© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1724543

Automated Classification of Subsurface Cracks in GPR B-Scans Using CNN and SVM

PRECIOUS NNAMDI IBE[1], MUBARAK ADEKUNLE JIMOH[2]
[1, 2] M2 MSc Advanced Electronic Systems Engineering, Université Bourgogne Europe, Dijon, France

Abstract- ***Ground Penetrating Radar (GPR) is widely used for non-destructive evaluation of pavements and concrete structures. However, manual interpretation of GPR B-scan images is time-consuming and subjective. This work proposes a machine learning framework combining Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) for automatic classification of subsurface crack sizes. Synthetic GPR data are generated using gprMax under controlled simulation conditions for different material permittivities and crack geometries. The CNN learns spatial features directly from B-scan images, while the SVM serves as a baseline classifier using handcrafted statistical features. Models are trained on discrete crack depths and evaluated on both seen and unseen datasets to assess generalization. Results show that the CNN significantly outperforms the SVM, especially on unseen crack depths, demonstrating its ability to learn physically meaningful features from GPR data.***

Index Terms- ***Convolutional Neural Networks, Crack Detection, Deep Learning, Ground Penetrating Radar, Non-Destructive Testing, Support Vector Machine***

I. INTRODUCTION

Ground Penetrating Radar (GPR) is a non-destructive testing technique widely used for subsurface inspection in civil engineering, particularly for the assessment of pavements and concrete structures. Subsurface cracks can significantly affect structural integrity, but their detection and characterization from GPR B-scan images are still largely based on manual expert interpretation. This process is time-consuming, subjective, and difficult to scale, especially when crack signatures are weak, noisy, or overlapping. As a result, there is a growing interest in automated and data-driven approaches to improve the reliability and efficiency of GPR data interpretation.

Recent advances in machine learning, and in particular deep learning, have shown strong potential for automatic analysis of complex imaging data. However, in the context of GPR-based crack characterization, many approaches either rely on handcrafted features or are evaluated only under limited and highly controlled conditions, with little emphasis on generalization to unseen defect geometries. This raises an important question: can a learning-based model not only classify known crack sizes, but also generalize to physically related but unseen crack depths?

In this work, we propose a machine learning framework for automatic classification of subsurface crack sizes from GPR B-scan images and compare a Convolutional Neural Network (CNN) with a classical Support Vector Machine (SVM) based on handcrafted statistical features. Synthetic GPR data are generated using the gprMax simulator under controlled and physically consistent conditions. The models are trained on discrete crack depths (seen data) and evaluated both on seen and unseen crack depths to explicitly assess their generalization capability. The results show that the CNN significantly outperforms the SVM, particularly on unseen crack depths, demonstrating its ability to learn physically meaningful spatial features from GPR data.

II. METHODOLOGY

Synthetic GPR data were generated using the open-source gprMax simulator, which is based on the finite-difference time-domain (FDTD) method. A two-dimensional model representing a pavement/concrete structure was constructed, consisting of a homogeneous host medium and an embedded vertical air-filled crack. The relative permittivity of the host medium was varied in the range 3.5–5.5 for pavement and 6.5–8.5 for concrete in order to account for material variability. The crack width was fixed at 4 mm, while the crack depth was varied to represent different defect sizes.

A standard GPR antenna configuration was used, and the scattered electromagnetic fields were recorded along a linear scan path above the surface. The total simulation time window was set to 4 ns, which was sufficient to capture the main reflections from the surface and the subsurface crack. For each configuration, the time-domain responses were assembled into B-scan images, representing two-dimensional radar cross-section views of the subsurface.

For the seen (training) dataset, crack depths of 1.0 cm, 3.0 cm, and 6.0 cm were considered, corresponding to small, medium, and large cracks, respectively. For each depth, multiple simulations were performed by varying the material permittivity values, resulting in a dataset composed of several B-scan images per class. To evaluate the generalization capability of the models, an unseen dataset was

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1724543

generated using intermediate crack depths of 1.5 cm, 3.3 cm, and 5.8 cm, which were not included during training. A summary of the datasets used in this study is provided in Table I.

TABLE I
SUMMARY OF THE DATASETS USED IN THIS WORK

Dataset	Classes	Number of Images
Seen (Training/Test)	Small, Medium, Large	189 (63 per class)
Seen (with No-Crack)	No-crack, Small, Medium, Large	252 (63 per class)
Unseen (Test Only)	Small, Medium, Large	189 (63 per class)

```
#title: B-scan
#domain: 1.00 0.24 0.002
#dx_dy_dz: 0.002 0.002 0.002
#time_window: 4e-9

#material: 1.0 0.0 1.0 0.0 air
#material: 3.5 0.0 1.0 0.0 pavement
#material: 6.5 0.0 1.0 0.0 concrete

#waveform: ricker 1 2.6e9 my_ricker
#hertzian_dipole: z 0.20 0.17 0 my_ricker
#rx: 0.25 0.17 0
#src_steps: 0.002 0 0
#rx_steps: 0.002 0 0

#box: 0.00 0.00 0.00   1.00 0.24 0.002   air
#box: 0.00 0.00 0.00   1.00 0.17 0.002   pavement
#box: 0.00 0.00 0.00   1.00 0.10 0.002   concrete

#box: 0.499 0.10 0.00   0.503 0.158 0.002   air

#geometry_view: 0 0 0   1.00 0.24 0.002   0.002 0.002 0.002   geometry n
```

Fig. 1. Example of the gprMax simulation input defining the computational domain, materials, source, and subsurface crack geometry used to generate B-scan data.

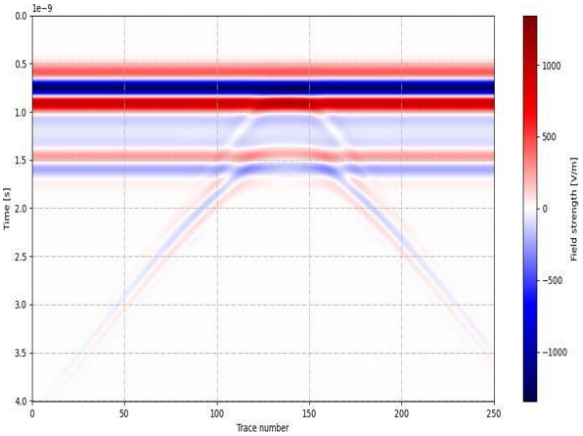


Fig. 2. Example of a simulated GPR B-scan image used as input to the learning models.

All B-scan images were processed using the same pre-processing pipeline. First, each image was cropped to retain only the region of interest containing the main reflections. The cropped images were then resized to a fixed resolution to ensure a consistent input size for the learning algorithms. Finally, per-image normalization was applied to scale the signal amplitudes and reduce the influence of absolute intensity variations between simulations. This pipeline was applied independently to the training, test, and unseen datasets to avoid data leakage and ensure a fair comparison between the models.

III. MACHINE LEARNING MODELS

The CNN model takes preprocessed B-scan images as input and automatically learns spatial features related to crack signatures, such as hyperbolic reflection patterns and amplitude variations. The network consists of multiple convolutional and pooling layers followed by a fully connected classification layer. In contrast, the SVM classifier uses handcrafted global statistical features extracted from each image, including mean, standard deviation, maximum intensity, signal energy, and intensity histograms. These features are standardized and used to train an SVM with a radial basis function (RBF) kernel.

A. Flowchart

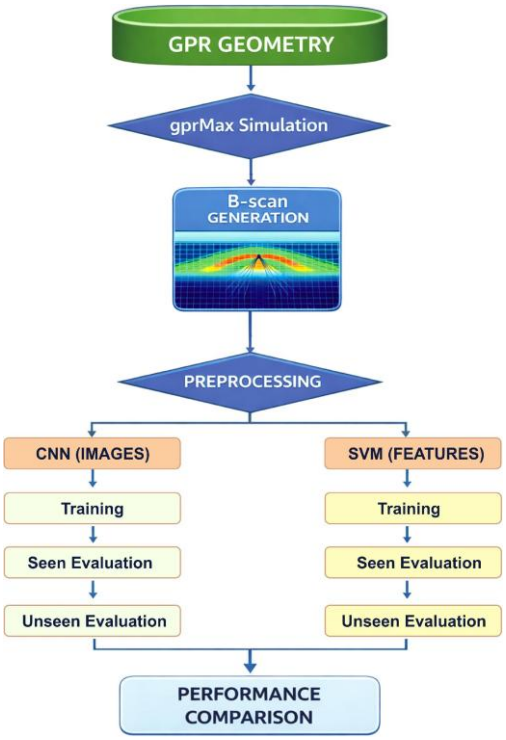


Fig. 3. Overview of the GPR simulation, preprocessing, and learning pipeline for the CNN and SVM models.

IV. EXPERIMENTS AND EVALUATION

The dataset generated from the seen crack depths was divided into training and test sets in order to evaluate the performance of the classifiers under controlled conditions. The training set was used to optimize the parameters of the Convolutional Neural Network (CNN) and the Support Vector Machine (SVM), while the test set was used only for performance evaluation. In addition, a separate unseen

dataset, constructed from intermediate crack depths, was reserved exclusively for assessing the generalization capability of the models.

The CNN was trained directly on the preprocessed B-scan images, whereas the SVM was trained using handcrafted global statistical features. These features included the mean, standard deviation, maximum intensity, signal energy, and intensity histograms, and were standardized before training.

Model performance was evaluated using standard classification metrics, including accuracy, precision, recall, and F1-score. Accuracy measures the overall proportion of correctly classified samples, while precision and recall quantify the reliability and completeness of the predicted classes, respectively. The F1-score provides a balanced measure that combines precision and recall into a single metric. Confusion matrices were also used to analyze class-wise performance and misclassification patterns.

The use of an unseen dataset is particularly important in this context, as it allows the evaluation of the ability of the models to interpolate between known crack depths and to generalize to new subsurface conditions not observed during training. This better reflects practical GPR inspection scenarios, where defect characteristics are not known in advance.

V. RESULTS AND DISCUSSION

The performance of the proposed Convolutional Neural Network (CNN) and the Support Vector Machine (SVM) classifiers was evaluated on both the seen and unseen datasets using accuracy, precision, recall, and F1-score, as well as confusion matrix analysis. These metrics provide complementary insights into the overall classification performance and the class-wise behavior of each model.

On the seen dataset, both models achieve high classification performance under controlled conditions, as the test samples follow the same distribution as the training data. The CNN results are shown in Table II, while the corresponding SVM results are presented in Table III. The CNN achieves near-perfect classification with minimal confusion between the small, medium, and large crack classes, indicating that it is able to learn highly discriminative spatial features directly from the B-scan images. The SVM also performs well on the seen dataset; however, its confusion matrix reveals a higher level of misclassification between neighboring crack classes, particularly between small and medium cracks. This behavior can be attributed to the use of handcrafted global statistical features, which are less effective at capturing subtle spatial and morphological differences in the radar signatures.

A more critical evaluation is provided by the results on the unseen dataset, which assess the ability of the models to generalize to crack depths not observed during training. In this case, the CNN maintains a high level of performance, demonstrating robust generalization and effective interpolation between learned crack representations. In contrast, the SVM exhibits a more pronounced performance degradation, with increased confusion between classes, especially for cracks of intermediate depth. This confirms that, while the SVM can perform adequately under controlled conditions, its reliance on handcrafted features limits its ability to generalize to new subsurface configurations.

In addition to the three-class scenario, the models were also evaluated on a more challenging four-class problem including a “No-crack” class, as shown in Tables IV and V. The CNN achieves perfect classification performance across all four classes, including the no-crack case, indicating that it can reliably distinguish between intact structures and different crack sizes under the simulated conditions. This suggests that the CNN has learned discriminative spatial features that capture both the presence and the morphology of crack-related reflections in the B-scan images.

In contrast, the SVM shows a significant drop in performance in the four-class scenario, with notable confusion between the no-crack and small crack classes. This behavior can be explained by the use of handcrafted global statistical features, which are often similar for weak crack signatures and no-crack cases, leading to overlapping feature distributions. While the SVM still classifies large cracks accurately due to their stronger and more distinctive signatures, it struggles to reliably separate subtle classes, such as no-crack and small cracks. These results further highlight the advantage of the CNN in learning spatially meaningful features directly from the B-scan images, especially in more realistic and challenging classification settings.

Overall, these results highlight the advantage of deep learning-based approaches for GPR data interpretation. By learning hierarchical and spatially meaningful features directly from B-scan images, the CNN is better suited to capture the complex reflection patterns associated with subsurface cracks. This leads to superior performance and improved robustness compared to the classical SVM approach, particularly in scenarios that require generalization to previously unseen conditions, which are common in practical GPR inspection applications.

TABLE II
DETAILED CLASSIFICATION PERFORMANCE OF THE CNN ON THE SEEN DATASET (3 CLASSES)

Class	Precision	Recall	F1-score	Support
Small	1.00	1.00	1.00	12
Medium	1.00	1.00	1.00	13
Large	1.00	1.00	1.00	13
Accuracy		1.00		38
Macro Avg	1.00	1.00	1.00	38
Weighted Avg	1.00	1.00	1.00	38

TABLE III
DETAILED CLASSIFICATION PERFORMANCE OF THE SVM ON THE SEEN DATASET (3 CLASSES)

Class	Precision	Recall	F1-score	Support
Small	0.62	0.67	0.64	12
Medium	0.69	0.69	0.69	13
Large	1.00	0.92	0.96	13
Accuracy		0.76		38
Macro Avg	0.77	0.76	0.76	38
Weighted Avg	0.77	0.76	0.77	38

TABLE IV
DETAILED CLASSIFICATION PERFORMANCE OF THE CNN ON THE SEEN DATASET (4 CLASSES)

Class	Precision	Recall	F1-score	Support
No-crack	1.00	1.00	1.00	13
Small	1.00	1.00	1.00	12
Medium	1.00	1.00	1.00	13
Large	1.00	1.00	1.00	13
Accuracy		1.00		51
Macro Avg	1.00	1.00	1.00	51
Weighted Avg	1.00	1.00	1.00	51

TABLE V
DETAILED CLASSIFICATION PERFORMANCE OF THE SVM ON THE SEEN DATASET (4 CLASSES)

Class	Precision	Recall	F1-score	Support
No-crack	0.50	0.46	0.48	13
Small	0.39	0.58	0.47	12
Medium	0.75	0.46	0.57	13
Large	1.00	1.00	1.00	13
Accuracy		0.63		51
Macro Avg	0.66	0.63	0.63	51
Weighted Avg	0.67	0.63	0.63	51

TABLE VI
DETAILED CLASSIFICATION PERFORMANCE OF THE CNN ON THE UNSEEN DATASET (3 CLASSES)

Class	Precision	Recall	F1-score	Support
Small	1.00	0.97	0.98	63
Medium	0.97	1.00	0.98	63
Large	1.00	1.00	1.00	63
Accuracy		0.99		189
Macro Avg	0.99	0.99	0.99	189
Weighted Avg	0.99	0.99	0.99	189

TABLE VII
DETAILED CLASSIFICATION PERFORMANCE OF THE SVM ON THE UNSEEN DATASET (3 CLASSES)

Class	Precision	Recall	F1-score	Support
Small	0.83	0.46	0.59	63
Medium	0.59	0.92	0.72	63
Large	1.00	0.87	0.93	63
Accuracy		0.75		189
Macro Avg	0.80	0.75	0.75	189
Weighted Avg	0.80	0.75	0.75	189

VI. CONCLUSION AND FUTURE WORK

This work demonstrated the effectiveness of CNN-based models for automatic classification of subsurface crack sizes from GPR B-scan images. Compared to the SVM approach, the CNN showed superior performance and better generalization to unseen crack depths. Future work will focus on validation using real GPR data, more complex simulation scenarios, and deeper network architectures.

ACKNOWLEDGMENT

The authors would like to thank Prof. Amine Ihamouten for his guidance, support, and valuable discussions throughout this work.

REFERENCES

[1] A. Giannopoulos, “Modelling ground penetrating radar by gprMax,” *Construction and Building Materials*, 2005. DOI: 10.1016/j.conbuildmat.2005.06.007

[2] A. Giannopoulos et al., “gprMax: Open source software to simulate electromagnetic wave propagation for GPR,” *Computer Physics Communications*, 2016. DOI: 10.1016/j.cpc.2016.08.020

[3] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, 1995. DOI: 10.1007/BF00994018

[4] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, 2015. DOI: 10.1038/nature14539

[5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. DOI: 10.5555/3086952.

[6] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006. DOI: 10.1007/978-0-387-45528-0.

[7] Y. Ma, Y. Wu, and J. Liu, “Automatic identification of subsurface defects using GPR and deep learning,” *NDT & E International*, 2018. DOI: 10.1016/j.ndteint.2018.05.004