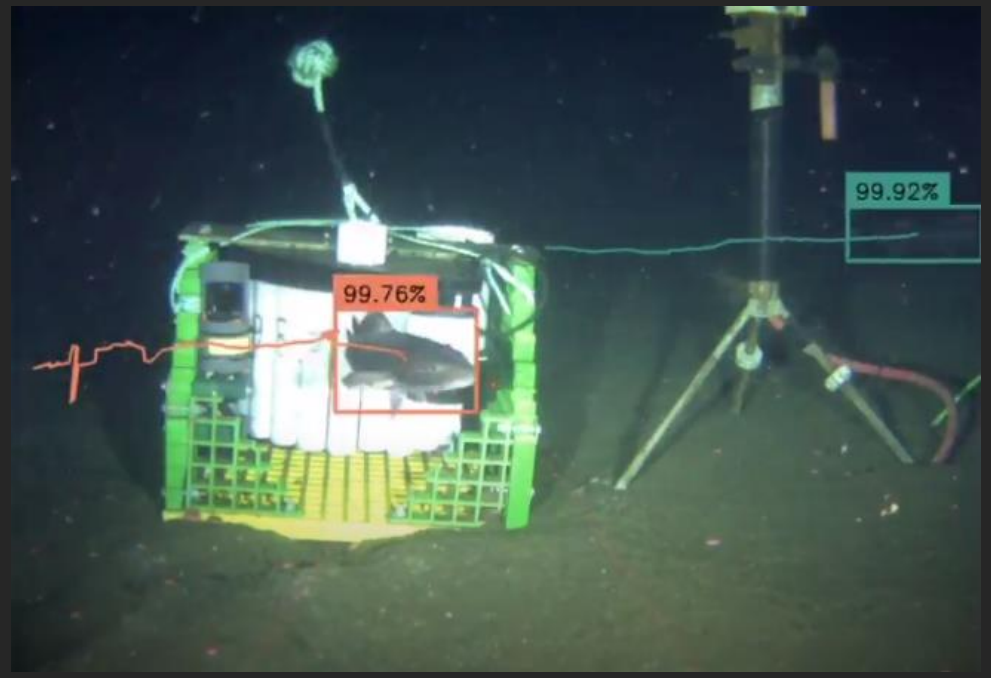


MOVEMENT TRACKS FOR THE AUTOMATIC DETECTION OF FISH BEHAVIORS IN VIDEOS

Authors

Declan McIntosh

Tunai Porto Marques

Alexandra Branzan Albu

Rodney Rountree

Fabio De Leo

Acknowledgements

Oceans Networks Canada

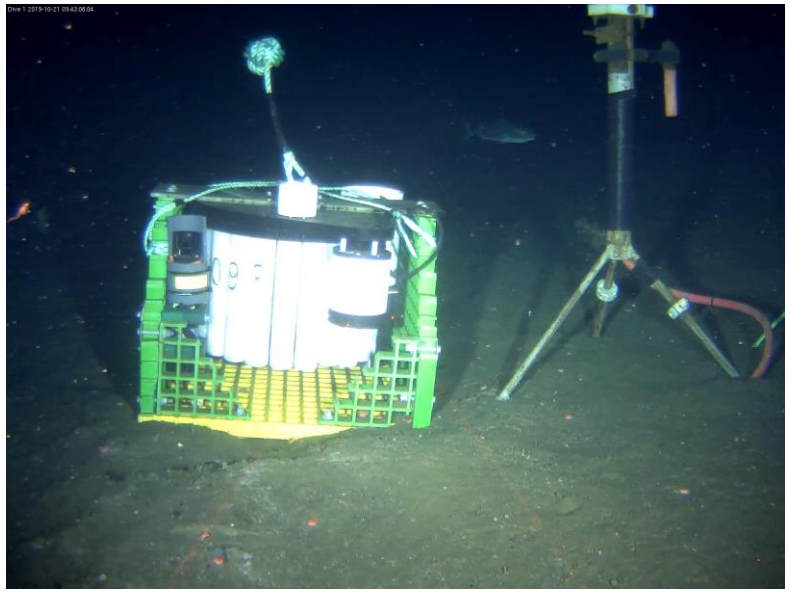
University of Victoria

Natural Science Engineering Research and Council of Canada, Undergraduate Student Research Award

McIntosh et al. (2020) Movement Tracks for the Automatic Detection of Fish Behavior in Videos at Tackling Climate Change with Machine Learning Workshop at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020).

INTRODUCTION

- Global warming, especially ocean acidification and warming can have significant effects on marine ecosystems [1, 2, 3]
- These changes can cause stresses to ecosystems and studies of ecological level behavior can give additional context to these changes [5]
- Manual annotating of the expansive amounts of underwater video for this purpose is prohibitively expensive [4, 5]
- We propose a novel end-to-end behavior detection framework which provides track-wise (can be down-sampled to clip-wise) detection of startle events
- We focus our efforts to sablefish (Anoplopoma fimbria) startle events for this study
- We also offer a dataset of sablefish startle events with multiple levels of data annotation



[1] Thomas F Stocker, Dahe Qin, G-K Plattner, Melinda MB Tignor, Simon K Allen, Judith Boschung, Alexander Nauels, Yu Xia, Vincent Bex, and Pauline M Midgley. Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of ipcc the intergovernmental panel on climate change, 2014.

[2] Nathaniel L Bindoff, Peter A Stott, Krishna Mirle AchutaRao, Myles R Allen, Nathan Gillett, David Gutzler, Kabumbwe Hansingo, G Hegerl, Yongyun Hu, Suman Jain, et al. Detection and attribution of climate change: from global to regional. 2013.

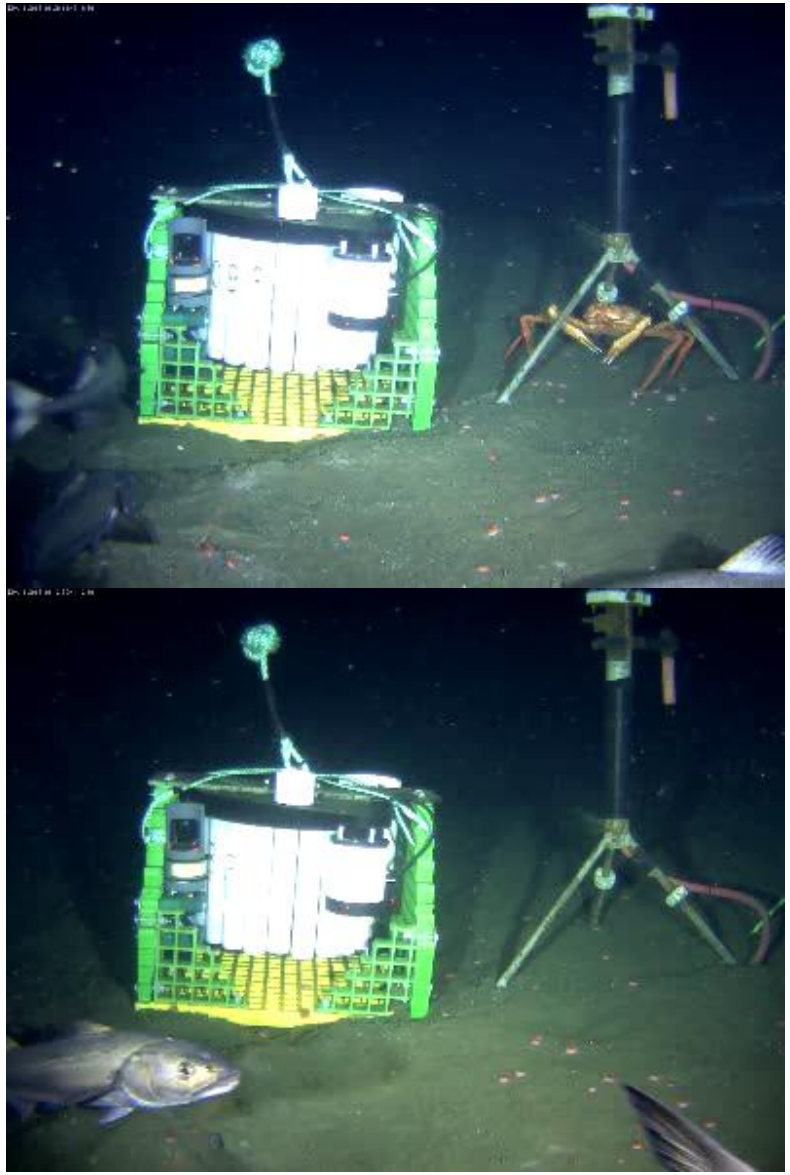
[3] Jacopo Aguzzi, Carolina Doya, Samuele Tecchio, Fabio De Leo, Ernesto Azzurro, Cynthia Costa, Valerio Sbragaglia, Joaquin del Rio, Joan Navarro, Henry Ruhl, Paolo Favali, Autun Purser, Laurenz Thomsen, and Ignacio Catalan. Coastal observatories for monitoring of fish behaviour and their responses to environmental changes. Reviews in Fish Biology and Fisheries, 25:463–483, 2015.

[4] Tunai Porto Marques and Alexandra Branzan Albu. L2uwe: A framework for the efficient enhancement of low-light underwater images using local contrast and multi-scale fusion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pages 538–539, 2020.

[5] Cosmin Ancuti, Codruta Orniana Ancuti, Tom Haber, and Philippe Bekaert. Enhancing underwater images and videos by fusion. In 2012 IEEE Conference on Computer Vision and Pattern Recognition, pages 81–88. IEEE, 2012.

RELATED WORKS

- Several works provide solutions for organism counting but these methods lack higher level understanding of organism behavior [6,7,8]
- Previous work on organism detection is not trivially extended to behavior detection
- Current event detectors, for example ReMotENet [9] do not provide instance-level behavior identification
- A system of abnormal event detection on intra-class domains, with similar difficulties to behavior detection, was offered by Ionescu et al. [10]



[6] YH Toh, TM Ng, and BK Liew. Automated fish counting using image processing. In 2009 International Conference on Computational Intelligence and Software Engineering, pages 1–5. IEEE, 2009.

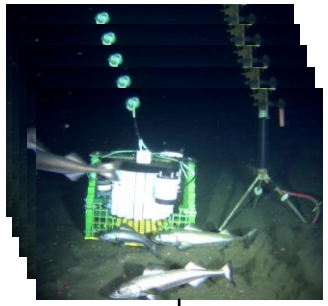
[7] Concetto Spampinato, Yun-Heh Chen-Burger, Gayathri Nadarajan, and Robert B Fisher. Detecting, tracking and counting fish in low quality unconstrained underwater videos. VISAPP (2), 2008(514-519):1,2008.

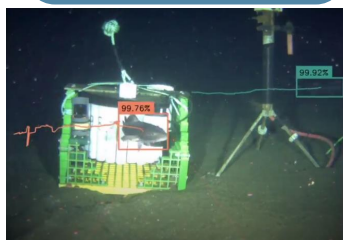
[8] Song Zhang, Xinting Yang, Yizhong Wang, Zhenxi Zhao, Jintao Liu, Yang Liu, Chuanheng Sun, and Chao Zhou. Automatic fish population counting by machine vision and a hybrid deep neural network model. Animals, 10(2):364, 2020.

[9] Ruichi Yu, Hongcheng Wang, and Larry S Davis. Remotenet: Efficient relevant motion event detection for large-scale home surveillance videos. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pages 1642–1651. IEEE, 2018.

[10] Radu Tudor Ionescu, Fahad Shahbaz Khan, Mariana-Iuliana Georgescu, and Ling Shao. Object-centric auto-encoders and dummy anomalies for abnormal event detection in video. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 7842–7851, 2019.

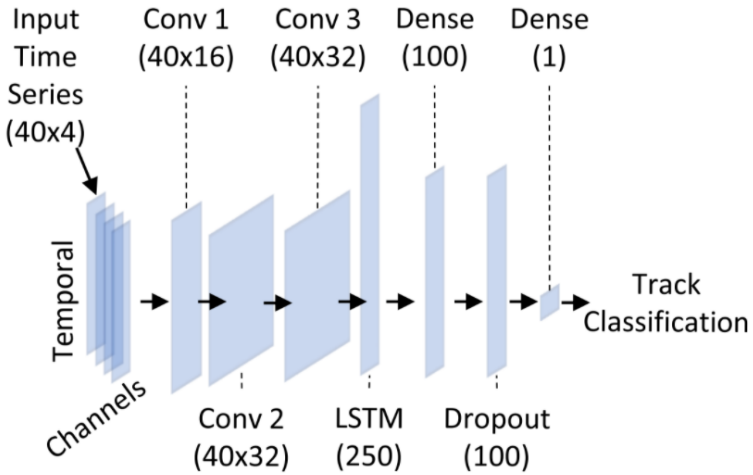
PROPOSED SOLUTION





Object Detection and Tracking with Domain Specific Metrics for LSTM classification

- We deploy a YoloV3[11] object detector to initially detect sable fish
- The Hungarian algorithm is used to generate loss minimizing associations as tracks
- A Long Short Term Memory (LSTM) classifier is used to categorize tracks based on 4 time series track metrics
- The LSTM classifier was chosen to use the temporal relationships of the metrics

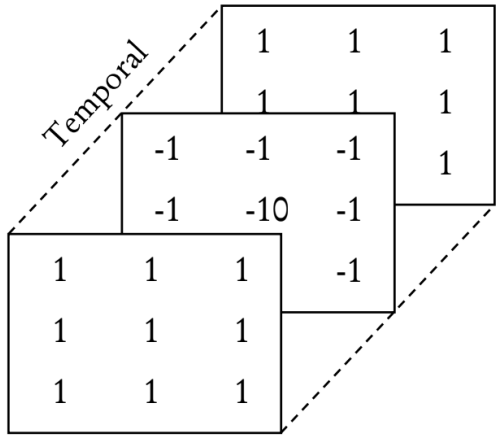


Proposed LSTM network for track classification

[11] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.

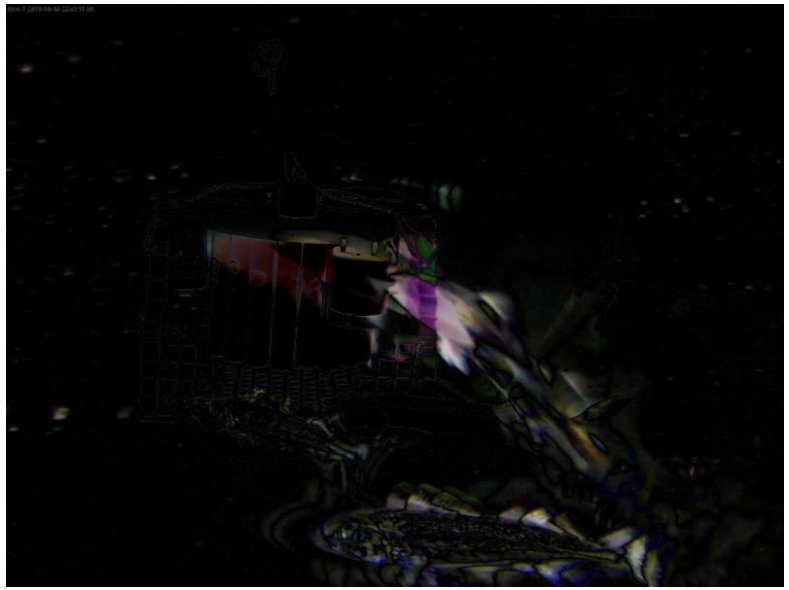
McIntosh et al. (2020) Movement Tracks for the Automatic Detection of Fish Behavior in Videos at Tackling Climate Change with Machine Learning Workshop at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020).

BEHAVIOR SPECIFIC FEATURES

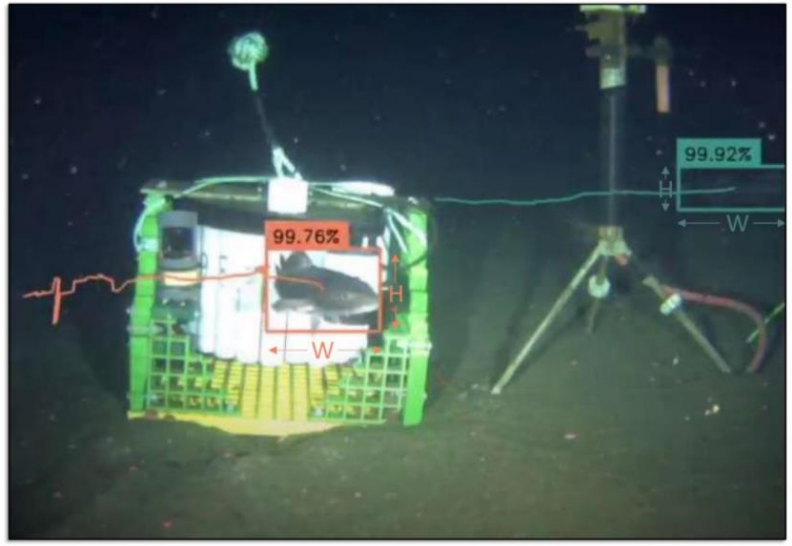


LMCM 3D (x, y, temporal) convolution kernel.

- We propose four domain specific metrics for the sable fish startle detection problem
  - Track speed
  - Track direction
  - Track detection aspect ratio
  - Local Momentary Change Metric (LMCM)
- These were found to be the minimal constraining metrics for the problem
- These metrics can be customized for specific problem domains



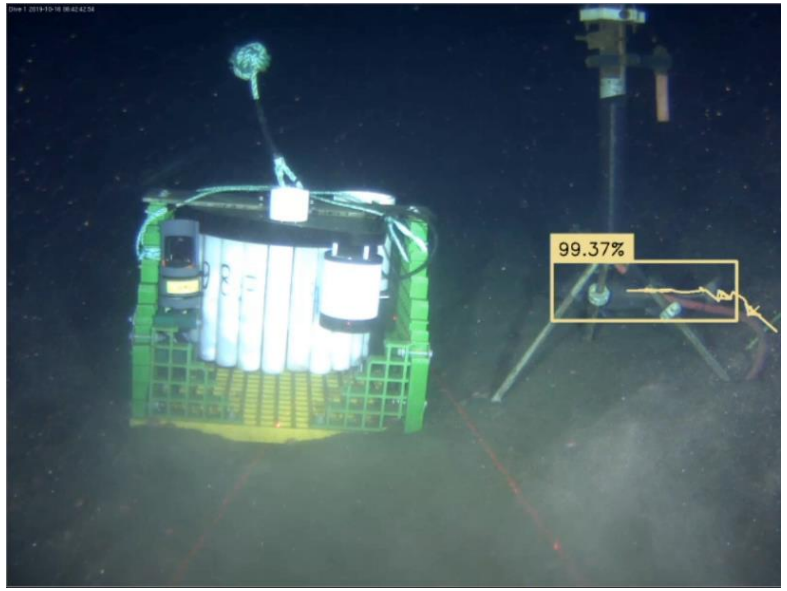
Example LMCM output on RGB image series.



Example tracks with width and heigh for track aspect ratio labeled.

McIntosh et al. (2020) Movement Tracks for the Automatic Detection of Fish Behavior in Videos at Tackling Climate Change with Machine Learning Workshop at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020).

SABLEFISH STARTLE DATASET



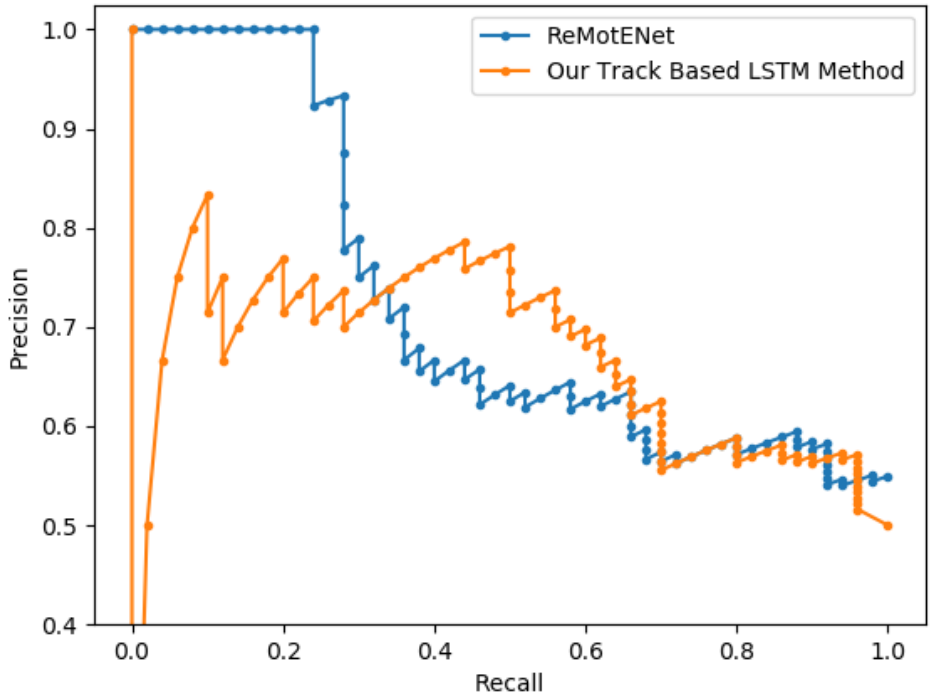
- The provided dataset contains 3 levels of annotation.
  - 600 single images, with sable fish detection ground truths
  - 892 4 second clips classified for the existence of any startle event
  - 2240 tracks classified for the existence of a startle event
- All tracks and individual images are generated from the 892 clips
- Tracks less than 2 seconds are discarded
- Videos are provided at 10 frames per second

Data Split	Clips	Startle Clips	Tracks	Startle Tracks
Train	642	321	1533	323
Validation	150	75	421	80
Test	100	50	286	50

RESULTS

Method	Track AP	Track BCE	Clip AP	Clip Recall
Ours	0.85	0.412	0.67	0.58
ReMotENet[15]	N/A	N/A	0.61	0.50

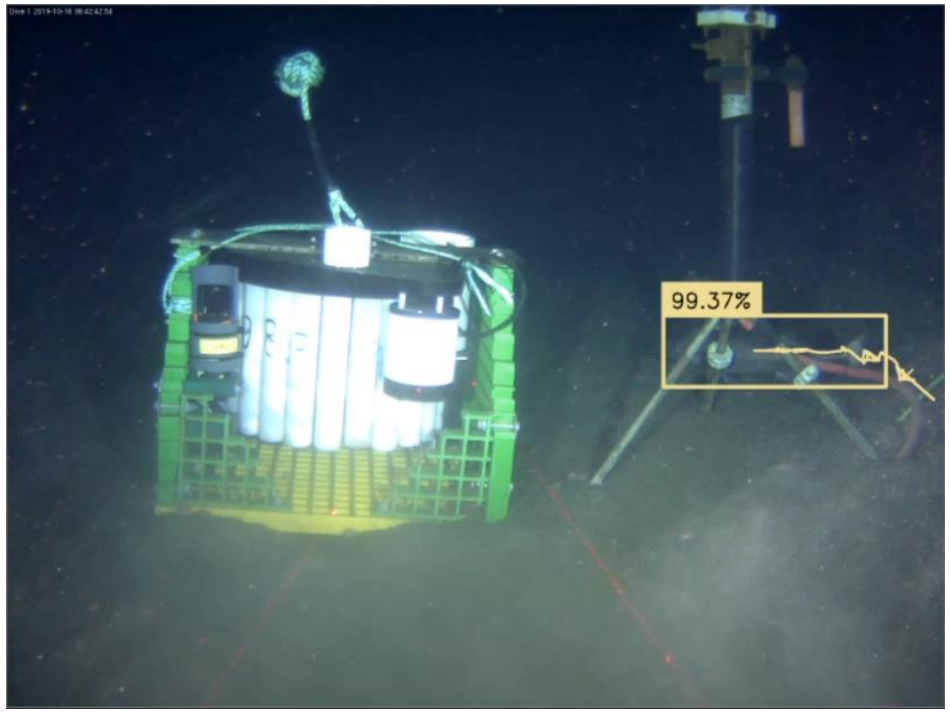
- We compare out network to a state of the art event detection method ReMotENet[9]
- ReMotENet cannot generate track-wise startle detections
- We provide our method's track-wise and down-sampled clip-wise classifications
- The degradation of track-wise AP to clip-wise AP is due to lost tracks and the high noise sensitivity of the maximal conversion



[9] Ruichi Yu, Hongcheng Wang, and Larry S Davis. Remotenet: Efficient relevant motion event detection for large-scale home surveillance videos. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pages 1642–1651. IEEE, 2018.

CONCLUSIONS

- Our proposed method generates semantically richer track-wise annotations
- We intend our methods to enable long term studies on fish behaviour over time for climate change related ecological information
- The generated dataset for sablefish behaviour provides multiple levels of annotation as a benchmark for organism behaviour detection
- Our method after down sampling outperforms an existing state of the art event detector ReMotENet[9]
- Future work will address more behaviours and associated track metrics



[9] Ruichi Yu, Hongcheng Wang, and Larry S Davis. Remotenet: Efficient relevant motion event detection for large-scale home surveillance videos. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pages 1642–1651. IEEE, 2018.