Ruihui PENG, Jie LAI, Xueting YANG, Dianxing SUN, Shuncheng TAN, Yingjuan SONG, Wei GUO, 2024. Camouflaged target detection based on multimodal image input pixel-level fusion. *Frontiers of Information Technology & Electronic Engineering*, 25(9):1226-1239.
https://doi.org/10.1631/FITEE.2300503

Camouflaged target detection based on multimodal image input pixel-level fusion

Key words: Camouflaged target detection; Pixel-level fusion; Anchor box optimization; Loss function; Multispectral dataset

Corresponding author: Jie LAI
E-mail: laijie@hrbeu.edu.cn
ORCID: https://orcid.org/0009-0005-7918-0941

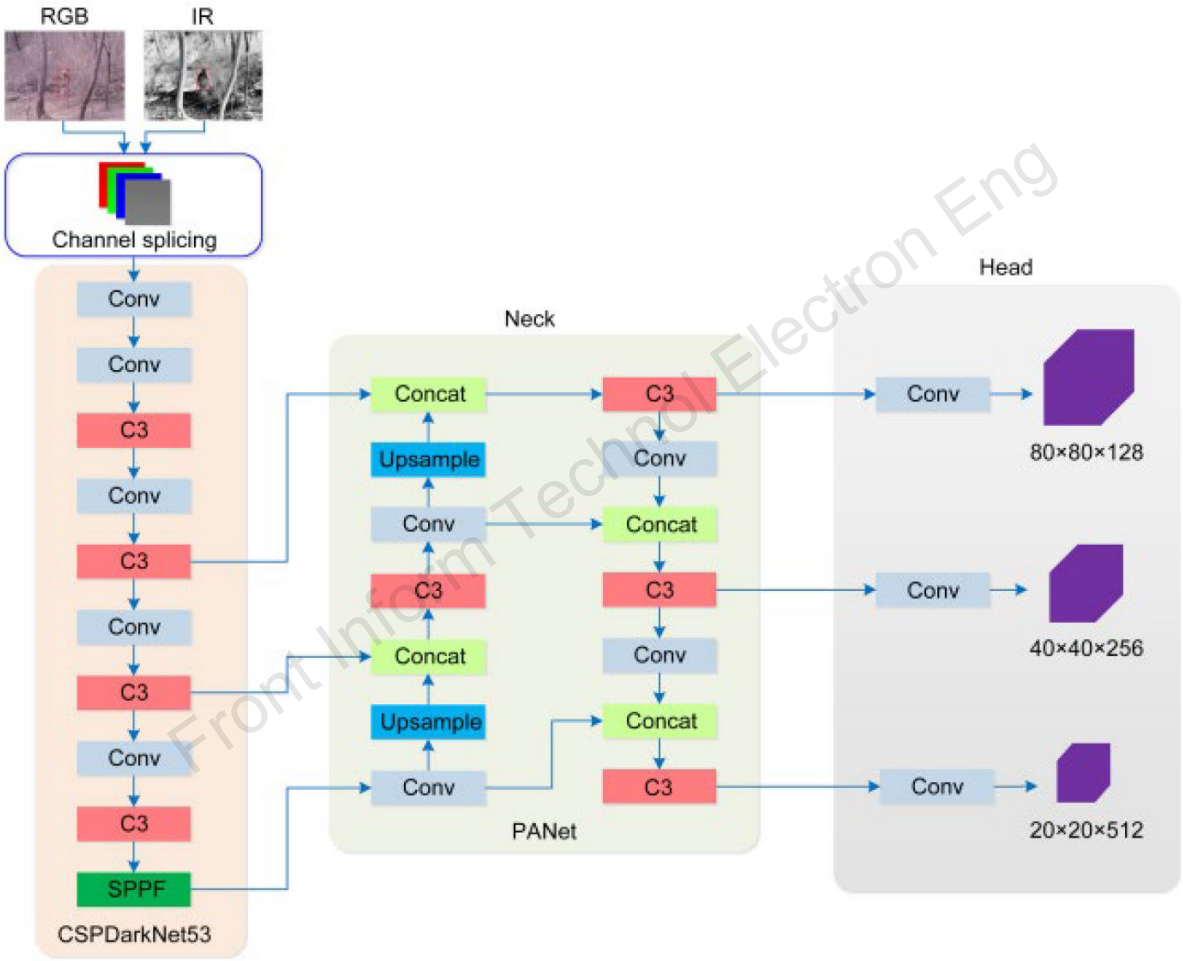
Motivation

- Camouflage means diversification. In today's battlefield, conventional single camouflage has difficulty in meeting the needs of practical applications. Military targets must counter not only visible light band reconnaissance, but also infrared (IR) and other band threats. It becomes important to reveal the deep mechanism of camouflaged targets efficiently and accurately.
- The visible light band camouflaged target has rich visual features such as brightness and texture, but the position information is blurred. The camouflaged target's position in the IR band is obvious, but textural information is scarce. It becomes important to use the complementary information between the two modal images to enhance the camouflage target detection.

Main idea

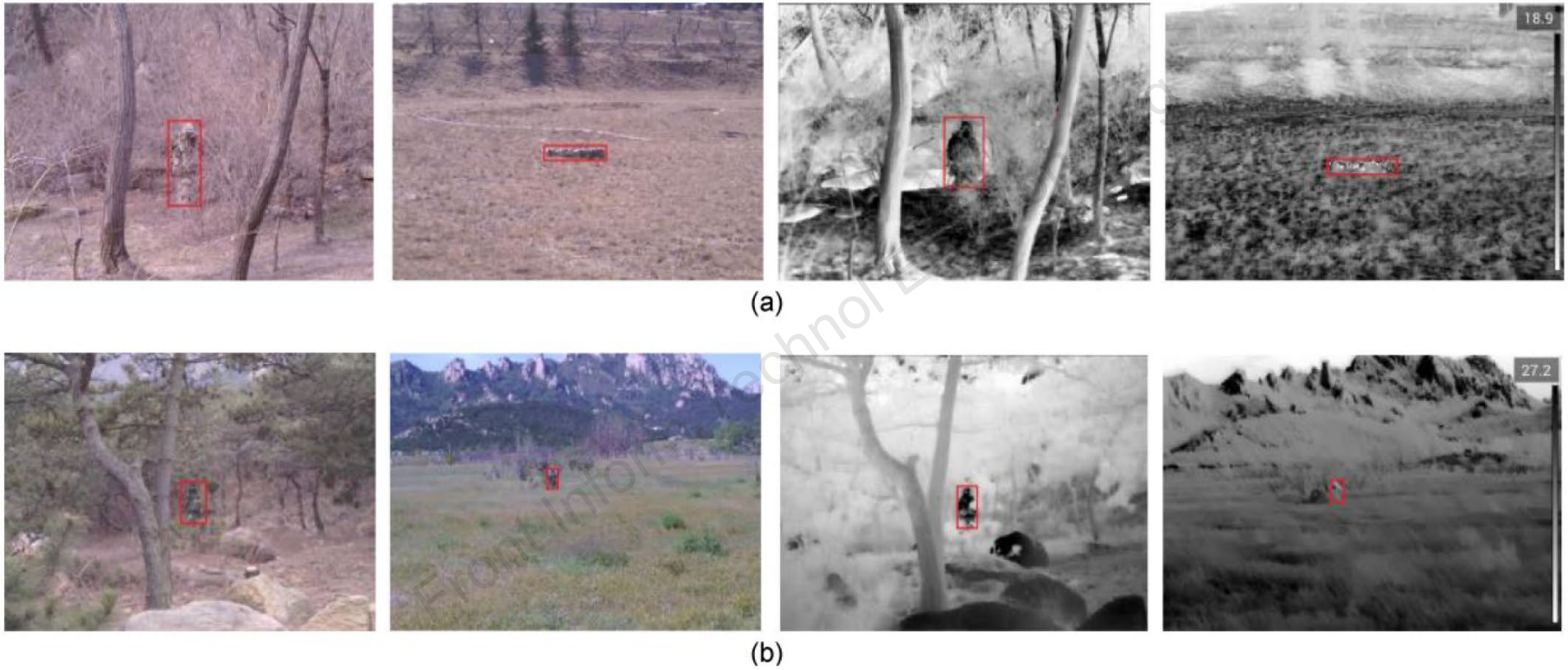
- Multispectral data collection is created which includes RGB and IR images of camouflaged targets in the same state space.
- An end-to-end multimodal image fusion technique is proposed, and a military camouflage target evaluation system is built.
- The loss function is created and the *K*-Means++ clustering technique is used to optimize the target anchor frame in the dataset to increase camouflage personnel detection accuracy and robustness.

Framework



Overview of the block diagram of the pixel-level fusion method based on the four-channel port (RGB: visible image; IR: thermal infrared image)

Multispectral dataset

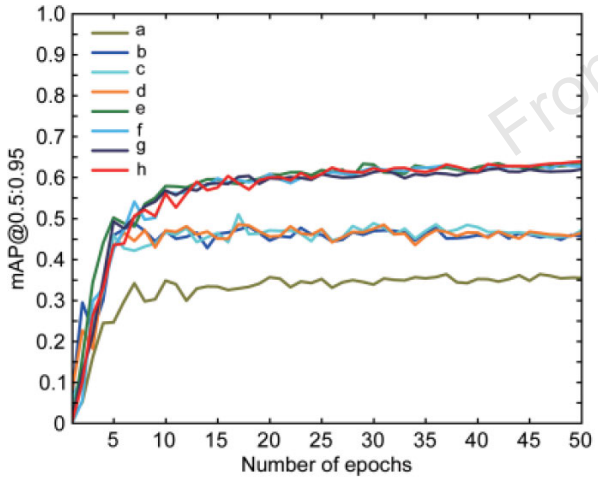


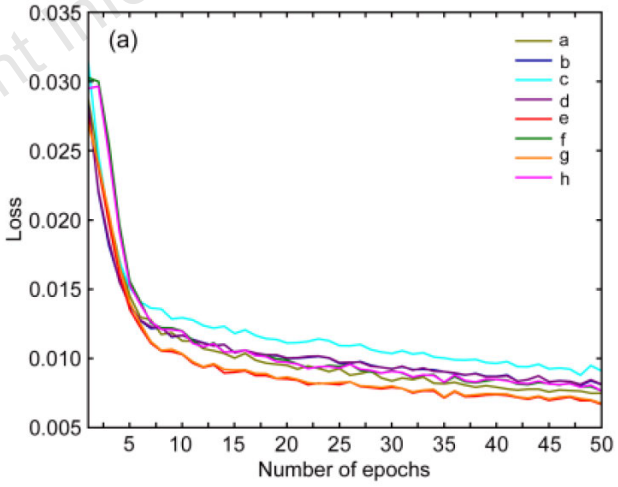
Dataset of multimodal image camouflaged targets: (a) desert setting; (b) woodland setting (left two, visible camouflaged targets; right two, infrared camouflaged targets)

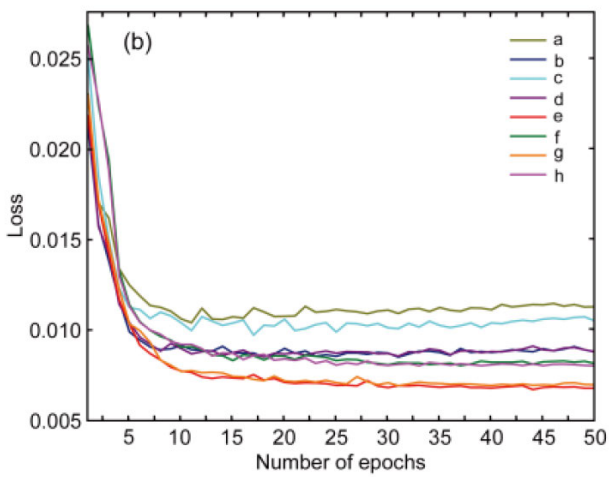
Major results

Table 3 Performance comparison of different experimental groups

Class	Number of parameters ($\times 10^6$)	GFLOPs	P	R	mAP@0.5:0.95	Frame rate (frame/s)	MA	FA
a	6.7	15.9	0.894	0.813	0.364	78	0.149	0.037
b	6.7	15.9	0.896	0.817	0.488	76	0.205	0.009
c	6.7	15.9	0.906	0.847	0.511	77	0.187	0.009
d	6.7	15.9	0.926	0.854	0.485	79	0.177	0.019
e	7.0	16.2	0.977	0.964	0.635	71	0.131	0.018
f	7.0	16.2	0.968	0.964	0.633	70	0.093	0
g	7.0	16.2	0.968	0.967	0.622	71	0.084	0.009
h	7.0	16.2	0.965	0.966	0.639	67	0.075	0.019







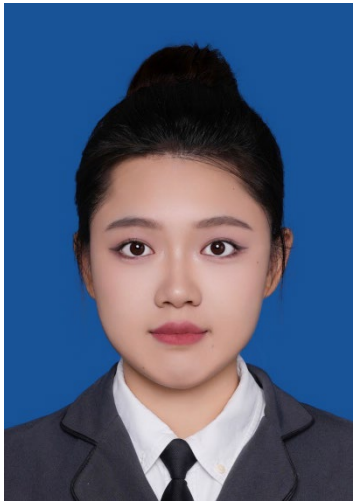
Conclusions

- Experimental results demonstrated that pixel-level fusion of optical and IR images in the same state space of the camouflaged target can effectively improve the feature information richness of the camouflaged target and fully use the target identification algorithm's potential.
- The anchor frame optimization and loss function were employed to further optimize the model's convergence process and improve the algorithm's robustness.

Ruihui PENG received his PhD degree in engineering from Naval University of Engineering in Wuhan, China, in 2010. Now he is a professional and technical information and communication engineering professor at Harbin Engineering University. His research interests include radar signal processing, radar target characteristics, and radar anti-jamming.

Jie LAI received his MS degree in electronic information from Harbin Engineering University in 2024 and is currently studying in Northwestern Polytechnical University. Currently, he is pursuing his PhD degree in electronic science and technology. His research interests include deep learning and multi-source information fusion.

Dianxing SUN received his PhD degree in information and communication engineering from Naval Aeronautical University in Yantai, China, in 2015. He is currently a professional and technical information and communication engineering professor at Harbin Engineering University. His research interests include radar signal processing, radar data processing, and radar anti-jamming.



Xueting YANG received her BS degree in communication engineering from Shenyang Aerospace University, China, in 2022. Now she is pursuing her MS degree in Harbin Engineering University. Her main research interests include deep learning and multi-modal image fusion.

Yingjuan SONG received her BS degree in communication engineering from Ludong University in 2022 and is currently pursuing her MS degree in information and communication engineering in Harbin Engineering University. Her research interests include radar signal processing and radar anti-jamming.

Wei GUO received his BS degree in electronic information engineering from Nanchang Hangkong University in 2022 and is currently pursuing his MS degree in information and communication engineering in Harbin Engineering University. His research interests include multi-sensor and targeting tracking.