Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria
 Published by Research Publishing, Singapore.
doi: 10.3850/ESREL2026061419_esrel26-p24059-cd

Unmanned Aerial Vehicle (UAV)-Based Identification of Pipeline High Consequence Areas (HCAs) in Few-Shot Scenarios: A Phased Transfer Learning Method

Huanyu Zhao

College of Artificial Intelligence, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, PR China. E-mail: 1213677646@qq.com

Haochong Li

College of Artificial Intelligence, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, PR China. E-mail: lehoc0401@163.com

Shiyuan Pan

College of Artificial Intelligence, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, PR China. E-mail: panshiyuan1@qq.com

Huaizhong Shi

College of Artificial Intelligence, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, PR China. E-mail: shz@cup.edu.cn

Rui Qiu

National Engineering Laboratory for Pipeline Safety/ Beijing Key Laboratory of Urban Oil and Gas Distribution Technology, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, PR China. E-mail: qiurui1996@126.com

Guangtao Fu

National Engineering Laboratory for Pipeline Safety/ Beijing Key Laboratory of Urban Oil and Gas Distribution Technology, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, PR China. E-mail: fgtao8783@163.com

Pengtao Niu

National Engineering Laboratory for Pipeline Safety/ Beijing Key Laboratory of Urban Oil and Gas Distribution Technology, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, PR China.

Yongtu Liang

College of Artificial Intelligence, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, PR China. E-mail: yongtuliang@126.com

High Consequence Area (HCA) identification is fundamental to pipeline integrity management because building distribution and population conditions along pipeline corridors directly shape consequence assessment and risk-informed decisions. In practical unmanned aerial vehicle (UAV) inspections, however, labeled samples are often scarce and corridor environments are highly heterogeneous, making building extraction prone to false positives and missed buildings and thereby weakening the reliability of automated HCA identification. To address this challenge, this study develops a phased transfer learning framework for building-footprint extraction from UAV imagery under few-shot conditions. Built on TransUNet, the proposed method integrates CNN-based local feature extraction with Transformer-based global context modeling and adapts the model to corridor scenes through staged transfer learning with limited annotations. Under spatially isolated five-fold cross-validation, the proposed approach achieves a Dice score of 82.7%, an intersection over union (IoU) of 79.4%, and a false positive rate of 1.2%. Compared with the widely used engineering baseline DeepLabV3+, Dice and IoU increase by 9.4% and 10.1%, respectively, while the false positive rate decreases by 29.4%. Two engineering cases further show that the extracted building inventories support reliable building counts, population estimation, and HCA grading, thereby providing a reliable basis for periodic HCA updating and quantitative risk assessment in pipeline safety management.

Keywords: High Consequence Areas, Few-Shot Learning, Transfer Learning, TransUNet, UAV Remote Sensing, Pipeline Integrity Management.

1. Introduction

Refined oil and gas pipelines are a critical component of national energy infrastructure, and their safe operation underpins energy security. When failures occur, their consequences are often amplified in High Consequence Areas (HCAs), where human activity and occupied assets concentrate along the corridor. Accurate and timely HCA identification is therefore a foundational requirement of pipeline integrity management.

Traditionally, HCA identification has relied on manual field surveys, which are labor-intensive, time-consuming, and costly. To improve efficiency, Geographic Information Systems supported by satellite imagery have been introduced. Yet satellite-based approaches remain constrained by limited spatial resolution and by environmental interference such as cloud cover. In recent years, Unmanned Aerial Vehicle (UAV) remote sensing has emerged as an effective alternative, offering flexible acquisition and high-resolution imagery for fine-grained corridor analysis.

Despite these advantages, practical UAV-based recognition still faces two persistent bottlenecks. First, high-quality annotated datasets are difficult to assemble because data collection is operationally constrained and labeling is expensive, producing a typical few-shot setting. Second, models trained on limited data often generalize poorly across heterogeneous corridor environments, from farmland to mixed residential areas. Under such conditions, buildings are readily confused with visually similar non-building objects, which drives false positives and weakens the reliability of downstream grading.

To address these challenges, this study develops a phased transfer learning method for intelligent building-footprint extraction. The approach is built on TransUNet, which couples Transformer-based global context modeling with CNN-based local feature extraction. Through staged fine-tuning for cross-domain adaptation, the method alleviates data scarcity and improves robustness in complex corridor scenes.

Accordingly, this study makes three contributions: (i) it establishes a UAV-image-based building-extraction framework for periodic pipeline HCA updating and risk assessment, clarifying the links among building-footprint extraction, building counting, population estimation, and HCA grading within pipeline safety management; (ii) it proposes a phased transfer learning strategy built on TransUNet for corridor scenes with scarce labels, heterogeneous backgrounds, and frequent false positives and missed detections, improving few-shot robustness by freezing low-level convolutional features while adapting high-level semantic representations; and (iii) it conducts a systematic evaluation using spatially isolated five-fold cross-validation and, through two engineering cases, verifies that the segmentation outputs support building statistics, population estimation, and HCA grading, thereby providing a reliable basis for periodic HCA updating, verification prioritization, and quantitative risk assessment.

2. Literature Review

Pipeline integrity management seeks to ensure safe and reliable operation through risk assessment and targeted mitigation. Within this framework, HCA identification provides the consequence-related information required for corridor risk evaluation and therefore depends on accurate recognition of buildings and other sensitive receptors. Related studies can be grouped into three themes: operating-condition intelligence from pipeline data, image-based building-footprint extraction, and learning strategies for limited annotated data.

The first theme concerns the analysis of operating-condition data for monitoring and anomaly detection. Early studies emphasized numerical modeling for scheduling optimization and thermal analysis (Zhang et al., 2016). With the rise of data-driven techniques, machine-learning and deep-learning frameworks such as DeepPipe were introduced to learn hierarchical representations from operating data for anomaly detection (Wang et al., 2022). More recently, spatiotemporally enhanced physics-informed neural networks have been developed for transient-flow simulation in multi-product pipelines, further illustrating

the value of embedding domain knowledge into industrial pipeline intelligence (Li et al., 2026). However, these approaches are designed primarily for one-dimensional operating states and therefore cannot provide the two-dimensional spatial semantics required for UAV-based HCA recognition.

The second and third themes concern corridor perception from imagery and learning under limited supervision. Semantic segmentation has become the dominant paradigm for building-footprint extraction, and UAV remote sensing has further enabled fine-grained scene understanding (Touzani and Granderson, 2021; Lyu et al., 2020). To overcome the limitations of purely local feature modeling, hybrid architectures that integrate Transformers with CNNs have been explored, demonstrating the benefit of combining global context with local feature extraction in challenging remote-sensing scenes (Zhang et al., 2022). At the same time, annotation scarcity remains a major practical barrier. Semi-supervised and few-shot strategies have therefore been investigated in both pipeline signal processing and remote sensing (Zheng et al., 2021; Chen et al., 2022; Sun et al., 2023). Despite this progress, phased transfer learning strategies tailored to TransUNet-style architectures for UAV-based HCA recognition remain underexplored, especially in corridor scenes with heterogeneous backgrounds and strong building-like distractors. This gap motivates the present study.

3. Methodology

3.1. *Problem formulation*

In intelligent HCA recognition, grade determination relies on quantitative estimates of population density and household counts within the potential impact radius. Automated extraction of building footprints from UAV orthomosaics is therefore not a stand-alone visual task, but a prerequisite for downstream consequence assessment. Given UAV orthomosaics $\mathscr{X}$, the objective is to predict a building distribution map $\mathscr{Y}$ while preserving the geometric fidelity required for reliable household and occupancy estimation.

The proposed workflow begins with UAV-based building-footprint segmentation, then derives building counts and population estimates within the impact buffer, incorporates place attributes for area-grade and HCA-grade determination, and ultimately supports periodic HCA ledger updates, inspection prioritization, and quantitative risk assessment (Fig. 1).

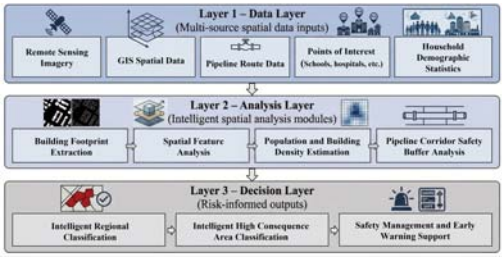

Fig. 1. Three-layer workflow linking building-footprint extraction to HCA classification and safety support.

Figure 1 situates the proposed method in a three-layer pipeline safety workflow, where multi-source spatial data feed building-footprint extraction and the resulting building counts, population estimates, and place attributes support regional classification, HCA classification, and management actions. For engineering deployment, the key challenge is robust transfer across corridor scenes with scarce labels and building-like distractors while keeping manual verification manageable.

To address this generalization bottleneck, we develop a unified building-footprint extraction framework with two tightly coupled components: a TransUNet backbone that integrates CNN-based local feature extraction with Transformer-based global context modeling, and a phased transfer learning strategy that adapts the model efficiently to corridor-specific UAV data under limited annotation.

3.2. *TransUNet backbone*

Building-footprint extraction in pipeline corridors requires both precise local delineation and broader scene awareness. We therefore adopt TransUNet (Fig. 2), which combines CNN-based local feature extraction with Transformer-based global context modeling.

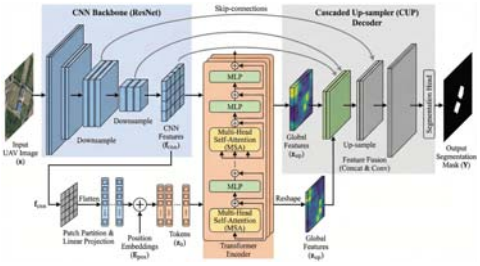

Fig. 2. TransUNet backbone for building-footprint segmentation.

Given an input image $\mathbf{x} \in \mathbb{R}^{H\times W\times 3}$, a ResNet-style CNN extracts multi-scale feature maps $\{\mathbf{f}^1, \mathbf{f}^2, \mathbf{f}^3, \mathbf{f}^4\}$. The deepest feature map is reshaped into a token sequence with $N = h \cdot w$ tokens:

$$\mathbf{X} = \text{Flatten}(\mathbf{f}^4) \in \mathbb{R}^{N\times c}. \tag{1}$$

Each token is then projected into a D-dimensional embedding space and augmented with positional information:

$$\mathbf{z}_0 = \mathbf{X}\mathbf{E} + \mathbf{E}_{pos}, \quad \mathbf{E} \in \mathbb{R}^{c\times D}, \mathbf{E}_{pos} \in \mathbb{R}^{N\times D}. \tag{2}$$

The Transformer encoder aggregates long-range context through self-attention. For a single head:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d_k}}\right)\mathbf{V}, \tag{3}$$

where $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ are linear projections of the input sequence and d_k is the key dimension. This global interaction improves discrimination between buildings and visually similar background objects.

High-resolution footprints are recovered through progressive upsampling with skip fusion. At decoding stage i:

$$\mathbf{y}^i = \text{Conv}\Big(\text{Concat}\big(\mathbf{z}^i_{up}, \mathbf{f}^i_{cnn}\big)\Big), \tag{4}$$

where $\mathbf{z}^i_{up}$ denotes the upsampled decoder features and $\mathbf{f}^i_{cnn}$ denotes the corresponding CNN feature map. This design preserves boundary detail while retaining corridor-level context.

3.3. *Two-stage transfer learning*

Labeled UAV samples for HCA recognition are typically scarce and geographically diverse. Training from scratch therefore tends to overfit, whereas naive full fine-tuning can overwrite transferable low-level cues. We therefore adopt a two-stage strategy (Fig. 3) in which generic visual primitives are learned during source pre-training and corridor-specific semantics are introduced during frozen fine-tuning.

Let the TransUNet be $f(\cdot;\theta)$ and the predicted building probability map be $\hat{y} = f(x;\theta)$. We partition parameters as:

$$\theta = \{\theta^{early}_{cnn}, \theta^{late}_{cnn}, \theta_{tr}, \theta_{dec}\}, \tag{5}$$

where θ^{early}_{cnn} and θ^{late}_{cnn} denote early and later CNN layers, θ_{tr} is the Transformer encoder, and θ_{dec} is the decoder head. Training minimizes a standard segmentation loss $\mathscr{L}_{seg}(\hat{y}, y)$.

3.3.1. *Stage 1: Source pre-training*

We first pre-train the network on a source dataset $\mathscr{D}_S = \big\{(x^s_i, y^s_i)\big\}_{i=1}^{N_s}$ to obtain a transferable initialization:

$$\theta_{pre} = \arg\min_{\theta} \frac{1}{N_s}\sum_{i=1}^{N_s} \mathscr{L}_{seg}\big(f(x^s_i;\theta), y^s_i\big). \tag{6}$$

This stage provides a transferable initialization for building-related texture and shape cues.

3.3.2. *Stage 2: Frozen fine-tuning on target HCA data*

We then adapt the model to the target dataset $\mathscr{D}_T = \big\{(x^t_i, y^t_i)\big\}_{i=1}^{N_t}$, where $N_t \ll N_s$. Under few-shot supervision, updating all parameters increases variance and can erase transferable low-level primitives. We therefore freeze the early CNN layers and update only higher-level parameters:

$$\begin{aligned} \theta_F &= \theta^{early}_{cnn}, \\ \theta_F &\leftarrow \theta_{pre,F}, \\ \theta_U &= \{\theta^{late}_{cnn}, \theta_{tr}, \theta_{dec}\}. \end{aligned} \tag{7}$$

The fine-tuning process is initialized with $\theta^{(0)}_U = \theta_{pre,U}$ and proceeds to minimize the regularized objective function $\mathscr{J}(\theta_U)$:

$$\mathscr{J}(\theta_U) = \frac{1}{N_t}\sum_{i=1}^{N_t} \mathscr{L}_{seg}\big(f(x^t_i;\theta_F,\theta_U), y^t_i\big) + \lambda\|\theta_U\|_2^2, \tag{8}$$

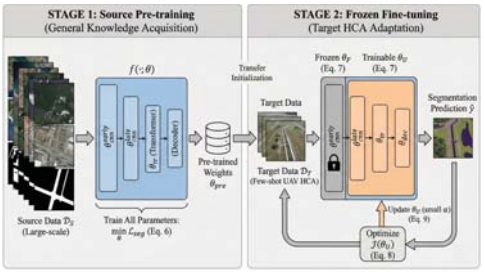

Fig. 3. Two-stage transfer learning strategy.

where λ is a weight-decay coefficient. Optimization updates only θ_U via gradient descent using a smaller learning rate α than in pre-training:

$$\theta_U^{(k+1)} = \theta_U^{(k)} - \alpha \nabla_{\theta_U} \mathscr{J}(\theta_U^{(k)}). \tag{9}$$

By constraining updates to high-level components, staged adaptation preserves transferable low-level features while adapting higher-level semantics to corridor scenes.

4. Experiments

4.1. *Dataset and experimental setup*

Experiments were conducted on a UAV dataset constructed from three high-resolution orthomosaics acquired along real pipeline corridors spanning different land-use settings. The orthomosaics were partitioned into fixed-scale image patches, and an automated quality-control procedure removed patches with insufficient optical information.

Because strong spatial autocorrelation is common in corridor imagery, random splitting can produce overly optimistic estimates. Following recommendations for spatial validation (Roberts et al., 2017), we adopted a cyclic band-wise split in which training, validation, and test subsets were strictly separated by spatial row groups. Five folds were generated by varying the spatial offset, and all reported results are fold averages. A minimal sample-balancing procedure ensured that each subset contained positive samples without duplication or leakage. The CNN encoder of TransUNet was initialized with ImageNet-pretrained weights, and all experiments were implemented in PyTorch on a workstation equipped with an NVIDIA RTX 3090 GPU.

4.2. *Evaluation metrics*

Segmentation performance was evaluated using standard pixel-level metrics, with predicted probability maps binarized at a threshold of 0.5. Because HCA identification is sensitive to missed buildings and underestimated population, Recall is emphasized as an indicator of detection completeness, whereas Precision reflects false-alarm control:

$$\text{Precision} = \frac{N_{TP}}{N_{TP} + N_{FP}} \tag{10}$$

$$\text{Recall} = \frac{N_{TP}}{N_{TP} + N_{FN}} \tag{11}$$

To quantify overall geometric agreement between predicted masks and ground-truth annotations, we adopt the Dice Similarity Coefficient and Intersection over Union:

$$\text{Dice} = \frac{2N_{TP}}{2N_{TP} + N_{FP} + N_{FN}}. \tag{12}$$

$$\text{IoU} = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}}. \tag{13}$$

Given the extensive non-building background area characteristic of pipeline corridors, we explicitly monitor the False Positive Rate (FPR):

$$\text{FPR} = \frac{N_{FP}}{N_{FP} + N_{TN}} \tag{14}$$

The training objective combines binary cross-entropy with a soft Dice loss to balance pixel-level classification accuracy and region-level overlap. In the following analysis, miss tendency and false-alarm burden are interpreted as implications for grading support rather than introduced as additional standalone metrics:

$$\mathscr{L} = \mathscr{L}_{BCE} + \left(1 - \frac{2\sum p \cdot t + \varepsilon}{\sum p + \sum t + \varepsilon}\right) \tag{15}$$

where p denotes predicted probabilities, t denotes ground truth labels, and $\varepsilon = 1.0$ is a smoothing constant.

4.3. *Results and analysis*

The results are discussed in three steps: comparative evaluation, transfer-learning ablation, and grading-oriented case validation.

Table 1. Quantitative comparison of segmentation models across five spatial folds.

Method	Dice	IoU	Precision	Recall	FPR
TransUNet (Ours)	**0.827**	**0.794**	**0.867**	**0.900**	0.012
CBAM-UNet	0.507	0.476	0.608	0.763	0.016
UNet++	0.567	0.556	0.846	0.667	**0.007**
UNet	0.536	0.497	0.612	0.828	0.028
DeepLabV3+	0.756	0.721	0.804	0.877	0.017
SegFormer-B0	0.610	0.594	0.832	0.660	0.010

4.3.1. *Comparative evaluation of segmentation architectures*

Table 1 compares TransUNet with representative segmentation architectures used for engineering building extraction, including UNet, UNet++, CBAM-UNet, DeepLabV3+, and SegFormer-B0 (Minaee et al., 2021). In this comparison, phased freezing and unfreezing were omitted so that performance differences primarily reflect architectural capability; phased transfer is examined separately in the ablation study.

Under the few-shot setting induced by the cyclic band-wise split, TransUNet achieves the best overall performance, reaching Dice 0.827 and IoU 0.794. Relative to DeepLabV3+, a well-established engineering baseline, it improves Dice and IoU by 0.071 and 0.073 while also increasing Precision and Recall.

UNet++ attains the lowest FPR, but only with substantially lower Dice and Recall, indicating overly conservative predictions. TransUNet instead preserves high Recall while reducing the FPR of DeepLabV3+ by 29.4%.

These differences mainly reflect contextual modeling capacity: pure CNN models are constrained by local receptive fields, DeepLabV3+ expands context but models long-range dependencies only indirectly, whereas TransUNet combines self-attention with multi-scale skip connections to delineate footprints more coherently in complex corridor backgrounds.

Table 2. Ablation study of transfer-learning strategies for TransUNet.

Strategy	Dice	IoU	Precision	Recall	FPR
A	0.701	0.662	0.742	0.781	0.021
B	0.827	0.794	0.867	0.900	0.012
C	**0.842**	**0.811**	**0.881**	**0.902**	**0.010**

Note: A: training from scratch without pre-training. B: direct fine-tuning of a pre-trained TransUNet with all parameters updated. C: proposed phased transfer learning strategy, where early CNN layers are frozen and only high-level components are fine-tuned on the target data.

4.3.2. *Ablation study on phased transfer learning*

Table 2 summarizes three training strategies for TransUNet under few-shot conditions. Training from scratch performs worst, confirming that the corridor labels are insufficient for stable optimization. Direct fine-tuning provides a strong baseline, and phased transfer further improves Dice and IoU while reducing FPR from 0.012 to 0.010, indicating that preserving early CNN layers while adapting higher-level semantics improves corridor-specific discrimination. The following subsection examines how these differences translate into grading support.

4.3.3. *Risk-oriented analysis and grading case validation*

In HCA identification, grading reliability is anchored in the quality of upstream building recognition, because building counts and their spatial distribution directly govern population estimation and area-grade determination. Building-footprint segmentation is therefore the entry point of the entire grading chain.

Segmentation errors also have direct engineering consequences: missed buildings can underestimate population and induce under-grading near decision thresholds, whereas false positives increase manual verification. A model for HCA identification must therefore maintain high detection completeness while controlling false alarms.

From Table 1, miss rate (defined as $1 - \text{Recall}$) decreases from 0.123 for DeepLabV3+ to 0.100 for TransUNet, while FPR falls from 0.017 to

0.012. For a 43-building segment such as Case 1, expected misses decrease from about 5.3 to 4.3, enough to affect grading stability near a threshold while reducing manual review.

To test whether these gains translate into correct grading, we examine two representative pipeline segments, denoted Case 1 and Case 2. Table 3 summarizes the key grading inputs. The table and Figs. 4–5 are taken from the intelligent identification system deployed by the pipeline management company using the proposed method as the core extraction module; the displayed masks are engineering outputs after post-processing rather than raw network predictions.

In Case 1, the model identifies 43 buildings, yielding an estimated population of 313 and a population density of 692 people/km^2. These inputs support Area Grade II; with flammable or explosive facilities in the corridor buffer, the final classification is HCA II.

Fig. 4. Case 1 intelligent-identification overlay from the deployed system.

In Case 2, the model identifies 45 buildings, corresponding to an estimated population of 62 and a population density of 126 people/km^2, again supporting Area Grade II. With a specified public place inside the buffer, the final grade becomes HCA I, consistent with the engineering record.

Fig. 5. Case 2 intelligent-identification overlay from the deployed system.

Table 3. System-derived grading evidence from the deployed identification system.

Attribute	Case 1	Case 2
Segment length	1.0 km	0.6 km
Detected buildings	43	45
Population / density (people/km^2)	313 / 692	62 / 126
Key place attribute	Flam./Expl. facility	Spec. public place
Area grade	II	II
HCA grade	II	I

These graded results further indicate that grade-aware outputs derived from the identification system can directly support differentiated UAV inspection strategies in practical corridor management. Once incorporated into the corridor ledger, segments assigned higher HCA grades or located near grading thresholds can be scheduled for denser revisit intervals, stricter analyst review, and earlier field verification, whereas lower-consequence segments may remain on routine patrol cycles. Such prioritization is valuable not only for improving the efficiency of inspection-resource allocation, but also for reducing the likelihood that consequence-sensitive environmental changes remain unnoticed between survey periods. Taken together, the two cases show that improvements in building-footprint segmentation are not confined to image-level accuracy metrics; they propagate into more reliable building inventories, more defensible population estimates, and more stable grading inputs. In this sense, improved segmentation quality strengthens the traceability and operational credibility of periodic HCA updating and provides a more dependable information basis for subsequent risk analysis and integrity-management decisions.

5. Conclusion

This study addresses UAV-based HCA identification under few-shot conditions through a

TransUNet-based phased transfer framework evaluated with spatially isolated validation. The method improves segmentation quality and reduces false alarms in corridor-specific UAV scenes.

The two engineering cases further show that, when the proposed method serves as the core extraction module of the pipeline management company's intelligent identification system and is coupled with engineering post-processing, it supports correct area-grade and HCA-grade determinations through population estimation and place-attribute rules. More importantly, it strengthens the evidential chain from imagery interpretation to grading decisions, making engineering records more consistent and auditable. This enables differentiated UAV inspection strategies: higher-grade or boundary-sensitive segments can be revisited more intensively, while lower-risk corridors remain on routine patrol cycles, and the resulting graded records can be used directly as structured inputs to quantitative risk assessment.

Remaining limitations include regional domain shift, annotation ambiguity, imagery timestamp mismatch, and threshold sensitivity; conservative review is still required for anomalous and near-threshold segments. Future work will extend multi-region validation and integrate auxiliary information for consequence assessment.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (52341203, 52502411).

References

Chen, Y., C. Wei, D. Wang, C. Ji, and B. Li (2022). Semi-supervised contrastive learning for few-shot segmentation of remote sensing images. *Remote Sensing 14*(17), 4254.

Li, H., J. Du, H. Zhao, K. Lu, J. Shen, Q. Liao, J. Zheng, and Y. Liang (2026). Spatiotemporally enhanced physics-informed neural network for transient-flow simulation in multi-product pipelines. *Chemical Engineering Research and Design*.

Lyu, Y., G. Vosselman, G.-S. Xia, A. Yilmaz, and M. Y. Yang (2020). UAVid: A semantic segmentation dataset for UAV imagery. *ISPRS Journal of Photogrammetry and Remote Sensing 165*, 108–119.

Minaee, S., Y. Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos (2021). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence 44*(7), 3523–3542.

Roberts, D. R., V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, D. I. Warton, B. A. Wintle, F. Hartig, and C. F. Dormann (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography 40*(8), 913–929.

Sun, Q., J. Chao, W. Lin, Z. Xu, W. Chen, and N. He (2023). Learn to few-shot segment remote sensing images from irrelevant data. *Remote Sensing 15*(20), 4937.

Touzani, S. and J. Granderson (2021). Open data and deep semantic segmentation for automated extraction of building footprints. *Remote Sensing 13*(13), 2578.

Wang, C., J. Zheng, Y. Liang, B. Wang, J. J. Klemeš, Z. Zhu, and Q. Liao (2022). Deeppipe: An intelligent monitoring framework for operating condition of multi-product pipelines. *Energy 261*, 125325.

Zhang, C., W. Jiang, Y. Zhang, W. Wang, Q. Zhao, and C. Wang (2022). Transformer and CNN hybrid deep neural network for semantic segmentation of very-high-resolution remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing 60*, 1–20.

Zhang, H.-R., Y.-T. Liang, Q. Xiao, M.-Y. Wu, and Q. Shao (2016). Supply-based optimal scheduling of oil product pipelines. *Petroleum Science 13*, 355–367.

Zheng, J., J. Du, Y. Liang, Q. Liao, Z. Li, H. Zhang, and Y. Wu (2021). Deeppipe: A semi-supervised learning for operating condition recognition of multi-product pipelines. *Process Safety and Environmental Protection 150*, 510–521.