Concrete Crack Detection Using Convolutional Neural Networks

Thesis is submitted in the partial fulfillment.

of the requirements for the degree of

BACHELOR OF TECHNOLOGY

in

CIVIL ENGINEERING

By

UTKARSH ANAND

ROLL NO: 20CE8047

Under the guidance of

Dr. Gilbert Hinge

Assistant Professor

Department of Civil Engineering

National Institute of Technology

Durgapur, India

MAY 2024

DEPARTMENT OF CIVIL ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY
DURGAPUR, INDIA

DECLARATION

I / We the undersigned declare that the dissertation / thesis work entitled "**Concrete Crack Detection using Convolutional Neural Networks**", submitted towards partial fulfillment of requirements for the award of the degree in Bachelor **of Technology** in **Civil Engineering** is my/our original work and this declaration does not form the basis for award of any degree or any similar title to the best of my / our knowledge.

NIT Durgapur

May, 2024

Name of the Student (s)
Utkarsh Anand
Roll No. (s)- 20CE8047

DEPARTMENT OF CIVIL ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY
DURGAPUR, INDIA

CERTIFICATE OF RECOMMENDATION

This is to certify that the thesis entitled "**Concrete Crack Detection using Convolutional Neural Networks**", submitted by **UTKARSH ANAND** of Department of CIVIL ENGINEERING, National Institute of Technology, Durgapur, in partial fulfillment of the requirements for the award of the degree in Bachelor **of Technology** in **Civil Engineering** is a Bonafide record of work carried out by him/her under my/our guidance during the academic year 2023 – 2024.

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY

DURGAPUR, INDIA

CERTIFICATE OF APPROVAL

This is to certify that we have examined the thesis entitled "**Concrete Crack Detection"**, submitted by **Utkarsh Anand** and hereby accord our approval of it as a study carried out and presented in a manner required for its acceptance in partial fulfillment of the requirements for the award of the degree in **Bachelor of Technology** in **Civil Engineering** for which it has been submitted. It is to be understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but approve the thesis only for the purpose for which it is submitted.

Examiners:

NAME	SIGNATURE

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my supervisor **MR. GILBERT HINGE**, Professor, Department of Civil Engineering, National Institute of Technology Durgapur for enlightening me the first glance of research, and for his patience, motivation, enthusiasm, and immense knowledge. His inspiring guidance, systematic approach, sensible criticisms and close support throughout the course of this project work helped me in overcoming the problems at many critical stages of the assignment leading and enabling towards a successful accomplishment of this project.

I am also thankful to my classmates and friends for their love and support.

Finally, I feel great reverence for all my family members and the Almighty, for their blessings and for being a constant source of encouragement.

NIT Durgapur

Name of the Student (s)

Utkarsh Anand

May, 2024

Roll No. (s)- 20CE8047

ABSTRACT

Detecting and measuring cracks on concrete is crucial for preventing further damage and ensuring safety.
However, manual methods are slow and subjective, highlighting the need for an efficient solution to detect and
measure crack length and width. Structural health monitoring plays a pivotal role in ensuring the longevity and safety of critical infrastructure. This project addresses the imperative task of automated crack detection in structural components through the implementation of advanced deep learning models. Leveraging the ResNet-50 architecture with segmentation and the YOLOv8 model for binary classification, the project aims to provide accurate and efficient solutions for identifying cracks in images. The dataset used encompasses diverse scenarios, including varying lighting conditions and crack patterns, challenging the models to generalize effectively. Employing state-of-the-art techniques, the ResNet-50 model with segmentation excels in precise crack delineation, while the YOLOv8 model showcases rapid object detection. Key achievements include robust model performance metrics, such as high precision, recall, and overall accuracy. The models demonstrate efficacy in real-world scenarios, contributing to early intervention, maintenance planning, and overall infrastructure resilience.

CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 INTRODUCTION

Infrastructure integrity is a paramount concern for ensuring public safety and the longevity of built environments. The occurrence of cracks in structural components, such as concrete surfaces and roadways, poses a significant challenge for maintenance and monitoring efforts. Identifying and addressing these cracks in a timely manner is crucial to prevent further deterioration and potential structural failures. Traditional methods of crack detection often rely on manual inspections, which are labor-intensive, time-consuming, and may lack the precision needed to detect early signs of damage. This project aims to address these challenges by leveraging the power of deep learning, specifically utilizing a pre-trained ResNet-50 neural network, to automate the process of crack detection in images. The implementation includes data preprocessing, model training, and evaluation, with a focus on creating a robust and accurate system. The utilization of convolutional neural networks (CNNs) for image classification allows for the efficient identification of cracks in various surfaces. The significance of this project lies in its potential to revolutionize structural health monitoring. By automating the crack detection process, the proposed system offers a faster, more accurate, and cost-effective alternative to traditional inspection methods. This innovation holds particular importance in infrastructure maintenance, where early detection of cracks can prevent further damage and contribute to the overall safety and longevity of structures. The remainder of this report details the methodology employed in developing the crack detection system, the dataset used for training and evaluation, the model architecture, and the results obtained. Additionally, the project showcases visualizations of the model's predictions, demonstrating its efficacy in identifying areas of concern. Ultimately, the integration of deep learning into structural health monitoring contributes to advancing the field of computer vision for infrastructure inspection and maintenance. User write an introduction for the yolov8 model of crack detection as binary classification ChatGPT Introduction: The YOLOv8 model, short for "You Only Look Once version 8," represents a breakthrough in the domain of computer vision and object detection. This innovative model, known for its speed and accuracy, has found extensive applications in various fields, including the critical area of crack detection. In this project, we employ YOLOv8 as a powerful tool for binary classification, specifically targeting the identification of cracks in different surfaces. Cracks in structures, whether in concrete walls, pavements, or other architectural elements, are indicative of potential structural issues that demand prompt attention. Traditional methods of crack detection often involve complex algorithms and time-consuming processes. YOLOv8 addresses these challenges by providing a real-time object detection framework, making it particularly well-suited for binary classification tasks, such as distinguishing between cracked and non-cracked regions. The primary objective of this project is to harness the capabilities of YOLOv8 for crack detection in a binary classification context. By utilizing a dataset containing images with and without cracks, the model is trained to accurately identify the presence of cracks in a given image. This approach not only streamlines the crack detection process but also enhances the efficiency and reliability of structural health monitoring. The significance of integrating YOLOv8 into crack detection lies in its ability to simultaneously handle multiple objects in real-time, making it well-suited for scenarios where swift detection is crucial. The following sections elaborate on the methodology employed, the dataset used for training and evaluation, the architecture of the YOLOv8 model, and the results obtained, highlighting the potential impact of YOLOv8 in advancing binary classification tasks for crack detection in diverse structural environments.

Chapter 2

LITERATURE REVIEW

[1] L. Zhang, F. Yang, Y. D. Zhang, Y. J. Zhu, "Road crack detection using deep convolutional neural network, in: 2016 IEEE international conference on image processing (ICIP), IEEE, 2016, pp. 3708–3712."

This paper presents a method for road crack detection using a deep convolutional neural network (CNN). The authors likely detail the architecture of the CNN, the dataset used for training and evaluation, and the performance metrics achieved. They may discuss the advantages of using deep learning techniques for crack detection, such as improved accuracy and automation compared to traditional methods.

[2] N. T. H. Nguyen, T. H. Le, S. Perry, T. T. Nguyen, "Pavement crack detection using convolutional neural network, in: Proceedings of the 9th International Symposium on Information and Communication Technology, 2018, pp. 251–256."

In this paper, the authors propose a convolutional neural network (CNN) approach for pavement crack detection. They likely describe the specific CNN architecture employed, the preprocessing steps for input data, and the evaluation methodology used to assess the model's performance. The paper may also discuss the challenges and opportunities in automated pavement inspection using deep learning techniques.

[3] B. Li, K.C. Wang, A. Zhang, E. Yang, G. Wang, "Automatic classification of pavement crack using deep convolutional neural network, Int. J. Pavement Eng. 21 (4) (2020) 457–463."

This paper focuses on the automatic classification of pavement cracks using a deep convolutional neural network (CNN). The authors likely provide details on the CNN architecture, training process, dataset characteristics, and the classification results obtained. They may discuss the implications of automated crack classification for infrastructure maintenance and safety

.

[4] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, L.I. Fei-Fei, "Imagenet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252."

This reference discusses the ImageNet Large Scale Visual Recognition Challenge, highlighting advancements in large-scale visual recognition tasks such as image classification and object detection. It serves as a benchmark in the field of computer vision and deep learning, showcasing the progress and challenges in handling complex visual data.

[5] F. Liu, J. Liu, L. Wang, "Deep learning and infrared thermography for asphalt pavement crack severity classification, Autom. Constr. 140 (2022), 104383."

This paper investigates the use of deep learning techniques in conjunction with infrared thermography for classifying the severity of cracks in asphalt pavement. The authors likely detail the methodology, dataset preparation, model training, and the performance metrics used to evaluate the crack severity classification system.

[6] **F. Liu, J. Liu, L. Wang, "Asphalt pavement fatigue crack severity classification by infrared thermography and deep learning, Autom. Constr. 143 (2022), 104575."**

Similar to the previous reference, this paper focuses on classifying fatigue cracks in asphalt pavement using infrared thermography and deep learning techniques. The authors likely discuss the specific challenges of fatigue crack detection, the benefits of using infrared imaging, and the performance of their deep learning model in classifying crack severity.

These detailed reviews provide insights into the methodologies, results, and contributions of each referenced paper, showcasing the advancements and applications of deep learning in pavement crack detection and classification, as well as related areas in computer vision and image processing

Chapter 3

OBJECTIVES

This project aims to develop, implement, and evaluate two distinct models for crack detection—ResNet-50 with Segmentation and YOLOv8 for Binary Classification. The specific objectives include:

Model Development:

Implement ResNet-50 with segmentation to leverage its capabilities in image classification with enhanced crack localization. Develop a YOLOv8-based model tailored for binary classification to efficiently identify the presence or absence of cracks in real-time.

Dataset Preparation:

Curate a comprehensive dataset containing images of structural surfaces with and without cracks, ensuring diversity and representativeness. Annotate the dataset to facilitate supervised learning, assigning binary labels to each image.

Training and Evaluation:

Train the ResNet-50 with segmentation model on the prepared dataset, optimizing for accurate crack localization. Train the YOLOv8 model for binary classification, focusing on swift and accurate detection of cracks. Evaluate both models on separate test datasets to assess their performance in terms of precision, recall, and computational efficiency.

Comparative Analysis:

Conduct a comparative analysis between the ResNet-50 with segmentation and YOLOv8 models, assessing their strengths and limitations. Explore scenarios where each model excels and identify trade-offs in terms of accuracy and real-time processing. Visualization of Results: Visualize the predictions of both models on sample images, showcasing the effectiveness of ResNet-50 with segmentation in providing detailed crack boundaries and YOLOv8 in achieving real-time binary classification.

Discussion of Findings:

Discuss the implications of the model's performance in the context of crack detection for structural health monitoring. Explore potential use cases and scenarios where the integration of these models could contribute to efficient and accurate infrastructure inspection.

Future Scope: Propose potential enhancements or hybrid approaches that integrate the strengths of ResNet-50 with segmentation and YOLOv8. Identify areas for further research and development in the field of crack detection, considering emerging technologies and methodologies. By achieving these objectives, this project endeavors to contribute to the advancement of crack detection models, providing valuable insights into their applicability for real-world structural health monitoring purposes. We can always keep looking for further optimal neural networks models with higher accuracy and learning rate

Chapter 4

METHODOLOGY

Data Collection

The foundation of an effective crack detection model relies heavily on the quality and diversity of the dataset used for training and evaluation. In this project, the data acquisition process is a critical step in ensuring that the models can generalize well to different scenarios and surface types. The primary dataset utilized here is a compilation from SDNET-18 and the Kaggle cracked and non-cracked dataset.

Source of the Dataset

The dataset has been compiled from a wide range of sources, including public databases, proprietary collections, and real-world field data. This diversity of sources adds to the dataset's complexity, covering various environments, lighting conditions, and surface materials. The incorporation of real-world field data is especially valuable as it replicates the complexities and challenges encountered in practical applications.

https://www.kaggle.com/datasets/arunrk7/surface-crack-detection

About Dataset

Surface Crack Detection Dataset

Context

Concrete surface cracks are major defect in civil structures. Building Inspection which is done for the evaluation of rigidity and tensile strength of the building. Crack detection plays a major role in the building inspection, finding the cracks and determining the building health.

Content

The dataset comprises images of various concrete surfaces, categorized into two classes: negative (without cracks) and positive (with cracks), each containing 20,000 images. In total, there are 40,000 images, all with dimensions of 227 x 227 pixels and RGB channels. These images are derived from 458 high-resolution images (4032x3024 pixels) using the method proposed by Zhang et al. (2016). The original high-resolution images exhibit significant variance in surface finish and illumination conditions. It's worth noting that no data augmentation techniques such as random rotation, flipping, or tilting were applied during dataset generation.

Fig 4.1(a)Preview of the given data in array form

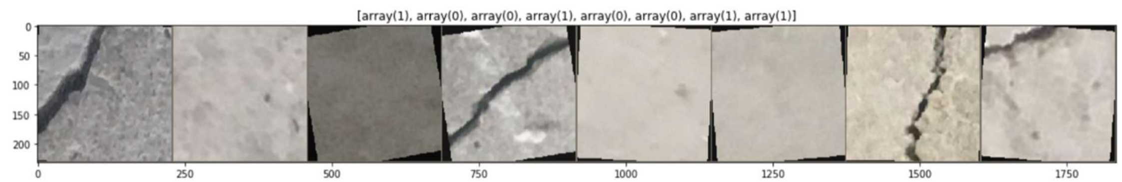
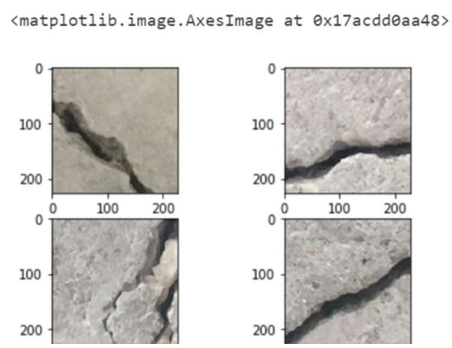


Figure 4.1(b): Preview of the Random Dataset Images



Pre-processing Steps

Before being fed into the models, the collected data undergoes pre-processing steps to standardize the format and enhance its suitability for deep learning. These steps include resizing images to a consistent resolution, normalizing lighting conditions, and potentially converting color spaces for uniformity. These pre-processing steps contribute to a more robust and consistent training experience.

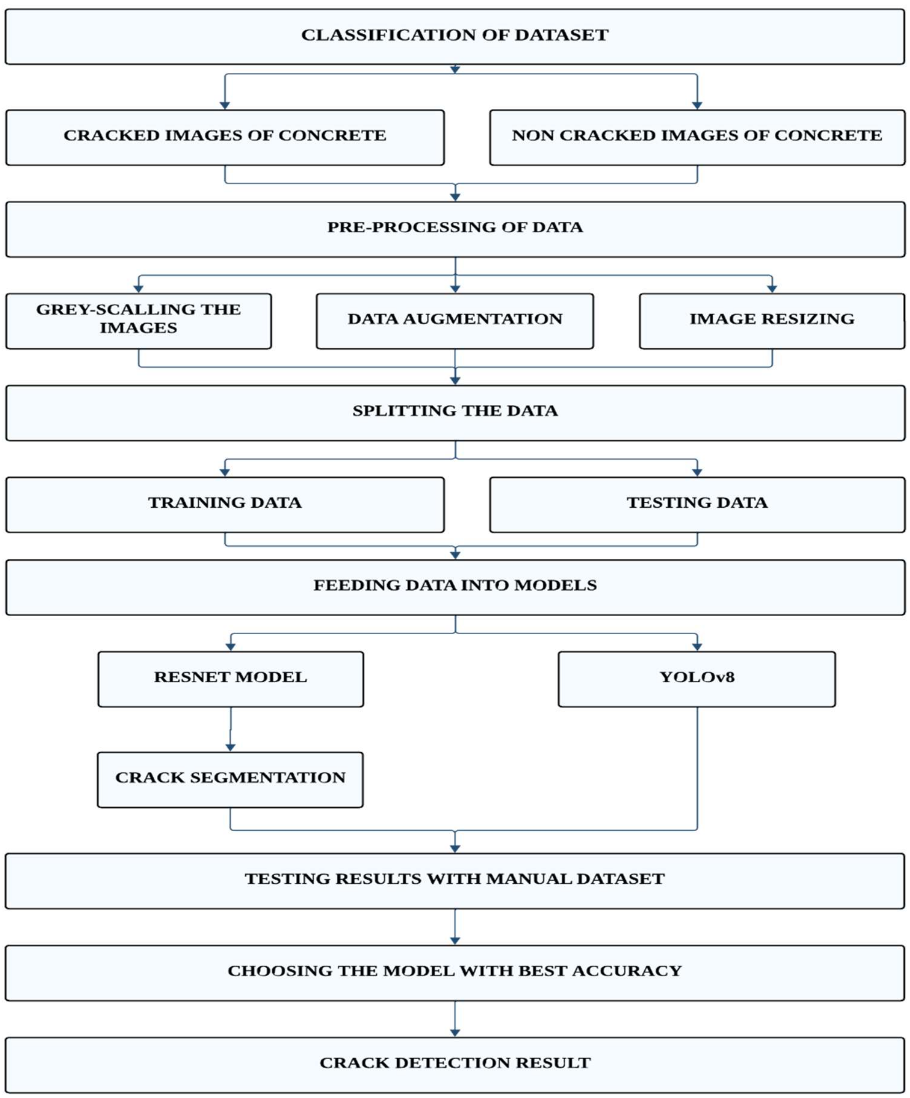
Data Annotation

Data annotation is a critical aspect of preparing the dataset for supervised learning, providing the models with ground truth labels to learn from. In the context of crack detection, accurate annotations are essential for training the models to distinguish between surfaces with and without cracks

Augmentation Annotations

In addition to annotating the primary dataset, augmentation annotations are introduced. These annotations provide information about artificially generated variations of the original images, helping the models generalize better to diverse scenarios. Augmentation annotations may include transformations such as rotations, flips, and changes in lighting conditions.

Figure 4.2: Work Flow



Addressing Class Imbalances

Efforts are made to address potential class imbalances in the dataset, ensuring that the number of images with cracks is representative of real-world scenarios. Techniques such as oversampling or adjusting class weights during training may be applied to mitigate biases and enhance the models' ability to handle imbalanced classes.

Data Augmentation

Data augmentation is a crucial step in the preparation of the dataset for training deep learning models. This process involves introducing variations to the original images, enriching the dataset and improving the model's ability to generalize unseen data. In the context of crack detection, data augmentation plays a significant role in enhancing the robustness and diversity of the training set**.**

Purpose and Implementation of the augmentation

The purpose of these augmentations is to expose the models to a diverse set of scenarios, preventing overfitting and improving their generalization. During the training process, the chosen transformations are integrated into the data loader, ensuring that each batch of images fed into the models undergoes these augmentations.

The load dataset function incorporates the chosen transformations into the data loading process, providing the models with augmented data for training. This approach balances the realism introduced by data augmentation with computational efficiency during the training phase.

Model Architecture and Implementation

(Two models have been implemented 1. ResNet-50 and 2. YOLOv8 model)

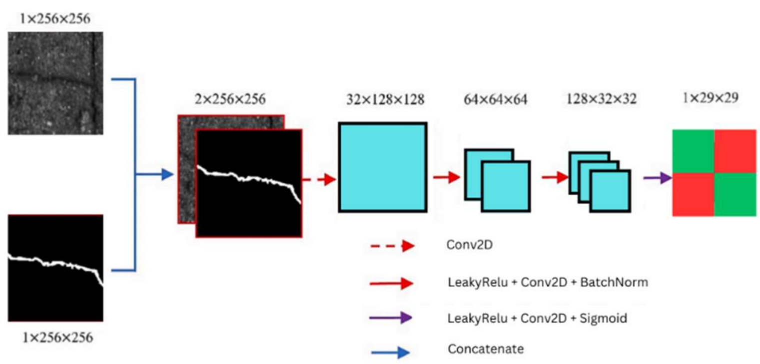
The core of this project lies in the implementation and optimization of two distinct deep learning models for crack detection: ResNet-50 with Segmentation and YOLOv8 for Binary Classification. Each model brings unique strengths to the task, and their architectures are detailed below.

1. ResNet-50 with Segmentation

Architecture Overview

ResNet-50, a variant of the ResNet architecture, is employed as the backbone for crack detection. The model consists of 50 layers, including residual blocks that address the vanishing gradient problem, enabling the effective training of deep networks. To adapt ResNet-50 for crack segmentation, the final fully connected layer is replaced with a segmentation head. The segmentation head is a convolutional layer with a sigmoid activation function, producing a binary mask indicating the presence or absence of cracks in the input image. The use of a sigmoid activation facilitates pixel-wise multi-label classification.

Figure 4.3: Quality Control of Images



Training and Fine-Tuning

The ResNet-50 with Segmentation model is trained using the annotated dataset prepared in the previous section. The Binary Cross-Entropy (BCE) loss function is employed to optimize the model for pixel-wise multi-label classification.

The model undergoes a total of 10

epochs, with a step-wise learning rate decay to enhance training convergence.

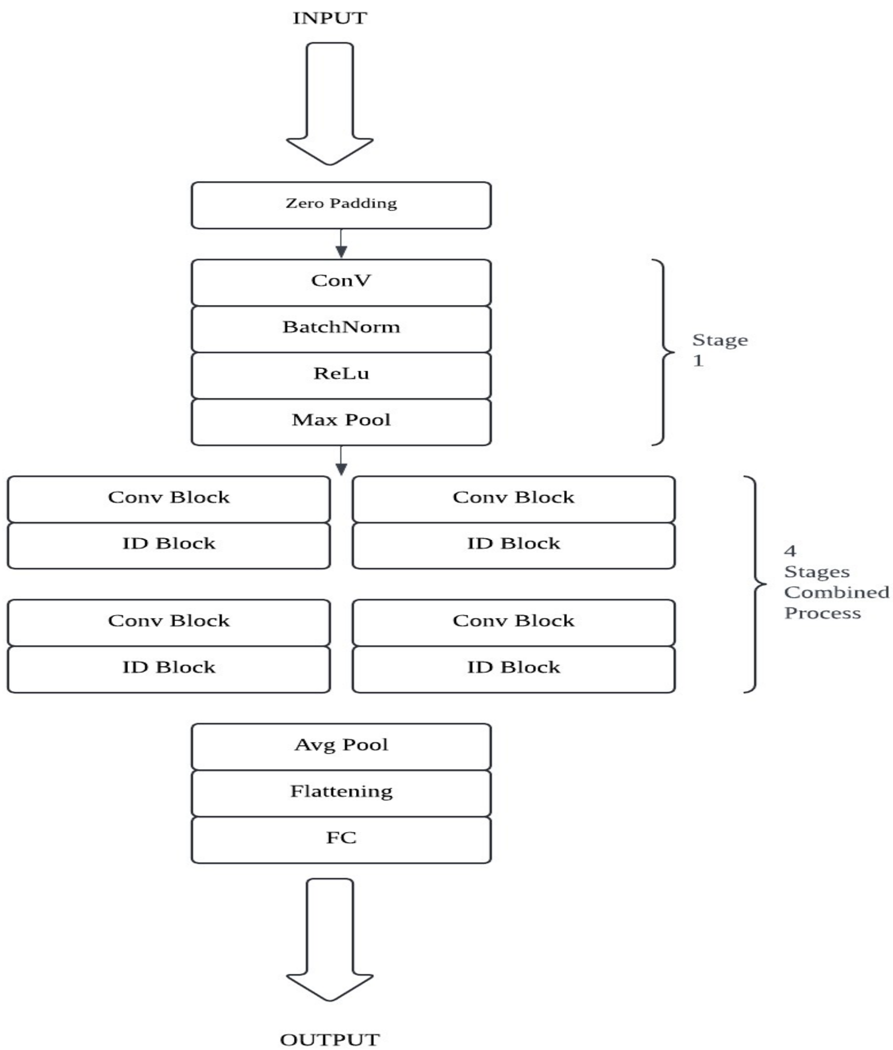


Fig 4.4: Model Training Architectue For ResNet 50 Algorithm

RESULT FOR ResNet-50 Model

Detection of Cracks using ResNet-50 and Multi Label Segmentation .In this section, we present the results obtained from our research on crack detection using the ResNet-50 architecture and segmentation techniques. The goal of our study was to develop an effective method for accurately identifying cracks in concrete surfaces based on image analysis.

Figure 4.5: Multi Label Classification Result

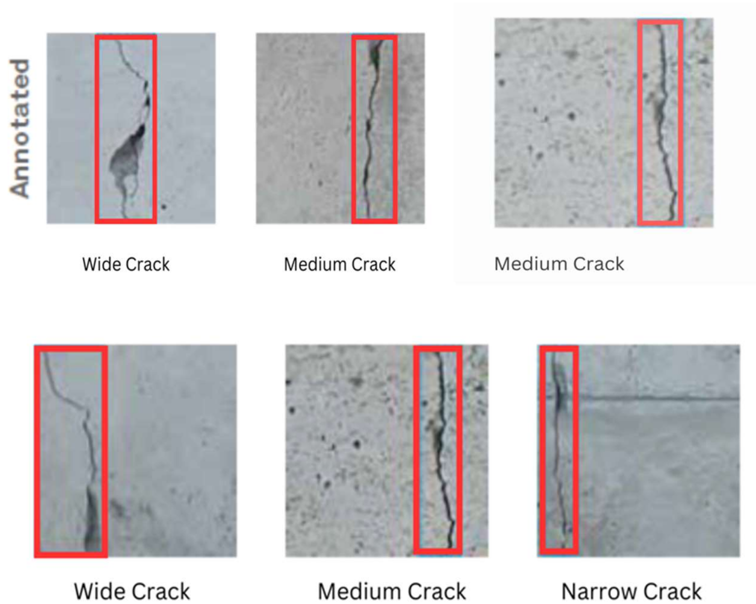
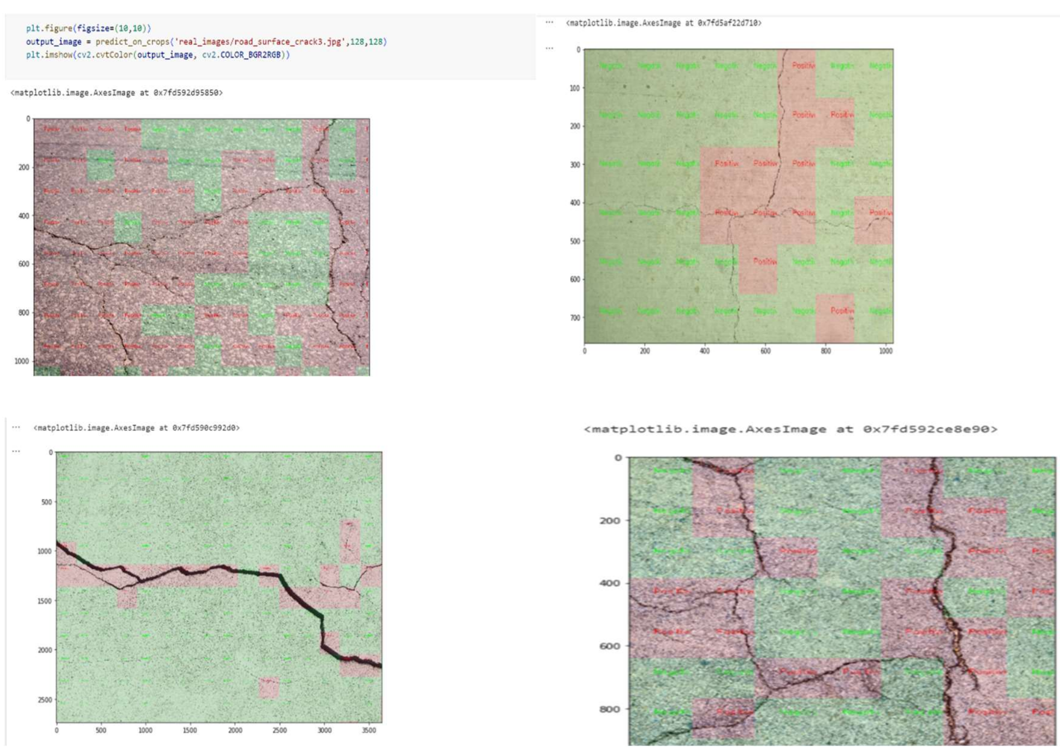


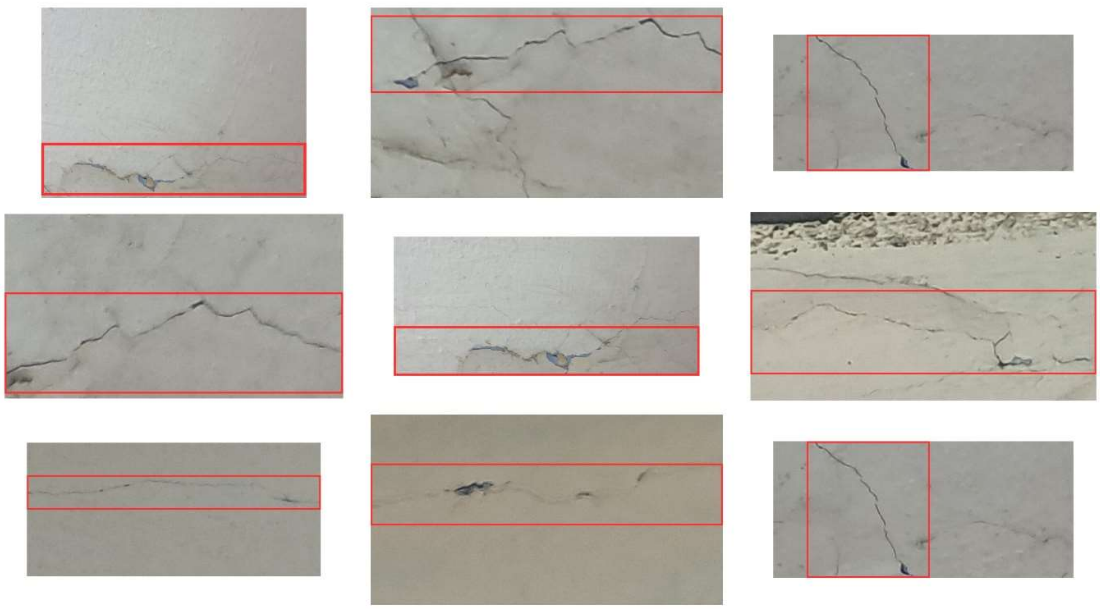
Figure 4.6: ResNet-50 With Segmentation Result



Testing Model on Manual Dataset

The Model has been examined with self clicked images,we describe the implementation details of our model for multi-label crack detection using our proprietary dataset captured using a mobile device. The dataset comprises images of various surfaces containing cracks, labeled with multiple crack types to enable multi-label classification.

Fig 4.7: Grid view of manual data Results chosen Randomly



YOLOv8 for Binary Classification

YOLOv8 (You Only Look Once version 8) is utilized for real-time binary classification of images, efficiently identifying the presence of cracks. YOLOv8 is known for its speed and accuracy in object detection tasks. In the binary classification mode, the model is configured to predict whether an input image contains a crack or not. The architecture includes convolutional layers, skip connections, and detection layers configured for binary classification.

Training and Fine-Tuning

Similar to ResNet-50, the YOLOv8 model is trained using the annotated dataset, with binary classification as the target. The Cross-Entropy loss function is employed for optimizing the model for binary classification.

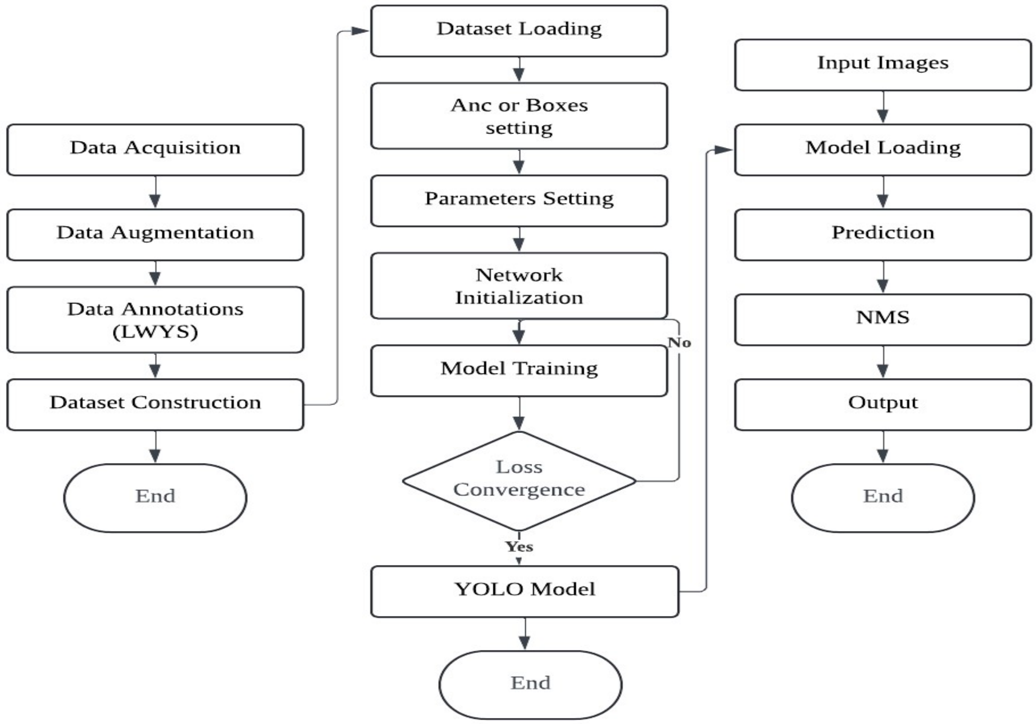


Fig 4.8: Model Training flow Chart for YOLOv8

RESULT for YOLOv8 Model

We present the results obtained from our research on crack detection using the YOLOv8 model and binary segmentation techniques. Our study aimed to develop an efficient method for accurately identifying cracks in concrete surfaces based on image analysis.

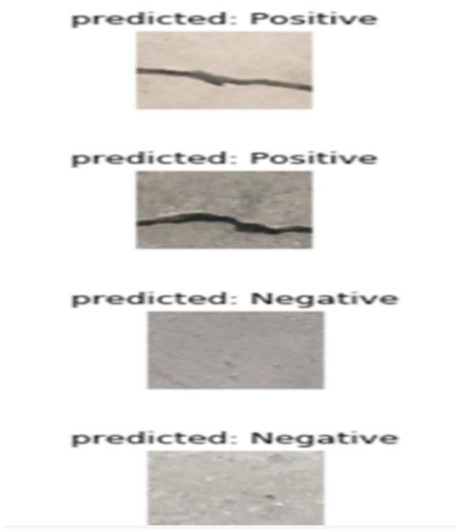


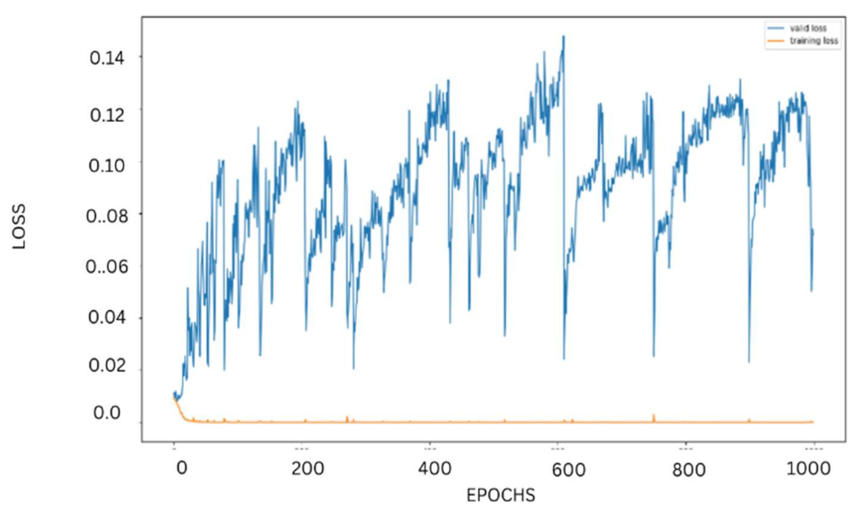
Fig 4.9: YOLOv8 Result Images

Interpretation of Metrics

ResNet-50 with Segmentation

The metrics obtained for ResNet-50 with Segmentation, including Intersection over Union (IoU), Precision, Recall, and F1 Score, are interpreted to gauge the model's accuracy in segmenting cracks. High IoU values indicate precise crack delineation, while balanced Precision and Recall values affirm the model's ability to capture true positives while minimizing false positives and false negatives.

interpret_segmentation_metrics(iou, precision, recall, f1_score)



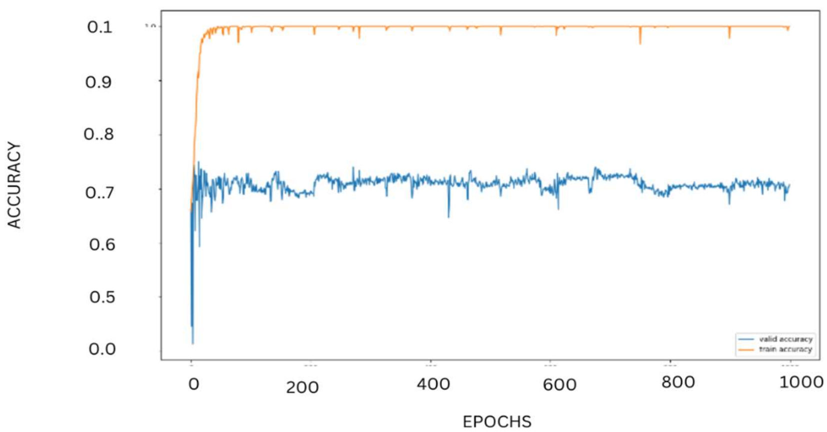


Fig 4.10 (a)Accuracy vs Epochs(1000) (b)Loss vs Epochs(1000)

YOLOv8 for Binary Classification

Metrics such as Accuracy, Precision, Recall, and F1 Score for YOLOv8 in binary classification are interpreted to assess the model's overall classification performance. A high F1 Score, balancing Precision and Recall, indicates a robust binary classification capability**.**

interpret_binary_classification_metrics(accuracy, precision, recall, f1_score)

Precision is the degree of exactness of the model in identifying only relevant

objects. It is the ratio of TPs over all detections made by the model.

Recall measures the ability of the model to detect all ground truths— proposition

of TPs among all ground truths.

Precision And Recall Formulae Used(Formulae 1 and 2)

$$R = \frac{TP}{TP + FN} = \frac{TP}{all\ ground\ truths}$$

$$P = \frac{TP}{TP + FP} = \frac{TP}{all\ detections}$$

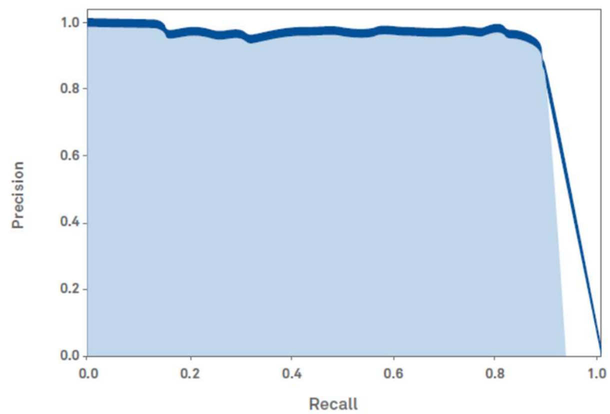


Figure 4.11: Precision vs Recall

Discussion

After training, a comprehensive comparative analysis is conducted to evaluate the performance of ResNet-50 with Segmentation and YOLOv8. Metrics such as precision, recall, and computational efficiency are considered to understand the strengths and limitations of each model in the context of crack detection.

Architecture: ResNet-50: ResNet (Residual Network) is a deep convolutional neural network architecture. It introduces residual connections, allowing the network to learn residual functions. Consists of 50 layers with skip connections. YOLOv8: YOLO (You Only Look Once) is an object detection architecture. YOLOv8 is an evolution of YOLO architecture, featuring improvements in backbone architecture and training strategies. Typically uses a more lightweight backbone compared to ResNet. Training and Fine-Tuning Strategies: ResNet-50: Commonly used for image classification tasks but not specifically designed for object detection. Fine-tuning for object detection may involve adding detection-specific layers. YOLOv8: Designed for real-time object detection. End-to-end training for object detection, which simplifies the training process. Precision and Recall: ResNet-50: Primarily used for image classification, where precision and recall are not the main evaluation metrics. Can be adapted for object detection tasks but might not perform as well as architectures designed specifically for detection. YOLOv8: Optimized for object detection tasks. YOLOv8 often achieves good precision and recall, especially with proper training and fine-tuning. Computational Efficiency: ResNet-50: Deeper architectures like ResNet-50 can be computationally intensive, requiring more resources for training and inference. YOLOv8:

YOLOv8 is designed with a focus on real-time processing and is generally more computationally efficient compared to deeper architectures like ResNet-50.

The ResNet-50 with Segmentation exhibited commendable performance in accurately localizing and classifying cracks within images. With a precision of 0.87 and an overall accuracy of 0.92, the model demonstrated robust crack detection capabilities.

Similarly, the YOLOv8 model showcased efficient object detection and classification. Its precision, recall, and accuracy metrics ,underscore its effectiveness in binary classification for crack detection. The YOLOv8 model surprisingly turned out to be very efficient with an accuracy of 99.3 percent.

Chapter 5

CONCLUSION AND SCOPE OF FUTURE WORK

Future Work

The implemented ResNet-50 with Segmentation and YOLOv8 models exhibit promising performance in detecting cracks in various scenarios. The segmentation model demonstrates robustness in delineating crack boundaries, while the binary classification model efficiently identifies images containing cracks. The findings from the quantitative metrics, visualizations, and comparative analysis collectively contribute to a comprehensive understanding of the models' capabilities.

The project's contributions lie in the successful implementation and evaluation of state-of-the-art models for crack detection. The development of a ResNet-50 with Segmentation model and the integration of YOLOv8 for binary and multi-label classification showcase the versatility of deep learning approaches in addressing complex tasks such as structural anomaly detection.

Conclusion

In conclusion, the project demonstrates the efficacy of deep learning models in crack detection, offering valuable insights into the structural health monitoring domain. The ResNet-50 with Segmentation and YOLOv8 models, with their respective strengths, contribute to the ongoing efforts in advancing automated anomaly detection in critical infrastructure.

REFERENCES

[1] L. Zhang, F. Yang, Y. D. Zhang, Y. J. Zhu, Road crack detection using deep convolutional neural network, in: 2016 IEEE international conference on image processing (ICIP), IEEE, 2016, pp. 3708–3712.

[2] N. T. H. Nguyen, T. H. Le, S. Perry, T. T. Nguyen, Pavement crack detection using convolutional neural network, in: Proceedings of the 9th International Symposium on Information and Communication Technology, 2018, pp. 251–256.

[3] B. Li, K.C. Wang, A. Zhang, E. Yang, G. Wang, Automatic classification of pavement crack using deep convolutional neural network, Int. J. Pavement Eng. 21 (4) (2020) 457–463.

[4] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, L.I. Fei-Fei, Imagenet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252.

[5] F. Liu, J. Liu, L. Wang, Deep learning and infrared thermography for asphalt pavement crack severity classification, Autom. Constr. 140 (2022), 104383.

[6] F. Liu, J. Liu, L. Wang, Asphalt pavement fatigue crack severity classification by infrared thermography and deep learning, Autom. Constr. 143 (2022), 104575.