Revolutionizing Building Safety with AI and Deep Learning-Based Crack Detection

Yati Kudtarkar
SCET, BTech CSE
MIT World Peace University
Pune, Maharashtra, India
1032211379@mitwpu.edu.in

Kamakshi Sarbhai
SCET, BTech CSE
MIT World Peace University
Pune, Maharashtra, India
1032211505@mitwpu.edu.in

Nicole Lobo
SCET, BTech CSE
MIT World Peace University
Pune, Maharashtra, India
1032211344@mitwpu.edu.

Shreya Deshpande
SCET,Btech CSE
MIT World Peace University
Pune, Maharashtra, India
1032211616@mitwpu.edu.in

Abstract— In the modern era of rapid urbanization, ensuring structural integrity is crucial for the safety and longevity of buildings. Cracks in infrastructures often go unnoticed until they pose significant risks, necessitating an advanced and proactive approach to detection. Traditional inspection methods rely heavily on manual assessment, which is time-consuming, prone to human error, and lacks scalability. This highlights the urgent need for an automated, precise, and efficient crack detection system driven by artificial intelligence and deep learning.

Our innovative approach integrates Convolutional Neural Networks (CNNs) for feature extraction with advanced deep learning architectures to enhance crack identification and classification accuracy. By leveraging AI-driven image processing techniques, our system efficiently detects cracks, differentiates their severity, and

provides real-time insights for maintenance and safety assessments. The fusion of deep learning and computer vision ensures high

precision in identifying structural vulnerabilities, reducing the dependency on manual inspections.

Our method offers a dual advantage. First, it enhances the speed and accuracy of crack detection, enabling early intervention to prevent structural failures. Second, the integration of AI-based predictive

analytics facilitates proactive maintenance, significantly improving building safety and durability. By revolutionizing traditional inspection practices, our system paves the way for smarter, AI-powered infrastructure monitoring, ensuring a safer and more resilient built environment.

Keywords— Deep Learning, Convolutional Neural Networks, Structural Health Monitoring, Crack Detection, AI-based Inspection, Infrastructure Safety.

1. INTRODUCTION

A. Motivation, Need, Concern, Background

In today's rapidly evolving urban landscape, ensuring the structural safety of buildings is of paramount importance. Cracks in infrastructures, if left undetected, can lead to severe consequences, including structural failures and financial losses. Traditional methods of crack detection rely on manual inspection, which is time-consuming, prone to human error, and lacks scalability. Moreover, these methods often fail to provide real-time insights, making it difficult to implement timely maintenance strategies.

The need for an automated, AI-powered crack detection system arises from the limitations of conventional techniques. Many existing solutions depend on predefined datasets, limiting their ability to generalize to unseen images. This restricts their effectiveness in real-world scenarios where variations in lighting, material textures, and environmental conditions can significantly impact detection accuracy.

Our research aims to address these challenges by integrating advanced deep learning techniques, including Convolutional Neural Networks (CNNs), for accurate crack identification and classification. Unlike traditional models, our system is designed to correctly predict both images from the training dataset and those outside it, ensuring robust generalization. By leveraging AI-driven image processing, our approach enhances precision, minimizes false positives, and enables real-time monitoring of structural health.

Through this project, we aim to revolutionize crack detection by developing an intelligent, scalable, and highly accurate system. Our goal is to enhance infrastructure safety by reducing human dependency, improving predictive maintenance, and ensuring timely intervention. By bridging the gap between conventional inspection methods and modern AI-driven solutions, our research contributes to a more resilient and sustainable built environment.

1. Research Objective

The main goal of this project is to develop an AI-driven crack detection system that leverages deep learning techniques to accurately identify and classify cracks in building structures,

ensuring enhanced safety and predictive maintenance. To achieve this overarching goal, the following specific research objectives have been outlined:

Accurate Crack Detection: To develop and implement a deep learning-based model that can precisely detect and classify cracks in various structural materials, ensuring high accuracy and reliability in real-world conditions.

Generalization and Robustness: To design a system capable of accurately predicting cracks not only in images from the training dataset but also in unseen images, ensuring adaptability to different environments, lighting conditions, and structural variations.

Advanced Deep Learning Integration: To enhance detection accuracy and efficiency by integrating Convolutional Neural Networks (CNNs) for feature extraction and other deep learning techniques for robust classification and prediction.

Automated and Scalable Monitoring: To create an AI-powered system that enables real-time crack detection, reducing dependency on manual inspections and improving the scalability of structural health monitoring.

User-Centric Implementation: To design an intuitive and user-friendly system that can be seamlessly integrated into building inspection workflows, enabling engineers and maintenance teams to make data-driven decisions for timely interventions.By addressing these research objectives, this project aims to revolutionize traditional crack detection methods, contributing to the field of structural health monitoring by offering an intelligent, scalable, and highly accurate AI-driven solution for ensuring infrastructure safety.

2. Literature Review

The study by Ko et al. (2021) [16] developed an Automated Building Exterior Crack Inspection System (ABECIS) using Google's Xception CNN and Özgenel's crack image dataset. This system integrates UAVs for autonomous image capture, deep learning for crack detection, and 3D modeling for crack visualization. The aim is to improve inspection efficiency and enhance worker safety while reducing data collection errors. However, challenges such as lighting variations, motion blur, occlusions, and high false positive rates limit its reliability. Additionally, the absence of a standardized dataset hinders model generalization, and the lack of predictive maintenance integration reduces its practical effectiveness.

The study by Yin et al. (2024) [15] systematically reviewed the application of deep learning in modular construction safety management, highlighting the use of CNNs, RNNs, GANs, Autoencoders, Deep Belief Networks (DBN), and Transformer models. The study found that deep learning significantly enhances safety risk management by improving hazard detection and safety compliance in construction sites. However, key challenges include the lack of interdisciplinary expertise in deep learning and construction engineering, along with insufficient modular construction safety datasets, which limit model scalability and real-world applicability. The study proposed strategies such as multi-modal data sharing platforms, guidance frameworks for AI applications in construction, and talent development programs to address these challenges.

The research by Jahanshahi et al. (2014) [14] introduced an image-based crack detection system utilizing Structure from Motion (SfM) for 3D reconstruction, morphological operations for segmentation, and classification using SVM, neural networks, and nearest neighbor classifiers. The model enables depth perception for accurate crack detection, making it adaptable for deployment on UAVs and autonomous robots. However, traditional visual inspections remain qualitative, non-destructive testing (NDT) methods lack adaptability, and manual interventions are still required for inaccessible regions. The study highlights the need for an automated, scalable, and real-world applicable crack detection approach.

Chen et al. (2023) [17] explored deep learning-based crack detection in building facades, employing CNN models with transfer learning (ResNet101). The study achieved a 94% accuracy improvement over standard CNNs, demonstrating the effectiveness of automating inspection processes for structural integrity assessments. However, research in building crack detection lags that of bridges and pavements, and manual inspections remain labor-intensive and subjective, indicating a gap in fully automated facade inspection methodologies.

The review by Chakurkar et al. (2023) [18] provided a systematic analysis of AI-based crack detection, comparing deep learning and traditional computer vision approaches. It highlighted several key challenges, including the lack of 3D crack datasets, limited adoption of unsupervised learning, high computational demands, and insufficient dataset diversity in pavement textures. Additionally, crack severity classification remains a challenge, necessitating further research in real-time AI-based defect assessment. The study emphasized the need for hybrid AI approaches, combining CNNs, GANs, and traditional vision techniques, to enhance detection accuracy and robustness.

Collectively, these studies highlight the evolution of AI-driven crack detection and the integration of UAVs, deep learning, and computer vision techniques. While significant advancements have been made in automating inspections and improving detection accuracy, challenges persist in dataset standardization, real-world adaptability, and predictive maintenance. Future research should focus on unsupervised learning, 3D dataset development, real-time processing, and integration with Building Information Modeling (BIM) systems to enhance infrastructure safety and reliability.

3. PROPOSED METHODOLOGY

The purpose of this paper is to create a concrete crack detection system using deep learning techniques. The goal of this technology is to assist engineers and infrastructure maintenance personnel by automatically detecting and analyzing cracks in concrete structures from images. The system leverages Convolutional Neural Networks (CNNs) for image processing and classification, using ResNet50 as a pre-trained model for feature extraction. Additionally, OpenCV-based image preprocessing techniques are employed to enhance crack visibility and filter out noise. The project also integrates Flask, a web framework for Python, to provide an interactive web-based platform for real-time crack detection and visualization. The ultimate objective is to develop an automated, efficient, and scalable solution for structural integrity assessment in civil engineering applications.

Deep learning methods—more specifically, Computer Vision (CV) and Convolutional Neural Networks (CNNs)—are utilized in this research. The system is designed to process raw images, extract features, and classify them as cracked or non-cracked. The ResNet50 model, a deep residual network known for its efficiency in feature extraction, is used as the backbone of the detection pipeline. The model is pre-trained on the ImageNet dataset and fine-tuned on a dataset of concrete surface images to improve crack detection accuracy.

The ResNet50 architecture consists of multiple residual blocks, which help mitigate the vanishing gradient problem and allow deeper network training. These residual blocks retain crucial low-level and high-level features essential for detecting thin, irregular cracks. The final fully connected layer of ResNet50 is replaced with custom dense layers to classify images into "cracked" or "non-cracked" categories. Global Average Pooling (GAP) is applied to reduce overfitting, and a softmax/sigmoid activation function is used depending on the classification type (binary or multi-class).

Furthermore, image preprocessing using OpenCV includes grayscale conversion, contrast enhancement, adaptive thresholding, and morphological operations to highlight crack structures before passing them to the model. The algorithm also employs contour detection for bounding box generation, allowing visualization of detected cracks. The system's final output includes an annotated image with detected cracks and a classification score.

To further enhance usability, the model is integrated into a web application using Flask, enabling users to upload images and receive instant results. The ResNet50-based approach ensures high accuracy, robustness, and generalization across different concrete textures and lighting conditions. Additionally, the OpenCV-based pipeline processes the images in real-time, ensuring efficient and reliable crack detection, reducing manual inspection time, and enhancing safety in construction and maintenance industries.

4. SYSTEM ARCHITECTURE

A deep learning image classification algorithm was developed and trained based on Google's Xception convolutional neural network . Özgenel's dataset [3], composed of 15,000 non-crack images and 15,000 crack images with 227×227 pixels resolution, was used to train the image classification algorithm. Figure 1 shows an example of the sample training data from the dataset used.

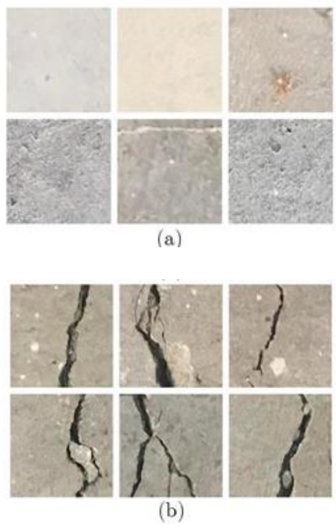

Figure 4.1. (a) Sample non-crack images from original dataset (b) Sample crack images from the original dataset (reproduced from [3])

The image classification algorithm was developed using Keras, an open-source Python deep learning library. The algorithm operates in the following steps:
1) load pre-trained model
2) take the input image (the texture generated by 3D photogrammetry),
3) split each image into an array of rectangular segments, 4) perform analysis on each segment and classify it as "crack" or "no-crack",
5) mark the segments which are classified as "crack" for visualization later. The process is visualized in Figure 2

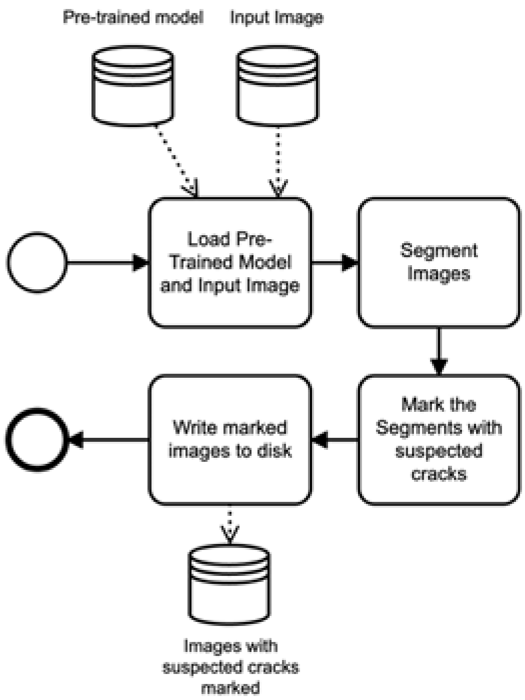

Fig.4.2. Deep Learning Image Analysis Algorithm(CNN)

ResNet50 is a deep convolutional neural network architecture designed for image classification and feature extraction. In the crack detection system, second deep learning model used was ResNet50 was utilized as a pre-trained model to extract high-level features from images, distinguishing between "crack" and "no-crack" regions. The architecture consists of multiple convolutional layers, identity blocks, and residual connections, which help in preserving essential image features and improving classification accuracy.

The process involves:

1. Input Processing – Images are preprocessed and passed through initial convolution, batch normalization, and ReLU activation.
2. Feature Extraction – ResNet50's deep layers extract features from the image using convolutional and identity blocks.
3. Classification – Extracted features are passed through an average pooling layer, flattened, and fed into fully connected (FC) layers to classify the image.

This method ensures efficient and accurate detection of cracks, leveraging residual learning to maintain performance while reducing computational complexity.The process is visualized in Figure 3.

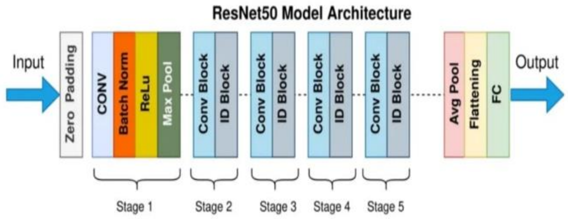

Fig4.3 Deep Learning Image Analysis Algorithm (ResNet50)

5.DISCUSSIONS AND FINDINGS

The main goal of our study is to detect cracks in building images using deep learning techniques. To achieve this, we utilize Convolutional Neural Networks (CNNs) for feature extraction and ResNet50, a powerful residual learning framework, to enhance the accuracy of crack detection. By leveraging pre-trained deep learning models, our system effectively analyzes structural images, identifying cracks with high precision and reliability.

Table 1 Accuracy and Loss Score

Model	Accuracy	Loss
CNN	0.9966	0.0123
ResNet50	0.9996	0.0025

Table 1 demonstrates the accuracy and loss scores for the ResNet50 and CNN models in crack detection. The results indicate that ResNet50 outperforms the traditional CNN model in terms of both accuracy and loss reduction. By leveraging residual learning, ResNet50 effectively mitigates the vanishing gradient problem, leading to improved feature extraction and better generalization. The evaluation metrics highlight that ResNet50 achieves higher accuracy and lower loss, making it a more reliable model for detecting cracks in building images.

Our findings indicate that transfer learning with ResNet50 enhances crack detection accuracy by effectively capturing intricate patterns. The pre-trained layers allow efficient feature extraction, improving classification performance with minimal training data.

Key visual features were extracted using ResNet50, followed by fine-tuning for classification. The model demonstrated robustness against variations in lighting and surface texture, ensuring reliable detection.

Our analysis showed that this approach streamlined the crack detection process, with ResNet50 efficiently distinguishing cracked and non-cracked surfaces. This significantly improved accuracy, making it a viable solution for structural health monitoring.

Fig 5.1 Original Cracked Image

Fig 5.2 Output image with boundary box and predicted label

Our findings indicate that deep learning significantly enhances crack detection in concrete walls. By leveraging a convolutional neural network (CNN) architecture, the model effectively captured key visual features, improving classification accuracy with minimal preprocessing.

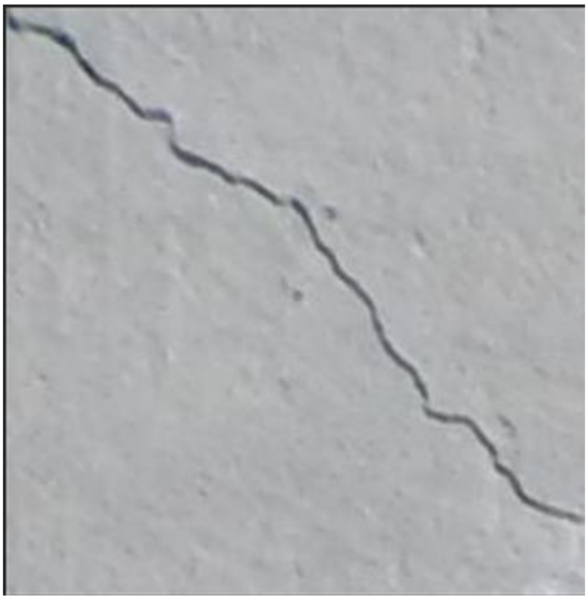

Fig 5.3 Original Cracked Image

Fig 5.4 Output image with boundary box and predicted label

CNN-based feature extraction allowed the system to distinguish cracked and non-cracked surfaces efficiently. The hierarchical feature learning enabled robust detection across varying lighting conditions, surface textures, and angles.

Our analysis demonstrated that this CNN-based approach streamlined the detection process. The model successfully identified cracks with high accuracy, making it a viable tool for automated structural health monitoring.

I. FUTURE SCOPE

Future research on crack detection using Faster R-CNN should focus on several key areas. Enhancing dataset diversity to include various building materials and crack types will improve system performance and generalizability. Optimizing model architectures and hyperparameters can further enhance detection accuracy and reduce computational costs. Integrating real-time processing capabilities will enable faster and more efficient crack detection in practical applications. Lastly, conducting extensive field tests on real-world structures will help validate the model's effectiveness and refine it based on real environmental conditions and structural variations.

II. CONCLUSION

Developed a system that detects cracks in building images using ResNet-50 and CNN models. We combined advanced image processing and deep learning techniques to accurately identify and classify cracks, ensuring reliable structural analysis. The model processes images efficiently, detecting even minute cracks with high precision. This work contributes to the automation of structural health monitoring, reducing manual inspection efforts and improving safety. By leveraging deep learning, the system enhances the accuracy and speed of crack detection, making it a valuable tool for infrastructure maintenance and safety assessments.

III. ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to our faculty advisor, Dr. Amit Savyanavar sir, for his invaluable guidance, support, and expertise throughout the duration of this natural language processing project. His mentorship has been instrumental in shaping our research approach and refining our methodologies.

5. REFERENCES

[1] A. M. Paterson, G. R. Dowling, and D. A. Chamberlain, "Building inspection: can computer vision help?," Automation in Construction, vol. 7, no. 1, pp. 13–20, 1997, doi: https://doi.org/10.1016/S0926-5805(97)00031-9. SMART-NYUAD/Concrete-Wall-Crack-Detection. S.M.A.R.T. Construction Research Group, 2021. Accessed: Jun. 02, 2021. [Online]. Available: https://github.com/SMART-NYUAD/Concrete Wall-Crack-Detection

[2] M. P. Ismail, "Selection of suitable NDT methods for building inspection," IOP Conf. Ser.: Mater. Sci. Eng., vol. 271, p. 012085, Nov. 2017, doi: 10.1088/1757-899X/271/1/012085.

[3] Ç. F. Özgenel and A. G. Sorguç, "Performance Comparison of Pretrained Convolutional Neural Networks on Crack Detection in Buildings," in Proceedings of the 35th International Symposium on Automation and Robotics in Construction (ISARC), Taipei, Taiwan, Jul. 2018, pp. 693–700. doi: 10.22260/ISARC2018/0094.
643
38th International Symposium on Automation and Robotics in Construction (ISARC 2021)

[4] S. Choi and E. Kim, "Building crack inspection using small UAV," in 2015 17th International Conference on Advanced Communication Technology (ICACT), Jul. 2015, pp. 235–238. doi: 10.1109/ICACT.2015.7224792.

[5] T. Rakha and A. Gorodetsky, "Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones," Automation in Construction, vol. 93, pp. 252–264, 2018, doi: https://doi.org/10.1016/j.autcon.2018.05.002.

[6] T. M. Cabreira, L. B. Brisolara, and P. R. Ferreira Jr., "Survey on Coverage Path Planning with Unmanned Aerial Vehicles," Drones, vol. 3, no. 1, 2019, doi: 10.3390/drones3010004.

[7] M. D. Phung, T. H. Dinh, V. T. Hoang, and Q. Ha, "Automatic Crack Detection in Built Infrastructure Using Unmanned Aerial Vehicles," Proceedings of the 34th International Symposium on Automation and Robotics in Construction (ISARC), Jul. 2017, doi: 10.22260/isarc2017/0115.

[8] J. Kim, S.-B. Kim, J.-C. Park, and J.-W. Nam, "Development of Crack Detection System with Unmanned Aerial Vehicles and Digital Image Processing," presented at the Advances in Structural Engineering and Mechanics, 2015.

[9] B. Rao, A. G. Gopi, and R. Maione, "The societal impact of commercial drones," Technology in Society, vol. 45, pp. 83–90, 2016, doi: https://doi.org/10.1016/j.techsoc.2016.02.009.

[10] "Mavic 2 - Product Information - DJI," DJI Official. https://www.dji.com/nl/mavic-2/info (accessed Apr. 17, 2021).