Creating Deep Learning Detectors in VIAME for Rare Objects in Marine Imagery

Alexander Lynch, Sarah Brockman, Christopher Funk, Roddy Collins, Bharadwaj Ravichandran, Matt Dawkins, Anthony Hoogs

Kitware, Inc

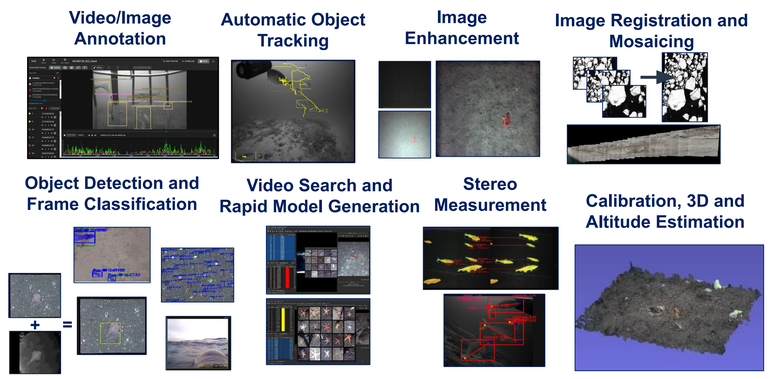
OCEAN SCIENCES MEETING | NEW ORLEANS LOUISIANA 18-23 FEBRUARY 2024

Sponsors

NOAA FISHERIES · DARPA · Office of Naval Research ONR Science & Technology

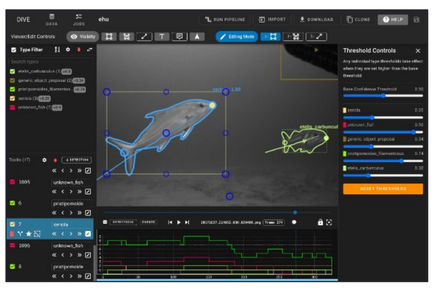
VIAME

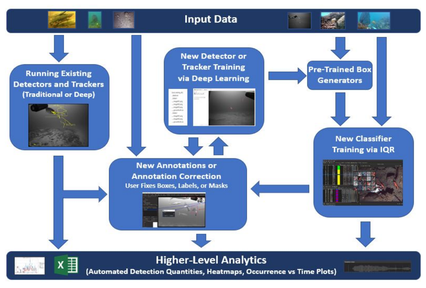
- Video and Imagery Analytics for Multiple Environments: a do-it-yourself AI toolkit for multiple types of imagery or video, with a marine emphasis
- Can be run by people with no programming or machine learning background in both web and desktop interfaces, while also containing command line interfaces (CLIs) and application program interfaces (APIs) for more advanced users
- Has been most commonly used for automating object detection and classification, but contains multiple features including:



Software

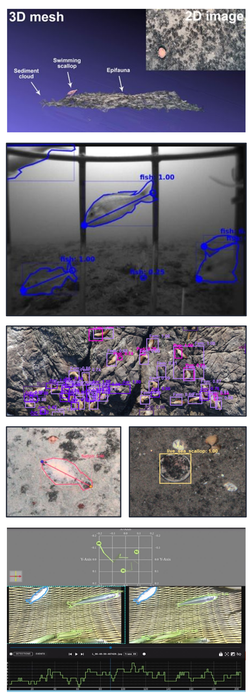
- Installers available at *viametoolkit.org* (on right) and *github.com/viame/viame*
- Free for use with highly permissive licensing
- Public web interface: *viame.kitware.com* (bottom left)
- Contains different workflows and models for varying amounts of training data (bottom right)





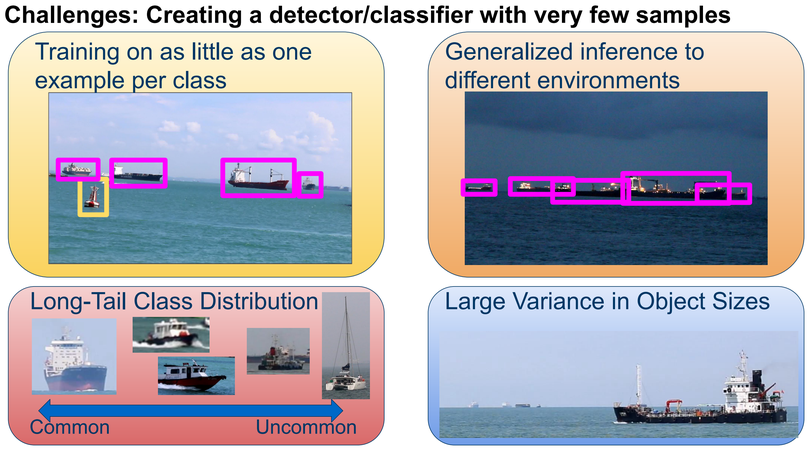
Recent and Future Additions

- **Anomaly detection and few-shot learning are now available**
- Recent publications:
  - "FishTrack23: An Ensemble Underwater Dataset for Multi-Object Tracking." IEEE/CVF Winter Conference on Applications of Computer Vision, 2024.
  - "Towards Depth Fusion into Object Detectors for Improved Benthic Species Classification." ICPR Workshops, 2022.
- Recent features:
  - New default fish, scallop, and sea lion detectors
  - Monocular, metadata-based size measurement
  - Automatic box to polygon converters
  - Additional scoring tools for computing detection precision-recall curves and tracking metrics such as MOTA and IDF1
  - Ensemble models for improving detection
  - 3D target localization using stereo cameras
- Upcoming features:
  - Fish head/tail keypoint localization
  - Additional box to polygon converters

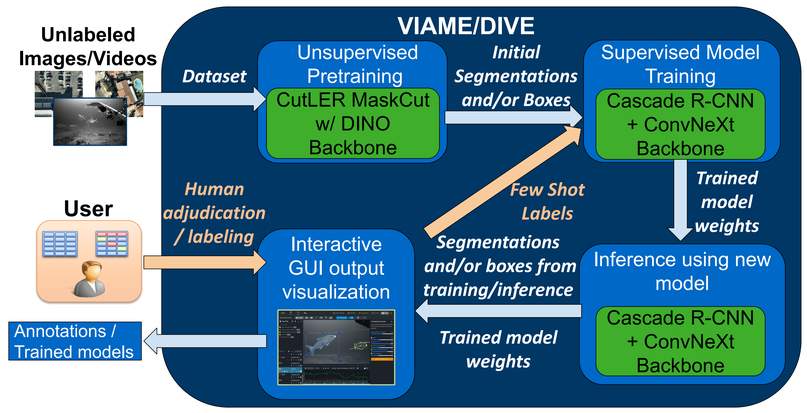


Low-Shot Learning

- **Challenges: Creating a detector/classifier with very few samples**

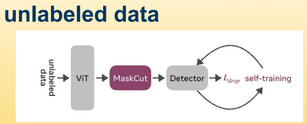


- **Approach: Unsupervised Pre-Training + VIAME IQR**



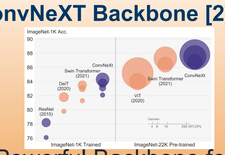
Developed on DARPA's Learning with Less Labeling program

CutLER [1] for leveraging unlabeled data



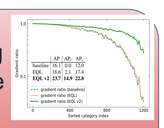
Unsupervised label creation with MaskCut for training

ConvNeXT Backbone [2]

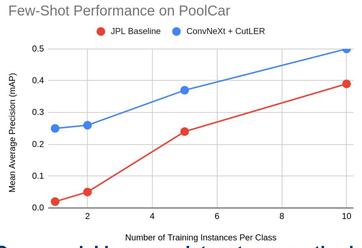


Powerful Backbone for improved generalization

Equalization Loss EQLv2 [3] Balancing positive and negative gradients to prevent discouraging gradients preventing any low-shot class predictions



Few-Shot Performance on PoolCar



On an aerial imagery dataset, our method significantly outperforms the DARPA LwLL baseline at low label counts

Low-Shot Experiment

unlabeled data for training CutLER (N = 12 images)

** also contained no ferries*

labeled data (no "ferry") for finetuning with labels (N = 12 images)

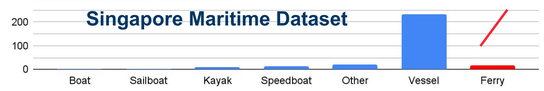
annotate new class ("ferry"); train and update model (N = 5 images)

inference to demonstrate new class (N = 6 images)

CutLER's MaskCut Pretraining

Reasonable accuracy (29%) achieved from 5 examples

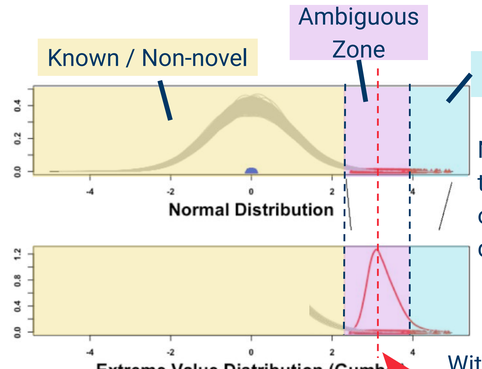
Training Data



LEARN System Results

Anomaly Detection

- **Challenge:** Anomalies are everywhere in natural environments. How do we find interesting, salient anomalies?
- Approaches typically assume enough data has been observed to build a complete generative model, but this is not usually the case
- Our solution leverages existing object detectors
  - Known classes have labeled annotations within training set
  - Unknown/novel/anomalous classes are only within the evaluation data
  - Salient novelties are similar to known classes and different from background
- Our novelty detectors were developed on the DARPA Science of AI and Learning for Open-world Novelty (SAIL-ON) program
  - Our method is theoretically grounded in Extreme Value Theory [5] (EVT), which provides a statistically-valid dissimilarity score for distinguishing between known classes, novelties and background
  - Other methods simply threshold the class probabilities, or perform logistic regression which requires training on known background data



Normal tail is very flat in the tough zone, lacking an obvious point for the decision boundary

EVT uses either Weibull, Gumbel, or Frechet (Generalized Pareto) depending on the type of tail

With a distribution fitted only to the tail, a meaningful decision threshold can be set

- Object detector is trained on 8 classes. Novelty detector identifies items similar to labeled items but not from a known class. Buoys are not in the training set.
- Object detector misclassifies buoy as a vessel. Buoy has a high novelty score.
- Novelty detector is robust against nuisance novelties such as fog, new views, new lighting.
- On xView dataset for satellite imagery, our method outperforms softmax thresholding by 15%.

References

[1] Wang, Xudong, et al. "Cut and learn for unsupervised object detection and instance segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.
[2] Liu, Zhuang, et al. "A convnet for the 2020s." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
[3] Tan, Jingru, et al. "Equalization loss v2: A new gradient balance approach for long-tailed object detection." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.
[4] Cai, Zhaowei, and Nuno Vasconcelos. "Cascade r-cnn: Delving into high quality object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
[5] Du, Dawei, Christopher Funk, Katarina Doctor, and Anthony Hoogs. "Novel Object Detection in Remote Sensing Imagery." In IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium, pp. 5798-5801. IEEE, 2023.
[6] Dawkins, Matthew et al. An open-source platform for underwater image and video analytics. In 2017 IEEE winter conference on applications of computer vision (WACV), 2017. Best Paper Honorable Mention winner.