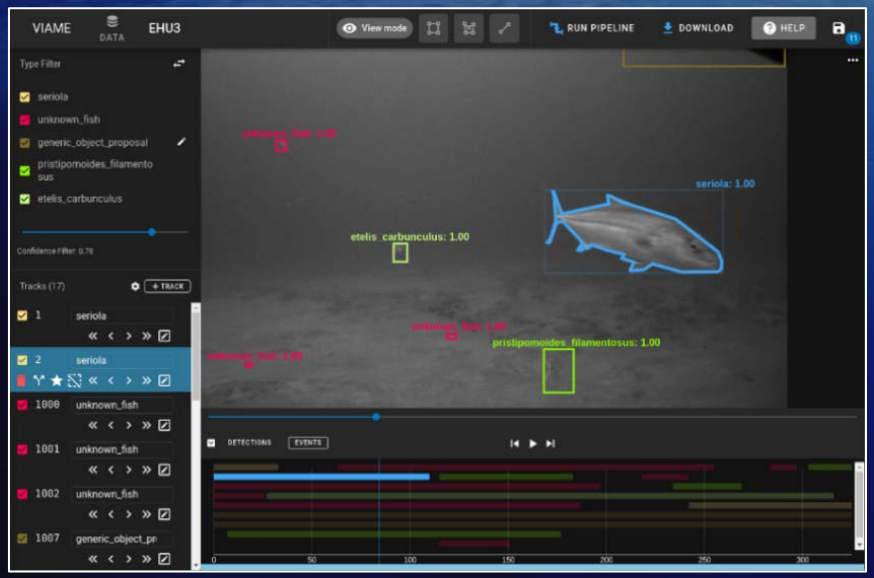
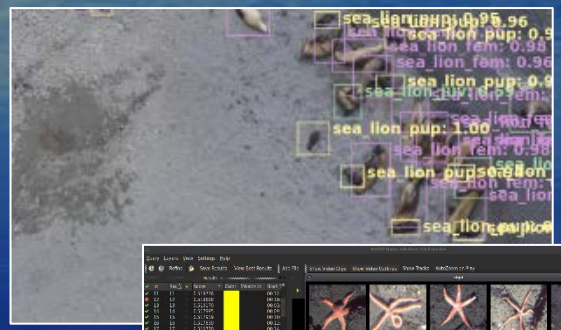
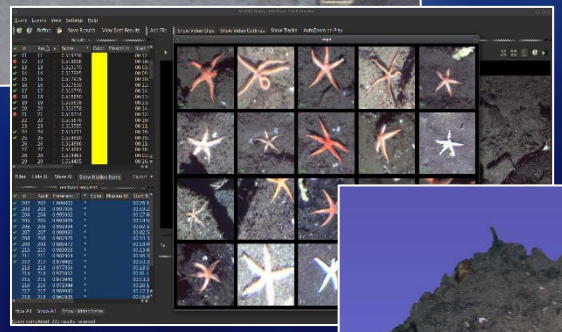






Video and Image Analytics for Marine Environments (VIAME)

An Open-Source, Do-It-Yourself AI Toolkit

Matt Dawkins (Speaker), Senior R&D Engineer, matt.dawkins@kitware.com

Other Developers: Aashish Chaudhary, Jon Crall, Matthew Woehlke, Bryon Lewis, Jacob Nesbitt, Brandon Davis, Neal Siekierski, Anthony Hoogs, Linus Sherrill, Matt Brown, Betsy McPhail, Kyle Edwards, Matt Leotta, Rusty Blue

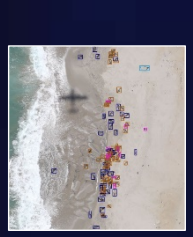
NOAA: Benjamin Richards, Dvora Hart, George (Randy) Cutter, Elizabeth Clarke, Charles Thompson, Kresimir Williams, Bill Michaels, Erin Moreland, Katie Sweeney, Abigail Powell

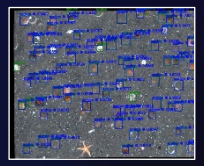
CFF: Liese Siemann

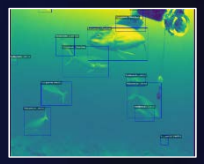
What is VIAME?

- A do-it-yourself (DIY) AI toolkit which can be applied to multiple types of imagery or video
- Can be run by people with no programming or machine learning background in both web and desktop interfaces
- Released as fully open-source with a permissive license, see viametoolkit.org
- Specializations to maritime processing such as motion fusion, stereo measurement, image enhancement, and object tracking which other software (e.g. Amazon SageMaker) lack

Object Detection



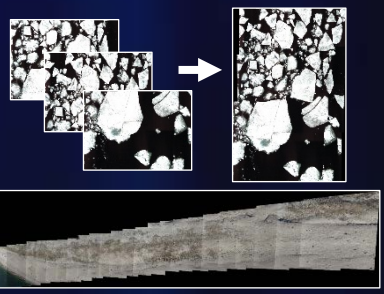




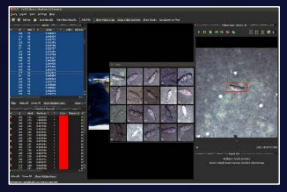
Object Tracking

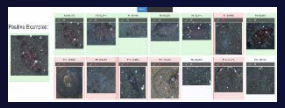
Image Enhancement

Image Registration and Mosaicing

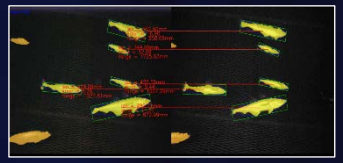


Video Search and Rapid Model Generation

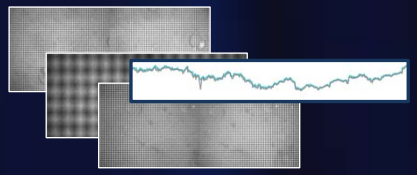




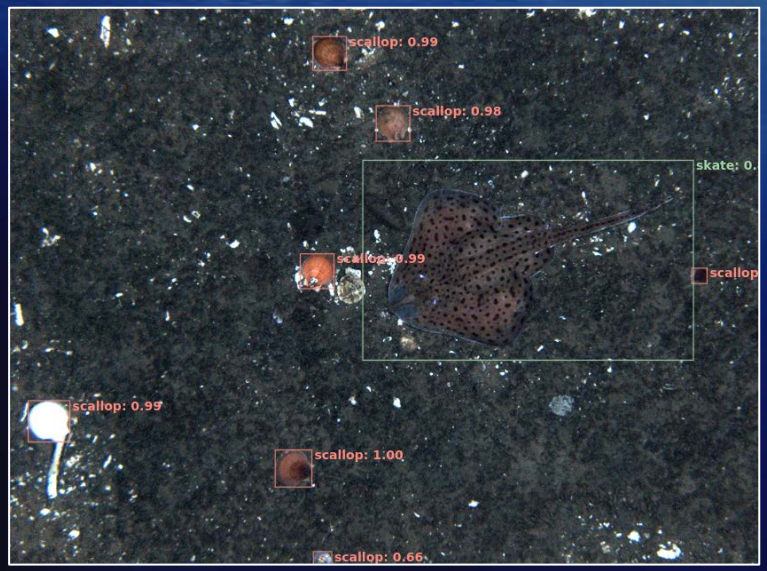
Stereo Measurement

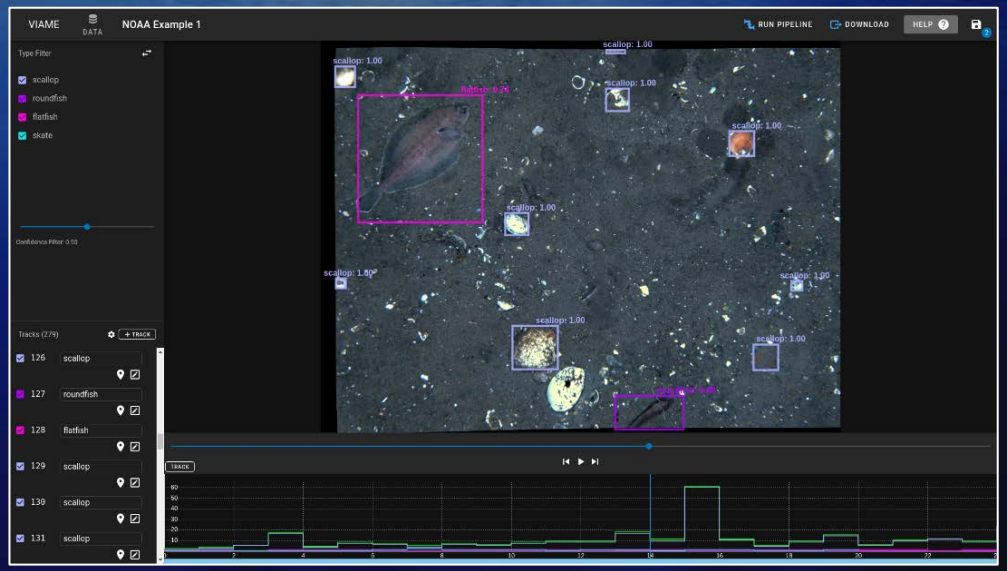


Calibration, 3D and Altitude Estimation

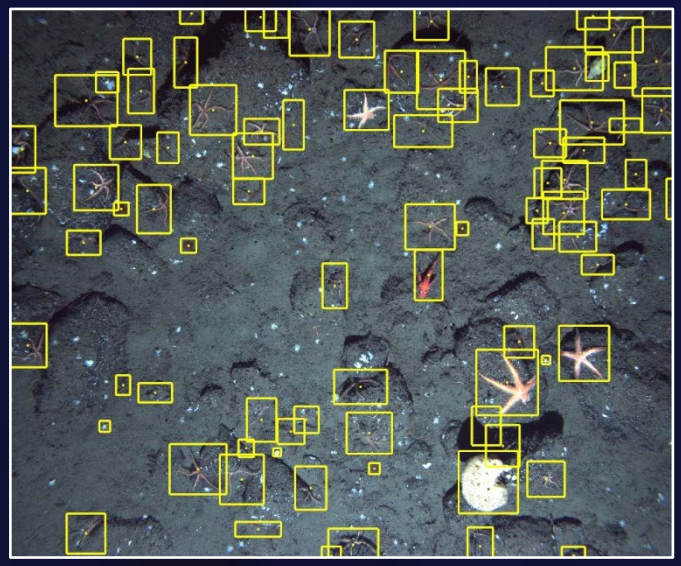


Downward-Facing Underwater Cameras (AUVs, Trawls)





Source: HabCam (NEFSC, CFF, WHOI)



Source: NWFSC AUV Data

Source: NWFSC AUV Data

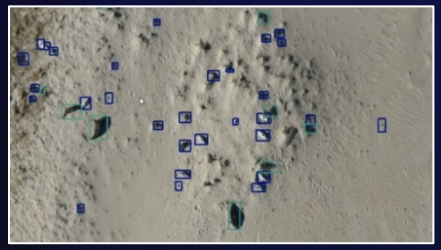
Source: UK P. Maximus Images

Aerial Surveys (Manned Fixed-Wings, UAVs)

Source: AKFSC MML Stellar Sea Lion

Source: AKFSC MML Arctic Seal

Source: AKFSC MML Stellar Sea Lion

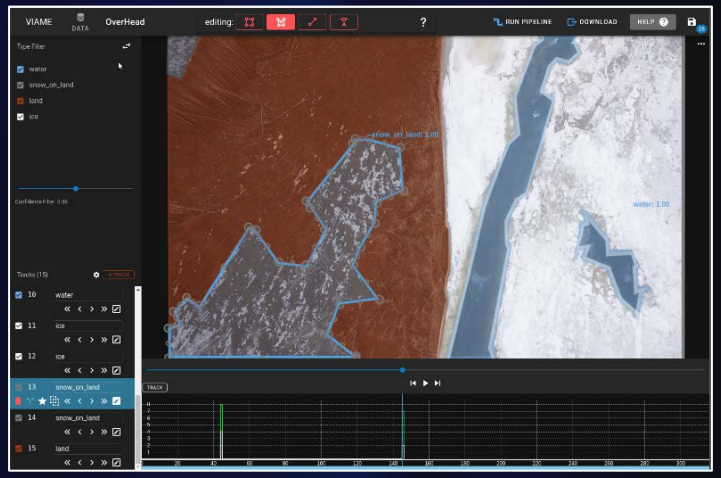


Source: NEFSC Gray Seal

Source: SWFSC

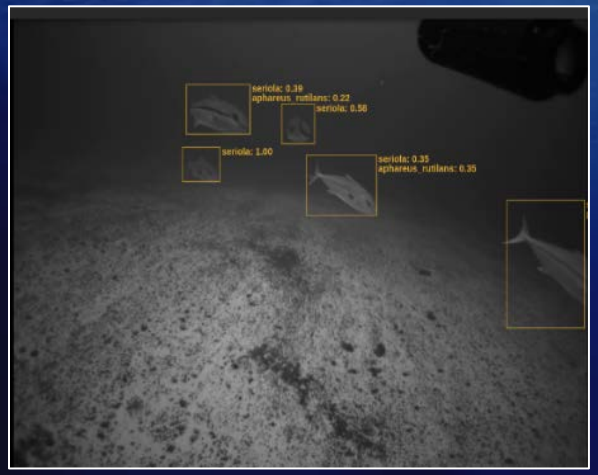
Source: NEFSC Harbor Seal

Source: SWFSC Penguin Aerial Data



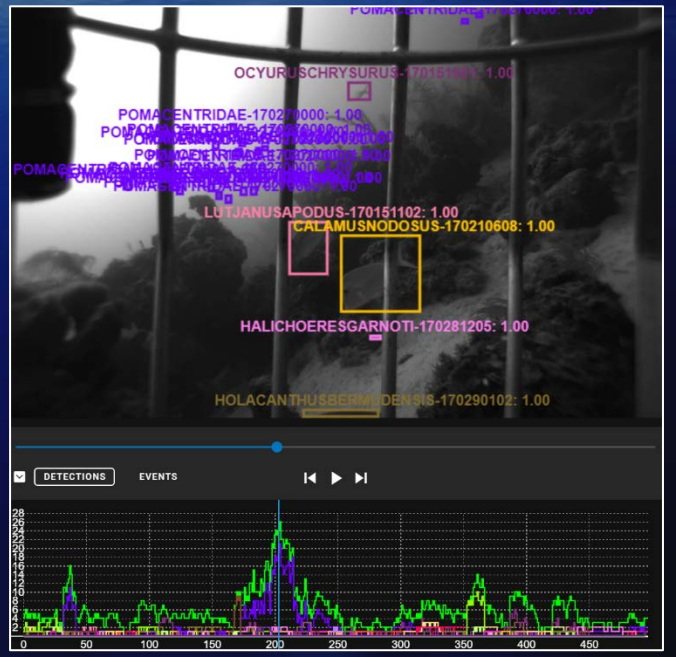
Sources: University of Alaska Data Scene Segmentation

Outward Facing Cameras (Underwater, Ship-Based)



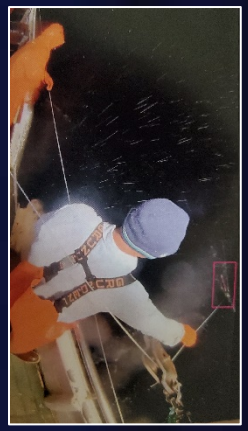
Source: PIFSC MOUSS

Source: SWFSC Penguin Cam



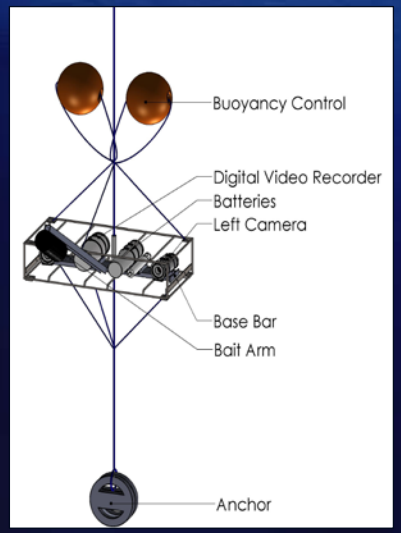
Source: SEFSC Quadcam

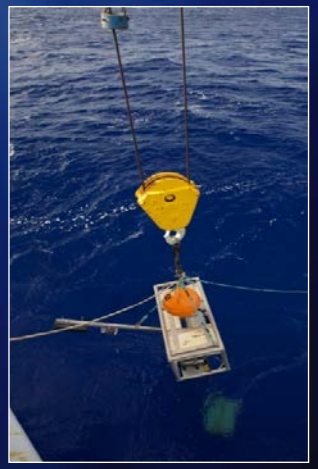
Source: PIFSC MOUSS



Source: PIFSC EM Data

Example Platforms





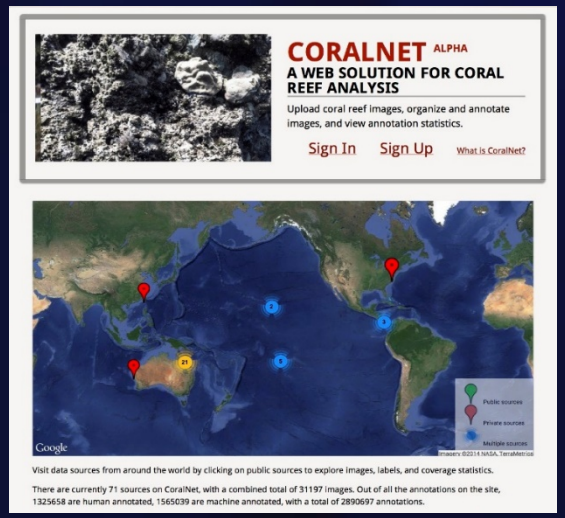
NMFS Strategic Initiative on Automated Image Analysis

Mission: Develop guidelines, set priorities, and fund projects to develop broad-scale, standardized, and efficient automated analysis of still and video imagery for use in underwater stock assessment

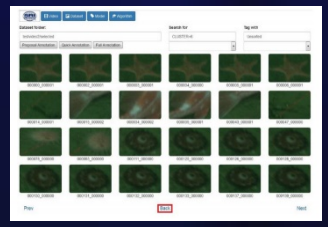
- Benjamin Richards *(chair)*
  NOAA Pacific Islands Fisheries Science Center
- Alexandra Branzan Albu
  University of Victoria
- *Elizabeth Clarke*
  NOAA Northwest Fisheries Science Center
- George "Randy" Cutter
  NOAA Southwest Fisheries Science Center
- Duane Edgington
  Monterey Bay Aquarium Research Institute
- Dvora Hart
  NOAA Northeast Fisheries Science Center
- Anthony Hoogs
  Kitware, Inc.
- David Kriegman
  University of California, San Diego

Funded VIAME and CoralNet from 2015 to present

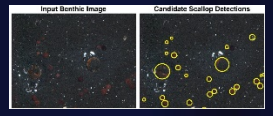
- *Clay Kunz*
  Google
- Michael Piacentino
  SRI International
- Lakshman Prasad
  Los Alamos National Laboratory
- Charles Thompson
  NOAA Southeast Fisheries Science Center
- Kresimir Williams
  NOAA Alaska Fisheries Science Center

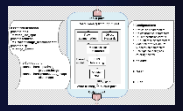


http://coralnet.ucsd.edu, D. Kreigman, UCSD



Flask, Michael Piacentino, SRI





Misc. Analytics (e.g. LANL)

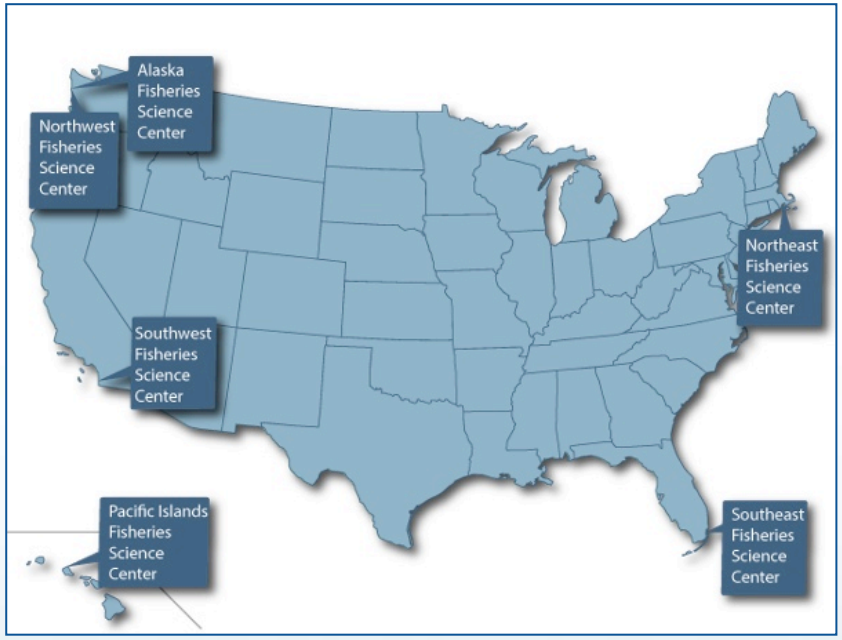
2019 Department of Commerce Gold Medal Awarded to NOAA Members of AIASI for VIAME and CoralNet

Other Funded Initiatives

Kitware

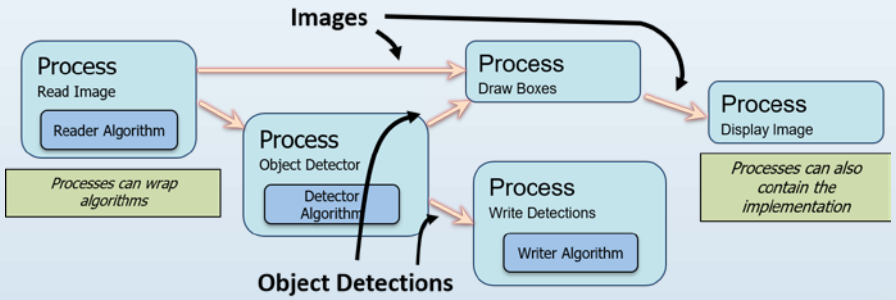
- Collaborative software R&D: algorithms & applications, image & data analysis, support & training
- Best known for open source toolkits and applications
- 150+ employees:
  - ⅓ masters
  - ⅓ PhD
- Founded in 1998
- Offices in Albany, NY; Chapel Hill, NC; Santa Fe, NM; Minneapolis, MN; Arlington, VA; Lyon, France

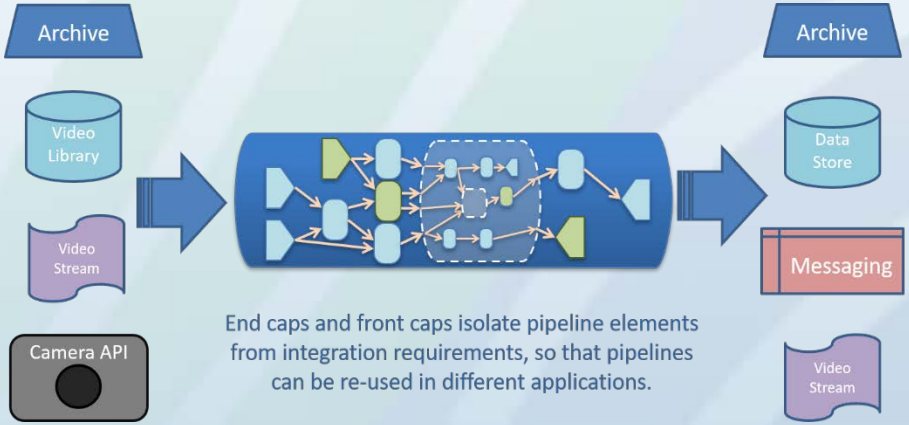
History – Algorithm Integration Platform



Capability	Primary data source	POC	Stereo calibration	Stereo processing	Video	Color, contrast correction	scallop detection	fish detection	fish length, sizing	fish tracking	fish classification	anomaly det.	habitat classification	image segmentation
NW SC CamTrawl	Cam Trawl	Williams	yes	yes	4 Hz	no, grayscale		yes	automatic	yes	yes			
ROV video fish detection and tracking	SWFSC ROV video	Cutter	no	no	30 Hz	yes		yes, DPM (UW)	no	yes (UW student)	desired	desired		
ROV stereo fish measurement	SWFSC ROV GigE stereo	Cutter	yes	yes	2-4 Hz	yes		no	manual	no				
WHOI/NEFSC scallop detector	HABCAM towed rig	Dvora	yes	yes	no	yes	yes							
RPI/Kitware scallop detector	HABCAM towed rig	Hoogs	no	no	no	yes	yes							
SRI fish detection, classification, size	PI FSC MOUSS/BotCam	Ben/Mike	yes, accept cal files	yes	30 Hz	no, grayscale		yes		yes	yes			
SEFSC stereo proc	Drop cams from SEFSC	Thompson	yes	yes	yes			yes, basic background	manual	no	no			
Toyon SBIR I	Drop cams from SEFSC	Thompson	yes	yes	yes			yes, basic HOG	manual	yes	yes			
LANL segmentation and shape analysis	HABCAM towed rig	Laksman	no	yes	no	no	yes	yes	no	no	yes	yes (image)	Yes	yes (polygonal)
Toyon SBIR II	Still Images AUV, drop, towed	Clarke	yes	yes	no	yes (Hanu)		yes	yes	no	yes			
WHOI/NEFSC habitat classifier	HABCAM towed rig	Dvora	yes	yes	no	yes							yes	
NWFSC clustering	AUV and MOUSS	Clarke	no		no	no						yes	partially	

Green: well-implemented; quantified, comparative performance assessment; ready for integration
Yellow: Existing implementation as mature research code; some performance quantification
Red: preliminary research code with ongoing work against major problems
Gray: idea or concept; no implementation

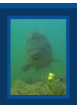




- Base classes for common operations (image filters, object detectors, trackers, …)
- Derivation of base classes in C/C++, CUDA, Python, or Matlab
- Backend coded in C++ for efficiency, automatically multi-threaded

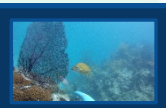
Kitware

Three Detection Workflows

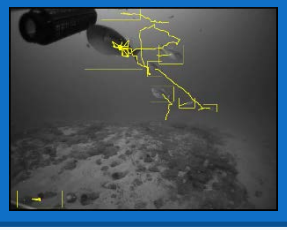




Input Data

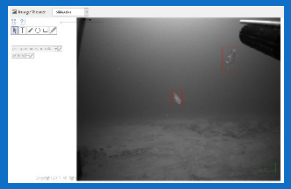




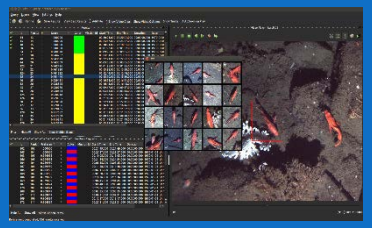
Running Existing Detectors and Trackers (Traditional or Deep)



New Detector Training via Deep Learning

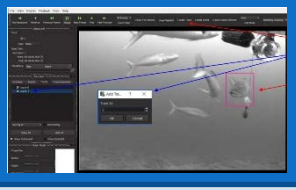



New Classifier Training via IQR



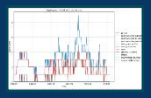
New Annotations or Annotation Correction

Fix Boxes, Labels, or Masks

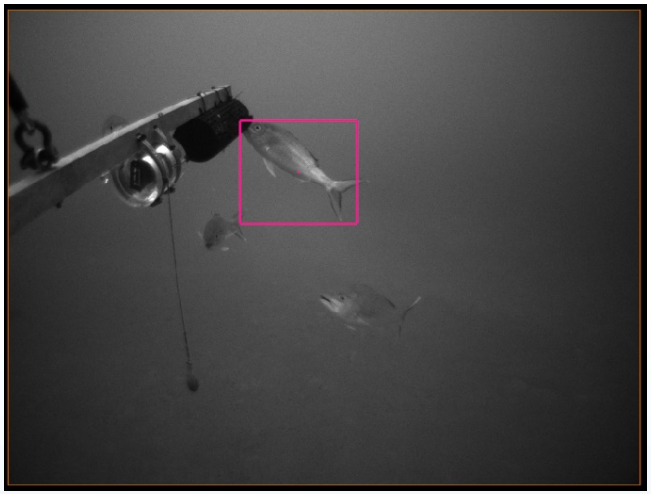


Higher-Level Analytics

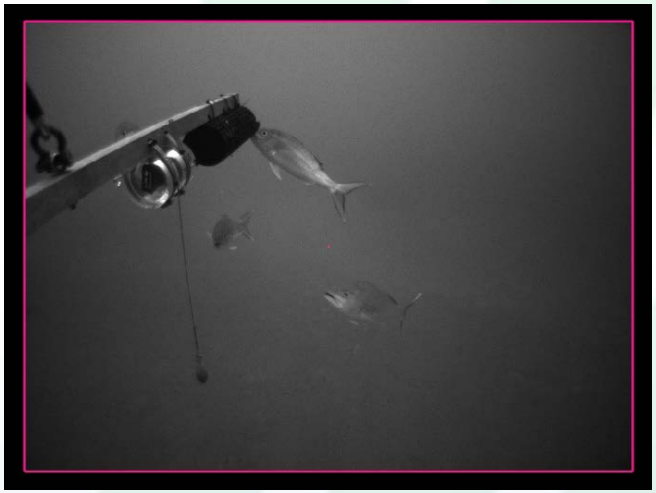
(Automated Detection Quantities, Heatmaps, Occurrence vs Time Plots)



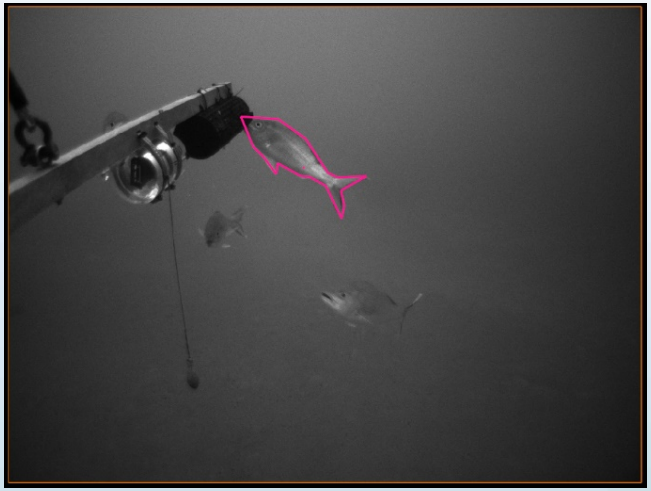
Types of Annotation and Detection Models



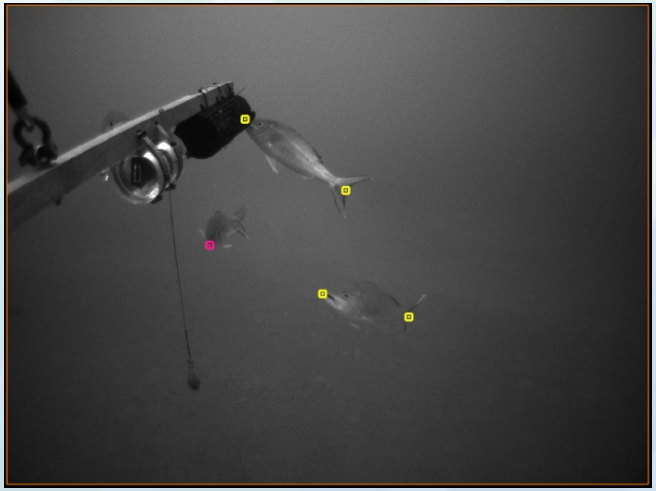
Box-Level



Frame-Level



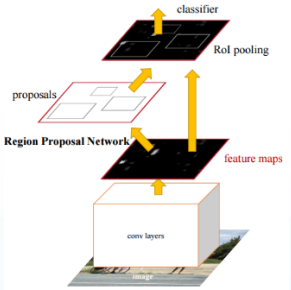
Pixel-Level

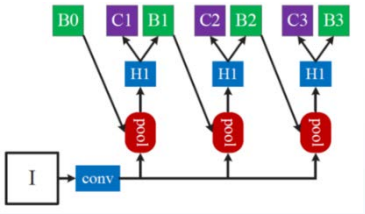


Keypoints

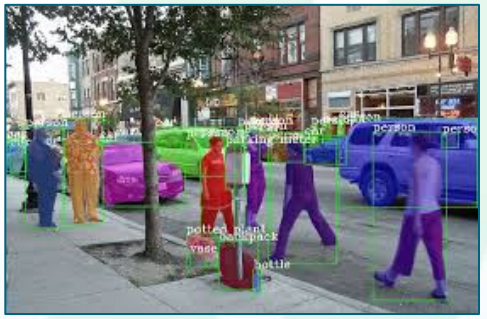
Baseline Object Detectors

Cascade Faster R-CNN [1]

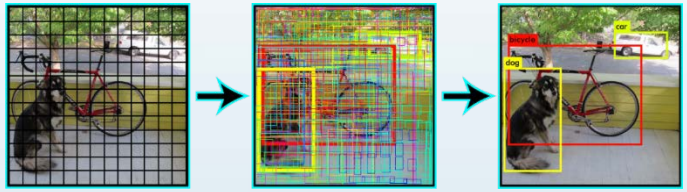




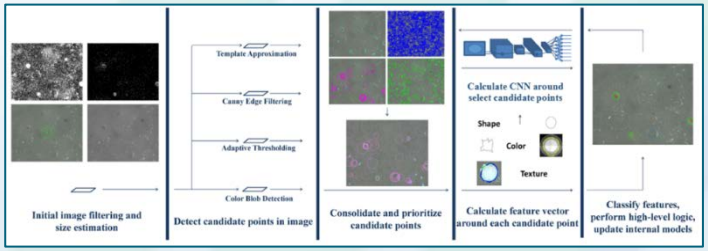
Cascade Mask Faster R-CNN [2]



YOLOv3 and v4 [3,4]



Scallop-TK



VIAME contains multiple baseline general purpose detectors from the larger computer vision community for wide applicability, but then specializations and other functionality added specific to domains of interest

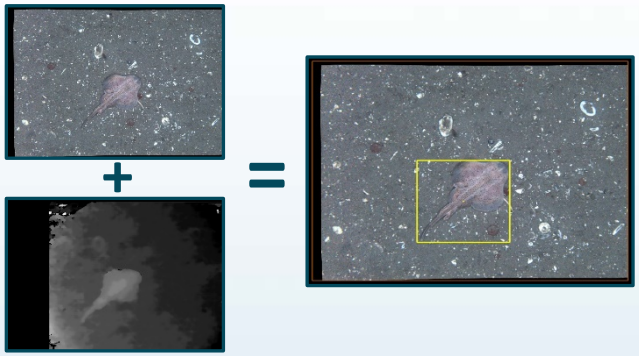
[1] Cai, Zhaowei, et al. "Cascade R-CNN: Delving into High Quality Object Detection." CVPR 2018.
[2] Chen, Kai et al. “MMDetection: Open MMLab Detection Toolbox and Benchmark.” arXiv preprint 2020.
[3] Redmon, Joseph, and Ali Farhadi. "YOLOv3: An Incremental Improvement." arXiv preprint 2018.
[4] Bochkovskiy, Alexey et al. "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv preprint 2020.

Automatic Parameter Optimizations

- Automatically handle LR stepping based on validation loss, running multiple hyperparameter sets, and early stopping criteria
  - NetHarn: https://gitlab.kitware.com/computer-vision/netharn
- Automatically choose whether to grid detectors over image
- Utilities to turn dot annotation into boxes

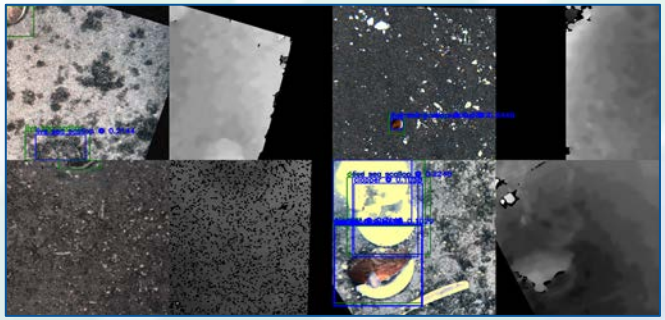
Auxiliary Data Fusion

Fuse depth and motion maps into object detectors



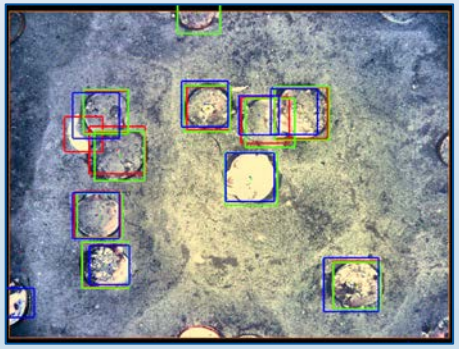
Extra Augmentation

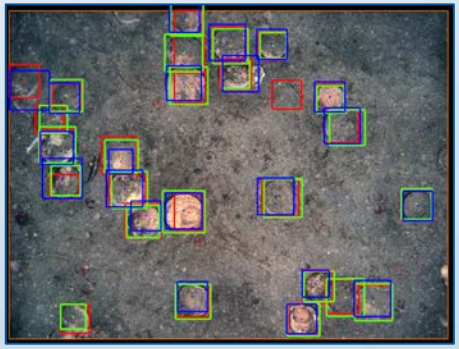
Selectively augment channels differently



Ensemble Classifiers

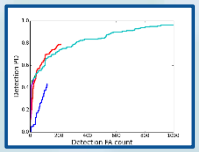
Fuse output of detectors from different frameworks





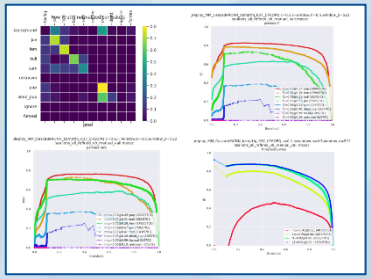
Scoring Utilities

Evaluation via ROCs, PRCs, track metrics, …



```
Detection-Pd: 0.791209
Detection-FA: 213
Detection-PFA: 0.515738
Frame-NFAR: not computed
Track-Pd: 0.791209
Track-FA: 213
Computed-track-PFA: 0.515738
Track-NFAR: not computed
Avg track (cont., purity ): 1.34, 1
Avg target (cont., purity ): 1.47, 0.79
Track-frame-precision: 0.5

DEPTH_Hash : "a2123cde"
```

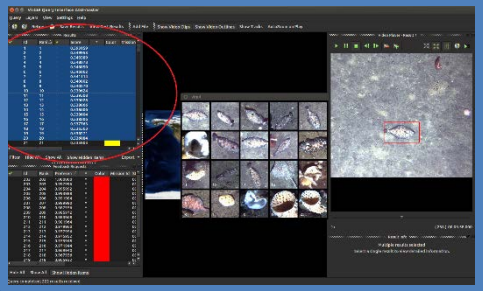


Kitware

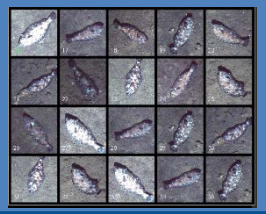
Interactive Search and Rapid Model Generation

Generate Image Descriptors around Detections

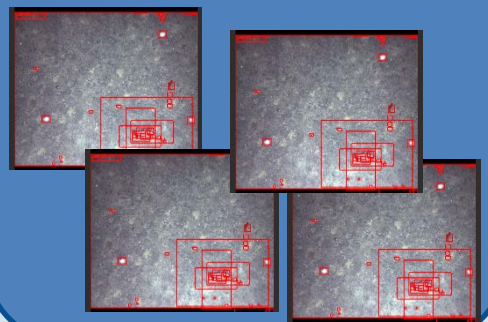
Query Image Archive using Image Exemplar



Create Classifier using Interactive Query Refinement



Generate Generic Object Detections

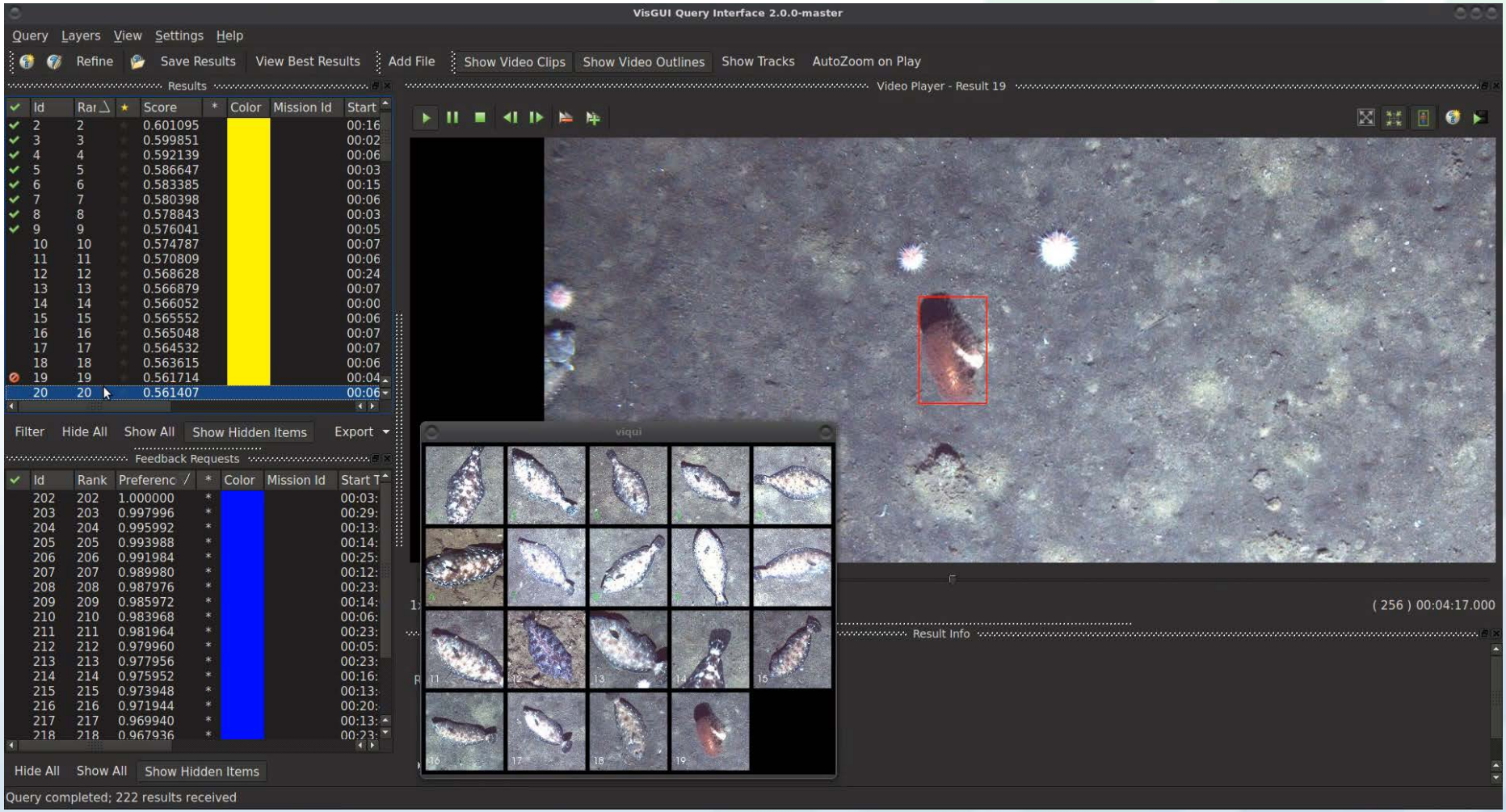


Data Storage via DB

Run Detector and Classifier on New Images

Kitware

Interactive Search and Rapid Model Generation

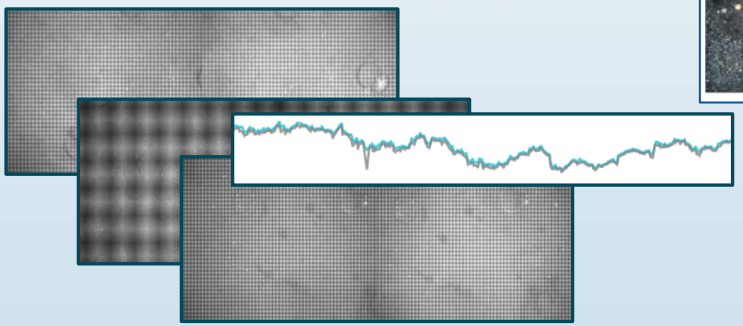


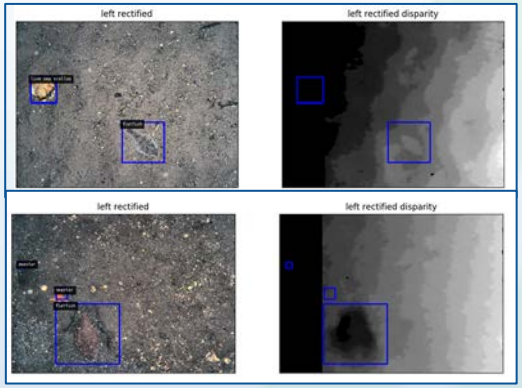
User provides initial image query:

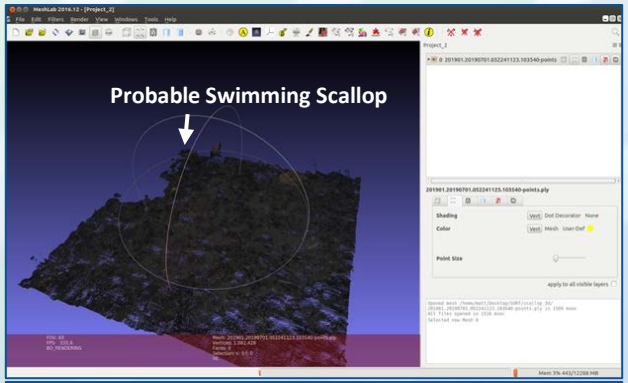
User corrects system returns on subsequent iterations through iterative query refinement (IQR)

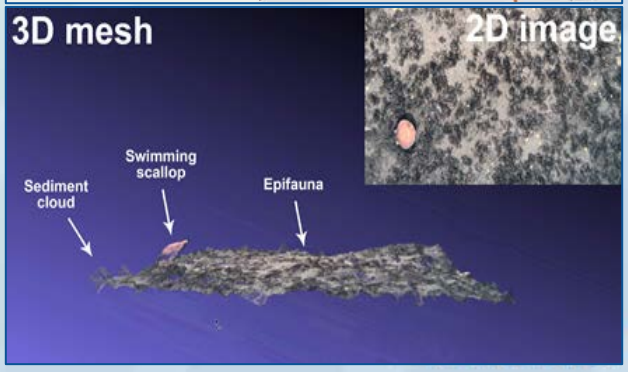
Kitware

Enhancement, Calibration and Depth Estimation



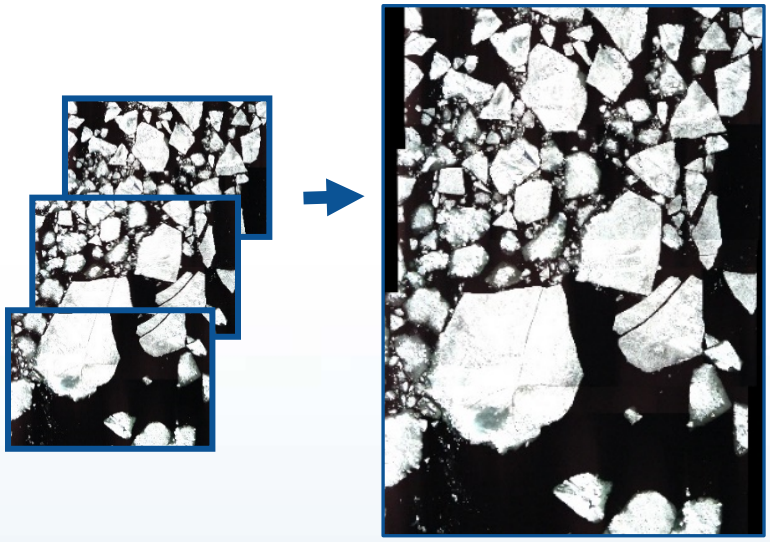


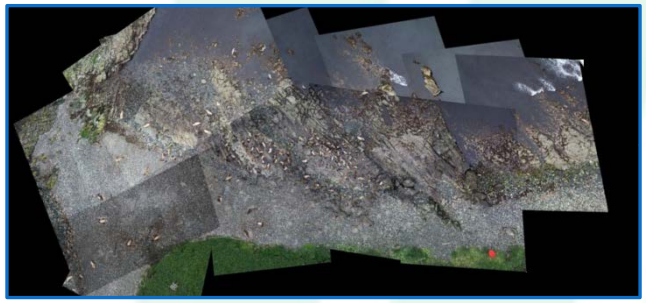


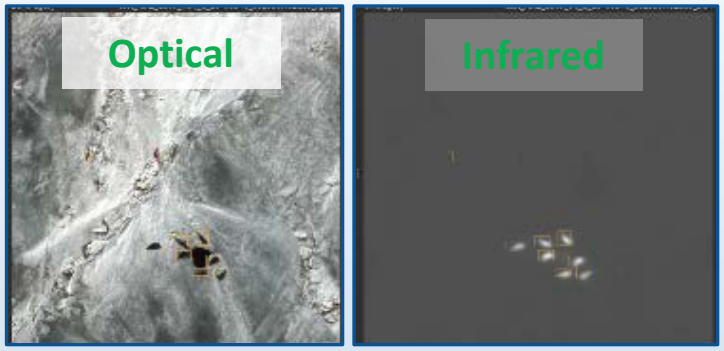


Work performed in conjunction with Coonamessett Farm Foundation (CFF) and NEFSC

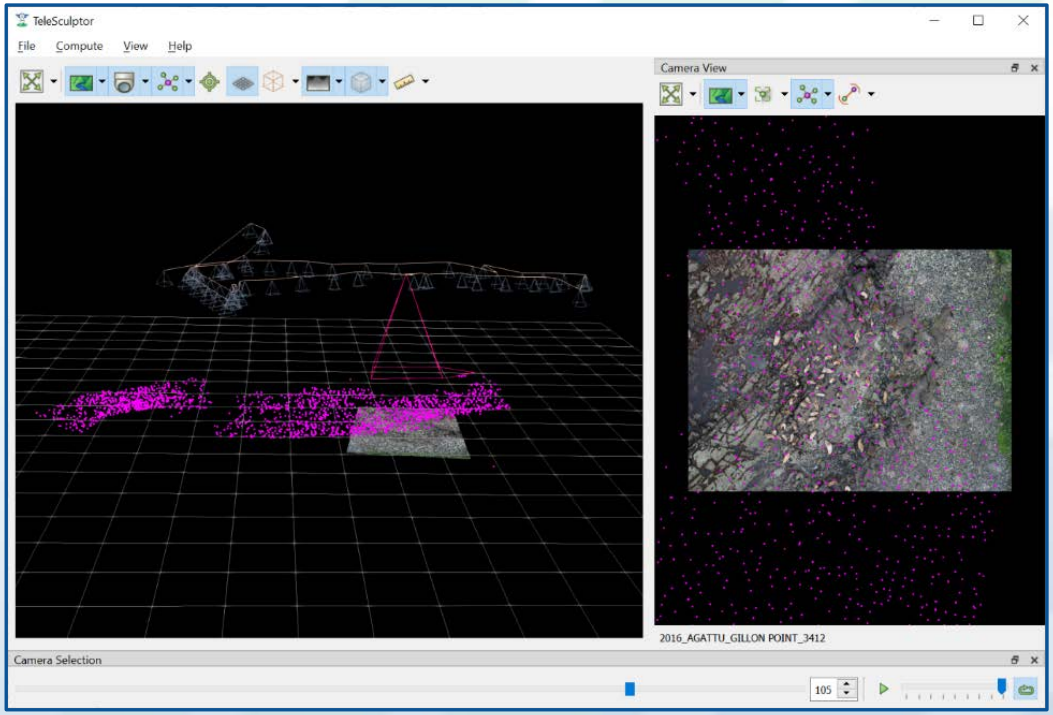
Registration and Mosaicing







Multi-Camera and Multi-Modality Registration



Above Source: TeleSculptor Open-Source 3D Geometry, https://github.com/Kitware/TeleSculptor

Contact: matt.leotta@kitware.com

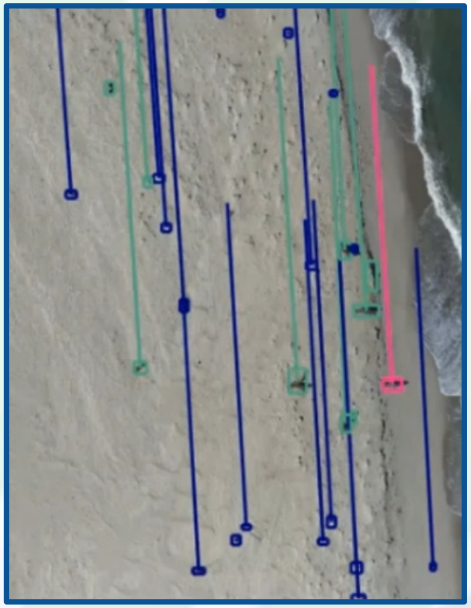
Work performed in conjunction with NOAA AKFSC Marine Mammals Lab

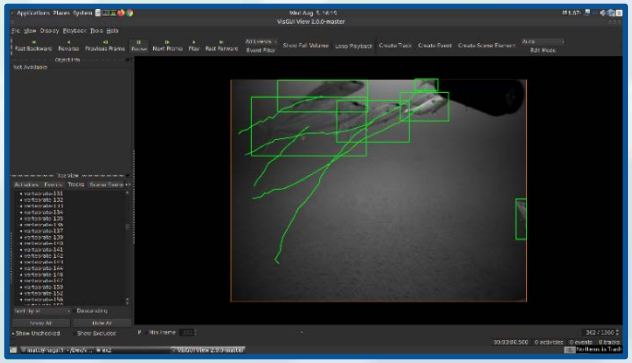
Kitware

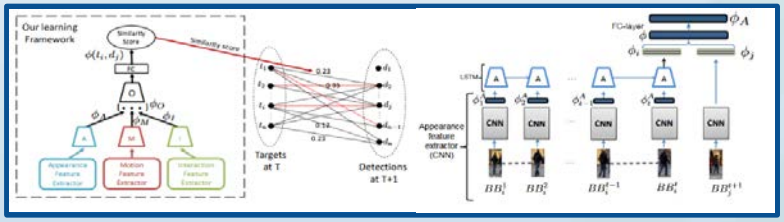
Object Tracking

Have integrated multiple trackers into VIAME:

- Registration-Only Based (Aerial Pinniped)
- Deep Learning Based LSTM Tracker Detector Linkers [5]
- Non-Deep Learning Kalman Filter Detector Linkers
- Single Target Trackers (e.g. SiamRPN++ [6]) for annotation assist (below)





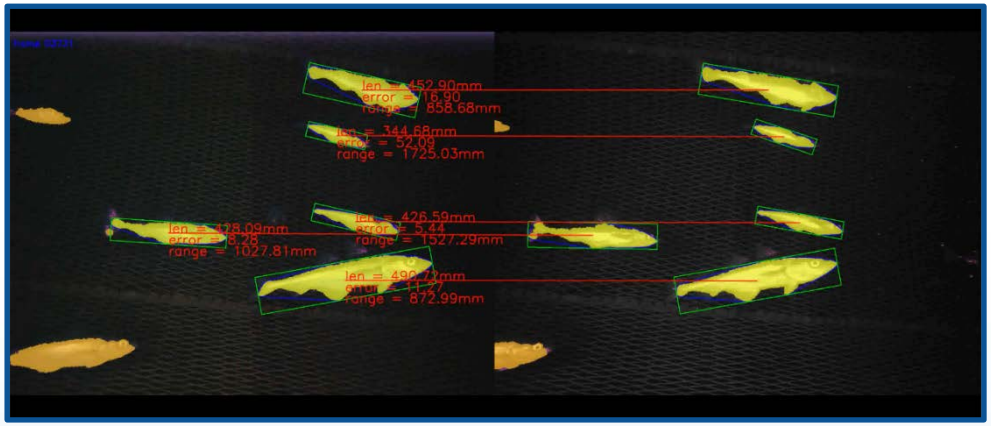


Code developed in conjunction with Air Force Research Laboratory (PRR# 88ABW-2019-4904, Distribution Statement A - Approved for public release: distribution unlimited).

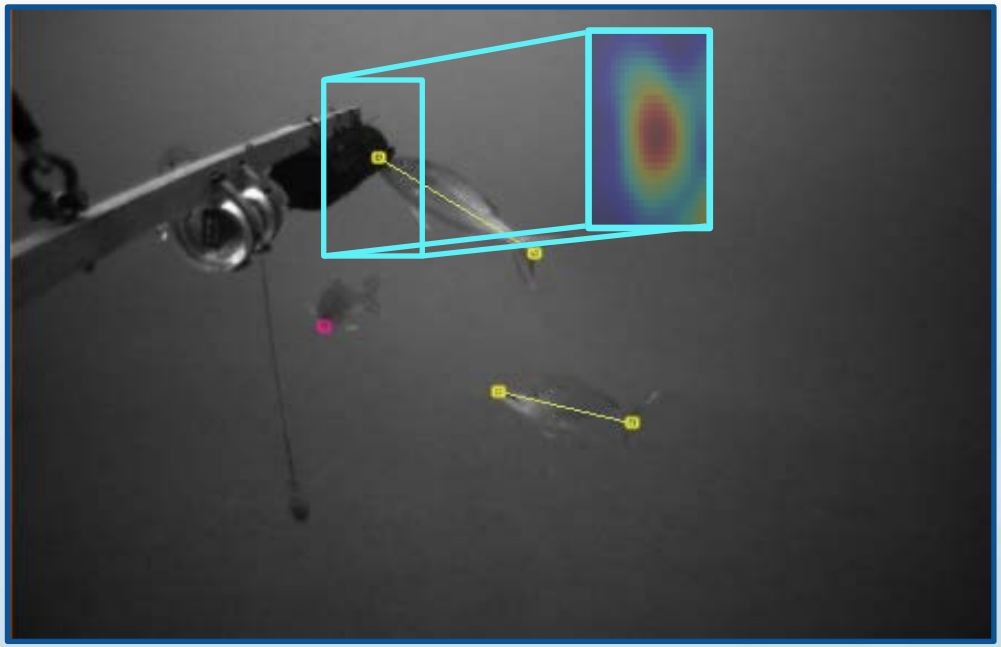
[5] Sadeghian, Amir, et al. "Tracking the untrackable: Learning to track multiple cues with long-term dependencies." CVPR 2017.
[6] Li, Bo, et al. "SiamRPN++: Evolution of siamese visual tracking with very deep networks." CVPR 2019.

Kitware

Stereo Measurement

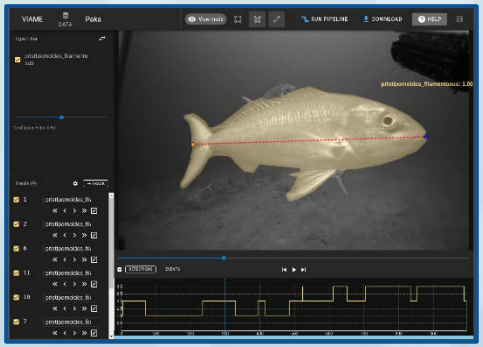


Method #1 - Modeled off camtrawl process (Williams et al 2010), python port of matlab code. Use centroids of smaller sides of oriented bounding boxes as head/tail positions.



Method #2 - Feature point detection using dedicated CNN keypoint detectors, either in the same network or separate dedicated network (e.g. heavily modified version of [7]).

Ongoing: collecting additional annotations.

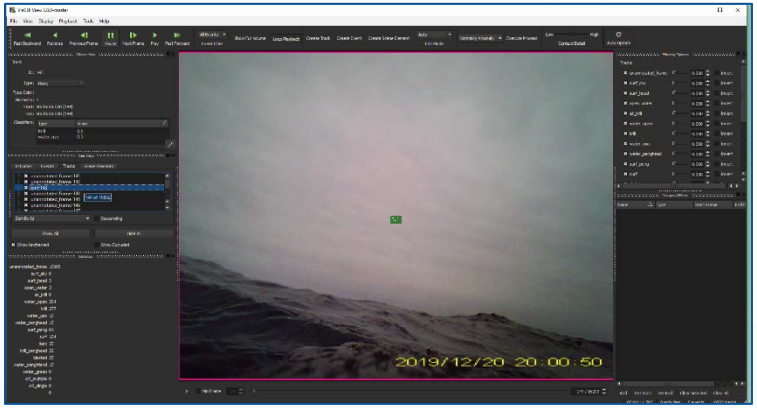


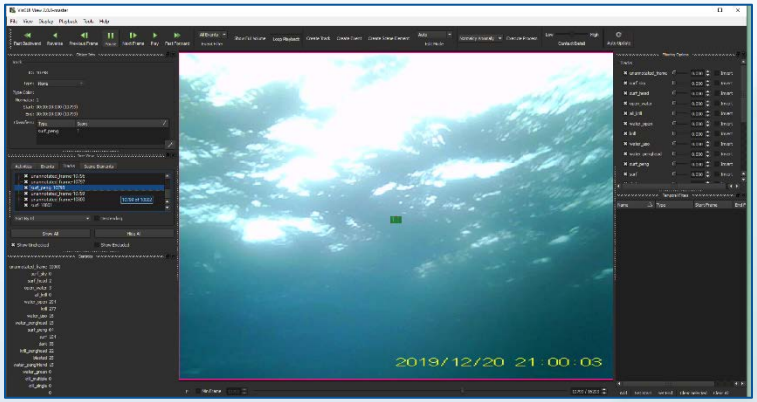
[7] Long, Jonathan, et al. "Fully convolutional networks for semantic segmentation." CVPR 2015.

Kitware

Full Frame Classification

unannotated_frame	10000
surf_sky	0
surf_head	2
open_water	3
all_krill	0
water_open	204
krill	277
water_uso	16
water_penghead	15
surf_peng	64
surf	104
dark	55
krill_penghead	22
blasted	25
water_pengfriend	15





Substrate = X

Substrate = Y

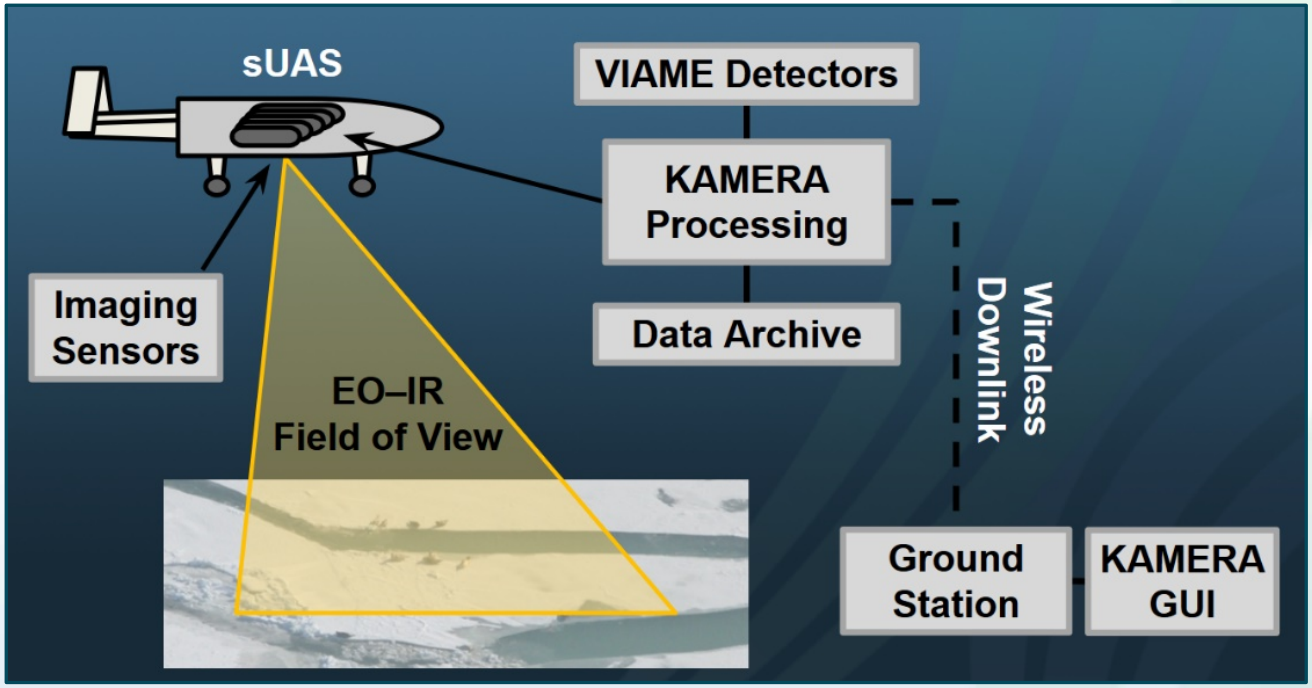
Method #1 – Typical deep training pipeline (ResNet50 [8] – better for cases that have a lot of manual groundtruth)

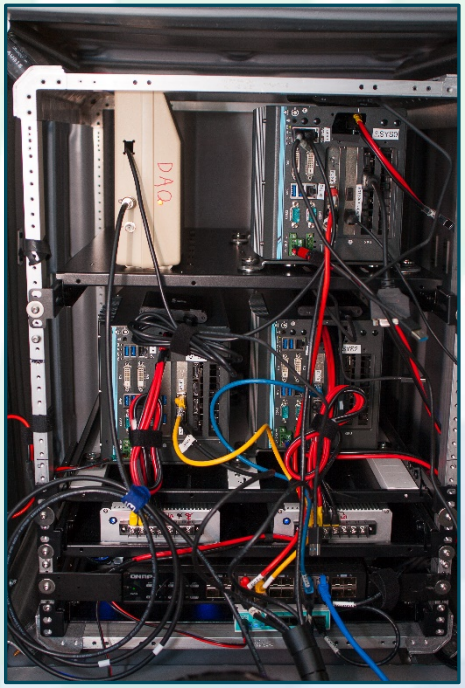
Method #2 – SVM on fixed feature vector, similar to image search and rapid model generation pipeline (better for less training samples)

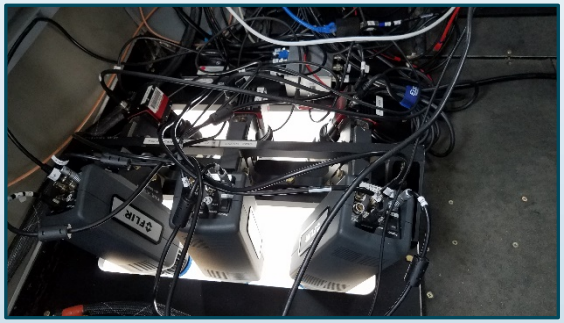
[8] He, Kaiming, et al. "Deep residual learning for image recognition." CVPR 2016.

Kitware

Embedded Processing





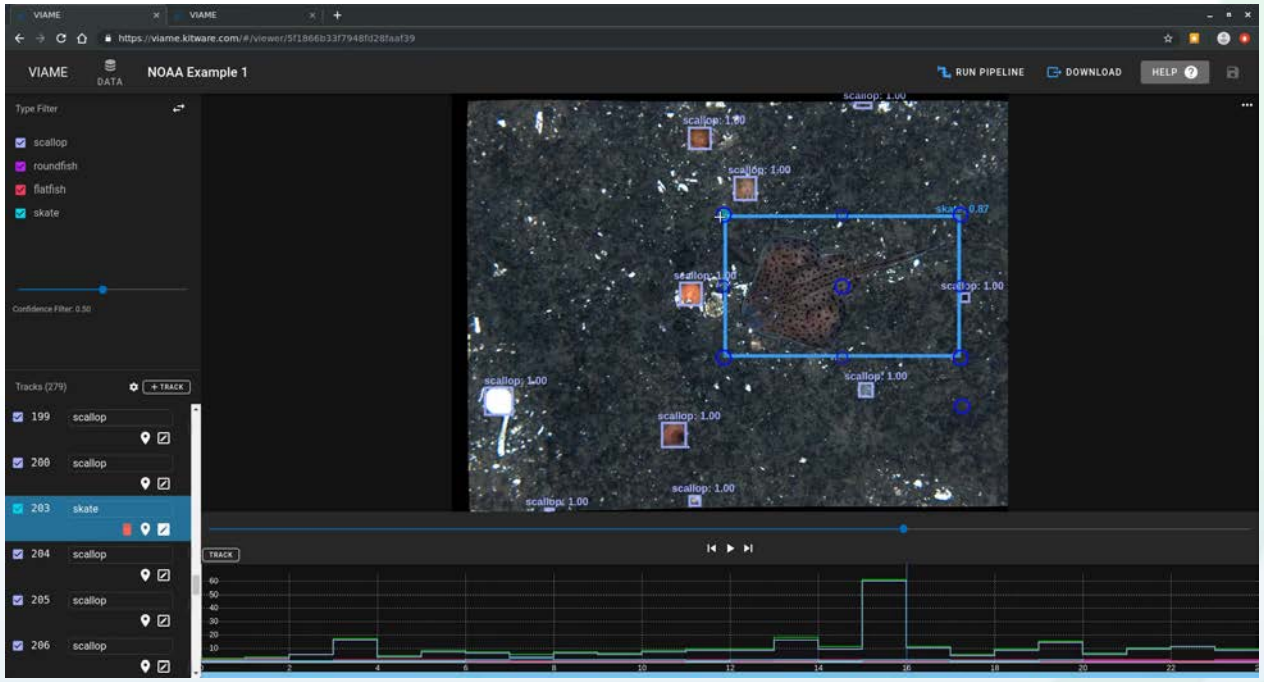


PI: matt.brown@kitware.com

Funded by NOAA ADAPT SBIR Phase I and NOAA AKFSC MML

Kitware

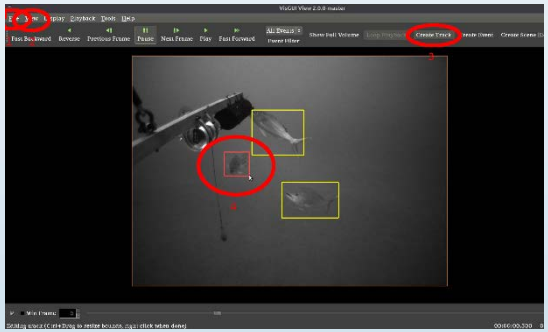
Graphical User Interfaces



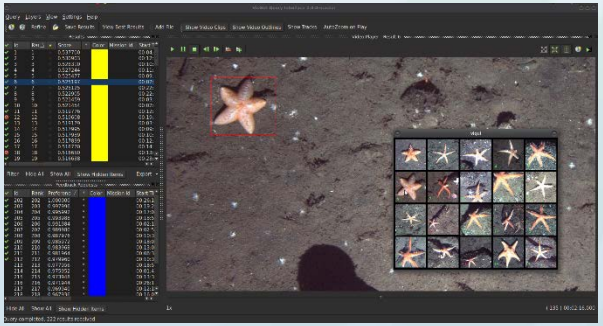
VIAME-Web Annotator

Web vs Desktop

- Both wrap arbitrary processing pipelines and have tracking support (split, merge)
- Web currently has better pixel classification annotation, though some in desktop
- Desktop can currently drive user initialized tracks, web can't, though adding shortly
- Desktop currently has search and rapid model generation and multi-view display, not in web
- Normalizing features across all versions in progress



Desktop Default Annotator

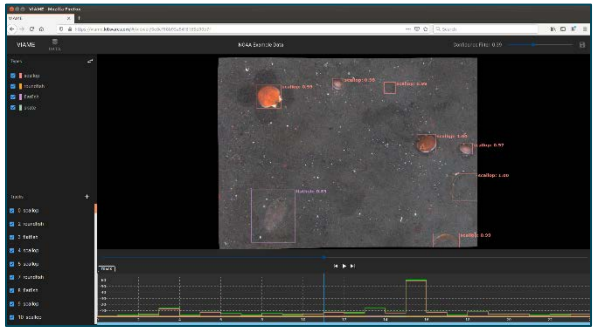


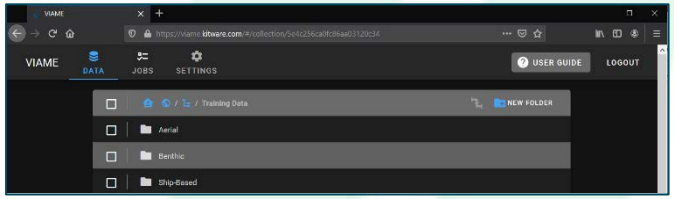
Desktop Search Engine

Stereo/Multi-View Annotator

Kitware

Web Application and Annotation Archive



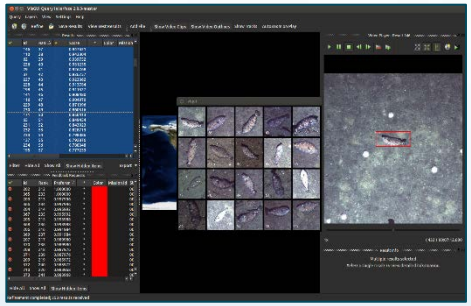


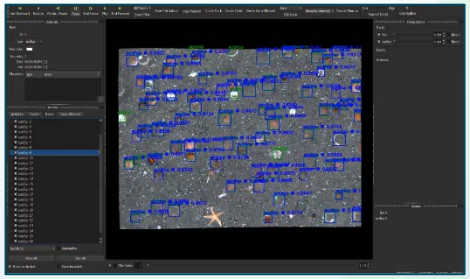
Online Example: **viame.kitware.com**

Server Manages Data and Annotations

Complexity

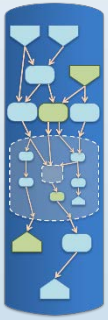
Desktop Applications

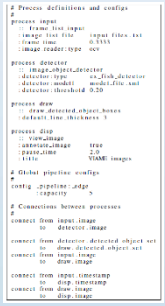


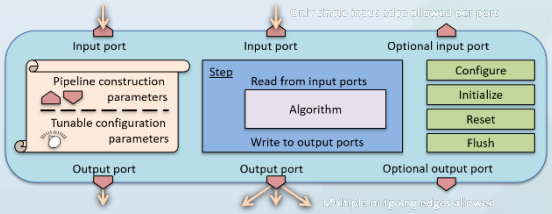


User Manages Data and Annotations

Command-Line Tools and APIs:







Full Feature Support

More Customization Ability

Useful for Embedded Platforms

Covered in Tutorial on Sept 22nd, 12 pm to 2 pm EST

Kitware

Public VIAME Server

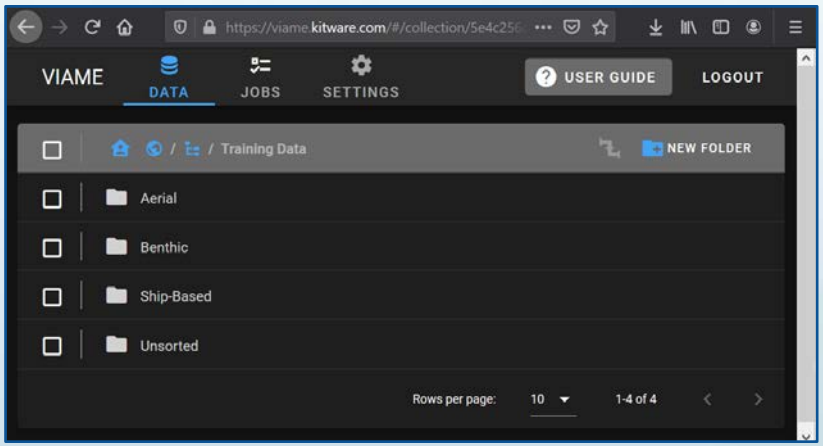
Example VIAME-Web instance: https://viame.kitware.com

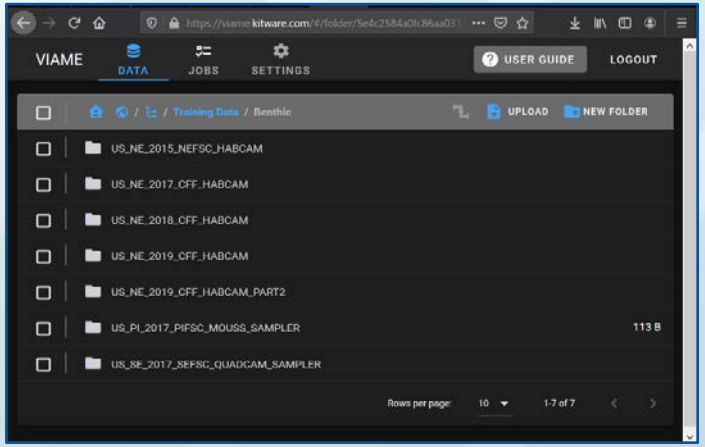
Public data and annotation store provided by Kitware for:

- Storing and Sharing of Annotations and Imagery related to VIAME
- Performing Annotation
- Running Detectors on Data
- Limited Detector Training

15 Tb open training data store (raided for backup), 2 GPUs

- 3 Tb used so far, though still sorting annotations to post online
- Users can annotate data, run multiple pre-trained detectors, and (shortly) use limited training capabilities (FIFO queue on 1 GPU) for new categories



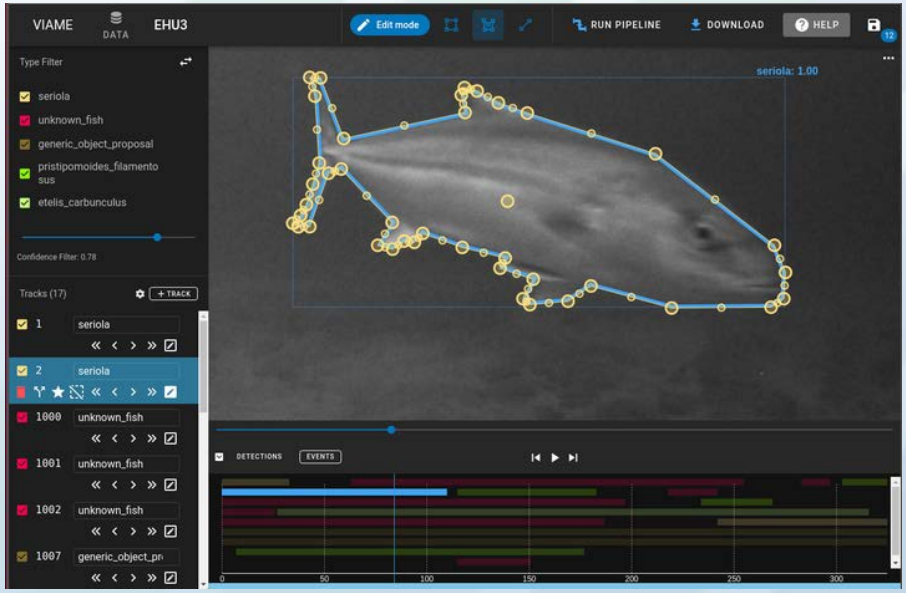


Code and docker containers: https://github.com/VIAME/VIAME-Web

Kitware

Current and Future Work

- Full feature support in web GUIs, improved desktop GUIs
  - Web GUI is very focused towards object detection and tracking, adding in support for other auxiliary features, such as mosaic generation, image enhancement, etc…
    - Make cross-use of other projects at Kitware performing similar work (IARPA, DARPA, DoD)
  - Pixel-classification utility pipelines (e.g. boxes to masks)
- Additional algorithm specializations
- More documentation
- Behavior and event detection
- Acoustic data processing
- Electronic monitoring

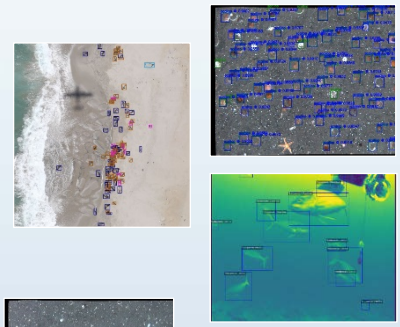


Kitware

Conclusions

- VIAME is a do-it-yourself (DIY) AI toolkit which can be applied to multiple types of imagery or video
- Can be run by people with no programming or machine learning background in both web and desktop interfaces
- Released as fully open-source with a permissive license
- Specializations to maritime processing such as motion fusion, stereo measurement, image enhancement, and object tracking which other software (e.g. Amazon SageMaker) lack

Object Detection



Object Tracking

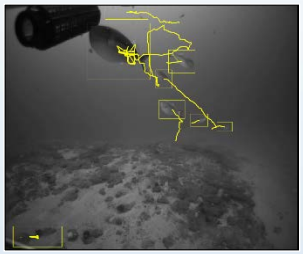
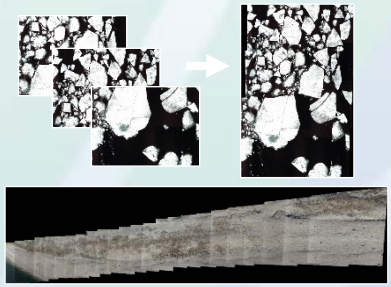
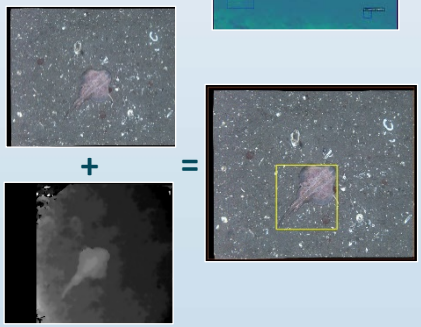


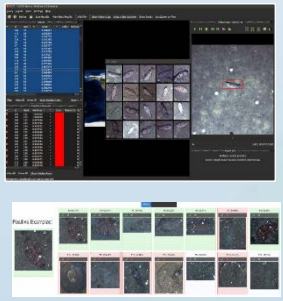
Image Enhancement

Image Registration and Mosaicing

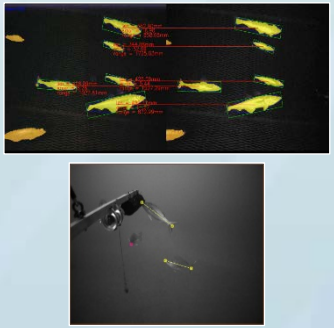




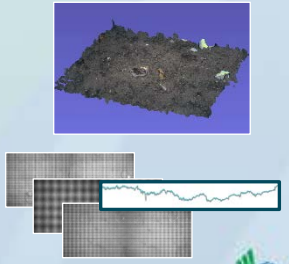
Video Search and Rapid Model Generation



Stereo Measurement



Calibration, 3D and Altitude Estimation



Kitware

Special Thanks to:

- NOAA AIASI, NOAA AKFSC MML and CFF for funding VIAME
- Image annotators and testers across various organizations

See tutorial (Sept 22nd) and https://viametoolkit.org/ for more information

Tutorial Agenda

12:00 to 12:15 pm	–	VIAME Overview Desktop Installation Web vs Desktop vs APIs
12:15 to 12:45 pm	–	VIAME Web Edition Types of Annotation Model Training
12:45 to 1:45 pm	–	VIAME Desktop Edition Core Functionalities Auxiliary Features
1:45 to 2:00 pm	–	APIs and Configs [Advanced]
2:00 to 2:30 pm	–	Open Discussion

Kitware