Available online at www.sciencedirect.com

ScienceDirect

Underground Space 25 (2025) 281–294

www.keaipublishing.com/undsp

Research Paper

Deep learning-based segmentation and detection of tunnel lining defects and components from GPR images using T-GPRMask

Jiahao Li [a], Hehua Zhu [a], Mei Yin [b,*]

[a] Department of Geotechnical Engineering, College of Civil Engineering, Tongji University, Shanghai 200092, China
[b] School of Civil Engineering, Southeast University, Nanjing 211189, China

Received 29 June 2025; received in revised form 23 July 2025; accepted 25 July 2025
Available online 26 August 2025

Abstract

Ground penetrating radar (GPR) has been extensively applied in tunnel engineering for the non-destructive assessment of lining structures. However, the interpretation of GPR images remains a time-consuming and expertise-dependent task. To address this challenge, this study proposes tunnel ground-penetrating radar mask region-based convolutional neural network (T-GPRMask), a deep learning-based instance segmentation model designed for the automated detection of tunnel lining defects and components. By integrating a convolutional block attention module (CBAM) and feature pyramid network (FPN), T-GPRMask enhances multi-scale feature extraction, enabling the detection of small, low-contrast defects that are commonly encountered in GPR images. The model was pretrained on a domain-specific dataset containing a diverse set of GPR images related to underground structures and then fine-tuned on a dataset specifically designed for tunnel inspections. The model achieved recognition accuracies of 83.18%, 88.24%, 92.84%, and 91.56% for detecting poor compactness, voids, steel arch supports, and initial lining thickness, respectively. A comparative study further demonstrated T-GPRMask's superior performance over traditional models, such as YOLOv7 and RetinaNet. Field experiments on real-world tunnel inspection data validated the model's high spatial accuracy and highlighted its practical applicability in tunnel maintenance.

Keywords: Ground penetrating radar; Tunnel lining inspection; Instance segmentation; Deep learning; Transfer learning

1 Introduction

The initial lining of tunnels, typically composed of flexible shotcrete, plays a crucial role in the load-bearing system of rock tunnels, working in conjunction with the surrounding rock to form a permanent structural framework. However, due to factors such as aging, construction defects, and insufficient maintenance, the initial lining often suffers from various hidden defects, such as insufficient lining thickness, poor compactness, excessive spacing between steel arch supports, and voids filled with air or water (as shown in Fig. 1). Timely detection and remediation of these defects during both the construction and operational phases are essential to ensure the structural safety and long-term stability of tunnels (Loupos et al., 2018). Among various non-destructive testing (NDT) methods, ground penetrating radar (GPR) has become widely used for evaluating tunnel lining quality, owing to its high spatial resolution and rapid data acquisition capabilities (Qin et al., 2020; Lai et al., 2018). Currently, the interpretation of GPR data largely depends on human–computer interaction, with its efficiency and accuracy heavily reliant on the operator's professional expertise and experience (De Coster et al., 2019; Guo et al., 2019). This process is often time-consuming and labor-intensive. For instance, GPR data collected in a single day using a vehicle-mounted system may require a week or more to process and interpret (Liu et al., 2021). These limitations significantly constrain the timeliness and accuracy of tunnel maintenance and repair decisions, thereby impeding the broader and more

* Corresponding author.
E-mail address: mei.yin@seu.edu.cn (M. Yin).
Peer review under the responsibility of Tongji University

https://doi.org/10.1016/j.undsp.2025.07.001

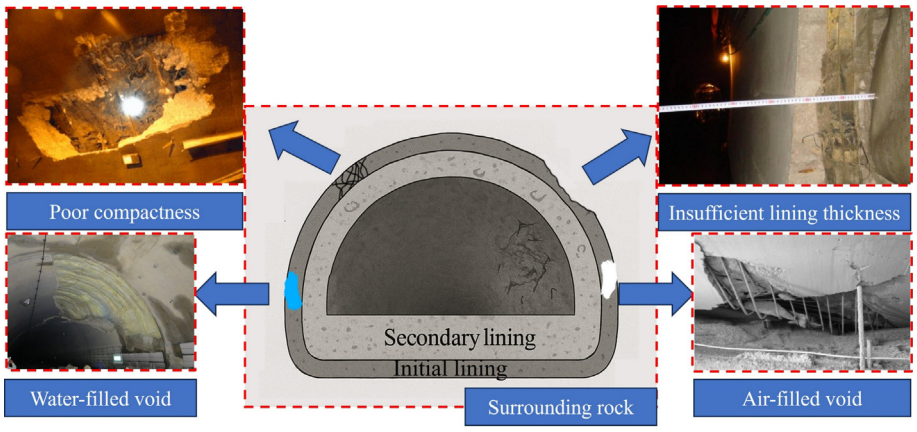


Fig. 1. Schematic illustration of hidden defects in tunnel linings.

efficient application of GPR technology. Consequently, there is an urgent need to develop automated recognition algorithms to substantially improve the interpretation efficiency of GPR data in tunnel inspections.

In recent years, various automated approaches have been introduced to improve the efficiency and accuracy of target recognition in GPR data interpretation. In the early stages, traditional image processing and machine learning techniques were extensively applied to subsurface utility detection tasks (Hoarau et al., 2017; Mertens et al., 2016), with the Hough transform (HT)-based hyperbola detection method being particularly representative (Borgioli et al., 2008; Maas & Schmalzl, 2013). As artificial intelligence has advanced, machine learning methods have been increasingly adopted for the automatic recognition and classification of GPR images (Youn & Chen, 2002). These approaches primarily rely on manual feature extraction and are well-suited for detecting geometrically regular targets (Lee & Mokji, 2014). However, in tunnel lining inspections—where high noise levels and complex backgrounds are prevalent—their recognition performance often deteriorates, increasing the likelihood of false positives and missed detections.

The rapid development of deep learning has opened new avenues for automatic target recognition in GPR images under complex conditions. Notably, object detection networks such as YOLO (Park et al., 2021; Qiu et al., 2022) and Faster R-CNN (Pham & Lefèvre, 2018; Lei et al., 2019) have been successfully applied in several studies for target detection in GPR imagery. These methods demonstrate significant potential in improving detection efficiency and achieving high accuracy, thereby offering promising prospects for the intelligent analysis of GPR data. However, despite these advantages, deep learning methods still face several challenges. Most existing models fail to fully integrate the engineering attributes of detection targets—such as tunnel lining defects and structural components—and typically provide only bounding box outputs, making it difficult to accurately describe the geometric shapes and structural characteristics of the targets. This limitation reduces their applicability and interpretability in practical engineering scenarios. Therefore, it is crucial to incorporate engineering semantics and geometric priors into both network architecture and data modeling to enhance the practicality and reliability of these models.

To tackle the aforementioned challenges, this study introduces a deep learning-based instance segmentation model, termed tunnel ground-penetrating radar mask region-based convolutional neural network (T-GPRMask), which incorporates an attention mechanism to enhance performance. The model is tailored to identify tunnel lining defects and components—such as poor compactness, voids, and steel arch supports—while also estimating lining thickness from GPR images. Building upon the mask region-based convolutional neural network (Mask R-CNN) framework, T-GPRMask integrates a convolutional block attention module (CBAM) to enhance its focus on critical target regions. It employs ResNet101 as the backbone for feature extraction and adopts a transfer learning strategy to improve training efficiency and overall performance. Initially, T-GPRMask is pre-trained on a GPR image dataset encompassing four typical categories: subsurface utilities, voids, intact areas, and simulated images. Subsequently, it is fine-tuned on a tunnel lining inspection dataset to enhance its adaptability and detection accuracy.

The primary innovations of T-GPRMask lie in its integration of CBAM and a feature pyramid network (FPN), which collectively address the challenges of multi-scale and low-contrast defect detection in GPR images. CBAM employs channel and spatial attention mechanisms to prioritize informative features while suppressing noise, thereby enhancing the model's ability to defect subtle signatures. Meanwhile, FPN facilitates effective fusion of high- and low-resolution features, enabling precise detection of small-scale defects that might be overlooked in traditional single-scale approaches.

The remainder of this paper is structured as follows: Section 2 reviews related work on GPR image interpretation and defect detection. Section 3 elaborates on the archi-

tecture and key components of the proposed T-GPRMask model. Section 4 details the construction of the tunnel lining GPR dataset used for training and evaluation, alongside the auxiliary datasets utilized in the transfer learning process. Section 5 presents the experimental results and representative case studies. Section 6 offers a comparative analysis of T-GPRMask with existing methods, while also addressing its limitations and potential directions for future research. Finally, Section 7 concludes the paper by summarizing the key findings.

2 Related work

Numerous studies have been conducted on the automatic interpretation of underground targets in GPR images, with the proposed solutions generally falling into two categories: (1) classical image processing and machine learning methods, and (2) deep learning methods.

2.1 Classical image processing and machine learning methods

For geometrically regular objects, such as underground pipelines and steel structures in tunnel linings, these targets typically appear as hyperbolic shapes in GPR images, making them key targets for early GPR image data interpretation. Capineri et al. (1998) proposed an approach based on an improved HT to detect straight lines and hyperbolas in GPR images, further estimating the location and shape of underground objects. Borgioli et al. (2008) introduced a weighted HT method to detect the position and burial depth of underground pipelines in GPR images. Maas and Schmalzl (2013) applied the Viola-Jones learning algorithm to automatically locate hyperbolic regions in GPR images, using HT to extract the precise position of hyperbolas. However, HT-based interpretation methods generally involve high computational costs and relatively high false positive rates (Li et al., 2016).

Clustering methods have been introduced for a more refined description of hyperbolic signals in GPR images. Dou et al. (2016) achieved hyperbolic feature recognition and parameter estimation in noisy environments by combining adaptive thresholding, the column-connection clustering (C3) algorithm, and a lightweight neural network classifier. Lei et al. (2019) achieved precise localization and curve fitting of hyperbolic features in both synthetic and real radar image data by combining a double-cluster seeking estimate (DCSE) algorithm with a column-based transverse filter points (CTFP) method.

2.2 Deep learning methods

With the continuous development of computer science and deep learning in computer vision, convolutional neural networks (CNNs) have been widely applied to classification, detection, and localization tasks for GPR images (Liu et al., 2023; Li et al., 2021). Qiu et al. (2022) improved the YOLOv5 network structure by integrating attention mechanisms, addressing the issues of false positives and false negatives in real-time GPR image detection. In the application of Faster R-CNN, Pham and Lefèvre (2018) used the Faster R-CNN framework to detect reflected hyperbolas in GPR images, enhancing performance through pretraining and fine-tuning on the CIFAR-10 dataset. Gong and Zhang (2020) validated the ability of Faster R-CNN for the automatic classification and recognition of GPR images using a simulated dataset generated by gprMax.

In tunnel lining detection, Guo et al. (2022) utilized the YOLOv3 model, combining both real and simulated data, to identify steel bars and voids in tunnel linings. Wang et al. (2022) introduced the rotational region deformation CNN (R2DCNN) model for detecting defects in tunnel linings from any direction. Qin et al. (2021) introduced an automatic recognition method based on Mask R-CNN, which effectively identifies defects such as steel bars and voids in tunnel linings, demonstrating high accuracy in both practical and numerical experiments. Yue et al. (2024) proposed an automatic defect recognition algorithm for railway tunnel GPR images based on the YOLOP model, which also estimates tunnel lining thickness. Recent studies in lining disease detection include Pan et al. (2025), who introduced EDeepLab for precise segmentation of tiny cracks in tunnel linings using a lightweight encoder-decoder architecture. Similarly, Zhou et al. (2023) proposed a pruned YOLOv4 model enhanced with Wasserstein generative adversarial network with residual enhancement data augmentation for efficient tunnel lining crack detection. Beyond GPR-based approaches, Afshani et al. (2019) explored infrared thermal imaging as a nondestructive method for detecting voids within tunnel linings, offering complementary insights to radar-based techniques. Gong et al. (2024) investigated the time-varying compressive properties of Ethylene–Propylene–Diene Monomer rubber materials used in tunnel gasketed joints, providing a foundation for understanding material degradation in structural components. Jin et al. (2025) developed a machine learning-based approach for intelligent safety evaluation of tunnel lining cracks, enabling quantitative assessment of crack stability and structural integrity.

Despite significant advancements in the automatic interpretation of tunnel lining GPR images, most methods still face limitations when dealing with complex image features and high-noise environments. Many models only output bounding box information and fail to fully incorporate the engineering attributes of the detected targets, limiting their practical application and interpretability in engineering scenarios.

3 T-GPRMask model

To precisely localize defects and components in tunnel linings from GPR images and generate high-quality segmentation masks, this study proposes an enhanced instance segmentation model called T-GPRMask. The complete

architecture of T-GPRMask is illustrated in Fig. 2, offering an overview of its structure. The overall architecture comprises three primary components: a convolutional backbone, a region proposal network (RPN), and a head architecture. Building upon the Mask R-CNN framework, T-GPRMask employs ResNet101 as the backbone for feature extraction, integrated with a FPN for multi-scale feature fusion. A CBAM is embedded within the backbone to enhance detection accuracy and robustness in complex subterranean environments. The RPN processes each scale of the feature maps output by the FPN (P2, P3, P4, P5) independently, rather than a single fused map. The model is specifically optimized to address the core challenges of strong clutter interference and varying defect scales in GPR images of tunnel linings.

For clarity, this section delineates the three core stages of the T-GPRMask instance segmentation process and their corresponding network modules: (i) A convolutional backbone network extracts structural features of the tunnel lining from the entire GPR image. (ii) An RPN generates candidate regions likely to contain tunnel lining structural features. (iii) A multi-task head network performs object classification, bounding box regression, and pixel-level mask prediction for each detected structural feature.

3.1 Stage 1: a convolutional backbone network extracts structural features of the tunnel lining from the entire GPR image

Deep CNNs are widely adopted as backbone architectures for extracting hierarchical features from images layer by layer. In this model, ResNet101 serves as the backbone, with FPN integrated to handle multi-scale features (not standalone). Due to variations in receptive fields and information content across convolutional layers, the feature maps produced by CNNs differ significantly in both spatial resolution and semantic richness. High-resolution feature maps typically capture low-level details such as edges and textures, whereas low-resolution feature maps contain more abstract, high-level semantic information that is advantageous for object recognition tasks. In earlier radar image interpretation studies, only high-level, low-resolution feature maps were often used for structural identification (Fig. 3(a)), which led to the neglect of high-resolution features essential for fine-grained target detection. This limitation becomes especially critical in GPR imagery, where target objects such as tunnel lining defects and components exhibit considerable variations in size. During the downsampling process, small targets are prone to semantic information loss in low-resolution feature maps.

To address this issue and enable the simultaneous detection of multi-scale targets—including poor compactness, steel arch supports, and voids—the model incorporates a FPN (Lin et al., 2017a) into the backbone architecture. The FPN fuses features from different convolutional layers via a top-down pathway and lateral connections, forming a multi-scale feature pyramid with strong semantic representation (Fig. 3(b)). The structure of the FPN is illustrated in the left part of Fig. 2. This architecture effectively preserves high-resolution spatial details without significantly increasing computational cost, thereby enhancing the model's capability in detecting small-scale objects. Consequently, it is particularly suitable for the identification of multiple categories of tunnel lining structural features in GPR images.

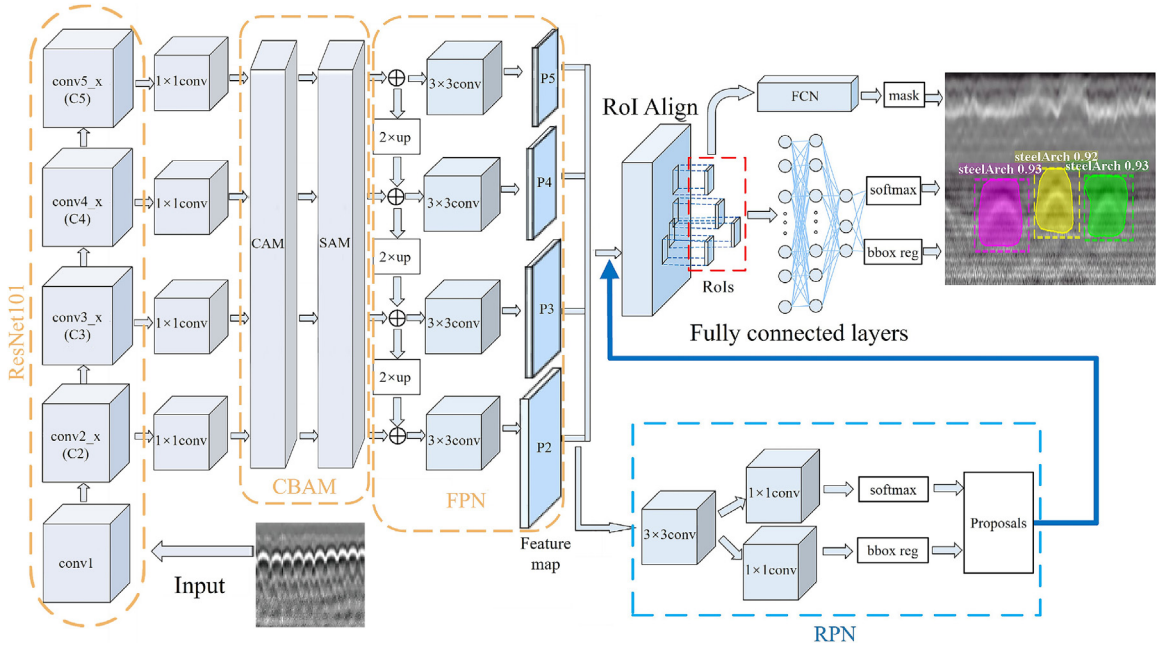


Fig. 2. Overall architecture of the T-GPRMask model (ResNet101 backbone integrated with FPN and CBAM).

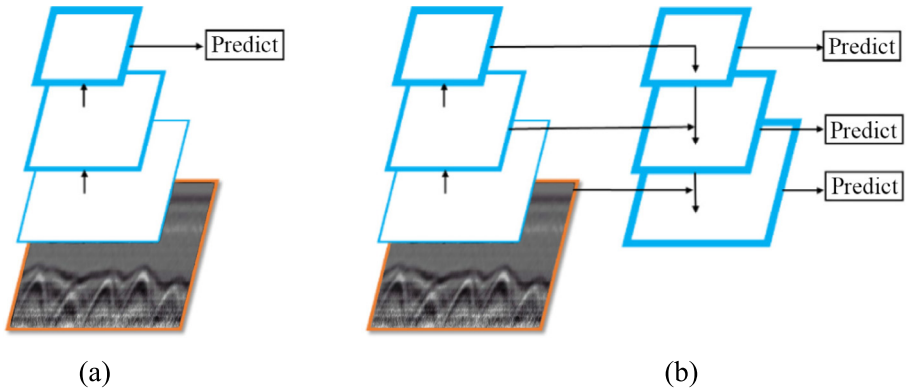


Fig. 3. Feature pyramid. (a) Single map, and (b) feature pyramid network. (Note: Thicker blue outlines denote features with richer information).

3.2 Stage 2: an RPN generates candidate regions likely to contain tunnel lining structural features

After feature extraction, object detection networks rely on region proposal algorithms to generate initial hypotheses regarding the locations of potential targets. This stage often becomes a major bottleneck in the inference speed of detection networks. To address this limitation, an RPN is employed to efficiently generate candidate regions. By sharing feature maps with the previously described convolutional backbone, the RPN introduces virtually no additional computational overhead.

The RPN functions as a lightweight subnetwork applied to the single-scale convolutional feature maps produced by the backbone. It evaluates objectness using dense 3×3 sliding windows. Specifically, a 3×3 convolutional layer is first applied to extract local features, followed by two parallel 1×1 convolutional layers responsible for binary classification (object vs. background) and bounding box regression, respectively (Fig. 4).

In Stage 1 of this model, an FPN is adopted to extract multi-scale features from GPR images, producing a set of feature maps at four levels: {P2, P3, P4, P5}. The RPN processes each scale of the FPN output feature maps separately to generate proposals, ensuring multi-scale sensitivity. During training, the model initially selects approximately 2000 candidate regions across all FPN levels. However, these initial RPN proposals are often highly redundant due to significant overlap. To address this, non-maximum suppression (NMS) algorithm based on class confidence scores (Ren et al., 2017) is applied to eliminate redundant proposals. The top 256 regions with the highest confidence scores are then retained for subsequent tasks, including object classification, bounding box regression, and mask segmentation.

3.3 Stage 3: a multi-task head network performs prediction for tunnel lining structural features

In Stage 2 of this model, the RPN generates a set of proposals corresponding to potential tunnel lining structural features based on the multi-scale feature maps. The NMS algorithm is then applied to eliminate redundant proposals. The remaining proposals, selected based on class confidence scores, are designated as regions of interest (RoIs) and forwarded to the RoI Align layer (He et al., 2017) to obtain spatially accurate feature representations. The RoI Align operation extracts fixed-size feature vectors from the multi-scale feature maps for each RoI, serving as input to subsequent classification and regression tasks.

These feature vectors are passed to the head architecture, which is constructed using an FCN, and processed by parallel task-specific branches responsible for object classification and bounding box regression. In parallel, a

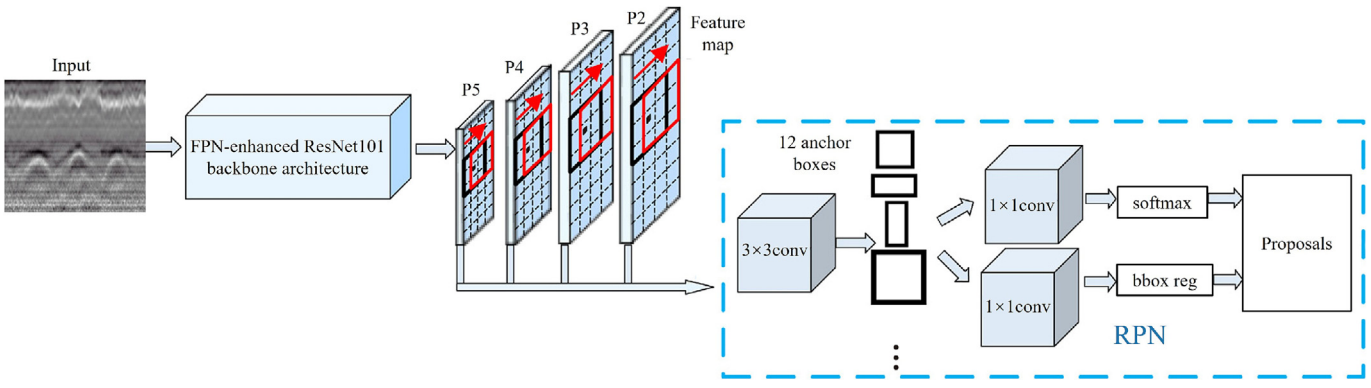


Fig. 4. RPN with an FPN-enhanced ResNet101 backbone architecture (RPN applied to multi-scale feature maps).

mask prediction subnetwork generates high-quality segmentation masks for each RoI, enabling precise delineation of tunnel lining defects and components.

3.4 Convolutional block attention module

The CBAM is incorporated into the T-GPRMask model to enhance feature extraction by selectively focusing on important spatial and channel-wise features, as shown in Fig. 5. CBAM operates in two stages: the channel attention module (CAM), which generates a weight map by applying global average and max pooling to emphasize the most informative channels, and the spatial attention module (SAM), which applies spatial attention to highlight critical regions within the feature maps (Woo et al., 2018). The integration of CBAM enables the model to suppress irrelevant information, thereby improving the accuracy of defect detection in tunnel lining GPR images. This attention mechanism is both lightweight and computationally efficient, offering an adaptive approach to prioritize key features without significantly increasing the model's computational burden. By incorporating CBAM, T-GPRMask is better equipped to capture subtle defects and complex patterns in GPR data, leading to enhanced performance in defect recognition tasks.

3.5 Loss function

During the training process of the instance segmentation model, a multi-task loss function is defined for each sampled RoI as follows:

$$L = L_{\text{cls}} + L_{\text{box}} + L_{\text{mask}}, \tag{1}$$

in which

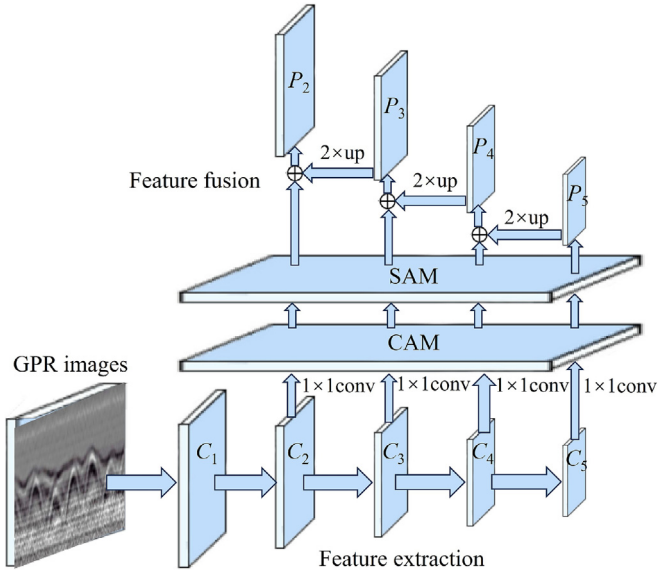


Fig. 5. Feature extraction network of T-GPRMask enhanced with CBAM.

$$\begin{cases} L_{\text{cls}} = -\ln p_u \\ L_{\text{box}} = \sum\limits_{i\in\{x,y,w,h\}} \text{smooth}_{L1}(t_i - t_i^*) \\ L_{\text{mask}} = \text{BCE}(m_k, \hat{m}_k) \end{cases}, \tag{2}$$

where L denotes the total loss, and L_{cls}, L_{box}, and L_{mask} represent the losses for classification, bounding box regression, and instance segmentation, respectively. Here, p_u is the predicted probability of the ground-truth class u (with u denoting the true class index); t_i and t_i^* represent the predicted and target bounding box regression values, respectively, where i iterates over the coordinates $\{x, y, w, h\}$ (representing center x-coordinate, y-coordinate, width, and height); smooth_{L1} is the smooth $L1$ loss function; BCE denotes the average binary cross-entropy loss per pixel; m_k and $\hat{m}_k$ are the ground-truth and predicted masks for the ground-truth class k (with k representing the true class index). Detailed descriptions of L_{cls} and L_{box} can be found in the work by Girshick (2015), and the binary cross-entropy loss is adopted for L_{mask} (He et al., 2017). This formulation enables the model to generate a mask for each class independently, without introducing competition among classes—an essential property for achieving high-quality instance segmentation results.

To better understand the impact of each component, the classification loss (L_{cls}) ensures accurate categorization of detected objects, which is crucial for distinguishing between various tunnel lining defects (e.g., poor compactness vs. voids) in noisy GPR images. The bounding box regression loss (L_{box}) refines the spatial localization of defects, enhancing the model's precision in identifying defect positions and sizes. Finally, the mask loss (L_{mask}) focuses on pixel-level segmentation, allowing for detailed contour delineation of defects and components, which directly contributes to improved interpretability and quantitative analysis in engineering applications. By balancing these losses, the model achieves robust performance across multi-scale and low-contrast targets typical in GPR data.

4 Datasets

Instance segmentation tasks rely heavily on large volumes of high-quality annotated image data and substantial computational resources. One of the primary challenges associated with conventional CNN-based approaches is their strong dependence on extensive manually labeled datasets and high computational costs. Although the development of annotation tools and advancements in graphics processing unit (GPU)-accelerated parallel computing have partially mitigated this issue, there remains a notable scarcity of publicly available, high-quality GPR image datasets with comprehensive annotations of tunnel lining structural features. Therefore, constructing a GPR dataset that encompasses typical tunnel lining defects and structural components holds significant practical value for achieving high-accuracy instance segmentation.

4.1 Dataset for network training

This study collected field GPR scan images from multiple railway tunnels located in southwestern China to construct a dataset for model training and testing (Fig. 6). GPR measurements were conducted along five longitudinal scan lines in each tunnel—covering the tunnel crown, left and right shoulders, and both sidewalls—ensuring comprehensive spatial coverage of the tunnel lining structures. To ensure image quality, the raw radar data underwent a standardized preprocessing pipeline before model training, including direct current removal, bandpass filtering, time-varying gain correction, background removal, and time-zero correction.

In total, 1262 GPR images were collected and manually annotated by experienced professionals with extensive field expertise. The original annotated dataset includes 492 instances of poor compactness, 291 voids, and 942 steel arch supports. Given the relatively limited number of samples in certain categories, systematic data augmentation was applied to the original GPR images to enhance the model's generalization ability and reduce the risk of overfitting. The augmentation strategy comprised two main components: geometric transformations and noise simulation. For geometric transformations, random rotations, scaling, and flipping were performed to simulate the variability in the appearance of tunnel lining defects and structural features under different orientations, positions, and scales. This improved the model's robustness to spatial deformations and positional variations. In parallel, random noise injection was employed to mimic common non-ideal factors in GPR data acquisition, such as sensor-induced noise and environmental interference. This enabled the model to maintain stable recognition performance under noisy input conditions. After argumentation, the sample distribution was expanded to 942 for poor compactness, 541 for voids, 992 for steel arch supports, and 1782 for lining thickness, as summarized in Table 1. Representative examples of each category are also presented. Some images contain multiple types of tunnel lining defects and components simultaneously. All annotated GPR images were divided into training, validation, and test sets in a ratio of 7:2:1 for model training, hyperparameter tuning, and evaluation.

4.2 Dataset for transfer learning

To enhance the target recognition performance of the model on tunnel lining GPR images, a transfer learning strategy was employed to pretrain and fine-tune the T-GPRMask model. The dataset for transfer learning included four typical categories of GPR images: subsurface utilities, voids, intact areas, and simulated images generated using the finite-difference time-domain (FDTD) method, as shown in Fig. 7. A total of 4168 GPR images were collected, including 786 images of subsurface utilities, 553 images of voids, 900 images of intact areas (Mojahid et al., 2025), and 1929 FDTD-simulated images.

All images in this dataset were accompanied by detailed manual annotations, which covered typical underground object contours, background noise characteristics, and radar response patterns to dielectric discontinuities. By pretraining the model on relevant GPR tasks and transferring the learned parameters to the tunnel lining defect detection task, this study effectively reduced the distribution gap between different image modalities and application domains, thereby significantly improving the model's generalization capability and accelerating training convergence.

5 Experiment

5.1 Training and testing

To modify and train the proposed T-GPRMask model, all experiments were conducted on a workstation equipped with an Intel Core i7-14700HX CPU, 64 GB of RAM, and an NVIDIA GeForce RTX 4060 GPU. The model was optimized using the Adam optimizer with a learning rate of 0.0005, a weight decay of 0.0001, and a first-moment exponential decay rate (β_1) of 0.9. Hyperparameter tuning was conducted using both the training and validation

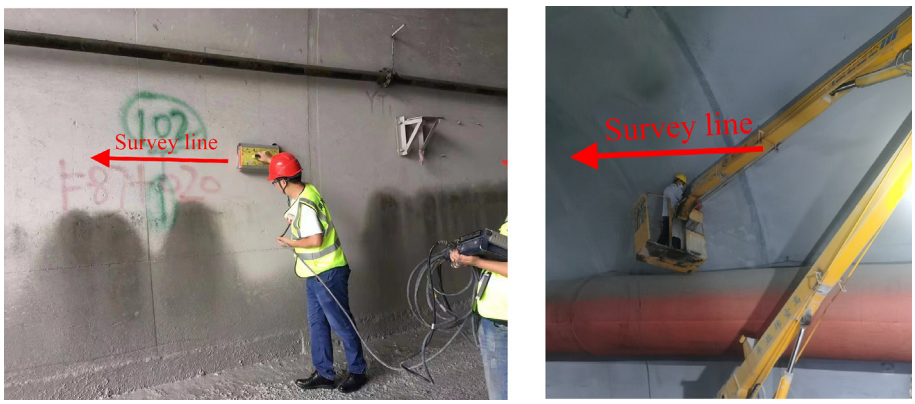
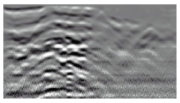
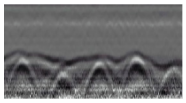
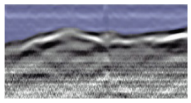


Fig. 6. GPR data acquisition in a rock tunnel.

Table 1
Data augmentation of the GPR dataset for tunnel lining inspection.

Feature type	Poor compactness	Voids	Steel arch supports	Lining thickness
Sample size	942	541	992	1782
Sample images				

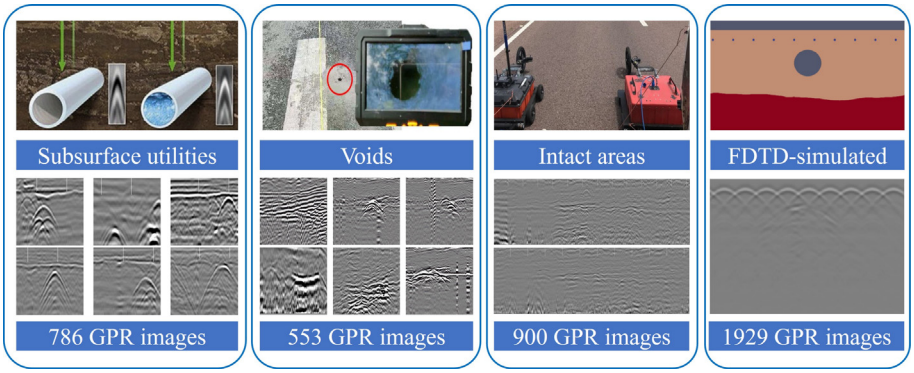


Fig. 7. GPR images used for transfer learning in T-GPRMask model pre-training.

datasets to evaluate the effects of different configurations on model performance. Based on the tuning results, the optimal hyperparameter set was selected to train the final model. Training was carried out with a batch size of 16 over 150 epochs. The initial loss started at 1.342 and dropped sharply to 0.196 within the first 20 epochs, followed by a gradual decline. By epoch 150, the loss had converged to 0.034, as illustrated in Fig. 8. The validation loss closely tracked the training loss throughout the process, indicating no signs of overfitting. The trained model weights were saved and subsequently used for performance evaluation and deployment.

In this study, we adopt the following metric to evaluate the overall recognition performance of the model:

$$\text{Recognition Accuracy} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}, \tag{3}$$

where TP (true positives) refers to the number of correctly identified targets, FP (false positives) refers to the number of incorrectly identified targets (false alarms), and FN (false negatives) represents the number of targets that were missed by the model. This metric jointly considers both types of recognition errors—misses and false detections—providing a balanced assessment of the model's detection capability.

On the dataset constructed in this study, the T-GPRMask model achieved average recognition accuracies of 83.18%, 88.24%, 92.84%, and 91.56% for poor compactness, voids, steel arch supports, and initial lining thickness estimation, respectively. The trained instance segmentation model was further applied to real GPR images

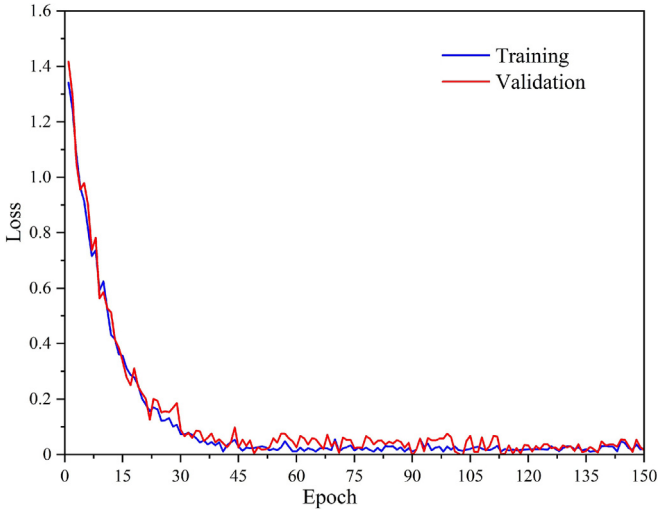


Fig. 8. Detection examples of tunnel lining structural features from GPR images by T-GPRMask model.

obtained from practical tunnel lining quality inspections, where it successfully and automatically extracted relevant defect and structural feature information from the data. The recognition results are visualized in Fig. 9: poor compactness is marked with yellow bounding boxes and masks, voids with red, steel arch supports with green, and initial lining with blue masks. The predicted results show strong agreement with manual annotations provided by experienced professionals, thereby validating the effectiveness and practical applicability of the T-GPRMask model in real-world tunnel inspection scenarios.

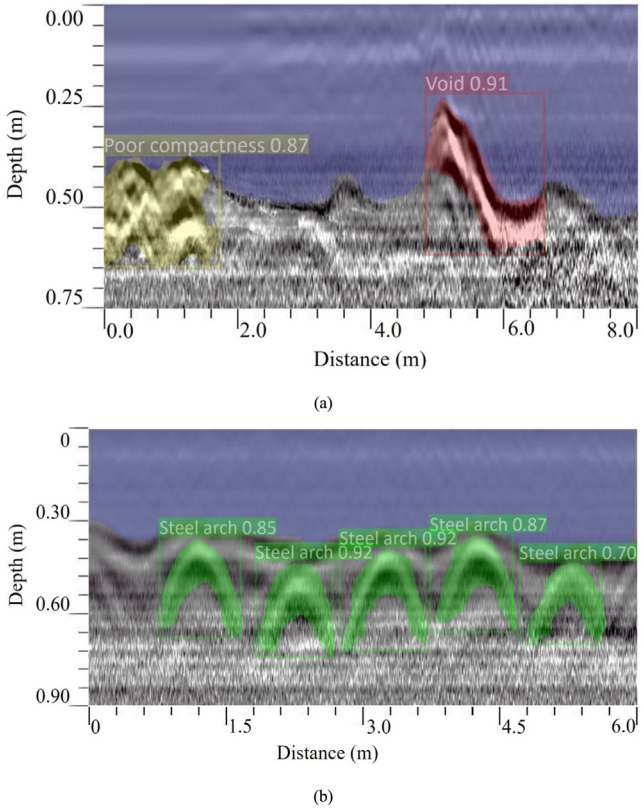


Fig. 9. Detection examples of tunnel lining structural features from GPR images by T-GPRMask model (a) Defects detection, and (b) components detection.

5.2 Ablation experiments

To validate the contributions of key modules in T-GPRMask, we conducted ablation experiments by removing CBAM and FPN individually and evaluating the impact on detection accuracy. The experiments were performed on the same dataset and hardware as described in Section 5.1. Table 2 summarizes the results, showing that removing CBAM leads to a 5%–8% drop in accuracy for low-contrast defects (e.g., voids), as it reduces the model's ability to focus on critical features. Without FPN, multi-scale detection suffers, resulting in a 7%–10% decline for small targets like poor compactness. These findings confirm the essential roles of CBAM and FPN in enhancing overall performance.

5.3 Field experiment

To validate the recognition capability of the T-GPRMask model in real-world engineering scenarios, it was applied to GPR images collected from a physical tunnel testing site containing embedded defects. The field experiment was conducted at the Long-term and Safe Physical Testing Site for Tunnel Engineering in Qingyuan, Guangdong Province, China. During the construction of the defect section, typical structural defects were simulated by embedding prefabricated concrete cavity models and inflatable fabric bags within the tunnel lining (as shown in Fig. 10). The embedded void measured 0.90 m × 0.20 m and 0.40 m in thickness.

As shown in Fig. 11(a), data acquisition was performed using a self-developed, vehicle-mounted phased-array radar system that traversed a 4.0 m measurement line over the defective section at a constant speed. The collected GPR data were processed using a standard preprocessing workflow and subsequently input into the trained T-GPRMask model for defect recognition.

Figure 11(b) illustrates the detection result on a representative GPR image using the T-GPRMask model from the defect section. At approximately 0.98 m along the measurement line, the model successfully identified a void with strong reflection characteristics and assigned it a high confidence score. By comparing the geometric center of the predicted segmentation mask with the known embedding location, it was found that the errors in both depth and horizontal position were within 0.04 m, indicating high spatial accuracy.

These field experiment results confirm that T-GPRMask can effectively detect concealed defects and components in tunnel linings and accurately characterize key parameters such as lining thickness. This demonstrates the model's strong engineering applicability and its potential for large-scale deployment in practical tunnel inspection tasks.

6 Discussion

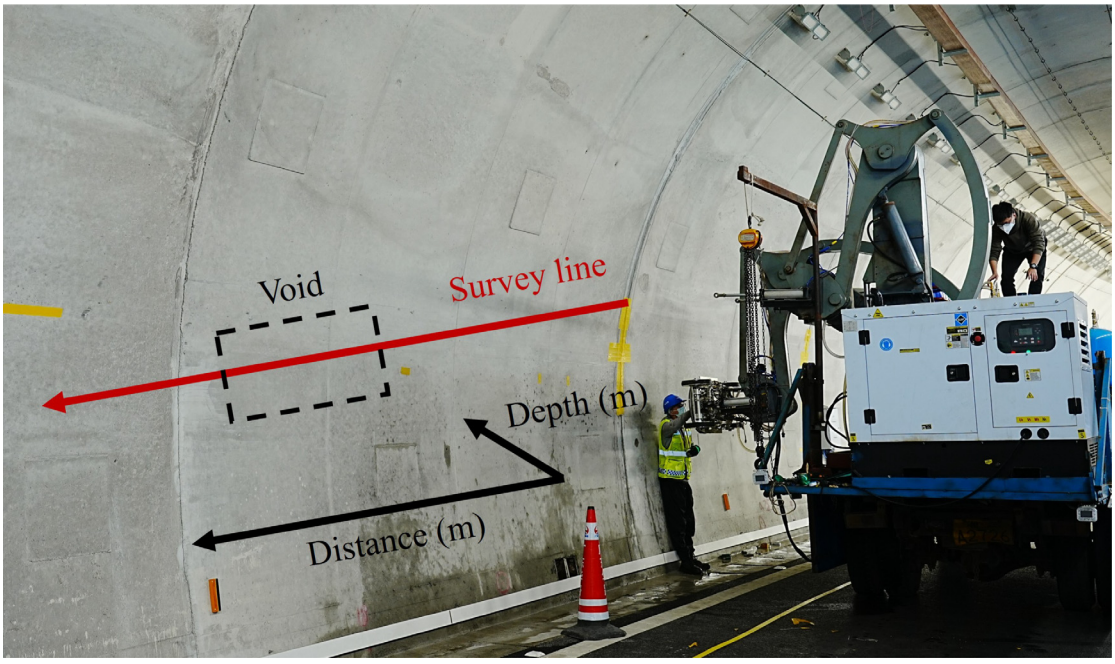
6.1 Comparative study

To comprehensively evaluate the performance of T-GPRMask, comparative experiments were conducted against two widely adopted object detection algorithms: YOLOv7 (Wang et al., 2023) and RetinaNet (Lin et al., 2017a, 2017b). All models were trained and tested on the same annotated GPR dataset. The recognition accuracies for each category of tunnel lining structural features are summarized in Table 3. The results demonstrate that T-GPRMask achieved the highest accuracy across all categories. Specifically, the recognition accuracy of T-GPRMask for poor compactness was 10.83 percentage

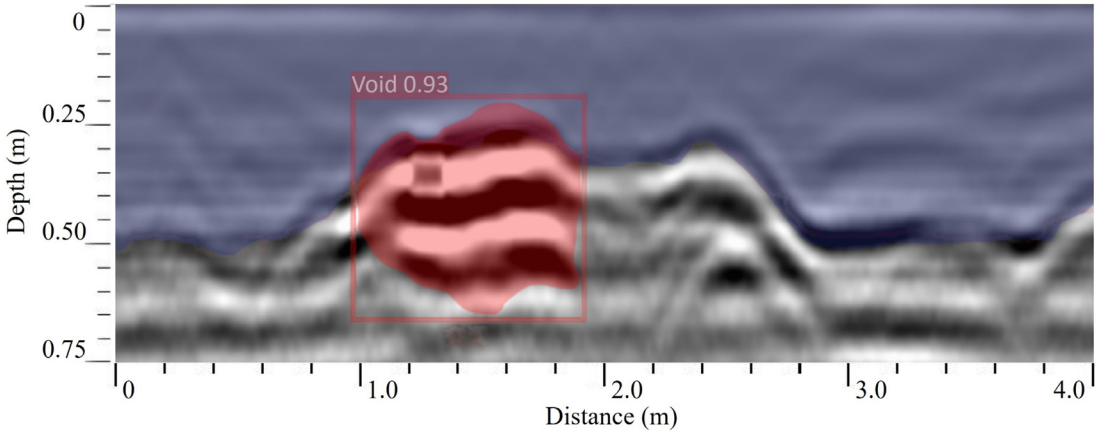
Table 2
Ablation experiment results on key modules.

Model variant	Recognition accuracy (%)			
	Poor compactness	Voids	Steel arch supports	Lining thickness
Without FPN	75.28	82.36	85.92	84.67
Without CBAM	77.50	80.12	88.76	87.45
T-GPRMask (full)	83.18	88.24	92.84	91.56

Fig. 10. Embedded structural defects in the tunnel lining at the physical testing site.



(a)



(b)

Fig. 11. Field data acquisition and defect detection with T-GPRMask model. (a) Data acquisition using vehicle-mounted phased-array radar system, and (b) detection result on a representative GPR image using the T-GPRMask model.

Table 3
Accuracy comparison of different recognition models.

Model	Recognition accuracy (%)			
	Poor compactness	Voids	Steel arch supports	Lining thickness
YOLOv7	72.35	80.16	85.82	–
RetinaNet	75.92	74.28	83.56	–
T-GPRMask	83.18	88.24	92.84	91.56

points higher than YOLOv7 and 7.26 percentage points higher than RetinaNet; for void detection, it was 8.08 percentage points higher than YOLOv7 and 13.96 percentage points higher than RetinaNet; and for steel arch support detection, it was 7.02 percentage points higher than YOLOv7 and 9.28 percentage points higher than RetinaNet.

Figure 12 presents the detection results of two comparative models on the same GPR image: YOLOv7 (Fig. 12 (a)), and RetinaNet (Fig. 12(b)), visually compared with the segmentation output of T-GPRMask (as shown in Fig. 9). It is evident that YOLOv7 failed to detect the poor compactness, and both comparison models missed certain instances of steel arch supports.

In terms of overall performance, T-GPRMask exhibited superior capability in detecting small-scale and low-contrast targets, particularly in recognizing weak-signal and blurry-boundary features such as voids and poor compactness in GPR images. This advantage is largely attributed to the integration of the FPN and the CBAM. The FPN improves detection of objects at different scales, while the CBAM improves the model's focus on critical target regions.

Moreover, T-GPRMask provides a significant functional advantage over YOLOv7 and RetinaNet: it incorporates a mask prediction branch that enables accurate segmentation of the initial tunnel lining. This functionality allows for precise delineation of the spatial boundaries of tunnel lining defects and components, providing refined data support for defect area estimation and quantitative analysis. These results highlight the strong potential of T-GPRMask for practical engineering applications.

In addition to accuracy, processing time is a critical factor for practical engineering applications in tunnel inspections, particularly for real-time or semi-real-time processing in vehicle-mounted GPR systems. We evaluated the average inference time per GPR image on the same hardware (NVIDIA GeForce RTX 4060 GPU). As shown in Table 4, YOLOv7 exhibited the fastest processing speed at approximately 78 ms per image (around 13 FPS), followed by RetinaNet at 156 ms per image (around 6 FPS). T-GPRMask, due to its additional mask prediction branch, required about 484 ms per image (around 2 FPS). Although slightly slower than the one-stage detectors, this processing time still meets the demands of most mobile detection scenarios, where semi-real-time analysis

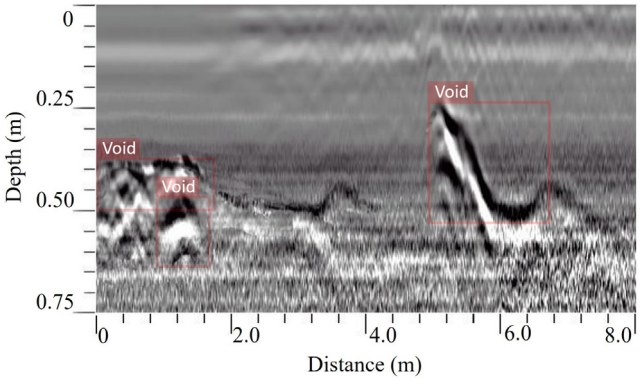


Fig. 12. Detection results of two comparative models. (a) YOLOv7, and (b) RetinaNet.

(e.g., processing during or immediately after scanning) is sufficient for engineering needs. For instance, in a typical tunnel scan at 10–20 km/h, T-GPRMask can handle data streams effectively without significant delays. For scenarios demanding higher throughput, optimizations such as model pruning or deployment on edge devices could be explored in future work.

6.2 Comparison of pre-training datasets

To evaluate the impact of different pretraining datasets on defect recognition model performance, this study used the COCO dataset, Open Images V6 dataset, and the GPR dataset introduced in Section 4.2. The T-GPRMask model pretrained on these three datasets was then trained on the tunnel lining detection GPR image dataset established in Section 4.1 for 150 epochs. The training loss curves are shown in Fig. 13, and Table 5 lists the final recognition accuracy of T-GPRMask under different pretraining datasets.

As shown in Fig. 13, the model pretrained on the GPR dataset converged the fastest, with the training loss steadily decreasing and reaching the lowest final loss after 150 epochs. As indicated in Table 5, the model pretrained on the GPR dataset significantly outperformed other pretrained models, achieving the highest recognition accuracy. This highlights that the pretraining dataset has a significant impact on the performance of the T-GPRMask model,

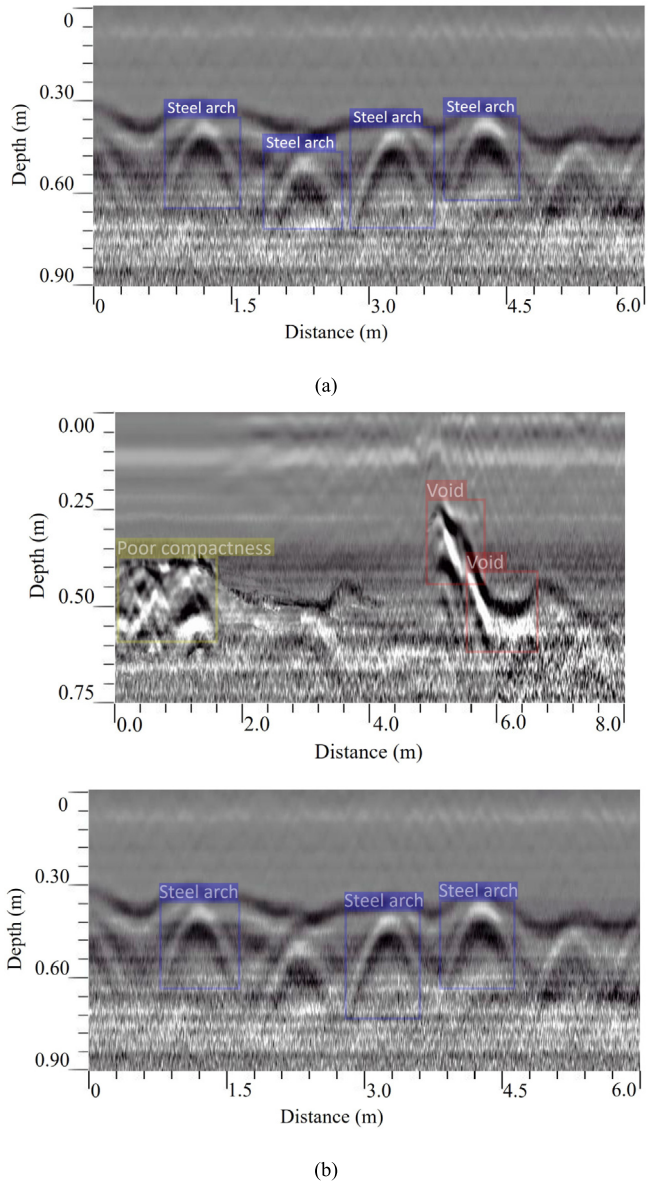


Fig 12. *(continued)*

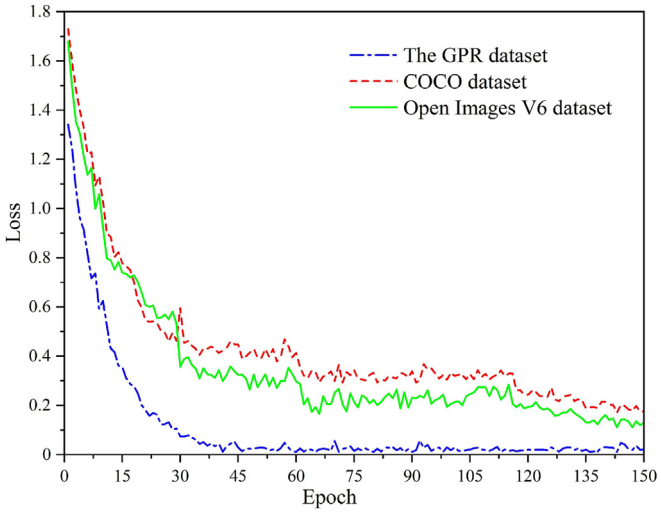


Fig. 13. Training loss curves of T-GPRMask pretrained on different datasets.

primarily due to the task relevance and image modality differences.

On one hand, GPR images typically exhibit low contrast, strong noise, and complex reflection patterns, while the COCO and Open Images V6 datasets mainly contain color natural images that lack features related to tunnel structures and underground defects. Therefore, models pretrained on these datasets perform poorly when adapting to GPR image characteristics, exhibiting slow convergence and higher final training loss. On the other hand, the GPR dataset specifically annotates features highly relevant to tunnel lining defects, and the FDTD simulation data covers extreme scenarios (e.g., deep-buried voids, multilayer structures), providing a more targeted pretraining foundation. This allows the model to converge faster during the fine-tuning phase and improves recognition accuracy.

The experimental results show that pretraining on domain-specific datasets significantly improved GPR image recognition performance, with accuracy increasing by approximately 6%–9%. This result validates that pretraining on domain-specific data allows the model to better generalize to the target task, facilitating faster adaptation and superior defect recognition performance. This is crucial for the application of deep learning in specialized engineering tasks such as tunnel lining detection.

6.3 Limitations and future work

Although the T-GPRMask model achieves high accuracy in detecting tunnel lining defects and components, it does not distinguish between air-filled and water-filled voids, treating them as a unified "voids" category. This limitation arises not from inherent constraints of the GPR method—which can differentiate void types through

Table 4
Processing time comparison of different models (per image, ms).

Model	Average inference time (ms)	Frames per second (FPS, approx.)
YOLOv7	78	13
RetinaNet	156	6
T-GPRMask	484	2

Table 5
Accuracy comparison of the T-GPRMask pre-trained on different datasets.

Pre-training datasets	Recognition accuracy (%)			
	Poor compactness	Voids	Steel arch supports	Lining thickness
COCO dataset	74.95	82.26	87.65	85.88
Open Images V6 dataset	76.74	80.42	88.38	84.46
The GPR dataset	83.18	88.24	92.84	91.56

variations in electromagnetic responses—but from the dataset design and task simplification in this study. Subclassification could enhance interpretability but would necessitate expanded annotations and potentially more advanced feature extraction techniques.

Future research will explore integrating void type differentiation and incorporating temporal GPR data for dynamic monitoring, such as tracking the evolution of defects like voids over time using sequential scans from vehicle-mounted systems. Additionally, we plan to adapt T-GPRMask for other types of infrastructure, including bridges and underground pipelines, by fine-tuning on domain-specific datasets and integrating multi-modal data (e.g., combining GPR with infrared or ultrasonic imaging). These extensions could address challenges in diverse environments, such as varying noise levels or structural complexities, further improving the model's generalization and practical utility.

7 Conclusions

In this study, we proposed the T-GPRMask model, a deep learning-based instance segmentation framework designed to accurately detect tunnel lining defects and components from GPR images. The model integrates a CBAM and an FPN to enhance multi-scale feature extraction and improve the detection of small-scale, low-contrast defects. The T-GPRMask model demonstrated superior performance in comparison with other popular object detection models such as YOLOv7 and RetinaNet, achieving high accuracy in recognizing tunnel lining defects like poor compactness, voids, and steel arch supports. Moreover, by leveraging transfer learning on a GPR-specific dataset, the model achieved a significant improvement in recognition performance, highlighting the importance of domain-specific pretraining.

The T-GPRMask model was further validated in real-world field experiments, where it effectively detected embedded defects in physical tunnel sites, proving its engineering applicability and potential for large-scale deployment. The results underscore the model's ability to assist in automatic, accurate, and efficient tunnel inspection, reducing the reliance on manual labor and increasing the efficiency of maintenance decisions. The proposed model represents a promising tool for future tunnel maintenance, with potential applications across various types of infrastructure inspections.

Future work will explore the incorporation of temporal GPR data for dynamic defect tracking, as well as the integration of domain-specific prior knowledge to improve model interpretability and generalization across diverse tunnel environments.

Data availability

The data that used in this study are available on https://github.com/ljhtj97/T-GPRMask.

CRediT authorship contribution statement

Jiahao Li: Writing – original draft, Visualization, Validation, Methodology, Data curation. **Hehua Zhu:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition. **Mei Yin:** Writing – review & editing, Investigation, Data curation.

Declaration of competing interest

Professor Hehua Zhu is an editor-in-chief for *Underground Space* and was not involved in the editorial review or the decision to publish this article. All authors declare that there are no competing interests.

Acknowledgement

This research is supported by the National Key Research and Development Program of China (Grant No. 2023YFC3009400).

References

Afshani, A., Kawakami, K., Konishi, S., & Akagi, H. (2019). Study of infrared thermal application for detecting defects within tunnel lining. *Tunnelling and Underground Space Technology, 86*, 186–197.

Borgioli, G., Capineri, L., Falorni, P., Matucci, S., & Windsor, C. G. (2008). The detection of buried pipes from time-of-flight radar data. *IEEE Transactions on Geoscience and Remote Sensing, 46*(8), 2254–2266.

Capineri, L., Grande, P., & Temple, J. A. G. (1998). Advanced image-processing technique for real-time interpretation of ground-penetrating radar images. *International Journal of Imaging Systems and Technology, 9*(1), 51–59.

De Coster, A., Medina, J. P., Nottebaere, M., Alkhalifeh, K., Neyt, X., Vanderdonckt, J., & Lambot, S. (2019). Towards an improvement of

GPR-based detection of pipes and leaks in water distribution networks. *Journal of Applied Geophysics, 162*, 138–151.
Dou, Q. X., Wei, L. J., Magee, D. R., & Cohn, A. G. (2016). Real-time hyperbola recognition and fitting in GPR data. *IEEE Transactions on Geoscience and Remote Sensing, 55*(1), 51–62.
Girshick, R. (2015). Fast R-CNN. In *Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)* (pp. 1440–1448).
Gong, C. J., Xie, C. R., Zhu, H. H., Ding, W. Q., Song, J. R., & Ge, Y. Y. (2024). Time-varying compressive properties and constitutive model of EPDM rubber materials for tunnel gasketed joint. *Construction and Building Materials, 433*, 136734.
Gong, Z. M., & Zhang, H. Q. (2020). Research on GPR image recognition based on deep learning. *MATEC Web of Conferences, 309*, 03027.
Guo, L., Cai, L. H., & Chen, D. J. (2022). Research on tunnel lining image target recognition method based on YOLOv3. In *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)* (pp. 1800–1805).
Guo, S. L., Deng, J., Li, W. W., & Zhang, X. Q. (2019). Ground penetrating radar multi-profile integrated interpretation method and application. *Progress in Geophysics, 34*(5), 2022–2029.
He, K. M., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 2961–2969).
Hoarau, Q., Ginolhac, G., Atto, A. M., & Nicolas, J. M. (2017). Robust adaptive detection of buried pipes using GPR. *Signal Processing, 132*, 293–305.
Jin, X., Qiao, K. K., Bu, M. H., Wang, J. M., Wang, M., & Fang, C. (2025). Intelligent safety evaluation of tunnel lining cracks based on machine learning. *Engineering Failure Analysis, 167*, 109082.
Lai, W. W. L., Dérobert, X., & Annan, P. (2018). A review of Ground Penetrating Radar application in civil engineering: A 30-year journey from locating and testing to imaging and diagnosis. *NDT & E International, 96*, 58–78.
Lee, K. L., & Mokji, M. M. (2014). Automatic target detection in GPR images using Histogram of Oriented Gradients (HOG). In *2014 2nd International Conference on Electronic Design (ICED)* (pp. 181–186).
Lei, W. T., Hou, F. F., Xi, J. C., Tan, Q. Y., Xu, M. D., Jiang, X. Y., Liu, G. Y., & Gu, Q. Y. (2019). Automatic hyperbola detection and fitting in GPR B-scan image. *Automation in Construction, 106*, 102839.
Li, S. W., Gu, X. Y., Xu, X. R., Xu, D. W., Zhang, T. J., Liu, Z., & Dong, Q. (2021). Detection of concealed cracks from ground penetrating radar images based on deep learning algorithm. *Construction and Building Materials, 273*, 121949.
Li, W. T., Cui, X. H., Guo, L., Chen, J., Chen, X. H., & Cao, X. (2016). Tree root automatic recognition in ground penetrating radar profiles based on randomized Hough transform. *Remote Sensing, 8*(5), 430.
Lin, T. Y., Dollár, P., Girshick, R., He, K. M., Hariharan, B., & Belongie, S. (2017a). Feature pyramid networks for object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 936–944).
Lin, T. Y., Goyal, P., Girshick, R., He, K. M., & Dollár, P. (2017b). Focal loss for dense object detection. In *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 2999–3007).
Liu, H., Shi, Z. S., Li, J. H., Liu, C., Meng, X., Du, Y. L., & Chen, J. (2021). Detection of road cavities in urban cities by 3D ground-penetrating radar. *Geophysics, 86*(3), WA25–WA33.
Liu, Z., Gu, X. Y., Chen, J. Q., Wang, D. Y., Chen, Y. H., & Wang, L. T. (2023). Automatic recognition of pavement cracks from combined GPR B-scan and C-scan images using multiscale feature fusion deep neural networks. *Automation in Construction, 146*, 104698.
Loupos, K., Doulamis, A. D., Stentoumis, C., Protopapadakis, E., Makantasis, K., Doulamis, N. D., Amditis, A., Chrobocinski, P., Victores, J., Montero, R., Menendez, E., Balaguer, C., Lopez, R., Cantero, S., Navarro, R., Roncaglia, A., Belsito, L., Camarinopoulos, S., Komodakis, N., & Singh, P. (2018). Autonomous robotic system for tunnel structural inspection and assessment. *International Journal of Intelligent Robotics and Applications, 2*, 43–66.
Maas, C., & Schmalzl, J. (2013). Using pattern recognition to automatically localize reflection hyperbolas in data from ground penetrating radar. *Computers & Geosciences, 58*, 116–125.
Mertens, L., Persico, R., Matera, L., & Lambot, S. (2016). Automated detection of reflection hyperbolas in complex GPR images with no *a priori* knowledge on the medium. *IEEE Transactions on Geoscience and Remote Sensing, 54*(1), 580–596.
Mojahid, A., El Ouai, D., El Amraoui, K., El-Hami, K., & Aitbenamer, H. (2025). Intelligent recognition of subsurface utilities and voids: A ground penetrating radar dataset for deep learning applications. *Data in Brief, 59*, 111338.
Pan, Z. T., Zhang, X. P., Jiang, Y. J., Li, B., Golsanami, N., Su, H., & Cai, Y. (2025). High-precision segmentation and quantification of tunnel lining crack using an improved DeepLabV3+. *Underground Space, 22*, 96–109.
Park, S., Kim, J., Jeon, K., Kim, J., & Park, S. (2021). Improvement of GPR-based rebar diameter estimation using YOLO-v3. *Remote Sensing, 13*(10), 2011.
Pham, M. T., & Lefèvre, S. (2018). Buried object detection from B-scan ground penetrating radar data using Faster-RCNN. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 6804–6807). IEEE.
Qin, H., Xie, X. Y., Tang, Y., & Wang, Z. Z. (2020). Experimental study on GPR detection of voids inside and behind tunnel linings. *Journal of Environmental and Engineering Geophysics, 25*(1), 65–74.
Qin, H., Zhang, D. H., Tang, Y., & Wang, Y. Z. (2021). Automatic recognition of tunnel lining elements from GPR images using deep convolutional networks with data augmentation. *Automation in Construction, 130*, 103830.
Qiu, Z., Zhao, Z. X., Chen, S. J., Zeng, J. Y., Huang, Y., & Xiang, B. R. (2022). Application of an improved YOLOv5 algorithm in real-time detection of foreign objects by ground penetrating radar. *Remote Sensing, 14*(8), 1895.
Ren, S. Q., He, K. M., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(6), 1137–1149.
Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 7464–7475).
Wang, J., Zhang, J. Q., Cohn, A. G., Wang, Z. F., Liu, H. C., Kang, W. Q., Jiang, P., Zhang, F. K., Chen, K. F., Guo, W., & Yu, Y. F. (2022). Arbitrarily-oriented tunnel lining defects detection from ground penetrating radar images using deep convolutional neural networks. *Automation in Construction, 133*, 104044.
Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In *Proceedings of the 15th European Conference on Computer Vision* (pp. 3–19).
Youn, H. S., & Chen, C. C. (2002). Automatic GPR target detection and clutter reduction using neural network. In *Ninth International Conference on Ground Penetrating Radar (GPR 2002)* (pp. 579–582). SPIE.
Yue, Y. P., Liu, H., Lin, C. D., Meng, X., Liu, C., Zhang, X. Y., Cui, J., & Du, Y. L. (2024). Automatic recognition of defects behind railway tunnel linings in GPR images using transfer learning. *Measurement, 224*, 113903.
Zhou, Z., Zhang, J. J., Gong, C. J., & Wu, W. (2023). Automatic tunnel lining crack detection via deep learning with generative adversarial network-based data augmentation. *Underground Space, 9*, 140–154.