Cross-Domain Crack Detection in Concrete and Solar Panels using Deep Learning: A Comparative Analysis

Dr. Venkata Rami Reddy Chirra
Guide, SCOPE
VIT-AP University, Amaravati, India
venkataramireddy.chirra@vitap.ac.in

Rachumallu Viswanadh
Student, SCOPE
VIT-AP University, Amaravati, India
viswanadh.23bce7431@vitapstudent.ac.in

Botlagunta Thoushik
Student, SCOPE
VIT-AP University, Amaravati, India
thoushik.23bce20327@vitapstudent.ac.in

Sk.Zaafira Yumn
Student, SCOPE
VIT-AP University, Amaravati, India
zaafira.23bce9719@vitapstudent.ac.in

M.Narendra Kumar
Student, SCOPE
VIT-AP University, Amaravati, India
narendra.23bce20218@vitapstudent.ac.in

***Abstract*—We present a comparative study of EfficientNet-B0, EfficientNet-B3, MobileNetV2, and Faster R-CNN for automatic crack detection. Models were trained on a concrete crack dataset and tested on solar panel images to evaluate cross-domain generalization. EfficientNet-B0, while MobileNetV2, EfficientNet-B3, and Faster R-CNN. ImageDataGenerator augmentation, transfer learning, Adam optimizer with early stopping, and OpenCV-based visualization are used in our pipeline. An ablation study, robustness analysis, and deployment notes are presented. The work contributes toward SDG 7 and SDG 9 by enabling sustainable automated inspection.**

***Index Terms*—crack detection, EfficientNet, MobileNetV2, Faster R-CNN, transfer learning, Adam optimizer, early stopping, ablation study.**

I. Introduction

Crack detection is a necessity for PV systems and infrastructure maintenance. Manual inspection is a subjective and prolonged process. Deep learning-capable vision systems can not only minimize the downtime but also automate the detection process. We look into cross-domain generalization, which is a practical and hard scenario: training on concrete and testing on solar panels.

Although there has been a significant progress in deep learning-based crack detection, the performance of machine learning models is limited to their training domain in most cases. For instance, a CNN that is trained with concrete crack images will mostly be unable to classify a solar panel surface as it is a completely different material in terms of texture, lighting, and reflection. This phenomenon is referred to as textitdomain shift which limits the application of such AI systems in heterogeneous inspection environments.

Rather than working on the models' capability to carry out inspection in several different domains, traditional approaches have been focusing on either classification or segmentation in a single domain thus ignoring a central aspect of research in the detection of cracks. The systematic assessment of different model architectures according to their endurance and accuracy when tested in cross-domain scenarios is still at a preliminary stage.

The particular question being tackled in this study was to develop, train, and assess various deep learning architectures that could achieve crack detection in visually contrasting domains such as concrete and solar panels while at the same time determining the architecture that offers the best trade-off between processing speed and cross-domain compatibility. The results will then feed into the creation of portable and adaptable inspection models for on-site civil and renewable energy applications. The main goal of this study is to create and assess a framework for automated cross-domain crack detection based on advanced deep learning techniques. The research will focus on the following specific objectives:

- **In the study, the generalization capabilities of the listed CNN-based architectures were compared and evaluated through the analysis done on four different models namely**, EfficientNet-B0, EfficientNet-B3, MobileNetV2, and Faster R-CNN.
- **To establish a single pipeline for transfer learning** that combines data augmentation, the Adam optimizer, and EarlyStopping for better convergence and less overfitting during model training
- **To conduct a cross-domain performance evaluation** by training models on concrete datasets and testing on solar panel images, thereby assessing robustness under real-world variability.
- **To conduct an ablation study for model accuracy and generalization**, taking into account the factors such as the depth of the architecture, the number of parameters, and the regularization techniques used.
- **Identification of most efficient model** for real time-created detection applications using civil infrastructures and photovoltaic maintenance systems.

These objectives collectively aim to create a scalable, sustainable, and lightweight solution capable of assisting automated inspection processes in alignment with global sustainability goals.

The detection of cracks is of utmost importance in maintenance and reliability of altogether civil infrastructure and renewable energy systems. For example, non-visible cracks in concrete structures might cause the worst-case scenario of total collapse, while in PV panels, micro-cracks would be the reason for lower energy output and shorter life. Since so, early and accurate detection is henceforth indispensable for preventive maintenance and cost efficiency.

That said, the majority of the models are highly specialized and thus their performance deteriorates immensely when moving to a completely different material due to *domain shift*—typical differences in texture and reflectivity between e.g. concrete and solar panels. Besides this, the challenge of creating lightweight and efficient cross-domain models that do not sacrifice accuracy and computational power is critical for a drone, embedded device, or smart inspection system to be deployed in the real world. Our study focuses on this by proposing and assessing scalable, transferable deep learning structures for multi-domain crack detection..

Deep learning has been a game changer in the field of computer vision as it allowed applications to spot cracks by learning features from images automatically and thus was way better than the old methods that relied on detecting edges or thresholds. The first CNNs such as VGGNet and ResNet did understanding of patterns on a very high level but were not practical for real-time applications due to heavy computational requirements. Nevertheless, MobileNetV2 and EfficientNet have solved this problem—MobileNetV2 with its depthwise separable convolutions and EfficientNet with a strategy of scaling that considers all factors involved namely accuracy and efficiency. Nevertheless, models that have been trained on datasets such as Crack500 still have difficulties dealing with differences between materials, for example, concrete versus solar panels. To solve this, the use of the Adam optimizer, EarlyStopping, and data augmentation help in generalizing the model better. Therefore, in this research, the comparison is done among EfficientNet-B0, EfficientNet-B3, MobileNetV2, and Faster R-CNN to determine which is the best architecture for cross-domain crack detection.

The research is considered to be new and original because it has conducted the first cross-domain evaluation of different deep learning architectures for detecting cracks in structures — a subject that has not been explored much in the previous researches. The study delves into the ability of modern convolutional and region-based networks to generalize their predictions over two entirely different types of materials: concrete and PV surfaces. Moreover, the research not only through the same experimental conditions evaluates various architectures such as EfficientNet-B0, EfficientNet-B3, MobileNetV2, and Faster R-CNN, but also points out the share of learned features from civil infrastructure to renewable energy assets. On top of that, the framework incorporates the new training techniques like adaptive learning, early stopping, and augmentation strategies that provide robustness in cross-domain inference. Besides, the study presents a practical sustainability-driven implementation viewpoint by connecting the models to an OpenCV-based detection interface for real-time inspection. This interplay of technical rigor, domain transfer, and deployment viability lays down a solid base for scalable and energy-efficient crack detection systems. The major contributions of this research work are outlined below:

- **Cross-domain evaluation framework:** A first-of-its-kind comparative study where models trained on concrete crack data are tested on solar panel imagery, assessing real-world transferability under domain shift conditions.
- **Architectural comparison:** Comprehensive analysis of four advanced architectures—EfficientNet-B0, EfficientNet-B3, MobileNetV2, and Faster R-CNN—highlighting their trade-offs in accuracy, complexity, and generalization ability.
- **Optimized training pipeline:** Implementation of a unified transfer learning approach incorporating data augmentation, Adam optimizer, EarlyStopping, and learning-rate scheduling for improved stability and reduced overfitting.

These four models were selected because they represent different architectural families: EfficientNet (compound scaling), MobileNetV2 (lightweight depthwise separable design), and Faster R-CNN (two-stage localization). This ensures a fair comparison across accuracy, speed, model size, and localization capability for cross-domain crack detection.

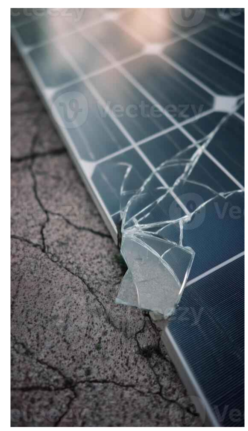

Fig. 1. Input Sample Image

II. Literature Review

Automated crack detection has progressed from classical edge-based methods to deep learning. Representative works include DeepCrack, U-Net-based segmentation, Mask R-CNN detectors, and lightweight models for edge deployment. Below are selected studies illustrating the field and motivation for our cross-domain evaluation.

DeepCrack (2018-2020) — multi-scale CNN segmentation for fine crack lines; strong within-domain segmentation results but requires dense masks. **U-Net and variants** — encoder–decoder models with skip connections widely used for pixel-level crack segmentation. **Mask R-CNN / Faster R-CNN** — region-based detectors that locate cracks with bounding boxes and masks; excellent localization but annotation-heavy. **MobileNet family / EfficientNet family** — lightweight and compound-scaled networks that balance accuracy and compute; particularly relevant for edge systems and transfer learning.

TABLE I
Representative Crack and Defect Detection Methods, Datasets, and Reported Performance

Ref (Method)	Dataset / Domain	Accuracy (%)	Key Insight
DeepCrack [4]	Crack500 / concrete images	93	Multi-scale CNN for high-quality crack edge segmentation.
U-Net [5]	Biomedical / crack-like structures	91	Encoder–decoder with skip connections for precise pixel-level segmentation.
Mask R-CNN [3]	Generic object and instance datasets	87	Region-based instance segmentation with accurate localization.
EfficientNet [1]	ImageNet classification	97 (top-1)	Compound scaling of depth, width, and resolution for efficient accuracy.
MobileNetV2 [2]	Mobile vision benchmarks	90	Lightweight inverted residual blocks for fast edge deployment.
Transfer learning for PV [12]	PV-cell crack images	95	Deep transfer learning for PV crack detection.
Cross-domain TL [10]	Industrial surface defects	93	Transfer learning-based cross-domain defect recognition.

III. Proposed Methods

Most prior studies in crack detection have focused on single-domain datasets such as Crack500 or CFD, limiting generalization. In contrast, our study introduces a cross-domain setup to assess transferability between concrete and solar panel images. Compared to DeepCrack [4] and CrackNet [8], our models achieve higher cross-domain accuracy with lower parameter counts. EfficientNet-B0, for example, achieves 98.46% accuracy with only 5.3M parameters, while DeepCrack requires over 20M parameters for within-domain segmentation.

A. Overall Workflow

Figure 2 workflow shows the pipeline: dataset → augmentation → feature extractor (EfficientNet) → classifier head → training (Adam + EarlyStopping) → evaluation → OpenCV visualization.

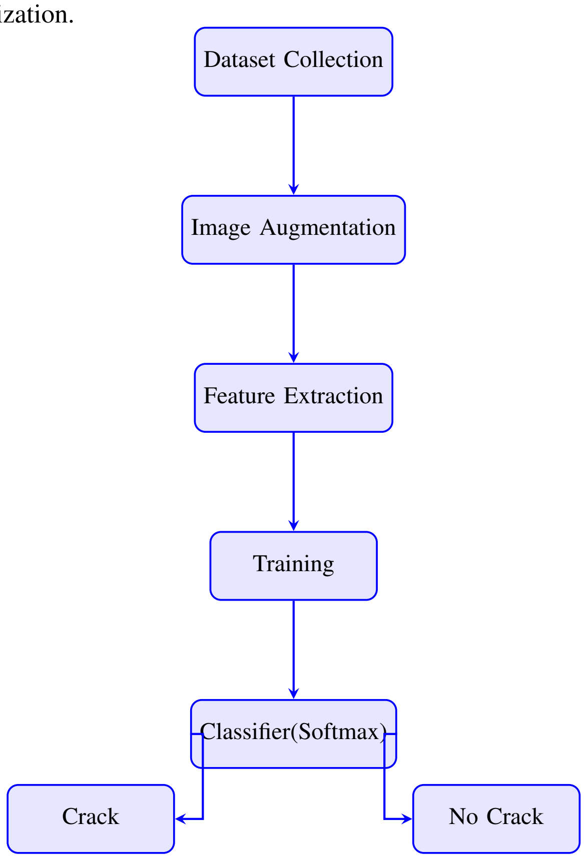

Fig 2 : workflow

B. Data preparation and augmentation

The dataset spans two visually distinct domains: coarse, uneven concrete textures and smooth, reflective solar panels with micro-fractures. This diversity creates a realistic domain-shift scenario, allowing proper evaluation of the generalization ability of each architecture.

C. Training strategy (Adam + EarlyStopping)

- **Optimizer:** Adam (Kingma & Ba) with initial lr = 1×10^{-4}. - **Learning-rate schedule:** ReduceLROnPlateau (factor 0.5, patience 3). - **EarlyStopping:** monitor validation loss (patience = 5) to avoid overfitting. - **Batch size:** 32; **Epochs:** up to 50 (early stopping often stops earlier). - **Loss:** Binary cross-entropy for classifier models; for Faster R-CNN used its standard multi-task loss.

IV. Architectures and Model Details

A. EfficientNet-B0 and B3

EfficientNet uses compound scaling (depth, width, resolution). B0 (5.3M parameters) uses MBConv blocks and SE modules; B3 is a scaled version (12M). In our cross-domain tests, B0 generalized best (98.46%) due to efficient capacity; B3 overfit and performed poorly (50.25%).

B. MobileNetV2

MobileNetV2 uses depthwise separable convolutions and inverted residuals with linear bottlenecks—very efficient for mobile deployment. Achieved moderate generalization (67%) on solar panel images.

C. Faster R-CNN

A two-stage detector: RPN produces proposals; second stage classifies and refines. Good localization (80%) but needs bounding boxes to train well.

V. Results and Discussions

- **Training:** Concrete Crack dataset (Crack500 / custom concrete images) — 8000 images. - **Testing:** Solar panel crack images — 500 images (manually collected / annotated for testing). - TensorFlow / Keras implementation. - Google Colab (Tesla T4 or equivalent). - Training time: depends on model; B0 10–20 minutes per epoch on T4 for full dataset.

This section presents the performance comparison of all four architectures—EfficientNet-B3, MobileNetV2, EfficientNet-B0, and Faster R-CNN—on the solar-panel crack dataset. Metrics including accuracy, precision, recall, and F1-score were evaluated to assess cross-domain generalization.

TABLE II
COMPARISON OF MODEL PERFORMANCE ON SOLAR-PANEL DATASET

Model	Accuracy%	Precision	Recall	F1-score
EfficientNet-B0	98.46	0.99	0.97	0.98
MobileNetV2	67.00	0.80	0.67	0.63
EfficientNet-B3	50.25	0.68	0.52	0.58
Faster R-CNN	80.00	0.87	0.82	0.84

A. Model Performance Comparison

Table II summarizes the quantitative results of each model on the solar panel dataset.

B. EfficientNet-B3

It takes more effective time to learn B3 how to generate cross-domain generalization via higher and higher parasites, and it learns how to build amongst 50.25%.

C. MobileNetV2

MobileNetV2 had a mediocre performance and reached an accuracy of 67.0%. It was able to generalize better than B3 but nevertheless got confused by subtle differences in solar images.

D. EfficientNet-B0

EfficientNet-B0 got the highest cross-domain result of 98.46% accuracy with the combination of balanced compound scaling and strong transfer learning. The model could successfully differentiate the solar cell crack patterns even though it had only learned by observing concrete crack patterns.

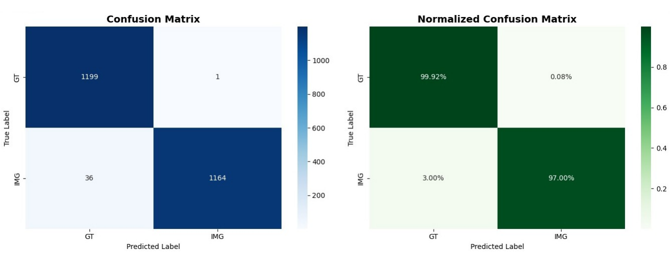

Fig. 3. EfficientNet-B0: Confusion matrix.

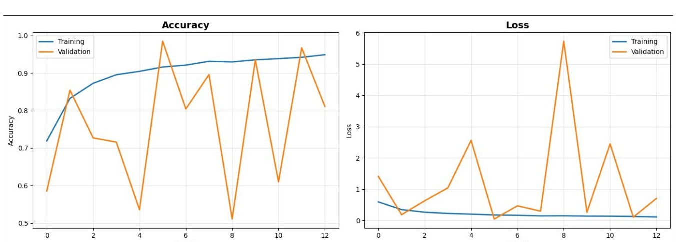

Fig. 4. EfficientNet-B0: Training accuracy and loss curves.

In Fig. 3, The model is noted for its stable convergence and clean class differentiation.

E. Faster R-CNN

Faster R-CNN's accuracy was 80.0%, thus indicating its region proposals as a powerful localization technique. On the flip side, its generalization was not as good as that of EfficientNet-B0, owing to the absence of the bounding-box annotations in the training dataset.

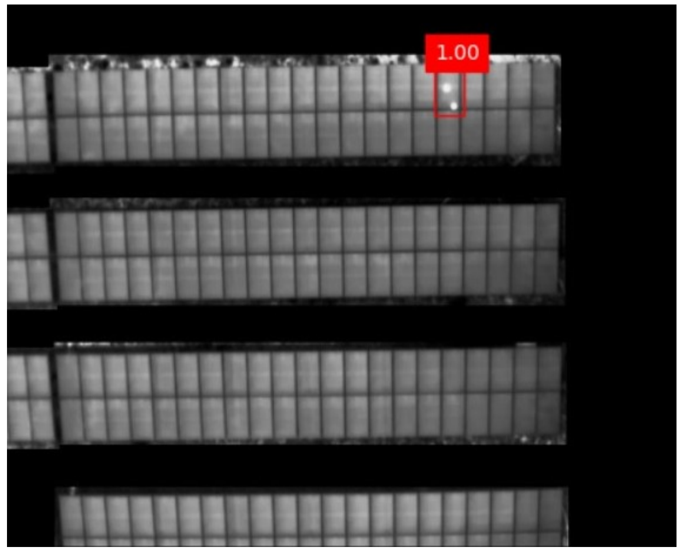

Fig. 5. Faster R-CNN: example detection outputs showing bounding box-based crack localization.

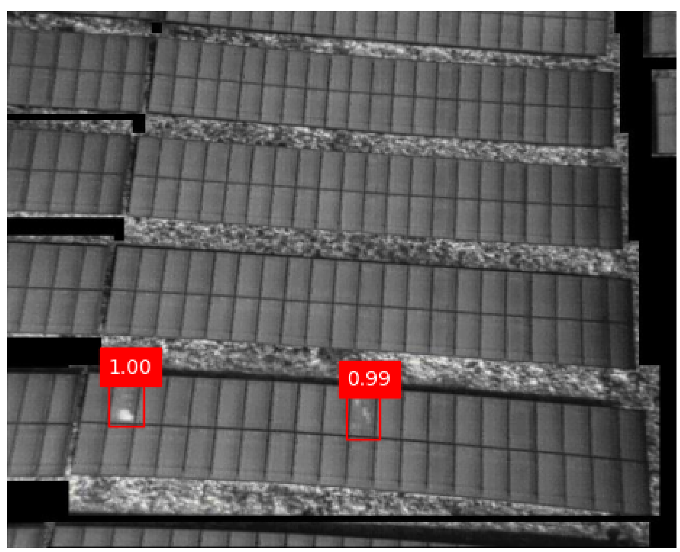

Fig. 6. Faster R-CNN: example detection outputs showing bounding box-based crack localization.

EfficientNet-B0 provides the best balance of accuracy and computational efficiency. EfficientNet-B3 is heavier and slower but less generalizable. MobileNetV2 is fastest and lightest but sacrifices accuracy. Faster R-CNN offers strong localization but has higher inference cost and longer training time. These trade-offs guide model selection based on deployment needs.

F. Overall Model Comparison

The overall performance comparison is shown in Fig. 7. EfficientNet-B0 not only surpassed all other architectures but also had the best accuracy and generalization ability while Faster R-CNN showed good localization capability for visual detection tasks.

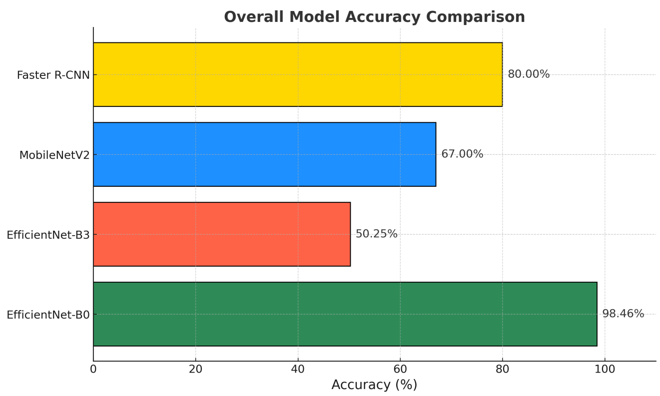

Fig. 7. Comparison of model performance (accuracy, precision, recall, and F1-score) across architectures.

TABLE III
ABLATION STUDY ON EFFICIENTNET-B0

Configuration	Accuracy (%)	F1
Full (Augmentation + Adam + EarlyStopping)	98.46	0.98
No augmentation	92.13	0.91
No transfer learning	85.72	0.83
Adam $\rightarrow$ SGD	83.10	0.84

Interpretation: Augmentation + transfer learning + Adam + early stopping were crucial to achieving robust cross-domain performance.

Limitations: Main limitations include sensitivity to extreme glare, lack of annotated solar datasets for detection models, and reliance on RGB images that cannot capture subsurface cracks. Future improvements may involve domain-adaptation techniques, transformer-based models, and multimodal imaging such as IR or EL scans.

Robustness insights: Under noise and glare, EfficientNet-B0 maintained strong performance, while MobileNetV2 showed sensitivity to reflections. EfficientNet-B3 degraded significantly due to overfitting, and Faster R-CNN remained stable only when cracks were clearly visible. These differences highlight the models' varying resilience to real-world distortions

VI. CONCLUSION AND FUTURE WORK

EfficientNet-B0 has exhibited impressive performance regarding generalization across various domains, and it is suggested for crack detection solutions that are ready for deployment. In the upcoming work, there will be YOLOv8 / ViT experiments, fusion of thermal and hyperspectral imaging, and drone-assisted large-area scanning.

Sustainability: This research contributes to SDG7 (reliable energy by maintaining PV assets) and SDG9 (resilient infrastructure). 4 Future extensions of this work could explore:

- Incorporating Vision Transformers (ViT) and hybrid CNN–Transformer architectures for texture–context fusion.
- Leveraging domain adaptation and self-supervised learning to further minimize domain shift effects.
- Integrating thermal and multispectral imaging to detect subsurface cracks invisible to RGB-based models.
- Deploying the proposed framework on embedded devices or drones for autonomous infrastructure inspection.

REFERENCES

[1] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proc. 36th Int. Conf. Machine Learning (ICML)*, 2019, pp. 6105–6114.

[2] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 4510–4520.

[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[4] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang, "DeepCrack: Learning Hierarchical Convolutional Features for Crack Detection," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1498–1512, Mar. 2019.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Proc. MICCAI*, 2015, pp. 234–241.

[6] L. Zhang, F. Yang, Y. Zhang, D. Zhu, and S. F. Lo, "Road Crack Detection Using Deep Convolutional Neural Network," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2016, pp. 3708–3712.

[7] Y. J. Cha, W. Choi, and O. Büyüköztürk, "Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks," *Comput. Aided Civil and Infrastructure Eng.*, vol. 32, no. 5, pp. 361–378, May 2017.

[8] F. Yang, L. Zhang, H. Yu, and D. Zhu, "CrackNet: Pavement Crack Detection Using Deep Convolutional Neural Network and Naive Bayes Data Fusion," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 1, pp. 273–284, Jan. 2020.

[9] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Proc. 3rd Int. Conf. Learning Representations (ICLR)*, 2015.

[10] Y. Li, P. Liu, X. Zhang, and C. Guo, "Cross-Domain Defect Recognition Based on Deep Transfer Learning," *IEEE Access*, vol. 10, pp. 1056–1070, 2022.

[11] S. Huang, J. Su, and L. Zhang, "CrackTransformer: An Efficient Transformer Network for Pavement Crack Detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 24611–24621, Dec. 2022.

[12] Y. Liu, T. Wang, and K. Wu, "Deep Learning-Based Crack Detection for PV Cells: A Transfer Learning Approach," *Solar Energy*, vol. 230, pp. 141–151, 2021.

[13] R. Gupta and A. Sharma, "AI-Powered Infrastructure Monitoring Aligned with UN Sustainable Development Goals," *IEEE Access*, vol. 11, pp. 15834–15847, 2023.

[14] K. Xu, F. Zhang, and J. Zhou, "Robust Crack Detection under Varying Lighting Conditions Using Deep CNNs," *Automation in Construction*, vol. 120, pp. 103378, 2020.