Check for updates

A few-shot oil spill segmentation network guided by multi-scale feature similarity modeling

Lingfei Shi[1,2], Xianhu Wei[3,4], Kun Yang[1,2]* and Gong Chen[1,2]

[1]Faculty of Geography, Yunnan Normal University, Kunming, China, [2]The Engineering Research Center of GIS Technology in Western China of Ministry of Education of China, Yunnan Normal University, Kunming, China, [3]Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, [4]Sino-Africa Joint Research Center, Chinese Academy of Sciences, Wuhan, Hubei, China

Segmentation of oil spills with few-shot samples using UAV optical and SAR images is crucial for enhancing the efficiency of oil spill monitoring. Current oil spill semantic segmentation predominantly relies on SAR images, rendering it relatively data-dependent. We propose a flexible and scalable few-shot oil spill segmentation network that transitions from UAV optical images to SAR images based on the image similarity of oil spill regions in both types of images. Specifically, we introduce an Adaptive Feature Enhancement Module (AFEM) between the support set branch and the query set branch. This module leverages the precise oil spill information from the UAV optical image support set to derive initial oil spill templates and subsequently refines and updates the query oil spill templates through training to guide the segmentation of SAR oil spills with limited samples. Additionally, to fully exploit information from both low and high-level features, we design a Feature Fusion Module (FFM) to merge these features. Finally, the experimental results demonstrate the effectiveness of our network in enhancing the performance of UAV optical-to-SAR oil spill segmentation with few samples. Notably, the SAR oil spill detection accuracy reaches 75.88% in 5-shot experiments, representing an average improvement of 5.3% over the optimal baseline model accuracy.

KEYWORDS

oil spill, UAV, deep learning, few-shot, segmentation

1 Introduction

With the intensification of marine resource exploitation and the expansion of marine transportation, the marine environment faces increasingly severe threats (Sun et al., 2023). Among them, marine oil spills have long-term impacts on the marine environment, are challenging to treat, and are difficult to detect (Habibi and Pourjavadi, 2022; Wang et al., 2023). Consequently, some scholars have focused on developing efficient and generalized oil spill detection models.

While traditional oil spill monitoring can accurately detect oil spills, it is limited in scope and time-consuming. Remote sensing technology is essential for detecting and monitoring oil spills ranging from microscopic to large-scale. A variety of satellite images, including those used for optical, microwave, and infrared remote sensing, are employed for this purpose (Ma et al., 2023; Mohammadiun et al., 2021; Al-Ruzouq et al., 2020). Optical remote sensing data can be acquired at a low cost; for example, drones can be used to acquire high-resolution regional oil spill imagery in near-shore waters of harbors. However, their data coverage is limited, making it challenging to utilize them in the complex meteorological environments of offshore oceans (Chen et al., 2024). Microwave and infrared remote sensing offer the advantage of wide-range and all-weather data acquisition. However, the resolution of the data is low, and the cost of acquisition is high, which makes data processing challenging (Chaturvedi et al., 2020; Song et al., 2020). Additionally, SAR data can be interfered with by various types of land features in harbor areas and on land, reducing the accuracy of oil spill detection (Fan and Liu, 2023). Hence, in the age of data-driven artificial intelligence, there is a need to investigate a deep learning based semantic segmentation network that performs well in oil spill detection using both UAV optical and SAR data with few shot samples. This study aims to enhance the accuracy and efficiency of oil spill extraction from multi-source, multi-modal remote sensing images while minimizing the model's data dependency, an area currently underexplored in oil spill detection research.

Data-driven detection methods, particularly semantic segmentation, aim to predict category labels for new objects from a pixel-level perspective (Wang et al., 2021). This task has numerous applications, including scene understanding, medical diagnosis, and remote sensing information extraction, among others (Yuan et al., 2021; Asgari Taghanaki et al., 2021). Unlike object detection methods for oil spill detection, semantic segmentation divides pixels of the oil spill area. This approach facilitates the measurement of the oil spill area's size and its impact on neighboring areas.

Following a literature review, current research on semantic segmentation of oil spills still primarily relies on SAR data. The main approach involves enhancing the model's feature extraction capability to improve segmentation accuracy (Soh et al., 2024; Hasimoto-Beltran et al., 2023; Fan et al., 2023; Bianchi et al., 2020). Chen (Chen et al., 2023) developed an effective segmentation network named DGNet for oil spill segmentation. This network incorporates the intrinsic distribution of backscatter values in SAR images. DGNet consists of two deep neural modules that operate interactively. The inference module infers potential feature variables from SAR images, while the generation module generates an oil spill segmentation map by using these potential feature variables as input. The effectiveness of DGNet is demonstrated through comparison with other state-of-the-art methods using a large amount of Sentinel-1 oil spill data. On the other hand, Lau (Lau and Huang, 2024) combines geographic information with remote sensing imagery for oil spill segmentation. The study explores a deep learning approach for oil spill detection and segmentation in nearshore fields, achieving overall accuracies of 77.01% and 89.02%, respectively. They

subsequently developed a Geographic Information System (GIS) based validation system to enhance the deep learning model's performance, leading to an overall accuracy of 90.78% for the Mask R-CNN model. However, the model's performance can be limited in some cases due to the unavailability of accurate geographic information under subjective and objective conditions. Some other researches have considered the remote sensing physical characteristics of marine oil spill in designing the model (Fan et al., 2024; Amri et al., 2022). For example, Chen (Chen et al., 2023) proposed a new segmentation method for marine oil spill SAR image, which considers the physical characteristics of SAR image and oil spill segmentation at the same time in order to realize the accurate oil spill segmentation. Secondly, a deep neural segmentation network is constructed by combining the features of oil spill backscattering values in SAR images, which is capable of segmenting oil spill SAR images including similar points. The experimental results show that the method can identify the oil spill area better with less samples. Li (Li et al., 2021) proposed a multi-scale conditional adversarial network (MCAN), designed for oil spill detection using limited data. This network employs a coarse-to-fine training approach to capture both global and local features, and it demonstrates accurate detection performance even when trained on only four oil spill observation images. Li (Li et al., 2022) also proposed an adversarial ConvLSTM network (ACLN) for correcting predicted wind fields to match reanalysis data. By using a correction-discriminator framework, the network improved the accuracy of real-time wind field estimation and demonstrated its effectiveness in oil spill drift prediction, significantly reducing errors compared to traditional methods. Overall, generative adversarial networks require sophisticated training techniques. However, current oil spill detection methods based on the physical mechanism of remote sensing have limited transferability across different domains, and these models are still reliant on data.

The acquisition cost of optical image is low and the visibility is strong. However, satellite based optical imaging is less used in oil spill studies than microwave imaging because it relies on favorable weather conditions. However, UAVs capture images at low altitudes with minimal cloud interference and high resolution, making them a crucial data source for oil spill detection. Zahra (Ghorbani and Behzadan, 2020) annotated UAV-acquired offshore oil spill images, created an optical semantic segmentation dataset, and utilized an improved Mask-RCNN to identify the oil spill area, achieving a 60% accuracy. Ngoc (Bui et al., 2024) collected patrol videos taken by UAV to create an oil spill optical dataset. This dataset was used to train the DaNet deep learning model for oil spill segmentation and classification. The results demonstrated that the DaNet model achieved an average accuracy of 83.48% and mIoU of 72.54%.

From another perspective, supervised learning-based oil spill detection is currently mainstream. However, its strong data dependence and high cost of data annotation pose challenges. Furthermore, the weak migration performance of models hinders timely access to oil spill information. As a result, some scholars are now focusing on weakly supervised oil spill detection methods. For example, Kang (Kang et al., 2023) proposed a self-supervised spectral-spatial transformer network (SSTNet) for hyperspectral oil spill detection. Initially, a transformer-based comparative

learning network extracts deep discriminative features, which are then transferred to the downstream classification network for fine-tuning. Experimental results on a self-constructed hyperspectral oil spill database (HOSD) demonstrate the method's superior performance over existing techniques in distinguishing various oil spill types, including thick oil, thin oil, sheen, and seawater. However, the network's fine-tuning process is complex.

Recently, few-shot semantic segmentation has emerged as a hot topic in current research due to its powerful cross-domain few-shot adaptive segmentation capability (Zhou et al., 2022). Existing research on few-shot learning primarily focuses on visual information extraction (Zheng et al., 2022; Kang et al., 2023). The main research methods include fine-tuning (Sun et al., 2023), metric learning (Jiang et al., 2021), and meta-learning (Khadka et al., 2022). Among them, meta-learning is currently the mainstream approach for few-shot semantic segmentation, typically employing a two-branch architecture consisting of a support branch and a query branch. The support branch extracts feature from the support image and its semantic mask in the source domain, while the query branch uses these features to match with the target domain image, thereby learning new knowledge and completing the cross-domain semantic segmentation with few samples (Lang et al., 2023).

Motivated by the latest research direction and considering the shortcomings of oil spill detection research, we proposed a study on multi-scale and few-shot oil spill semantic segmentation by combining UAV optical and SAR images. Utilize the distinctive image features of oil spill areas in UAV optical and SAR images. Leverage the deep learning modeling capabilities to guide the model in establishing feature correlations of oil spill areas in both image types. Enhance the model's cross-domain oil spill detection capabilities.

Since UAV optical images can provide more spatial and semantic information, UAV optical images are used as support sets for few-shot sample training, and SAR images are used as query sets for testing. Firstly, a cross-domain adaptive module is designed to create and update feature templates, updating respective category features through similarity matching with the support and query sets. Secondly, to fully leverage multi-scale features, we employ the Adaptive Feature Enhancement Module for high-level and low-level features. Finally, the Weighted Multiscale Feature Fusion Module is designed to fuse the multi-scale features.

The main contributions of this paper are as follows:

1. We propose a multi-scale cross-domain few-shot oil spill detection network to quickly self-learn the common features of oil spills in UAV optical and SAR images. This is achieved by designing corresponding adapter modules in the few-shot semantic segmentation task. We experimentally demonstrate that the architecture can be integrated into different feature extraction networks, improving the network's cross-domain adaptive capability. In addition, our network is flexible. First, different types of feature extraction modules could be selected according to the actual situation of oil spill detection, making the entire network easier to deploy on edge devices. Secondly, the available data can flexibly select the training mode of the model, which can either choose the full supervision of a single branch or the two-branch training of knowledge transfer with few samples. In this way, the oil spill detection requirements of different types of images such as UAV optical and SAR can be effectively met, and the detection efficiency is greatly improved.

2. We designed the cross-domain Adaptive Feature Enhancement (AFEM) module and the multi-scale feature fusion module (FFM) to enhance the extraction of multi-scale cross-domain similar features of the model.

3. In this paper, through the innovative method of setting the similarity matching of oil spill template library, the knowledge transfer from UAV optical to SAR oil spill image is realized efficiently by learning the common features of UAV optical and SAR oil spill images, and the oil spill segmentation of few-shot SAR images is completed. Moreover, the segmentation accuracy is higher than that of the current mainstream few-shot semantic segmentation benchmark model.

The rest of the paper is organized as follows: Section 2 will introduce the dataset used for the study and the processing methods. Section 3 will present the specific design details of the few-shot oil spill detection network. Section 4 will present the experimental design details and experimental results. Section 5 will discuss the design ideas and future research directions of the network in light of the experimental results. Finally, Section 6 will summarize the entire research.

2 Datasets

2.1 UAV optical oil spill dataset

Thomas (De Kerf et al., 2024) recently released the UAV-acquired harbor oil spill optical dataset, comprising 1268 images. As shown in Figure 1, the dataset is categorized into oil, water, and other categories, and it ensures that each sample image contains all semantic categories, thereby ensuring the balance of the samples. This dataset fills a gap in publicly available oil spill optical semantic segmentation datasets and serves as a crucial resource for environmental protection in harbor environments. To maintain consistency with the semantic segmentation categories of the SAR oil spill detection dataset, we harmonized the category segmentation of the optical dataset, merging waters, other categories, and unknown categories into a single "other" category.

2.2 SAR oil spill dataset

The first oil spill detection SAR image dataset used in this study was developed by Krestenitis (Krestenitis et al., 2019), as shown in Figure 2. The dataset includes Sentinel-1 SAR image representations of offshore oil spills and corresponding labels. The training set comprises 1002 images, and the test set contains 110 images with corresponding labels. The spatial resolution of the image is 10m. To
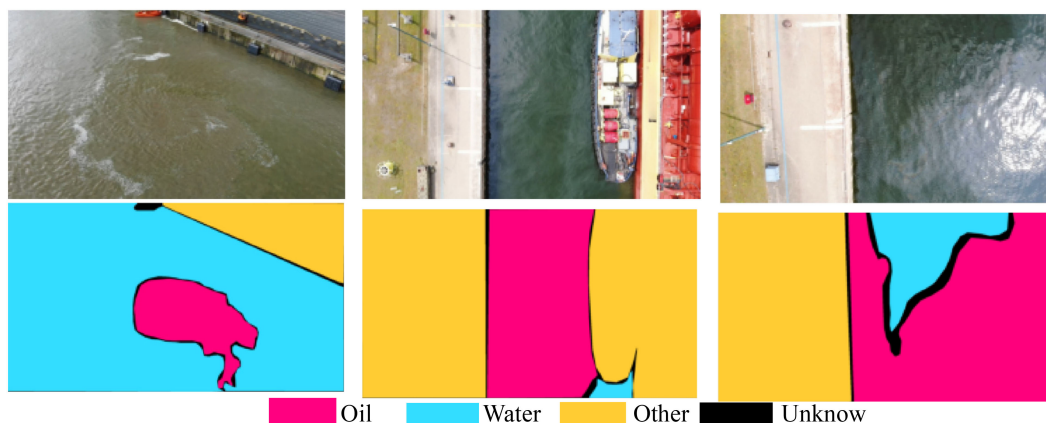
FIGURE 1
Semantic segmentation dataset of UAV optical images.

emphasize the recognition of oil spill areas, this study merged the semantic segmentation task into two categories: oil spill areas and other. Data augmentation was employed to increase the training dataset size, with randomized horizontal and vertical flips applied only to the training set. The images were normalized using the mean and standard deviation of the training set before being used as inputs to the various models in this study.

In Figure 3, to verify the robustness of the model, the SOS oil spill detection dataset released by Zhu (Zhu et al., 2022) was selected for this study. It comprises two parts: (1) 14 PALSAR images of the Gulf of Mexico region with a pixel spacing of 12.5 m and HH polarization; and (2) 7 Sentinel 1A images of the Persian Gulf region, with a spatial resolution of 5 m × 20 m and VV polarization. To enhance dataset diversity, original images underwent operations such as rotation and random cropping, resulting in 6456 oil spill images from the 21 original images. Subsequently, the data were manually labeled and subjected to expert sampling, resulting in a final dataset of 8070 labeled oil spill images.

Finally, to improve the training efficiency of the model, the dimensions of the images and annotations in all three datasets were resized to 256×256.

Figure 4 presents a typical example of UAV optical oil spill imagery. A comparison of Figures 3, 4 shows that the oil spill areas in both SAR and UAV optical images display a patchy distribution and are distinctly different from the ocean background. These characteristics align with the similarity criteria necessary for transfer learning in computer vision across different image sources.

3 Methods

3.1 Overview of few-shot semantic segmentation network

In this section, we first introduce the overall network for semantic segmentation of oil spills in UAV optical and SAR
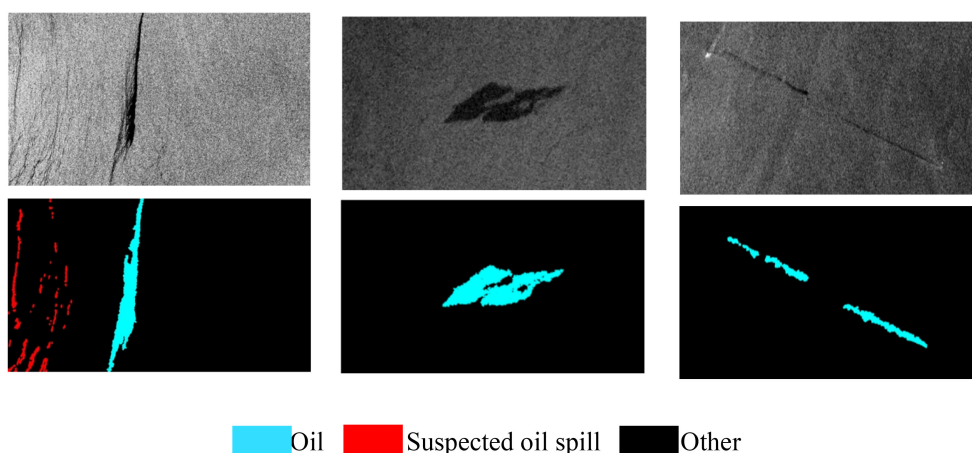

FIGURE 2
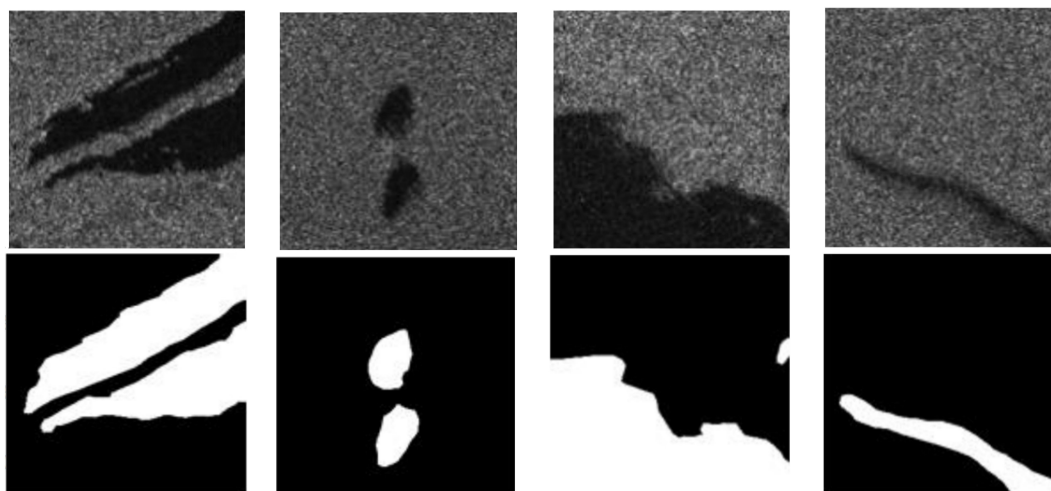Semantic segmentation dataset of the first SAR images.

FIGURE 3
Semantic segmentation dataset of the second SAR images.

images, and then focus on two main components of the network: the Adaptive Feature Enhancement Module (AFEM) and the Feature Fusion Module (FFM).

The semantic segmentation network for UAV optical and SAR images with few samples designed in this study mainly focuses on the detection of oil spill categories, which belongs to the one-way k-shot detection task, meaning that only k samples are used for training. Therefore, in the model training stage, c support set samples $\{I_S^c, M_S^c\}_{c=1}^c$ are selected in the UAV optical oil spill dataset $D_{optical}$, c is the number of training sets for visible oil spill images. Moreover, one query set sample $\{I_q, M_q\}$ is selected in the SAR oil spill dataset D_{SAR}. The model obtains the ability to predict M_q by training on the support ensemble sample and I_q. The UAV optical and SAR image less sample oil spill semantic segmentation network, as shown in Figure 5, feeds the UAV optical oil spill semantic segmentation support sample set $\{I_S^c, M_S^c\}_{c=1}^c$ and the SAR oil spill semantic segmentation query sample I_q into the encoder of ResNet50 for feature extraction, respectively. Since the low-level features retain more structural features such as texture and edges of the image while the high-level features retain more abstract semantic relational features, all of the information are important for sample less semantic segmentation. Thus, the model obtains the encoder's low-level features F_S^{low}, F_q^{low} and high-level features F_S^{high}, F_q^{high} at two levels respectively. Then, F_S^{low}, F_q^{low} and M_S are fed into our proposed adaptive feature enhancement module for feature interaction, thereby obtaining the class-specific features $F_S^{low\star}$ and

$F_q^{low\star}$ of the oil spill. Subsequently, the original features are merged with the oil spill-specific features by means of residual concatenation to maximize information retention. For the high-level features F_S^{high} and dF_q^{high}, the same approach of feature interaction and feature merging is used. Finally, the low-level and high-level features after the adaptive feature enhancement module and feature merging are fed into the feature fusion module for the interaction of high-level and low-level features, which makes full use of the advantages of the different levels of features of the oil spill area in the UAV optical and SAR images, and improves the model's performance of semantic segmentation of the oil spill area with few samples.

3.2 Adaptive feature enhancement module

The Adaptive Feature Enhancement Module proposed by us is shown in Figure 6, which consists of generating and updating the oil spill class feature template library and oil spill feature enhancement.

Here we take the low-level feature as an example, and the high-level feature performs the same operation. For the oil spill semantic segmentation task of low-sample SAR images with c UAV optical support samples, the first task is to generate the oil spill feature template database. The support set feature $F_S^{low} \in U^{c \times d \times h \times w}$ and the support set image mask $M_S \in U^{c \times h \times w}$ (where d represents the dimension of the feature, h and w represent the height and width of
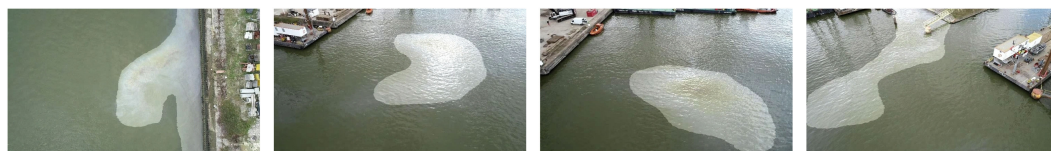

FIGURE 4
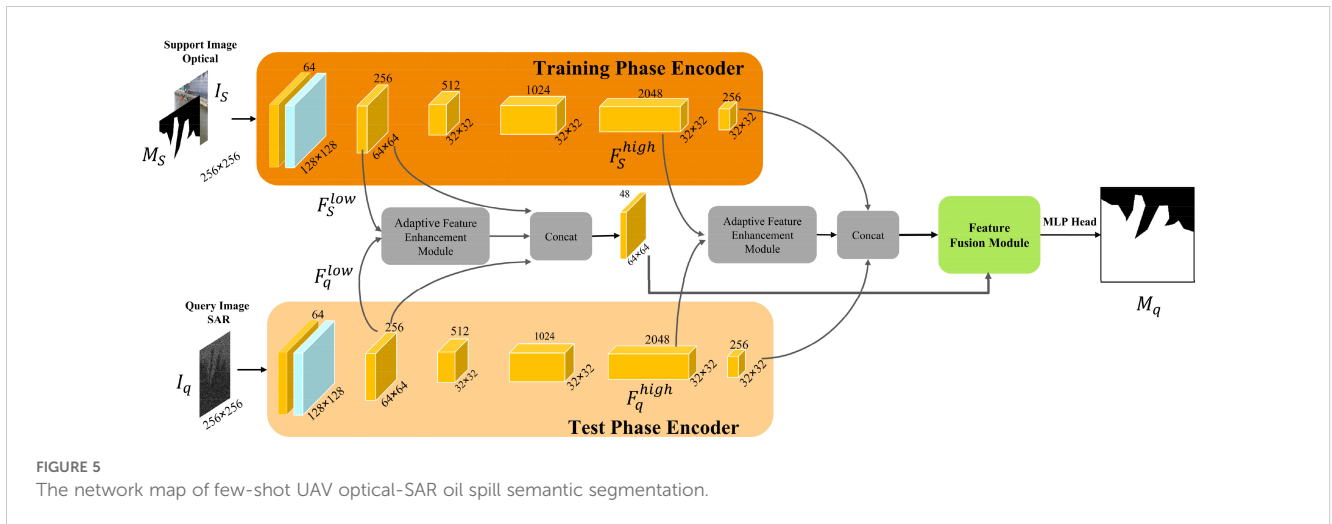Example of typical UAV optical oil spill images.

FIGURE 5
The network map of few-shot UAV optical-SAR oil spill semantic segmentation.

the feature) are used to generate a specific class of features for the oil spill area. Initially, an initial SAR image oil spill template library is obtained after the branch training of the query set. $P = \{P_1, P_2, ..., P_m\} \in U^{m \times d}$, where m represents the number of categories of templates, which is mainly related to the morphological categories of oil spill areas in SAR images. The specific type of oil spill template P_t^{low} is obtained by supporting set features F_S^{low} and M_S through pixel-by-pixel multiplication in spatial dimension, and the calculation formula is as follows:

$$P_t^{low} = Mean(F_S^{low} . M_S) \qquad (1)$$

Since we set the semantic category of the non-oil spill area in mask M_S to 0 and the semantic category of the oil spill area to 1, only the features of the oil spill area are retained after the template generation operation. This way, interference from the background on the oil spill feature recognition can be avoided. In the model test stage, similarity matching was used to update the template P_m of different oil spill types, and the specific calculation formula was as follows.

$$P_m^{train} = (1 - \alpha) \times l_2(P_t^{low}) + \alpha \times l_2(P_m) \qquad (2)$$

$$P_m^{test} = Argmax(l_2(P_m^{train}) \cdot l_2(P_t^{low})) \qquad (3)$$

Finally, according to experimental experience, using the top five templates with the highest similarity can obtain the most comprehensive features more efficiently, and can prevent feature overfitting. Therefore, only the top five templates with the highest similarity are selected as the final UAV optical-SAR oil spill template library. This way, each oil spill template P_m will be able to express the features of different oil spill types well in the current feature representation space.

In addition, the model can learn the invariant oil spill template feature when transforming from UAV optical to SAR oil spill feature space. We can calculate the similarity between the UAV optical support set feature F_S^{low} and the oil spill class template P_m obtained earlier, and the calculation formula is as follows.

$$E_m^{low} = Min(6, Max(0, l_2(F_S^{low}) \cdot l_2(P_i) \times \sqrt{d})) \qquad (4)$$

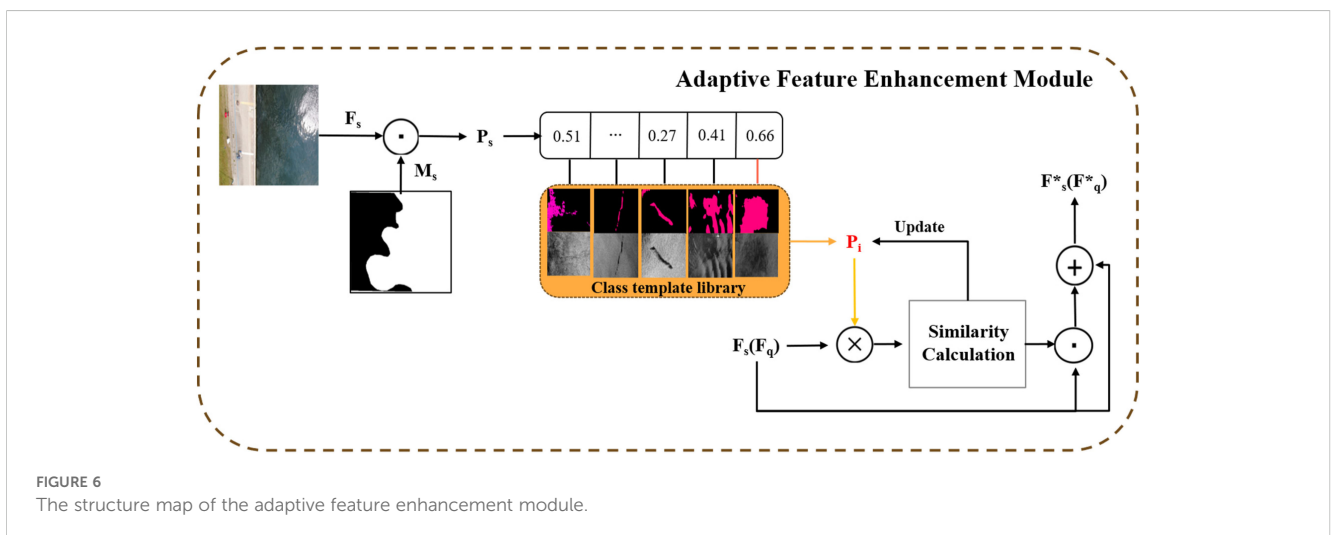$$F_S^{low\star} = E_m^{low} \cdot F_S^{low} + F_S^{low} \qquad (5)$$

FIGURE 6
The structure map of the adaptive feature enhancement module.

After the above steps, the support set feature F_S^{low} is constrained by the previously obtained oil spill template, and here only the support set feature F_S^{low} is described as an example, the same procedure will be applied to the query set feature F_q^{low} based on the SAR oil spill image, and to the encoder's high-level feature outputs F_S^{high} and F_q^{high}.

3.3 Feature fusion module

Although UAV optical images and SAR images have similar features in the oil spill area, the two different data characteristics pose great challenges to transfer learning of few-shot oil spill segmentation. Furthermore, the distribution of the enhanced features $F_S^*(F_q^*)$ in the high-level and low-level features and the original features $F_S(F_q)$ has changed significantly after the previous feature enhancement. Based on this, instead of fusing only single features of the support set and query set images as in previous sparse sample semantic segmentation studies, this study fully considers the contour information in the low-level features and the abstract semantic information in the high-level features (Hyperspectral image instance segmentation using spectral-spatial feature pyramid network). An adaptive weighted high and low-level feature fusion module is proposed, and the specific structural diagram is shown in Figure 7.

First, a new feature C is generated from the merged high-level feature B after upsampling and the low-level feature A. The number of channels in C consists of A and B'. Then, C is reduced by 1×1 convolution to obtain D. Next, two projective fully connected layers with weight parameters W_{down} and W_{up} are used to compress and expand the channel dimension of the feature, with a ReLU activation function used between the two fully connected layers to improve the nonlinear characteristics of the feature. Finally, the output features are obtained by taking the dot product with D to obtain E. Feature E is then connected with feature A using residuals to obtain the final fusion feature F. The overall calculation formula is as follows:

$$F = ReLU(Conv_{1\times1}(\text{UP}(B) \cdot A) \cdot W_{down})$$
$$\cdot W_{up} * Conv_{1\times1}(\text{UP}(B) \cdot A) + A \qquad (6)$$

4 Experiment

4.1 Experimental setup

Our network is built on PyTorch 2.0 and trained on an RTX 3090 GPU with 24GB of memory. 70% of the UAV optical dataset (887) are selected as the training set and 10% (126) are selected as the validation set. The initial model weights and oil spill template library are obtained by training in the support set branches of the model. The training batch size for the experiment is set to 8. As shown in Equation 7, we choose the weighted binary cross entropy as the loss function. The importance of the loss in the oil spill area is increased by 20%, making the model pay more attention to identifying the oil spill area. The model uses AdamW as the optimizer, with an optimization momentum of 0.95, a learning rate of 0.001, and a weight decay coefficient of 0.001. After training the model support branches using UAV optical images, we begin the second phase of transfer learning training. At this stage, the model weights trained on the UAV optical oil spill dataset were transferred to the feature extraction branch of the SAR oil spill model. Only 1, 3, and 5 samples are randomly selected from the SAR oil spill data for the few-shot query branch fine-tuning training and last test, respectively, and fine-tuned together with the previous support branches. Simultaneously, the oil spill template library is updated, and fine-tuning is limited to no more than 1500 iterations each time. Each trial is evaluated based on the results of five random sample trials, and the results of 1-shot and 5-shot SAR oil spill detection with small samples are obtained.

$$L_{wbce} = -1.2 \times y_{true}log(y_{pred}) - (1 - y_{true})log(1 - y_{pred}) \qquad (7)$$

4.2 Experimental result

In order to comprehensively evaluate the effectiveness of the network proposed in this study, we selected four mainstream few-shot semantic segmentation networks: MSANet (Iqbal et al., 2022), PFENet (Tian et al., 2022), HSNet (Min et al., 2021), and BAM

FIGURE 7
The structure map of the feature fusion module.
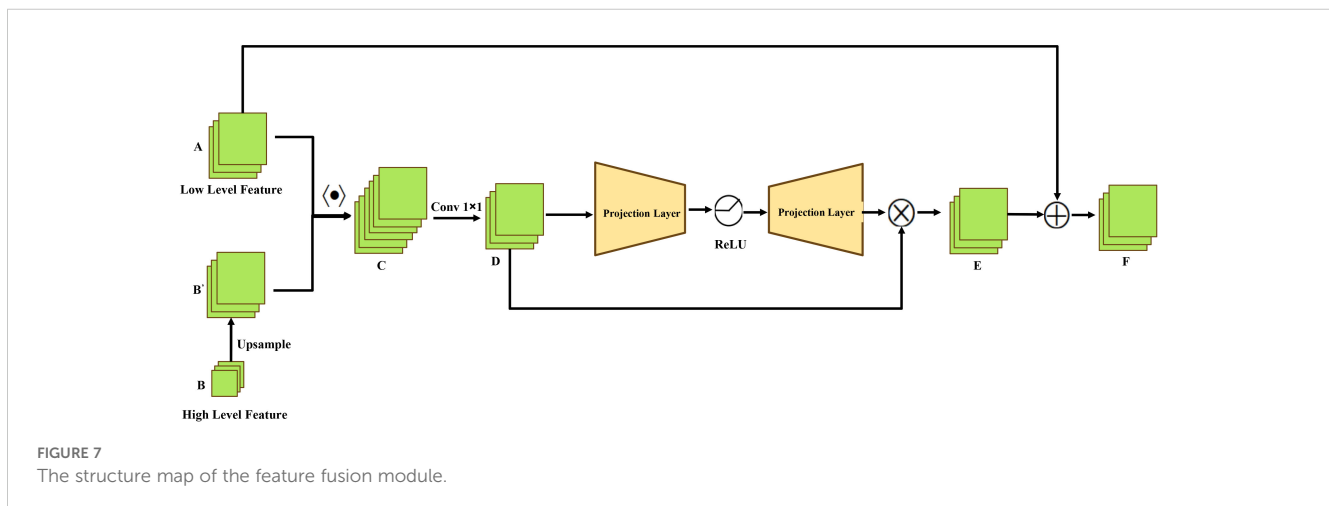
(Lang et al., 2022). VGG, ResNet, and DenseNet are then used as encoders for the network, respectively. The performance of the model is evaluated from both quantitative and qualitative perspectives.

In terms of quantitative evaluation, as shown in the Table 1, our network outperforms other small-sample semantic segmentation models in UAV optical-SAR oil spill detection tasks in both 1-shot and 5-shot settings. Compared with the SOTA model MSANet, the proposed network in this study achieves an average IoU improvement of 3.7%, 6.5%, and 8.2% in the 1-shot configuration of VGG16, ResNet50, and DenseNet121 backbone oil spill detection tasks, respectively. IoU increases averaged 3.5%, 6.2%, and 6.1% in the 5-shot configuration, respectively.

In terms of qualitative evaluation, as shown in Figure 8, we first generate an attention heat map for the task of SAR oil spill detection with few samples. It can be observed that after training with the support of the ensemble of UAV optical oil spill data, along with the adaptive feature enhancement module and the multi-scale feature fusion module, the output features can effectively focus on the oil spill areas distributed in different locations in the SAR image. This indicates that the two modules we designed facilitate feature migration, enabling the model to achieve a relatively strong ability to detect oil spills based on learning only a small amount of SAR oil spill data. This demonstrates the efficacy of the two modules in facilitating the migration of oil spill features, thereby enhancing the model's oil spill detection capability with few shot SAR oil spill data.

On the other hand, according to the previous experiments result, we found that the results of the 5-shot experiment were the most representative, so we chose the test results of the 5-shot experiment as the benchmark for model comparison. The Figure 9 illustrates the comparison results between several benchmark models and our

proposed network in the 5-shot SAR oil spill semantic segmentation task. It is evident from the figure that our model closely approximates the real label and effectively segments different types of oil spill areas. Overall, our model effectively focuses on the boundaries of oil spill regions with varying shapes, enabling more accurate extraction of the spill areas. Figure 9 shows that our model accurately extracts oil spill regions, with no internal holes appearing in the detected spill areas. This improves the overall integrity of the detection, and the spill boundaries are more precise without being overly smoothed.

We conducted ablation experiments to investigate the effects of the key components of the model, the Adaptive Feature Enhancement Module (AFEM) and the Feature Fusion Module (FFM), on the model's performance, as shown in the Table 2, Figure 9. The feature enhancement using the AFEM module resulted in the most significant model accuracy improvement of 1.12% compared to the strategy of directly merging features from both branches. This indicates that the interaction of oil spill features from UAV optical and SAR can fully learn the respective features of the two data types to enrich the model's information and improve its migration learning ability. Additionally, Utilizing the Oil Spill Template Library (OTL) fully explores the correlation between deep model feature modeling. It guides the establishment of feature similarity matching criteria between UAV optical and SAR images, thereby enhancing the few-shot SAR image oil spill detection capability. As shown in Table 2, the detection accuracy was increased by 1.55% after OTL was used. Additionally, the loss variation curve of the ablation model training function in Figure 10 shows that the model utilizing feature enhancement modules and multi-scale feature oil spill templates most effectively improves the cross-domain few-shot oil spill detection capability. This model also achieves the fastest training and convergence speed.

TABLE 1 IoU results of the few-shot oil spill semantic segmentation tasks.

Backbone	Method	1-shot (IoU)				5-shot (IoU)			
		Flod-0	Fold-1	Fold-2	Fold-3	Flod-0	Fold-1	Fold-2	Fold-3
VGG 16	MSANet	60.11	68.6	62.31	57.16	64.28	70.47	63.91	58.76
	PFENet	55.13	64.23	57.17	53.6	57.09	66.24	60.09	57.92
	HSNet	52.19	61.8	55.24	54.83	55.86	64.13	61.53	56.08
	BAM	60.05	65.72	58.1	56.4	62.11	69.78	62.28	62.27
	Ours	**63.11**	**69.65**	**62.56**	**61.86**	**64.22**	**72.63**	**65.08**	**64.38**
ResNet 50	MSANet	61.37	70.46	63.4	58.67	65.78	71.47	64.89	59.27
	PFENet	56.3	68.23	59.17	55.39	58.76	68.11	61.17	58.39
	HSNet	55.6	63.8	57.51	55.13	57.46	66.08	63.51	57.08
	BAM	62.18	69.72	60.24	58.18	62.59	71.27	64.19	62.87
	Ours	**64.69**	**72.65**	**66.78**	**68.78**	**66.09**	**75.43**	**67.08**	**69.83**
DenseNet 121	MSANet	61.92	70.88	62.67	59.34	66.13	71.8	65.14	62.59
	PFENet	57.15	69.27	61.05	54.89	58.8	70.56	61.33	60.89
	HSNet	56.05	63.36	58.21	56.08	57.96	65.17	63.21	58.6
	BAM	63.1	70.14	60.68	59.22	63.98	70.61	64.75	61.22
	Ours	**65.95**	**73.94**	**67.72**	**69.89**	**67.13**	**75.88**	**68.52**	**70.18**

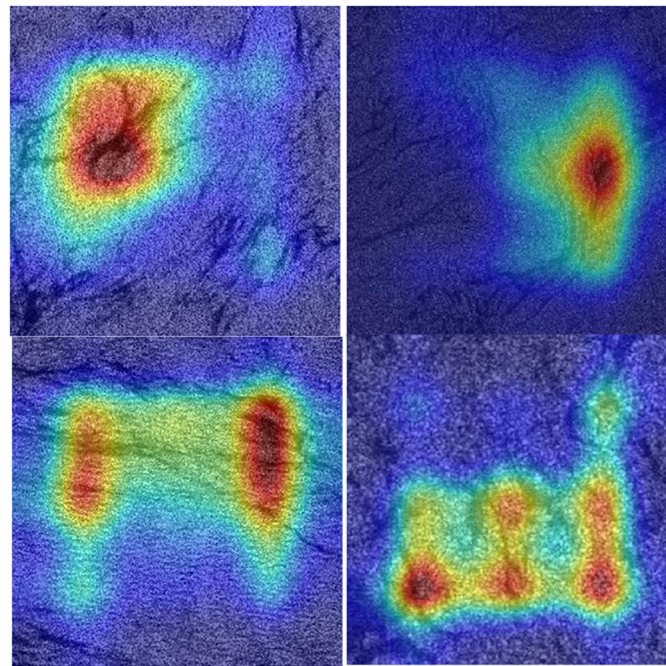The performance metrics of our proposed model are shown in bold.

FIGURE 8
The attention heat map of SAR oil spill images.

FIGURE 9
Visual comparison results of the first disaster semantic segmentation task.

TABLE 2 Ablation study with 5-shot semantic segmentation on SAR images.

AFEM	FFM	OTL	IoU%
	✓	–	70.61
✓		–	71.73
✓	✓	–	73.84
✓	✓	✓	75.39

5 Discussion

Based on the previous experimental results and analysis, this section discusses the effectiveness of the few-shot oil spill semantic segmentation network proposed in this thesis for UAV optical-SAR images from several aspects.

During the research of the paper, we identified similarities between UAV optical and SAR oil spill images, such as the patchy distribution of oil spill regions and the distinct texture and color differences between the oil spills and the background. Our model design is based on these principles, resulting in a framework for knowledge transfer and template matching from UAV optical to SAR images. Additionally, we observed that differences in image resolution between the two types of data also affect the knowledge transfer performance of the model. For instance, the optical images we used have a sub-meter resolution, suitable for detecting small sized oil spills, while SAR images have a 10-meter resolution, suitable for large-scale oil spill detection in adverse weather conditions. We took this into account in our model design by incorporating modules for extracting and fusing low-level and high-level features, which provides the model with receptive fields at various scales and enhances its ability to perceive oil spills of different sizes.

To further validate the transfer learning capability of our proposed model, we conducted an experiment in contrast to the previous one. We collected UAV oil spill images from the previous dataset and input them into the model trained on SAR oil spill data for fine-tuning and testing. The results, shown in the Figure 11, demonstrate that the model effectively extracts oil spill regions from
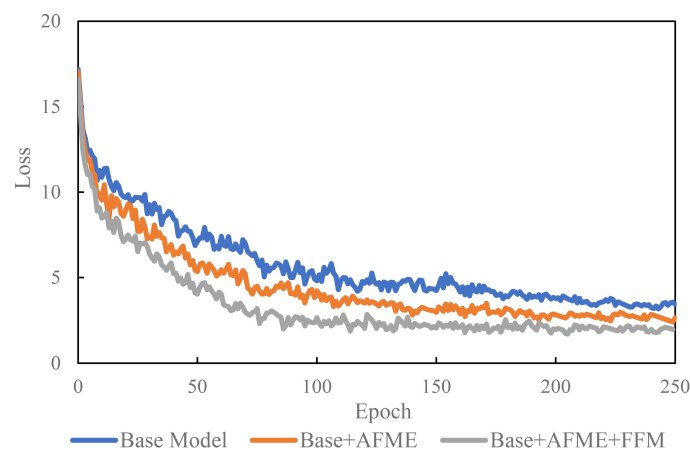

FIGURE 10
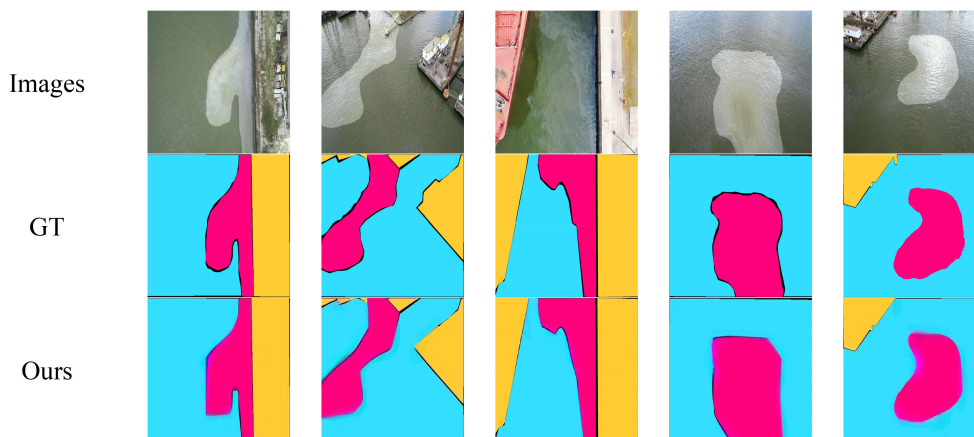Training loss curve of the ablation models.

FIGURE 11
Transfer learning results from SAR to UAV optical images.

UAV optical images. This indicates that our proposed model exhibits some transfer learning capabilities.

During the model design phase, this study first adopts the current mainstream two-branch architecture for few-shot semantic segmentation and sets up replaceable feature extraction encoders. It then considers the similarity of oil spill features in optical and SAR images, designing an optical-SAR adaptive feature enhancement module with an updatable oil spill template library. This module enables the formation of an optimal oil spill template library through few-shot SAR images, guiding cross-data-modality few-shot oil spill segmentation. Secondly, inspired by the concept of encoder feature extraction, this study fully exploits the information of low-level and high-level features. It designs a multi-scale feature fusion module that allows the model to adaptively learn the importance of features in different dimensions through weighting. This approach helps retain the most important features while removing redundant ones. The effectiveness of the network is demonstrated through few-shot semantic segmentation experiments.

Of course, our model has limitations. Firstly, although it establishes a connection between UAV optical and SAR oil spill images through similarity matching templates, improving the few-shot SAR oil spill detection capability, its segmentation accuracy is still lower than that of fully supervised oil spill semantic segmentation models. Additionally, due to the lack of incorporation of physical mechanism knowledge, the model's interpretability is relatively weak, and its cross-regional transferability has not been validated. Additionally, we found that the model's transfer learning capability between optical oil spill images from Sentinel and Landsat platforms and SAR images is relatively limited. This is because optical remote sensing images are more easily affected by complex ocean environments and weather conditions like clouds. In our experiments, the transfer model often misclassified waves and clouds as oil spills. This issue requires further investigation in future research.

In the future, we plan to design a few-shot semantic segmentation network based on the physical mechanisms of SAR oil spill imaging to improve cross-modal oil spill detection performance. Additionally, we will continuously collect oil spill semantic segmentation data from various regions for training and leverage the latest large models (SAM) to design prompt learning methods that incorporate expert knowledge, thereby enhancing model robustness. This network can serve as a reference for future research on few-shot semantic segmentation methods.

6 Conclusions

In this paper, we propose a new flexible semantic segmentation network for few-shot oil spills. Based on the current mainstream two-branch structure, we design the Adaptive Feature Enhancement Module, which can update the oil spill template library, to accurately guide and enhance features highly correlated with few-shot objects according to the oil spill image features of UAV optical and SAR. Additionally, we design the Feature Fusion Module for weighted feature fusion, fully considering the advantages of encoder low-level and high-level features, which improves the efficiency of feature information utilization. Finally, through comparison experiments with mainstream few-shot semantic segmentation models under different backbones, the results show that our network improves oil spill detection accuracy by an average of 5.3% over the state-of-the-art benchmark model. The model reduces the misclassification of oil spills with slender bar shapes.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

LS: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. XW: Writing – review & editing, Supervision. KY: Writing – review & editing, Supervision, Funding acquisition. GC: Writing – review & editing, Formal analysis.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer TM declared a shared affiliation with the author XW to the handling editor at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Al-Ruzouq, R., Gibril, M. B. A., Shanableh, A., Kais, A., Hamed, O., Al-Mansoori, S., et al. (2020). Sensors, features, and machine learning for oil spill detection and monitoring: A review. *Remote Sens.* 12, 1–42. doi: 10.3390/rs12203338

Amri, E., Dardouillet, P., Benoit, A., Courteille, H., Bolon, P., Dubucq, D., et al. (2022). Offshore oil slick detection: from photo-interpreter to explainable multi-modal deep learning models using SAR images and contextual data. *Remote Sens.* 14. doi: 10.3390/rs14153565. MDPI.

Asgari Taghanaki, S., Abhishek, K., Cohen, J. P., Cohen-Adad, J., and Hamarneh, G. (2021). Deep semantic segmentation of natural and medical images: A review. *Artif. Intell. Rev.* 54, 137–178. doi: 10.1007/s10462-020-09854-1. *Springer Science and Business Media B.V.*

Bianchi, F. M., Espeseth, M. M., and Borch, N. (2020). Large-scale detection and categorization of oil spills from SAR images with deep learning. *Remote Sens.* 12. doi: 10.3390/rs12142260

Bui, N. A., Oh, Y., and Lee, I. (2024). Oil spill detection and classification through deep learning and tailored data augmentation. *Int. J. Appl. Earth Obs. Geoinform.* 129, 103845.

Chaturvedi, S. K., Banerjee, S., and Lele, S. (2020). An assessment of oil spill detection using sentinel 1 SAR-C images. *J. Ocean Eng. Sci.* 5, 116–135. doi: 10.1016/j.joes.2019.09.004

Chen, F., Balzter, H., Zhou, F., Ren, P., and Zhou, H. (2023). "DGNet: distribution guided efficient learning for oil spill image segmentation," in *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 61. doi: 10.1109/TGRS.2023.3240579

Chen, Y., Yu, W., Zhou, Q., and Hu, H. (2024). "A novel feature enhancement and semantic segmentation scheme for identifying low-contrast ocean oil spills," in *Marine Pollution Bulletin*, vol. 198. (Elsevier Ltd). doi: 10.1016/j.marpolbul.2023.115874

De Kerf, T., Sels, S., Samsonova, S., and Vanlanduit, S. (2024). Oil spill drone: A dataset of drone-captured, segmented RGB images for oil spill detection in port environments. *arXiv preprint arXiv:2402.18202.* doi: 10.48550/arXiv.2402.18202

Fan, J., and Liu, C. (2023). "Multitask GANs for oil spill classification and semantic segmentation based on SAR images," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 16. 2532–2546. doi: 10.1109/JSTARS.2023.3249680

Fan, J., Sui, Z., and Wang, X. (2024). Multiphysical interpretable deep learning network for oil spill identification based on SAR images. *IEEE Trans. Geosci. Remote Sens.* 62, 1–15. doi: 10.1109/TGRS.2024.3357800

Fan, J., Zhang, S., Wang, X., and Xing, J. (2023). Multifeature semantic complementation network for marine oil spill localization and segmentation based on SAR images. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 16, 3771–3783. doi: 10.1109/JSTARS.2023.3264007

Ghorbani, Z., and Behzadan, A. H. (2020). "Identification and instance segmentation of oil spills using deep neural networks," in *5th World Congress on Civil, Structural, and Environmental Engine* (Avestia Publishing), 140-1-140–148. doi: 10.11159/iceptp20.140

Habibi, N., and Pourjavadi, A. (2022). Magnetic, thermally stable, and superhydrophobic polyurethane sponge: A highly efficient adsorbent for separation of marine oil spill pollution. *Chemosphere* 188, 114651. doi: 10.1016/j.chemosphere.2021.132254

Hasimoto-Beltran, R., Canul-Ku, M., Díaz Méndez, G. M., Ocampo-Torres, F. J., and Esquivel-Trava, B. (2023). "Ocean oil spill detection from SAR images based on multi-channel deep learning semantic segmentation," in *Marine Pollution Bulletin*, vol. 188. (Elsevier Ltd). doi: 10.1016/j.marpolbul.2023.114651

Iqbal, E., Safarov, S., and Bang, S. (2022). MSANet: multi-similarity and attention guidance for boosting few-shot segmentation. *arXiv:2206.09667. arXiv preprint*, June. doi: 10.48550/arXiv.2206.09667

Jiang, W., Huang, K., Geng, J., and Deng, X. (2021). Multi-scale metric learning for few-shot learning. *IEEE Trans. Circuits Syst. Video Technol.* 31, 1091–1102. doi: 10.1109/TCSVT.2020.2995754

Kang, X., Deng, B., Duan, P., Wei, X., and Li, S. (2023). Self-supervised spectral-spatial transformer network for hyperspectral oil spill mapping. *IEEE Trans. Geosci. Remote Sens.* 61. doi: 10.1109/TGRS.2023.3260987

Kang, D., Koniusz, P., Cho, M., and Murray, N. (2023). Distilling self-supervised vision transformers for weakly-supervised few-shot classification & segmentation. *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognition*, 19627–19638. doi: 10.1109/CVPR52729.2023.01880

Khadka, R., Jha, D., Hicks, S., Thambawita, V., Riegler, M. A., Ali, S., et al. (2022). Meta-learning with implicit gradients in a few-shot setting for medical image segmentation. *Comp. Biol. Med* 143, 105227. doi: 10.1016/j.compbiomed.2022.105227

Krestenitis, M., Orfanidis, G., Ioannidis, K., Avgerinakis, K., Vrochidis, S., and Kompatsiaris, I. (2019). Oil spill identification from satellite images using deep neural networks. *Remote Sens.* 11. doi: 10.3390/rs11151762. MDPI AG.

Lang, C., Cheng, G., Tu, B., and Han, J. (2022). "Learning what not to segment: A new perspective on few-shot segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 8057–8067.

Lang, C., Cheng, G., Tu, B., Li, C., and Han, J. (2023). Base and meta: A new perspective on few-shot segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 45, 10669–10686. doi: 10.1109/TPAMI.2023.3265865

Lau, T. K., and Huang, K. H. (2024). A timely and accurate approach to nearshore oil spill monitoring using deep learning and GIS. *Sci. Total Environ.* doi: 10.1016/j.scitotenv.2023.169500

Li, Y., Lyu, X., Frery, A. C., and Ren, P. (2021). Oil spill detection with multiscale conditional adversarial networks with small-data training. *Remote Sens.* 13, 2378. doi: 10.3390/rs13122378

Li, Y., Huang, W., Lyu, X., Liu, S., Zhao, Z., and Ren, P. (2022). An adversarial learning approach to forecasted wind field correction with an application to oil spill drift prediction. *Int. J. Appl. Earth Observation Geoinformation* 112, 102924. doi: 10.1016/j.jag.2022.102924

Ma, X., Xu, J., Pan, J., Yang, J., Wu, P., and Meng, X. (2023). Detection of marine oil spills from radar satellite images for the coastal ecological risk assessment. *Journal of Environmental Management* 325, 116637. doi: 10.1016/j.jenvman.2022.116637

Min, J., Kang, D., and Cho, M. (2021). "Hypercorrelation squeeze for few-shot segmentation," in *Proceedings of the IEEE/CVF international conference on computer vision.*, 6941–6952.

Mohammadiun, S., Hu, G., Gharahbagh, A. A., Li, J., Hewage, K., and Sadiq, R. (2021). Intelligent computational techniques in marine oil spill management: A critical review. *J. Hazardous Materials.* doi: 10.1016/j.jhazmat.2021.126425. Elsevier B.V.

Soh, K., Zhao, L., Peng, M., Lu, J., Sun, W., and Tongngern, S. (2024). SAR marine oil spill detection based on an encoder-decoder network. *Int. J. Remote Sens.* 45, 587–608. doi: 10.1080/01431161.2023.2299274

Song, D., Zhen, Z., Wang, B., Li, X., Gao, L., Wang, N., et al. (2020). A novel marine oil spillage identification scheme based on convolution neural network feature extraction from fully polarimetric SAR imagery. *IEEE Access* 8, 59801–59820. doi: 10.1109/ACCESS.2020.2979219

Sun, X., Fu, H., Bao, M., Zhang, F., Liu, W., Li, Y., et al. (2023). Preparation of slow-release microencapsulated fertilizer-biostimulation remediation of marine oil spill pollution. *J. Environ. Chem. Engineering.* 11 (2), 109283. doi: 10.1016/j.jece.2023.109283

Tian, Z., Zhao, H., Shu, M., Yang, Z., Li, R., and Jia, J. (2022). Prior guided feature enrichment network for few-shot segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 44, 1050–1065. doi: 10.1109/TPAMI.2020.3013717

Wang, C., Liu, X., Yang, T., Sridhar, D., Algadi, H., Xu, B. B., et al. (2023). An overview of metal-organic networks and their magnetic composites for the removal of pollutants. *Separation Purification Technol.* 320, 124144.

Wang, W., Zhou, T., Yu, F., Dai, J., Konukoglu, E., and Van Gool, L. (2021). "Exploring cross-image pixel contrast for semantic segmentation," in *Proceedings of the IEEE/CVF international conference on computer vision.*, 7303–7313.

Yuan, X., Shi, J., and Gu, L. (2021). A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Syst. Appl.* 169, 114417. doi: 10.1016/j.eswa.2020.114417

Zheng, W., Tian, X., Yang, B., Liu, S., Ding, Y., Tian, J., et al. (2022). A few shot classification methods based on multiscale relational networks. *Appl. Sci.* 12. doi: 10.3390/app12084059. MDPI.

Zhou, T., Wang, W., Konukoglu, E., and Van Gool, L. (2022). "Rethinking semantic segmentation: A prototype view," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 2582–2593.

Zhu, Q., Zhang, Y., Li, Z., Yan, X., Guan, Q., Zhong, Y., et al. (2022). "Oil spill contextual and boundary-supervised detection network based on marine SAR images," in *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60. doi: 10.1109/TGRS.2021.3115492