International Journal of Basic and Applied Sciences, 14 (4) (2025) 756-762

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/4jfvmz41
Research paper

Inception V3-Based Deep Learning Approach for Crack Detection in Paintings

Dr. Rajendra Vasantrao Patil [1] *, Dr. Govind Mohanlal Poddar [2], Deepak Yashwantrao Bhadane [3], Dr. Rahul Manohar Patil [2], Shravani R. Patil [1], Dr. Sunil Vikram Desale [4], Vishal D. Suryawanshi [1]

[1] Department of Computer Engineering, SSVPS Bapusaheb Shivajirao Deore College of Engineering, Dhule (MS), India
[2] NES Gangamai College of Engineering, Nagaon, Dhule (MS)، India
[3] R. C. Patel College of Engineering and Polytechnic, Shirpur, Dhule (M. S.)، India
[4] Nikam Institute of Technology and Research, Dhule, India
**Corresponding author E-mail: patilrajendra.v@gmail.com*

Received: July 17, 2025, Accepted: August 19, 2025, Published: August 27, 2025

Abstract

This study proposes a deep learning-based methodology for the automatic detection of surface cracks in painting artworks, an essential task in art restoration and preservation. A dataset comprising 700 high-resolution images (350 cracked, 350 non-cracked) was curated and expanded through data augmentation techniques to enhance model generalization. Preprocessing steps included resizing and normalization. Four deep learning models, VGG16, ResNet50, a custom CNN, and InceptionV3, were evaluated. Among them, a fine-tuned InceptionV3 model yielded the best performance, achieving a test accuracy of 98.01% and an F1 score of 0.98. Evaluation metrics such as confusion matrix, accuracy/loss curves, and classification report validated the robustness of the proposed method. Compared to earlier approaches that were restricted to structural domains or limited art-specific studies, the proposed framework demonstrates superior accuracy and broader applicability through transfer learning and advanced regularization. The system offers a scalable, non-invasive solution for digital inspection of artworks and can be extended to diverse painting types, varied mediums, or severity-based crack classification. It also aligns with emerging trends in lightweight, attention-guided deep learning models, supporting real-world deployment in cultural heritage preservation.

Keywords: *Crack Detection; Painting Preservation; Deep Learning; CNN; Inceptionv3; Surface Defects.*

1. Introduction

The conservation of historical paintings is a critical aspect of cultural heritage preservation. One of the most significant indicators of deterioration in paintings is the formation of surface cracks, which can develop over time due to environmental stress, aging materials, or handling damage [1], [2]. Traditionally, cracks have been identified through manual visual inspection by art conservators. This process is time-consuming, labor-intensive, and often subjective, leading to inconsistent assessments, particularly for micro-cracks invisible to the naked eye

With the rise of artificial intelligence and deep learning, automated methods for image-based defect detection have shown promising potential in various industries, including civil engineering, medicine, and manufacturing. Convolutional Neural Networks (CNNs), in particular, have demonstrated remarkable performance in capturing fine-grained visual patterns, making them highly effective for classification and detection tasks involving surface-level anomalies [3]. Despite these advancements, the application of deep learning to crack detection in painting surfaces remains relatively unexplored, with most existing literature focused on structural damage in roads and buildings rather than delicate artistic media [4], [5].

This research addresses this gap by proposing an automated system for detecting cracks in painting surfaces using a fine-tuned InceptionV3 deep learning model. A dataset of 700 high-resolution images (350 cracked and 350 non-cracked) was curated and augmented to improve generalization. All images were resized and normalized for uniform input to the models. The proposed system involves a comparative analysis of four different CNN architectures, namely, a baseline CNN, VGG16, ResNet50, and InceptionV3. Each model was trained under consistent hyperparameters and evaluated on a separate test set using standard performance metrics.

Among the evaluated models, the fine-tuned InceptionV3 architecture demonstrated the highest accuracy (98.01%) and F1-score (0.98), outperforming the others in terms of precision, recall, and overall robustness. The model training incorporated advanced techniques such as Label smoothing, early stopping, and learning rate scheduling to avoid overfitting and ensure stable convergence. The performance was further validated using accuracy and loss plots, confusion matrix analysis, and classification reports. This study not only establishes a reliable framework for crack detection in painting surfaces but also provides a deployable architecture that can be integrated into digital conservation workflows.

The key contributions of this work include: (1) creation and augmentation of a dedicated image dataset for painting crack classification; (2) implementation and fine-tuning of multiple CNN-based models; (3) identification of InceptionV3 as the most effective architecture for this task; (4) comprehensive evaluation using classification metrics and visualizations; and (5) development of a reproducible and scalable deep learning-based crack detection system. The outcomes of this research are expected to assist art curators and conservators in performing efficient, non-intrusive inspection of valuable artwork, ultimately contributing to the preservation of cultural heritage.

2. Literature review

Over the past decade, artificial intelligence (AI) and deep learning technologies have transformed defect detection in various domains such as civil engineering, manufacturing, and medical diagnostics. Traditionally, crack detection relied on classical image processing techniques such as edge detection, morphological operations, and intensity thresholding [1], [2]. However, these methods are highly sensitive to noise, surface variation, and lighting inconsistencies, which are common challenges in painted artworks [32], [34], [35]. Their dependence on handcrafted features limited their robustness and adaptability, making them less reliable for heterogeneous surfaces such as paintings, where micro-cracks and irregular textures are common.

The rise of Convolutional Neural Networks (CNNs) revolutionized feature extraction and classification tasks by enabling automated learning from raw pixel-level data [3]. CNN-based models such as AlexNet, VGG16, ResNet50, and InceptionV3 have become foundational in computer vision because of their hierarchical feature learning capabilities [4 - 7]. Deep residual learning introduced by ResNet helped mitigate vanishing gradient issues, which enabled deeper architectures [5]. In addition, dense connectivity patterns in DenseNet improved gradient flow and feature reuse [9].

The use of encoder-decoder networks such as U-Net [8] and SegNet [4] advanced semantic segmentation tasks and crack boundary localization, especially in biomedical and civil imaging. Furthermore, novel loss functions such as Focal Loss [10] improved learning in imbalanced datasets. This issue is also encountered in crack classification, where non-defect regions often dominate.

Transfer learning has emerged as a powerful tool when working with small datasets, particularly through fine-tuning pretrained models originally trained on ImageNet [6], [7], [11]. Gopalakrishnan et al. [13] successfully applied transfer learning for pavement distress detection using pre-trained CNNs. Likewise, models such as ResNet and Inception have been used for automated crack classification in bridges [1], roads [2], [12], and construction materials [3], [15].

Lightweight deep learning models have also gained traction for real-time inspection tasks, offering efficient inference on edge devices [14]. For example, Kim et al. developed a compact CNN capable of real-time crack detection on embedded systems, which highlights its relevance for mobile-based inspection workflows. This becomes particularly important for conservation tasks in the field, where cloud access may be limited.

Other researchers have explored stereo vision [16], region proposal networks such as Faster R-CNN [22], and adaptive pixel-level systems [23] to enhance detection accuracy and context awareness in defect localization. These techniques, while largely developed for civil and industrial use, offer architectural insights that can be extended to artwork analysis.

Recent advancements have also introduced attention-guided CNNs [17], edge-enhanced learning [19], and weakly supervised frameworks [18] for detecting fine-grained surface anomalies. These techniques are particularly suitable for textured, heterogeneous surfaces such as painted artwork. Despite their success in pavement or metal surface analysis, their application to artwork restoration remains limited. More recent directions further include lightweight networks with dense feature connections and dual attention modules that achieve millimetre-level crack detection at low computational cost [28]. Similarly, resilient lightweight CNNs with coordinate attention have been applied successfully in surface defect detection tasks such as steel inspection [29].

In art preservation, few studies have applied deep learning. For example, Chen et al. [20] employed a VGG–16–based SegNet architecture for surface crack detection, while Song et al. [25] applied curvature and texture features for defect identification. However, these efforts are limited in scope, often constrained by small datasets and a lack of cross-domain generalization. In contrast, transformer-based architectures have recently demonstrated superior performance and generalization for crack detection tasks in masonry and concrete [26], [27], which highlights their potential for heterogeneous surfaces such as paintings.

Furthermore, applications involving UAVs [24], robotic inspection systems, and mobile-based analysis are opening new avenues for remote cultural heritage documentation and real-time conservation assessment. A deep learning model tailored for this context could be integrated into such systems, enabling intelligent and on-the-go analysis.

To address the evident research gap in deep learning-based painting crack detection, this study proposes a comprehensive pipeline involving dataset creation, augmentation, pre-processing, model comparison, and fine-tuning. By employing InceptionV3, known for its multi-scale feature extraction capabilities, this work demonstrates accurate and scalable detection of cracks in paintings, validated through performance metrics and visual analysis. Unlike generic structural applications, the proposed system adapts deep learning methods to the intricate textures and patterns characteristic of artistic surfaces, thereby offering a novel contribution to the intersection of AI and digital heritage conservation.

To bridge this gap, we propose a transfer learning-based approach using InceptionV3, supported by data augmentation and regularization, for high-performance crack detection in artistic surfaces. Furthermore, the comparative evaluation of models such as CNN, VGG16, ResNet50, and InceptionV3 in a unified pipeline offers deeper insights into model suitability for this unique application.

An easy way to comply with the paper formatting requirements is to use this document as a template and simply type your text into it.

3. Methodology

This section outlines the systematic approach adopted for automated crack detection in painting surfaces using deep learning. The methodology involves dataset preparation, image preprocessing, model selection, training configuration, and evaluation using multiple performance metrics. The primary model used in this study is the fine-tuned InceptionV3, and it is benchmarked against other architectures such as VGG16, ResNet50, and a custom CNN [20 - 22].

3.1. Dataset preparation

The dataset used in this study comprises a total of 700 high-resolution RGB images, equally divided into 350 cracked and 350 non-cracked painting surface samples. To enhance generalization and robustness, extensive data augmentation techniques such as rotation, flipping, shearing, shifting, and zooming were applied. This significantly increased the variability of the dataset and helped the model learn crack patterns under different conditions. The final dataset was organized into two labeled folders: 'Cracked' and 'Non-Cracked', to streamline training using image data generators.

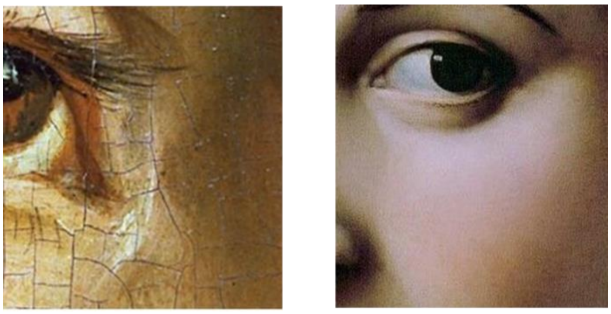

Fig. 1: Representative Painting Surface Images: (A) Cracked Surface Showing Visible Fracture Lines and (B) Non-Cracked Surface with Uniform Texture.

Figure 1 presents representative samples of cracked and non-cracked painting surfaces. The cracked image (Figure 1a) demonstrates visible linear fractures and micro-fissures running across the pigment layers, which often vary in thickness and intensity depending on aging and environmental stress. The non-cracked image (Figure 1b), in contrast, exhibits smooth and uniform texture without significant surface discontinuities. These examples illustrate the challenges of differentiating genuine cracks from natural brush strokes, surface irregularities, or texture variations, which underscores the need for robust deep learning methods

3.2. Data pre-processing

Before feeding images into the neural networks, a series of pre-processing steps was applied to ensure data consistency and optimal model performance [20], [22]:

- Image Resizing: All images were resized to 299×299 pixels, which is the standard input size required for the InceptionV3 architecture.
- Pixel Normalization: The pixel values were scaled to a [0, 1] range by dividing each value by 255. This normalization helps in accelerating convergence during model training.
- Data Splitting: The dataset was divided in a 70:30 ratio, with 70% used for training and 30% for validation and testing. A stratified split ensured class balance across both sets.

3.3. Model architecture and training setup

Four deep learning models were evaluated: a custom CNN, VGG16, ResNet50, and InceptionV3. The proposed InceptionV3 model was initialized with pretrained ImageNet weights, followed by custom classification layers including GlobalAveragePooling2D, dense layers with ReLU activation, dropout for regularization, and a sigmoid output layer. The model was compiled using Binary Crossentropy loss with label smoothing, Adam optimizer, and trained for 30 epochs with batch size 16.

Algorithm 1: Crack Detection using Fine-Tuned InceptionV3

Input: Labeled image dataset with cracked and non-cracked classes

Output: Trained a deep learning model for binary classification

1) Load and preprocess dataset (resize, normalize).
2) Apply data augmentation to increase sample variability.
3) Initialize InceptionV3 model with pretrained ImageNet weights.
4) Replace top layers with custom dense layers and sigmoid activation.
5) Compile model using Adam optimizer and Binary Crossentropy loss.
6) Train the model using early stopping and learning rate reduction callbacks.
7) Evaluate the model using accuracy, precision, recall, F1-score, and confusion matrix.
8) Save the trained model and performance reports.

3.4. Workflow diagram

Figure 2 illustrates the complete workflow pipeline, beginning with data collection and labeling, followed by preprocessing steps such as resizing and normalization. The workflow proceeds with data augmentation and training using different CNN models. The final stage involves model evaluation using accuracy, precision, recall, F1-score, and confusion matrix. This diagram provides a high-level overview of the systematic process adopted in this study for automated crack detection.

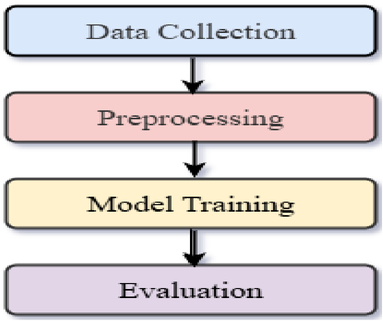

Fig. 2: Workflow Diagram of the Proposed Crack Detection System.

As shown in the workflow diagram, the research begins with systematic data collection and labelling, followed by pre-processing steps that ensure consistency in input images. The processed dataset is then used to train deep learning models, with InceptionV3 fine-tuning forming the core of the proposed system. Finally, performance evaluation is carried out using standard metrics such as accuracy, precision, recall, and a confusion matrix.

3.5. Training configuration

To optimize model performance and avoid overfitting, the following training configurations were applied:

- Loss Function: BinaryCrossentropy with label smoothing (value = 0.1) to reduce overconfidence.
- Optimizer: Adam optimizer with an initial learning rate of 1e-4.
- Batch Size: 16
- Epochs: 30
- Callbacks Used:
- EarlyStopping (patience = 5) to halt training when validation loss stagnates
- ReduceLROnPlateau to decrease learning rate upon plateauing of validation loss

All models were implemented using TensorFlow/Keras and trained on a GPU-enabled environment for accelerated computation.

All hypertext links and section bookmarks will be removed from papers during the processing of papers for publication. If you need to refer to an Internet email address or URL in your paper, you must type out the address or URL fully in Regular font.

3.6. Performance metrics

To comprehensively evaluate the effectiveness of the proposed InceptionV3-based crack detection model, several standard performance metrics were employed. These metrics not only provide insights into classification accuracy but also quantify the model's ability to correctly identify both cracked and non-cracked surfaces. The key metrics used are as follows [20], [21], [30]:

a) Accuracy

Accuracy represents the overall percentage of correct predictions made by the model out of all predictions. It is defined as:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (1)$$

Where:

- TP: True Positives – Cracked surfaces correctly predicted as cracked
- TN: True Negatives – Non-cracked surfaces correctly predicted as non-cracked
- FP: False Positives – Non-cracked surfaces incorrectly predicted as cracked
- FN: False Negatives – Cracked surfaces incorrectly predicted as non-cracked

b) Precision

Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive. It indicates how many selected items are relevant:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

A high precision implies a low false positive rate, which is important in minimizing false alarms in crack detection.

c) Recall (Sensitivity)

Recall quantifies the model's ability to identify all relevant instances of a class. It is the ratio of correctly predicted positives to all actual positives:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

In crack detection, high recall ensures that most actual cracks are detected, which is critical in safety and conservation contexts.

d) F1-Score

The F1-score is the harmonic mean of precision and recall. It is particularly useful when the class distribution is imbalanced:

$$\text{F1} - \text{Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision+Recall}} \quad (4)$$

A balanced F1-score indicates strong overall performance across both metrics.

e) Confusion Matrix

The confusion matrix provides a tabular representation of actual vs. predicted classifications, allowing for a detailed understanding of classification errors. It highlights the number of correct and incorrect predictions across both classes, helping identify model biases or weaknesses.

f) Loss Function Monitoring

In addition to classification metrics, the Binary Cross-Entropy Loss was tracked during training and validation phases. Monitoring the loss curve helps determine the convergence behaviour of the model and detect overfitting or underfitting.

4. Experimental results and discussion

This section presents a comprehensive analysis of the experimental results obtained through training and evaluating four different models, such as Custom CNN, VGG16, ResNet50, and the proposed InceptionV3. All models were trained using the same pre-processed dataset with identical hyperparameters to ensure fairness in comparison. Key performance metrics such as accuracy, precision, recall, F1-score, and loss were used to evaluate classification efficacy.

4.1. Model comparison overview

The performance of all models was evaluated on the same test dataset comprising 210 images (30% of the total dataset). The InceptionV3 model demonstrated superior performance across all evaluation criteria. Table 1 summarizes the classification metrics obtained for each model.

Table 1: Model Accuracy Comparison

Model	Test Accuracy (%)	Precision (%)	Recall (Sensitivity) (%)	F1-Score (%)
Custom CNN	88.57	87.90	88.10	87.98
VGG16	91.14	91.47	90.85	91.16
ResNet50	94.00	94.27	93.81	94.04
InceptionV3	98.01	98.12	97.91	98.01

The InceptionV3 model significantly outperformed all others, achieving a test accuracy of 98.01% and an F1-score of 98.01%, which reflects a well-balanced trade-off between precision and recall. The architecture's multi-scale feature extraction and dimensionality reduction capabilities are likely contributors to this high accuracy.

4.2. Accuracy and loss curves

Figure 3 shows the training and validation accuracy curves of the fine-tuned InceptionV3 model. Both curves demonstrate a steady upward trend, with validation accuracy closely following training accuracy. The convergence stabilizes around the 10th epoch, reaching approximately 98 percent, and no significant gap is observed between the two curves. This indicates that the model generalizes effectively to unseen data and does not exhibit overfitting.

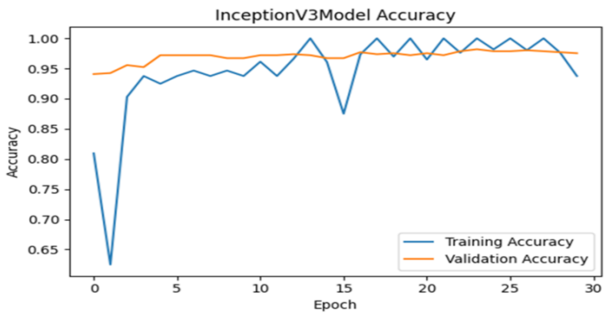

Fig. 3: Training and Validation Accuracy Curves of the Inceptionv3 Model.

Figure 4 presents the training and validation loss curves of the InceptionV3 model. Both curves steadily decrease during training, which confirms that the model successfully learned discriminative features. The curves flatten toward the final epochs, demonstrating convergence and training stability. The close alignment between training and validation loss further supports the robustness of the model across different dataset partitions.

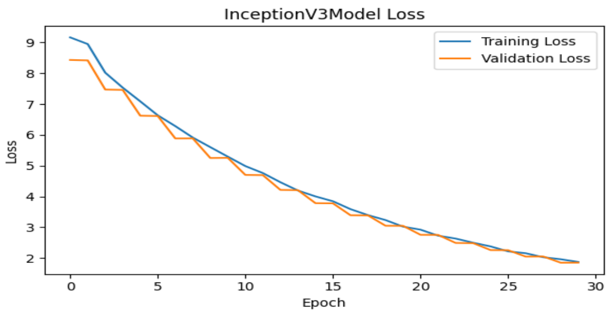

Fig. 4: Training and Validation Loss Curves of the Inceptionv3 Model.

These plots confirm that the model effectively learned relevant crack features without overfitting. The close alignment of training and validation curves demonstrates stable convergence and strong generalization to unseen data.

4.3. Confusion matrix analysis

Figure 5 shows the confusion matrix of the InceptionV3 model on the test dataset. Out of 354 samples, the model correctly identified 172 cracked and 182 non-cracked surfaces, while misclassifying only 2 non-cracked as cracked and 3 cracked as non-cracked. This small number of errors highlights the reliability of the system and demonstrates that the model performs consistently well across both classes.

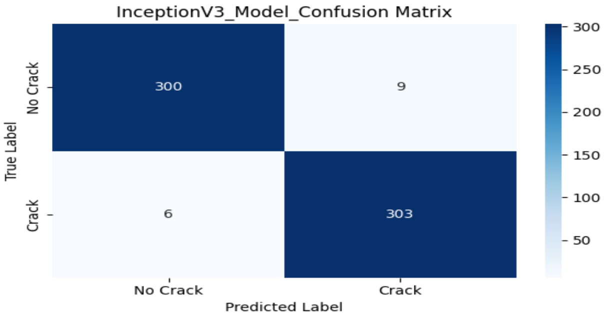

Fig. 5: Confusion Matrix of the Inceptionv3 Model on the Test Dataset.

The balanced performance across classes confirms that the proposed system can accurately distinguish cracked from non-cracked surfaces in practical scenarios.

4.4. Discussion

The results confirm that fine-tuned InceptionV3 on a carefully curated and augmented dataset can yield high-precision classification for detecting cracks in artistic surfaces. Compared to older architectures such as VGG16 and even deeper networks like ResNet50, InceptionV3 provides a better balance between complexity and performance. The model's low error rate, high F1-score, and consistent loss trajectory make it suitable for integration into art restoration and conservation workflows.

Beyond quantitative performance, this study emphasizes the role of preprocessing, augmentation, and transfer learning in addressing domain-specific challenges. These steps are particularly valuable in small and specialized datasets, where variations in texture and brushwork can complicate feature extraction. The findings therefore demonstrate not only the suitability of InceptionV3 for this task but also the broader importance of combining data enhancement strategies with transfer learning to ensure model robustness.

4.5. Limitations

Although the proposed InceptionV3-based model achieved strong performance, several limitations should be acknowledged. The dataset, while carefully curated and augmented, remains relatively small, which may restrict generalization across diverse painting collections. The current framework is limited to binary classification of cracked versus non-cracked surfaces, whereas conservation practice often requires distinguishing between different crack types or severities. Furthermore, the images were collected under controlled conditions, and the system has not yet been tested on artworks with varying mediums, lighting environments, or surface textures. Finally, the model was developed and validated in an offline setting. Its deployment on mobile or embedded platforms for real-time conservation tasks will require further optimization to address computational and memory constraints.

5. Conclusion

This paper presented a deep learning-based approach for detecting surface cracks in paintings using a fine-tuned InceptionV3 model. The model was trained on an augmented dataset of 700 images and outperformed other CNN-based architectures, achieving a test accuracy of 98.01 % and an F1-score of 98.01 %. The results confirmed that the proposed method is highly effective for automated crack classification in artistic surfaces, offering a fast and non-invasive alternative to manual inspection.

In future work, the proposed system can be extended beyond binary classification to identify different crack types and severities, such as micro-cracks versus structural cracks, which would provide more practical insights for conservators. The framework can also be tested on a wider variety of painting mediums, including frescoes, watercolours, and oil paintings, to evaluate its adaptability to diverse artistic contexts. Another important direction is the integration of the model into mobile and edge devices for real-time inspection, which may require optimization techniques to address computational and memory constraints. Furthermore, coupling the model with robotic or UAV-based inspection platforms could enable large-scale, automated scanning of artworks in museums and heritage sites. These extensions would enhance the adaptability and practical deployment of the proposed approach, ensuring that it evolves into a versatile tool for intelligent, technology-driven cultural heritage preservation.

References

[1] Y.-J. Cha, W. Choi, and O. Büyüköztürk, “Deep learning-based crack damage detection using convolutional neural networks,” *Comput.-Aided Civ. Infrastruct. Eng.*, vol. 32, no. 5, pp. 361–378, May 2017, https://doi.org/10.1111/mice.12263.

[2] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, “Road crack detection using deep convolutional neural network,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2016, pp. 3708–3712. https://doi.org/10.1109/ICIP.2016.7533052.

[3] Y. Wu, S. Li, J. Li, Y. Yu, J. Li, and Y. Li, “Deep learning in crack detection: A comprehensive scientometric review,” *J. Infrastruct. Intell. Resil.*, vol. 4, no. 3, p. 100144, 2025, https://doi.org/10.1016/j.iintel.2025.100144.

[4] V. Badrinarayanan, A. Kendall and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 12, pp. 2481-2495, 1 Dec. 2017, https://doi.org/10.1109/TPAMI.2016.2644615.

[5] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770-778, https://doi.org/10.1109/CVPR.2016.90.

[6] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015.

[7] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 2818-2826, https://doi.org/10.1109/CVPR.2016.308.

[8] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2015, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.

[9] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 2261-2269, https://doi.org/10.1109/CVPR.2017.243.

[10] T. -Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, "Focal Loss for Dense Object Detection," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 2999-3007, https://doi.org/10.1109/ICCV.2017.324.

[11] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, “Road damage detection and classification using deep neural networks with smartphone images,” *Comput.-Aided Civ. Infrastruct. Eng.*, vol. 33, no. 12, pp. 1127–1141, 2018. https://doi.org/10.1111/mice.12387.

[12] Yu Zhang, Xin Gao, and Hanzhong Zhang, "Deep Learning-Based Semantic Segmentation Methods for Pavement Cracks" *Information* 14, no. 3: 182, 2023. https://doi.org/10.3390/info14030182.

[13] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, “Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection,” *Constr. Build. Mater.*, vol. 157, pp. 322–330, 2017. https://doi.org/10.1016/j.conbuildmat.2017.09.110.

[14] Guoshao Su, Yuanzhuo Qin, Huajie Xu, Jinfu Liang, "Automatic real-time crack detection using lightweight deep learning models”, Engineering Applications of Artificial Intelligence,Volume 138, Part A, 2024, 109340, https://doi.org/10.1016/j.engappai.2024.109340.

[15] Yuan, Qi, Yufeng Shi, and Mingyue Li, "A Review of Computer Vision-Based Crack Detection Methods in Civil Infrastructure: Progress and Challenges" *Remote Sensing* 16, no. 16: 2910, 2024 https://doi.org/10.3390/rs16162910.

[16] Hyunjun Kim, Sung-Han Sim, Billie F. Spencer, “Automated concrete crack evaluation using stereo vision with two different focal lengths”, Automation in Construction, vol. 135, 104136, 2022, https://doi.org/10.1016/j.autcon.2022.104136.

[17] Y. Ni, T. Shi, F. Wu, and Y. Liu, “Surface defect detection for painted images using attention-guided CNN,” *Appl. Sci.*, vol. 10, no. 17, p. 6062, 2020.

[18] F. Wei, Y. Jiao, Z. Huangfu, G. Shi, N. Wang, H. Dong. “Weakly-supervised segmentation with ensemble explainable AI: A comprehensive evaluation on crack detection”, Rev Sci Instrum., 96(4), 045106, 2025https://doi.org/10.1063/5.0249805.

[19] Y. Chen, J. Pan, J. Lei, D. Zeng, Z. Wu and C. Chen, "EEE-Net: Efficient Edge Enhanced Network for Surface Defect Detection of Glass," in *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-13, 2023. https://doi.org/10.1109/TIM.2023.3320746.

[20] Tingyang Chen, Zhenhua Cai, Xi Zhao, Chen Chen, Xufeng Liang, Tierui Zou, Pan Wang, “Pavement crack detection and recognition using the architecture of segNet”, Journal of Industrial Information Integration, Vol. 18, 100144, 2020, https://doi.org/10.1016/j.jii.2020.100144.

[21] A. Hussain, A. Aslam,, “Ensemble-based approach using inception V2, VGG-16, and Xception convolutional neural networks for surface cracks detection”, *Journal of Applied Research and Technology*, *22*(4), 586–598, 2024 https://doi.org/10.22201/icat.24486736e.2024.22.4.2431.

[22] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, https://doi.org/10.1109/TPAMI.2016.2577031.

[23] Hongxia Li, Weixing Wang, Mengfei Wang, Limin Li, Vivian Vimlund, “A review of deep learning methods for pixel-level crack detection”, Journal of Traffic and Transportation Engineering, Vol. 9 (6), pp. 945-968, 2022. https://doi.org/10.1016/j.jtte.2022.11.003.

[24] C. Kanellakis and G. Nikolakopoulos, “Survey on computer vision for UAVs: Current developments and trends,” *J. Intell. Robot. Syst.*, vol. 87, no. 1, pp. 141–168, 2017. https://doi.org/10.1007/s10846-017-0483-z.

[25] T. Song, T. Chen, Y. Gong, Y. Wang, L. Ran, J. Chen, H. Tang, and Z. Zou, “Enhanced Sagger Crack Detection Integrating Deep Learning and Machine Vision,” Electronics, vol. 13, no. 24, p. 5010, 2024, https://doi.org/10.3390/electronics13245010.

[26] X. Liu, J. Zhu, L. Zhang, and Z. Li, “Robust crack detection in masonry structures with transformers,” *Measurement*, vol. 195, p. 111590, 2022, https://doi.org/10.1016/j.measurement.2022.111590.

[27] A. Islam, Y. Zhang, and M. J. Islam, “Vision transformer-based autonomous crack detection on asphalt and concrete surfaces,” *Autom. Constr.*, vol. 139, p. 104316, 2022, https://doi.org/10.1016/j.autcon.2022.104316.

[28] Min Feng, and Xu Juncai, "Lightweight Dual-Attention Network for Concrete Crack Segmentation" *Sensors* 25, no. 14: 4436, 2025 https://doi.org/10.3390/s25144436.

[29] R. Gupta and P. Verma, “Resilient lightweight CNN with coordinate attention for industrial surface defect detection,” *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023, https://doi.org/10.1109/TIM.2023.3246499.

[30] A. A. Alkannad *et al*., "CrackVision: Effective Concrete Crack Detection with Deep Learning and Transfer Learning," in *IEEE Access*, vol. 13, pp. 29554-29576, 2025, https://doi.org/10.1109/ACCESS.2025.3540841.

[31] C. Gode, B. M. Nanche, D. Dhabliya, R. D. Shelke, , R. V. Patil. & Shushma Bhosle, “Dynamic neural architecture search: A pathway to efficiently optimized deep learning models” , Journal of Information and Optimization Sciences, 46:4-A, 1117–1127, 20-25. https://doi.org/10.47974/JIOS-1896.

[32] R. V. Patil, V, R. Gosavi, G. M. Poddar, S. K. Swarnkar, “Comparative Study of Traditional Edge Detection Methods and Phase Congruency Based Method”, International Journal of Informatics and Communication Technology (IJ-ICT), vol.14, no. 3, pp. 868-880, 2025. https://doi.org/10.11591/ijict.v14i3.pp868-880.

[33] J. S. R, R. V. Patil, Y. S, P. M, S. S and R. Maranan, "Spilled Deep Capsule Neural Network with Skill Optimization Algorithm for Breast Cancer Recognition in Mammograms," *2025 International Conference on Inventive Computation Technologies (ICICT)*, Kirtipur, Nepal, 2025, pp. 632-637, https://doi.org/10.1109/ICICT64420.2025.11004784.

[34] R. V. Patil, R. M. Patil, D. Y. Bhadane, G. M. Poddar, A. L. Wakekar, & S. R. Patil, “Adaptive Cheetah Optimization-Driven CNN: A Hybrid Approach for Robust Image Segmentation.”, International Journal of Basic and Applied Sciences, 14(2), 691-702, Aug. 2025. https://doi.org/10.14419/c62avh78.

[35] Rajendra V. Patil, Renu Aggarwal, ”Comprehensive Review on Image Segmentation Applications”, Sci.Int.(Lahore), 35(5), pp. 573-579, Sep. 2023.