Visual Diver Recognition for Underwater Human-Robot Collaboration

Youya Xia and Junaed Sattar
University of Minnesota, Twin Cities, USA

Interactive Robotics and Vision Laboratory (IRVLab)

ICRA 2019 International Conference on Robotics and Automation
May 20-24, 2019
Montreal, Canada

Contributions

Overview:

- A method for visually detecting and identifying divers using deep-learned detection models and feature-based learning of individual appearances

Experiments:

- Evaluation of the accuracy and performance of the proposed algorithm on both the open-water (ocean) and closed-water (pool) datasets
- A real-time implementation of the said algorithm to run on-board a mobile robot

Motivation and Challenges

Motivations:

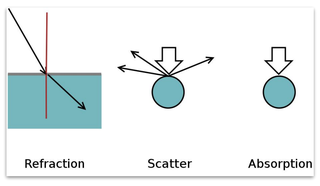
- Facilitating the direct communication between divers and autonomous underwater vehicles (AUVs)
- Enabling AUVs to interact only with particular users, for example, to interact with or follow specific people

Limitation of the existing methods:

- The existing diver detection methods do not detect specific divers

Challenges of vision-based methods underwater:

- Optical distortion
- Color absorption
- Scattering



Methodology

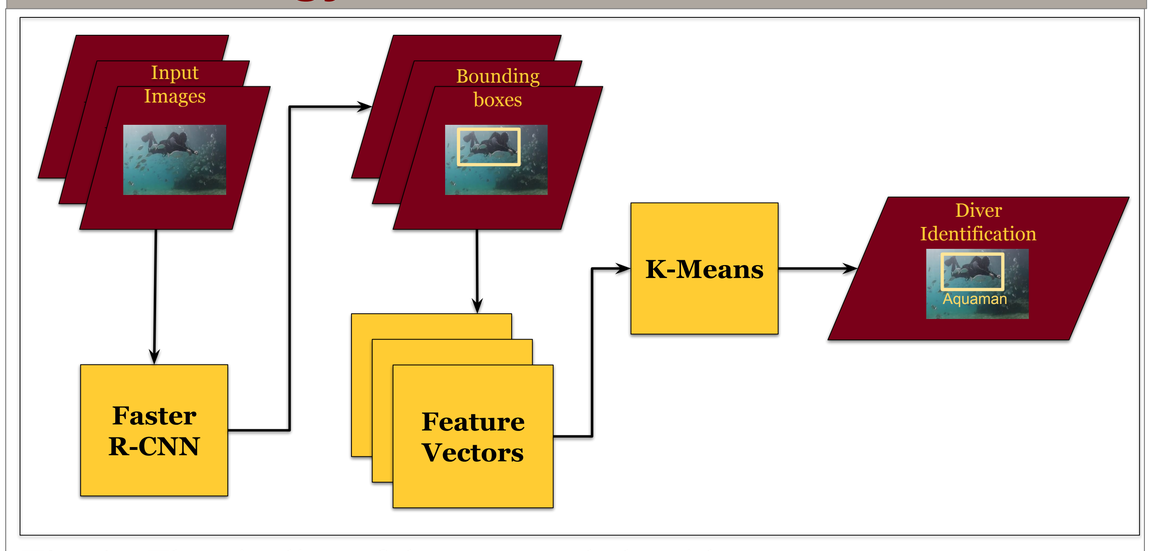


Fig. 1: The pipeline of the proposed algorithm

- **Diver Detection with Faster R-CNN:**
  - Outperforms the majority of other deep learning neural networks in accuracy (*e.g,* about 10 percent more accurate than YOLO)
  - Also outperforms the majority of other Regional Convolutional Neural Networks (R-CNNs) in speed (*i.e.*, about 20 times faster than other R-CNN models)
- **K-Means clustering:** Choosing five different features from each diver detected by Faster R-CNN
  - Convex hull of shapes of divers
  - Seven Hu's Moments
  - Average color distribution
  - Amplitude of Spatial Frequency Distribution
  - Canny edge features

Training and Evaluation

Training:

- Faster R-CNN was trained on a quad NVIDIA GTX 1080 system with 2,000 images
- Used the pretrained Neural Network distributed by Tensorflow to speed up training

Evaluation:

- Performance was evaluated on:
  - GPU (NVIDIA GTX 1080)
  - CPU (Intel i3-6100U)
- Conducted on different scenarios in the pool
- Also conducted on the ocean datasets collected during AUV field trials at the Bellairs Research Center in Barbados

Implementation and Experimental Evaluation

Scenario	Accuracy(%)	Missed Identification(%)	Wrong Identification(%)
Scenario 1: two divers, no flippers, one diver exits scene	100	0	0
Scenario 2: two divers, no flippers, one diver exits scene and later reenters	96.8	0	3.2
Scenario 3: two divers, with flippers, one diver exits scene	94.9	0.3	4.8
Scenario 4: two divers, with flippers, one diver exits scene and later reenters	90.8	2.2	7
Scenario 5: three divers, no flippers, one diver exits scene	77.5	1.4	21.1
Scenario 6: three divers, with flippers, one diver exits scene	80.7	0	19.3
Scenario 7: two divers, no flippers, free-form swim	90.5	0	9.5
Scenario 8: two divers, ocean waters ,full-body dive suit and flippers	96.07	0	3.93

Fig.2: The experimental results of the 8 different scenarios

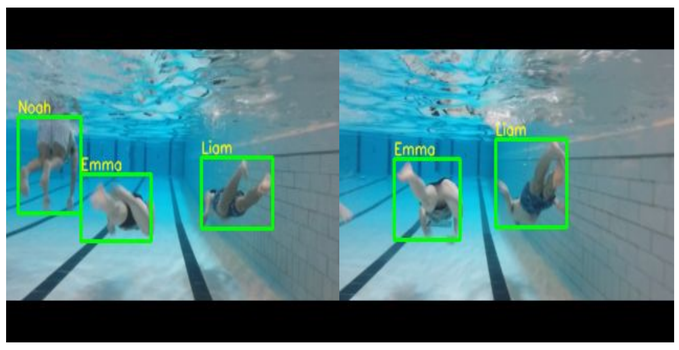


Fig.3: Three divers, no flippers, one diver exits scene

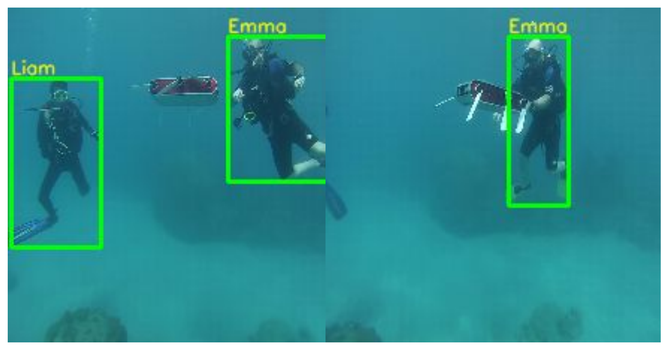


Fig.4: Two SCUBA divers , one exits the scene

Conclusions and Work-in-progress

Conclusions:

- We have proposed the first vision-based algorithm for underwater autonomous robots to uniquely identify swimmers and divers
- Part of a broader human-robot collaboration framework

Future work:

- Incorporating pose detection models (*e.g*, OpenPose) into the feature vector to enhance identification accuracy
- Integrating gesture-based communication and diver-following abilities with the diver-identification algorithm

Resources

- **Code**: https://github.com/xiaxx244/diver_detection
- **Demo**: youtu.be/rbsOBoG2QeM
- **Contact**: *xiaxx244@umn.edu*
- **More**: visit us at irvlab.cs.umn.edu

Code

Project page