

SMALL OBJECT DETECTION ON THE WATER SURFACE BASED ON RADAR AND CAMERA FUSION

Qiancheng Wei¹, Xiaoping Jiang¹, Ying Liu^{1,*}, Qiya SU², Muyao YU²

¹ University of Chinese Academy of Sciences, China

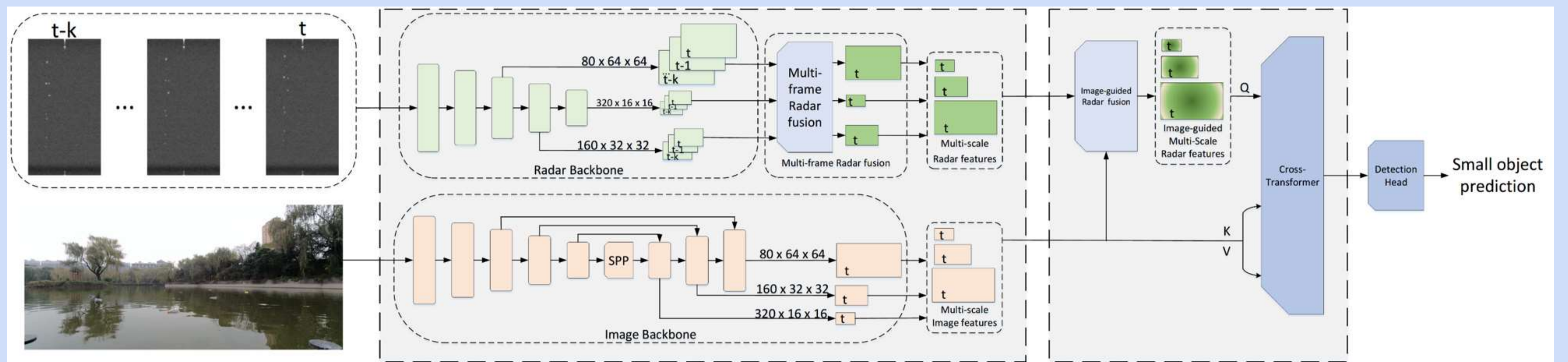
² Beijing Institute of Remote Sensing Equipment, China

Introduction

In this study, a radar-camera fused object detection method, RCFNet, is proposed for small object detection on the water surface.

A radar-camera fused object detection method (RCFNet):

- RGB image and RD image feature extraction
- Multi-frame radar features fusion module
- Image-guided radar feature by fusing semantic information
- Cross-transformer for radar and RD features fusion

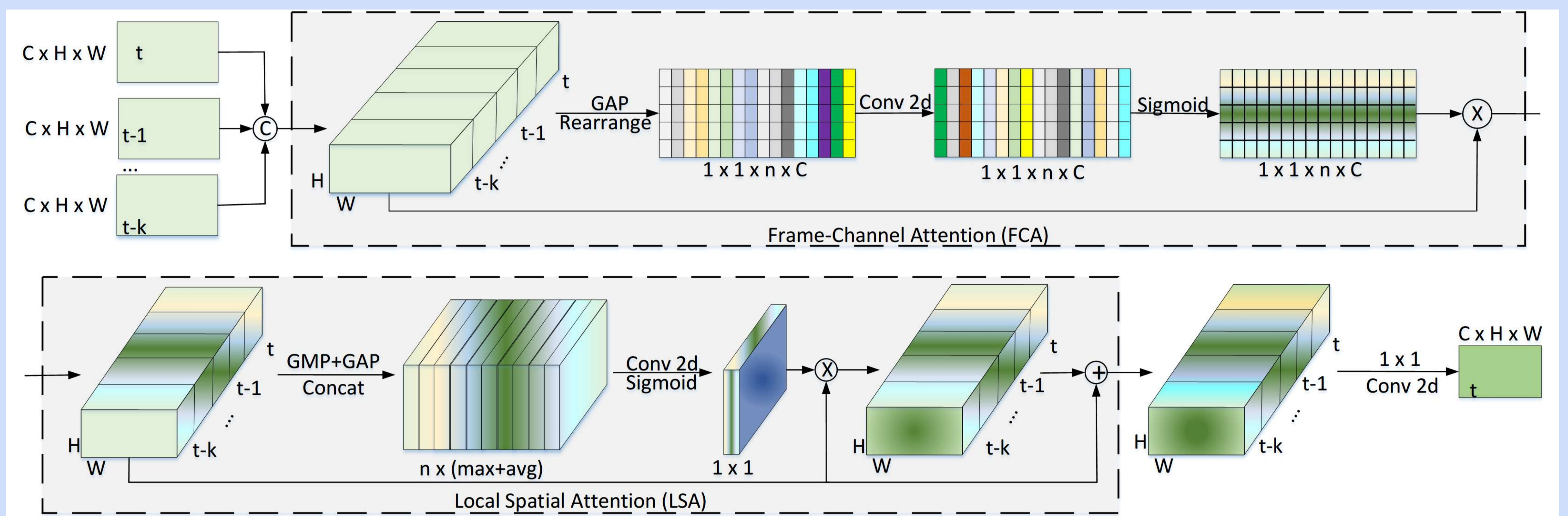


The overview of RCFNet framework.

Materials

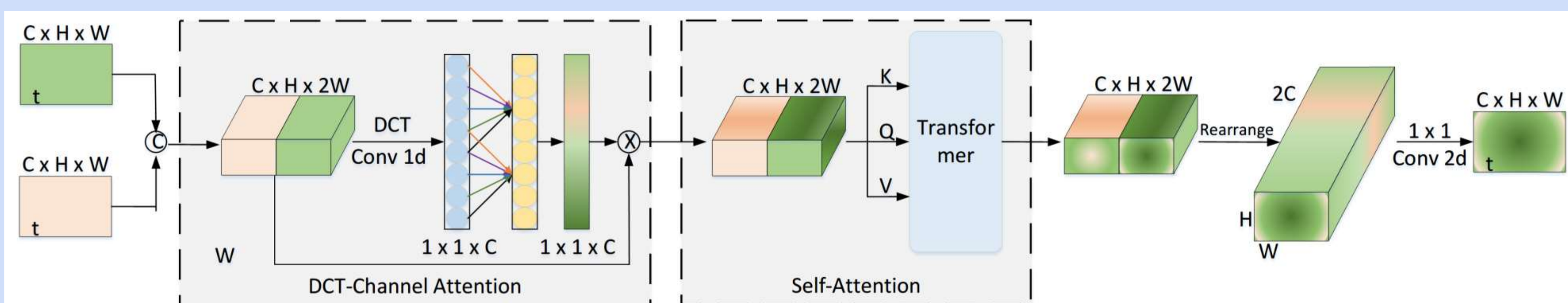
A multi-frame radar feature fusion module (MRF):

- Temporal-position encoding
- Frame-channel attention (FCA)
- Local spatial attention (LSA)



The structure of MRF module.

EXPERIMENT



The structure of VGRF module.

VGRF module can suppress noise and learn semantic alignment information.

Performance	AP50	AP50-90	FPS
Achelous	0.879	0.358	33.67
RCDPT	0.754	0.279	32.98
TransFuser	0.783	0.276	13.18
RCFNet (ours)	0.932	0.447	25.94

Results on Flow dataset using our method and advanced radar-visual fusion methods.

Performance	AP50	AP50-90
(w/o) FCA	0.922	0.445
(w/o) FDA	0.923	0.437
(w/o) MRF	0.920	0.430
(w/o) VGRF	0.921	0.443
RCFNet	0.932	0.447

Results of ablation experiments on Flow dataset.



The detection result of RCFNet.

Conclusions

In this paper, we propose a radar-camera fused object detection method for small object detection on the water surface. The model first extracts features from both modalities and then enhances the radar features in two stages. The first stage is to fuse multi-frame radar data, which is used to enrich the radar features. The second stage is to use attention to align the image semantic information with the radar features to generate more robust radar features. Finally, the cross-transformer is used to deeply fuse the two modalities in a global manner to generate object features for detection. Experimental results on a dataset of water surface floaters show that RCFNet achieves significant improvements in detection performance compared to single-modality detection methods and other fusion methods. Ablation experiments verify the effectiveness of the proposed modules.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant #62176247 and research grant #2020-JCJQ-ZD-057-00. It was also supported by the Fundamental Research Funds for the Central Universities. Thanks Dr. Cheng Wang and other researchers in Beijing Raying Technologies, Inc. for their great support in radar data analysis.