Issues in Information Systems
Volume 27, Issue 4, pp. 686-694, 2026

DOI: https://doi.org/10.48009/4_iis_2026_157

Comparative Study of Deep Learning Models in GPR-Based Structural Health Monitoring

Yongjun Lee, *Korea Advanced Institute of Science and Technology (KAIST),* yongjun.lee@kaist.ac.kr
Joong-Lyul Lee, *Gyeongsang National University,* joonglyul.lee@gnu.ac.kr

Abstract

Structural Health Monitoring (SHM) plays a critical role in ensuring the safety and longevity of civil infrastructure by supporting the interpretation of subsurface conditions such as embedded reinforcement, delamination, voids, and other anomalies. Among non-destructive testing techniques, Ground Penetrating Radar (GPR) is widely used for subsurface inspection, where embedded objects and anomaly-related responses often appear as hyperbolic patterns in B-scan images. However, accurate interpretation of GPR data remains challenging because it depends heavily on expert knowledge and labor intensive manual analysis. This study presents a comparative analysis of three object detection models YOLOv8, YOLO11, and Faster R-CNN for automatic hyperbola detection in GPR B-scan images. The experiments were conducted using a GPR B-scan object detection dataset consisting of 512 images and 6,138 labeled instances. The dataset was divided into 328 training images and 184 validation images, and the three models were independently trained and evaluated under the same class definition. The evaluated classes were noise, hyperbola, and hyperbola line. The results show that YOLOv8 achieved the best overall AP based detection performance, with an mAP50 of 0.824 and an mAP50–95 of 0.503. YOLO11 achieved the highest precision and recall, with values of 0.754 and 0.770, respectively, while also providing the fastest inference speed. Faster R-CNN trained for 15,000 iterations showed stable loss convergence but achieved lower validation performance, with an mAP50 of 0.468 and an mAP50–95 of 0.262. These results indicate that YOLO-based one stage detectors are more suitable than the evaluated Faster R-CNN baseline for noisy and geometrically complex GPR B-scan hyperbola detection.

Keywords: Structural Health Monitoring, Deep Learning, Ground Penetrating Radar Image Analysis

Introduction

Structural Health Monitoring (SHM) is an essential component in the maintenance and management of civil infrastructure, providing the means to detect, assess, and monitor structural integrity over time. As infrastructures such as bridges, buildings, and highways age, timely interpretation of internal conditions such as reinforcement layout, corrosion related changes, delamination, and voids becomes increasingly important for public safety and maintenance planning. Among non-destructive testing (NDT) techniques, Ground Penetrating Radar (GPR) has gained substantial attention for subsurface inspection (D. J. Daniels, 2005). In GPR-based inspection, B-scan images provide visual representations of subsurface responses that can be further analyzed for structural condition assessment. However, practical use of GPR data in SHM still requires reliable and scalable interpretation methods.

Recent advances in computer vision, particularly Convolutional Neural Networks (CNNs), have enabled breakthroughs in automated image analysis across various domains. In GPR-based inspection, B-scan

images provide visual representations of subsurface responses that can be further analyzed for structural condition assessment. However, practical use of GPR data in SHM still requires reliable and scalable interpretation methods.

Building upon these insights, this study systematically compares YOLOv8, YOLO11, and Faster R-CNN for automatic hyperbola detection in GPR B-scan images. The experimental dataset used in this work consists of 512 B-scan images with 6,138 labeled object instances. The dataset was split into 328 training images and 184 validation images. Each model was trained independently, and the final comparison was conducted using the validation set. This comparative analysis aims to identify a practical deep learning model for noisy GPR B-scan interpretation in SHM applications.

Contributions of this paper:

• Systematic comparison of detection models: We conducted a side-by-side evaluation of YOLOv8, YOLO11, and Faster R-CNN for automatic hyperbola detection in GPR B-scan images.
• Dataset aware evaluation: We evaluated the models on a 512-image GPR B-scan dataset containing 6,138 labeled instances with severe class imbalance across noise, hyperbola, and hyperbola line classes.
• Practical SHM interpretation: We analyzed detection accuracy, recall behavior, localization performance, learning stability, and computational efficiency to clarify the practical strengths and limitations of each model for automated GPR-based inspection.

The remainder of this paper is organized as follows. Section II reviews the related literature. Section III presents the proposed methodology. Section IV describes the evaluated models, training settings, and evaluation metrics, and presents the experimental results and discussion. Finally, Section V concludes the paper.

Literature Review

Daniels, 2005 summarized the theoretical and practical foundations of Ground Penetrating Radar (GPR), including its use for subsurface sensing and infrastructure inspection. In GPR-based structural health monitoring, B-scan images are commonly used to examine subsurface responses, and embedded objects, reinforcement, voids, or other localized reflectors often appear as hyperbolic signatures. However, identifying these patterns remains challenging because GPR B-scan images can contain dense reinforcement responses, overlapping reflections, low-contrast targets, and background clutter. These factors make manual interpretation time-consuming and expert-dependent, while also motivating the development of automated detection methods.

Traditional hyperbola detection methods have often relied on signal processing, image processing, and pattern recognition. Maas and Schmalzl, 2013 used pattern recognition to automatically localize reflection hyperbolas in GPR data, while Lei et al., 2019 proposed an automatic hyperbola detection and fitting approach for GPR B-scan images to reduce manual parameter selection and expert intervention. These approaches provide useful foundations for structured hyperbolic responses, but their robustness may decrease when B-scan images include strong noise, irregular shapes, partially visible hyperbolas, or overlapping target signatures.

Recent studies have shown that deep learning can improve automated GPR image interpretation by learning visual features directly from B-scan data. Ozkaya et al., 2020 analyzed GPR B-scan images using deep learning methods and demonstrated the potential of learned features for subsurface interpretation. Barkataki et al., 2021 applied deep learning techniques to classify soil types from GPR B-scans, further showing the

feasibility of CNN-based GPR analysis. Liu et al., 2023 combined B-scan and C-scan GPR images using multiscale feature fusion for pavement crack recognition, indicating that multiple GPR representations can support more effective feature extraction for infrastructure assessment. In addition, Pham and Lefèvre, 2018 adapted Faster R-CNN for detecting buried objects represented by hyperbola reflections in B-scan GPR images and used simulated radargrams to address the limited availability of real GPR training data.

In GPR-related infrastructure inspection, Khedr et al., 2025 compared multiple deep learning models for GPR-based rebar detection and reported that YOLOv8 outperformed Faster R-CNN and YOLOv7. More recent GPR B-scan object detection research has also emphasized the importance of preserving hyperbolic geometric features while reducing computational cost, as shown by the HFL-YOLOv8 model proposed for object detection in GPR B-scan images (Lei et al., 2026).

Although prior studies have demonstrated the usefulness of classical hyperbola detection, Faster R-CNN-based GPR detection, and YOLO-based GPR applications, direct comparison of YOLOv8, YOLO11, and Faster R-CNN under the same noisy GPR B-scan hyperbola detection setting remains limited. Therefore, this study compares these three object detection models using the same dataset, class definitions, training-validation split, and evaluation metrics. The goal is to clarify the practical strengths and limitations of one-stage and two-stage detectors for automated hyperbola detection in GPR-based structural health monitoring..

Methodology

The dataset consisted of 512 GPR B-scan images with 6,138 labeled object instances across three classes: noise, hyperbola, and hyperbola line. The dataset was divided at the image level into 328 training images and 184 validation images, corresponding to approximately 65% and 35% of the full dataset, respectively. The split was performed before data augmentation to prevent augmented versions of validation images from appearing in the training set. The same validation set was used for all three models to ensure a consistent comparison. The validation subset was kept relatively large to provide a stable comparison under noisy B-scan conditions and class imbalance. This research evaluates the effectiveness of three deep learning-based object detection models–YOLOv8, YOLO11, and Faster R-CNN–for automatic hyperbola detection in GPR B-scan images. CNN-based object detectors generally consist of feature extraction modules and detection heads that predict object classes and bounding-box coordinates. These components enable the models to learn spatial and hierarchical features from image data.

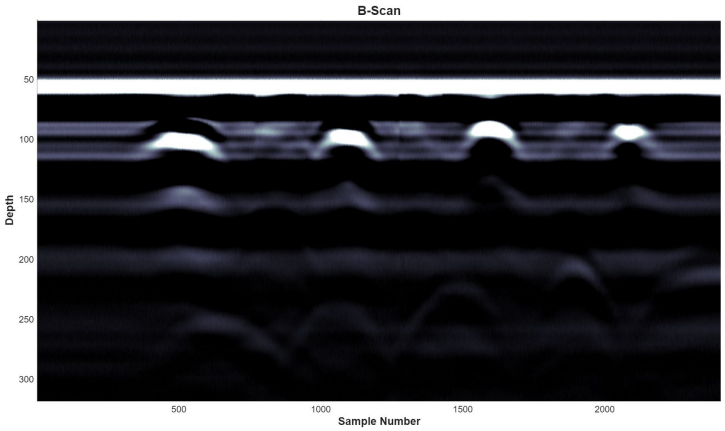


Figure 1: example of GPR B-scan when the object is scanned by GPR.

Figure 1 shows an example of a real GPR B-scan image and illustrates how the image appears when an internal defect is scanned using GPR. For GPR B-scan image analysis, object detection models are particularly well suited to structural health monitoring (SHM) because they can simultaneously localize and classify target responses in B-scan images.

The evaluated methods are categorized into one-stage and two-stage detectors. YOLOv8 and YOLO11 are one-stage detectors from the YOLO family (G. Jocher et al., Ultralytics, 2023). They process an image in a single forward pass and simultaneously perform classification and localization. This design is advantageous for real-time or near-real-time applications. Faster R-CNN (S. Ren et al., 2015) is a representative two-stage detector. It first generates region proposals through a Region Proposal Network (RPN), and then performs object classification and bounding-box refinement. This additional proposal stage can improve detection quality in some tasks, but it generally increases computational cost. The main experimental configuration for the three models is summarized in Table I.

Table I. Model Configuration Used in This Study

Model	Variant/backbone	Input Size	Training
YOLOv8	Large	640×640	100 epochs
YOLO11	Medium	640×640	100 epochs
Faster R-CNN	ResNet50	800×800	15k iterations

The YOLOv8 and YOLO11 models were trained for 100 epochs. The Faster R-CNN model used a ResNet50 backbone and was trained for 15,000 iterations. Learning rate decay was applied to Faster R-CNN at 10,000 and 13,000 iterations to support stable convergence. Data augmentation methods such as rotation, scaling, and translation were used to reduce overfitting and improve robustness under variations in GPR image appearance.

A. Evaluation Metrics

The validation set was used to evaluate all models. Precision, recall, mean Average Precision at IoU 0.50 (mAP50), and mean Average Precision averaged from IoU 0.50 to 0.95 (mAP50–95) were used as the main metrics. For object detection, a prediction is treated as a true positive when its class label is correct and its bounding box sufficiently overlaps with the corresponding ground-truth box based on an IoU threshold.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$IoU = \frac{|B_p \cap B_g|}{|B_p \cup B_g|} \tag{3}$$

$$mAP50 = \frac{1}{C}\sum_{c=1}^{C} AP_c(0.5) \tag{4}$$

$$mAP_{50:95} = \frac{1}{10C}\sum_{t\in\{0.5,0.55,\ldots,0.95\}}\sum_{c=1}^{C} AP_c(t) \tag{5}$$

In (1)–(5), TP , FP , and FN denote True Positives, False Positives, and False Negatives, respectively. B_p and B_g denote the predicted and ground-truth bounding boxes, and C denotes the number of classes. For the Faster R-CNN result, the precision-related value is reported using the AP50-based metric, and recall is reported using average recall (AR), following the evaluation output used for that model.

Experimental Results and Discussion

Table II. Corrected Validation Metrics Across Models

Metric	YOLOv8	YOLO11	Faster R-CNN (15k)
Precision	0.746	0.754	0.468
Recall / AR	0.754	0.770	0.341
mAP50	0.824	0.804	0.468
mAP50-95	0.503	0.474	0.262

Note: Faster R-CNN precision is AP50-based; recall is AR-based.

A. YOLOv8 Performance

As shown in Table II, the YOLOv8 model demonstrated strong detection capability throughout training. After 100 epochs, YOLOv8 achieved the highest overall AP-based detection accuracy among the evaluated models, with an mAP50 of 0.824. The model also reached a precision of 0.746 and a recall of 0.754, indicating a balanced ability to suppress false detections while maintaining reliable target detection. Under the stricter mAP50–95 criterion, YOLOv8 achieved 0.503, showing that the model was not only capable of detecting GPR targets but also able to localize them with reasonable spatial accuracy.

A notable observation from the YOLOv8 results is its strong performance on challenging minority patterns. In particular, the model showed competitive behavior for the hyperbola line class, which is characterized by elongated, thin, and geometrically irregular patterns. This result suggests that YOLOv8 can learn both local texture cues and broader geometric features in GPR B-scan images. Such capability is important for automated bridge inspection, where object-related or anomaly-related responses may appear as subtle and non-uniform hyperbolic patterns.

B. YOLO11 Performance

The YOLO11 model showed competitive detection performance while providing improved computational efficiency. As shown in Table II, YOLO11 achieved a precision of 0.754 and a recall of 0.770, slightly outperforming YOLOv8 in these two metrics. The model obtained an mAP50 of 0.804 and an mAP50–95 of 0.474, which are slightly lower than those of YOLOv8 but still substantially higher than those of Faster R-CNN.

The main advantage of YOLO11 lies in its efficient medium-scale configuration. YOLO11 maintained detection accuracy close to YOLOv8 while requiring less training and inference time. This indicates that YOLO11 provides an effective balance between detection accuracy and computational cost. For practical bridge inspection scenarios, where GPR data may need to be processed rapidly in the field, this efficiency is particularly valuable.

C. Faster R-CNN Performance

The Faster R-CNN model with the ResNet50 backbone was trained for 15,000 iterations, corresponding to approximately 90 epochs. Learning rate decay was applied at 10,000 and 13,000 iterations to support stable convergence. During training, the total loss decreased substantially from 3.19 at the beginning of training to 0.31 at the end, indicating that the model converged successfully.

However, despite this stable convergence, Faster R-CNN showed limited detection performance on the GPR B-scan dataset. As shown in Table II, the model achieved a precision-related AP50 value of 0.468, an AR value of 0.341, an mAP50 of 0.468, and an mAP50–95 of 0.262. Although Faster R-CNN was able to recognize generic noise patterns and major hyperbolic responses, it struggled to detect subtle and elongated hyperbola line features. This result suggests that the evaluated two-stage, anchor-based detection mechanism is less suitable for the irregular and horizontally stretched patterns commonly observed in the GPR B-scan images used in this study.

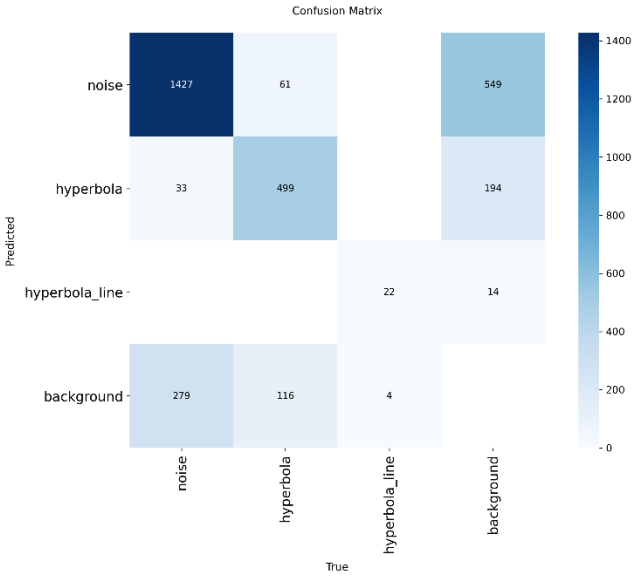


Fig 2. Confusion matrix of YOLOv8 on the validation set. Rows indicate predicted classes and columns indicate true classes.

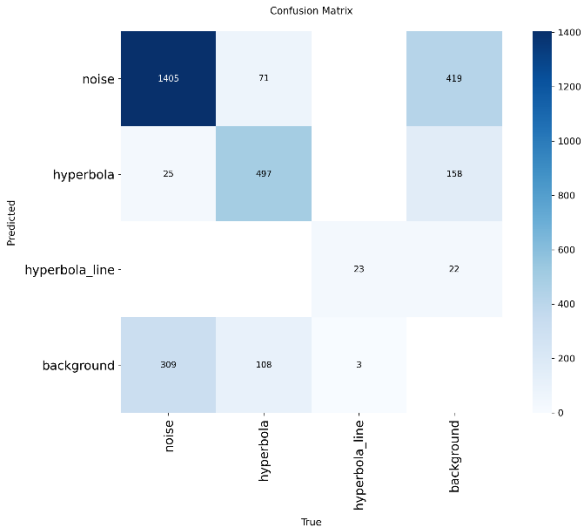


Fig 3. Confusion matrix of YOLO11 on the validation set. Rows indicate predicted classes and columns indicate true classes.

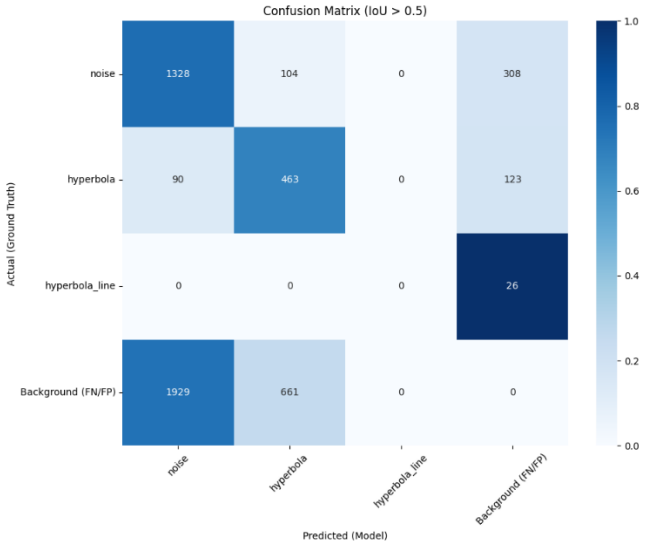


Fig 4. Confusion matrix of Faster R-CNN on the validation set. Rows indicate predicted classes and columns indicate true classes, using the same convention as the YOLO matrices. The background column summarizes false positives and the background row summarizes false negatives.

The confusion matrices in Figures. 2–4 are used to examine the main error patterns of the three detectors. These matrices show class level confusion among noise, hyperbola, hyperbola line, and background predictions, allowing false positives and false negatives to be interpreted in addition to the aggregate mAP values.

The error distributions indicate that many mistakes occur between visually similar high-intensity noise responses and true hyperbolic responses. This behavior is expected in noisy GPR B-scan images because clutter and overlapping reflections of GPR scan can resemble target hyperbolas due to the electromagnetic characteristic (M. M. Riaz et al. 2012**)**. Therefore, although the YOLO- based models achieved high recall and AP performance, additional threshold calibration or post processing may be useful for field deployment where false alarms should be minimized.

Conclusion

This study compared YOLOv8, YOLO11, and Faster R-CNN for automatic hyperbola detection in noisy GPR B-scan images. The experiments were conducted using 512 images and 6,138 labeled instances, divided into 328 training images and 184 validation images. Based on the experimental results, YOLOv8 achieved the strongest AP-based detection performance among the evaluated models, with an mAP50 of 0.824 and an mAP50–95 of 0.503. These results indicate that YOLOv8 provided the best overall detection and localization accuracy on the validation dataset.

YOLO11 also showed competitive performance, achieving the highest precision and recall values of 0.754 and 0.770, respectively. Although its mAP50 of 0.804 and mAP50–95 of 0.474 were slightly lower than those of YOLOv8, YOLO11 still substantially outperformed Faster R-CNN in AP-based evaluation. Therefore, YOLO11 can be considered a practical alternative when high precision, high recall, and efficient model performance are prioritized.
In contrast, Faster R-CNN with a ResNet-50 backbone showed stable training convergence but achieved lower validation performance than both YOLO-based models. Faster R-CNN obtained an mAP50 of 0.468

and an mAP50–95 of 0.262, indicating that it was less effective for the noisy, elongated, and geometrically irregular hyperbola-related patterns in the GPR B-scan dataset used in this study.

Overall, YOLOv8 is recommended when AP-based detection and localization accuracy are the primary requirements, while YOLO11 is a strong alternative when precision, recall, and computational efficiency are important. The results suggest that YOLO-based one-stage detectors are more suitable than the evaluated Faster R-CNN baseline for automated GPR B-scan hyperbola detection in structural health monitoring applications. Future work will include larger and more diverse GPR datasets, independent test-set evaluation, and additional post-processing strategies to reduce false detections caused by clutter and overlapping reflections.

References

D. J. Daniels, "Ground Penetrating Radar," in Encyclopedia of RF and Microwave Engineering, Wiley, 2005.

U. Ozkaya, F. Melgani, M. B. Bejiga, L. Seyfi, and M. Donelli, "GPR B scan image analysis with deep learning methods," Measurement, vol. 165, p. 107770, 2020.

C. Maas and J. Schmalzl, "Using pattern recognition to automatically localize reflection hyperbolas in data from ground penetrating radar," Computers & Geosciences, vol. 58, pp. 116–125, 2013.

W. Lei, F. Hou, J. Xi, Q. Tan, M. Xu, X. Jiang, and Q. Gu, "Automatic hyperbola detection and fitting in GPR B scan image," Automation in Construction, vol. 106, p. 102839, 2019.

Z. Liu, X. Gu, J. Chen, D. Wang, Y. Chen, and L. Wang, "Automatic recognition of pavement cracks from combined GPR B-scan and C-scan images using multiscale feature fusion deep neural networks," Automation in Construction, vol. 146, p. 104698, 2023.

N. Barkataki, S. Mazumdar, P. B. D. Singha, J. Kumari, B. Tiru, and U. Sarma, "Classification of soil types from GPR B scans using deep learning techniques," in Proc. International Conference on Recent Trends on Electronics, Information, Communication & Technology, 2021, pp. 840–844.

G. Jocher, A. Chaurasia, and J. Qiu, "YOLO by Ultralytics," 2023. [Online]. Available: https://github.com/ultralytics/ultralytics

Ultralytics, "YOLO11," 2024. [Online]. Available: https://docs.ultralytics.com/models/yolo11/

S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in Advances in Neural Information Processing Systems, vol. 28, 2015.

M.-T. Pham and S. Lefèvre, "Buried Object Detection from B-Scan Ground Penetrating Radar Data Using Faster-RCNN," in Proc. IGARSS 2018, Valencia, Spain, 2018, pp. 6804–6807, doi: 10.1109/IGARSS.2018.8517683.

M. Khedr, M. Metawie, and M. Marzouk, "Integrated ground penetrating radar and deep learning approach for rebar diameter classification in concrete elements," Frontiers of Structural and Civil Engineering, vol. 19, no. 4, pp. 524–540, 2025.

W. Lei, S. Yu, and T. Zhang, "HFL-YOLOv8: A hyperbolic feature-enhanced lightweight network for object detection in ground penetrating radar images," Applied Soft Computing, vol. 188, Art. no. 114403, 2026.

P. Jafary, D. Shojaei, and K. A. Ehinger, "Deep Learning and Geometric Modeling for 3D Reconstruction of Subsurface Utilities from GPR Data," Sensors, vol. 25, no. 20, Art. no. 6414, 2025.

M. M. Riaz and A. Ghafoor, "Information Theoretic Criterion Based Clutter Reduction for Ground Penetrating Radar," Progress In Electromagnetics Research B, vol. 45, pp. 147–164, 2012.