional (2D) images. Stereoscopic vision in humans allows the brain to extract three-dimensional (3D) information from the environment, which allows for the perception of depth. In machine vision systems, stereo images provide the information necessary from paired 2D images to accurately measure an object's size and distance from the camera (Stevens et al. 2013). Obtaining size measurements from stereo imagery has been realized by other applications, such as CamTrawl (Williams et al. 2013), and at the start of EMI's research program in 2015, this system configuration was the most promising to develop. As there were no stereo camera systems on the open market specifically for marine applications at the relevant distances and resolutions, this project began exploring automated measurement with the development of custom hardware and software.

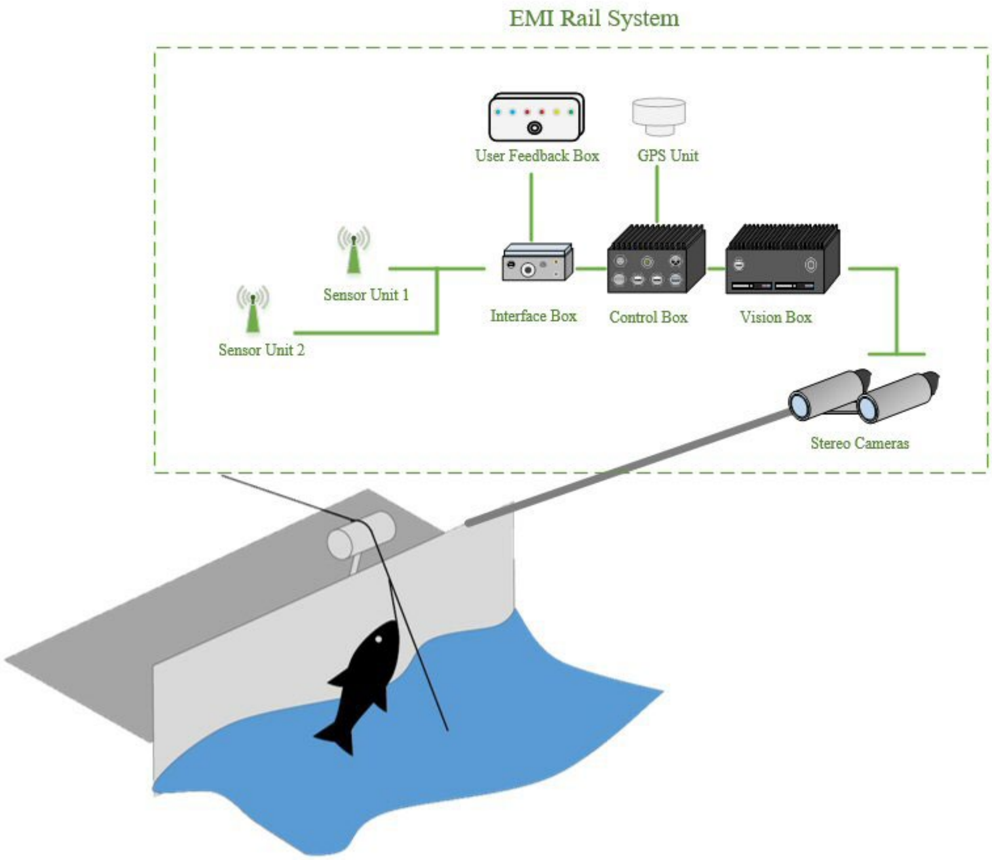



Figure 3. -- Components of the EMI Rail System.

The EMI Rail system consists of hardware and software components that are installed and deployed on vessels. Hardware includes machine vision cameras, computers, GPS unit and sensors, while software includes the stereo acquisition application and system control application (Fig. 3).

The EMI Rail systems is iteratively designed, built, and deployed to satisfy several requirements and constraints (Table 2). The data acquired by EMI Rail system is then cataloged, labeled, and annotated, and these annotations are then used to train and build the machine learning algorithms.

Table 2. -- EM camera system requirements for the Rail system.

| Design requirement/constraint | Description |
|---|---|
| **3.1.1 - Stereo Image Acquisition** | This system needs to perform in the same manner as current EM systems in the industry but using stereo machine vision cameras. Images need to be acquired in stereo pairs with clear views of the rail. Image quality requirements are based on the needs of algorithm development (resolution, angle, distance). |
| **3.1.2 - Environmental Constraints** | The deployed hardware needs to be able to withstand challenging environmental surroundings. The design and build needs to accommodate power restrictions as frequent power outages are expected. Maintenance of the hardware needs to be as simple as possible as well. |
| **3.1.3 - Support For Autonomous Collection** | The system needs to operate without the need for input action from the vessel crew. Continuous recording of the entire trip is not possible; therefore, the system must be able to determine periods of data collection based on vessel hauling activity. |
| **3.1.4 - Haul Activity Logging** | Along with haul imagery, a log of vessel hauling activity needs to be recorded, including vessel permit ID, trip numbers, vessel location, haul start and end time |
| **3.1.5 - Deployments** | Vessel and administration buy-in is required with adequate deployment plans |

### 3.1.1 Stereo Camera Image Acquisition

Initial development of the stereo camera system leveraged the research conducted for CamTrawl, a stereo camera acquisition system used for underwater research (Williams et al. 2010). The CamTrawl system includes a custom printed circuit board (PCB) and includes a microcontroller for sensor and camera triggering, power management electronics and enclosures for the computer and cameras. Development of the EMI Rail system follows this same design pattern. Custom PCB boards were designed to support the camera triggering sensors and power management. Custom camera enclosures were also built for the MV cameras. Dedicated stereo image acquisition software was developed to manage the recording and capture of the image frames.

GigE vision cameras (MV cameras) were selected and implemented as part of the EMI Rail system as they allowed for some degree of programmable flexibility. GigE vision is an interface standard for high-performance industrial cameras and provides a framework for transmitting high-speed video and related control data over Ethernet networks.

The design and deployment of the camera housing went through multiple iterations. In early designs, custom housings were built, but small camera movements would lead to frame loss and calibration issues. The material used for fabrication of the housing was difficult to work with and the design was too rigid for trying different camera angles. It was also expensive to produce and would be cost-prohibitive for deploying at a large scale. As such, it was determined that more affordable off-the-shelf housing units that allow for standard camera orientations would be a better option.

Waterproofing of custom designs also proved challenging. To prevent immersion of cameras due to vessel movements, cameras would have to be angled up from the vessel deck. Vibration prevented some hardware from maintaining waterproof integrity and some camera lens settings were also negatively affected. These issues were addressed in later designs, with one of the solutions being the use of Loctite in all hardware and another being the addition of an 'air cushion', a mechanism to keep the camera lenses from fogging up. The images depicted below in Figure 4 highlight the evolution in camera housing design.

Figure 4. -- Evolution of stereo camera housing design with the left-most photo showing the initial camera housing through to the current housing design.

For stereo image acquisition and collection, a software application building upon the CamTrawl framework was developed. This application is responsible for acquiring image pairs at a specified frame rate and recording them to disk. While initially built using the GigE standard to allow for easy adaptation to different camera products, this eventually revealed itself to not be necessary. It was determined that variations between the MV camera vendors were greater than anticipated and priority was given to camera pairs that proved more robust during deployments and whose firmware was well supported. This, however, does limit the current stereo image collection hardware to only those camera types. Further research and development in 3D camera stereo acquisition will need to address this issue and there are more low-cost options appearing on the market to collect stereo images as the 3D information is useful for many computer vision algorithms.

In early builds and deployment of Rail systems, image pair syncing was also an issue. This was mostly caused by moving the camera triggering from a hardware-based to a software-based triggering mechanism. While this removed the need for a custom micro-controller, the signal conditioning required between the camera pairs increased the complexity of custom built boards that utilize the internal opto-isolated trigger lines. Exposure settings can also cause the software-triggered cameras to get out of synchronization with each other. This is difficult to detect without direct comparison of images displaying a unique event that can be assessed to the millisecond.

The camera pair is connected to a waterproof PC via Ethernet. This PC is configured to run the acquisition application on system startup and begin acquiring images until the PC is shut down. The shutdown signal is triggered from the controller application running on a second, lower powered, waterproof PC (see Section 3.1.4 for further details on the control PC). Figure 5 below shows an image pair acquired through the EMI Rail system.


Figure 5. -- Example of synced stereo images acquired with the EMI Rail system.

MV cameras with infrared (IR) capabilities were also deployed on study research vessels using the same configurations. In these deployments four cameras are installed, a pair of standard MV stereo cameras and a pair of IR cameras. Two sets of image pairs are acquired simultaneously.

While the strategy to deploy MV cameras was chosen with the intention of providing high resolution images for algorithm development, it became unnecessary as latter training and development with lower resolution images produced equivalent results when compared to the training of higher ones. The high resolution affects the size of the image files, and the image data volume negatively impacts real-time processing and storage options. Internet Protocol (IP) video cameras, like those used in standard EM deployments, capturing lower resolution images have been integrated into the 2020 stereo acquisition collection systems to enable testing and development of single camera algorithms for use with Rail data.

### 3.1.2 Environmental Constraints

There are significant challenges when deploying EM systems onboard fixed gear vessels for use at sea. The two most significant of these are the environment, ensuring electronics are protected from the ocean elements, and managing the power constraints and restrictions on the vessels.

It was established early in the system development life-cycle that a waterproof PC would be required. Proof of concept PCs were built with this aim in mind using custom fabricated housings. Early deployments were made with these systems, but it became apparent that these components would not be robust enough to withstand the 16–20-hour collections per day and overheating and data storage issues needed to be addressed. Custom-built PCs were also not sustainable as the cost was not in line with the strategic goal of having cost effective systems that vessels could buy and install on their own. To overcome these issues, off the shelf fan-less IP67 rated waterproof PCs were obtained. While system uptime increased significantly, other issues such as endpoint corrosion were unavoidable. Vibration was also an issue in some deployments, affecting the ability to maintain waterproof integrity and negatively impacting camera cabling connections. Mounting the PC units onto plates helped mitigate some of these issues. Due to the variation to the location of PC units onboard the vessels, unique configurations were developed for each deployment. Therefore, a uniform design could not be developed across vessel installations. Images in Figure 6 show the evolution of the PC components.


Figure 6. -- EM Rail system hardware evolution. Left image depicts initial design with right image being the most recent deployed version.

Power consumption onboard the deployed vessels was also a major concern. Power cuts can be a frequent occurrence and protecting system components with individual fuses became necessary given the environment and complexity of the wire runs. It was also unrealistic for the cameras and PC to be continuously powered, especially when there is no fishing activity, or to require the vessel crew to turn on the cameras and PC every time they begin to retrieve gear as this can lead to inconsistency and misrepresentation. Hence, triggers and autonomous collection strategies were needed.

### 3.1.3 Support for Autonomous Collection

The EM Rail System needs to be autonomous; it should run without the need for user input. This is to allow for minimal impact to the fishing activities of the crew as well as to keep accurate records of hauling activities. To detect fishing activity sensors were placed onboard to detect proper intervals for image data collection. Some sensors, like hydraulic sensors and GPS, were like those used by standard EM camera systems that monitor gear and vessel activity. Other sensors monitor human behavior or location by utilizing acoustic- or infrared- derived proximity measurements. Sensors indicate when the haul starts and ends. The sensors tested were connected to a microcontroller unit and low-power control PC. The control PC monitors the sensor data and sends a Wake-on-LAN (WoL) signal to the connected high power acquisition PC which in turn powers up the cameras and acquisition software and begins acquiring images. Once the sensors indicate that hauling activity has ceased, a shutdown signal is sent to the camera PC.

Sensors tested include hydraulic sensors, acoustic, infrared light sensors, LIDAR sensors, and IP cameras. Hydraulic sensors proved to be difficult to work with and the exclusive use of these sensors to identify hauling periods and trigger cameras proved to be unreliable for precise identification of the start and end of fishing periods. Additionally, hydraulic sensors are highly variable between vessels due to differences in individual vessels' dockside and at-sea power supplies. Secondary sensors were developed to identify fishing activity with more precision. Radio-frequency identification (RFID) tags and acoustic proximity sensors were deployed to supply more precise haul time data, but both proved to be unable to withstand the marine environment for extended periods of time. Infrared (IR) sensors were tested as drum rotation sensors and as proximity sensors; both performed well but ran the risk of getting blocked by gear after installation. As a drum rotation sensor, they performed well and required less waterproofing than the haul sensors. Garmin Lidar proximity sensors were also tested and were found to be sufficiently robust for deployment.

Current research aims to incorporate region-of-interest IP cameras and the detection of a crew member (presence/absence detection) as a trigger to start monitoring systems recording. Section 7.5 elaborates on this detection trigger.

#### 3.1.4 Haul Activity Logging

To collect useful data for management, metadata is automatically recorded for each haul as part of the system. This data consists of information such as: vessel permits, trip number, vessel location and haul start and end times. For positional and time information, a GPS unit is connected to the continuously powered control PC. GPS coordinates are logged and the start and end times for each haul are recorded when the connected sensors indicate when to start and stop the haul monitoring. Timestamps are also included on all images that are acquired, aiding with the timekeeping reports when needed. When combined with the analysis results from the haul imagery, a holistic dataset can be obtained of what fishing activities occurred for each haul.

### 3.1.5 Deployments

Deployments were conducted on vessels targeting Individual Fishing Quota (IFQ) Halibut and Sablefish, the AFSC's Auke Bay Laboratory's Sablefish Longline survey, and the International Pacific Halibut Commission's (IPHC) setnet survey. Prior to each deployment, the MV cameras and image quality settings were configured to the specific environment on the vessel. Infrared cameras were field tested for the capture and analysis of fishing events in low light conditions. For certain deployments, it was only necessary to test the GPS unit and sensors for identifying precise haul start and stop time without cameras. These deployments were specifically made to test drum rotation and proximity sensors. This sensor testing system is referred to as EM Lite. Table 3 lists the deployments made to date for the rail stream.

Table 3. -- EMI Rail system deployments.

| Year | Deployments | Notes |
|---|---|---|
| 2015 | 3 vessel deployments<br>• FV *Bold Pursuit*<br>• FV *LaPorsche*<br>• FV *Northern Endurance* | 44 hauls including 35 paired hauls<br>634 GB of data |
| 2016 | 3 vessel deployments<br>• FV *Kariel*<br>• FV *Marilyn J*<br>• FV *Middleton* | 120 hauls including 97 paired hauls<br>1340 GB of data |
| 2017 | 4 vessel deployments<br>• FV *Kariel*<br>• FV *Middleton*<br>• FV *Pender*<br>• FV *Van Isle* | 104 hauls including 84 paired hauls<br>Infrared cameras were tested |

Table 3. -- Continued

| | | |
|---|---|---|
| 2018 | 5 vessel deployments<br>• FV *Alaskan Leader*<br>• FV *Defender**<br>• FV *Kariel*<br>• FV *Middleton*<br>• FV *Predator* | 254 hauls including 173 paired hauls<br>Infrared cameras were tested<br>*EM Lite test deployment |
| 2019 | 7 vessel deployments<br>• FV *Defender**<br>• FV *Kariel*<br>• FV *Kema Sue*<br>• FV *Middleton*<br>• FV *Ocean Prowler*<br>• FV *Pacific Surveyor*<br>• FV *Predator* | 274 hauls including 71 paired hauls<br>Infrared cameras were tested<br>*EM Lite test deployment |
| 2020 | 3 vessel deployments<br>• FV *Alaskan Leader*<br>• FV *Defender**<br>• FV *Middleton* | *EM Lite test deployment |

The image data that are acquired and collected during the deployments were catalogued and annotated for algorithm training.

### 3.2 EMI Rail Annotation

Image annotation is the process of manually defining regions in an image and creating a textual description of those regions. Annotations can then be used to train machine learning algorithms for computer vision applications. For the training of detection and classification algorithms used with the Rail system, annotated Rail images are needed. Images are acquired as per Section 3.1 and annotated using 2D bounding boxes. Bounding boxes are rectangular points of reference used to designate object locations for image processing.

Bounding boxes are drawn over an image, shape, or text to define its X and Y coordinates. This is the start of the process of training a machine to recognize distinct types of objects. They are essential for object identification tasks. Table 4 below lists the annotation bounding box requirements for the training of the rail algorithms.

Table 4. -- Annotation requirements for EMI Rail

| Annotation requirement | Description |
|---|---|
| **3.2.1 - Species Detection and Classification Annotation** | The algorithm required at least 3,000 bounding box annotations for each species to provide useful confidence levels. For each vessel, 6 hauls are selected, with each containing a section of 10,000 images for annotation. Those images are used to train the algorithm. |
| **3.2.2 - Tracking and Quality Assurance Annotation** | Annotations needed for tracking algorithm |

### 3.2.1 Species Detection and Classification Annotation

To train rail detection and classification algorithms, the analysts indicated that at least 3,000 annotations (bounding boxes) per class would be required to account for the various appearances, background, poses and transformations the object can take. To tailor the learning to multiple angles and lighting changes, annotated images are required from different vessels.

When training algorithms for computer vision, annotation tools allow human annotators to move, transform, rotate, and scale the bounding boxes. The current annotation tool is the open source LabelImg software application (Lin 2016). LabelImg allows for object hierarchies and is used for annotating species identification purposes. A species hierarchy is used for annotating species labels to the lowest level possible. For instance, if the fish is an unidentified flatfish, the image is labeled flatfish unidentified. If the annotator can identify a species to Pacific halibut, the image is labeled as “Pacific halibut”. Both are classified as flatfish when run through the classification algorithms, with identification to species level directly related to user defined confidence levels. A user can then assess and define the quality of the algorithm derived data of species classification and select the taxonomic identification and associated confidence most appropriate for management needs.

Due to the vast amount of rail imagery acquired per year, an annotation protocol was developed and maintained for consistent algorithm training. For each year imagery was acquired, and for each vessel, six hauls were randomly selected for annotation. From each haul a random segment of 10,000 images were then selected to be annotated. These 10,000 images were reviewed, and bounding boxes were created for each fish, bird, and piece of fishing gear (hooks, buoys, etc.) that came into frame for the entire duration of time that target was in frame. Oftentimes multiple targets are in each frame as birds are often near the fishing gear. All fish and birds in frame are labeled this way and notes are made if the target is not on a hook. Table 5 below highlights the number of annotations made between 2015 and 2019. Due to time and resource constraints bounding box annotation for 2017 was skipped in favor of data from later years.

Table 5. -- Summary of annual Rail system image annotations.

| Year | Application | # Images reviewed | # of Bounding boxes |
|---|---|---|---|
| 2015 | Rail Camera | 527,276 | 1,560 |
| 2016 | Rail Camera | 50,000 | 31,001 |
| 2017 | Rail Camera | 50,000 | none |
| 2018 | Rail Camera | 156,7411 | 257,293 |
| 2018 | Rail Infrared Camera | 75,000 | 5,670 |
| 2019 | Rail Camera | 269,028 | 18,878 |

The LabelImg software outputs an xml file that includes the image name, species classification, and bounding box information (Fig. 7).

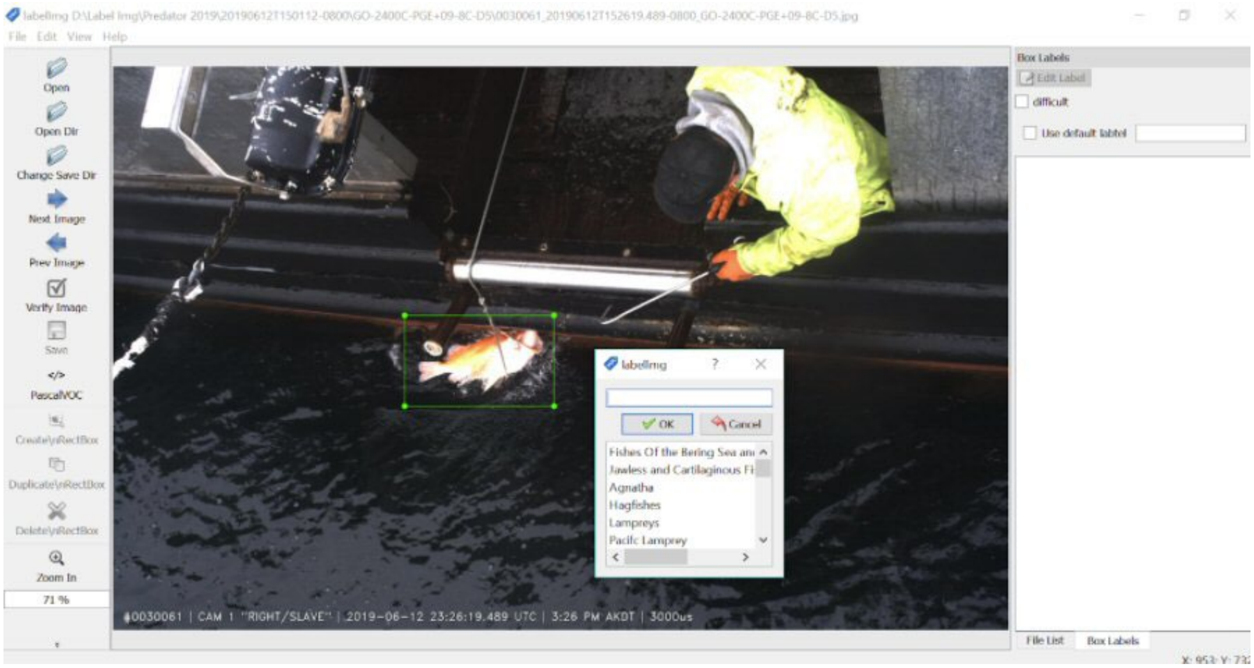



Figure 7. – The LabelImg user interface, with manual bounding box and label annotation.

### 3.2.2 Tracking and Quality Assurance Annotation

'Tracking' refers to the tracking of the detected object from one frame to another in a set of sequential frames. Accurate tracking is required to determine the count of unique objects. For example, in an hour of video footage there could be 1,000 fish detections, but this does not mean that 1,000 fish were caught, as the same fish will appear multiple times throughout the sequence of frames of the video. Tracking allows for the determination of one object from another in a sequence. By tracking the direction of the movement of the object, in this case a fish, tracking also allows for determining if the fish is coming up into the vessel or going back into the direction of the ocean. This direction tracking allows for the determination of whether the fish was discarded or retained.

Training of the tracking algorithms is accomplished by checking the output, correcting missed detections through annotation, and retraining the tracking model. Each track is checked for completeness to ensure there are no missed detections resulting in multiple tracks for the same object and consequently incorrect object counts. For this tracking annotation, the LabelTrack software program (Huang, 2017) is used. LabelTrack takes as input the output from the tracking algorithm which includes species classification, tracking identifiers (IDs), and bounding box detections and was developed specifically for this project. Each track is given a confidence score so that the tracks can be sorted by confidence level. This allows the reviewer to focus on a certain range of accuracy while bypassing the highly confident tracks. Figure 8 depicts a screenshot of the LabelTrack application and illustrates the difference between LabelTrack and LabelImg above.

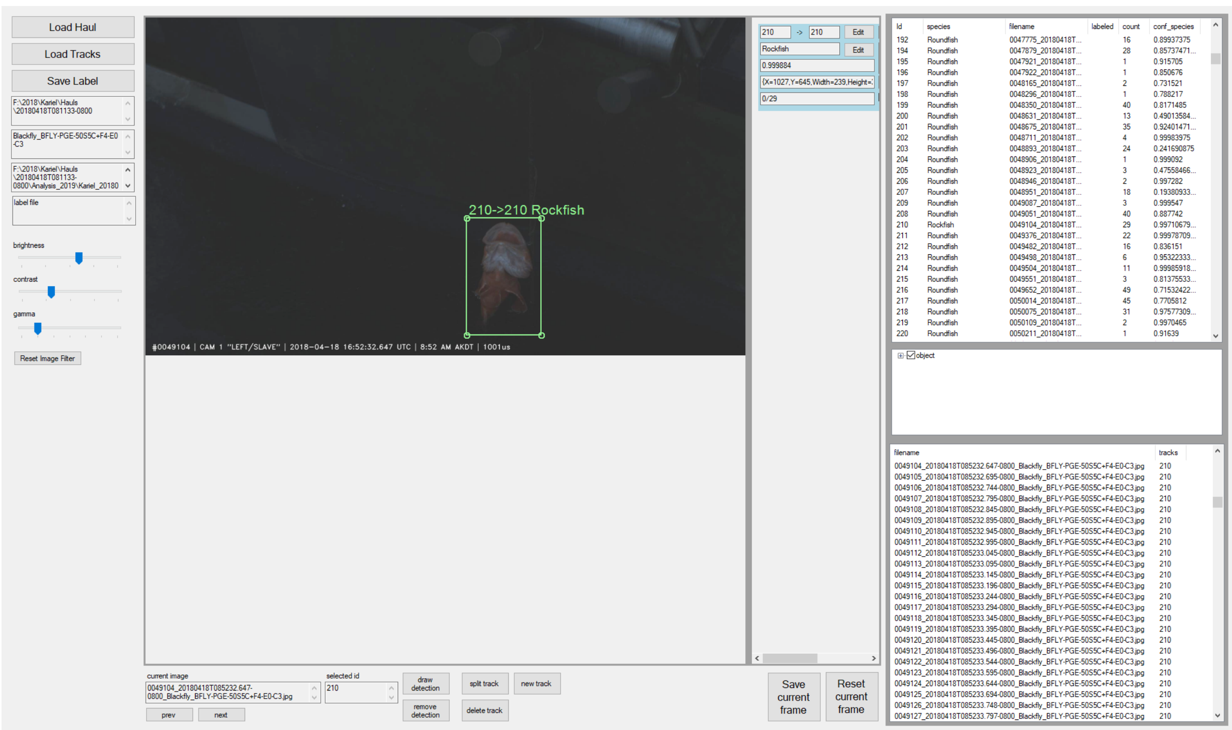



Figure 8. -- LabelTrack being used to check the output of the tracking algorithm; annotations (bounding boxes) are used to correct and retrain the tracking algorithm.

Tracking annotation review includes checking for species classification accuracy and ensuring tracks are connected through an entire catch event. This manual review is resource intensive as bounding boxes need to be created around each target object for the duration of time the object is visible, manually correcting each track per frame. Initial review of hauls took approximately 80 hours per person per haul to correct but as the accuracy of the algorithms improved with each subsequent retraining, the number of needed corrections decreased making the review process more efficient over time.

The output annotations from both LabelImg and LabelTrack are used for training the Rail detection, tracking, and classification algorithms.

### 3.3 EMI Rail Computer Vision Algorithms

The goal of the computer vision algorithms developed for fixed gear hook-and-line fisheries is to identify the species of each fish, count the number of each species, determine its length, and determine if the fish was retained or discarded from imagery collected at the rail. To accomplish this, the algorithms need to be able to detect the fish in the image frame, track the movement of the fish from one frame to another, and then classify the fish to species and estimate its length. Figure 9 depicts the steps involved from inputting stereo images to obtaining data suitable for catch accounting.

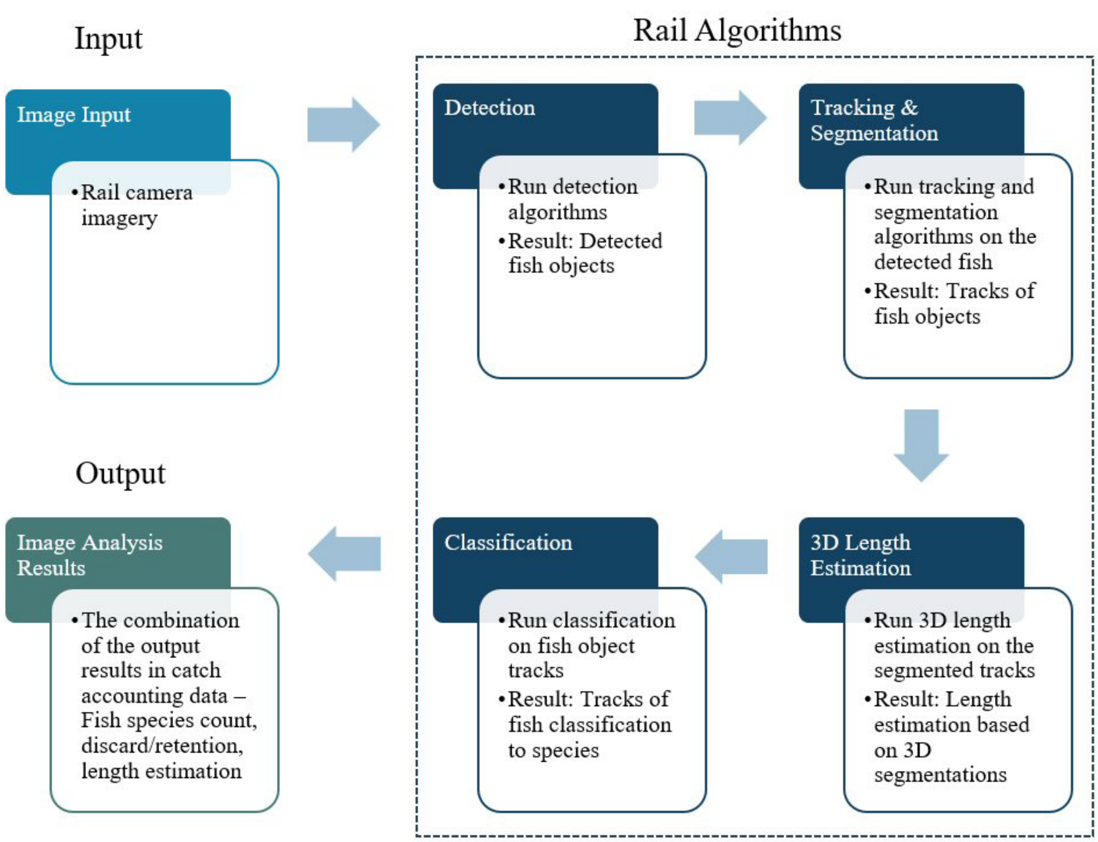



Figure 9. -- Flowchart showing the Rail system data inputs, algorithm sequences, and final output data.

There are significant challenges when attempting to automate the detection, tracking, and classification of fish at the rail. Deformable objects (objects whose shape changes as it moves in 3D space), noise from the sea surface, and dynamic image background all make conventional tracking, segmentation (separation of the object from the background of the image), and classification methods unreliable (Huang et al., 2019). As a fish on the longline gear (hook) moves from the waterline to the rail, it quickly changes its appearance, bending and deforming its shape, which can lead to lost detections during the tracking process. This deformation and pose variation can change the fish's visual features and make species classification challenging. Added to this the dynamic sea water background can cause problems in object segmentation and the estimated range information. To overcome these challenges, a 3D tracking and segmentation system for stereo video-based monitoring of Rail fish caught with a sea surface background was developed. Table 6 below highlights the utility (functional requirements[1]) for each of the EMI rail computer vision algorithms.

Table 6. -- Descriptions of Rail system computer vision algorithms.

| Functional algorithm process | Description |
|---|---|
| **3.3.1 -Stereo Camera Calibration** | Used for intrinsic and extrinsic calibration of the stereo cameras. |
| **3.3.2 - Rail Detection** | Used for finding the fish/item in the frame. The detector takes as input a list of images, for the stereo camera systems the list is of images acquired from the left camera. |
| **3.3.3 - Rail Tracking and Segmentation** | Used for tracking the detected fish/item from one frame to another. Segmentation refers to subtracting the background from the frame detection to perform further analysis |
| **3.3.4 - Rail Length Estimation** | Used for calculating length estimations using stereo input |
| **3.3.5 - Rail Classification** | Used for classifying the detected fish to species. |

### 3.3.1 Stereo Camera Calibration

By using the stereo camera system, it is possible to acquire range information to facilitate tracking and measurement of catch items. For reliable stereo range estimation, the cameras need to be calibrated. The calibration algorithm calculates the camera matrix using the extrinsic and

[1] The system functional requirements define the basic system behavior; what the system must and must not do.

intrinsic parameters through use of checkerboard calibration pattern before the fishing vessel goes to sea for fishing activities. The resulting calibration file is used as input for stereo tracking and segmentation (Figure 10).


Figure 10. -- An example of the checkerboard used in calibration of the stereo Rail system.

### 3.3.2 Rail Algorithms - Detection

Object detection is used to assess whether an object of interest is in the frame. An object detection model is trained to detect the presence and location of multiple classes of objects. Given an image or a video, the object detection model can then identify which of a known set of objects might be present and provide information about their positions within the image. When images are provided to the model, a list of the objects it detects is outputted. Output consists of a bounding box location that contains each object and a score that indicates the confidence of the detection.

For fish detection at the rail, two different object detectors are applied: SSD (Single Shot MultiBox Detector) (Liu et al., 2016) and YOLOv3 (You Only Look Once) real-time object detector (Redmon 2018). SSD detects objects in images using a single deep neural network, discretizing the output space of bounding boxes into a set of default boxes over different aspect ratios and scales per feature map location. The network generates scores for the presence of each object category in each default box and produces adjustments to the box to better match the object shape (Liu, et al. 2016). The YOLOv3 real-time object detector applies a single neural network to the full image. This network divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the prediction probabilities.

The training of both detectors is based on 10,000 labeled frames from video sequences, in which 80% of the labeled frames are used as the training set and the remaining 20% as the validation set. In its current state the detector is trained to detect the following classes of fish:

- Flatfish *(Pleuronectiformes)*
- Invertebrates *(Invertebrata)*
- Rockfish *(Scorpaenidae)*
- Roundfish (fish other than flatfish or rockfish)
- Sharks *(Selachii)*
- Skates *(Rajiformes)*

Once the model is trained, a list of images can then be run through the detector. In the case of stereo haul images, the list consists of the images acquired from only one of the cameras. A trained weightings file and a configuration (config) file with classifications are also used as input. Weights are the learnable parameters of a machine learning model that control the strength of the connection between two neurons. A weight decides how much influence the input will have on the output.

Once the detection process is completed a csv file is output containing the name of the image, the bounding box coordinates of the detection, the confidence rating, and the detection classification (Fig. 11). An example of the bounding box of an output detection is depicted on an image in the corresponding Figure 12 below. These detection results are then processed through the tracking and segmentation algorithms to determine the number of catch items that are detected.

Rail Detection Output Sample.csv - Notepad

File Edit Format View Help

| Image Filename | X Coordinate | Y Coordinate | Width | Height | Confidence | Detection Class |
|---|---|---|---|---|---|---|
| ... | ... | ... | ... | ... | ... | ... |
| 0049103.jpg | 1029.992676 | 670.139282 | 1276.687256 | 1031.82788 | 0.99984 | Rockfish |
| 0049104.jpg | 1045.981079 | 656.571838 | 1248.254517 | 1012.48773 | 0.99989 | Rockfish |
| 0049105.jpg | 1050.976929 | 617.862427 | 1247.705688 | 994.61243 | 0.99963 | Rockfish |
| ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... |

Figure 11. -- Example of the output detection CSV file (Rockfish with 0.99 certainty)


Figure 12. -- A catch item with the first level detection (Rockfish) bounding box.

### 3.3.3 Rail Algorithms - Tracking and Segmentation

The fish objects that were discovered through the detector need to be tracked to determine unique fish catch events to provide reliable count and discard/retention information for estimation of at-sea discards for the hook and line fisheries. The detector output will indicate that an object of a certain class (such as 'Rockfish' in the example in 3.3.2) was found; however, in the case of continuous input frames it needs to be determined whether the detected object in the current frame is the same as the detected object in the previous frame. In the cases where there are multiple fish detections in the same frame, these need to be tracked and segmented. Added to this complexity is the tracking of highly deformable objects in a noisy environment. When a fish is being pulled up from the sea, the view of the fish changes shape relative to the camera making tracking difficult.

Segmentation and length measurement face challenges from the noisy and dynamic sea surface environments in the background of the image. Traditional segmentation methods using background subtraction in color images cannot work because the abrupt white-water noise and strong shadows can merge to the foreground. The segmentation process consists of three steps: background plane clustering, pixelwise classification and global refinement.

For tracking of highly deformable objects in a noisy environment, the method combines a deep convolutional neural network (D-CNN) image object detector with a Kalman filter in 3D. The tracking method consists of four main steps: object proposals, proposal re-scoring, tracking and learning parameter estimation, and generating weighting constants (Fig. 13). Object proposals come from the 2D object detector (see Section 3.3.2). These 2D object proposals are then projected to 3D objects using foreground segmentation in RGB-D (red, green, blue, depth). The position (X, Y, Z) and size (W, H, D) of the 3D object proposal are then passed to a Kalman filter for re-scoring. Scoring is also called prediction and is the process of generating values based on a trained machine learning model, given some new input data. The values or scores that are created represent predictions of future values. The Kalman filter tracks the objects and predicts the locations in 3D. Tracking is the result of associating the current predicted tracks with the proposals by matching the highest score above a predetermined threshold. If a track cannot be associated with any object proposal, it is assumed to be temporarily missing but if a track is missing for several frames, then that track is stopped and a new track starts for the next object. The parameters used in the proposal rescoring can be systematically learned from the training data. This results in a set of tracks of object detections, or in this case, a set of unique fish detections (Fig. 13).

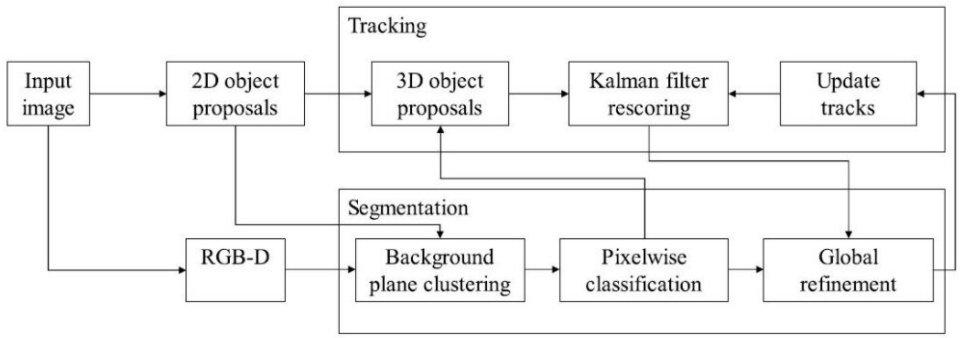
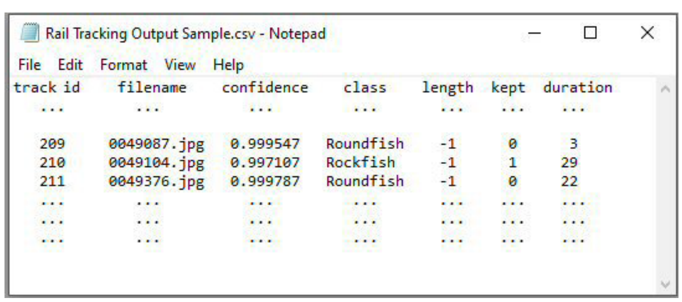


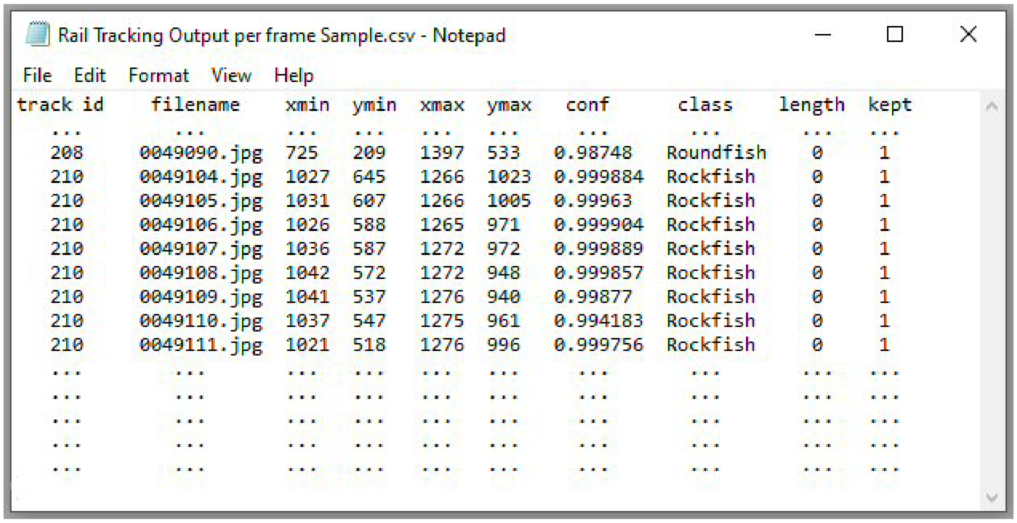

Figure 13. -- Flow chart of the object tracking and segmentation process.

The tracking and segmentation process outlined above results in two csv files. The first is a ‘summary tracks file’ that contains the track information such as its object class, aggregate confidence score, retention/discard information (kept), aggregate length estimation, and track duration (number of frames that make up the track). The second output file is the ‘frames’ csv file which contains the detailed information for each frame. Using the same set of data and detection output from the example in Section 3.3.2, from the earlier example, image detection is part of track 210 and consists of 29 consecutive images (see Fig. 14: track csv file and Fig. 15: frame csv file). The image sequence in Figure 16 illustrates the movement and subsequent tracking of a detected Rockfish.


Rail Tracking Output Sample.csv - Notepad
File Edit Format View Help

| track id | filename | confidence | class | length | kept | duration |
|---|---|---|---|---|---|---|
| ... | ... | ... | ... | ... | ... | ... |
| 209 | 0049087.jpg | 0.999547 | Roundfish | -1 | 0 | 3 |
| 210 | 0049104.jpg | 0.997107 | Rockfish | -1 | 1 | 29 |
| 211 | 0049376.jpg | 0.999787 | Roundfish | -1 | 0 | 22 |
| ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... |

Figure 14. -- Example of tracking csv file showing one record per track. Length estimation is not available at this time and is marked as -1.





| track id | filename | xmin | ymin | xmax | ymax | conf | class | length | kept |
|---|---|---|---|---|---|---|---|---|---|
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 208 | 0049090.jpg | 725 | 209 | 1397 | 533 | 0.98748 | Roundfish | 0 | 1 |
| 210 | 0049104.jpg | 1027 | 645 | 1266 | 1023 | 0.999884 | Rockfish | 0 | 1 |
| 210 | 0049105.jpg | 1031 | 607 | 1266 | 1005 | 0.99963 | Rockfish | 0 | 1 |
| 210 | 0049106.jpg | 1026 | 588 | 1265 | 971 | 0.999904 | Rockfish | 0 | 1 |
| 210 | 0049107.jpg | 1036 | 587 | 1272 | 972 | 0.999889 | Rockfish | 0 | 1 |
| 210 | 0049108.jpg | 1042 | 572 | 1272 | 948 | 0.999857 | Rockfish | 0 | 1 |
| 210 | 0049109.jpg | 1041 | 537 | 1276 | 940 | 0.99877 | Rockfish | 0 | 1 |
| 210 | 0049110.jpg | 1037 | 547 | 1275 | 961 | 0.994183 | Rockfish | 0 | 1 |
| 210 | 0049111.jpg | 1021 | 518 | 1276 | 996 | 0.999756 | Rockfish | 0 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

Figure 15. -- Example of tracking csv file showing one record per frame.

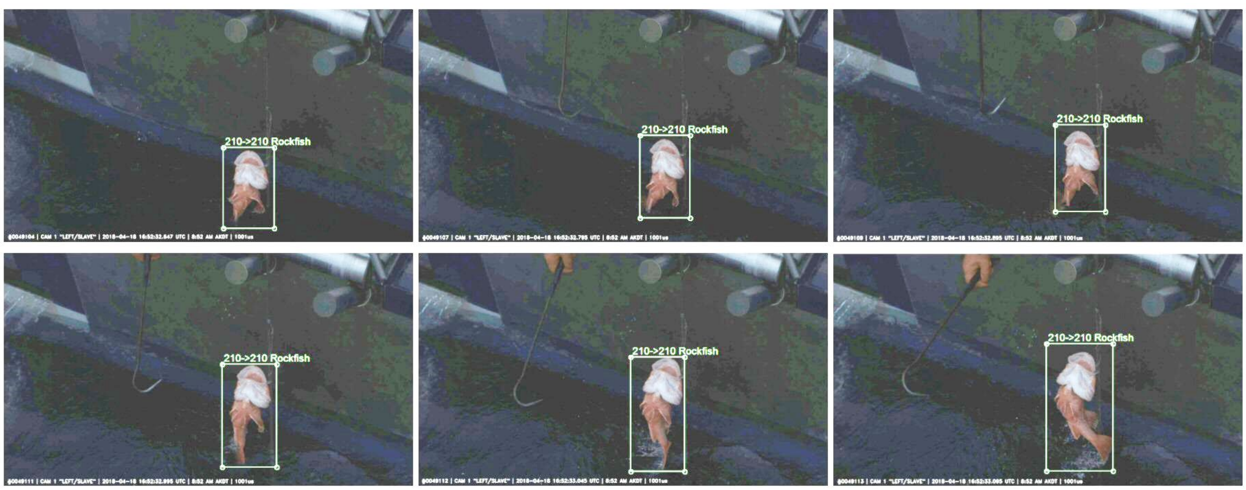




Figure 16. -- A series of images showing the tracking movement of a detected Rockfish. Track numbers and classification are indicated above the bounding box in each frame.

### 3.3.4 Rail Algorithms - Length Estimation

The development of length estimation algorithms further builds upon the 3D projections created by the tracking and segmentation algorithms. Two methods to estimate fish length were evaluated: 3D midline calculation and estimation of lengths from multiple frames. To measure the fish length more accurately when the fish body is curved, the 3D midline of the fish body is found based on the back-projected point cloud of the fish body in 3D (Fig. 17). The fish body is first separated equally into $H$ bins along the major axis given by performing principal component analysis (PCA). For each bin, the geometric center of the point cloud is found and then the center points of each bin are connected to form the 3D midline and measure the midline length of the fish. Currently, evaluations of the 3D midline calculation are ongoing, and efforts are expanding to determine how to achieve these same results using 2D monoscopic camera images.

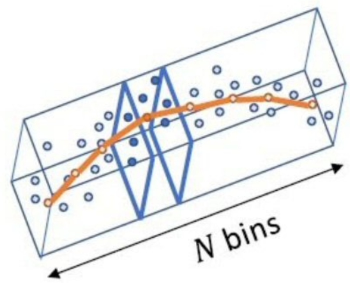



Figure 17. -- Schematic of the midline calculation length estimation method. The 3D midline is acquired by connecting the center point of each bin along the major axis.

3.3.5 Rail Algorithms - Classification

Recognizing fish species from a video image when the fish is captured live (is moving in 3D space) against a sea surface background is challenging due to the deformation of fish shape, self-occlusion of body parts, and similar texture between different fish classes (Huang 2019). Prior work on fish species classification usually relies on hand-crafted features, saliency part association or codebook learning, but most of these are for use in controlled or stable environments where fish are in similar poses or shapes (Huang 2019). In the case of fish at the vessel's rail, fish can change shape and orientation freely, resulting in dramatically different visual features and self-occlusions. Additionally, many species of fish share similar colors and textures, making discriminant features difficult to identify. Varying lighting conditions also makes it difficult to distinguish similar species (see Fig. 18 for examples where the same species of fish can look vastly different from one image to another).

Feature extraction and classification algorithms for use with Rail system data need to be robust to account for fish in any orientation and in different poses while at the same time being able to find the difference between similar looking species. To address these issues, a fine-grained image classification method based on a deep convolutional neural network (CNN) trained by an innovative metric learning scheme with a temporal constraint was developed (Huang et al. 2019).

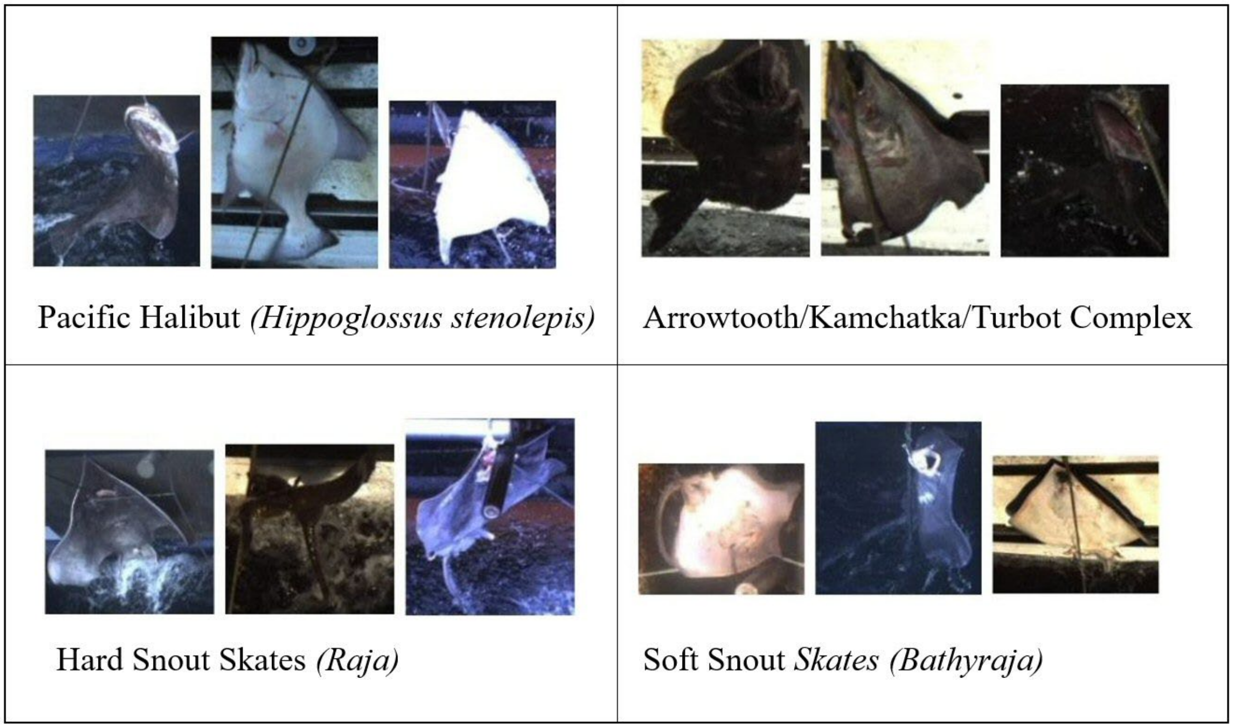



Figure 18. -- Images showing variations in features between fish of the same species.

The fish species classification model takes the tracked fish from the video as input (the tracking output from Section 3.3.3) and uses the temporal information between frames to extract useful features. By applying deep metric learning with a temporal constraint, the model is forced to learn the closeness between temporal neighbor frames. Based on the temporal constraint, two classification models (representative feature classifier and semantically-decoupled temporal attention) were investigated.

The representative feature classifier discriminatively learns the representative features of each class (species) and uses them for feature aggregation during prediction. The semantically-decoupled temporal attention model learns which multiple attention groups to focus on for different feature dimensions. A diversity constraint is applied to make different attention groups focus on different frames and feature dimensions. The experimental results show that this approach outperforms the conventional softmax classification on the current rail-fishing dataset (Huang 2019).

This proposed method was evaluated on an initial dataset consisting of cropped image frames of 17 classes. The dataset was captured on cameras from different fishing vessels (see Section 3.1). The bounding boxes of the fish and the fish species were labeled as part of the annotation process (see Section 3.2). The training set consisted of 135,150 images of fish in 17 classes. These 17 classes are subclasses of the initially detected classes in Section 3.3.2 and are listed in Table 7.

Table 7. -- List of Rail species classes, class types, and numbers of images in the training dataset.

| **Species class** | **Type class** | **Images Trained** |
|---|---|---|
| Anemones *(Actiniaria)* | Invertebrates *(Invertebrata)* | 667 |
| Flathead sole *(Hippoglossoides elassodon)* | Flatfish *(Pleuronectiformes)* | 100 |
| Grenadier *(Macrourinae)* | Roundfish (fish other than flatfish or rockfish) | 6946 |
| Hard snout skate *(Raja)* | Skate *(Rajiformes)* | 2225 |
| Kamchatka/arrowtooth/turbot complex | Flatfish *(Pleuronectiformes)* | 1221 |
| Octopus *(Octopodiformes)* | Invertebrates | 169 |
| Pacific cod *(Gadus macrocephalus)* | Roundfish (fish other than flatfish or rockfish) | 2152 |
| Pacific halibut *(Hippoglossus stenolepis)* | Flatfish | 41311 |
| Redbanded rockfish *(Sebastes babcocki)* | Rockfish *(Scorpaenidae)* | 1062 |
| Sablefish *(Anoplopoma fimbria)* | Roundfish | 19374 |
| Shortraker, rougheye, blackspotted rockfish, unidentified (Sebates borealis, aleutianus, melanostictus) | Rockfish | 1348 |
| Soft snout skate *(Bathyraja)* | Skate | 1080 |
| Spiny dogfish (*Squalus suckleyi)* | Shark *(Selachii)* | 51371 |
| Spotted ratfish *(Hydrolagus colliei)* | Shark | 193 |
| Thornyhead unidentified *(Sebastolobus)* | Rockfish | 4340 |
| Yellow Irish lord *(Hemitripterus bolini)* | Roundfish | 58 |
| Yelloweye rockfish *(Sebastes ruberrimus)* | Rockfish | 1533 |

Training of the dataset is ongoing; however, in its current configuration, running the rail classifier updates the tracking csv with these classes, the bounding box coordinates, and confidence level. Following on the examples from 3.3.2 and 3.3.3, Figure 19 below shows the updated extract of the tracking csv file. This last step of the Rail algorithms shows the frame where the fish was detected and tracked and the final classification to thornyhead (Fig. 20). This combination of the output allows the estimation of the number of each species needed for catch accounting.

Rail Classfication Output Sample.csv - Notepad

File Edit Format View Help

| track id | filename | xmin | ymin | xmax | ymax | conf | class | length | kept |
|---|---|---|---|---|---|---|---|---|---|
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 210 | 0049104.jpg | 1027 | 645 | 1266 | 1023 | 0.999884 | Thornyhead Unidentified | 0 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

Figure 19. -- Sample of species identification output per track csv file showing the track ID, image filename, and species identification (class).

While positive results have been achieved with the current model, research into the classification for the Rail system data is ongoing. Current investigation and research include extending and streamlining the classification hierarchy to increase the number of classes as well as applying the species classifier developed for use with controlled environments (Section 4.3) to the Rail.


Figure 20. -- Sample image of rockfish in Figure 12 classified to Thornyhead.

### 3.4 EMI Rail Algorithm Review and Results

The success factors for the development and implementation of algorithms for automated catch accounting data for fish caught at the rail are twofold:

- Accuracy. The algorithms need to be able to make predictions with a high degree of confidence while achieving as few false positives as possible.
  Runtime processing performance. The time it takes to process data automatically should be as fast as or faster than the time it takes to complete the same task by a human reviewer.

### 3.4.1 Rail Algorithms - Accuracy Results

Research results for detection, tracking, classification, and length measurements for fish at the rail have been positive. Testing and development of the algorithms have primarily been conducted on controlled datasets (see Section 3.3.5). The classes included are of species with enough image data available to train the algorithms, they do not include all species encountered particularly rare species, and some classes are species groupings that mirror visual identification protocols for observers for catch accounting collections. The following results of classifications and tracking for fish at the rail have been achieved using these trained datasets. Based on this testing, detection, tracking, and classification accuracy is high for four species; Pacific halibut, sablefish, spiny dogfish, and grenadier, mostly giant grenadier (Table 8) which is expected since there are more images of those species of fish in the training and testing datasets.

For each species class in the dataset several pre-identified tests were selected. These images were then run through the classifier to achieve the per-image accuracy. Accuracy was measured by the number of correct classifications that were achieved with a high confidence level. False classifications and missed classifications were counted against the accuracy score. For evaluating the accuracy of the tracker, the same dataset was run through the tracker and the resulting tracks were run through the classifier.

Fish species classifications were split into a hierarchy-grouped into a more generalized set of classes including sharks, skates, flatfish, rockfish, roundfish, and invertebrates, then classified to species within these groupings. The classes included are of species with enough image data available to train the algorithms; they do not include all species encountered, particularly rare species. One species class includes the rougheye, blackspot, and shortraker rockfish species, which are grouped to mirror visual identification protocols for fisheries observers.

Table 8. -- Classification and tracking algorithm results for a controlled dataset showing the species, number of images, and accuracy at the image and track levels.

| Species | Number of images tested | Accuracy % per-image | per-track |
|---|---|---|---|
| Anemones *(Actiniaria)* | 20 | 70.3 % | 80 % |
| Flathead sole *(Hippoglossoides elassodon)* | 6 | 68 % | 66.7 % |
| Grenadier | 147 | 90.9 % | 93.2 % |
| Hard snout skate *(Raja)* | 25 | 88.2 % | 91.7 % |
| Kamchatka/arrowtooth/turbot complex | 37 | 79.7 % | 83.3 % |
| Octopus *(Octopodiformes)* | 4 | 82.1 % | 100 % |
| Pacific cod *(Gadus macrocephalus)* | 40 | 83.9 % | 90 % |
| Pacific halibut *(Hippoglossus stenolepis)* | 686 | 96.8 % | 98.3 % |
| Redbanded rockfish *(Sebastes babcocki)* | 23 | 80.2 % | 90.9 % |
| Sablefish *(Anoplopoma fimbria)* | 420 | 95.5 % | 96.7 % |
| Shortraker, rougheye, blackspotted rockfish unidentified (Sebates borealis, aleutianus, melanostictus) | 36 | 65 % | 72.2 % |
| Soft snout skate *(Bathyraja)* | 22 | 68.7 % | 54.5 % |
| Spiny dogfish (*Squalus suckleyi)* | 923 | 97.9 % | 98 % |
| Spotted ratfish *(Hydrolagus colliei)* | 5 | 55.2 % | 50 % |
| Thornyhead unidentified *(Sebastolobus)* | 88 | 86.7 % | 90.9 % |
| Yellow Irish lord *(Hemilepidotus jordani)* | 2 | 44.8 % | 0 % |
| Yelloweye rockfish *(Sebastes ruberrimus)* | 10 | 68 % | 80 % |
| **Complete dataset** | **2494** | **77.8 %** | **78.6 %** |

To further assess the accuracy of the algorithms in a real world scenario, testing is ongoing using data from the EMI Rail System initially used for acquiring training images. Additional fixed gear haul data collected via traditional human reviewed EM systems is also being tested. Data analysis of data from these tests will form a better understanding of how the Rail algorithms will perform in an operational environment. Measuring the length estimation accuracy for fish caught at the rail has been difficult to measure due to the lack of ground-truth observations to compare with the test results. This testing and retraining of Rail algorithms against ground-truth observations based on standard EM video review is a continuing effort as more one to one data are collected and annotated.

### 3.4.2 Rail Algorithms - Processing Results

Initial analysis and testing of the EMI Rail algorithms have proved positive; however, the time it takes to process and analyze the results to produce data suitable for use in catch accounting is equally important. Any cost-benefit advantages of implementing automated solutions could be minimal or negated if automation processes take longer to produce results than human review, then the. Initial processing results have shown the automated process to be capable of analyzing data close to real time human analysis, and algorithms can be automated to run 24 hours a day, so a time savings is possible. The current review program can take up to two weeks before data are available to FMA.

To obtain species specific catch data, imagery from a haul needs to be run through the detection algorithm, followed by the tracking and segmentation algorithms, and then the classification algorithm. The detection and tracking algorithms make use of the computer's central processing units (CPU) processing power while classification algorithms require a dedicated graphics processing unit (GPU). For performance testing, imagery from select hauls was processed end-to-end using a high-powered CPU and GPU computer. The following results in Table 9 were attained using this machine as a benchmark.

Table 9. -- Algorithm processing benchmark results for processing 9,000 images, representing approximately 30 minutes of catch retrieval time.

| Algorithm | Number of images | Images represent real world catch retrieval time | Average time to process |
|---|---|---|---|
| Detection | 9000 | 30 min | 23 min |
| Tracking and Segmentation (2D) | 9000 | 30 min | 10 secs |
| Classification | 9000 | 30 min | 17 min |
| Total analysis runtime | 9000 | 30 min | 40 min |

Performance results will vary depending on hardware capabilities and the number of detections in the imagery. For example, in one set of images there may be more hauling activity than in another set thus requiring more detections to be processed; this makes it difficult to fully gauge the time it takes to analyze a complete haul. Only by running further hauls will a baseline be established.

The results of this analysis then need to be human reviewed for quality assurance and accuracy measurement. This quality assurance and final reporting is in its preliminary stages with further results yet to be produced. By combining the results of analysis with data extracted from the EM Rail system haul log data from when the haul imagery was acquired (vessel information, GPS coordinates, haul start/end time etc.) a holistic view of the catch can be obtained.

### 3.5 EMI Rail Discussion and Operational Readiness

The long-term goal of the EMI Rail research is to implement a real-time EM system, acquiring and analyzing the image data at the same time. This analysis can then be transmitted to and integrated into the Observer Program data reporting stream. While this real-time analysis is not possible yet due to current technical and development limitations, it remains one of the research programs goals for catch reporting for fixed gear vessels. In the meantime, while the EMI Rail stereo system and algorithms continue to be refined and developed with this integrated solution in mind, the EMI Rail system and EMI Rail algorithms can now be applied to other research and development efforts.

Implementation strategies will need to be defined for both the EMI Rail algorithms and EMI Rail camera system before any production-level implementation can take place. These strategies should define how algorithms will be run in the field and how they will scale across fisheries as well as the operational (business) processes related to integration with existing EM and observer data streams.

### 3.5.1 Rail Algorithms - Implementation and Collaboration

In the absence of real-time processing of the algorithms on board vessels, the standalone algorithms can be applied to the current review data cycle to augment the human review process. The current reviewer software isolates hauling periods but does not isolate each fish caught on the line, so reviewers currently scan the entire haul period for catch events of fish caught. By processing hauls through the algorithms, catch events can be detected before the human review process starts, the human reviewer will then only need to review the automated detections instead of scanning the entire video record for catch events (detections), potentially decreasing review time. Integrating the EMI Rail algorithms into the current data review process would have the following benefits:

- Increased productivity of human reviewers by reducing their workload by finding and labeling detections.
- Increased number of video haul catch reports since more video could be reviewed (for a given amount of time).
- Every haul can be 'pre-processed' through the algorithms whereas currently not all hauls are reviewed.
- Running of the algorithm application on hauls can occur concurrently on a 24hr per day cycle limited only by computing power.
- Video review using standard (human-based) methods can be better prioritized. If certain events are automatically detected in the haul, priority can be given for standard-review of that specific haul
- If more events are detected in certain hauls over others, it can be predetermined which hauls are more likely to take longer for standard review than others.

To achieve this sort of integration, the algorithms would need to produce detections and events like that of the human reviewer. Detection, tracking (counts), and species classification algorithms are currently available; however, these algorithms would need to be expanded to include the following:

- Fate Disposition: Determination of whether catch items were retained, discarded by the crew, dropped off the gear.

- Crew Activity (monitoring for illegal on deck actions).
- Depredation by marine mammals and other predators.
- Gear Performance (gear entanglement, parted lines, loss of gear).
- Handling of prohibited species (halibut viability and release method).

For these analysis activities, training, input, and annotation requirements will need to be defined with each function being a separate research target. Therefore, annotation software is currently being updated to include depredation, fish retention, halibut release method, halibut injury, the location of target (on the line, in water, in air), sex, and gear type information.

The current detection, tracking, and classification algorithms will also need to perform at an acceptable success rate before they can be integrated into production. Additionally, the EM systems currently in use by the fleet (to collect imagery for standard review) are monocular camera systems; to obtain length estimation for these types of systems, the algorithms will need to be adjusted. Adaptation of Rail algorithms for use with monocular cameras is currently being researched. Testing activities with production data will be crucial to achieve implementable results.

While the EMI Rail algorithms are still under development, the current goal is to package and release them under the open-source model to encourage open collaboration. This package would also require applications and guides for algorithm model retraining and annotation protocols.

### 3.5.2 Rail Camera System - Implementation and Collaboration

Lowering the costs associated with the collection, transfer, storage, and analysis of event-based image data allows for greater fisheries monitoring rates over a wider range of vessel types and sizes, particularly those vessels where it is impractical to place an observer. By leveraging the latest developments in computer vision, cost-effective and timely extraction of scientific data from images will provide greater certainty for resource management and support sustainable fishing practices.

In developing the EM Rail system, the focus has been on designing and deploying durable stereo camera systems for data collection. Technological advancements in GPU processing for running machine learning algorithms have increased since the start of the project, with newer GPU components becoming available with greater processing abilities while requiring less power to operate. Integration of these low powered GPUs into the EM Rail system would allow for integration of algorithms to reduce data storage, report image quality issues in real time, speed up review times by identifying fish catch events in advance, and could even at some point lead to the reality of real-time catch accounting data analysis. Currently, the fan-less waterproof PCs deployed as part of the EM Rail system do not have this ability and therefore real-time cloud computing would not be a suitable solution; vessels that will use these systems will not have access to continual dedicated internet connectivity in the near future. Additional research needs to be conducted to develop the best practices for this integration, balancing and benchmarking the tradeoffs of performance against component costs.

Before the Rail camera system can be integrated into the current pool of monitoring options available, additional hardware and software development will be required. The current EM Rail System relies on use of a specific type of stereo machine vision cameras that can acquire the imagery required for fish length estimation based on stereo imagery. The data collected via this system has been invaluable in the development of the length estimation algorithm for the fish caught at the rail. However, in its current form the EM Rail system is impractical to deploy at a scale appropriate for fisheries monitoring. Research is being conducted to make use of less costly components with this in mind, switching from expensive machine vision cameras to IP and Raspberry Pi cameras.

Traditional EM systems used in fisheries for compliance monitoring make use of single view IP cameras, capturing lower resolution images at lower frame rates than the EM Rail system. Processing the data from these systems through the algorithms developed for the Rail system gave us a better understanding of implementation trade-offs for the EM Rail system. If these trade-offs are acceptable, then the EM Rail system can switch to an IP camera strategy. Likewise, length estimation algorithms for use with imagery collected using monocular camera systems are currently being researched (using existing stereo camera imagery). If the existing length estimation algorithms can be adapted for use with monocular camera images, then the EM

Rail system can switch to a single camera strategy that more closely resembles the current EM monitoring systems.

Research should continue into the development of robust haul start and end sensors. Testing and implementing an automated computer vision region-of-interest trigger would prove to be a better trigger than the electrical sensors currently deployed.

Improvements in communication and data transmission components should also be investigated. When real-time processing becomes a reality, there will be a need to transmit the resulting information from the vessel to EM data managers as quickly as possible to allow for use in in-season quota monitoring. Application development will be required to integrate these results into existing data streams.

### 3.6 EMI Rail Summary

In summary, the goal of this research was to develop algorithms for automating the detection, tracking, estimated length, and species identification of fish caught on fixed gear vessels. This is a challenging task due to the complexity of the imagery from the algorithm's perspective; deformed and bending fish makes identification and length estimation difficult due to the multitude of variations that can occur. The lighting conditions and background with the fish silhouetted against is also complex and changes throughout the collections.

To accomplish this task, dedicated EMI rail systems were iteratively built and deployed to collect needed training imagery. The camera system deployments had to overcome environmental challenges including power restrictions and electronic robustness. Many challenges were overcome, and successful data collection deployments were achieved. Training imagery consists of multiple species types, multiple angles of the catch, and stereo imagery for length estimation. A considerable proportion of the training imagery was annotated for detection, tracking and identification purposes.

Algorithms were developed and have achieved high accuracy in detecting, tracking, and identifying the predominant Alaskan species. The track identification accuracy of: Pacific halibut was 98.3%; spiny dogfish 98%; sablefish 96.7%; and grenadier, 93.2%. Overall accuracy of less common species can improve with additional training imagery. Stereo-derived length measurements were achieved; however, further development will be required to refine length estimation algorithms.

Since the project's initiation, deployments of EMI rail systems in the field have become standard for image acquisition, with custom strategies in continued development for precise fishing activity monitoring and post-processing of imagery. Current commercial EM systems designed to collect video data for later human review are simple in nature compared to the stereo camera systems built as part of this research stream. The developed algorithms are being investigated for use in commercial EM imagery review as an automated 'pre-reviewing' of the data collected by these systems, but current predictions are not robust enough to be useful for optimizing human review time at this time. Images from the commercial systems need to be incorporated into the currently available Rail algorithms to improve predictions, and this is the step where commercial vendors can take our base open-source algorithms and customize it as desired for their proprietary review product. To achieve this, priority focus will be given to further testing the existing algorithms and establishing training dataset baselines to complement the release of open-source algorithms. These baselines will establish the retraining requirements of the detection, tracking, and classification datasets and subsequent annotation labeling will be conducted on this production data. Further algorithm development to address additional requirements such as automating the detection of fate disposition and depredation by marine mammals has also been initiated. While the EMI Rail algorithms are still under development, the goal is to package and release these in their current form under the open-source model, with descriptions of successful annotation and training techniques, to encourage open collaboration. Direct integration of algorithms into commercial systems is dependent on the algorithms being made publicly available

For an integrated real-time automated EM system capable of species identification, length measurement, and at-sea transmission of data to be ready for deployment into the Alaska fixed-gear fisheries our team plans to 1) adapt existing (or create new) single camera algorithms for species identification and length measurements 2) data transmission protocol development needs to be completed, and 3) full-system field testing needs to be completed inclusive of in situ data verification studies.

# 4. SPECIES DETECTION, IDENTIFICATION, AND MEASUREMENT -- CONTROLLED ENVIRONMENTS

Many of the difficulties encountered in the rail environment can be reduced for other electronic monitoring applications by controlling lighting and fish orientation to the camera. Variation in the imaging environment, particularly lighting and background, complicates image analysis for species classification by creating uncertainties as to whether different colors and patterns are due to fish characteristics or lighting differences. Differences in fish posture and orientation to the camera can be similarly problematic. Image analysis for fish classification or measurement were simplified by enclosing the imaging area and limiting lighting to consistent artificial sources, as well as reducing pose and orientation variability by imaging the subject against a flat, monochrome surface. Applications where fish can be counted, identified, and measured by being passed through an imaging enclosure include monitoring discards where retained catch components are accounted for during delivery or where measurements of particular species need to be made before they are returned to the ocean.

Such EM applications depend on supplying sufficient images to train classification models to recognize the relevant fish species. Species classification imagery was collected aboard the AFSC Gulf of Alaska and Aleutian Islands trawl surveys in 2015, 2016, and 2019, providing images of a wide range of groundfish and invertebrate species for training classifier algorithms. The wider the variety of species and the more images of those species that can be used to train the algorithm, the better it will perform. Figure 21 below shows the iterative development cycle for species classification in a controlled environment.

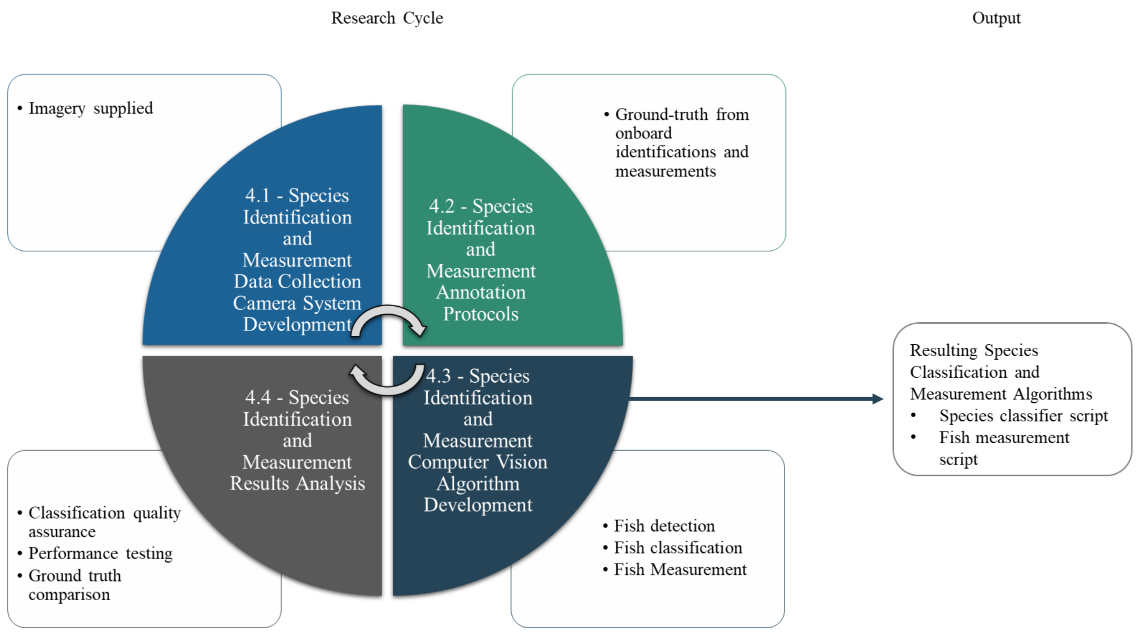
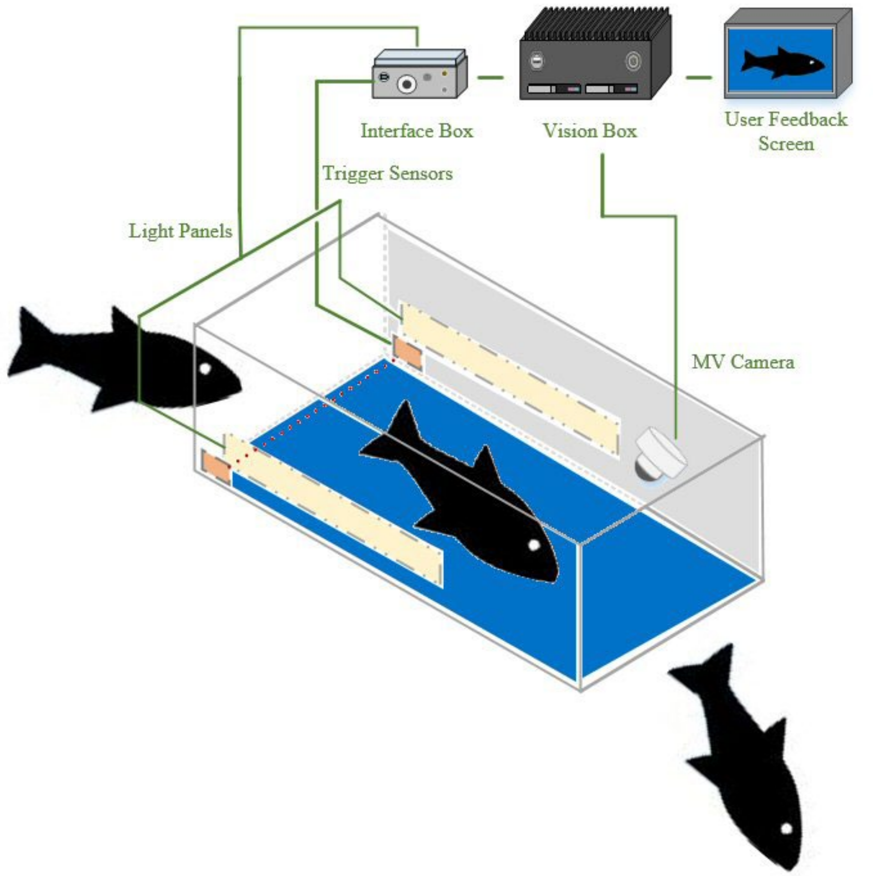


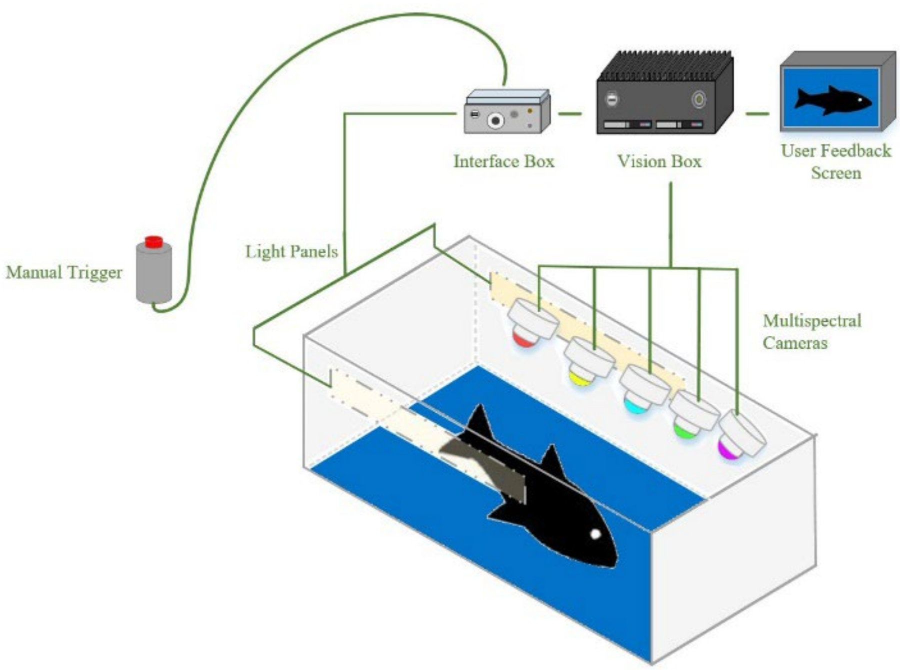

Figure 21. -- Schematic of the development lifecycle for species classification and measurement in controlled environments.

## 4.1 Controlled Environments - Data Collection Camera Systems

To collect imagery for the number of distinct species needed to train the classification algorithm, initial collections of these images occurred during the AFSC Gulf of Alaska and Aleutian Islands trawl surveys. Three camera systems were used to collect the needed imagery. In 2015, fish were slid through a sloped enclosure (chute) and single images were acquired of those fish using machine vision cameras (MVChute). To assess the potential of separating light frequencies, the 2016 system was a Multispectral System, which used multiple cameras, each sensitive only to a narrow light frequency range, and LED lighting providing those frequencies. Those systems were configured as a 'photo booth', where each fish was placed in a box, posed, and a photo was manually triggered. The 2019 system (VideoChute) used an Internet Protocol (IP) camera to capture a video sequence of each fish as they passed through a chute from one end to the other, allowing for the rapid collection of live fish imagery. These systems are collectively called enclosed camera systems (Fig. 22).

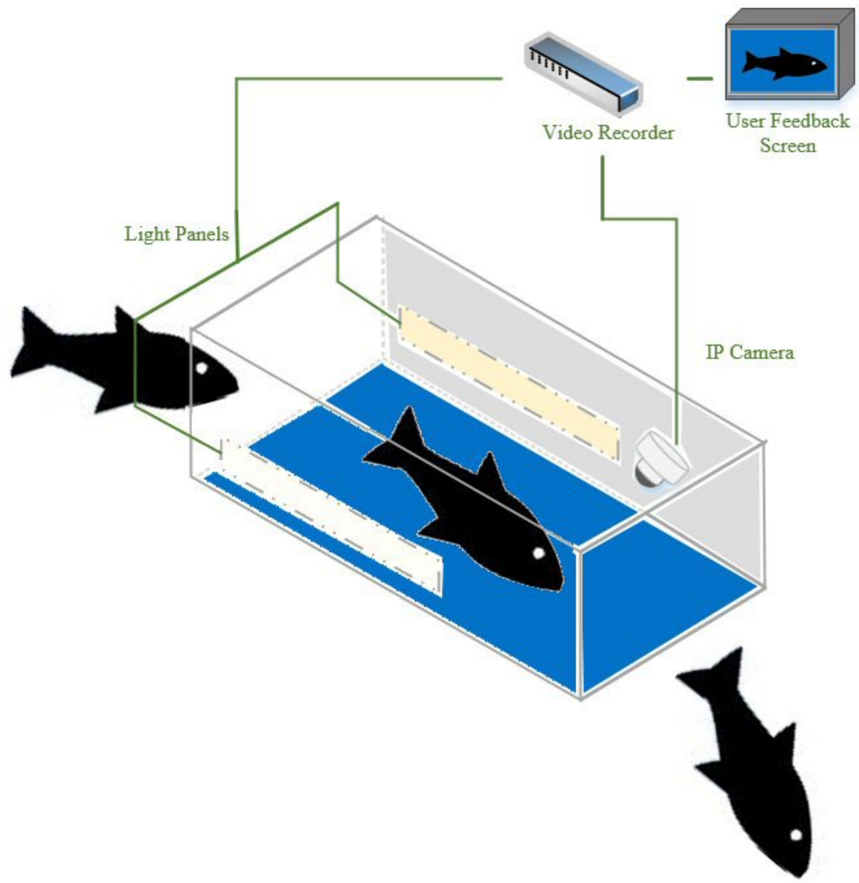



Chute style system deployed in 2015 - Fish are sent into one end and out the other end, sliding over a sloped surface – (MVChute)




Multispectral imaging system deployed in 2016 – Fish is placed one at a time in an enclosed box – (Multispectral System)

Figure 22. -- Schematic depicting the three enclosed camera systems, Machine Vision (MVChute) Multispectral System, IP video (VideoChute).




Chute style system deployed in 2019- Fish are sent into one end and out the other end, sliding over a sloped surface – (VideoChute)

Figure 22. Continued. --Schematic depicting the three camera chute systems (Multispectral, Machine Vision, IP.)

Table 10. – Descriptions of camera chute system design requirements.

| Design requirement | Description |
|---|---|
| **4.1.1 – Camera Requirements** | A range of MV and IP cameras, lenses and filters were used. Wide angle lenses were necessary, due to the proximity of cameras to fish. Multispectral system used filters to limit sensitivity to narrow frequency slots The enclosures need to exclude external light, allowing artificial lights to provide consistent imagery. Chutes required bright lighting to allow fast exposures of moving fish. |
| **4.1.2 – Autonomous Collection** | The multispectral photo booth used manual triggering, while the camera chutes were triggered automatically. |

| | |
|---|---|
| **4.1.3 – Environmental, Size and Maneuverability Limitations** | Deployments aboard fishing vessels put a premium on camera systems occupying minimal volumes. |
| **4.1.4 – Deployments** | Deployments during AFSC trawl surveys provided a wide range of species under more controlled conditions than during commercial fishing. |

### 4.1.1 Camera Systems

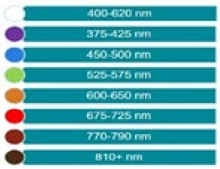
Several requirements led to the design, development and deployment of these collection systems and are outlined in Table 10. Machine vision (MV) cameras, like those used in the stereo EMI Rail system (see Section 3.1.1), were used in the initial chute designs (Fig. 23). Enclosures excluded external light and light arrays provided very bright white lighting, allowing rapid exposures of rapidly moving fish. The chutes' floors were rigid plastic sheets, providing a contrasting background for imaging. With the acquisition software already having been developed for the cameras on the Rail system, the same software was incorporated into the camera chute design to acquire imagery of fish as they were passed through the chute. The software was modified to capture the single, triggered images taken by those systems. For further details on the evolution of these chute designs see Section 5 which describes their deployment as halibut bycatch chutes.


*Machine Vision Camera Chute System evolution*

Figure 23. -- MV camera chute (MVChute) systems design, showing the progressive design evolution (upper three panels).


Figure 24. -- An example of an acquired image from the MVChute.

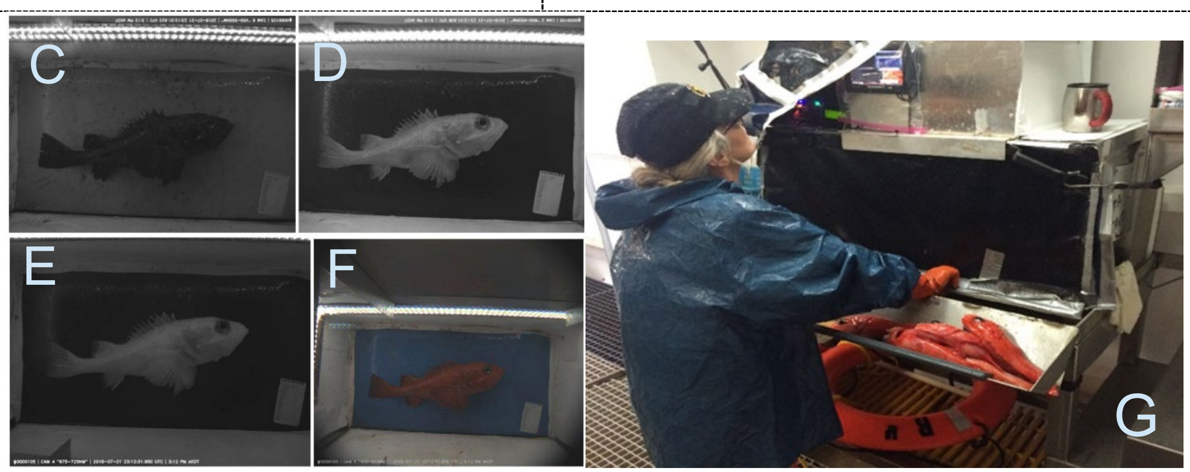
To explore the potential of combinations of specific light frequencies to improve fish species identifications, a specialized system was built with seven cameras, each of which was equipped with a slot filter limiting it to a narrow light frequency band. There was also one unfiltered camera. A dedicated lighting source was added to provide enough light power at specific frequencies sampled, as well as the broad band camera (Fig. 25). Fish placed in the enclosure were stationary during imaging, allowing longer exposures than the two camera chutes. This Multispectral System was deployed during the AFSC 2016 trawl survey of the Aleutian Islands area. Multiple fish species images were collected and provided along with identification data to develop and test species identification algorithms using the multi-frequency data. Results of that development and testing indicated that the multi-spectral information only marginally improved species identification accuracy. Given the complexity of the multispectral chute, data, and analysis, this development direction was not pursued further. However, the data collected from the multispectral camera chute was beneficial for seabird experiments that were conducted (see Section 8: Seabird Species Identification Experiment). Figure 25 shows the multispectral camera setup.




*Lens filter values applied to the machine-vision cameras*

*Multispectral camera array - 8 cameras with different filters mounted above surface*


*Multispectral system enclosure – LED light strips were mounted around the upper sides (not visible)*




*Example of rockfish images captured with the Multispectral system and onboard operation*

Figure 25. -- Multispectral Imaging System showing camera array (A), imaging booth (B), and examples of acquired images using different frequency bands (C-F) and onboard operation (G).

As the algorithms evolved, tests with reduced resolution determined that image resolution requirements (megapixel size) could be relaxed from those initially conceived (from 2.8 to 0.7

megapixels). This allowed for the implementation of IP cameras instead of the dedicated machine vision cameras used in previous development stages. This in turn allowed for application development to shift from dedicated purpose-built camera systems to standard, off-the-shelf IP cameras. A video surveillance-type IP camera was implemented into a chute design for 2019 species image collections during the AFSC survey (Fig. 26). This system recorded video as opposed to still image frames. Multiple images also provide more opportunities for accurate classifications and measurements. This system proved to be the simplest in terms of usability, durability, and ease of use.

*IP camera chute build and finished design*


*Example of a salmon image captured with the IP camera chute*

Figure 26. -- Portable IP camera chute system design showing a side view of the uncovered chute (upper left), the full (covered) chute (upper right), and an example of an acquired image (lower).

System usability is dependent on the environment in which a camera chute system is deployed. In the case of the Multispectral System, this system would typically be used in a controlled environment with no power or size restraints with a full monitor, keyboard and mouse attached. This is not the case for the other chute systems, where they could be used in the field where time and physical space will be limited. These chutes require autonomous image acquisition and must conform to certain design restrictions such as power and size limitations.

### 4.1.2 Autonomous Collection

These camera systems required devices to trigger the cameras to acquire images. Because the operator of the Multispectral System posed each fish, a manual trigger was used. The MVChute was triggered by a light beam interruption sensor near the chute's exit. However, the durability and consistency of these sensors was insufficient, producing false triggers, missed triggers and system failures. For the Video Chute, autonomous collection was triggered by built-in motion detection capabilities of the IP camera, with time buffers before and after motion detections to assure all fish passage events were fully recorded.

### 4.1.3 Environmental, Size and Maneuverability Limitations

The size of the camera chutes needed to conform to the specific environment (vessel) where they are being deployed and space is often limited aboard fishing vessels. To reduce enclosure volume, cameras were placed above the exit end of the chutes, providing an oblique view, and requiring less enclosed space than mounting the camera directly above the center of the chute. The MVChute would typically be affixed to the side of the vessel, with fish passing through it sent back into the ocean. This required the physical design of this chute having to support environmental issues such as water proofing components and enduring vibration issues. This chute system was also not ideal for spontaneous image collection as the setup requires secure installation. The VideoChute, however, was smaller and lighter, and easily moved around. Since there were no sensors on this chute, it was also easier to maintain.

For the reasons outlined above it was determined that the VideoChute was the best chute design for quick autonomous species classification data collection. However, the lessons learned from the Multispectral System and MVChute were still invaluable when developing and deploying all camera chute systems.

#### 4.1.4 Deployments

Species classification imagery and associated length and species data were collected aboard the AFSC Gulf of Alaska and Aleutian Islands trawl surveys in 2015, 2016, and 2019 (Table 11). For all these deployments, an onboard scientist was responsible for selecting fish after the survey sampling was complete to put through the chute. Fish were put through the chute in species groups, as they had already been sorted. The scientists recorded which species were put through in which order and the lengths of selected fish.

Table 11. -- Survey deployments by year with numbers of species and images or video clips collected.

| Year | Deployments | System | Notes |
|---|---|---|---|
| 2015 | FV *Alaska Provider* | MV camera chute | 58 species, 18,815 total specimen images |
| 2016 | FV *Alaska Provider* | Multispectral camera system | 145 species, 6,740 total specimen images |
| 2019 | FV *Ocean Explorer* | IP camera chute | 88 species, 888 total specimen video clips |

### 4.2 Controlled Environments - Annotation

Annotation requirements for the species classification algorithm were quite different from those for the Rail system. There were no requirements for detection or tracking, as each fish only had a single image. Annotation for species classification for controlled environment images only required fish species identifications and these were recorded by scientists during the survey. As survey image collections used fish that had been pre-sorted for survey purposes, fish were put through the imaging system in same-species batches, with species recorded for each batch.

## 4.3 Controlled Environments - Algorithm Development

### 4.3.1 Species identification

To perform species classifications on fish image sets from our controlled environment enclosures and Cam-Trawl (Williams et al. 2010) imagery, Wang et al (2016) developed a two-level codebook algorithm. This unsupervised classifier uses two steps to identify the most informative features to discriminate between species. The 2016 multispectral image set was one of the examples used in their paper. While the single-image collections from 2015 and 2016 did not require detection or tracking, the videos from 2019 did require those processes. Detection, segmentation, and tracking used similar processes to those processes already described for the rail analyses, albeit much simplified by the controlled environment (e.g., 2D tracking instead of 3D).

### 4.3.2 Length measurement

Measuring fish lengths from images of fish sliding over a well-lit monochrome surface reduces or eliminates many of the problems encountered for the rail analyses. However, several problems remain, including locating the nose and tail measuring points, adjusting for camera and perspective distortions, and following the curve of fish that are not posed in a straight line. Huang et al. (2016) developed an effective algorithm to measure fish lengths from images of fish placed on a calibrated flat surface. This algorithm rectified the image, segmented the fish from the background, and located the fish midline through recursive morphological operations on the segmented fish outline (Figure 27).

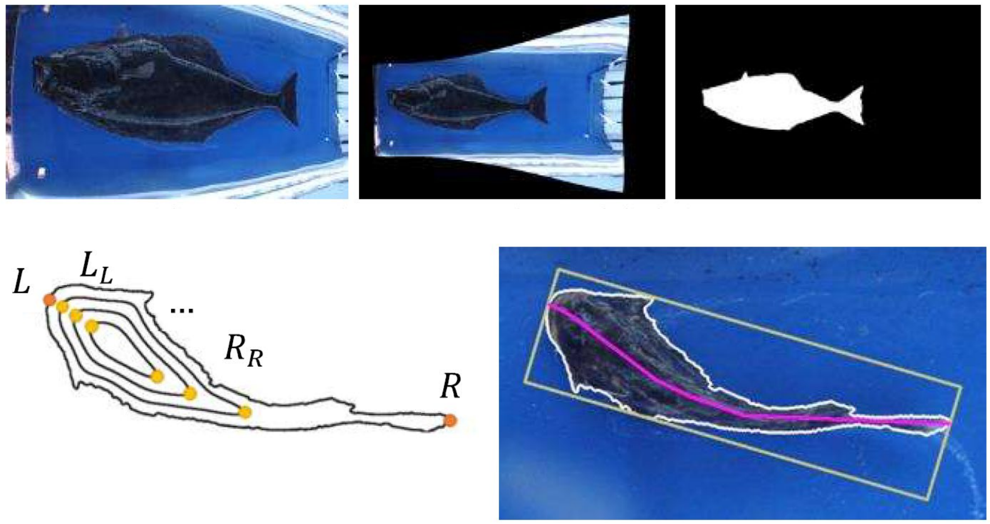



Figure 27. -- Method to measure lengths of fish from images (Huang et al. 2016)

4.3.3 Active learning

In applying a species classifier to fisheries monitoring, new image sets could have enough differences from the training set that performance would be significantly degraded. Active learning is a case of machine learning in which a learning algorithm can interactively query a user to label new data points with the desired outputs. Unlabeled data are often abundant but manual labeling is usually expensive. This is the scenario for fish species classification for fisheries monitoring where there will be lots of unlabeled images of fish compared to the manually labeled training dataset. In such a scenario, learning algorithms can actively query the user for labels. This type of iterative supervised learning is called active learning. Wang et al. (2019) developed an active learning algorithm that indicates which images would most improve the classification model if identifications were provided by a reviewer. The algorithm then retrains the model by including those identifications.

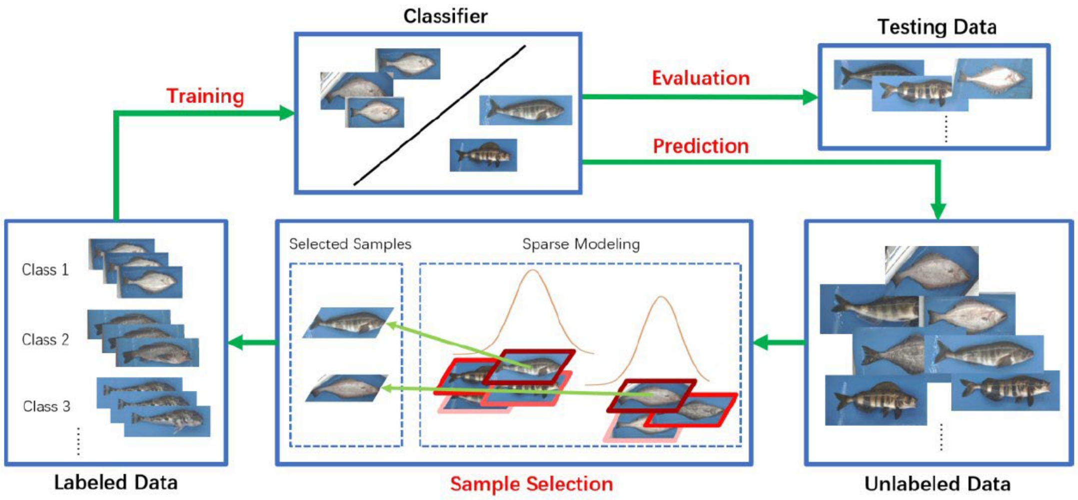



Figure 28. -- Schematic showing the active learning model.

The framework for this classification model is depicted in Figure 28 above. Starting in the lower left of the image, labeled data are used to train the classifier which, in turn, is evaluated against a testing data set. The classifier is used with an unlabeled dataset to classify images to (predict) species, and based on a sample of the predictions, the labeled dataset is updated, and the model retrained. First, a multi-class classifier is initially trained on the labeled data. A trained classifier is then applied on the unlabeled data. Based on the classifier predictions, sparse modeling via Gaussian kernels is used for sample selection. These selected samples are then labeled and moved from unlabeled sets to labeled sets. At the end of each iteration, the classifier is re-trained with the updated labeled set. Finally, the performance of active learning is evaluated on an independent testing dataset.

## 4.4 Controlled Environment Algorithms - Review and Results

### 4.4.1 Species Identification

The classification algorithm was trained and tested with 6,740 images from the 2016 survey (Wang et al. 2016). Testing was done with 10-fold cross validation, excluding all testing images from the training used to predict their classification. Separate classifiers were trained with a combination of the seven images from the narrow-frequency-band, filtered cameras and with the unfiltered camera alone. The multispectral classifier produced results with 97% accuracy at identifying 42 taxonomic classes, plus an 'other' class for species with < 20 images (Table 12). Identification accuracy was affected by how many images were available for each species and whether species were similar in appearance (e.g., arrowtooth and Kamchatka flounders). The classifier trained on images from the unfiltered color camera performed with a very minimal accuracy reduction (94%). Given the complexity of the Multispectral system, its further use was not pursued.

Table 12. -- Species included in the 2016 classification training dataset along with the number of identified images used in the initial training model and the accuracy of classifications from the resulting model (Wang et al. 2016, see Section 4.4).

| Common name | *Scientific Name* | Number | Accuracy (%) | Common name | *Scientific Name* | Number | Accuracy (%) |
|---|---|---|---|---|---|---|---|
| Northern Rock Sole | *Lepidopsetta polyxystra* | 753 | 100 | Sea Urchin | *Strongylocentrotus sp.* | 57 | 98.3 |
| Kamchatka Flounder | *Atheresthes evermanni* | 553 | 91.5 | Dover Sole | *Microstomus pacificus* | 55 | 87.3 |
| Flathead Sole | *Hippoglossoides elassodon* | 511 | 98.2 | Prowfish | *Zaproridae* | 54 | 90.7 |
| Arrowtooth Flounder | *Atheresthes stomias* | 436 | 88.1 | Sablefish | *Anoplopoma fimbria* | 50 | 100 |
| Rex Sole | *Glyptocephalus zachirus* | 433 | 98.6 | Mud Skate | *Bathyraja taranetzi* | 48 | 91.7 |
| Northern Rockfish | *Sebastes polyspinis* | 385 | 99.7 | Scissortail Sculpin | *Triglops forficatus* | 47 | 78.7 |
| Atka Mackerel | *Malacocottus zonurus* | 307 | 100 | Brown King Crab | *Lithodes aequispinus* | 45 | 68.9 |
| Walleye Pollock | *Gadus chalcogrammus* | 280 | 100 | Poacher | *Leptagonus frenatus* | 44 | 79.6 |
| Darkfin Sculpin | *Malacocottus zonurus* | 279 | 98.9 | Armorhead sculpin | *Gymnocanthus galeatus* | 42 | 95.2 |
| Pacific Ocean Perch | *Sebastes alutus* | 227 | 98.7 | Giant Grenadier | *Albatrossia pectoralis* | 40 | 100 |
| Pacific Cod | *Gadus macrocephalus* | 223 | 97.8 | Ebony Eelpout | *Lycodes concolor* | 34 | 85.3 |
| Shortspine Thornyhead | *Sebastolobus alascanus* | 213 | 99.5 | Golden King Crab | *Lithodes aequispinus* | 33 | 66.7 |
| Yellow Irish Lord | *Hemilepidotus jordani* | 193 | 97.4 | Toad Lumpsucker | *Eumicrotremus phrynoides* | 32 | 87.5 |
| Shortraker Rockfish | *Sebastes borealis* | 154 | 100 | Basket star | *Gorgonocephalus eucemis* | 31 | 100 |
| Black Spotted Rockfish | *Sebastes melanostictus* | 139 | 96.4 | Pacific Octopus | *Enteroctopus dofleini* | 31 | 80.7 |
| Red Squid | *Berryteuthis magister* | 131 | 100 | Sea cucumber | *Cucumaria fallax* | 27 | 92.6 |
| Searcher | *Bathymaster signatus* | 97 | 94.9 | Sturgeon Poacher | *Podothecus accipenserinus* | 24 | 62.5 |
| Spectacled Sculpin | *Triglops scepticus* | 76 | 92.1 | Leopard Skate | *Bathyraja panthera* | 22 | 36.4 |
| Dusky Rockfish | *Sebastes variabilis* | 67 | 95.5 | Black Rockfish | *Sebastes melanops* | 21 | 61.9 |
| Bubblegum Coral | *Paragorgia arborea* | 61 | 95.1 | Harlequin Rockfish | *Sebastes variegatus* | 21 | 76.2 |
| Pacific Halibut | *Hippoglossus stenolepis* | 59 | 88.1 | Whiteblotched Skate | *Bathyraja maculata* | 20 | 30 |
| | | | | Other | | 385 | 78.2 |

### 4.4.2 Length measurement

Huang et al. (2016) compared 3,571 automated length measurements from images taken during the 2015 survey with manual measurements of the same fish taken onboard during collection (Table 13). Overall mean absolute errors were 1.49% of the measured length. While the maximum error rate was 2.7% for Pacific ocean perch, all but 3 of 11 species had error rates at or below 1.5%.

Table 13.-- Mean absolute error (%) of automated measurements of groundfish.

| **Common name** | **Scientific Name** | **Number** | **Mean absolute error (%)** |
|---|---|---|---|
| Arrowtooth Flounder | *Atheresthes stomias* | 722 | 1.7 |
| Flathead Sole | *Hippoglossoides elassodon* | 450 | 1.1 |
| Pacific Cod | *Gadus macrocephalus* | 282 | 1.1 |
| Pacific Halibut | *Hippoglossus stenolepis* | 213 | 1.3 |
| Pacific Ocean Perch | *Sebastes alutus* | 156 | 2.7 |
| Rex Sole | *Glyptocephalus zachirus* | 178 | 1.5 |
| Shortspine Thornyhead | *Sebastolobus alascanus* | 210 | 2 |
| Southern Rock Sole | *Lepidopsetta bilineata* | 316 | 1.5 |
| Walleye Pollock | *Gadus chalcogrammus* | 839 | 1.3 |
| Yellow Irish Lord | *Hemilepidotus jordani* | 71 | 1.1 |
| Yellowfin Sole | *Limanda aspera* | 134 | 1.1 |
| | | 3571 | 1.49 |

### 4.5 Controlled Environments - Rockfish Identification Experiment

With the accuracy of rockfish identification proving to be quite high, a separate model specifically for rockfish species identification was developed and tested for species that field biologists misidentify at higher rates. Specialized imagery collections (beyond those conducted on trawl surveys, Table 14) were completed for three rockfish species with remarkably similar appearances, blackspotted rockfish, rougheye rockfish, and shortraker rockfish (Fig. 29), including genetic sampling to assure accurate identifications. As these species can be difficult to distinguish in the field, they are grouped as part of observer program visual identification policies to ensure accurate collections. The lighting and background were tightly controlled to allow consistent levels across light spectrums in the multispectral chute. Fin clip samples were collected from all imaged fish and analyzed for genetic markers to verify accurate species identification. Over two survey periods, 1 Hybrid, 55 shortraker rockfish, 104 rougheye rockfish, and 130 blackspot rockfish were imaged with the species identification verified by genetics. These species can be difficult for trained scientists to distinguish in the field.

Table 14. -- Specialized annual rockfish deployments conducted as part of the rockfish identification experiment and summary of images collected.

| Year | Deployments | Notes |
|---|---|---|
| 2017 | Auke Bay Sablefish Survey (FV *Alaska Leader)* | blackspotted, shortraker and rougheye rockfish. 1,050 images with paired genetics (Multiple images of the same specimens collected) |
| 2018 | Auke Bay Sablefish Survey (FV *Ocean Prowler)* | blackspotted, shortraker and rougheye rockfish. 124 images with paired genetics |
| 2019 | Observer collected specimen | blackspotted and shortraker and rougheye rockfish. 14 images with paired genetics |

blackspotted rockfish (*Sebastes melanostictus*) | shortraker rockfish *(Sebastes borealis)* | rougheye rockfish *(Sebastes aleutianus)*

Figure 29. -- Examples of rockfish species images from the controlled environment (chute) camera systems.

The goal was to evaluate the controlled environment species identification algorithm to identify similar looking species. If successful, it can be used on other difficult to distinguish species, such as juvenile arctic cod and juvenile pollock, while out in the field by observers or by scientists on research cruises to augment or expand collections. Images collected during these deployments were annotated and 60% were randomly selected for training the model with the remaining 40% reserved for testing for accuracy (Table 15). Testing resulted in 93% accuracy of species identification of the testing are depicted in the confusion matrix in Table 16. A confusion matrix is a table that is used to describe the performance of the classifier on a set of test data for which the true values are known. The results indicated that tight control over lighting and background combined with a low number of classes allowed for a relatively small training set for these species.

Table 15. -- Rockfish species dataset.

| **Rockfish species** | **# Training images** | **# Testing images** |
| --- | --- | --- |
| blackspotted rockfish (*Sebastes melanostictus*) | 305 | 77 |
| rougheye rockfish *(Sebastes aleutianus)* | 166 | 42 |
| shortraker rockfish *(Sebastes borealis)* | 217 | 55 |
| **Totals** | **688** | **174** |

Table 16. -- Testing results for the rockfish identification experiment. The matrix of labels predicted by the algorithm against actual species identification (true label) using the rockfish classification test set. Overall accuracy of predictions was 93%.

| | | **Predicted label** | | |
| --- | --- | --- | --- | --- |
| | | **Blackspotted** | **Rougheye** | **Shortraker** |
| **True label** | **Blackspotted 98.7% accurate** | **76** | 1 | 0 |
| | **Rougheye 76.19% accurate** | 3 | **32** | 7 |

## 4.6 Controlled Environments: Species Identification - Discussion, Operational Readiness and Summary

The development of a species identification algorithm is one of the underlying tenets of research conducted by EM Innovation. This algorithm has many potential applications. As a standalone program incorporating different training datasets, the algorithm can be used in various situations where species identification from images is needed. As an integrated solution for real-time species identification, the algorithm can be incorporated into a camera system such as the camera chute systems that were designed for this research. The species identification algorithm also has the potential to be applied to other research areas such as species identification of discarded bycatch, or species identification on a processing plant belt or where fish are caught at the rail. In all cases, training datasets will need to be defined for the specific purpose.

### 4.6.1 Standalone Application

As a standalone application the fish species identification classifier can be applied to a multitude of research efforts where fish identification is a requirement. The application approach can now be utilized with other species by collecting an image data set, selecting a percentage for training and testing, and increasing the training as needed to refine the model and achieve the desired results. An example of such efforts would be the Rockfish experiment as mentioned in Section 4.5 where species identification is needed for species that are similar visually.

In its current form, the species identification classifier was developed as a set of Python scripts for use in a MS Windows® environment. These scripts were run as part of the training and testing of the algorithm and can be released to the community for further development. Release of this type will require formal documentation and rights management.

To operate effectively as a standalone application that can be used by non-technical researchers, the scripts will need to be integrated into an application with intuitive user interfaces. This can either be in the form of a desktop application running natively or as an online browser-based application hosted with cloud computing capabilities. Both options have their own benefits and shortcomings: a desktop application would allow for running of the algorithms in the field where internet coverage is not available while a browser-based system would allow for lower system requirements and be more available to users.

Whichever development path is chosen, this application will need to clearly indicate which training datasets are used and how to train new datasets so that the application is useful in new research situations. The current training dataset is based on species from the Alaska region and therefore, further training will be required to adapt the algorithms to identify species from other regions. Due to the nature of the active learning model, this retraining should not require a major number of resources; further testing and analysis of the existing algorithm scripts will inform this determination.

#### 4.6.2 Integrated Real-time Species Identification Camera System

Integrating the species identification classification algorithm(s) into a camera hardware/software system will allow for real-time identification of fish species in the field in a controlled environment. An image of the fish can be captured and immediately analyzed provided certain input criteria are met, such as clear backgrounds, adequate lighting etc. Integrating the algorithms back into the camera acquisition systems used to collect the training data would allow for real-time identification. As a fish slides through the chute, it can be analyzed and identified. This ‘Species Identification Chute’ would be beneficial for automated bycatch monitoring and reporting.

A specialized version of this kind of bycatch monitoring system was developed specifically for a single species, halibut (see Section 5). The halibut bycatch chute is an example of an integrated real-time system, extending this for use with the species identification classifier would allow for automated reporting of various species. To achieve this type of system, application development is planned for a standalone species identification application capable of running on the chute camera system PC with further research and development to add to length estimation reporting. The purpose of this Species Identification Chute would be to collect fish imagery to be analyzed by a species identification classifier and fish measurement tools, thereby providing automated species composition and length data on monitored trips. Physical hardware development conducted for the Halibut bycatch chute can be used in this endeavor; however as with development of the application, baseline training datasets will have to be acquired and algorithms will need to be retrained where appropriate. Tools for labeling and retraining will have to be effectively communicated and tested before this can be achieved.

### 4.6.3 Applying the Species Identification Classifier in Uncontrolled Environments

Figure 30. -- Examples of different species in uncontrolled capture environments on the rail and in open air chute-like capture poses and uncontrolled lighting and backgrounds

Research is currently being conducted for species identification in uncontrolled environments for chute-like images of fish on tables without lighting and background control. Images and genetic samples of the rougheye, shortraker, and blackspot rockfish species group are currently being collected by observers during commercial fishing operations. These images will be tested on the model and verified with genetic identifications (Section 4.5) to assess if it can perform as well without strict background and light controls in place without a complete retraining effort. A new model trained with salmon images with the same algorithmic approach will also be tested in an uncontrolled lighting environment to identify salmon sorted at processing plants to species after the belt detector described in Section 6 has detected the salmon. The salmon in the experiment will be placed flat on a table to capture images for the species identification model.

The model developed for chute images cannot be applied directly to rail images because the training images between the two image types are much too different to have the models directly transfer. Species identification of fish caught at the rail of fixed gear vessels (Section 3), presents an uncontrolled background and lighting environment, and hence have proven to require more training images to achieve accurate identification to species. For species identification at the rail, a specialized classification model was developed with a training dataset of 17 classes, but with the success of the chute's active learning model a similar strategy is planned for the rail applications. Applying the active learning classifier model strategy on the Rail data has the potential to expand species identification beyond these initial classes by automatically training the model with newly learned classes. Without active learning the Rail classification model would have to be completely re-trained for any new class. The concurrent development of species identification models for controlled chute and uncontrolled rail images has made clear that training data utilized for any application must have similar complexity in terms of background and lighting to the image data targeted for analysis.

## 5. HALIBUT BYCATCH COUNT AND LENGTH ESTIMATION – TRAWL FISHERY

Bycatch of Pacific halibut is an important management issue for most Alaska trawl fisheries and is difficult to monitor with precision, particularly for individual vessel bycatch. Therefore, it was an application pursued earliest for the controlled-environment chutes described in Section 4. Halibut are required to be returned to the sea quickly. The length-weight relationship for halibut is well established, so measuring and counting halibut discarded through a chute would provide good data on the weights and numbers discarded.

For Bering Sea bottom trawl fisheries, vessel-specific halibut bycatch mortality can be a limiting criterion. Vessels participating in that fishery can sort halibut from catches (deck-sorting) before transferring those catches into holding bins to release halibut sooner and reduce their mortality with requirements to assure that the weight and condition of the released halibut is recorded (https://www.federalregister.gov/documents/2019/10/15/2019-22198/fisheries-of-the-exclusive-economic-zone-off-alaska-halibut-deck-sorting-monitoring-requirements-for). The EMI project has been developing and testing camera chutes to automatically measure such halibut as they are released during deck-sorting.

During deck-sorting operations, the codend is opened on deck and halibut are sorted out for release as catch is moved to holding bins. Deck-sorting is monitored by observers who must leave factory sampling to occupy on-deck stations where they sample halibut as they are released. This stops processing lines and takes time away from other observer duties. An automated alternative for this monitoring of deck-sorting activity would be the implementation of a camera chute system with integrated machine vision algorithms to estimate the halibut bycatch lengths from images taken as halibut are released through the camera chute. Mortality can be estimated using the time of halibut release (time out of water) and would be recorded by the system. Deployment of these chute systems would expedite handling with halibut returning to the ocean sooner.

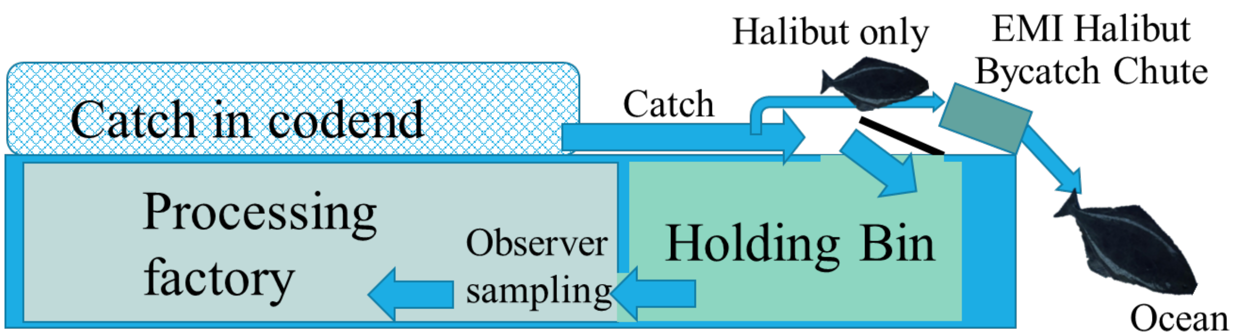



Figure 31. -- Schematic showing where the halibut bycatch chute fits into the deck sorting process; catch is brought on deck (in the codend) and halibut are sorted from the catch and discarded through the EMI halibut bycatch chute. The remaining catch is stored in holding bins before being transferred to the vessel's factory where the observer samples the catch before it is processed.

The first halibut bycatch chute system was built and deployed in 2014. Since then, chutes have been deployed nine more times on volunteer catcher-processor (CP) vessels conducting deck-sorting operations. (This included operations under an exempted fishing permit as deck-sorting methods were developedand since the regulations implementing the deck-sorting program were adopted). Chute systems were installed on deck between the observer sampling table and the discard chute. Deck-sorted halibut were passed through the chute for image collection and processing as the halibut were discarded. Imagery collected from these deployments was used to train and test the halibut length measurement algorithms. The main application goal for these chutes is to enable rapid discard and enumeration of deck-sorted halibut in support of on-board observer data collection and halibut bycatch estimation. The figure below summarizes the iterative research cycle of that EMI chute system.

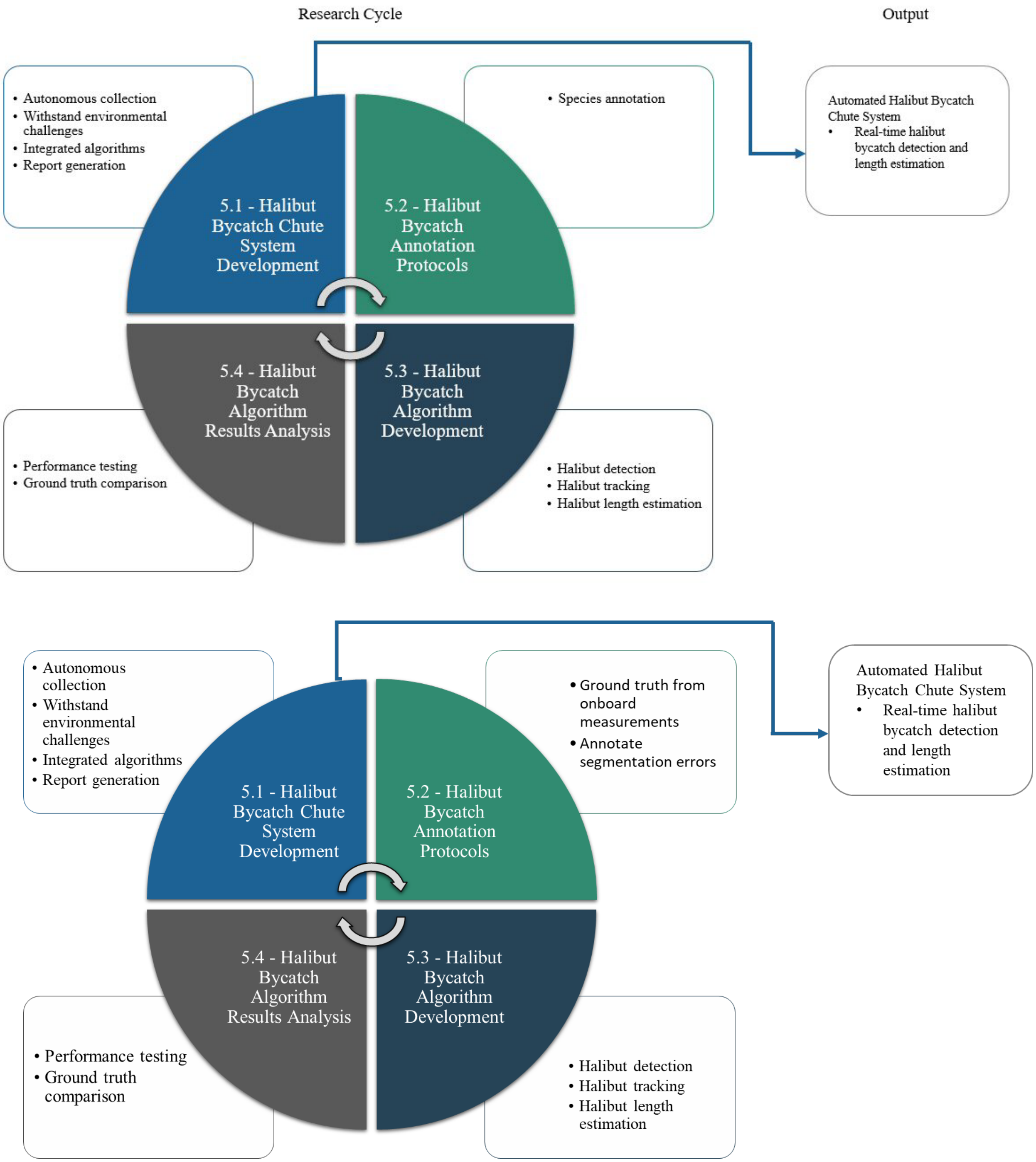



Figure 32. -- Research cycle for the halibut bycatch chute.

## 5.1 EMI Halibut Bycatch Chute System - Camera System

The design for the halibut bycatch chute follows on from the camera chutes developed for data collection activities. Halibut are passed through the chute as they are sorted from the catch on-deck and discarded back to the ocean. While passing through the chute, imagery is captured of the fish (Fig. 33). The system consists of a metal box, a camera, light panels, and computer components.

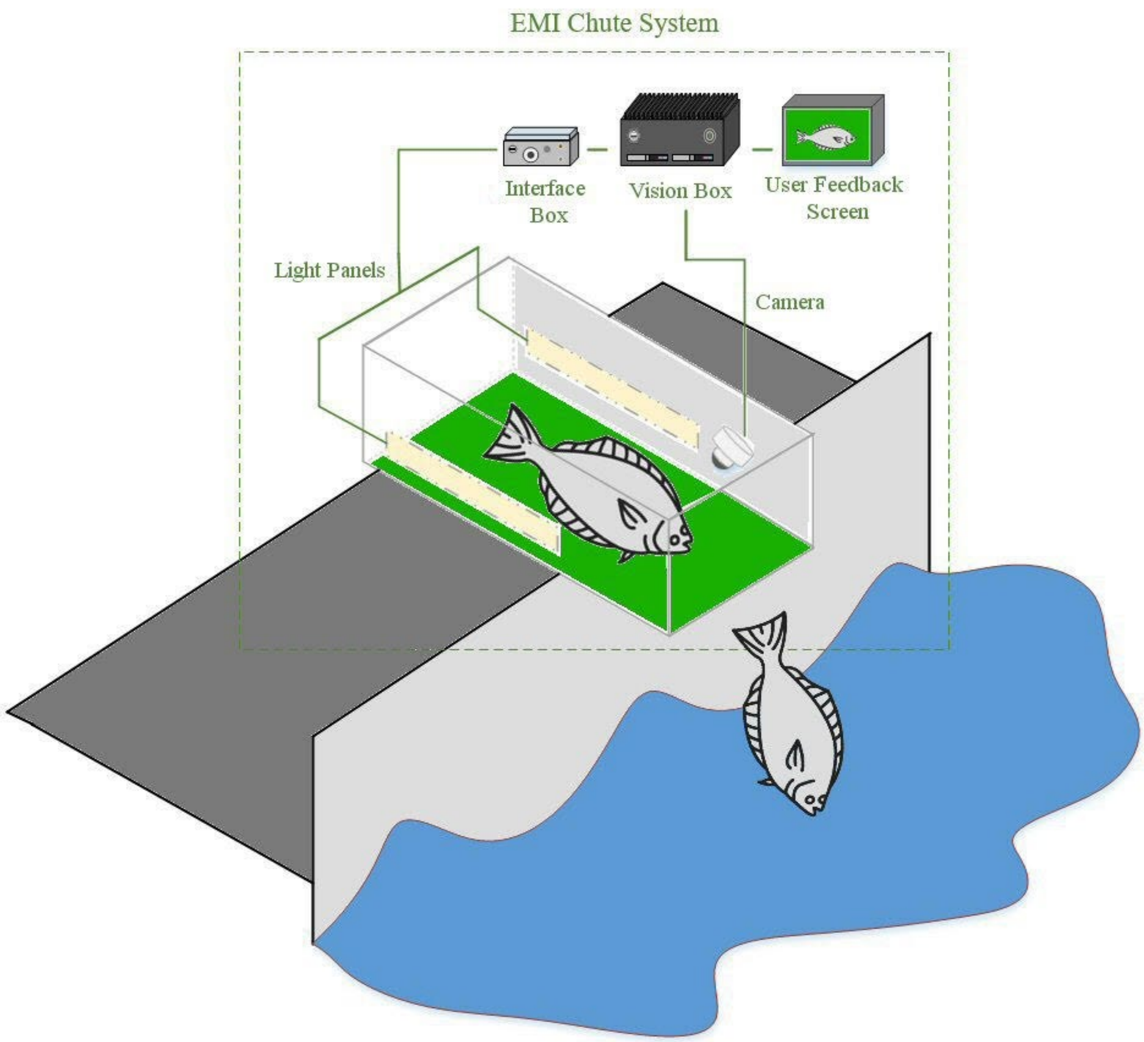



Figure 33. -- Halibut bycatch chute system.

The halibut bycatch chute system has been iteratively designed, building upon the lessons learnt from development of other camera chute systems. Design requirements specific to the halibut chute are listed below in Table 17.

Table 17. -- Descriptions of halibut bycatch chute camera system design requirements and constraints.

| Design requirement/Constraint | Description |
|---|---|
| **5.1.1 - Hardware Requirements and Environmental Constraints** | The halibut chute needs to withstand the outdoor deck environment in the trawling industry. |
| **5.1.2 - Image Acquisition and Image Quality** | Images need to be clear for halibut measurement, including halibut fish shape algorithms. Consistent background required for background segmentation. Lighting must be bright enough to allow fast exposures, keeping images of sliding fish clear. |
| **5.1.3 - Autonomous Collection and Integrated Real-time Algorithm Analysis** | As the halibut is discarded through the chute it needs to be recorded. The only action that is required from the observer or crew is to pass the fish through the chute. Image analysis occurs as the halibut is discarded through the chute. |

### 5.1.1 Hardware Requirements and Environmental Constraints

The original halibut bycatch chute paralleled the MVC chute design seen in Section 4. The chute is typically affixed to the side of the vessel, with fish passing through it sent back into the ocean. As the chute would be deployed at sea in the open on the deck, the build needs to withstand these environmental conditions. Early deployment of the MV camera chute to function as the halibut bycatch chute led to vessel personnel indicating that chute size, particularly height, would be a problem on many vessels. This MV camera chute was also not sustainable for long trips with water and light intrusion issues becoming more prevalent the longer it was deployed.

The redesign included fabricating an aluminum body and reducing the chute dimensions (Fig. 34). Height was reduced by putting the camera above the exit end of the chute, instead of above the center. This oblique view puts the camera directly above the fish snout, which is commonly several cm above the chute surface, which would cause an overestimate of length from the central position. Water intrusion that damaged electronics was solved by identifying the sources of the intrusion and components and replacing them with waterproof computers (the same model used in the rail system in Section 3.3). Moving to a solid enclosure helped address the light intrusion issues as did adding rigid entry and exit door panels with rubber edge flaps. It was also discovered that occasionally fish could stick to the floor of the chute depending on the angle at which the chute was installed and on the nature of the fish. To make it easier for the fish to slide through the chute, a rinse water hose was installed to wet and wash down the chute floor making it easier for fish to move through the chute. The chute redesign also addressed vibrational issues which were causing electrical connections to fail or become intermittent by deploying dedicated solid state circuit boards and soldered or screw-down/clamping connections.

Figure 34. -- Evolution of the physical chute design showing the design improvement from the early MV camera chute (left) to the halibut bycatch chute design (right).

### 5.1.2 Image Acquisition and Image Quality

The camera deployment strategy used in the MV camera chutes was adopted for the halibut bycatch chute deployments. MV cameras used in the stereo rail system were deployed using a modified version of Rail Acquisition software application to acquire the images. The camera is installed in a fixed position within the enclosed chute which meant the camera housings did not require the same intensity of waterproofing as that of the Rail system. The camera is connected to a waterproof PC and separately provided 12-volt power. A light beam interruption sensor was connected to the PC as well and, when triggered, a signal was sent to the camera to strobe the lights and acquire an image. Because these chute systems were deployed on larger vessels, the power restrictions were not as restrictive as those encountered on vessels for the Rail system, and thus the camera and PC could be left powered on throughout the duration of the trip. Lighting power was minimized because lights were triggered as rapid strobes.

While this strategy was able to acquire images, there were significant challenges that had to be overcome. The sensor was vulnerable to water intrusion over periods of weeks to months, largely because they had to be installed at the exit end of the chutes where they were exposed to substantial amounts of water. This resulted in either missed triggers or a significant amount of false triggering which in turn would clog the PC storage system with wasted images. Maintaining image quality over the duration of a trip was also challenging using this camera system. Keeping the settings of the camera coordinated over prolonged periods of time was difficult due to the MV cameras rebooting and reverting to their factory settings. Fixed camera settings that were optimal when the chute was deployed could also become poor when lighting partially failed and resulted in images with inconsistent clarity and color.

Camera faceplate fouling was also an issue (see Fig. 35 for examples). Soiling of the camera viewport occurs over time due to water spots and flapping fish. Fish flapping was strong enough to move the camera, which invalidated calibration and reduced views of the chute floor. Water drops are regularly splattered onto the camera port, clouding parts or all of the image. A blower was installed to blow air at the camera port, triggered by the fish passage sensor, and while it did remove water drops it also left a salt film residue. Freshwater spray was partially effective when combined with a blower.

Maintaining consistent lighting was also difficult. Overhead lighting created strong reflections on chute surfaces. Placing strips of LEDs close to the chute surface along the sides prevented most reflections. Painting all interior surfaces white augmented reflected light. Light needed to be strong enough to allow image acquisition durations below 1/500 sec to limit motion blurring. The painted chute bottom surface began to peel over time, hampering object detection algorithms. To overcome paint peel, colored plastic sheets were used instead of paint allowing for more longevity and durability. Additionally, green plastic was selected to replace blue floor color it provided a brighter background for subtraction and was available in a more durable sheeting.

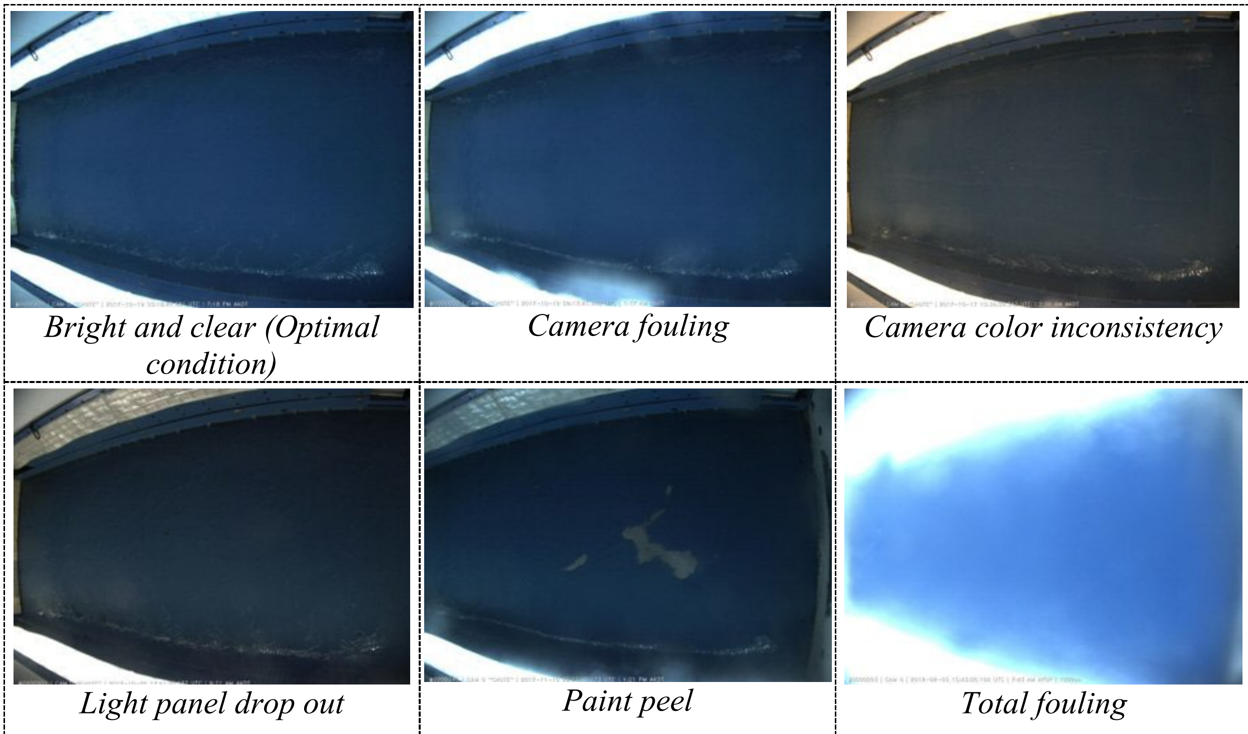



Figure 35. -- Examples of various physical issues occurring in image acquisition ranging from no issues (clear view, upper left) to a totally fouled view (lower right).

As the halibut chute and associated algorithms evolved, tests with reduced resolution determined that image quality requirements (megapixel size) could be relaxed from those initially used (from 2.8 to 0.7 megapixel). This meant dedicated MV cameras could be replaced with IP-type cameras. This was a significant breakthrough since standard off-the-shelf cameras and IP camera acquisition software could be used. These types of IP cameras are typically used in surveillance systems and have built-in motion detection triggers. This availability combined with the elimination of the complexity of the light beam triggering system was a substantial part of the decision to change to IP cameras.

Software issues to implement a system based on IP cameras were addressed by moving away from the custom-built image collection application and making use of the open-source video surveillance application iSpy to control and acquire the needed imagery. This in turn led to more consistent image acquisition and offered more control in terms of camera settings.

The IP cameras continuously monitor for movement in the chute, providing a live view of what the cameras are seeing. Adding a feedback monitor screen to the system allows the vessel operators and system users to see when the camera faceplate needs cleaning. This makes cleaning of the interior of the chute easier since the effect of cleaning the chute floor and faceplate can be seen on captured imagery. Images in Figure 36 below depict the IP camera installation in the chute as well as the video monitor as well as an example of a video frame captured through the camera in Figure 37.

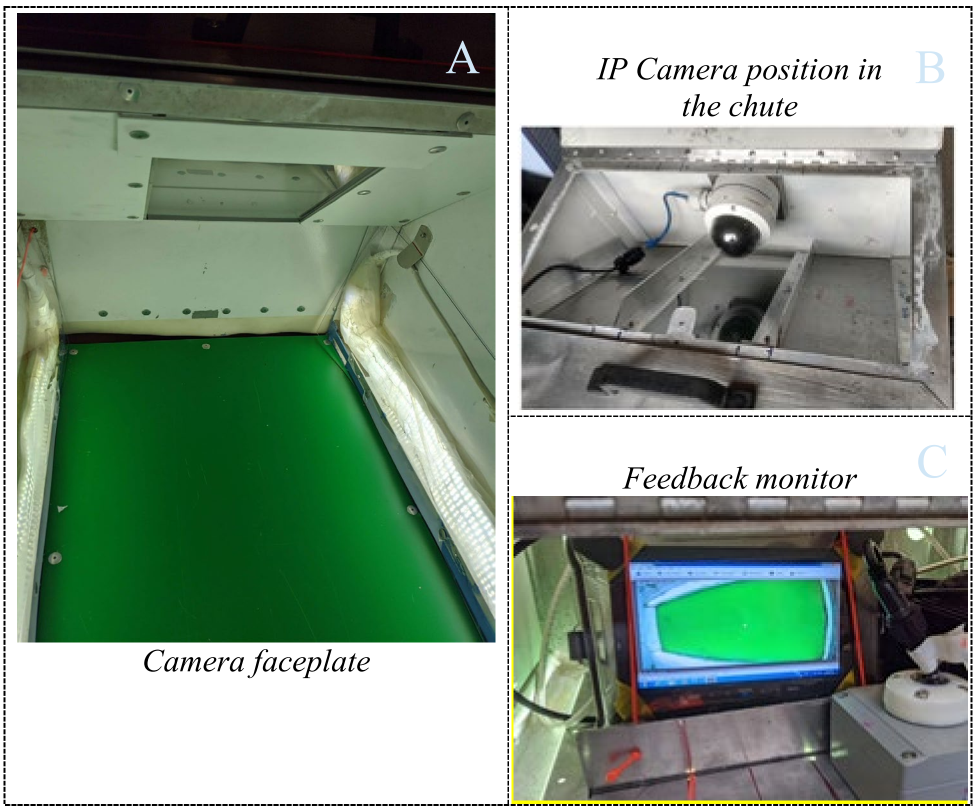



Figure 36. -- View of camera plate from inside the chute (A), IP Camera installation (B), and the feedback monitor showing an image of the chute interior (C).


Figure 37. -- Image captured with the improved IP camera version of the halibut bycatch chute

### 5.1.3 Autonomous Collection and Integrated Real-time Algorithm Analysis with IP cameras

For the halibut bycatch chute to be effective, it needs to be able to operate without obstructing the current duties of the observers and the discard activities should also not be slowed. Fish should be sent through the chute with minimal effort from the user. The user operations should be limited to turning the light panels on, passing the fish through the chute, and turning the light panels off at the end of sorting activities. During early deployments, the observer entered their manually measured halibut lengths into our system using a keypad, but this proved to be cumbersome, and the keypad electronics eventually failed from water intrusion. Keeping the observer-measured lengths and the automated estimated lengths matched for a single halibut was also challenging.

As the development of the detection, tracking, segmentation, and length estimation of the halibut improved, it became feasible to integrate the algorithms directly into the camera chute system, allowing on-board analysis during image collections. Initial integration of IP video into the acquisition application proved difficult since an application interface had not been developed., and the algorithm classification code had been directly incorporated into the MVChute acquisition application. As the algorithms were refined and newer versions became available, the IP system could not be easily updated. Attempts were made to develop integrated algorithms with IP camera acquisition software, but this method proved too restrictive. Instead, it was decided to develop a workflow control application that would pass the imagery from the acquisition application to the standalone algorithm application (Fig. 38). This method allowed for more flexibility, catering for both image and video files to be able to be recorded and for those files to also be processed in parallel using multiple instances of the algorithm application independent of the camera.

While the length estimation occurs in near real-time using this method, analysis of the collected data can only occur once the results are available to fisheries analysts. Designs were considered that increased how quickly and easily the results could be transferred. Provision was made for the transfer of the length estimate results from the chute system PC to the onboard observer's PC via USB transmission. Once on the observer PC, these results were added to the observer data transmission application (ATLAS) and transferred to ASFC FMA databases together with standard observer data. This proof of concept design is under review for further analysis.

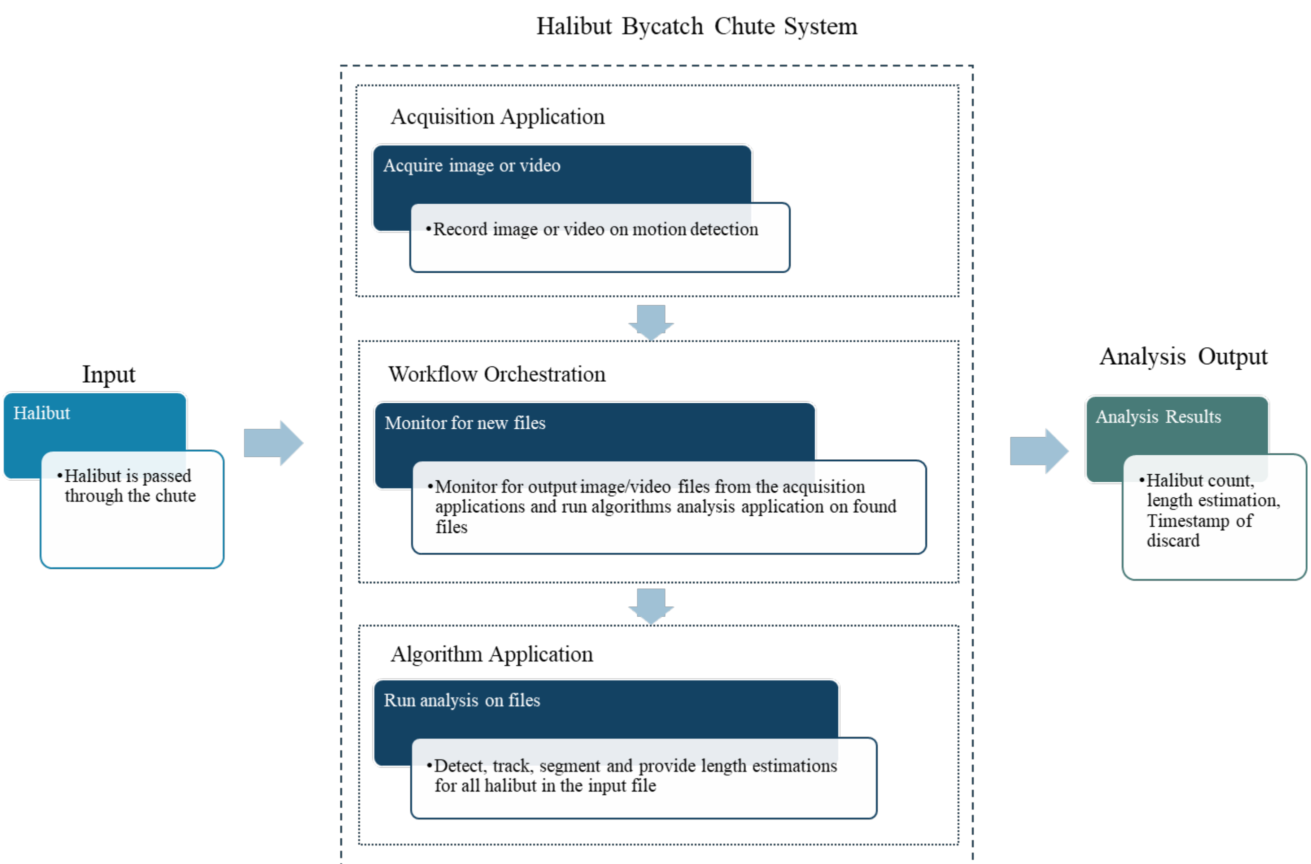



Figure 38. -- Schematic of the integrated acquisition and analysis design using workflow orchestration.

### 5.1.4 Deployments

From 2014 to 2020 halibut bycatch chute deployments were made on 14 trips aboard seven Gulf of Alaska catcher vessels and during 12 extended (2 weeks – 3 months) sessions aboard four Bering Sea catcher-processors. (Table 18).

Table 18. -- Summary of halibut bycatch chutes deployments. * indicates Bering Sea catcher/processors, no * indicates Gulf of Alaska catcher vessels. Numbers in parentheses refer to the number of deployments made that year, if more than one.

| Year | Deployments | Notes |
|---|---|---|
| 2014 | • FV *Constellation** | Proof of concept |
| 2015 | • FV *Arica** | Data collection, vessel operator feedback, hardware development and automated analysis - MVChute |
| 2016 | • FV *Katie Ann**<br>• FV *Laura*<br>• FV *Marathon* (3) | Data collection, vessel operator feedback, hardware development and automated analysis - MVChute |
| 2017 | • FV *Arica**<br>• FV *Cape Kiwanda* (3)<br>• FV *Elizabeth F*<br>• FV *Excalibur*<br>• FV *Marathon* (3)<br>• FV *Nichole*<br>• FV *Sea Mac*<br>• FV *Seafreeze America**<br>• FV *Topaz* | Data collection, vessel operator feedback, hardware development and automated analysis - MVChute |
| 2018 | • FV *Arica**<br>• FV *Marathon*<br>• FV *Nichole* (2)<br>• FV *Seafreeze America** | Data collection, vessel operator feedback, hardware development, and automated analysis - MVChute image frames |
| 2019 | • FV *Arica* (2)* | Data collection, vessel operator feedback, hardware development, and automated analysis - VideoChute collection on second deployment |
| 2020 | • FV *Arica**<br>• FV *Seafreeze America** | Data collection and automated analysis, trial system for data transfer - VideoChute |

## 5.2 Halibut Bycatch Chute Annotation

Halibut length-measurement algorithms were based on differentiating fish from the background (segmentation) and calibration geometry, assuming that halibut were in full contact with the chute surface. However, some fish could not be accurately measured, as the halibut were bent or parts were above the chute surface due to the fish flapping as it passed through the chute. To identify when this occurred, some annotation was needed to identify images where fish position or posture prevented accurate measurement from the image. Several parameters were extracted from image segmentations, including total area, tail area, boundary length, and width and ratios of these were tested to see which best identified images from which length measurements were inaccurate. The ratio of tail area to total area found most of those images, as flapping fish usually had tails closer to the camera, hence appearing much larger. Later, the ‘Fit Ratio’ from the halibut shape fitting algorithm (see Section 5.3.3) provided better discrimination of images likely to be inaccurately measured. Annotation was also important to identify image segmentations affected by problems such as material on the chute surface or droplets on the camera lens. Segmentation parameters were adjusted, and the shape-fitting algorithm was developed to minimize the effects of such problems on length measurements. Ground-truth testing was done by comparisons with observer measurements of each halibut.

## 5.3 Halibut Bycatch Chute Algorithm Development

The computer vision algorithms aim to count and estimate length of each halibut as it passes through the halibut bycatch chute. To do this, the algorithms needed to segment and separate the detection from the background and then to measure its length. When there were multiple images of each fish (video), the algorithms also needed to detect the fish in each image frame and track the movement of that fish from one frame to another before segmentation (Fig. 39; Table 19).

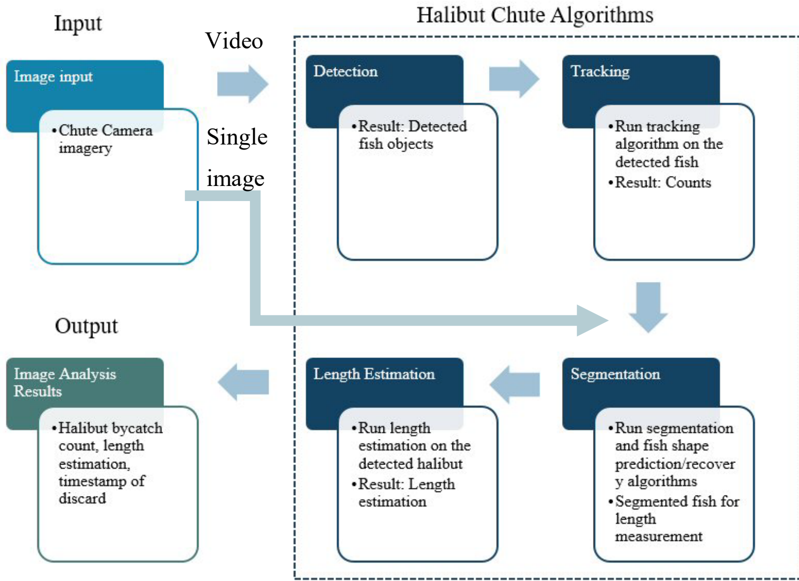



Figure 39. -- Schematic depicting the steps involved in automated halibut length estimation from input of video or images (upper left) through a series of algorithms, to output of final results.

Table 19. -- Description of algorithms used in estimation of halibut length and count; additional details can be found in the referenced sections.

| Research deliverable | Description |
|---|---|
| **5.3.1 - Camera Calibration** | Adjusts images for perspective and any camera or lens distortions. This results in a calibration parameter file used during fish measurement and identification. |
| **5.3.2 - Halibut Detection and Tracking** | Tracks the detected halibut from one frame to another |
| **5.3.3 - Segmentation, Shape Prediction and Length Estimation** | Distinguishes fish from chute background to isolate fish information for further analysis. This process determines where the fish is in each frame from individual images or video.<br><br>Extrapolates the shape of halibut detected in the frame. If there is only a partial view of the halibut, the process estimates the shape of the full fish. |

### 5.3.1 Camera Calibration

The calibration algorithm adjusts the images for perspective and any camera or lens distortions, providing a calibration parameter file for fish measurement and identification. With calibration parameters, rectified images can be produced having a consistent relationship between pixel dimensions and real-world dimensions across the imaged surface and in all directions, allowing accurate measurements and consistent fish shapes. The calibration algorithm derives the camera matrix and its relationship to the chute surface from images of a board with a checkerboard pattern with known dimensions that is moved across the whole chute surface. Collections of such calibration images were completed before the systems were deployed. Any changes in the camera view relative to the chute surface invalidate the calibration parameters, so additional sets of calibration images, a two-to-five-minute process, were taken periodically during and after deployments.

### 5.3.2 Halibut Detection and Tracking

When used as halibut bycatch chutes, MVChutes did not require detection and tracking, as only a single image was taken of each fish. As VideoChutes provided video clips, including multiple images of each fish, their use required detection and tracking algorithms. A Gaussian Mixture Model (GMM) is used for detection and tracking objects in the video imagery. GMMs have been used extensively in object tracking of multiple objects where the number of mixture components and the mean of the mixture components predict object locations at each frame in a video sequence. For tracking, bounding box detections from the GMM detector are connected if their overlapping Intersection-over-Union (IoU) is above the threshold. IoU is an evaluation metric used to measure the accuracy of an object detector on a particular dataset. As there is only one fish per frame in previous data, this detection and tracking is quite straight forward. The images in Figure 40 depict the detection and tracking of a halibut in a video recorded within a halibut bycatch chute.

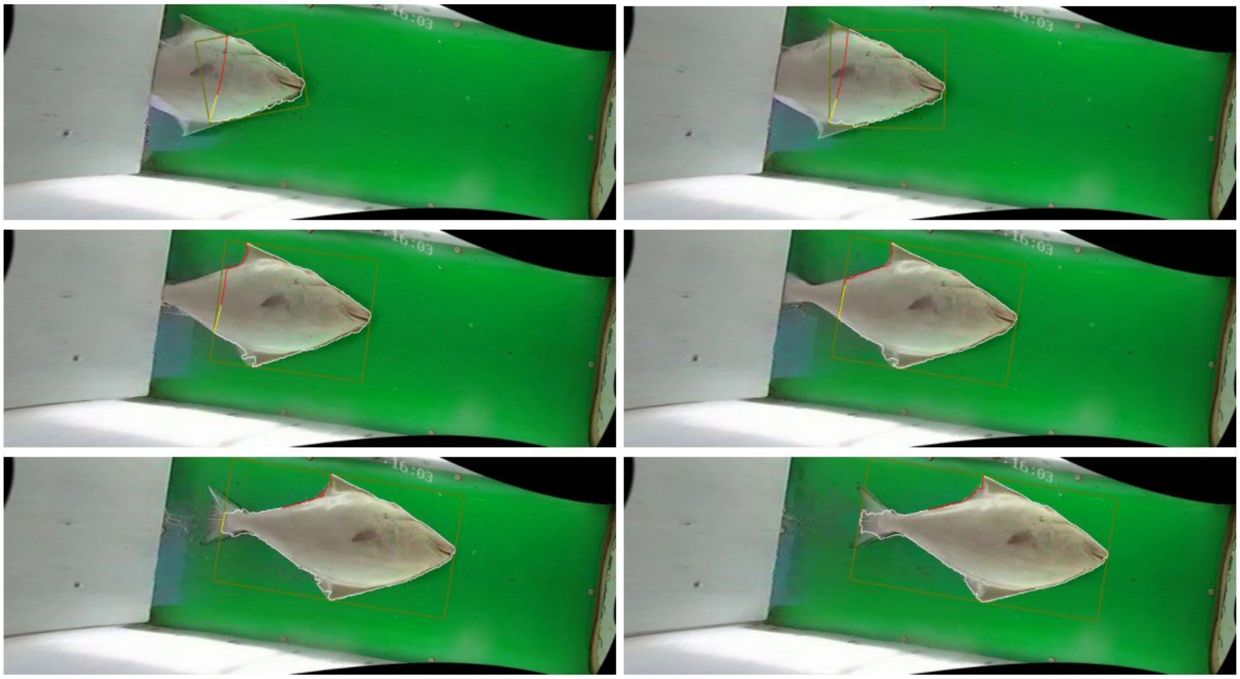


Figure 40. -- Detection and tracking example. Frames are transposed relative to camera calibration.

### 5.3.3 Segmentation, Shape Prediction and Length Estimation

Segmentation and measurement of early image collections used the methods developed in Huang et al. (2016) and described above in Section 4. While these methods were effective for most fish images, major challenges identified with some fish segmentations included: images blurred due to water drops on the camera lens, some part of the fish body outside of the camera view and fouling or glare on the chute surface being mistaken for part of the fish. These issues motivated development of an algorithm that fit the segmentation outline to general halibut shapes, allowing well-segmented sections to overcome segmentation flaws.

For the algorithm requirement, a coarse-to-fine contour-based method for segmentation refinement and missing body recovery for chute-based halibut images was developed. Several segmentation approaches were applied to the raw fish images to get the initial segmentation mask which is then treated as the input for the developed refinement system. At the beginning, the initial segmentation contour is aligned with pre-trained representative contours using an affine transform, constituting the coarsest level for entire contour alignment. Then the contour segments are refined iteratively to represent the poorly segmented or missing portions of the testing image (Fig. 41). From coarse to fine, the segmentation refinement progressively focuses more on local parts of the fish, allowing for more variation of the fish shape (Wang et al. 2018).

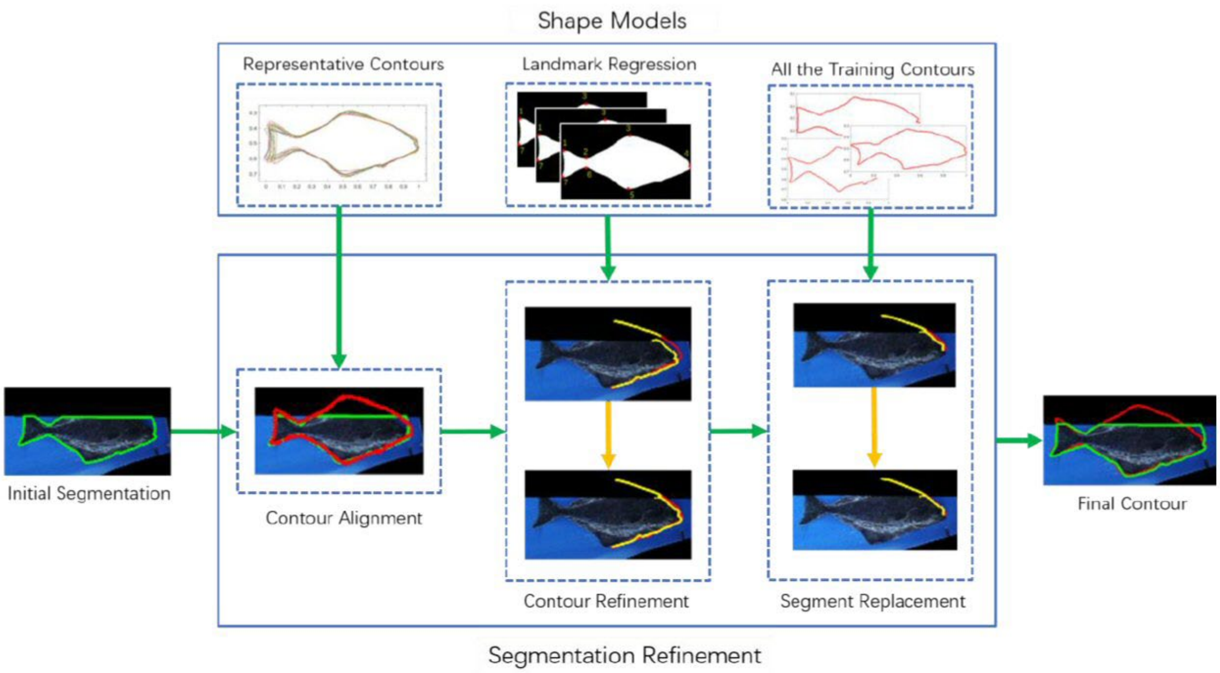
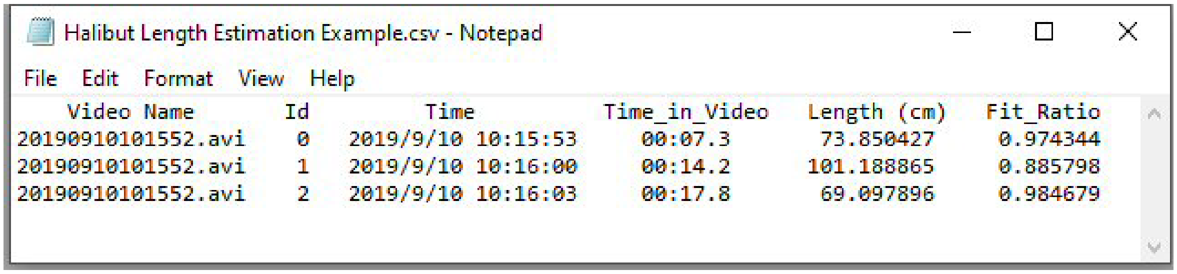



Figure 41. -- Schematic showing the fish shape prediction and recovery algorithm process starting with initial segmentation through to estimation of the final contour of the imaged fish.

Measuring the fish size and length requires a robust segmentation approach which was generally well-accomplished by the shape-fitting algorithm. The Huang et al. (2016) length-measurement algorithm derived lengths by locating the tip of the snout, the middle point of the tail, and generating a series of line segments between the two with intermediate points located along the middle of the fish shape. The simplest and most accurate method for halibut was found to be using only a single intermediate point at the center of the narrowest part (caudal peduncle) of the segmented fish profile and adding the lengths of the two line segments between that point and the snout and the tail. The halibut length estimation algorithms produced an output csv which includes the video file reference, a reference to the time on the video, the estimated fish length, and a Fit Ratio (Fig. 42). Low Fit Ratio (< 0.65) identified fish that could not be reliably measured due to bad segmentation or pose. For image ID 2 (Fig. 42), images of the input image and recovered shape are given in Figure 43.



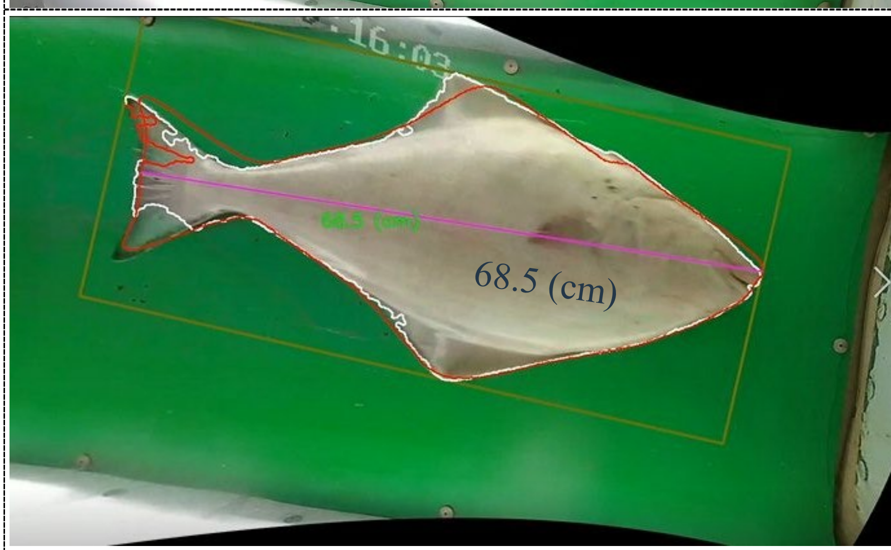
Halibut Length Estimation Example.csv - Notepad

File Edit Format View Help

| Video Name | Id | Time | Time_in_Video | Length (cm) | Fit_Ratio |
|---|---|---|---|---|---|
| 20190910101552.avi | 0 | 2019/9/10 10:15:53 | 00:07.3 | 73.850427 | 0.974344 |
| 20190910101552.avi | 1 | 2019/9/10 10:16:00 | 00:14.2 | 101.188865 | 0.885798 |
| 20190910101552.avi | 2 | 2019/9/10 10:16:03 | 00:17.8 | 69.097896 | 0.984679 |

Figure 42. -- Example of algorithm CSV output showing a sample of halibut length estimation csv file showing the video filename, image ID, time references, estimated length, and fit ratio.


*Original video frame*




*Output analysis rectified for perspective (calibration), shape recovery (red outline) and length estimation (purple line)*

Figure 43. -- Example photos showing input image and final recovered shape and estimated length.

## 5.4 Halibut Bycatch Chute Algorithm Review and Results

The success factors for the development and implementation of algorithms for automated halibut bycatch length estimation are two-fold:

1. Measuring accuracy results. The algorithm results need to be accurate to a certain measure of confidence. The measure of accuracy is determined by how closely the output length measurements are compared to measured lengths.
2. Measuring processing performance results. The processing time needed to run the algorithms should not be longer than it takes to manually review and extract the needed data. The measure of processing performance is determined by the processing time benchmarks for running algorithms and determined by how long it takes to run an analysis application, relative to a human review.

### 5.4.1 Halibut Bycatch Algorithms - Accuracy Results

The initial study conducted to evaluate the halibut bycatch chute camera system demonstrated that the algorithms could achieve highly accurate length estimates (Fig. 44). In 2014, 183 halibut were deck sorted, measured, and discarded through a camera chute deployed aboard the FV *Constellation*. Comparing the ground truth measurements to the automated length estimates demonstrate that image-based measurements a) produce similar estimated length compositions across the full collection, b) are highly correlated and c) have low error and minimal bias (Fig. 44). The mean difference between image-based length estimates and measured lengths was 0.07 cm, with a standard deviation of 0.85 cm. Seventy-nine percent of fish measured from images were within 1 cm of the physical measurements. Error rates did not appear to be length dependent. The system performed equally well regardless of whether the eyed (dark) or blind (white) side of the halibut was presented (mean difference for individual fish = 0.13 cm) (Wallace et al., 2015).

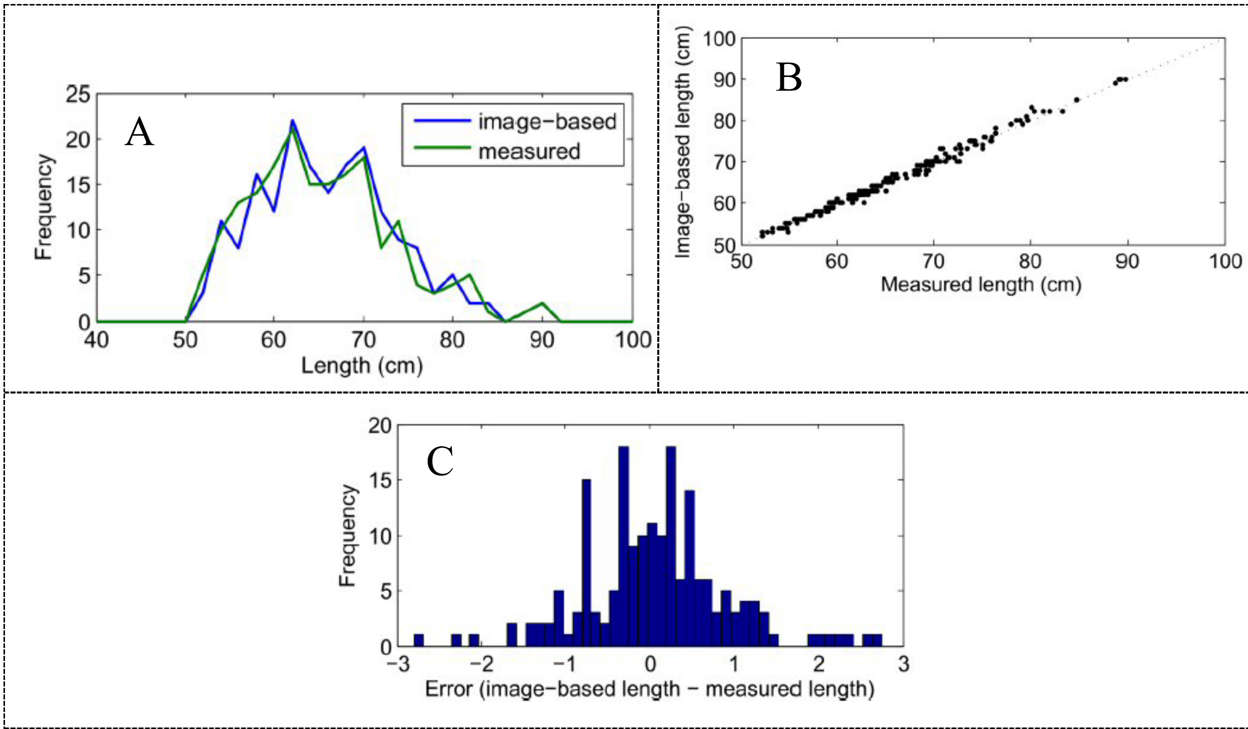



Figure 44. -- Comparisons of halibut bycatch chute (image-based) length estimates relative to onboard physical length measurements for 183 halibut from 2014 deployment aboard the FV *Constellation*; frequency distributions for estimated (blue) and measured (green) lengths (A), estimated and measured lengths were close to equal (B), and histogram of differences between estimated length and measured length with mean difference of 0.07 cm and standard deviation of 0.85 (C).

Data were analyzed from collections during a two-month deployment aboard the FV *Arica* in the fall of 2017, where halibut were deck-sorted and discarded through a halibut bycatch chute and onboard observers collected measurements for about 20% of the discarded halibut. Halibut images with Fit Ratios below 0.65, indicating segmentation errors or poses that would produce high measurement errors, were eliminated from the analysis. In practice, the sizes of such fish would need to be inferred from the remainder of the collection. This provided 498 observer measurements for comparison with measurements obtained through the length estimation algorithms. The shape-fitting algorithm was used for segmentation and greatly improved segmentation and measurements for images with lighting, clarity, and fish-position issues. Nearly all measurement errors were smaller than 2 cm. original images had 2.8 megapixel (mP) resolution. However, after image resolutions were reduced by half in both dimensions (0.72 mP) and then by half again (0.18 mP) before analysis, length measurement accuracies were unchanged, indicating that high resolution images are not necessary for this application.

Comparisons of algorithm estimates with actual measurements show similar length distributions, with regression analysis showing greater than 96% of variance explained and an average difference for image measurements of less than 0.5 cm (Fig. 45). Further analysis of observer measured lengths compared to the automated analysis for 2019/2020 (the first years of IP camera chute data) is ongoing.

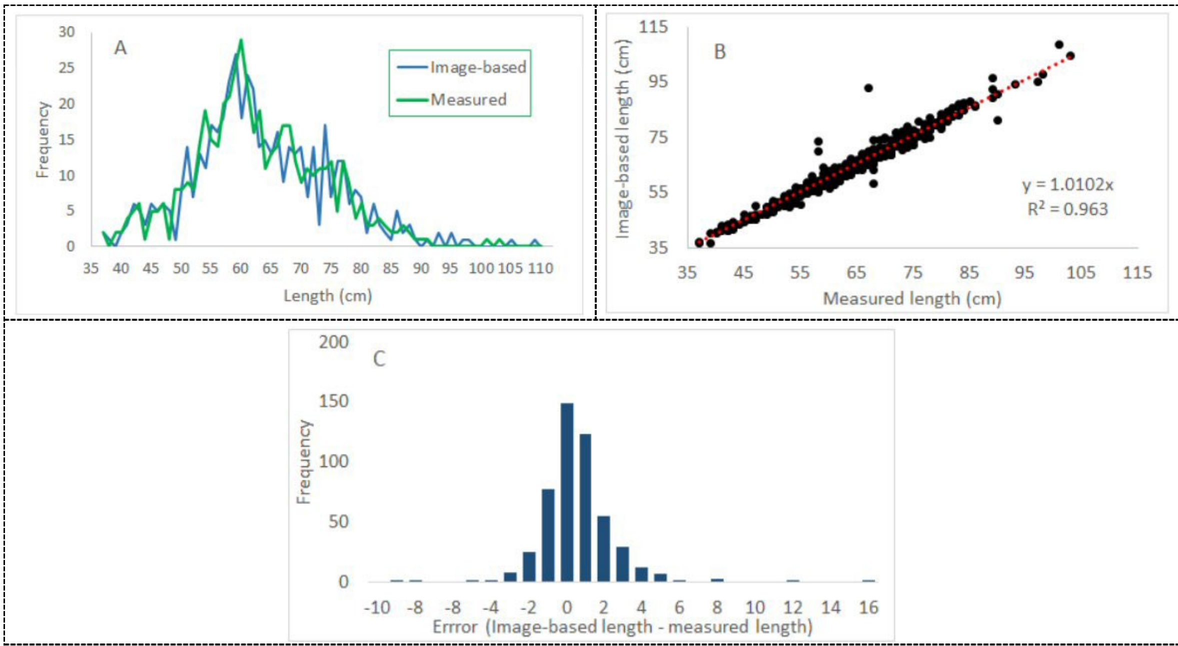



Figure 45. -- Comparison of 498 image-based and observer-measured halibut length measurements from the FV *Arica* between September 8 to November 5, 2017 (90 images (15%) with Fit Ratio less than 0.65 had been eliminated). Frequency distribution of estimated (blue) and measured (green) lengths (A), regression of estimated lengths against measured lengths showing $R^2$ and regression equation (B), and histogram of errors (differences between estimated and measured length (C).

### 5.4.2 Halibut Bycatch Algorithms - Performance Results

Halibut bycatch algorithms were packaged into a single executable application. The input to the application is either a list of images or a video file. Processing time is based on the number of images or frames in the video; with shape fitting, processing takes a few seconds to process each image, on a computer without dedicated GPU processing. As each fish provides 4-8 images, processing can be faster or slower than halibut arrive, depending on the sorting operation. This allows for the algorithms to be integrated directly on the camera PC with analysis usually occurring directly after the images are acquired. If a new video is acquired while analysis is still being conducted, that second video gets placed in a queue to be analyzed once the original is completed. The performance results and the relatively low processing power for running the algorithms allows for this near real-time analysis. As deck-sorting occurs only once for every haul and lasts less than 30 minutes, and hauls take hours to be completely processed, there would be time to complete analyses before the next haul is sorted.

## 5.5 Halibut Bycatch Open Air Detection

Implementing the halibut bycatch chute can be cumbersome for vessels with limited space on deck and eliminating the enclosure could improve acceptance and implementation. Such ‘open-chutes’ have been tried in other projects including EFP 2007-02 (Bonney 2008) and initial trials for UW’s Advanced Physics Laboratory (APL) (Brodsky 2017). The EMI team advised the UW APL on later stages of their project that used an enclosed chute. Challenges with open-air chutes have included obstruction by arms and hands of personnel, and the highly variable lighting conditions, including shadows and glare, encountered on deck with the open chutes (Figure 46). Overcoming these issues is possible, however more work is needed to fully develop this potential.

To evaluate such an alternative, video was collected in 2019 to investigate the possibility of obtaining lengths in an uncontrolled open-air environment. Vessels conducting halibut deck-sorting were required to have video cameras monitoring the sorting process, including cameras viewing the tables where observers collect data on sampled halibut. Video from an IP camera monitoring the sorting process was acquired from a vessel (FV *Cape Horn)* which had a close and unobstructed view of their sampling table. Sorting of two hauls were recorded and used for demonstration and feasibility trials and preliminary algorithm development. Detector training required annotation of bounding boxes around the halibut. Initially a small batch of annotation labels were created for demonstration purposes. More extensive annotations were completed but those bounding box labels have not yet been used for training the model. Approximately 17,775 images were separated for annotation purposes and 2,228 rectangular bounding boxes were created. Preliminary algorithms were successful at detecting and tracking the halibut even with considerable interference from the observer's hands, arms, and sampling equipment, which rendered standard segmentation tools (separating background from foreground) ineffective. Developing this configuration toward implementation would require trials over a wider range of lighting conditions and testing the application of shape fitting tools to measure those halibut. While open-air chutes may prove feasible, further pursuit of this project has been deferred in favor of work with more controlled environments.

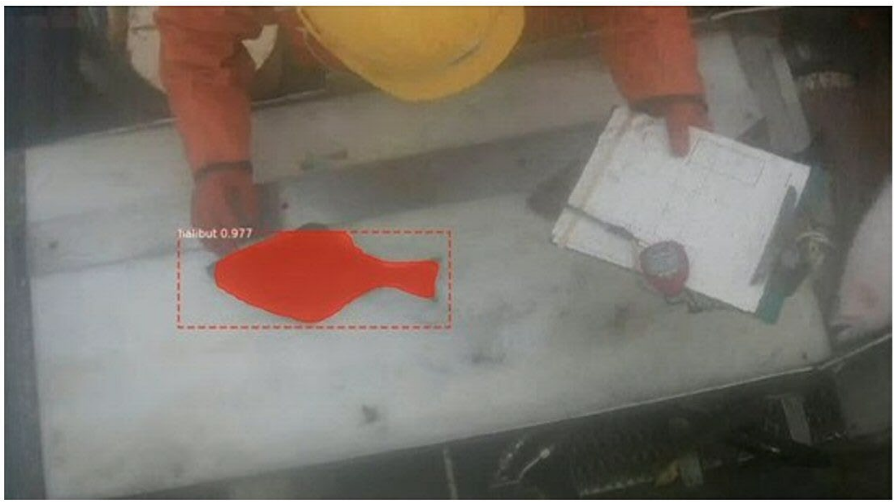



Figure 46. -- Image of a halibut detection on an open-air table and showing the bounding box for the detected halibut.

### 5.6 Halibut Count and Length Estimation Discussion and Operational Readiness

Halibut counting and length estimation systems have been developed to acquire and analyze image data on the vessel. Analysis results can be transmitted to and fed into the data reporting stream for timely catch accounting analysis. While this real-time analysis has been developed, the analysis communication methods and the integration of transmitted data with management systems still needs to be designed and tested. Improvements in both collection and analysis systems will emerge from ongoing deployments and will need to be incorporated into future designs for implementation into the fisheries.

Implementation strategies will need to be defined for both the EMI analysis algorithms and EMI controlled environment camera system before any production implementation can take place.

These strategies should define how algorithms will be run in the field and how they will scale across fisheries.

### 5.6.1 Halibut Count and Length Algorithms - Implementation and Collaboration

Halibut counting and length measurement algorithms have been thoroughly developed and tested, using IP video clips triggered by motion detection, with objects being detected as they are acquired and stored, and routed to the analysis routine. Vessels conducting halibut deck-sorting are already required to have cameras installed and there are established vendors that provide and maintain those IP systems. Instead of introducing a new stand-alone system, implementation could be efficiently achieved by integrating the analysis capability developed here alongside existing video collection systems, with an IP camera in the chute set for motion detection and added to the data collection array. Length measurements would be converted to weights and summed to provide the needed estimates of halibut weight discarded from the deck. Collaboration with vessel video vendors would be needed to find the best way to link analysis capabilities with their systems. The proportion of halibut that could not be measured would be assigned the average weight. Prototype tools have been deployed that provide these weight estimates onboard and transmit data to the AFSC for use in management. These prototypes will need to be improved and tested for full implementation.

### 5.6.2 Halibut Count and Length Camera Systems - Implementation and Collaboration

While the IP camera would be a component of the vessel's video monitoring system, the camera chute has many other important characteristics needed for effective discard monitoring. These include an enclosure to exclude exterior light, artificial lighting arranged to illuminate the chute surface without reflections or glare, a chute surface that allows easy fish passage and a good background for segmentation, fish entrance and exit openings to allow fish passage but exclude external light, and supports to protect, position, and orient the camera to the chute surface. The current chute design achieves all of these but may not be optimal for installation on every vessel. Providing functional requirements, the example of the current system, and availability of project personnel to answer questions would allow vessels to adapt the chute to their vessel and sorting procedures. Specific operational procedures are also critical to chute function, including monitoring the camera port for drops or other fouling and cleaning when needed.

### 5.7 Halibut Count and Length Estimation - Summary

The ability to automatically count and measure halibut as they are being discarded would replace a time-consuming observer task on Being Sea vessels conducting halibut deck-sorting operations and would make vessel-specific bycatch monitoring more precise should such management be implemented for partial-coverage or limited-discard fisheries. Sophisticated, robust algorithms have been developed and proven for measuring halibut for both single-image and video collections. While analysis tools have developed quickly, maintaining system operations for months in the challenging, at-sea environment has been the main roadblock to a system ready for implementation. The shift from sensor-triggered single images to motion-triggered video collection with IP cameras has advanced system durability. These are the same challenges and advances experienced for the species identification chute.

While the halibut-measurement application of the camera chute is mostly developed, several EMI tasks are necessary before implementation, including the following:

- The code for the algorithms needs to be made available to potential users and documented.
- Collaboration with vendors that supply video systems to the Bering Sea fleet is needed to determine how best to link the analysis code to their video-collection systems.
- A thorough and specific list of necessary chute characteristics needs to be developed and communicated to fleet participants, followed by any consultation needed to help them design and install chutes on their vessels.
- The 2020 deployment results need to be analyzed.
- A paper documenting the accuracy of chute-generated weight estimates, needs to be prepared for publication in a peer reviewed journal and provided to both fisheries management and fishery participants. This is needed to achieve buy-in and to justify the necessary implementation and approval efforts.

Halibut bycatch mortality is a driving consideration for many of the management decisions made for Alaska marine fisheries and a significant limitation for many of those fisheries. The camera chutes developed by EMI provide an efficient alternative for measuring halibut discards. They can also measure the time of air exposure for each halibut, which is a significant factor affecting the probability of mortality (Rose et al. 2019). The most immediate application of halibut bycatch chutes would be the deck-sorting operations of Bering Sea bottom trawl fisheries, but they would also be applicable to Gulf of Alaska bottom trawlers, should their management be shifted to require accounting of halibut bycatch at the vessel level. While further advancements such as the open-air chute could make a measurement chute easier to deploy, these chutes have thus far only shown potential at every preliminary level. More development and testing would be necessary if those directions were indicated to be useful, as well as acceptance by the relevant fleet and fisheries management and incorporation into their systems and operations.

# 6 SALMON DETECTION AND TRACKING - TRAWL DELIVERY PLANTS

Observer monitoring for bycatch, particularly for salmon, in catcher vessel deliveries at shore side processing plants is time-consuming, tedious, and is done in parallel with the sorting conducted by plant workers. This monitoring situation was identified as having potential for EM detection and associated species identification algorithms to be used to validate the plant's sorting process. In 2018 and 2019, imagery was collected by deploying IP cameras and making use of cameras already available at four plants in Kodiak. These monitored the flow of catch from rockfish-predominant deliveries into the sorting area and at a designated location where salmon were presented when sorted from the catch. To increase the number of training images, marked salmon were inserted into the flow of fish, in addition to the salmon bycatch in those deliveries. Analyses were developed to monitor the plant personnel's ability to accurately detect, sort, and report salmon bycatch in deliveries from trawlers. Videos were also collected from a few deliveries of walleye pollock at one Kodiak plant in February 2019 to develop detection models to separate salmon from pollock catches. The brief pollock collection mostly used salmon introduced into the catch. If EM can successfully validate that plant sorters remove all salmon from deliveries and correctly account for them, salmon bycatch can then be monitored and salmon counted with delivery tickets already being created and reported by the plants, alleviating the observer workload (see Fig. 47 for the research development cycle).

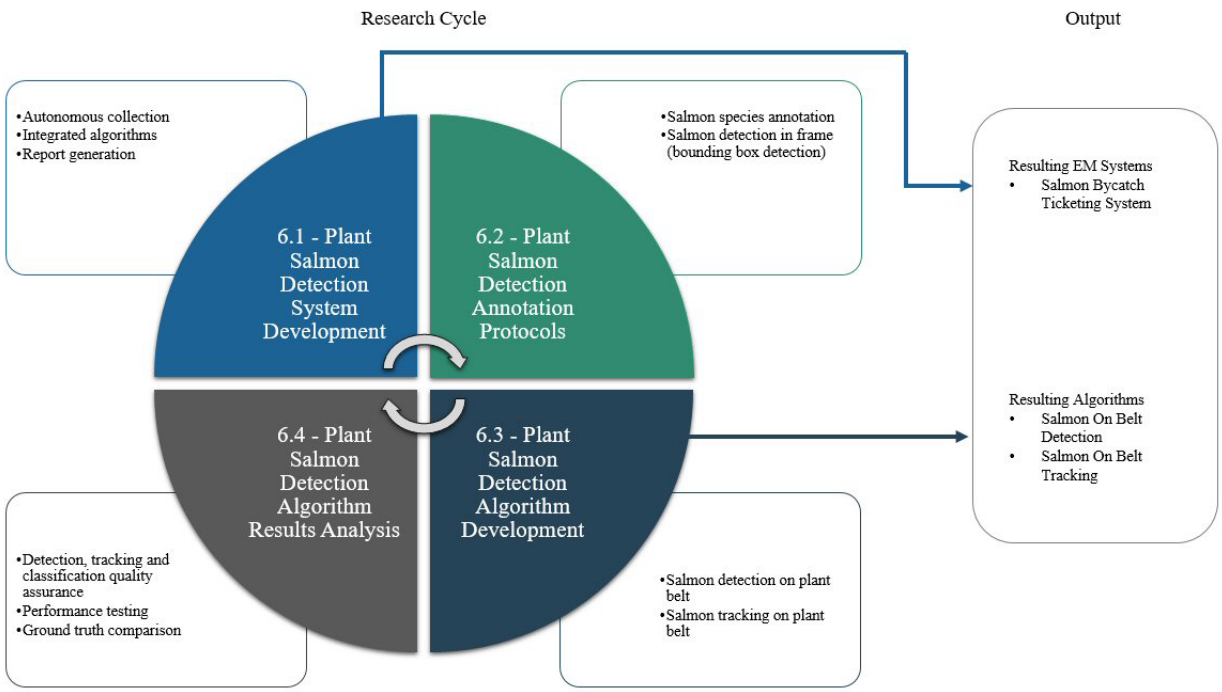



Figure 47. -- Schematic showing the research cycle for detecting and tracking salmon in processing plants (on a conveyor belt system).

## 6.1 Salmon Detection and Tracking on Plant Belts - Camera System

The objective of the camera system in fish processing plants is to validate industry salmon bycatch reporting. Analysis of ‘Entry’ video detects salmon contained in trawl deliveries as the fish move on a belt into the sorting area of the processing plant. In addition, images are acquired by using a ‘Check-In’ camera that salmon are placed under when they are sorted from the catch by plant personnel. Salmon sorting efficiency can be estimated as the proportion of salmon detected by analyzing Entry belt video that are soon thereafter detected on the Check-In video. This does not require all salmon in the delivery to be detected in the Entry belt video. Check-In salmon images are also required for training the machine learning algorithms that will automate species identification. For salmon identification, multiple images of different salmon species were required. Images were initially acquired with the IP camera chute in 2018 (see Section 4.1) and collected from designated locations near the sorting belt in 2019. IP surveillance type cameras, attached to network video recorders, were placed in the sorting plants with views of the moving belts (Fig. 48). Correct species identification is evaluated by comparing automated identifications with salmon reported on the industry reports (plant's fish ticket) for each delivery.

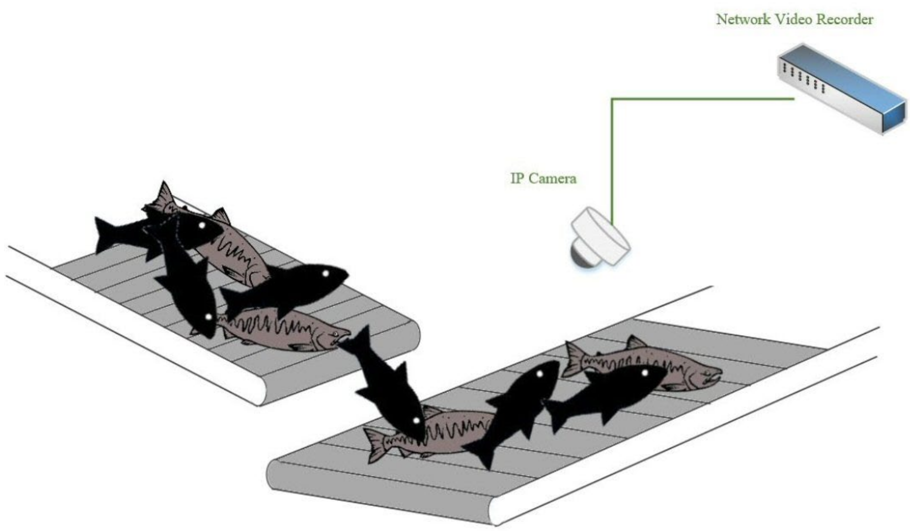



Figure 48. -- Diagram showing delivered catch on a processing plant belt system and salmon ID camera placement. Salmon bycatch are light grey while other fish species in the catch are colored black. IP cameras and video recorders are placed above the conveyor belts.

### 6.1.1 Data Collection for Salmon Detection

The training data consists of video imagery of salmon moving on the plant belt together with other species of fish, namely rockfish and walleye pollock which are separated into image frames for annotation (Figure 49). For data collection, IP camera network video recorder systems were installed at participating plants in 2018 and 2019 (Table 20).

Table 20. -- Summary of salmon detection data collections in shoreside processing plants over a 3-year period.

| **Year** | **Plant** | **Delivery type** | **Total images extracted** |
|---|---|---|---|
| 2018 | **Processing Plants 1, 2, 3, 4** | Rockfish | 1,106,983 |
| 2019 | **Processing Plants 1, 2, 3, 4** | Rockfish | 1,180,148 |
| 2019 | **Processing Plant 1** | Walleye Pollock | 222,327 |

Figure 49. -- Belt image extracted from video collected by a camera system in a shoreside processing plant showing industry personnel sorting salmon from a delivery of rockfish.

Data that was collected as part of the species identification algorithm development in Section 4.1 was used for development of salmon species identification algorithms.

## 6.2 Salmon Detection and Tracking on Plant Belts - Annotation

Two types of labeled datasets are needed to develop algorithms for monitoring salmon sorting and identification: salmon detection in rockfish and pollock deliveries and salmon species identification. Image frames from video acquired from the plants were extracted and annotated. The total number of images per video determined the annotation sample section. Two types of belt videos were annotated: salmon in rockfish deliveries and salmon in pollock deliveries.

Annotation was completed using LabelImg with 34,666 rectangular bounding boxes created for salmon detected in rockfish deliveries. For salmon detection in pollock deliveries, a total of 222,327 images with 33,225 rectangular bounding boxes were created (Fig. 50).


Figure 50. -- Photo extracted from video recording of a pollock delivery in a shoreside processing plant. The bounding box in the center of the image (shaded blue) shows an annotated salmon detection.

Salmon species ID annotation was conducted as part of the controlled environment species identification stream (Section 4.2). Salmon specific annotation consisted of identifying and cataloging salmon tails and salmon bodies to species (Table 21).

Table 21. -- Listing of the total number of salmon annotations per year and application type.

| Year | Application | Images reviewed | # of Bounding boxes |
|---|---|---|---|
| 2018 | Salmon Plant Belt | 23258 | 11,454 |
| 2019 | Salmon Plant Belt | 523069 | 79,345 |
| 2015 | Salmon ID Chute Tail | 156 | 156 |
| 2015 | Salmon ID Chute Body | 156 | 156 |
| 2018 | Salmon ID Chute Tail | 29 | 28 |
| 2018 | Salmon ID Chute Body | 29 | 29 |
| 2019 | Salmon ID Chute Tail | 49501 | 893 |
| 2019 | Salmon ID Chute Body | 49501 | 1,616 |

## 6.3 Salmon Detection and Tracking on Plant Belts - Algorithm Development

The goal of the detection and tracking computer vision algorithms developed for processing plants is to detect salmon from multiple species of fish moving on a belt and record the time that each salmon passed. To accomplish this the algorithms need to be able to detect the fish in the image frame and track the movement of the fish from one frame to another (Figure 51). Another algorithm classifies salmon appearing on a second camera to species and record the times that those fish appear. Research for this classification makes use of the species identification algorithms developed in Section 4.

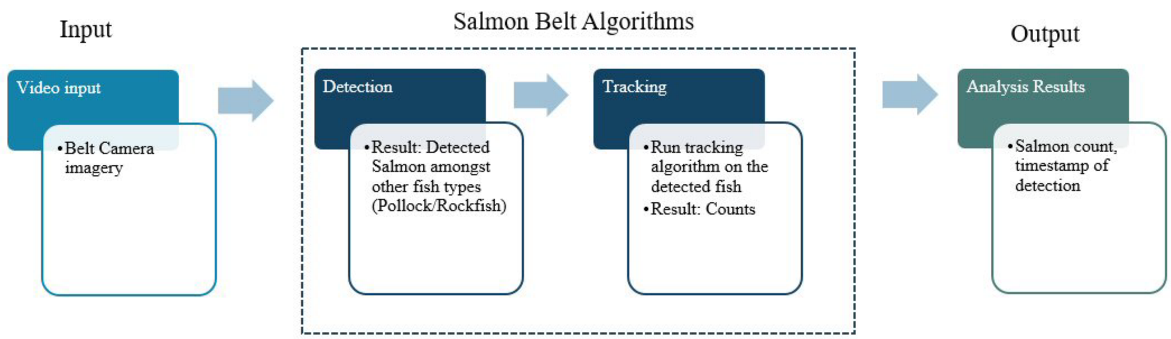



Figure 51. -- Schematic showing salmon detection and tracking algorithms, data inputs, and final output.

The detection algorithm is based on the Faster Region-based Convolutional Neural Network (RCNN) method for higher accuracy and for robustness. Faster RCNN is an object detection architecture and uses convolution neural networks like YOLO (You Only Look Once) and SSD (Single Shot Detector). Faster RCNN is composed from three parts, convolution layers, a Region Proposal Network (RPN), and classes and bounding boxes prediction.

For the tracking, the initial approach made use of DeepSORT, a tracking-by-detection algorithm that considers both the bounding box parameters of the detection results, and the information about appearance of the tracked objects to associate the detections in a new frame with previously tracked objects. However, this learning-based tracking algorithm is more suitable for unpredictable trajectories and clustering is offline and slow.

A new tracking algorithm TrackletNet Tracking (TNT) was applied to automatically use the frame rate information from the video input file to adjust thresholds and the overall confidence to filter tracks. In this framework, detections are associated based on the CNN appearance-feature similarity and overlapping between frames. After achieving passage-based tracking results, video visualizations and results in csv format are available for users in the post-processing step (Fig. 52 and 53.

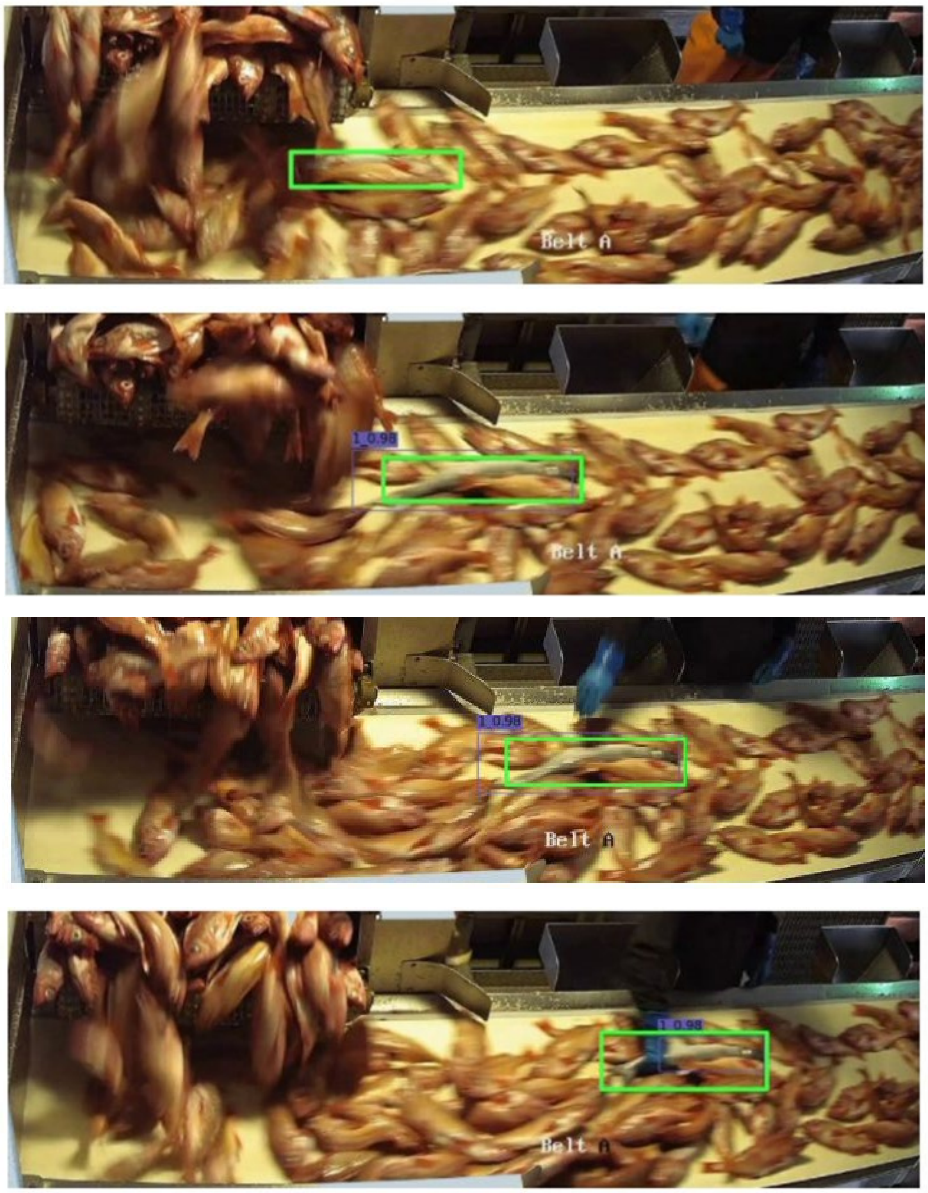
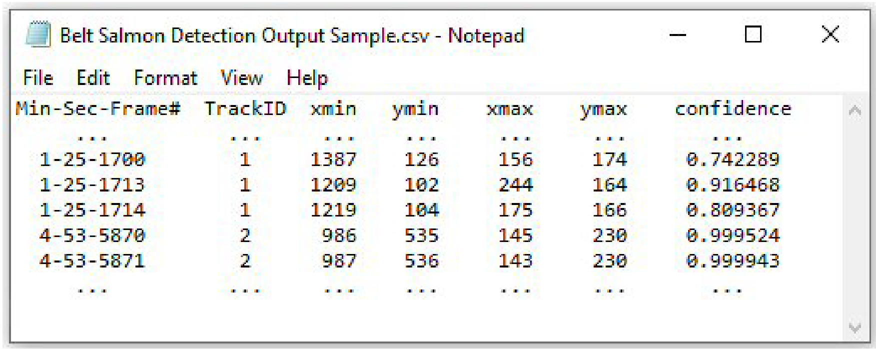



Figure 52. -- Photo images from cropped input video showing salmon detection and tracking; green bounding boxes show a detected salmon as it is tracked (top to bottom).


Belt Salmon Detection Output Sample.csv - Notepad

File Edit Format View Help

| Min-Sec-Frame# | TrackID | xmin | ymin | xmax | ymax | confidence |
|---|---|---|---|---|---|---|
| ... | ... | ... | ... | ... | ... | ... |
| 1-25-1700 | 1 | 1387 | 126 | 156 | 174 | 0.742289 |
| 1-25-1713 | 1 | 1209 | 102 | 244 | 164 | 0.916468 |
| 1-25-1714 | 1 | 1219 | 104 | 175 | 166 | 0.809367 |
| 4-53-5870 | 2 | 986 | 535 | 145 | 230 | 0.999524 |
| 4-53-5871 | 2 | 987 | 536 | 143 | 230 | 0.999943 |
| ... | ... | ... | ... | ... | ... | ... |

Figure 53. -- Example of algorithm CSV output showing a sample of salmon detection and tracking output file (csv file) including the video frame rate and track ID.

The salmon identification algorithm was trained on 162 images of chinook and chum salmon, with 55 images held for testing. These image sets were assembled from multiple collections and lighting differences between collections complicated the discrimination of full body images. Images limited to the tails were found to provide better separation, with only one of the 55 testing images misclassified (one Chinook predicted as chum) for a 98.2% accuracy. An algorithm to isolate salmon tails from salmon images was developed for use with the salmon classification algorithm.

### 6.4 Salmon Detection and Tracking on Plant Belts - Algorithm Review and Results

The success factors for the development and implementation of algorithms for automated salmon detection and tracking in a plant environment are twofold:

1. Measuring accuracy results. The algorithm results need to be accurate to within a predetermined measure of confidence.
2. Measuring processing performance results. The measure of processing performance is determined using the processing benchmarks for running algorithms and the time needed to run the analysis application.

### 6.4.1 Salmon on Belt Algorithm - Accuracy Results

Salmon detector algorithms had false positive detections when analyzing individual frames and each salmon passage generates several frames. To overcome these issues, tracking routines were developed to exclude frame detections that were not part of a sequential series of detections and to combine frames from each salmon into a single event. Focusing on tracks greatly reduced the number of false positives. The initial detector detected about 70% of the frames with salmon when applied to rockfish deliveries and less than 40% when applied to pollock deliveries (Table 22). That algorithm was too slow to consider real-time analysis, so a faster version that also incorporated detection improvements was developed in 2020. That version increased recall and precision at both frame and track levels, except for a drop in the precision of track detections in rockfish deliveries from 98% to 90%. A slower tracking algorithm (TNT) improved those detections to 95%.

Table 22. -- Training and testing results for detecting and tracking salmon in rockfish and pollock deliveries.

| Delivery species | Algorithm | Frame annotations | Frame recall | Frame precision | Track annotations | Track recall | Track precision |
|---|---|---|---|---|---|---|---|
| Rockfish | Initial | 4,194 | 68.3% | 73.2% | 85 | 74.1% | 98.4% |
| Rockfish | Fast | 4,194 | 77.3% | 77.2% | 85 | 94.7% | 89.5% |
| Pollock | Fast | 8,917 | 54.1% | 88.2% | 85 | 76.1% | 73.1% |

For species identification, algorithms focusing on salmon tails distinguished Chinook salmon from chum salmon with 92.8% accuracy with images from a range of collections, including controlled environment chutes. Further development and research will be needed to achieve these types of results with imagery from new views or plants.

### 6.4.2 Salmon on Belt Algorithm - Performance Results

The salmon detection and tracking algorithm application makes use of dedicated GPU processing. Performance was improved by skipping real-time visualization, removing video to image conversion and ROI cropping, and setting a detection interval that skips frames until a detection occurs and then examines all the next 20 frames. With the latest version, 2,391 s video with 15 frames per second (36k frames) took 265 s to process, while an 850 s video with 20 fps (17k frames) took 336 s. While this range of 11% to 40% of processing time to running time ratios indicates considerable variation, both being much less than running time indicates real-time detection may be feasible. However, further testing of the application with in-plant installations is needed to establish real-world performance benchmarks.

## 6.5 Salmon Detection and Tracking on Plant Belts – Discussion, Operational Readiness and Summary

EMI developed species detection and species identification algorithms to validate the plant personnel's bycatch sorting of catcher vessel deliveries at shore side processing plants. These algorithms provide verification of industry bycatch sorting and processing, allowing bycatch reports from plant fish tickets to be used to monitor salmon bycatch. This could replace time consuming and tedious delivery-belt-monitoring by observers that is done in parallel with the plant's sorting or salmon sampling aboard catcher vessels. In this system, one camera detects most of the salmon entering the sorting area. Sorters put all salmon under a second camera, which confirms that all detected salmon are sorted and identifies the species of all salmon. Placement of these two cameras and recording of their video streams would be a minor addition to monitoring systems that most plants already have in place. Recordings from both cameras can be triggered by motion detection to limit storage and processing.

EMI work needed to proceed toward implementation includes running the existing algorithms on video collected from 2018, 2019, and from real-time collections during fishery deliveries. Opportunities for these real-time trials include deliveries in the rockfish and pollock fisheries, either in Dutch Harbor or Kodiak, AK. A key test will be determining whether and how much retraining is needed to apply the detector to new plants.

The processing algorithms currently consist of a linked series of Python code routines and further application development is required for a general application that analyzes both detection and identification video streams. This solution would provide an estimate of the sorting efficiency and a count of salmon by species for all sorted salmon. Output can then be compared to the plant's report for confirmation of events and potential errors. The results of these comparisons will be summarized in a peer-reviewed paper and presented to fisheries management and participants for their considerations of potential implementation. Key implementation questions include who runs the analysis software, how many deliveries need to be analyzed and how those are selected, and what are the performance criteria for use of plant report numbers.

# 7 CREW ON DECK DETECTION AND ACTIVITY TRACKING

Monitoring the location and activity of the crew on the deck of vessels is an important function in fisheries compliance reporting. Crew are surveilled not only to prevent invalid access to restricted areas but also to monitor unauthorized activity such as discarding fish. Surveillance video is recorded on board vessels as part of EM across all gear types, with manual video review being conducted to find and report on potential marine resource violations or to collect catch accounting data. This video review process can be a laborious task with days of footage required to be reviewed. EMI collaborated with staff at the NOAA NMFS West Coast Region (WCR) in 2019 to develop computer vision algorithms to detect crew presence in an area of interest in the video frame and track crew movement to report on activity. Regions of interest include areas with high traffic and those near discard points such as hatches, scuppers, and rails.

Computer vision human presence detection is a mature field (Zhang et al. 2019) allowing for rapid prototyping of the crew detection algorithm. The research method and development lifecycle for this application prototype is like the methods used in other EMI streams but no camera acquisition system was needed to be built or deployed (Fig. 54). Data was acquired through collaborative efforts, labeled and annotated, and algorithms were developed and trained with the labeled data

This detection algorithm has potential applications including but not limited to real-time compliance monitoring, real-time activity monitoring, real-time human presence/absence triggering.

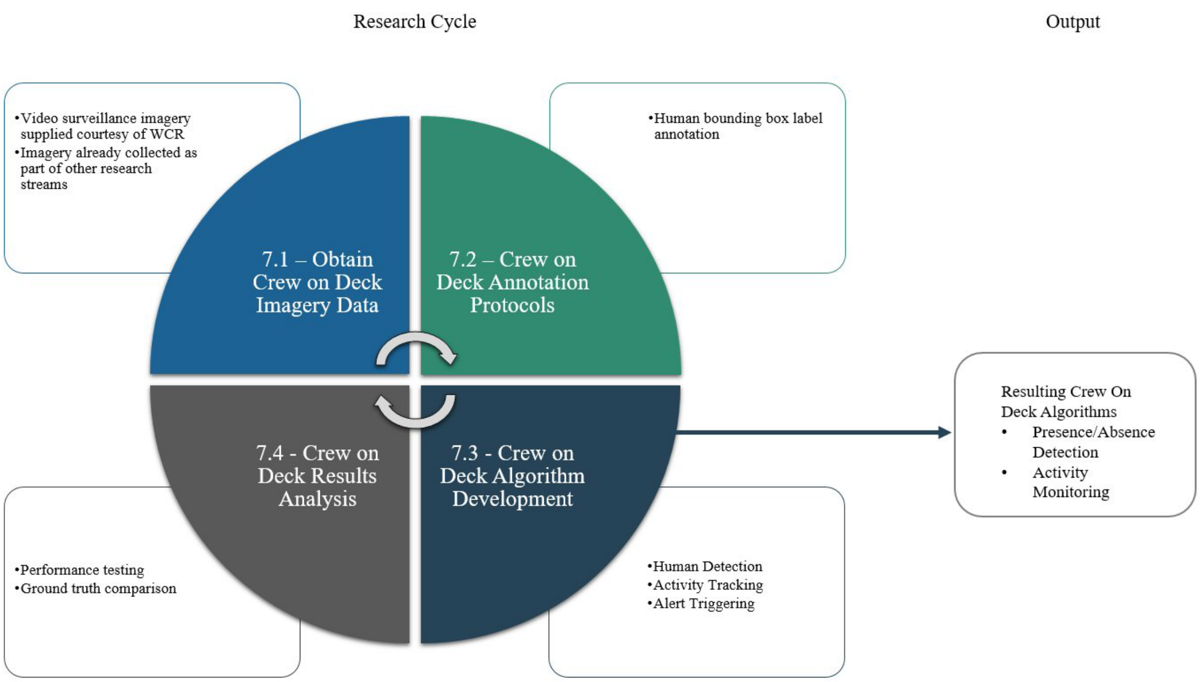



Figure 54. -- Schematic showing the research cycle for detecting crew activity on the deck of fishing vessels.

## 7.1 Crew on Deck Training Data

Training data was supplied by the WCR and consisted of surveillance video from multiple trawls vessels with multiple views and angles of the vessel deck visible in the videos, all from vessels under 100 ft. These videos were acquired from IP cameras. Additionally, it was determined that existing data collected as part of the data collection efforts for EMI Rail (Section 3) could also be utilized for crew detection algorithm training since crew members are standing at the rail in these images. These images were acquired via MV cameras from 2019 EMI Rail deployments aboard FV *Kariel*, FV *Ocean Prowler* and FV *Predator*.

Since there is no single standard surveillance video input type, frame rate, or image resolution deployed in the field, this combination of both the IP camera and MV camera imagery allowed for the developed algorithms to be flexible in its input requirements, supporting multiple input resolutions and framerates.

## 7.2 Crew on Deck Annotation

There was a total of 655,714 single frames of imagery with 33,050 bounding boxes labeled for training: 7,668 frames and 3,047 bounding box labels from WCR data from trawl decks, and 648,046 frames and 30,003 bounding box labels from EMI Rail deployments.

## 7.3 Crew on Deck Algorithm Development

Algorithm development for the Crew on Deck stream consists of detecting the presence of a human within a specific region of interest, and once detected, determining what actions or activity that person is performing.

For the detection of crew in a specified region of interest, the YOLO object detector model was implemented. YOLO ("You only look once", Redmon et al. 2016 and 2018) is a convolutional neural network (CNN) for achieving object detection in real-time. The algorithm applies a single neural network to the full image, and then divides the image into regions and predicts bounding boxes (weighted by the predicted probabilities) for each region. The detector is fast and lightweight, which enables real-time processing and deployment on any system with a graphics processing unit (GPU) A region of interest (ROI) is defined before running the detector. If detections occur outside that ROI, then a trigger alert occurs.

Tracking the activity of a human is more complex that simply detecting human presence. As a proof of concept, the OpenPose model was used to perform human pose estimation. OpenPose is a real-time multi-person system to jointly detect human body, hand, facial, and foot key-points on single images. Limbs and joints of the body are mapped allowing for activity and movement tracking of the crew on deck. While this proof of concept was promising, further research and development is needed. Actions, such as throwing a fish overboard or opening a hatch will need to be defined and modeled with specific training data needed for each action.

## 7.4 Crew on Deck Detector Review and Results

To test the detection algorithm, 19,903 images from six camera angles were run through the algorithm. The human detector achieves 99% precision and 98% recall. Real-time processing was achievable using both high powered GPU machines and NVidia Jetson. Jetson is a low-power system designed for accelerating machine learning applications. Figure 55 below depicts an example of a detection made by the detector. The output of the Crew on Deck detection algorithm is in the form of a csv file, and details the timestamp and the detection bounding box


Figure 55. -- Image frame from an example of Crew on Deck detection. The light blue bounding box is the ROI with orange bounding box being the detection.

## 7.5 Crew on Deck Haul Station Sensor Experiment

With the promising results achieved from the Crew on Deck detection algorithm, an experiment was conducted to use this detector as a sensor trigger for the EMI Rail system (Section 3.1). In developing the EM Rail system to acquire the images for the Rail system data collection, a need arose for a more efficient way of triggering the start and end of hauling activity. While the deployed sensors work to a certain degree, proximity and gear-based sensor results varied between each vessel and season and a standardized hardware strategy could not be achieved between vessels. The addition of a video detection trigger would allow for a standard hardware strategy. For this reason, the logic and output of the Crew on Deck ROI presence/absence detector was expanded for use as a sensor trigger.

When the detection occurs in a ROI for a set amount of time, an alert can be triggered to indicate that the activity (hauling in this case) has started, and when no crew are present for a set amount of time, this can signal that the activity has ended. Camera imagery is already being recorded as part of EMI Rail system; however, this haul sensor would take a real-time video feed from a low powered IP camera as input independent from the EMI Rail stereo cameras. Further research and testing will determine the best implementation and integration strategy for this sensor.

### 7.6 Crew on Deck Discussion, Operational Readiness and Summary

The Crew on Deck detector has shown that achieving real-time human detections in areas of interest on board fishing vessels is achievable with relatively low development costs. While collaboration with WCR ended on the delivery of this proof-of-concept detector, there are still multiple research and operational opportunities available. The first of these being the incorporation of this detector into the video data review stream to monitor compliance. Secondary to this endeavor would be expanding upon the activity tracking capabilities to report on specific movements of the crew. Movements such as discard motions, setting and hauling operations, offload and backload operations can be defined and tracked.

Integrating the detector into existing data review streams will require collaboration and input from data reviewers. The detector in its current form is a script-based application. It is not anticipated that further training of the algorithm will need to be conducted; however, this is dependent on the view and angle of input data. If views are vastly different from those of the training data, then retraining will need to be completed, along with labeling efforts. For direct integration into existing data review streams, the resulting csv file output types need to be defined and managed. A typical use case of the Crew on Deck detector would be to run existing surveillance video and then human reviewing the actions of the crew based on the timestamp detections made. This would be a beneficial use of the detector for monitoring and reporting on specific restricted areas where it would not be common for a detection to occur.

To achieve real-time video analysis on a live feed, collaboration will need to be conducted with system owners on how to integrate the application into those types of live environments. The findings of experiments conducted using the detector as a haul station sensor would provide insight into this type of real-time reporting.

The benefits of researching crew motion activities would benefit the monitoring reporting currently being human reviewed. The ability to automatically track, interpret or predict the actions of the crew on deck allows for true automated crew monitoring; however, this will be a resource-heavy endeavor. It will require defining the types of actions that need to be monitored and will require training imagery on each action and variations of those actions. These images would, in turn, need to be labeled and trained. This data would either need to be mimicked and acquired for further collaboration. Currently, deck video is being reviewed as part of the fixed-gear EM review process to track discard events. Combining this crew activity tracking with the fish detection and tracking for the rail would yield powerful automation for discard tracking.

# 8 SEABIRD SPECIES IDENTIFICATION EXPERIMENT

The purpose of this experiment is to automate identification of birds near and on vessels as well as birds captured on hook and line gears. The outcomes of this experiment can be applied to hook and line fishery, at the rail, deck strikes, and during setting procedures.

Two experiments were conducted using the algorithms previously developed for use on trawl vessels (chute system) and on hook-and-line vessels (Rail system). These experiments focused on using the existing algorithms to identify species of incidentally caught seabirds. The goal of the first experiment was to identify bird species in a controlled environment (the multi-spectrum chute system), while the second experiment investigated identifying seabirds that were near the fishing gear or caught on the line (Rail system).

The multi-spectrum chute system consists of eight machine vision cameras each equipped with a band-pass filter to limit each individual camera to a specific light frequency. This includes the standard RGB, Infrared (IR), and UV light frequencies. Images were captured of 15 different species of seabirds and used to train the existing algorithms. A total of 1,837 images were used for training the system and 213 images were set aside for testing (Figure 56). Based on the test images, identification accuracy was 93% (Table 23). This includes 100% accuracy for commonly caught species including Black-footed Albatross, Northern Fulmar, and Laysan Albatross. Training and testing were conducted using images captured by the standard RGB cameras; images captured under other light frequencies have not been examined up to this point. The training data set is small and dominated by the species commonly encountered in the fishery. A larger variety and quantity of specimens would be needed to generate a more robust algorithm.


Figure 56. -- Example seabird image captured using the multispectral camera chute

Table 23. -- Seabird species identification results (Controlled Environment) showing overall results in addition to species specific testing results.

| **Summary results** | | | |
|---|---|---|---|
| **Mean per class accuracy** | 93% | | |
| **Number of training images** | 1837 | | |
| **Number of testing images** | 213 | | |
| **Higher classification** | **Species** | **Images tested** | **Accuracy** |
| Ardenna | Shearwater Unidentified | 1 | 0% |
| Ardenna | Short-tailed shearwater *(Puffinus tenuirostris)* | 6 | 100% |
| Ardenna | Sooty shearwater *(Ardenna grisea)* | 2 | 50% |
| Fulmar (Fulmarus) | Northern fulmar *(Fulmarus glacialis)* | 53 | 100% |
| Gulls (Larus) | Glaucous gull *(Larus hyperboreus)* | 1 | 0% |
| Gull | Glaucous-winged gull *(Larus glaucescens)* | 3 | 66.7 |
| Gull | Gull unidentified | 1 | 0% |
| Gull | Herring gull *(Larus argentatus)* | 1 | 0% |
| Gull | Large immature gull | 2 | 100% |
| Kittiwakes (Rissa) | Black-legged kittiwake *(Rissa tridactyla)* | 2 | 50% |
| Murre (Uria) | Common murre *(Uria aalge)* | 6 | 100% |
| Murre | Murre unidentified | 3 | 66.70% |
| Murre | Thick-billed murre *(Uria lomvia)* | 2 | 50% |
| North Pacific albatross *(Phoebastria)* | Black-footed albatross *(Phoebastria nigripes)* | 53 | 100% |
| North Pacific albatross | Laysan albatross *(Phoebastria immutabilis)* | 52 | 100% |

After successful completion of the pilot studies for identifying seabirds in the multispectral chute, a mock rail system was used at AFSC to simulate seabirds being caught on the longline and test the rail system's ability to detect and classify seabird bycatch. The goal was to identify bird species using algorithms developed for stereo camera usage. Although birds are not caught in high abundance, when they are caught, it is essential to identify them to species. Since few imagery data existed for instances of seabirds being caught on the line, this scenario needed to be simulated and captured using the EM Innovation stereo Rail system.

For the simulation, 17 different species of birds were used to train and test the system. These included species of Albatross, Northern Fulmar, Shearwater species, and miscellaneous incidental seabirds that had been brought back from the field by observers. Of the total 538,056 images recorded, 18,878 rectangular bounding boxes were created. The hook-and-line detector algorithm was trained with a subset of the dataset; the testing dataset consisted of 89 tracks and 8,868 images and was held out of training data. The identification accuracy for this simulation was 93.25% overall with commonly caught species identified with near 100% accuracy (Table 24).

Moving forward, collecting more specimens will further improve the accuracy of this system. At that point, the bird species identification will be integrated into the standard stereo systems for identifying birds seen on the water near or on the gear. Results of these experiments were presented at the Ninth Meeting of the Seabird Bycatch Working Group (Fitzgerald, 2019).

Figure 57. -- Example seabird image captured using the EMI Rail system at the AFSC.

Table 24. -- Seabird species identification results (Rail system) showing overall results in addition to species specific testing results.

| **Summary results** | | |
|---|---|---|
| **Mean per class accuracy** | 93.25% | |
| **Number of training images** | 8868 | |
| **Number of testing tracks** | 89 | |
| **Species** | **Tracks** | **Accuracy** |
| Black-footed albatross *(Phoebastria nigripes)* | 7 | 100% |
| Black-legged kittiwake *(Rissa tridactyla)* | 5 | 60% |
| Cassin's auklet *(Ptychoramphus aleuticus)* | 6 | 100 |
| Common murre *(Uria aalge)* | 5 | 100% |
| Crested auklet *(Aethia cristatella)* | 10 | 90 |
| Emperor goose *(Chen canagica)* | 2 | 100 |
| Fork-tailed storm petrel *(Oceanodroma furcata)* | 4 | 100 |
| Gull unidentified | 5 | 100% |
| King eider *(Somateria spectabilis)* | 5 | 100 |
| Large immature gull | 4 | 100% |
| Laysan albatross *(Phoebastria immutabilis)* | 5 | 100% |
| Leach's storm petrel *(Oceanodroma leucorhoa)* | 4 | 100% |
| Northern fulmar *(Fulmarus glacialis)* | 7 | 100% |
| Parakeet auklet *(Aethia psittacula)* | 6 | 83% |
| Shearwater unidentified | 10 | 100% |
| Tufted puffin *(Fratercula cirrhata)* | 1 | 100% |
| Whiskered auklet *(Aethia pygmaea)* | 3 | 33% |

# 9 ACKNOWLEDGMENTS

The authors gratefully acknowledge the algorithm development contributions of our partners at the UWEE Information Processing Laboratory; Prof. Jenq-Neng Hwang, Dr. Tsung-Wei Huang, Dr. Gaoang Wang, Jiarui Cai and Jie Mei.

In addition, the authors thank Kresimir Williams, Rick Towler, Matt Dunlap, and Shannon Fitzgerald for their insights, assistance, and input, as well as the staff at the NOAA-AFSC-FMA (Observer Program) and Pacific States Marine Fisheries Commission (PSMFC) EM Program for their continued support.

None of this work would have been possible without the support of and feedback from the captains and crew members of participating vessels and the processing plant owners and staff who assisted in field testing of the EMI systems.

This work was funded in part through NOAA grant Award NA14NMF4370237 to PSMFC.

Much of the funding was provided by the NOAA Fisheries Information System Program.

# 10 CITATIONS


Alaska Fisheries Science Center and Alaska Regional Office. 2019. North Pacific Observer Program 2018 Annual Report. AFSC Processed Rep. 2019-04, 148 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Bonney, J., and K. McGauley. 2008. Testing the use of electronic monitoring to quantify at-sea halibut discards in the central Gulf of Alaska rockfish fishery. Alaska Groundfish Databank. EFP 07-02 Final Report. Available at https://www.fisheries.noaa.gov/webdam/download/84770167

Brodsky, P. 2017. Automated Measurement of Bycaught Halibut from Realtime Video, Alaska Marine Science Symposium, January 2017, Anchorage, AK. http://www.apl.washington.edu/project/projects/fish_measurement/pdfs/AMSS_2017_Brodsky_Poster.pdf

Fitzgerald, S., F. Wallace, K. Magrane. 2019. Improving seabird species identification in electronic monitoring applications using machine learning systems. In: ACAP - Ninth Meeting of the Seabird Bycatch Working Group. ACAP SBWG9 Inf 21, Florianópolis, Brazil.

Huang, T., J. Hwang, S. Romain and F. Wallace. 2019. Fish tracking and segmentation from stereo videos on the wild sea surface for electronic monitoring of rail fishing. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 10, pp. 3146-3158, Oct. 2019, doi: 10.1109/TCSVT.2018.2872575.

Huang, T., J. Hwang and C. S. Rose. 2016. Chute based automated fish length measurement and water drop detection. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016. pp. 1906-1910, doi: 10.1109/ICASSP.2016.7472008.

Huang, T., J. Hwang, S. Romain and F. Wallace. 2019. Recognizing fish species captured live on wild sea surface in videos by deep metric learning with a temporal constraint. *IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, 2019. pp. 3407-3411, doi: 10.1109/ICIP.2019.8803592.

Huang, T. 2019. Automatic Video Analysis for Electronic Monitoring of Fishery Activities. (Accession No. 2019-10-15T22:57:37Z) Doctoral dissertation, University of Washington. http://hdl.handle.net/1773/44796

Lin, T. 2016. LabelImg - A graphical image annotation tool and label object bounding boxes in images. GitHub repository; https://github.com/tzutalin/labelImg

Liu, W. 2016. SSD: Single Shot MultiBox Detector." European Conf. on Computer Vision (ECCV), 2016.

Loefflad, M. R., F. R. Wallace, J. Mondragon, J. Watson, and G. A. Harrington. 2014. Strategic plan for electronic monitoring and electronic reporting in the North Pacific. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-276, 52 p. Available at https://repository.library.noaa.gov/view/noaa/4691.NMFS (National Marine Fisheries Service). 2017. 2018 Annual Deployment Plan for Observers and Electronic Monitoring in the Groundfish and Halibut Fisheries off Alaska. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. Available at https://www.fisheries.noaa.gov/resource/document/2018-annual-deployment-plan-observers-and-electronic-monitoring

NMFS 2020. Fisheries of the United States, 2018. U.S. Department of Commerce, NOAA Current Fishery Statistics No. 2018 Available at: https://www.fisheries.noaa.gov/national/commercial-fishing/fisheries-united-states-2018 A

NOAA 2019. NOAA Fisheries Strategic Plan 2019-2022. Available at: https://media.fisheries.noaa.gov/dam-migration/noaa_strategicplan_2019_singlesv5.pdf

Redmon, J., and A. Farhadi (2018) YOLOv3: An Incremental Improvement. Computer Science, arXiv: 1804.02767. Available at http://arxiv.org/abs/1804.02767

Redmon, J., S. Divvala, R. Girshick and A. Farhadi, (2016). You only look once: Unified, real-time object detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

Redmon, J., and A. Farhadi. (2018) YOLOv3: An Incremental Improvement. Computer Science, arXiv: 1804.02767. Available at http://arxiv.org/abs/1804.02767

Rose, C.S., J. Nielsen, J. Gauvin, T. Loher, S. Sethi, A. Seitz, M. Courtney, and P. Drobny 2019. Survival outcome patterns revealed by deploying advanced tags in quantity (160): Pacific halibut (*Hippoglossus stenolepis*) survivals after release from trawl catches through expedited sorting. Can. J. Fish. Aquat. Sci. 76: 2215–2224.

Stevens, G.B., G.A. Clendening W. Waters, S. Weintraub, W. Redden, C. Redden, R.W. Srack, and J.M. Srack, 2013. Stereoscopic measurement system and method US 8,345,953 B2 Available at https://patents.google.com/patent/US8345953

Wallace, F., K. Williams, R. Towler, and K. McGauley. 2015. Innovative Camera Applications for Electronic Monitoring. In: G.H. Kruse, H.C. An, J. DiCosimo, C.A. Eischens, G.S. Gislason, D.N. McBride, C.S. Rose, and C.E. Siddon (eds.), Fisheries Bycatch: Global Issues and Creative Solutions. Alaska Sea Grant, University of Alaska Fairbanks. http://doi.org/10.4027/fbgics.2015.06

Wang, G., J.-N. Hwang, K. Williams, F. Wallace, and C. S. Rose 2016. Shrinking encoding with two-level codebook learning for fine-grained fish recognition, in *Proc. 2nd Workshop Compututer Vision for Analysis of Underwater Imagery (*CVAU*I)*, Dec. 2016, pp. 31–36. DOI 10.1109/CVAUI.2016.18.

Wang, G., J. Hwang, C. Rose and F. Wallace. 2019. Uncertainty-based active learning via sparse modeling for image classification. *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 316-329, Jan. 2019, doi: 10.1109/TIP.2018.2867913.

Wang, G., J. Hwang, Y. Xu, F. Wallace and C. S. Rose. 2018. Coarse-to-fine segmentation refinement and missing shape recovery for halibut fish. *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Anaheim, CA, USA, 2018, pp. 370-374, doi: 10.1109/GlobalSIP.2018.8646442.

Wildlife and Fisheries, 50 C.F.R. § 679.51 (2021)

Williams, K., R. Towler, and C. Wilson.  2010. Cam-trawl: a combination trawl and stereo-camera system. Sea Technology, 51(12), pp.45-50.

Williams, K., R.H. Towler, P. Goddard, R.E. Wilborn and C.N. Rooper. 2016. *Sebastes* Stereo Image Analysis Software. AFSC Processed Rep. 2016-03, 42 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Zhang, H.-B., Y.-X  Zhang,; B.-N. Zhong, Q. Lei, L. Yang, J.-X. Du, D.-S. Chen. 2019. A Comprehensive Survey of Vision-Based Human Action Recognition Methods- Sensors. Available at https://www.mdpi.com/1424-8220/19/5/1005

U.S. Secretary of Commerce
Gina M. Raimondo

Under Secretary of Commerce for
Oceans and Atmosphere
Dr. Richard W. Spinrad

Assistant Administrator,
National Marine Fisheries Service.
Janet Coit

March 2023

www.fisheries.noaa.gov

OFFICIAL BUSINESS

National Marine
Fisheries Service
Alaska Fisheries Science Center
7600 Sand Point Way N.E.
Seattle, WA 98115 6349