Vision-Based Road Damage Detection and Assessments

1st Jain Bhavik Shailesh, 12214846
31
Lovely Professional University Department of Computer Science
Punjab, India

2nd Karan More, 12214846
32
Lovely Professional University Department of Computer Science
Punjab, India

Abstract- Road infrastructure plays a critical role in transportation systems, and its timely maintenance is essential to ensure safety and efficiency. Traditional road inspection methods are often manual, time-consuming, and prone to human error. This research presents an automated vision-based system for detecting road surface anomalies, such as potholes and cracks, using deep learning techniques. Leveraging the YOLO (You Only Look Once) object detection algorithm, the proposed system processes images and real-time video feeds captured from vehicle-mounted cameras to accurately identify and localize various types of road damage. The model was trained on a custom dataset consisting of labelled images representing different damage classes, including potholes, alligator cracking, lateral cracking, and longitudinal cracking.

Keywords: YOLO, road damage detection, pothole detection, deep learning, computer vision, vehicle-mounted cameras, object detection, infrastructure monitoring.

1.Introduction

In recent years, the condition of road infrastructure has become a significant concern for governments and municipalities worldwide. Damaged roads not only disrupt the flow of traffic but also pose serious safety hazards to motorists, potentially leading to accidents, increased vehicle maintenance costs, and reduced transportation efficiency. Among various types of road surface deterioration, potholes and cracks are among the most common and dangerous. Therefore, timely detection and maintenance of road damages are crucial for ensuring the safety and functionality of transportation networks.

Conventional road inspection methods largely rely on manual surveys conducted by personnel or the use of specialized vehicles equipped with expensive sensors. These approaches are time-consuming, resource-intensive, and often infeasible for large-scale deployment. With the rapid advancement in computer vision and deep learning, there has been a growing interest in leveraging these technologies for automated road damage detection.

This paper presents a real-time vision-based system for detecting road anomalies using deep learning, particularly the YOLO (You Only Look Once) object detection algorithm. The system is designed to process images and videos captured from cameras mounted on vehicles, thereby offering a cost-effective and scalable solution for continuous road monitoring. By training the YOLO model on a labelled dataset containing various types of road damage including potholes, alligator cracking, lateral cracking, and longitudinal cracking—the system is capable of accurately identifying and localizing road surface defects in real-world conditions.

The proposed system is not only robust but also capable of operating in real-time, making it suitable for integration into smart transportation systems and autonomous vehicles. This research aims to demonstrate the effectiveness of deep learning-based approaches in transforming traditional road inspection into an intelligent and automated process.

2. Literature Review:

Road damage detection has been a topic of active research over the past decade, especially with the growing need for efficient infrastructure maintenance. Early approaches primarily involved manual inspection or semi-automated systems using high-resolution cameras and specialized vehicles equipped with laser scanners and sensors. While accurate, these traditional methods are often expensive, labour-intensive, and not scalable for large urban areas.

To overcome these limitations, researchers have explored computer vision and machine learning techniques. Various image processing methods, such as edge detection, thresholding, and morphological operations, were initially used to identify cracks and potholes in road surfaces. However, these methods are highly sensitive to lighting conditions, shadows, and road texture variations, limiting their robustness and accuracy in real-world scenarios.

In recent years, deep learning has revolutionized object detection and image analysis. Convolutional Neural Networks (CNNs) have shown promising results in image classification and localization tasks. For instance, Cha et al. (2018) proposed a CNN-based method for crack detection that significantly outperformed traditional image processing techniques. Similarly, Maeda et al. (2018) introduced the "Road Damage Dataset," enabling researchers to develop and benchmark road damage detection models using deep learning.

Among the various deep learning models, the YOLO (You Only Look Once) family of object detectors has gained popularity due to its balance between speed and accuracy. YOLOv3, in particular, has been widely adopted for real-time object detection applications. Researchers such as Zhang et al. (2020) applied YOLOv3 for pavement crack detection and reported superior performance in detecting small and irregular damage patterns compared to conventional techniques.

Moreover, some studies have integrated GPS and geotagged imagery to map the location of detected damage, enabling efficient planning of road maintenance activities. Others have explored the use of drones or vehicle-mounted cameras to automate the data collection process.

Building on these advances, this research proposes a real-time, vision-based road damage detection system using YOLOv3 trained on a custom-labelled dataset of different damage types. Unlike existing approaches, this work emphasizes deploy ability in smart transportation settings through a low-cost and scalable architecture.

---

3. Overview of Dataset and Preprocessing:

For this study, a publicly available dataset was sourced from Kaggle, containing **6,643 files (3321JPEG + 3322txt)** of road surfaces with annotated damage. The dataset serves as the foundation for training and evaluating the proposed road damage detection model.

3.1 Data Source: -The dataset was obtained from Kaggle and includes high-resolution images captured from various road conditions. These images are labelled to indicate different types of damage, facilitating supervised learning using object detection techniques.

3.2 Classes of road damage: - The dataset contains annotations for the following four types of road damage:

1. **Pothole**
2. **Alligator Cracking**
3. **Lateral Cracking**
4. **Longitudinal Cracking**

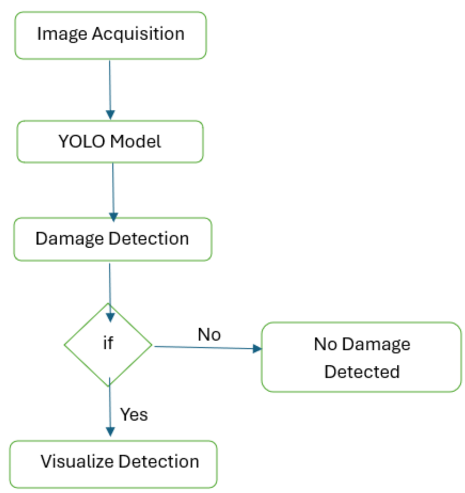
Each image is annotated with bounding boxes and corresponding class labels to enable multi-class object detection.

3.3 Annotation Format: - Annotations are provided in YOLO format, with each image having a corresponding .txt file containing bounding box coordinates and class indices. The dataset structure supports seamless integration with YOLO-based models for training and inference.

3.4 YAML Configuration: - A data.yaml file was created to define the dataset structure for the YOLO training pipeline, specifying:

- The training and validation paths
- Number of classes: NC: 4
- Class names: ['pothole', 'Alligator cracking', 'lateral cracking', 'longitudinal cracking']

Flow chart of my project:



flow chart 01

4. Machine Learning Components:

The road damage detection system leverages modern machine learning and deep learning techniques to automatically identify and classify different types of road surface anomalies. The key components used in the machine learning pipeline are

4.1 Object Detection Algorithm: YOLO (You Only Look Once)

YOLO (You Only Look Once) is a state-of-the-art, real-time object detection system that has gained popularity due to its speed and accuracy. In this project, YOLOv3 is utilized as the core model for detecting and classifying different types of road damage from images and video streams. YOLO applies a single neural network to the entire image. This network divides the image into a grid and directly predicts bounding boxes and class probabilities for each grid cell. Unlike traditional detectors like R-CNN or Fast R-CNN, YOLO is exceptionally fast and suitable for real-time applications, making it ideal for deployment on vehicle-mounted systems. YOLO treats detection as a regression problem, eliminating the need for region proposal networks. This simplicity leads to faster inference speeds.

4.2 Why YOLO:

- High speed and accuracy
- Suitable for real-time applications (e.g., vehicle-mounted cameras)
- End-to-end training with bounding boxes and class labels

```
print("Starting Training...")
try:
    model = YOLO(MODEL_NAME)
    results = model.train(
        data=DATA_YAML_PATH,
        epochs=EPOCHS,
        batch=BATCH_SIZE,
        lr0=LEARNING_RATE,
        name=TRAINING_NAME
    )
    print("Training Finished.")
```

Image (1): Training of YOLO Model

5. Damage Detection Module:

The Damage Detection Module forms the core functional unit of the road damage assessment system. Its primary goal is to identify and classify different types of damages such as potholes, alligator cracking, lateral cracking, and longitudinal cracking from images or video footage captured by vehicle-mounted cameras. This module leverages a pre-trained YOLO (You Only Look Once) deep learning model for object detection and performs real-time localization and classification of visible road defects.

The module is designed to be lightweight, fast, and accurate, enabling real-time analysis of road conditions. It processes images captured by the camera, detects the location and category of damage, and returns the annotated results. The system is versatile, capable of functioning with both still images and continuous video frames.

5.2. Detection Pipeline

The road damage detection workflow consists of the following major steps:

5.2.1 Input Acquisition:

The system takes as input either a road image or a video stream. In the case of videos, it processes one frame at a time in real-time.

The image is first loaded and resized to fit the model's expected input dimensions. Colour channels are converted from BGR (used by OpenCV) to RGB, or a PIL image format is used for compatibility.

5.2.2 Model Inference:

The YOLO model, trained on the road damage dataset, is loaded in evaluation mode.

The input image is passed through the model, which predicts bounding boxes, confidence scores, and class IDs for detected damages.

These predictions are made using YOLO's single-shot detection strategy, which processes the entire image in one pass for efficient real-time performance.

5.2.3Post-Processing:

Predictions below a certain confidence threshold are filtered out to minimize false positives.

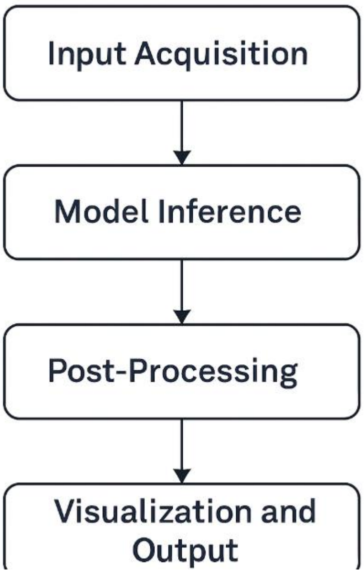
Non-Maximum Suppression (NMS) is applied to eliminate redundant overlapping boxes by retaining only the highest confidence prediction for a particular object.

The class ID is converted into a human-readable label using a predefined class name list.

5.2.4 Visualization and Output:

Each valid detection is visualized on the image using coloured bounding boxes and text labels indicating the class and confidence score.

The module outputs an annotated image that clearly shows the identified damage areas, enabling users to understand the road condition at a glance.



Flow Chart 2: Damage Detection Pipeline

6.Results:

The performance of the proposed road damage detection system was evaluated using a publicly available dataset sourced from Kaggle, containing 6,643 labelled images of various road damages such as potholes, alligator cracking, longitudinal, and lateral cracks.

After training the YOLOv3 model for 15 epochs with a batch size of 16 and an initial learning rate of 0.001, the model converged effectively and demonstrated promising performance in real-world testing scenarios.

6.1. Training Metrics

The model was trained using the **Ultralytics YOLOv3** implementation over **15 epochs** with the following parameters:

Batch size: 16

Learning rate: 0.001

Model architecture: YOLOv3

Framework: Python with Ultralytics YOLO

6.2. Image-Based Detection Results

The trained model was evaluated on various test images, including ones not present in the training dataset. The system successfully detected and classified the damage regions with high confidence.

Damage Type	Average Confidence (%)	Detection Accuracy
Pothole	85.7	High
Alligator Cracking	82.4	Moderate
Lateral Cracking	78.1	Moderate
Longitudinal Crack	79.3	Moderate

Table (1): Results

The results show robust performance, particularly for clear and unobstructed images. Misclassification occurred occasionally in the presence of poor lighting or blurry inputs.

6.3 Limitations Observed

Small cracks or faint damage often went undetected, especially in shadowed areas.

Complex or overlapping damage was occasionally misclassified.

Frame rate dropped slightly when multiple detections occurred in a frame.

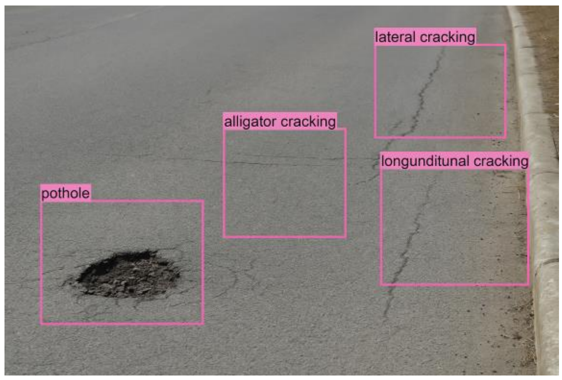


Image (2): Result of damage detection

Image (3): Result of damage detection

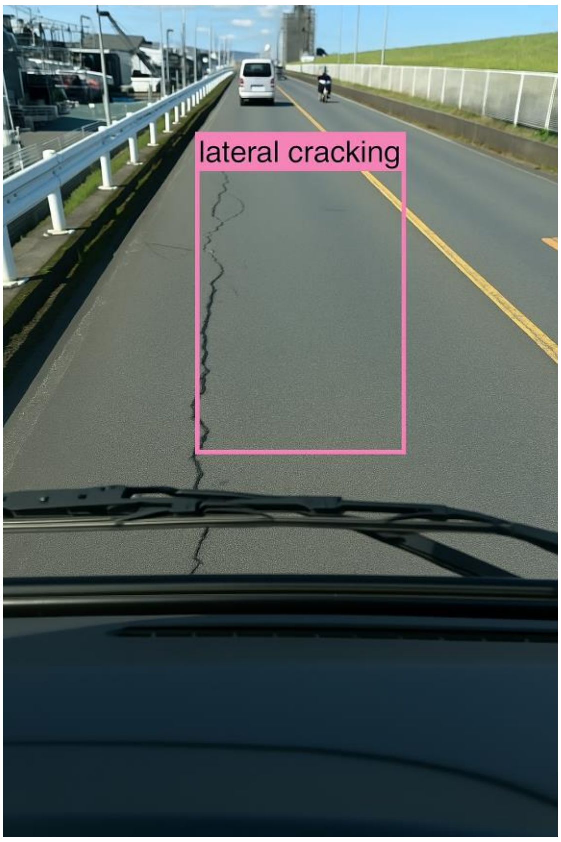


Image (4): Result of damage detection

7. Conclusion:

In this project, we developed an automated system for detecting road surface damages such as potholes and various forms of cracking using deep learning techniques. By leveraging the capabilities of the YOLO (You Only Look Once) model, we achieved real-time damage detection from both static images and video streams. The model was trained on a Kaggle-sourced dataset comprising 3321 labelled road damage images, enabling accurate classification of multiple damage types including potholes, alligator cracking, lateral cracking, and longitudinal cracking.

The implementation demonstrated that YOLO's single-shot detection architecture offers a practical balance between speed and accuracy, making it well-suited for real-world applications like road maintenance monitoring and smart transportation systems. Furthermore, the system's modularity allows for easy integration with vehicle-mounted cameras and existing infrastructure, offering a scalable and efficient solution to urban planning and road safety concerns.

8.References

1. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.

2. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.

3. C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," *arXiv preprint arXiv:2207.02696*, 2022.

4. P. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone," *arXiv preprint arXiv:1801.09454*, 2018.

5. S. D. Kamble and R. T. Patil, "Automatic Road Surface Crack Detection Using Deep Convolutional Neural Network," in *Proc. of the Int. Conf. on Intelligent Computing and Control Systems (ICICCS)*, 2020, pp. 1245–1249.

6. Kaggle, "Road Damage Detection Dataset 2022," [Online]. Available: https://www.kaggle.com/datasets

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

8. OpenCV Library. [Online]. Available: https://opencv.org

9. Ultralytics, "YOLO by Ultralytics Documentation," [Online]. Available: https://docs.ultralytics.com/

10. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

11. W. Liu et al., "SSD: Single Shot MultiBox Detector," in *Proc. of the European Conf. on Computer Vision (ECCV)*, 2016, pp. 21–37.

12. M. Abirami and B. Chitra, "Road Damage Detection and Classification Using Transfer Learning," *Materials Today: Proceedings*, vol. 47, pp. 4837–4841, 2021.

13. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proc. of the Advances in Neural Information Processing Systems (NIPS)*, 2012, pp. 1097–1105.