# Zero-Shot Instance Segmentation of Shield Tunnel Linings via Adaptive SAM and Generative Inpainting

**Chengxin Shi[1,2], Jiepeng Liu[1,2], Dongsheng Li[1,2], Pengkun Liu[1,2,3*], Yongjing Wang[4], Shan Huang[4]**

[1]School of Civil Engineering, Chongqing University, Chongqing 400045, China
[2]State Key Laboratory of Safety and Resilience of Civil Engineering in Mountain Area, Chongqing 400045, China
[3]Hong Kong Center for Construction Robotics, The Hong Kong University of Science and Technology, Hong Kong, China
[4]China MCC5 Group Corporation Ltd., Chengdu 610000, China

scx@stu.cqu.edu.cn, liujp@cqu.edu.cn, lds@cqu.edu.cn, pengkunliu@cqu.edu.cn, 892238201@qq.com, 475347170@qq.com

**Abstract -**
**Precise segmentation of shield tunnel linings is a prerequisite for structural health monitoring and digital twin construction. However, existing methods are heavily relying on extensive annotated datasets, which are scarce in tunneling scenarios. To address this, we propose a novel zero-shot framework for the adaptive recognition and instance segmentation of tunnel linings. The proposed method starts by unfolding tunnel point cloud data to an intensity image, enabling the use of a Vision-Language Model (VLM) and a prompt-based Segment Anything Model (SAM). To enhance performance, we employ (i) an adaptive hybrid prompt generation mechanism, integrating VLM with pixel intensity priors to automatically generate bounding box and point prompts for SAM and (ii) a Stable Diffusion inpainting module to restore missing regions caused by occlusions, enabling SAM's zero-shot capability to achieve precise instance-level segmentation of tunnel linings. Experimental results in two different tunnel datasets, compared to the original baseline SAM without prompts, demonstrate the robustness of the framework, achieving a mean Intersection over Union (mIoU) of 90.58% in standard circular tunnels. This approach effectively overcomes data scarcity issues and provides an efficient, largely automated solution for intelligent tunnel maintenance.**

**Keywords -**
**Shield Tunnel; Point Cloud; Zero-Shot Segmentation; Segment Anything Model; Vision-Language Models**

## 1 Introduction

With the acceleration and deepening of urbanization, the effective utilization of urban underground space has become a key way to address issues such as traffic congestion, population density, and resource constraints [1] . However, such facilities typically exhibit characteristics that include diverse facility types, complex environments, and extensive coverage areas. Consequently, the challenges to ensuring tunnel operational safety and quality inspection remain worthy of attention.

Currently, conventional operational maintenance is largely based on manual inspections and fragmented digital recording. This approach mainly involves maintenance personnel performing periodic assessments on-site of facility conditions, with reports compiled manually [2, 3, 4]. Moreover, the tunnel environment is characterized by poor implementation, high humidity, and hazardous conditions, making manual inspection labor-intensive and risky [5]. To address these challenges, Yi et al. [6]extract segments from point clouds to construct the corresponding three-dimensional segment models, thereby obtaining precise structural models for subsequent digital twin-based tunnel operations and maintenance. Feng et al. [7] achieve a precise segmentation of the segments to allow detailed monitoring, such as calculating the position and orientation of the lining ring and analyzing joint deformation. These techniques offer valuable information for the comprehensive management of tunnels throughout their lifecycle.

However, both constructing digital twin geometric models and monitoring deformation fundamentally require precise segmentation of structural components. Existing tunnel lining segmentation methods primarily involve direct point cloud segmentation and projection-based segmentation. However, both approaches require extensive datasets for training, imposing significant limitations in data-scarce scenarios like tunnels.

The Segment Anything Model (SAM) proposed by Meta AI [8], demonstrates unprecedented zero-shot segmentation capabilities. It can precisely segment any object within an image based on simple user prompts (e.g., points, boxes, rough boundaries). Although this model exhibits strong generalizability, its performance may vary across specific scenarios. Preliminary research has explored the application of SAM's to remote sensing image analysis, medical image segmentation, and industrial defect detection[9], achieving favorable results due to SAM's robust pre-training foundation in image segmentation.

This paper proposes a novel point cloud segmentation method for tunnel linings, leveraging SAM's zero-shot segmentation capability to achieve high-precision automated segmentation. We employ (i) Vision-Language Model (VLM) with pixel intensity prior information to automatically generate bounding boxes and point prompts, aiming to enhance SAM's segmentation capabilities. (ii) We introduce the Stable Diffusion inpainting module to complete missing image portions, thereby improving the quality of images input to SAM. The final results demonstrate that our method performs effectively.

## 2 Related Work

Numerous research achievements have emerged in the field of point cloud segmentation, laying a solid foundation for technological advancement. Within the domain of

tunnel point cloud segmentation, approaches primarily fall into two categories.

The first category comprises direct segmentation methods based on point clouds, exemplified by Zhou et al.[10], whose research validated the application of deep learning in point cloud segmentation. The second category involves projection-based segmentation approaches, which reduce the dimensionality of three-dimensional data to two-dimensional images for processing. This reduction in dimensionality simplifies the segmentation challenge. For example, Zhang et al. [11] utilise traditional image processing techniques for segmentation, suitable for straightforward point cloud data.

However, current deep learning-based segmentation in tunneling still faces specific limitations, primarily in relation to scarce training data. Consequently, SAM's flexible prompting mechanism has been extended to specific engineering scenarios. Jing et al. [12] proposed a segmentation framework that integrates depth map information with visual SAM features, significantly improving the accuracy of the location of the edge for engineering components. Ye et al. [13] generated template prompts centered on initial points based on lining design parameters, supplemented by coarse mask prompts using polygonal envelopes, to achieve instance segmentation in shield-tunnel construction. Xu et al. [14], focusing on the implementation of engineering automation, further optimized the efficiency of inference and the compatibility of the deployment of large visual models. Their lightweight segmentation architecture achieved a real-time balance between component segmentation and condition monitoring in construction scenarios, providing technical support for intelligent supervision of construction processes.

Existing research indicates that segmentation methods based on large visual models are progressively transitioning from general-purpose segmentation to engineering-specific segmentation. However, existing SAM-based methods mainly rely on standard prompts and do not explicitly address the challenges of missing data and structural discontinuities in tunnel inspection images. To address these limitations, This study introduces an adaptive prompt generation module and an image restoration model, offering an enhanced solution for tunnel lining segmentation and advancing progress in this field.

# 3 Methodology

The proposed method is shown in Figure 1, consisting of four core steps: data preprocessing, PCD image to intensity and completion, SAM integration and optimization, and reprojection. Specifically, we preprocess the raw point cloud and perform a center axis fitting. Subsequently, the dimensionality reduction is achieved through cylindrical unfolding to generate intensity images. Finally, adaptive bounding boxes and point prompts are generated using VLMs and gray-value-driven techniques, which guide the SAM to perform zero-shot segmentation.

## 3.1 Data Preprocessing

Raw Point Cloud Data (PCD) from shield tunnels often contains non-structural artifacts such as equipment mounts and pipelines. To ensure data quality, we employ a statistical outlier removal filter combined with manual refinement using CloudCompare to eliminate noise while preserving the tunnel geometry.

Precise axis estimation is critical for subsequent unfolding. We propose an iterative least-squares fitting algorithm. First, Principal Component Analysis (PCA) provides an initial axis $L_0$. In each iteration $i$, the tunnel is sliced into normal planes $S_i$ at equal intervals along $L_i$. For each plane, the cross-section center $(a_i, b_i)$ and radius $R_i$ are estimated by minimizing the geometric distance:

$$E(a_i, b_i, R_i) = \sum_{j=1}^{N_i} \left( \sqrt{(x_j - a_i)^2 + (y_j - b_i)^2} - R_i \right)^2 \tag{1}$$

The optimal centers are then fitted to form an updated spatial curve $L_{i+1}$. This process repeats until the deviation $d(L_{i+1}, L_i)$ falls below a convergence threshold $\varepsilon$, yielding the final spatial axis $L$ and radius parameters.

## 3.2 PCD to Intensity Image and Completion

To take advantage of 2D foundation models, the 3D tunnel surface is projected onto a 2D plane. Based on the fitted axis $L$, we construct a local orthogonal basis $\{\mathbf{U}, \mathbf{E}, \mathbf{N}\}$ for each point $p_i$. The 3D coordinates are transformed into longitudinal distance $u_i$ and azimuth angle $\theta_i$:

$$u_i = (p_i - \mathbf{c}) \cdot \mathbf{U}, \quad \theta_i = \text{atan2}(\delta_i \cdot \mathbf{N}, \delta_i \cdot \mathbf{E}) \tag{2}$$

where $\mathbf{c}$ is the reference center, and $\delta_i$ is the radial vector.

The unfolded point set $P' = \{(u_i, \theta_i, I_i)\}$ is then rasterized in a 2D grid. The image width $W$ and height $H$ are determined by the angular resolution $\Delta\theta$ and the axial resolution $\Delta u$. To handle non-uniform point density, we employ a mean aggregation strategy where the pixel value $I_{\text{final}}[y, x]$ is calculated as the average intensity of all points falling within that grid cell:

$$I_{\text{final}}[y, x] = \frac{\sum I_i}{C[y, x]} \quad \text{if } C[y, x] > 0 \tag{3}$$

However, the generated intensity image inevitably contains void regions due to scanner blind spots, equipment occlusions, or extreme sparsity in the point cloud. These discontinuities disrupt the geometric integrity of the segments and can hinder the recognition performance of SAM.

To address this, we employ a generative inpainting approach based on the Stable Diffusion model. Specifically, we first generate a binary mask that identifies the missing regions. The binary mask for missing regions is automatically generated by identifying pixels with zero intensity (or below a sparsity threshold) resulting from the unfolding process. The incomplete intensity image and the mask are then fed into the Stable Diffusion Inpainting module. Using its powerful latent generative priors, the model synthesizes realistic textures to fill the masked areas based on contextual information from valid surrounding pixels. This process effectively restores the tunnel surface, providing a complete and high-quality input for the downstream segmentation task.

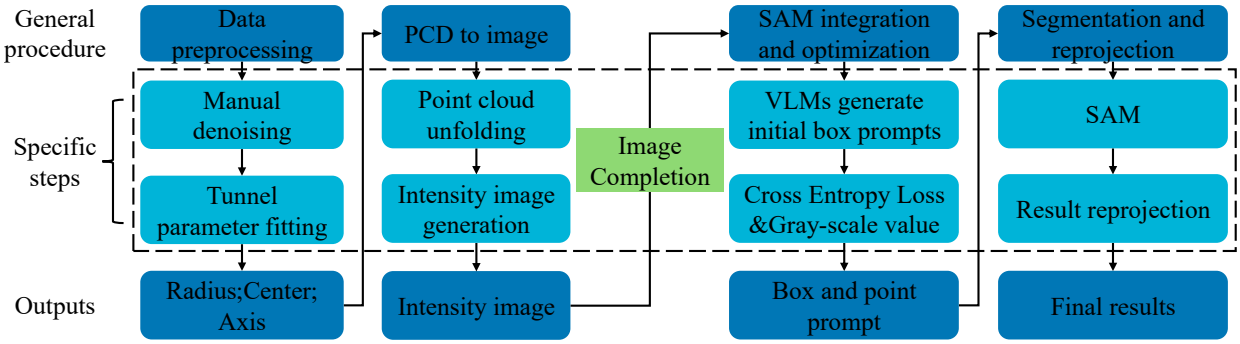


Figure 1. The framework of the proposed adaptive segmentation.

### 3.3 SAM integration and optimization

To enable precise zero-shot segmentation of shield tunnel segments, we propose an adaptive hybrid prompt generation module. Unlike standard SAM approaches that rely solely on generic visual foundation models, our method integrates open-vocabulary detection with domain-specific intensity priors to construct a robust prompt set:

$$P = \{V_{\text{box}}, V_{\text{point}}\} \tag{4}$$

for the SAM encoder. The specific model framework is shown in Figure 2.

#### 3.3.1 Bounding Box Generation and Optimization

For the bounding box prompts $V_{\text{box}}$, we utilize the DINO-X model, which accepts the unfolded intensity image $I$ and a text prompt $T_k$ = "rectangles" as input. The model encodes image features $E_I(I)$ and text features $E_T(T_k)$, fusing them through a cross-modality decoder to predict initial candidate boxes. However, raw outputs often suffer from fragmentation. To address this, we introduce an optimization strategy based on cross-entropy scoring.

Specifically, we extract the classification logits from the detector head, which correspond to the cross-entropy loss used during pre-training. These logits serve as confidence scores $S_{\text{conf}}$ for each candidate box. We first filter candidates based on a confidence threshold, and then apply an iterative merging algorithm. Boxes with high spatial overlap (IoU) are merged, prioritized by their cross-entropy scores, to form a purified set of bounding boxes $V_{\text{box}}$. These optimized boxes provide the coarse localization for the target segments.

#### 3.3.2 Intensity-Driven Point Prompt Sampling

Although bounding boxes provide spatial limits, relying solely on them for segmentation often results in the background being segmented, affecting the segmentation accuracy. We formulate a heuristic sampling strategy to generate point prompts $V_{\text{point}}$. For each optimized bounding box, we analyze the pixel intensity distribution. Pixels falling within the background value range are sampled as negative prompts to suppress background noise. In contrast, pixels within the median intensity range of the segment are sampled as positive prompts. We randomly sample $N$ positive points and $M$ negative points per instance. Finally, the optimized boxes $V_{\text{box}}$ and the intensity-driven points $V_{\text{point}}$ are concatenated and fed into SAM to generate the final high-quality segmentation masks.

### 3.4 Reprojection

Upon leveraging the zero-shot capabilities of SAM to obtain high-quality 2D instance segmentation masks, the final step is to transfer these labels back to the original 3D point cloud space. This process closes the loop of our framework, allowing for the direct segmentation of the raw tunnel data.

Since our unfolding process (described in Section 3.2) maintains a deterministic mapping between the 3D points and the 2D grid pixels, the reprojection is formulated as a direct index lookup operation.

For every point $p_i$ in the original point cloud $P$, we retrieve its corresponding pixel coordinates $(x_i, y_i)$ calculated during the rasterization stage. The semantic label $L_i$ for the 3D point $p_i$ is then assigned based on the value of the corresponding pixel in the 2D mask:

$$L_i = \mathcal{M}_{SAM}(y_i, x_i) \tag{5}$$

By iterating through all points in $P$, we obtain a fully segmented 3D point cloud where each point carries an instance label consistent with the result of the 2D segmentation. This approach ensures that the precise boundaries delineated by SAM in the image domain are accurately preserved in the 3D geometric space.

## 4 Experiments and Results

### 4.1 Dataset and Setup

To evaluate the precision and robustness of the proposed method, we selected experimental data from two distinct tunnels. As shown in Figure3, Tunnel A is a circular shield tunnel in Chongqing, while Tunnel B originates from the publicly available Seg2Tunnel T1 dataset[15], representing a horseshoe-shaped tunnel. For specific parameters, see Table1.

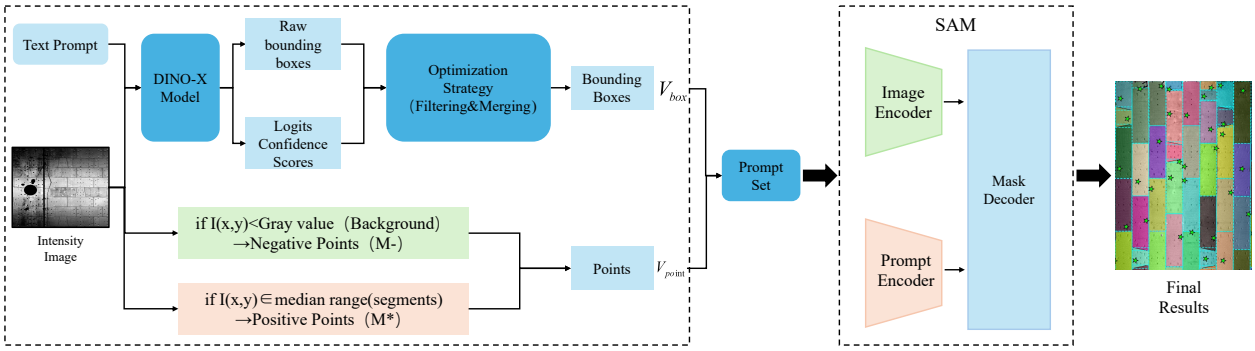


Figure 2. Optimized SAM Framework

Table 1. Key parameters of the shield tunnel datasets

| Parameter | Tunnel A | Tunnel B |
|---|---|---|
| Design inner diameter (m) | 7.5 | 5.5 |
| Segments per ring | 7 | 6 |
| Segment width (m) | 1.6 | 1.2 |
| Total point count | 16 million | 14 million |

All experiments were conducted with an NVIDIA GeForce RTX 3090 GPU (24GB VRAM). The model is implemented with the PyTorch 2.0 framework running on CUDA 11.8.

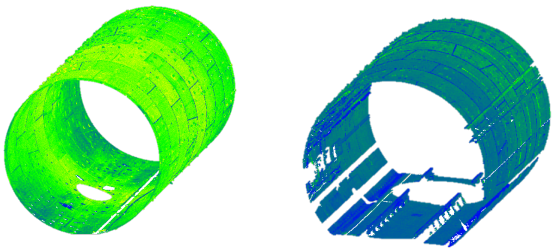


Figure 3. Overview of the experimental datasets.

### 4.2 Evaluation method

We offer Precision, Recall, Intersection over Union (IoU), and Mean Intersection over Union (mIoU) to quantify segmentation performance. The specific calculation formulas are as follows:

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i} \tag{6}$$

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i} \tag{7}$$

$$\text{IoU} = \frac{|M_{\text{GT}} \cap M_{\text{Pred}}|}{|M_{\text{GT}} \cup M_{\text{Pred}}|} \tag{8}$$

where:

- $TP_i$ denotes the true positives (TP) for class $i$,
- $FN_i$ denotes the false negatives (FN) for class $i$,
- $FP_i$ denotes the false positives (FP) for class $i$,
- $M_{\text{GT}}$ denotes the ground truth mask of the segment instance,
- $M_{\text{Pred}}$ denotes the predicted mask of the segment instance.

To evaluate the overall quality of the segmentation, we calculate the average value of IoU in all annotated segment instances, i.e., mIoU:

$$\text{mIoU} = \frac{1}{N} \sum_{i=1}^{N} \text{IoU}_i \tag{9}$$

where:

- $N$ is the total number of annotated segment instances,
- $\text{IoU}_i$ is the IoU value of the $i$-th segment instance.

For instance-level evaluation, predicted instances are matched to ground-truth instances using a one-to-one IoU-based strategy. A prediction is considered a true positive if its IoU with the best-matching ground-truth instance exceeds 0.5. Otherwise, it is counted as a false positive. Precision and Recall measure instance detection performance, whereas IoU evaluates the mask quality of matched instances.

### 4.3 Results and Analysis

For qualitative results, as illustrated in Figure4, we select Tunnel A as an example, presenting the unfolded intensity image, the segmentation predictions of the instance at the pixel-level, and the predicted point cloud reprojected. Among these, we evaluate the original SAM as a baseline to quantify . In this baseline, we employ the official model to perform fully automatic instance segmentation on the unfolded intensity images, without using any external prompts or inpainting modules. The ground truth for Tunnel A was manually annotated using the open-source software LabelMe, with labels numbered sequentially according to the number of segments, where 0 denotes background.

Table 2 summarizes the quantitative performance of our method in two datasets.

For circular Tunnel A, the method achieves a high mIoU of 90.58, with a precision of 87.95 and a recall of 97.33. The high values for precision and recall indicate that the model effectively identified most tunnel segment instances

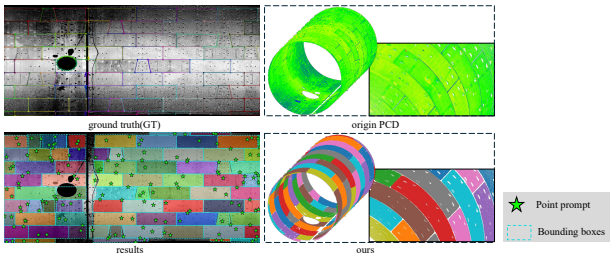


Figure 4. The segmentation results from tunnel A

(TP=73), with minimal false positives (FP=10) or false negatives (FN=2).

In contrast, performance for the horseshoe-shaped Tunnel B declined, achieving an mIoU of 75.59. While recall remained robust at 89.58, indicating the model successfully located most segments, precision dropped significantly to 75.44. This directly reflects a higher number of false positives (FP=14).

As shown in Figure5, the primary cause of this precision degradation is the extensive occlusion introduced by the equipment and cables attached to the tunnel lining. Unlike Tunnel A, these obstructions in Tunnel B result in large-scale data voids within the original point cloud.

Although our repair module mitigated this issue to some extent, the sheer size of the missing regions and the complex visual texture of the utility lines themselves posed significant challenges. Our method may misclassify missing background regions in intensity images as potential segmentation instances, thereby increasing the false positive (FP) rate. Nevertheless, recall remains robust at 89.58, demonstrating that even under severe occlusion, our method can still correctly segment a large number of tunnel segments, exhibiting significant potential for generalization.

Compared with the prompt-free SAM baseline, the proposed framework shows clear quantitative gains on both tunnels. The baseline only achieves 34.17% precision, 90.67% recall, and 80.31% mIoU on Tunnel A, and 31.48% precision, 70.83% recall, and 62.22% mIoU on Tunnel B, indicating that directly applying SAM tends to produce many spurious instances, whereas our method substantially improves precision and overall segmentation quality without sacrificing recall.

To analyze the impact of the inpainting module, Fig.6 presents a qualitative comparison of the unfolded intensity images before and after diffusion-based completion. While the inpainting process improves the visual continuity by filling missing regions, it may occasionally introduce synthetic textures that resemble structural boundaries. These pseudo-structures can mislead SAM and lead to false-positive detections, especially when large void regions must be synthesized with limited contextual information.

The proposed framework integrates three key components to achieve robust tunnel segment segmentation. The diffusion-based inpainting module improves the continuity of the unfolded intensity images by restoring missing regions caused by occlusions or sparse scanning. The box prompts generated by the vision-language model provide coarse instance localization, guiding SAM to focus on candidate segment regions. In addition, intensity-driven point

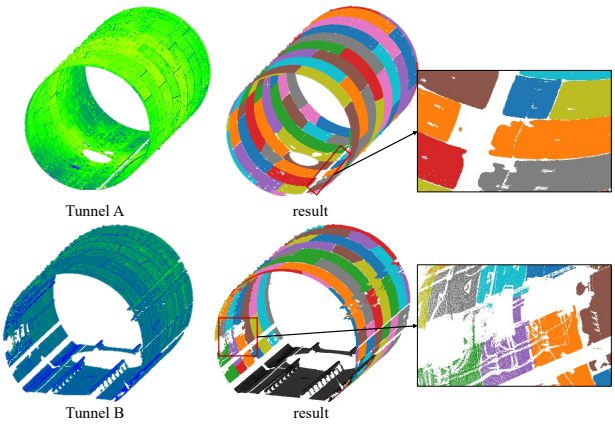


Figure 5. View of Tunnel Segmentation Results

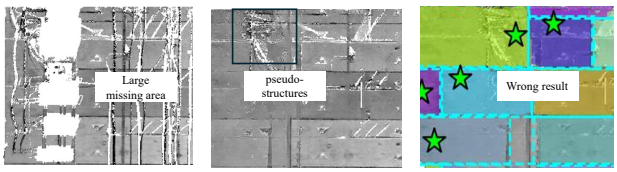


Figure 6. Error analysis of Tunnel B using the diffusion-based inpainting module.

prompts refine the segmentation boundaries by leveraging intensity variations along segment joints. These components jointly contribute to the overall segmentation performance of the framework.

Table3 presents the processing time of Tunnel A in the pipeline. It contains nine rings and containing 10 million points.

In summary, this paper proposes an automated zero-shot framework for instance-level segmentation of shield tunnel linings in 3D point clouds. Quantitative results in Table 2 show that the framework achieves high-quality instance segmentation for multi-type segments, accurately localizing most instances even under unfavorable conditions and validating its strong generalization potential.

## 5 Discussion

This paper constructs a complete closed-loop system. Experimental results demonstrate that this method preserves the continuity and topological consistency of segmented annular structures within circular tunnels, thereby reducing the reliance on complex three-dimensional models and extensive training datasets.

The comparison with the SAM baseline further confirms that directly applying a generic segment-everything model is insufficient in the presence of geometric distortions and severe occlusions: SAM tends to generate fragmented and spurious masks, particularly around repaired or missing regions. This approach relies entirely on pre-trained models and automated prompt-generating workflows, enabling zero-shot instance segmentation within minutes to

Table 2. Results of the Tunnel A and Tunnel B datasets

| Tunnel | Method | Prec(%) | Rec(%) | mIoU(%) | TP | FP | FN |
|---|---|---|---|---|---|---|---|
| A | SAM (baseline) | 34.17 | 90.67 | 80.31 | 68 | 131 | 7 |
| A | **Ours** | 87.95 | 97.33 | 90.58 | 73 | 10 | 2 |
| B | SAM (baseline) | 31.48 | 70.83 | 62.22 | 34 | 74 | 14 |
| B | **Ours** | 75.44 | 89.58 | 75.59 | 43 | 14 | 5 |

Table 3. Runtime analysis of the proposed framework

| Module | Time (s) |
|---|---|
| Data preprocessing | 120 |
| Unfolding and Inpainting | 102 |
| SAM integration and optimization | 40 |
| Reprojection | 37 |
| Total | 299 |

meet the efficiency demands of engineering inspection and maintenance.

However, when geometric structures deviate from circularity, unfolded images may exhibit distortion, which can affect boundary recognition in SAM model. Although SAM effectively addresses boundary adaptation issues, such distortions and occlusions still diminish segmentation accuracy, particularly in tunnels with complex geometries.

However, due to the generative nature of diffusion models, the restored textures may not perfectly correspond to the true tunnel geometry. Although the inpainting process improves the visual continuity of the unfolded intensity images, it may also introduce synthetic patterns that resemble structural edges or segment boundaries. These pseudo-structures can interfere with SAM's boundary recognition and occasionally lead to false-positive detections. This phenomenon becomes particularly pronounced in tunnels with heavy equipment occlusions, where large missing regions must be synthesized with limited contextual cues. This behavior is consistent with the experimental observations in Tunnel B, where complex occlusions increase the likelihood of misclassification. Despite these limitations, the inpainting module significantly improves the overall structural continuity of the input images, contributing to more stable and reliable segmentation results in most scenarios.

This approach relies entirely on pre-trained models and automated workflows, enabling zero-shot instance segmentation within minutes to meet the efficiency demands of engineering inspection and maintenance.

## 6 Conclusion and Future Work

This paper addresses the typical engineering scenario of segment segmentation in shield-tunnel construction by proposing a zero-shot adaptive prompt-generation segmentation framework. By establishing a closed-loop workflow 3D - 2D - 3D, it converts unstructured tunnel point clouds into two-dimensional intensity unfolded images, effectively bridging the gap between 3D engineering data and 2D foundation models. Using a general-purpose visual model and gray values for adaptive prompt generation, it achieves high-precision segmentation (90.58% mIoU) without requiring task-specific custom training. Experimental results provide preliminary evidence of the applicability of the proposed method across different tunnel geometries and missing areas, particularly in maintaining the continuity and topological consistency of ring-shaped structures within standard circular tunnels.

These findings validate the feasibility of the proposed approach, offering a practical technical path to reduce annotation costs and improve the efficiency of model deployment. Concurrently, it establishes a foundational data set for subsequent tunnel deformation monitoring and the construction of digital twin models.

However, this method also has certain limitations. Future research should focus on two primary areas:

- Fine-tuning image restoration models: Current general models exhibit insufficient learning of texture features in intensity images, resulting in poor performance when faced with large-scale voids. Subsequent work will address this issue to achieve enhanced robustness.
- Hierarchical prompting: Explore a hierarchical approach where a primary model identifies larger structures (e.g., tunnel rings), while a secondary, fine-tuned VLM targets detailed components (e.g., structural elements), potentially using 3D bounding boxes derived directly from fitted axes as more robust prompts.

## 7 Acknowledgments

This study was financially supported by the National Key Research and Development Program of China (Grant No. 2023YFC3805901), Postdoctoral Fellowship Program of CPSF (Grant No. GZC20242128), and China Postdoctoral Science Foundation (Grant No. 2024M763870) to which the authors are very grateful. Pengkun Liu extends sincere appreciation to The Hong Kong Center for Construction Robotics (InnoHK center supported by the Hong Kong ITC, InnoHK-HKCRC) for their invaluable support.

## References

[1] Jian Zhou, Hongning Qi, Kang Peng, Yulin Zhang, and Manoj Khandelwal. Comprehensive review and future perspectives on prediction and mitigation of tunnel-induced ground settlement: A bibliometric analysis and methodological overview (2002–2022). *Tunnelling and Underground Space Technology*, 154:106081, 2024. ISSN 0886-7798. doi:https://doi.org/10.1016/j.tust.2024.106081. URL `https://www.sciencedirect.com/science/article/pii/S0886779824004991`.

[2] J.A. Richards. Inspection, maintenance and repair of tunnels: International lessons and practice. *Tunnelling and Underground Space*

*Technology*, 13(4):369–375, October 1998. ISSN 0886-7798. doi:10.1016/s0886-7798(98)00079-0. URL http://dx.doi.org/10.1016/s0886-7798(98)00079-0.

[3] X. Jiang, Y. Yuan, and X. Liu. *Performance-based Predictive Maintenance of Shield Tunnels*, page 1007–1016. CRC Press, October 2014. ISBN 9780429227196. doi:10.1201/b17618-147. URL http://dx.doi.org/10.1201/b17618-147.

[4] A. Haack, J. Schreyer, and G. Jackel. Report to ita working group on maintenance and repair of underground structures: State-of-the-art of non-destructive testing methods for determining the state of a tunnel lining. *Tunnelling and Underground Space Technology*, 10(4):413–431, October 1995. ISSN 0886-7798. doi:10.1016/0886-7798(95)00030-3. URL http://dx.doi.org/10.1016/0886-7798(95)00030-3.

[5] Leanne Attard, Carl James Debono, Gianluca Valentino, and Mario Di Castro. Tunnel inspection using photogrammetric techniques and image processing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144:180–188, October 2018. ISSN 0924-2716. doi:10.1016/j.isprsjprs.2018.07.010. URL http://dx.doi.org/10.1016/j.isprsjprs.2018.07.010.

[6] Cheng Yi, Dening Lu, Qian Xie, Shuya Liu, Hu Li, Mingqiang Wei, and Jun Wang. Hierarchical tunnel modeling from 3d raw lidar point cloud. *Computer-Aided Design*, 114: 143–154, September 2019. ISSN 0010-4485. doi:10.1016/j.cad.2019.05.033. URL http://dx.doi.org/10.1016/j.cad.2019.05.033.

[7] Yong Feng, Xiao-Lei Zhang, Shi-Jin Feng, Yong Zhao, Qing-Zhao Kong, and Shu-Lin Zhu. Metro tunnel deformation detection based on laser scanning robot and point cloud semantic segmentation. *Tunnelling and Underground Space Technology*, 166:106966, December 2025. ISSN 0886-7798. doi:10.1016/j.tust.2025.106966. URL http://dx.doi.org/10.1016/j.tust.2025.106966.

[8] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, page 3992–4003. IEEE, October 2023. doi:10.1109/iccv51070.2023.00371. URL http://dx.doi.org/10.1109/iccv51070.2023.00371.

[9] Jay N. Paranjape, Shameema Sikder, S. Swaroop Vedula, and Vishal M. Patel. *S-SAM: SVD-Based Fine-Tuning of Segment Anything Model for Medical Image Segmentation*, page 720–730. Springer Nature Switzerland, 2024. ISBN 9783031723902. doi:10.1007/978-3-031-72390-2_67. URL http://dx.doi.org/10.1007/978-3-031-72390-2_67.

[10] Yunxiang Zhou, Ankang Ji, Limao Zhang, and Xiaolong Xue. Attention-enhanced sampling point cloud network (aspcnet) for efficient 3d tunnel semantic segmentation. *Automation in Construction*, 146:104667, February 2023. ISSN 0926-5805. doi:10.1016/j.autcon.2022.104667. URL http://dx.doi.org/10.1016/j.autcon.2022.104667.

[11] Yuxian Zhang, Xuhua Ren, Jixun Zhang, and Zichang Ma. A method for deformation detection and reconstruction of shield tunnel based on point cloud. *Journal of Construction Engineering and Management*, 150(3), March 2024. ISSN 1943-7862. doi:10.1061/jcemd4.coeng-14225. URL http://dx.doi.org/10.1061/jcemd4.coeng-14225.

[12] Yixiong Jing, Cheng Zhang, Haibing Wu, Guangming Wang, Olaf Wysocki, and Brian Sheil. Infradiffusion: zero-shot depth map restoration with diffusion models and prompted segmentation from sparse infrastructure point clouds, 2025. URL https://arxiv.org/abs/2509.03324.

[13] Zehao Ye, Wei Lin, Asaad Faramarzi, Xiongyao Xie, and Jelena Ninić. Sam4tun: No-training model for tunnel lining point cloud component segmentation. *Tunnelling and Underground Space Technology*, 158:106401, 2025. ISSN 0886-7798. doi:https://doi.org/10.1016/j.tust.2025.106401. URL https://www.sciencedirect.com/science/article/pii/S0886779825000392.

[14] Gang Xu, Yingshui Zhang, Qingrui Yue, and Xiaogang Liu. Automated detection and quantification of structural component dimensions using segment anything model (sam)-based segmentation. *Automation in Construction*, 176:106304, August 2025. ISSN 0926-5805. doi:10.1016/j.autcon.2025.106304. URL http://dx.doi.org/10.1016/j.autcon.2025.106304.

[15] Wei Lin, Brian Sheil, Pin Zhang, Biao Zhou, Cheng Wang, and Xiongyao Xie. Seg2tunnel: A hierarchical point cloud dataset and benchmarks for segmentation of segmental tunnel linings. *Tunnelling and Underground Space Technology*, 147:105735, May 2024. ISSN 0886-7798. doi:10.1016/j.tust.2024.105735. URL http://dx.doi.org/10.1016/j.tust.2024.105735.