Joint CSCE Construction Specialty & CRC Conference 2025
Conférence conjointe spécialisée en construction de la SCGC et CRC-2025

Montreal, Quebec
July 28-31, 2025 / *28-31 juillet 2025*

ADVANCING CULVERT INSPECTION AUTOMATION USING DEEP LEARNING AND DATASET OPTIMIZATION

P. Mohammadi[1], S. Asgari[2] and A. Rashidi[1]

[1] Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, UT, USA
[2] Department of Civil and Environmental Engineering, Merrimack College, North Andover, MA, USA

ABSTRACT: The Utah Department of Transportation (UDOT) manages over 120,000 culverts and storm drain pipes critical to the state's transportation infrastructure. Manual inspection methods for culverts, relying on visual assessments, are time-intensive, inconsistent, and prone to human error. This paper explored the use of advanced computer vision techniques to automate culvert inspections. Specifically, we employed YOLOv11, known for its speed, efficiency, and real-time capabilities in computer vision tasks. We collected a diverse dataset of culvert inspection videos and images from UDOT's (Utah Department of Transportation) records, incorporating various data sources such as CCTV and zoom camera footage. Initially, the dataset faced significant class imbalance, leading to suboptimal model performance. To address this issue, we refined the data by removing redundant frames and applying data augmentation techniques. The final dataset of 12,320 labeled images enabled the YOLOv11 model to achieve an mAP of 75%, a precision of 79.5%, a recall of 67.9%, and an F1 score of 73%. Our study highlights the potential of intelligent inspection systems to transform culvert maintenance by improving inspection efficiency, accuracy, and reliability. By integrating such systems with the existing asset management platforms, UDOT can better allocate resources, mitigate risks, and enhance the resilience of its transportation network. Future work includes expanding dataset diversity by incorporating underrepresented defect categories and employing active learning for improved accuracy.

1. INTRODUCTION

Culverts are critical components of transportation infrastructure, facilitating water flow beneath roadways, railways, and other pathways to prevent flooding, soil erosion, and structural damage. Like many infrastructure assets, culverts are aging, and their deterioration poses significant risks to roadway safety and functionality (Stoner et al., 2019). For example, the Utah Department of Transportation (UDOT) manages over 120,000 culverts and storm drain pipes, which are vital to the state's transportation network. Ensuring their integrity through timely inspection and maintenance is essential for preventing failures that could lead to costly repairs, road closures, and environmental hazards (Mohammadi et al., 2024).

Regular inspection of culverts is a cornerstone of effective asset management, allowing transportation agencies to detect defects early and implement preventive maintenance strategies. Traditional culvert inspections rely primarily on manual visual assessments, often conducted through digital video inspections. However, these methods are time-intensive, inconsistent, and susceptible to human error due to variations in inspector experience and judgment (Youngblood & Smith, 2017). Given the scale of culvert networks like

UDOT's, manual inspection approaches struggle to provide the efficiency and accuracy required for proactive maintenance planning.

To address these limitations, this study explores the potential of automating culvert inspections using advanced computer vision and deep learning techniques. Specifically, we leverage YOLOv11, a state-of-the-art object detection model known for its real-time processing capabilities and efficiency in identifying structural defects. Our approach integrates a diverse dataset of culvert inspection videos and images from UDOT's records, including footage from CCTV (Closed-Circuit Television) and zoom cameras. One of the primary challenges encountered was the class imbalance within the dataset, where corrosion defects were disproportionately represented. To enhance model generalization and performance, we refined the dataset through frame selection, data augmentation, and label consolidation, ultimately achieving a significant improvement in defect detection accuracy.

This research contributes to the growing body of work on infrastructure inspection automation by demonstrating the applicability of deep learning in culvert defect detection. While previous studies primarily focus on sewer pipelines or general pavement inspection, our study tailors deep learning methodologies to culvert assessment, addressing specific challenges such as dataset imbalance and defect classification optimization. By refining the dataset and leveraging data augmentation techniques, we improve upon prior efforts in automated damage detection of culverts, paving the way for more reliable and scalable inspection frameworks. By integrating automated inspection methods with existing asset management systems, UDOT and similar agencies can enhance culvert maintenance efficiency, minimize risk, and optimize resource allocation. This study represents a significant step toward transforming traditional culvert inspection practices through intelligent, data-driven solutions.

The remainder of our paper is organized as follows: Background provides a review of existing work in damage detection and condition assessment using computer vision and deep learning. Methodology details our proposed approach, including the use of dataset optimization strategies and deep learning techniques. Results & Discussion presents the performance evaluation of our model, highlighting improvements achieved through dataset refinement. Conclusion summarizes our findings, discusses limitations, and outlines potential directions for future research.

2. BACKGROUND

The assessment and maintenance of infrastructure assets, including culverts, pipelines, and storm drains, have traditionally relied on manual inspections. These inspections often involve on-site visual assessments or digital video evaluations, both of which are labor-intensive and susceptible to inconsistencies due to human subjectivity (Meegoda et al., 2019). With the increasing need for efficient and reliable inspection methods, researchers have turned to artificial intelligence, particularly machine learning (ML) and computer vision, to automate damage detection and condition assessment. Recent advancements in deep learning have enabled automated systems to analyze inspection videos and images with high accuracy, reducing the reliance on manual interpretation (Zuo et al., 2025).

Hawari et al., (2018) developed an automated defect detection tool for sewer pipelines using a combination of image-processing algorithms and segmentation tools applied to CCTV footage. Their study assessed four types of defects: cracks, settled deposits, ovality, and displaced joints. The model performed well in detecting ovality but struggled with other defects, highlighting the need for larger and more diverse datasets to improve performance. (Yin et al., 2020) applied deep learning for real-time sewer pipe defect detection, leveraging the YOLOv3 object detection model. They trained their model on a dataset containing 3,664 images extracted from 63 CCTV videos, encompassing six defect types: holes, fractures, cracks, deposits, broken sections, and root intrusions. The model achieved an 85.37% mean average precision (mAP) and F1 scores of 87.6% and 88.2% for testing and validation sets, respectively. This study demonstrated the effectiveness of YOLO-based models for automated damage detection. Kumar et al., (2020) compared the performance of Faster R-CNN, YOLOv3, and SSD (Single Shot Detector) for detecting sewer pipe defects such as root intrusions and deposits. Their findings indicated that Faster R-CNN achieved the highest accuracy but required longer processing times, whereas YOLOv3 provided a more balanced trade-off

between speed and accuracy, making it more suitable for real-time applications. In a subsequent study, Yin et al., (2021) advanced sewer pipe assessment automation by developing a Video Interpretation Algorithm for Sewer Pipes (VIASP). This system integrated YOLO-based defect detection with text recognition algorithms to identify defect locations and generate structured assessment reports. Their testing results showed that VIASP could achieve an F1 score of 75%, demonstrating its potential to automate the defect interpretation process.

These studies collectively illustrate the increasing adoption of deep learning for pipeline inspection, enabling faster and more consistent defect detection compared to manual methods. However, while extensive research has been conducted on sewer pipelines, relatively few studies have focused on the automated inspection of culverts. Culverts share many similarities with sewer pipelines but present unique challenges due to their varying structural materials (metal, concrete, plastic), environmental exposure, and diverse defect types. Despite their critical role in transportation infrastructure, deep learning applications for culvert inspections remain underexplored. Recent research has attempted to bridge this gap. Meegoda et al. (2019) introduced 360-degree camera technology for culvert inspections, allowing for improved defect visualization. However, their approach still required manual review of inspection footage, limiting its scalability for large transportation agencies such as the UDOT.

Our study builds upon previous research by tailoring YOLOv11 for culvert defect detection. While past studies primarily focus on sewer pipelines or bridges, our work is among the first to develop a deep learning-based automated inspection framework specifically for culverts. Key contributions of our study include: (1) Dataset Optimization: We address the issue of class imbalance in culvert defect datasets by refining data selection, applying augmentation techniques, and consolidating defect labels into five main categories (corrosion, deformation, cracks, break/collapse, and joint issues). (2) Model Selection and Performance Improvements: By employing YOLOv11, we extend   previous applications of machine learning in culvert inspection   processes in Utah by achieving high accuracy and real-time inference capabilities.

3. METHODOLOGY

This study aims to develop an automated culvert inspection framework utilizing deep learning and computer vision techniques. Our approach leverages YOLOv11, a state-of-the-art object detection model, to detect structural defects in culverts from CCTV and zoom camera inspection videos provided by UDOT. The methodology consists of four key components: data collection, data preprocessing and annotation, model development and training, and performance evaluation.

3.1   Data Collection

The dataset used for this study was sourced from UDOT's existing culvert inspection records, comprising zoom camera and CCTV inspection videos, along with additional culvert images collected from field inspections. Also, we included images of sewer and storm pipes in the dataset to help address the imbalance in the culvert dataset provided by UDOT. The data collection process included:

- Zoom camera inspection videos: This dataset included over 2,000 video files captured using telescopic pole-mounted cameras without requiring prior cleaning of the culverts, with 80% showing no structural defects.
- CCTV inspection videos: This dataset included 197 videos collected from culverts along the 1-80 highway. Among these, only 66 contained visible structural defects. The defect rating system used in these videos was converted from the old four-digit UDOT rating scale to the new five-point rating scale for consistency.
- Culvert images from field inspections: UDOT personnel contributed an additional 450 images of culverts, enhancing dataset diversity. There were only 197 images with structural defects among them. These images were manually labeled according to UDOT's hybrid condition rating system.
- Sewer and storm pipe images: This dataset contained over 2,000 images of sewer and storm pipes exhibiting structural defects. The images were captured using CCTV cameras. These images were included in our dataset because they primarily feature concrete pipes similar in structure to culverts.

This addition helped address the issue of limited images of concrete pipe structural defects in the UDOT dataset.

3.2   Data Preprocessing and Annotation

Given the high volume of raw data, preprocessing was essential to ensure balanced, high-quality input for deep learning training. The following steps were performed:

- Frame extraction and filtering: Videos were converted into static image frames, with highly similar frames removed to reduce redundancy. This process reduced the dataset to 3,722 labeled images.
- Dataset imbalance adjustment: The initial dataset was dominated by corrosion defects, leading to poor model generalization. To improve data balance, we implemented data augmentation techniques such as rotation, exposure adjustment, blurring, and noise addition.
- Label consolidation: The original dataset contained 20 defect labels, but due to the underrepresentation of some classes, these were consolidated into five main categories: corrosion, deformation, crack/fracture, break/collapse, and joint issues. This approach enhanced model learning by ensuring a sufficient number of images for each defect type.
- Manual annotation with Computer Vision Annotation Tool (CVAT): The CVAT (Sekachev et al., 2020) was used to label the dataset with bounding boxes around defects. Bounding box coordinates and class labels were exported in a format compatible with YOLOv11.

3.3   Model Development

For defect detection, we selected YOLOv11, the latest version of the You Only Look Once (YOLO) object detection model, known for its high-speed, real-time detection capabilities. The model architecture consists of Convolutional Neural Networks (CNNs) for extracting key defect features from images, advanced feature extraction layers to enhance small-object detection and improve accuracy in complex culvert environments, and a refined anchor box mechanism that dynamically adjusts to different defect scales and aspect ratios, ensuring more precise localization of structural issues (Jocher et al., 2023).

The model was trained using a holdout cross-validation approach, where the dataset was divided into: 70% training set, 20% validation set, and 10% test set. To optimize model learning, we employed transfer learning, where the model was first pre-trained on large, generic datasets before being fine-tuned on the UDOT dataset to enhance defect detection capabilities. Additionally, batch normalization and data augmentation techniques were applied to improve generalization and prevent overfitting, ensuring the model could effectively recognize various defect types under different conditions. Finally, we implemented adaptive learning rate scheduling, dynamically adjusting the learning rate during training to facilitate efficient model convergence and achieve optimal performance.

3.4   Model Performance Evaluation

To assess the model's effectiveness, we evaluated its performance using industry-standard metrics for object detection. Precision was measured as the proportion of correctly predicted defects among all defect predictions, while recall represented the proportion of actual defects that were correctly detected (Mohammadi et al., 2023). Additionally, we utilized Mean Average Precision (mAP) to assess accuracy across different Intersection over Union (IoU) thresholds, providing a comprehensive evaluation of the model's detection capabilities (Zhang et al., 2023). Finally, confusion matrices were employed to visualize classification performance across various defect types, offering insights into the model's strengths and areas for improvement (Mashhadi et al., 2022).

4. RESULTS & DISCUSSION

This study aimed to enhance the inspection process of culverts by leveraging advanced computer vision techniques, particularly through the use of deep learning. We developed a YOLOv11 model trained on a

diverse dataset consisting of images of culverts, along with additional storm and sewer pipe images. The inclusion of these additional images was critical in addressing the class imbalance observed in the original dataset provided by UDOT. A major challenge we faced was that corrosion defects were overrepresented, leading to poor model performance when trying to detect other types of defects, such as cracks and joint issues. To mitigate this issue, we applied several data augmentation techniques, such as rotation, noise addition, exposure adjustments, and blurring, which helped diversify the dataset and balance the representation of defect types.

As a result of these augmentation techniques, the number of images in the dataset increased substantially, from 3,722 to 12,320 images. This expansion provided the model with more varied examples, helping it generalize better across different types of defects and culvert conditions. These efforts were crucial in improving the model's ability to detect less common defects with higher accuracy. The performance of the YOLOv11 model was evaluated based on standard object detection metrics such as precision, recall, and Mean Average Precision (mAP). Table 1 shows the results for two datasets: one with augmentations and one without augmentations.

Reviewing the results presented in Table 1, it is evident that the performance of the YOLOv11 model improved significantly after applying augmentation techniques to the training set. These enhancements demonstrate the model's increased ability to generalize across a wider range of defect types. The increase in precision indicates that the model was more successful at correctly identifying defects, while the improvement in recall demonstrates a better detection rate for all true defects, reducing the chances of missed detections. The substantial increase in mAP further validates the effectiveness of the augmented dataset, indicating that the model is now capable of more accurately detecting a variety of culvert defects.

Table 1: Performance of the YOLOv11 model on the dataset with and without augmentations

Model	Number of Classes	Dataset Size	Precision (%)	Recall (%)	F1 Score (%)	mAP (%)
YOLOv11 + Augmentation	5	12320	79.5	67.9	73	75
YOLOv11	5	3722	70.4	56.7	62.8	60.6

These results underscore the importance of dataset optimization for training deep learning models, especially in the context of infrastructure inspection. By addressing the class imbalance and expanding the dataset, we significantly enhanced the model's ability to detect common defects like corrosion and rarer defects like cracks and joint issues. This improvement is critical for automating the inspection of culverts, as it allows for more accurate and efficient identification of structural issues, ultimately reducing human error and speeding up the inspection process.

Table 2 presents a comparison of the developed models' performance across various defect classes. For YOLO + Augmentation, the model shows notable improvements in precision, recall, and mAP across most classes compared to the standard YOLO model. Specifically, the Break/Hole/Collapse/Kink class, with YOLO + Augmentation, achieves a mAP of 78.5%, significantly higher than the 52.4% mAP achieved by the standard YOLO model. Similarly, for the Joints class, YOLO + Augmentation shows a substantial increase in performance, reaching an mAP of 76.2%, compared to 54% with the standard YOLO. The Deformation/Shape class also benefits from augmentation, with YOLO + Augmentation achieving an mAP of 89.8%, outperforming the 91% achieved by YOLO, which already had a strong baseline. While the Surface Damage Corrosion class still shows relatively lower performance, augmentation helps boost the mAP from 47.6% to 57.5% compared to the normal dataset. Overall, data augmentation has had a clear positive impact, especially on underrepresented classes, improving the model's ability to generalize and detect distress types more accurately.

Table 2: Performance of the developed models across different defect classes

| Model	Class	Number of Images in Test Set	Precision (%)	Recall (%)	F1 Score (%)	mAP (%)
YOLOv11 + Augmentation	Break/Hole/Collapse/ Kink	27	81.8	71.1	76.3	78.5
	Crack/Fracture	158	79.1	66.7	72.1	72.1
	Deformation/Shape	37	94.1	84.2	88.9	89.8
	Joints	120	80.7	65.2	72.1	76.2
	Surface Damage Corrosion	178	61.2	52.8	56.5	57.5
YOLOv11	Break/Hole/Collapse/ Kink	21	71.4	46.4	56	52.4
	Crack/Fracture	135	64.6	56.8	60.5	57.8
	Deformation/Shape	38	94.3	84.5	89.4	91
	Joints	74	62	49.8	55	54
	Surface Damage Corrosion	153	59.8	46	56.5	47.6
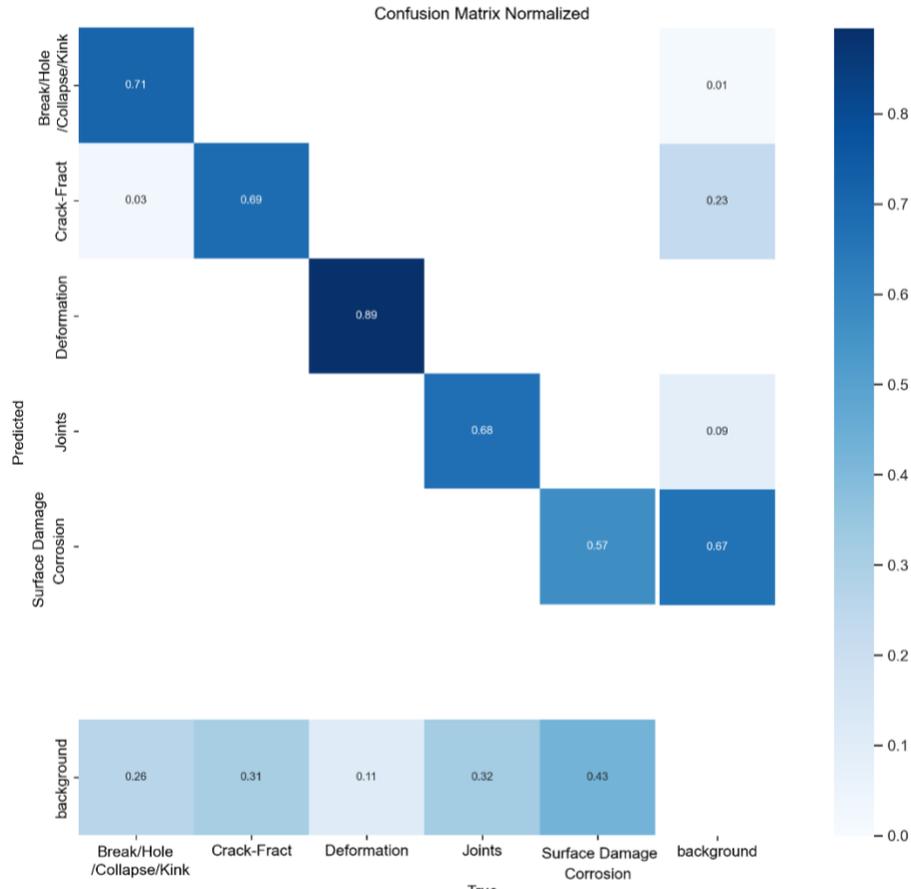
Figure 1: Confusion matrix of the YOLOv11 model using augmented data

The confusion matrix results from the YOLOv11 model, presented in Figure 1, provide valuable insights into the model's performance across different defect types. The model performed particularly well in detecting Surface Damage Corrosion, achieving high precision and recall, as reflected by the large values along the diagonal for this class. Similarly, the Break/Hole/Collapse/Kink defects were accurately detected,

with the model showing strong performance in this category. However, the model's performance was less satisfactory for Surface Damage Corrosion and Crack/Fracture defects, as evidenced by the lower values along the diagonal for these categories, indicating that these defects were more challenging for the model to identify. Additionally, the background class, which represents areas without defects, showed significant confusion with other defect classes, suggesting that the model occasionally misclassified defect-free areas as defects (as shown in Figure 2). This highlights the need for further model refinement, particularly in improving the detection of less obvious defect types and minimizing background misclassification.
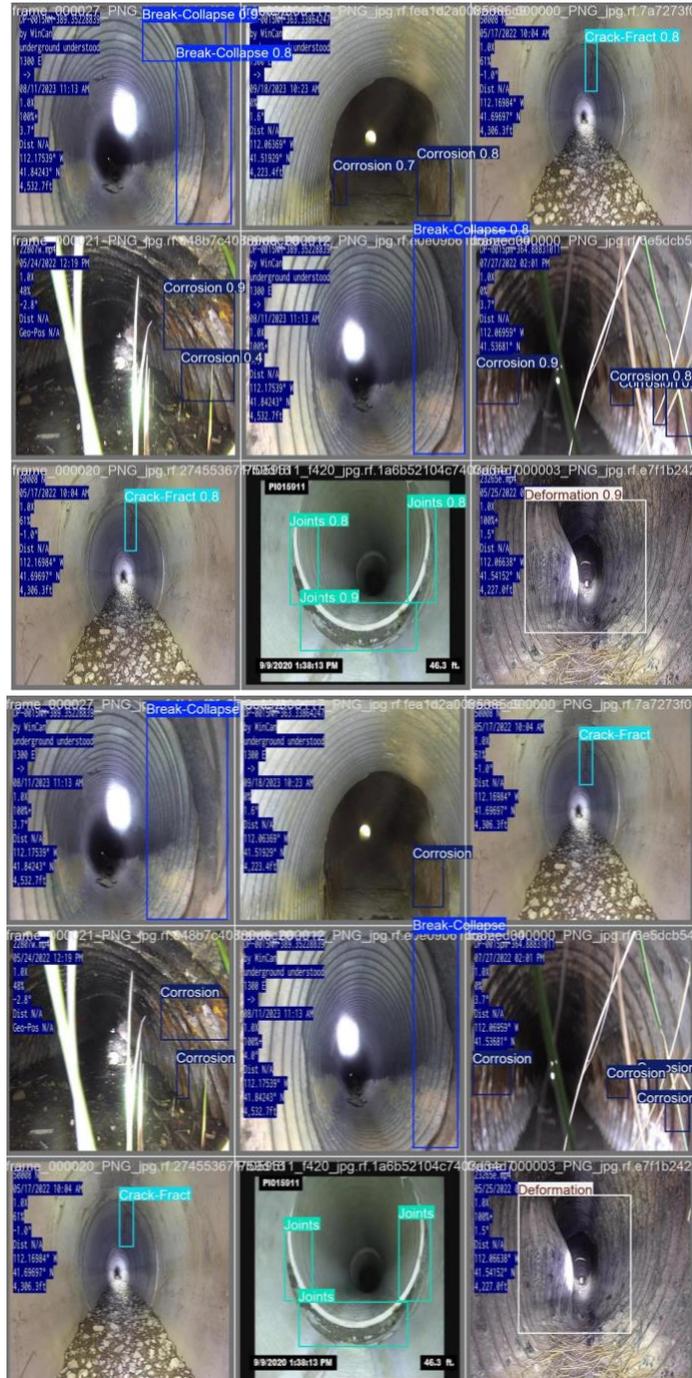

Figure 2 - Examples of model predictions (top images) and true labels (bottom images) for a batch of testing set

To address the limitations, we are actively working to collect additional training data specifically focused on these underrepresented defect types. Additionally, we plan to explore more advanced ML techniques, such as active learning and federated learning, to improve the model's ability to generalize and detect these defects more reliably. By refining the dataset further and utilizing state-of-the-art ML techniques, we aim to enhance the model's performance and achieve more accurate, real-world results for automated culvert inspections.

5. CONCLUSIONS

This study demonstrates the effectiveness of using YOLOv11 in combination with dataset optimization techniques to automate and improve the culvert inspection process. By addressing the challenges of class imbalance and data variability, we significantly enhanced the model's ability to detect various structural defects in culverts, such as corrosion, cracks, deformation, and joint issues. The application of data augmentation and label consolidation played a pivotal role in improving the model's performance, particularly in detecting underrepresented defect types. The YOLOv11 model achieved impressive precision and recall, as well as high Mean Average Precision (mAP), highlighting its suitability for real-time infrastructure inspection tasks.

The findings of this study validate the use of deep learning for automated culvert inspections, showcasing its potential to increase both efficiency and accuracy compared to   manual inspection methods. Integration of this technology into existing infrastructure management systems—such as those used by UDOT—   can significantly streamline inspection workflows, reduce operational costs, and improve decision-making regarding maintenance priorities.

However, there remain challenges, particularly with the detection of subtle defects like cracks and joint issues, which require further refinement of the model. The primary limitation of this study was the imbalance in the dataset, especially regarding certain defect classes. We are actively collaborating with UDOT to collect more diverse data and are exploring additional data sources to address this issue. By expanding our dataset and improving data quality, we aim to enhance the model's performance in detecting underrepresented classes. Future work will focus on expanding the dataset, incorporating higher-quality images, and exploring advanced ML techniques to enhance detection performance. This research marks a significant step toward transforming infrastructure maintenance through intelligent, automated solutions that can help improve the longevity and safety of critical transportation assets.

REFERENCES

Hawari, A., Alamin, M., Alkadour, F., Elmasry, M., & Zayed, T. (2018). Automated defect detection tool for closed circuit television (cctv) inspected sewer pipelines. *Automation in Construction*, *89*, 99–109.

Jocher, G., Chaurasia, A., & Qiu, J. (2023). *YOLO by Ultralytics*.

Kumar, S. S., Wang, M., Abraham, D. M., Jahanshahi, M. R., Iseley, T., & Cheng, J. C. P. (2020). Deep learning–based automated detection of sewer defects in CCTV videos. *Journal of Computing in Civil Engineering*, *34*(1), 4019047.

Mashhadi, A. H., Markovic, N., & Rashidi, A. (2022). Estimating Construction Work Zones Capacity Using Deep Neural Network. *Construction Research Congress 2022*, 98–107.

Meegoda, J. N., Kewalramani, J. A., & Saravanan, A. (2019). Adapting 360-Degree Cameras for Culvert Inspection: Case Study. *Journal of Pipeline Systems Engineering and Practice*, *10*(1).

https://doi.org/10.1061/(asce)ps.1949-1204.0000352

Mohammadi, P., Rashidi, A., & Asgari, S. (2024). Privacy-preserving culvert predictive models: A federated learning approach. *Advanced Engineering Informatics*, *61*, 102483.

Mohammadi, P., Rashidi, A., Malekzadeh, M., & Tiwari, S. (2023). Evaluating various machine learning algorithms for automated inspection of culverts. *Engineering Analysis with Boundary Elements*, *148*, 366–375. https://doi.org/https://doi.org/10.1016/j.enganabound.2023.01.007

Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., Hoff, B., TOsmanov, Kruchinin, D., Zankevich, A., DmitriySidnev, Markelov, M., Johannes222, Chenuet, M., a-andre, telenachos, Melnikov, A., Kim, J., Ilouz, L., Glazov, N., … Truong, T. (2020). *opencv/cvat: v1.1.0*. Zenodo. https://doi.org/10.5281/zenodo.4009388

Stoner, M., Pang, W., & Piratla, K. (2019). Predicting Culvert Deterioration Using Physical and Environmental Time-Independent Variables. *Journal of Pipeline Systems Engineering and Practice*, *10*(4), 04019035. https://doi.org/10.1061/(asce)ps.1949-1204.0000409

Yin, X., Chen, Y., Bouferguene, A., Zaman, H., Al-Hussein, M., & Kurach, L. (2020). A deep learning-based framework for an automated defect detection system for sewer pipes. *Automation in Construction*, *109*, 102967.

Yin, X., Ma, T., Bouferguene, A., & Al-Hussein, M. (2021). Automation for sewer pipe assessment: CCTV video interpretation algorithm and sewer pipe video assessment (SPVA) system development. *Automation in Construction*, *125*, 103622.

Youngblood, D., & Smith, C. D. M. (2017). *Enhanced culvert inspections-best practices guidebook*. Minnesota. Dept. of Transportation. Research Services & Library.

Zhang, L., Ding, G., Li, C., & Li, D. (2023). DCF-Yolov8: an improved algorithm for aggregating low-level features to detect agricultural pests and diseases. *Agronomy*, *13*(8), 2012.

Zuo, X., Sheng, Y., Shen, J., & Shan, Y. (2025). Multilabel Sewer Pipe Defect Recognition with Mask Attention Feature Enhancement and Label Correlation Learning. *Journal of Computing in Civil Engineering*, *39*(1), 4024050.