Fei WANG, Wanyu LI, Miao LIU, Jingchun ZHOU, Weishi ZHANG, 2023. Underwater object detection by fusing features from different representations of sonar data. *Frontiers of Information Technology & Electronic Engineering*, 24(6):828-843.
https://doi.org/10.1631/FITEE.2200429

Underwater object detection by fusing features from different representations of sonar data

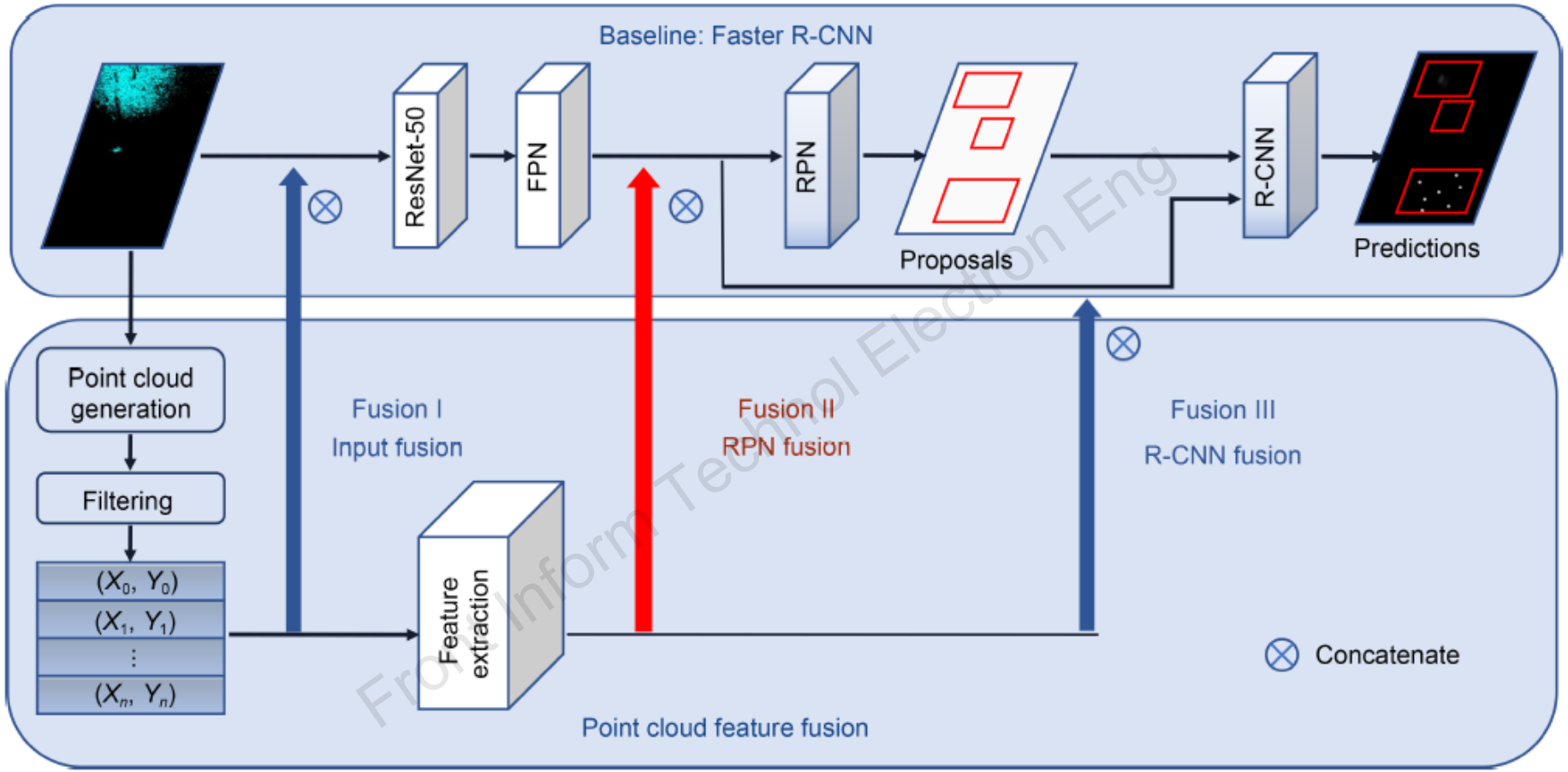
Key words: Underwater object detection; Sonar data representation; Feature fusion

Corresponding author: Jingchun ZHOU
E-mail: zhoujingchun@dlmu.edu.cn
ORCID: https://orcid.org/0000-0002-4111-6240

Motivation

- Underwater sonar images are low in resolution and high in noise. The accuracy of detection results is not good.
- Existing methods have focused mainly on improving the model structure and training procedure. The influence of sonar data representation on the detection results has seldom been investigated.
- Two kinds of distortion in sonar data are discussed, including projection distortion and representation distortion. The first is unrecoverable due to the limitation of sonar sensors, while the second is introduced by data representations. To reduce the influence of object distortion in sonar data, three sonar data representations are presented and analyzed.

Model structure



- Three strategies to fuse features from polar image and point cloud representations of sonar data are presented, including input fusion, fusing before region proposal network (RPN fusion), and fusing before the region of interest (ROI) pooling and ROI head of the R-CNN (R-CNN fusion).

Major results

1. mAP results of our model and related methods in underwater object detection

Table 4 Mean average precision of approaches without/with feature fusion

Representation	Mean average precision (%)					
	mAP@[.5, .95]	mAP@0.5	mAP@0.75	mAP(S)	mAP(M)	mAP(L)
Polar image*	50.8	93.8	49.7	18.5	51.0	45.5
Point cloud*	50.8	95.4	48.4	24.4	50.9	46.2
Input fusion#	51.8	95.6	50.0	24.4	51.8	48.7
RPN fusion#	**52.2**	**95.8**	**51.4**	**26.1**	**52.1**	**49.3**
R-CNN fusion#	51.5	94.3	50.5	24.4	51.6	47.5

Best results are in bold. * without feature fusion; # with feature fusion

Major results (Cont'd)

2. Average recall results of our model and related methods in underwater object detection

Table 8 Average recall of proposal generation without/with feature fusion

Representation	AR_{100}	AR_{300}	AR_{1000}	AR(S)	AR(M)	AR(L)
Polar image*	61.1%	61.1%	61.1%	24.4%	61.3%	61.4%
Point cloud*	60.8%	60.8%	60.8%	32.2%	61.0%	60.8%
Input fusion#	**61.7%**	**61.7%**	**61.7%**	32.2%	**61.7%**	**64.9%**
RPN fusion#	61.5%	61.5%	61.5%	**34.4%**	61.6%	62.4%

Best results are in bold. * without feature fusion; # with feature fusion

Conclusions

- A summary of different representations for sonar data is first presented, including polar image, Cartesian image, and point cloud.
- Three feature fusion strategies are presented in the framework to investigate the impacts of feature fusion on different components of the detection pipeline.
- A series of experiments are conducted on a public sonar dataset to demonstrate the effectiveness of our framework.

Fei WANG received the PhD degree in Control Theory and Engineering from Dalian University of Technology, China in 2019. He obtained the BSc and MSc degrees in Computer Science and Technology from Dalian Maritime University, China, in 2012 and 2015 respectively. He is currently a lecturer in College of Information Science and Technology at Dalian Maritime University, China. His research interests include robotics, deep learning, 3D data processing, and semantic scene understanding.

Wanyu LI received the BSc degree in English and Software Engineering from Dalian Jiaotong University, China in 2020. She is currently a post-graduate student in College of Information Science and Technology at Dalian Maritime University, China. Her research interests include robotics, deep learning, and object detection.

Miao LIU is currently an undergraduate student majoring in traffic management at Dalian Maritime University. She has been involved in deep learning and related research projects.

Jingchun ZHOU received his MS degrees in software engineering from Beijing University of Posts and Telecommunications (BUPT) and computer information technology from Northern Arizona of University (NAU). He received his PhD degree in computer applications at Dalian Maritime University, China, in 2021. He is currently a post-doctoral researcher with Dalian Maritime University and Hong Kong Polytechnic University (PolyU), in China. He is a reviewer for *IEEE TIP, TGRS*, *SPL*, *JOE*, *AI*, *EAAI*, *Information Fusion*, *Neurocomputing*, etc. His research interests include computer vision, deep learning, and underwater image enhancement.

Weishi ZHANG received his PhD degree in computer science from University of Munich, Germany. He is currently a professor in College of Information Science and Technology at Dalian Maritime University, China. His primary research involves computer vision, pattern recognition, software engineering, and software architecture.