RESEARCH

UDOT

Keeping Utah Moving

Utah Department of Transportation - Research Division
4501 South 2700 West - P.O. Box 148410 - SLC,UT 84114-8410

Report No. UT-25.20

AUTOMATED INTERPRETATION OF CULVERT INSPECTION VIDEOS USING AI AND COMPUTER VISION

Prepared For:

Utah Department of Transportation
Research & Innovation Division

Final Report
August 2025

DISCLAIMER

The authors alone are responsible for the preparation and accuracy of the information, data, analysis, discussions, recommendations, and conclusions presented herein. The contents do not necessarily reflect the views, opinions, endorsements, or policies of the Utah Department of Transportation or the U.S. Department of Transportation. The Utah Department of Transportation makes no representation or warranty of any kind, and assumes no liability therefore.

ACKNOWLEDGMENTS

The authors acknowledge the Utah Department of Transportation (UDOT) for funding this research, and the following individuals from UDOT on the Technical Advisory Committee for helping to guide the research:

- Abdul Wakil
- Brad Loveless
- Kevin Nichol
- Greg Merrill
- Sean Berry
- Keith Meinhardt
- Chris Whipple
- Jeff Erdman
- Brandon Cox

The authors also thank Dr. Khalid Kaddoura and his former company, AECOM, for providing data that was essential to this research.

TECHNICAL REPORT ABSTRACT

1. Report No. UT-25.20	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A
4. Title and Subtitle Automated Interpretation of Culvert Inspection Videos Using AI and Computer Vision		5. Report Date July 2024 6. Performing Organization Code
7. Author(s) Pouria Mohammadi, Abbas Rashidi		8. Performing Organization Report No.
9. Performing Organization Name and Address The University of Utah Department of Civil and Environmental Engineering 201 Presidents Circle Salt Lake City, Utah 84112		10. Work Unit No. 5H094 24H 11. Contract or Grant No. 24-8333
12. Sponsoring Agency Name and Address Utah Department of Transportation 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410		13. Type of Report & Period Covered Final August 2023 to August 2025 14. Sponsoring Agency Code UT-25.20
15. Supplementary Notes Prepared in cooperation with the Utah Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration		

16. Abstract

Culvert assets play a critical role in ensuring the safe operation of highways. UDOT maintains over 120,000 drainage culverts and storm drain pipes along state highways. To maintain these assets optimally and prevent failures, UDOT must collect comprehensive information about all culverts across the state and inventory them in the ATOM system. Accurate identification of culverts needing repair, rehabilitation, or replacement necessitates thorough and well-documented inspections. Traditional culvert inspection is slow and prone to subjectivity, leading to inconsistencies in the assessment of culvert conditions. This project focused on automating the interpretation of culvert inspection videos using advanced computer vision and deep learning techniques. Beginning with a small and imbalanced dataset, the team expanded the data through additional data collection and augmentation, followed by manual labeling and annotation of structural defects. Three model types were developed to support different stages of inspection analysis: a binary classification model to identify defective frames, multiclass image classification models to classify five major defect types, and an object detection model capable of localizing and classifying defects. To bridge the gap between model output and practical deployment, graphical user interfaces (GUIs) were created for each model type, enabling UDOT staff to analyze inspection videos, receive condition ratings, and generate detailed summary reports without technical expertise. When tested on 56 real-world videos, the object detection GUI correctly assessed culvert conditions in 84% of cases. The system offers a scalable, cost-effective, and objective approach to culvert inspection, reducing manual workload and increasing the consistency and accuracy of infrastructure condition assessments.

17. Key Words Culvert, Computer Vision, Object Detection, Automatic Interpretation		18. Distribution Statement Not restricted. Available through: UDOT Research Division 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410 www.udot.utah.gov/go/research		23. Registrant's Seal N/A
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 62	22. Price N/A	

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

UNIT CONVERSION FACTORS

No unit conversions

LIST OF ACRONYMS

FHWA	Federal Highway Administration
UDOT	Utah Department of Transportation
NASSCO	National Association of Sewer Service Companies
CVAT	Computer Vision Annotation Tool
YOLO	You Only Look Once
IoU	Intersection over Union
CNN	Convolutional Neural Network
Faster R-CNN	Faster Region-Based Convolutional Neural Network

EXECUTIVE SUMMARY

The Utah Department of Transportation (UDOT) is responsible for maintaining over 120,000 drainage culverts and storm drain pipes across state highways. Ensuring the integrity and functionality of these culverts is crucial for preventing flooding, sinkholes, and road damage, thereby safeguarding transportation infrastructure. Traditional inspection methods, primarily relying on manual visual assessments, are time-consuming, prone to human error, and lack consistency. To address these challenges, this project aims to develop an automated interpretation system for culvert inspection videos using advanced computer vision and deep learning technologies, specifically leveraging image classification and object detection algorithms.

The project utilized a diverse set of culvert inspection videos and images from UDOT's database, including video data from CCTV and zoom camera inspections. However, a significant challenge was the imbalanced dataset, dominated by corrosion defects in metal culverts. Therefore, we labeled more data collected from different sources and applied data augmentation techniques, including rotation and adding noise. Based on Utah's pipe rating system, the structural defect categories used in this study include Crack-Fracture, Break-Hole-Collapse-Kink, Corrosion, Deformation-Shape, and Joints, which represent the most observed issues in Utah's culvert inspection data. Each structural defect category originally contained up to five distinct classes; however, due to class imbalance and overlapping visual features, we merged these into smaller groups of two or three classes per category to improve model performance and ensure more consistent training.

After finishing the labeling and annotating of the collected images, we tailored the latest computer vision algorithms to each type of model we aimed to develop. For the binary classification task, we trained models to distinguish between defective and non-defective frames, achieving a high accuracy of 91%, which proved effective in filtering out defective frames for further analysis. Next, we developed multiclass classification models for each structural defect category, enabling the system to identify and categorize specific types of defects within a frame. These models achieved accuracies ranging from 77% to 96%, depending on the complexity and balance of the defect class distributions. To take the analysis a step further, we implemented object detection models capable of both localizing and classifying defects within each frame using

bounding boxes. The final object detection model achieved an average mean Average Precision (mAP) of 78%, offering detailed spatial insights necessary for assigning condition ratings.

To make these models accessible and usable by non-technical personnel, we developed intuitive Graphical User Interfaces (GUIs) for each model type. These GUIs allow UDOT employees to upload culvert inspection videos, run automated analyses, and receive detailed outputs without requiring programming knowledge. When evaluated on 56 real inspection videos, the object detection GUI accurately predicted the condition of 84% of the culverts, while the multiclass classification GUI achieved 75% accuracy. These results highlight the reliability and practical value of the system in real-world conditions.

For UDOT, these tools offer significant benefits: They reduce manual inspection time, improve consistency and objectivity in assessments, and help prioritize maintenance based on automated condition ratings. Ultimately, this system has the potential to streamline culvert management workflows, minimize human error, and lower operational costs while supporting timely, data-driven infrastructure decisions.

1.0 INTRODUCTION

1.1 Problem Statement

Ensuring the safety and functionality of existing transportation infrastructure, including roads, bridges, and culverts, is a top priority for engineers. Culverts, often hidden underground, play a crucial role in stormwater management by acting as channels that allow water to flow beneath various transportation structures [1]. They are vital for preventing flooding, sinkholes, and road damage. Regularly monitoring infrastructure is essential to catching small problems before they become expensive and time-consuming repairs. Therefore, transportation agencies send inspectors to inspect culverts on a regular basis. However, visually inspecting culverts presents a unique challenge due to their being buried and often located in areas with limited access [2].

To address these challenges, the use of digital video inspections has emerged as a valuable tool in assessing culvert conditions. Digital video inspections involve deploying cameras to capture images and videos of the interior of culverts. This method offers several advantages, including the ability to collect comprehensive data without the need for extensive excavation or physical entry into the culvert. Inspectors can review the footage to identify defects, assign condition ratings, and document inventory, significantly improving the accuracy and thoroughness of the inspection process. Remote cameras can also help inspectors to enhance safety by avoiding potentially hazardous conditions inside culverts, such as confined spaces, unstable structures, or exposure to toxic substances [3]. However, there are also disadvantages to consider. The initial investment in video inspection equipment and technology can be substantial, which may be a barrier for some transportation agencies. Moreover, the process still requires trained personnel to operate the equipment and interpret the data accurately. Despite automation advancements, the manual review of footage remains time-consuming and labor-intensive [4].

The Utah Department of Transportation (UDOT) maintains over 120,000 drainage culverts and storm drain pipes along state highways in Utah, making the importance of culverts more evident. Despite this, UDOT lacks a comprehensive inventory of these assets [5]. Establishing a detailed culvert inventory is crucial for predicting future performance and developing effective

maintenance strategies [6]. To this end, UDOT has to collect information about all culverts across the state and inventory them in ATOM as quickly as possible.

Identifying whether culverts need repair, rehabilitation, or replacement requires comprehensive and well-documented inspections. The current culvert inspection practice at UDOT is based on digital video inspection and relies heavily on human interpretation and defect identification. Inspectors collect videos on-site and later review them off-site, a process that includes defect identification, condition rating assignment, and inventory documentation. Each video interpretation takes approximately 10 to 12 minutes, leading to inefficiencies and time-consuming procedures. Moreover, because these video interpretations rely on human judgment, they are prone to subjectivity.

1.2 Objectives

The rapid advancements in computing technologies have significantly enhanced computer vision and deep learning models' capabilities in various fields, including the inspection of infrastructures. Key innovations in computer vision and deep learning have paved the way for automated systems that can precisely identify and evaluate defects in infrastructure components such as bridges, pipes, and roads. These advancements are reshaping traditional inspection methods, offering improved accuracy, efficiency, and cost-effectiveness.

Since UDOT currently relies on manual post-video interpretation for culvert inspections, this project aims to optimize the process by employing advanced technologies. Manual interpretation, while useful, is time-consuming, prone to human error, and can vary significantly based on the inspector's experience and subjectivity. To address these limitations, this project will review the most recent defect detection models developed for the assessment of culverts or pipes and develop a state-of-the-art deep learning model tailored to UDOT's specific needs. The model will be customized based on UDOT's culvert-condition rating criteria.

The automated system will also enable early detection of potential issues, facilitating proactive maintenance and reducing the risk of costly infrastructure failures. It would also allow seamless import of inspection data into ATOM UDOT Maintenance Management. This project aligns with UDOT's commitment to leveraging innovative technologies to enhance infrastructure

management and ensure the safety and functionality of the state's transportation network. In conclusion, this project represents a significant step forward in modernizing UDOT's culvert management system. By leveraging the latest advancements in deep learning and computer vision, UDOT can achieve a more efficient, accurate, and reliable system for assessing culvert conditions and maintaining critical infrastructure.

1.3 Scope

Research Tasks include:

- Conducting a comprehensive literature review to find the most advanced algorithm for defect detection
- Collecting available culvert inspection video data and converting them into images
- Labeling and annotating the images based on UDOT's culvert-condition rating criteria for model training purposes
- Developing a deep learning model that can interpret culvert inspection videos and assign a condition rating
- Comparing the results with the ground-truth data to evaluate the performance of the developed model
- Developing a Graphical User Interface (GUI) for an automated culvert inspection interpretation framework

1.4 Outline of Report

- Introduction
- Research Methods
- Data Collection
- Data Evaluation & Results
- Conclusions

2.0 RESEARCH METHODS

2.1 Overview

This chapter presents the methodology employed to automate the interpretation of culvert inspection videos through advanced computer vision and deep learning techniques. The methodology encompasses both traditional approaches to culvert inspection and cutting-edge applications of artificial intelligence in infrastructure assessment. We begin with an examination of conventional culvert inspection practices, highlighting their limitations and the need for automated solutions, followed by a review of recent advances in computer vision applications for infrastructure inspection.

The technical methodology covers three primary approaches: image classification for overall condition assessment, object detection for precise defect localization and classification, and data augmentation strategies for enhancing model robustness and addressing dataset limitations. Special attention is given to the latest developments in object detection architectures, including YOLO and RF-DETR models, each offering unique advantages for culvert defect detection tasks. The data augmentation section addresses critical challenges in developing robust models with limited training data, providing strategies for improving model generalization while maintaining the integrity of defect characteristics essential for accurate assessment.

2.2 Background

2.2.1 Traditional Culvert Inspection

Traditional culvert inspection methods have historically relied on visual assessments conducted by trained inspectors who physically examine culvert structures to identify defects and assess structural integrity [7]. These conventional approaches involve inspectors entering culvert systems when accessible or using basic optical equipment to evaluate visible portions of the infrastructure. However, traditional inspection methods face significant limitations, particularly when dealing with confined spaces, hazardous environments, or lengthy culvert systems that extend beyond safe human access [4].

The introduction of digital video inspection technology has emerged as a significant advancement over purely manual visual assessments. This method employs specialized cameras mounted on remotely operated vehicles or cable systems to capture comprehensive footage of culvert interiors. Digital video inspection allows for thorough documentation without requiring inspectors to enter potentially dangerous confined spaces, thereby improving safety while enabling more detailed condition assessments [4]. Despite these technological improvements, the interpretation of inspection footage remains largely dependent on human expertise, making the process time-intensive and subject to variability based on inspector experience and judgment.

2.2.2 Applications of Computer Vision in Infrastructure Inspection

The rapid advancement of computer vision and deep learning technologies has opened new possibilities for automating infrastructure inspection processes. Recent research has demonstrated the effectiveness of machine learning algorithms in detecting and classifying various types of structural defects with high accuracy and consistency [8]. Computer vision applications in infrastructure inspection have shown particular promise in analyzing large datasets of images and videos to identify patterns and anomalies that might be missed or inconsistently assessed through manual inspection.

Hawari et al. [9] developed an automated defect detection system for sewer pipelines using image processing algorithms applied to CCTV footage, focusing on four primary defect types: cracks, settled deposits, ovality, and displaced joints. Their study demonstrated varying performance levels across different defect types, with ovality detection showing superior results compared to other defect categories. The research highlighted the importance of comprehensive datasets for improving detection capabilities and recommended incorporating larger image collections to enhance system performance.

Yin et al. [10] advanced the field by implementing a real-time automated defect detection system using YOLOv3 architecture for sewer pipe assessment. Their model was trained on 3,664 images extracted from CCTV videos, encompassing six defect categories including holes, breaks, deposits, fractures, cracks, and root intrusion. The balanced distribution of defects in their dataset contributed to achieving an impressive 85.37% mean Average Precision (mAP) and F1 scores exceeding 87% for both testing and validation sets.

Kumar et al. [11] conducted a comparative analysis of multiple object detection frameworks, including Faster R-CNN, YOLOv3, and Single Shot Detector (SSD), for detecting sewer pipe deposits and root intrusion. Their findings indicated that while Faster R-CNN achieved superior overall performance, YOLOv3 provided a more balanced trade-off between detection speed and accuracy, making it suitable for real-time applications. Yin et al. [12] further expanded automated assessment capabilities by developing the Video Interpretation Algorithm for Sewer Pipes (VIASP), which integrated defect detection with location identification and report generation, achieving an F1 score of 0.75.

According to past studies, deep learning algorithms have shown significant potential in enhancing the defect detection process for pipelines, leading to more consistent and reliable results. Research has demonstrated that these algorithms can effectively identify and classify defects such as cracks, corrosion, and joint misalignments with a high degree of accuracy, surpassing traditional manual inspection methods. Despite these advancements, there has been a notable gap in applying deep learning techniques specifically to culvert inspections. Our project aims to address this gap by employing novel deep learning algorithms tailored for defect detection and condition assessment of culverts in Utah.

To address the limitations of traditional culvert inspection methods currently used by UDOT, we propose an automated system for interpreting culvert inspection videos using deep learning. Manual inspection of video footage is time-consuming, labor-intensive, and prone to subjectivity. Our goal is to streamline and standardize this process by developing an intelligent pipeline that can detect, localize, and assess defects automatically. The automation process is divided into three key phases:

1. In the first phase, we use advanced image classification models to analyze individual frames extracted from culvert inspection videos. These models are trained to identify frames that show visible signs of damage, such as cracks, joint misalignments, corrosion, or surface deformation. This step effectively filters out non-defective frames, allowing the system to focus on areas that actually require further analysis, thereby improving efficiency and reducing processing time.
2. Once defective frames are identified, they are passed through object detection models. These models are trained on a curated dataset of annotated images containing various types of culvert defects. They are capable of not only locating defects within a frame

but also categorizing them into specific classes (e.g., crack/fracture, break/hole, and deformation). This step provides both spatial and contextual information about each defect, which is essential for comprehensive analysis.

3. In the final phase, the localized and categorized defects are evaluated to determine their severity. This is done using advanced object detection models trained on a large dataset that is annotated based on UDOT's culvert rating system. The model estimates factors such as defect color, shape, and texture, and assigns a condition rating based on learned patterns in the training data. These ratings can then be used to inform maintenance decisions and prioritize repairs.

To implement this pipeline, we utilize two main classes of deep learning algorithms: image classification and object detection. Image classification models are used in the first phase to detect frames likely to contain defects, while object detection models are employed to combine the second and third phases to precisely locate, label, and assess each defect. By automating the interpretation of culvert videos, this system promises to significantly enhance inspection accuracy, reduce human error, and streamline the maintenance workflow for UDOT. The result is a faster, more consistent, and more scalable approach to culvert condition assessment.

2.3 Image Classification

Image classification represents a fundamental computer vision task that involves categorizing entire images into predefined classes based on their visual content [13], [14]. In the context of culvert inspection, image classification can be employed to automatically determine the overall condition rating of culvert segments or to classify images based on the presence or absence of specific defect types. This approach differs from object detection in that it assigns a single label to the entire image rather than localizing and identifying specific defects within the image [15].

The image classification methodology typically employs Convolutional Neural Networks (CNNs) that learn hierarchical feature representations from training data. These networks progressively extract features from low-level edge and texture information to high-level semantic representations that enable accurate classification decisions [16]. For this project we developed two types of image classification models:

1. A binary classification model which efficiently filters non-defective frames from culvert video footage.
2. A multi-class classification model that assigns a condition rating (on a scale from 1 to 5) to each frame in the video footage, based on the severity level of identified defect categories.

We employed CNNs, such as ResNet and EfficientNet, which have demonstrated high performance in various image recognition tasks. These models are trained on a labeled dataset consisting of both defective and non-defective culvert frames. Through supervised learning, the model learns to distinguish visual patterns associated with defects such as cracks, joint separations, and corrosions. The output is a binary or multi-class label indicating whether a given frame should be flagged for further inspection.

The classification process begins with preprocessing steps, including image normalization, resizing, and potential augmentation to improve model robustness. CNN architecture processes these images through multiple convolutional layers, pooling operations, and fully connected layers to produce probability distributions across the target classes. Training involves optimizing the network parameters using labeled examples and validation on separate datasets to ensure generalization capability. Performance evaluation typically employs metrics such as accuracy, precision, recall, and F1-score to assess the model's effectiveness in correctly classifying culvert conditions.

2.4 Object Detection

YOLOv8 (You Only Look Once, Version 8) represents one of the latest iterations in the evolution of the YOLO family of real-time object detection models. Building on the strengths of its predecessors, YOLOv8 incorporates several architectural enhancements to achieve improved accuracy and speed in detecting and classifying objects within images and videos. At its core, YOLOv8 retains the fundamental principle of YOLO: treating object detection as a single regression problem, predicting bounding boxes and class probabilities directly from full images in one evaluation. This end-to-end approach ensures that YOLOv8 remains efficient and fast, making it particularly suitable for real-time applications [17].

Recent developments in YOLO architecture have continued to push the boundaries of object detection performance. YOLOv11, released as an advancement over YOLOv8, introduces

enhanced feature pyramid networks and improved anchor-free detection mechanisms that provide better handling of multi-scale objects and reduced computational overhead [18]. The architecture incorporates advanced attention mechanisms and optimized backbone networks that significantly improve detection accuracy while maintaining real-time processing capabilities.

YOLOv12 represents the most recent iteration in the YOLO family, featuring revolutionary architectural improvements including dynamic head structures and advanced multi-scale fusion techniques. This version introduces novel training strategies and loss functions that enhance the model's ability to detect small objects and handle complex scenes with multiple overlapping instances [19].. The model demonstrates superior performance in challenging scenarios common in infrastructure inspection, where defects may appear at various scales and orientations.

RF-DETR (Real-time DEtection TRansformer) represents a significant departure from traditional CNN-based detection approaches by employing transformer architectures for object detection tasks. This model leverages self-attention mechanisms to capture long-range dependencies and spatial relationships more effectively than conventional approaches. RF-DETR demonstrates strength in detecting complex defect patterns and spatial relationships that are crucial for accurate culvert condition assessment [20].

The architecture of YOLOv8 (Figure 1) introduces key innovations that enhance its performance. One of the primary improvements is the incorporation of advanced CNN layers that optimize feature extraction. These layers are designed to capture more detailed spatial information, allowing the model to detect smaller objects and distinguish between closely spaced objects with greater precision. YOLOv8 also utilizes advanced activation functions and normalization techniques that enhance the model's learning capability and stability during training. Additionally, the model benefits from a refined anchor box mechanism, which dynamically adjusts to different object scales and aspect ratios, further improving detection accuracy. The architecture is also designed to be modular, facilitating the integration of additional components such as attention mechanisms or more sophisticated loss functions, thus enabling further customization and optimization for specific tasks. These architectural advancements collectively contribute to YOLOv8's superior performance in object detection, making it a powerful tool for various applications, from autonomous driving to surveillance and beyond.

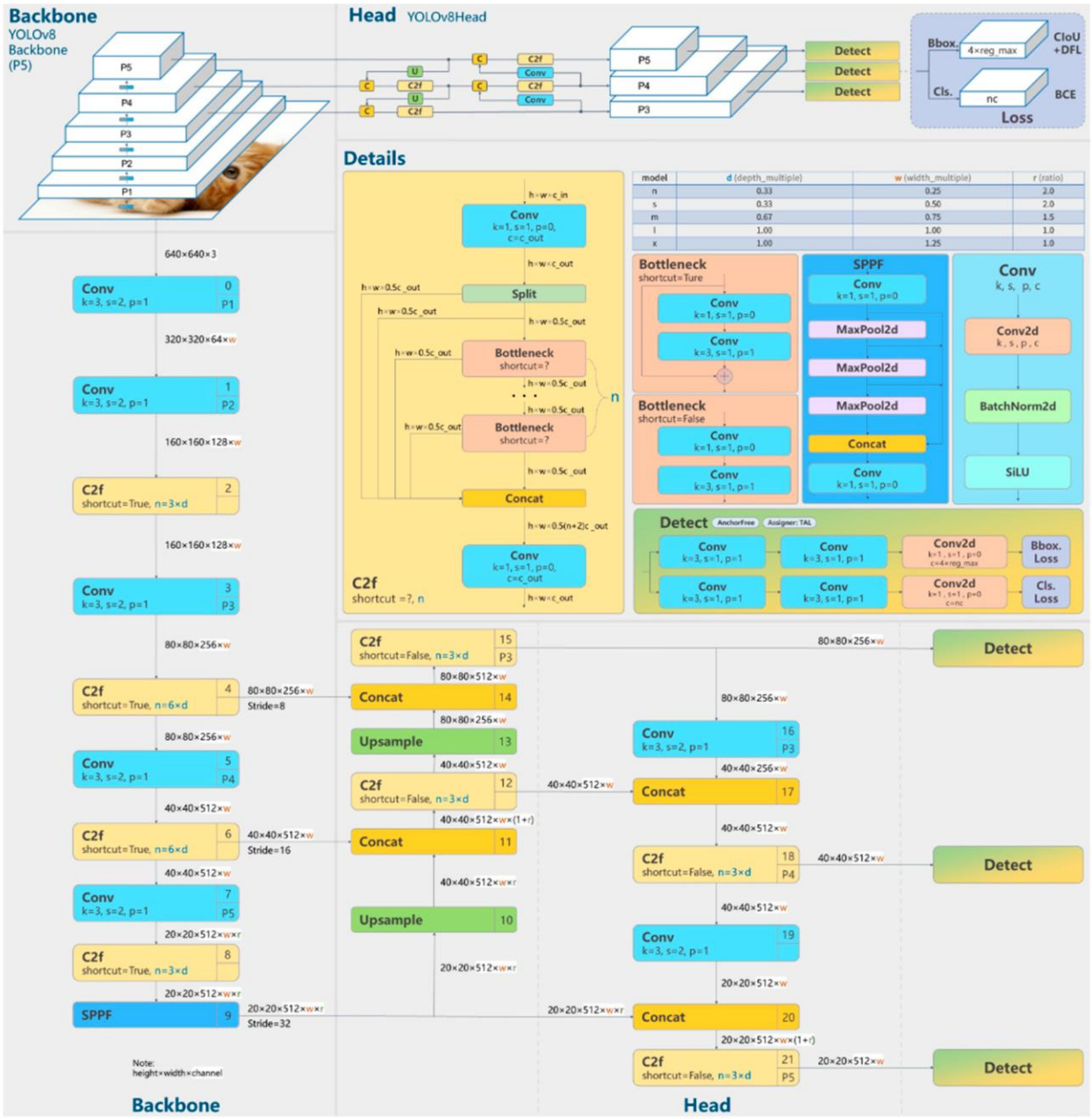

Figure 1-YOLOv8 architecture [21]

For training and testing our model, we employed a holdout cross-validation technique, wherein the dataset was divided into two distinct subsets: 70% for training, 20% for validation, and 10% for testing. This approach ensures that the model is trained on a substantial portion of the labeled data while reserving a separate set of data to evaluate its performance on unseen samples. By doing so, we can obtain a more accurate assessment of the model's generalization capabilities and its ability to detect defects in new, unobserved culvert inspection videos. This method helps

to prevent overfitting and provides a realistic measure of the model's effectiveness in real-world scenarios, ensuring that it reliably performs when deployed in UDOT's inspection processes [22]. For this project we developed two types of object detection models:

1. An object detection model for detecting and localizing defect categories without rating scales.
2. An object detection model for detecting, localizing, and classifying defects with rating scales

2.5 Data Augmentation

Data augmentation represents a critical technique in deep learning that artificially expands training datasets by applying various transformations to existing images while preserving their semantic content and class labels [6], [23]. In the context of culvert inspection, data augmentation serves multiple purposes: addressing dataset imbalances, improving model generalization, and enhancing robustness to variations in imaging conditions encountered during field inspections. The technique is particularly valuable when working with limited datasets, as is often the case in specialized infrastructure inspection applications where collecting comprehensive labeled data can be time-consuming and expensive [24].

Common augmentation techniques applicable to culvert inspection include geometric transformations such as rotation, scaling, translation, and horizontal flipping, which help the model become invariant to different camera orientations and positions during inspection. Photometric augmentations, including brightness adjustment, contrast modification, color jittering, and noise addition, are essential for handling varying lighting conditions encountered in different culvert environments. More advanced techniques such as cutout, mix-up, and mosaic augmentation can further enhance model robustness by forcing the network to rely on multiple visual cues rather than focusing on specific image regions [25].

The implementation of data augmentation requires careful consideration of the specific characteristics of culvert inspection imagery. For instance, excessive rotation may not be appropriate as culvert orientations are typically constrained, while brightness and contrast adjustments are crucial given the challenging lighting conditions often present in underground infrastructure [23]. To increase the size of our training dataset, we applied a diverse set of data

augmentation techniques to each original image, generating seven augmented outputs per example. The augmentations used are as follows:

- **Flip**: Applied horizontal and vertical flips to introduce directional variability.
- **90° Rotation**: Included both clockwise and counter-clockwise 90-degree rotations to account for different viewing angles.
- **Crop**: Performed random cropping with a minimum zoom of 0% and a maximum zoom of 20% to simulate partial views of defects.
- **Rotation**: Applied random rotations ranging between -15° and +15° to account for slight camera tilts.
- **Shear**: Introduced horizontal and vertical shear transformations up to ±10° to mimic perspective distortion.
- **Saturation Adjustment**: Randomly varied image saturation between -25% and +25% to account for lighting differences and material surface changes.
- **Brightness Adjustment**: Modified brightness levels within a range of -15% to +15% to simulate varied lighting conditions.
- **Exposure Adjustment**: Altered exposure levels from -10% to +10% to reflect overexposed or underexposed footage.
- **Blur**: Added Gaussian blur with a maximum radius of 1.8 pixels to simulate motion blur or low focus.
- **Noise**: Introduced random noise affecting up to 0.22% of pixels to improve robustness against video compression artifacts.

These augmentations were carefully selected to preserve the semantic integrity of the defects while improving the model's generalization to real-world variations in culvert inspection videos.

3.0 DATA COLLECTION

3.1 Overview

Collecting the necessary input data is the first crucial step in developing a robust deep learning model. For this project, we gathered extensive culvert video inspection data from UDOT's database. Then, we had to label the collected data to develop a deep learning model. Data labeling for object detection involves annotating images or videos by drawing bounding boxes around objects of interest and assigning a specific class to each box. This precise annotation allows the model to learn how to identify and classify various objects accurately, which is crucial for developing an effective and reliable object detection system.

3.2 Collected data

We collected four distinct categories of data for this project: zoom camera inspection videos, CCTV culvert inspection videos, culvert images taken by cell phone, and CCTV sewer pipe inspection images. These diverse data sources provided a comprehensive view of culvert conditions, capturing various perspectives and levels of detail. By utilizing these varied data types, we ensure that our model is trained on a rich dataset, enhancing its ability to accurately detect and assess a wide range of defects in different inspection scenarios.

3.2.1 Zoom Camera Inspection Videos

UDOT provided us with its available culvert inspection data, a vital resource for our deep-learning model development. A significant subset of this dataset consists of culvert inspection videos collected by Consor Company. Consor employed a method of video inspection using zoom cameras mounted on the ends of telescopic poles. This technique allows for a detailed examination of the culverts without the need for prior cleaning, which can save considerable time and resources. The zoom camera inspection provides high-resolution footage of the culvert interior, capturing minute details that are crucial for accurate defect detection and assessment. Consor Company shared over 2000 inspection video files with us, encompassing a wide range of culvert conditions and types. These videos cover inspections conducted in Region One of Utah (Figure 2).

Figure 2-Inspected culverts by Consor in Region One of Utah

For culvert inspection, Consor Company utilized the NASSCO rating system, which is comparable to UDOT's hybrid rating system, eliminating the need for conversion between rating systems in this project. Out of the extensive dataset of culvert inspection, only 1094 inspection reports were available. Among these reports, approximately 22% indicated that the inspected culverts had no structural defects, further limiting the amount of data containing observable issues and reducing the number of samples useful for training defect detection models. This limited number is due to inconsistencies in the dataset; some culverts had multiple videos linked to the same pipe, while others lacked corresponding video footage altogether. As a result, the number of usable video-report pairs was significantly reduced.

We converted the video files into images. These images capture detailed visual information from the culvert inspections, providing a rich dataset for training our deep learning model. However, a critical step in utilizing these images is the annotation process. We meticulously annotated the defects in each image as part of data labeling, which is essential for training the model to accurately detect and classify defects. This labor-intensive process ensures that the model learns from high-quality, labeled examples, improving its ability to generalize and perform effectively on unseen data.

3.2.2 CCTV Culvert Inspection Videos

Another significant subset of culvert inspection data available in UDOT's database comprises CCTV inspection video files collected by Horrocks Company, specifically from the culverts along the I-80 highway. We accessed this data through UDOT's R2 culvert rating app website (Figure 3). This dataset includes 2000 data rows, but only 259 of these entries have corresponding video files. Among these 259 videos, merely 59 exhibit structural defects, providing a more focused dataset for our defect detection model.

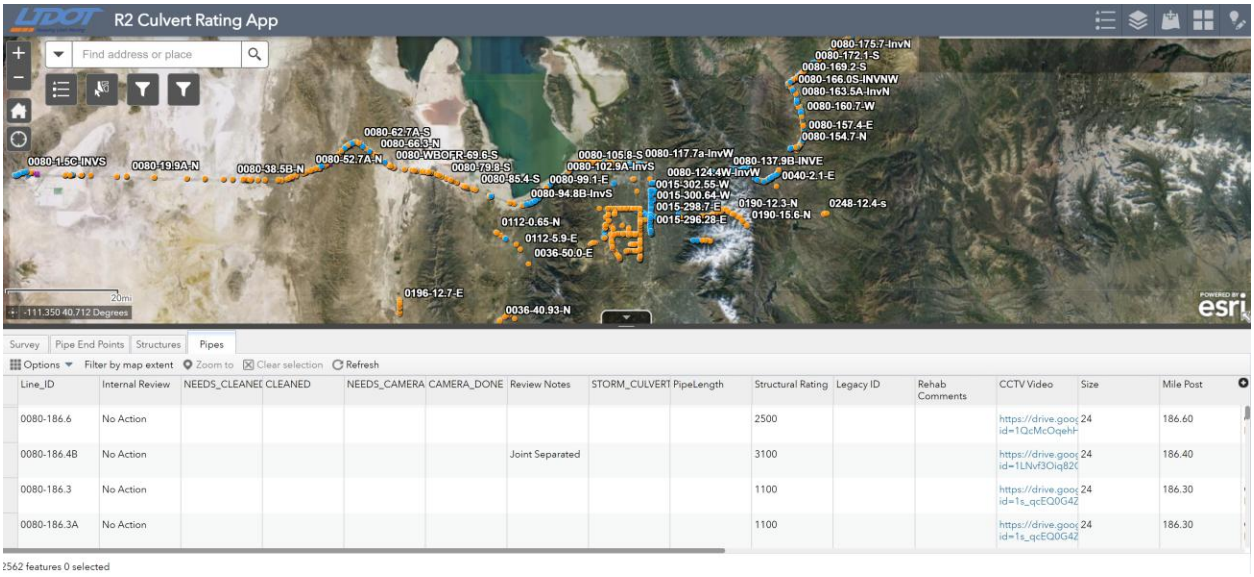

Figure 3-Inspected culverts by Horrocks along the I-80 highway

The data collected by Horrocks Engineers utilized the old four-digit rating scale previously used by UDOT. To ensure consistency and compatibility with our model, we needed to convert these ratings to the new 5-point rating scale that UDOT currently employs. This conversion was necessary to standardize the data and make it suitable for training our deep learning model. The conversion process involved using a predefined table (Table 1) to translate the old ratings into the new scale, ensuring that the defect severity and conditions are accurately represented in the updated system.

Table 1-Data conversion

5-point rating scale	Four-digit rating scale
1	<1000
2	1000-1999
3	2000-2999
4	3000-3999
5	>=4000

3.2.3 Culvert Images Taken by Cell Phone

Another subset of data we collected consisted of culvert images taken by UDOT employees during their field visits. Recognizing that our initial dataset was unbalanced, with a risk of overfitting the model to specific defect classes, we needed to incorporate additional data to ensure a more comprehensive training set. To achieve this, we reached out to all UDOT employees, requesting that they share any culvert images they had captured in the field. This initiative resulted in the collection of 450 additional culvert images. Out of the 450 images, only 193 contained visible structural defects.

These images, however, had not been labeled with condition ratings or defect annotations. Therefore, as part of our data preparation process, we should assess each culvert image and assign a condition rating based on UDOT's new hybrid culvert condition rating system. This meticulous labeling process is crucial for creating a high-quality dataset that accurately represents a wide range of culvert conditions and defects.

By integrating these additional images into our dataset, we aim to enhance the model's ability to generalize across different defect types and conditions, reducing the likelihood of overfitting. The diverse and balanced dataset that results from this effort will provide a solid foundation for training our deep learning model, ultimately improving its accuracy and reliability in detecting and assessing culvert defects.

3.2.4 CCTV Sewer Pipe Inspection Images

The final type of data we collected consisted of images captured from inside sewer pipes. These images were extracted as frames from CCTV inspection videos and were sourced from three different repositories.

The first batch was obtained from the Roboflow website [26], which provided approximately 1,500 unlabeled images of sewer interiors. Since these images had no annotations, we manually labeled them using UDOT's culvert defect rating system to ensure consistency with the rest of our dataset.

The second batch came from a former employee of AECOM, who shared a collection of 45,000 sewer pipe images labeled using the NASSCO defect rating system. However, only about 3,000 of these images contained visible structural defects. After further review and filtering to exclude defects irrelevant to culvert inspection, such as those found in pipes made from materials like vitrified clay, we narrowed this batch down to 591 usable images. To maintain a unified labeling scheme across all data sources, we converted the NASSCO ratings to UDOT's rating system.

The third batch was sourced from the Kaggle website and consisted of an augmented dataset of sewer pipe images. Initially, it contained 22,120 images, but after removing the augmented duplicates and retaining only the original frames, we were left with 5,530 annotated images. These images were labeled with six types of defects: Deformation, Obstacle, Rupture, Disconnect, Misalignment, and Deposition. Following a thorough review and the conversion of these labels into UDOT's classification framework, this batch was reduced to 3,216 relevant and consistently labeled images.

In total, these three sources contributed to a diverse and standardized dataset of sewer and culvert defect images, all aligned under the UDOT rating system to support robust training and evaluation of our defect detection models.

3.3 Data Labeling

In this project, our goal is to develop two types of supervised learning models specifically for classification and object detection. Supervised learning models require labeled data to learn meaningful patterns, but the type and complexity of labeling differ significantly between

classification and object detection tasks. Object detection models, such as YOLO, require detailed annotations, including bounding boxes around each object of interest and corresponding class labels. This makes the annotation process both time-consuming and labor-intensive. In contrast, supervised image classification models like ResNet and EfficientNet operate on labeled images without the need for precise localization. They simply require a label for the entire image, making the data preparation process considerably easier and faster.

In this study, we approached the labeling process in stages, progressing from simpler to more complex tasks. We began by labeling images for a binary classification model, distinguishing between defective and non-defective culvert frames. This involved extracting individual frames from culvert inspection videos and manually assigning a binary label (defective and non-defective) to each (Figure 4).

Figure 4-Assigning a binary label to images

Next, we labeled the same dataset for a multi-class classification model, assigning specific defect labels (e.g., corrosion-3 or joint-5) to the images. In the development of our multi-class classification models, we designed a modular approach by creating five separate models, each specialized for a distinct category of structural defects commonly found in culvert inspections. This strategy allowed us to fine-tune each model for the unique visual characteristics and classification challenges associated with different defect types, rather than training a single model to handle all defect categories simultaneously. The five categories we focused on were: **Break/Hole/Collapse/Kink, Corrosion, Crack/Fracture, Deformation**, **Joint Offset.**

For each of these categories, we curated a tailored subset of images from our dataset and applied class labels specific to the types of defects within that group. These class labels, along with the corresponding categories, are detailed in Table 3. This categorization not only enhanced the performance of each model by reducing label noise and inter-class confusion but also enabled

more targeted training and evaluation. By isolating defect types, we improved the models' sensitivity to subtle variations within each defect class, leading to more reliable and interpretable classification outcomes in real-world culvert assessments. Since the number of images in some classes was very limited, we merged similar classes within each defect category to create a more balanced dataset. For example, in the corrosion category, we combined severity levels 2 and 3 into a single class labeled as "Corrosion-3" due to the low number of samples in level 2. This merging strategy helped address class imbalance and improved the reliability of model training. As a result, some categories ended up with two or three consolidated classes, depending on the distribution of available data.

Table 2-Multi-class labels by defect type and severity

Break/Hole/ Collapse/Kink	Corrosion	Crack/Fracture	Deformation	Joint Offset
Non-bre-hol-col-kin-1	Non-corrosion-1	Non-crack-frac-1	Non-deformation-1	Non-joints-1
	Corrosion-2	Crack-Fract-2	Deformation-2	Joints-2
	Corrosion-3	Crack-Fract-3	Deformation-3	Joints-3
Bre-Hol-Col-Kin-4	Corrosion-4		Deformation-4	Joints-4
Bre-Hol-Col-Kin-5	Corrosion-5		Deformation-5	Joints-5

Finally, for the object detection model, we annotated the frames with bounding boxes around visible defects and assigned each region a corresponding class label from Table 3. This structured, step-by-step annotation strategy allowed us to build and evaluate models of increasing complexity, leveraging the same video data across multiple learning tasks.

3.3.1 Data Annotation with CVAT

To prepare our data for object detection, we need to go through a detailed annotation process. This involves labeling each image or video frame by drawing bounding boxes around the objects of interest, such as culvert defects, and assigning a specific class to each box. For instance, if a culvert image contains cracks, corrosion, or joint misalignments, each of these defects must be identified with a bounding box and labeled with the appropriate class. This precise annotation is crucial as it provides the model with the necessary information to distinguish between different types of defects during training.

The labeling process must be thorough and consistent to ensure the model learns from high-quality examples. Each annotated image or video frame helps the object detection model understand the features and patterns associated with various defects. Once the model is trained, its performance will be evaluated on a labeled test set, which was not used during the training phase. This evaluation will help us measure the model's accuracy in detecting and classifying defects, ensuring that it performs reliably on unseen data.

For this task, we used CVAT a powerful open-source tool specifically designed for annotating image and video datasets [27]. Using CVAT's user-friendly interface (Figure 4), annotators manually draw bounding boxes around each defect in the images or video frames. Each bounding box is labeled with the corresponding class. CVAT provided various tools to streamline the annotation process, such as auto-segmentation, interpolation for video frames, and copy-paste functions for repetitive objects.

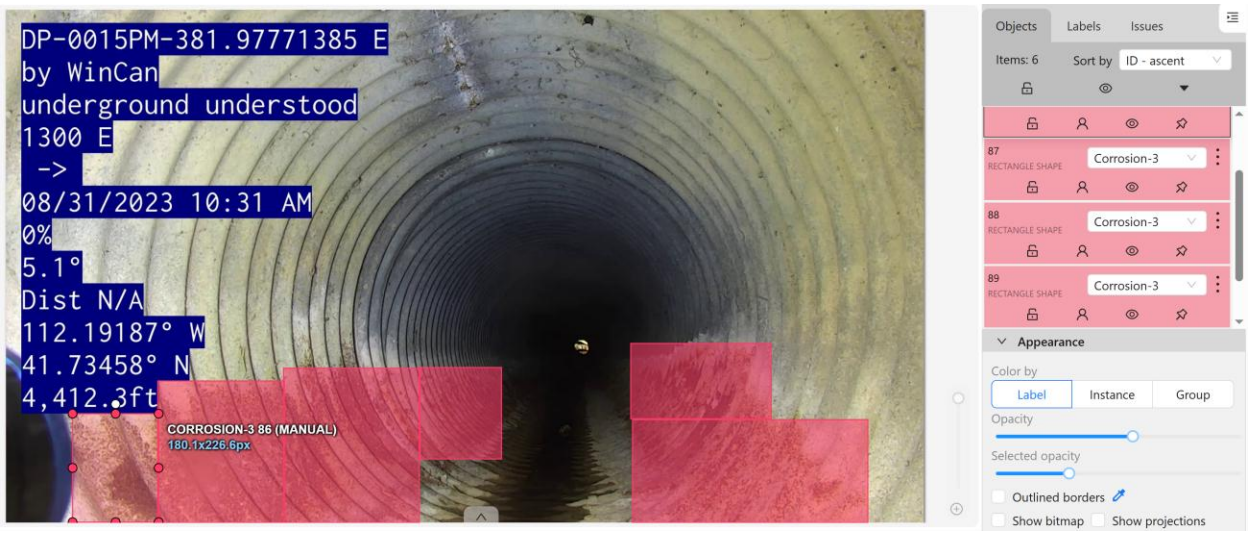

Figure 5-CVAT's interface during data labeling

For annotation, we utilized 16 distinct labels, as illustrated in Table 3, in accordance with UDOT's culvert condition rating system. Once the annotation process was completed, the labeled data was exported in a format compatible with the specific model we want to use. CVAT supports various export formats, including YOLO's native format. The exported annotations include the coordinates of the bounding boxes and the class labels for each object.

Table 3 illustrates the distribution of defects boxes across the 4863 labeled images. According to the table, it indicates that a significant portion of UDOT's culverts are metal pipes afflicted with corrosion defects. This imbalance in our dataset reveals that certain defect types are underrepresented, with some labels appearing in fewer than 200 boxes. To address this issue, we have been actively working to augment our dataset by incorporating additional data. This effort aims to balance the representation of defect types, thereby enhancing the overall performance and accuracy of our model.

Table 3-Distribution of defect boxes across labeled data

#	Defects	Box count
0	Bre-Hol-Col-Kin-4	309
1	Bre-Hol-Col-Kin-5	1333
2	Corrosion-2	586
3	Corrosion-3	2803
4	Corrosion-4	78
5	Corrosion-5	90
6	Deformation-2	193
7	Deformation-3	610
8	Deformation-4	302
9	Deformation-5	478
10	Crack-Fract-2	646
11	Crack-Fract-3	150
12	Joints-2	423
13	Joints-3	801
14	Joints-4	328
15	Joints-5	211
Sum		9341

4.0 DATA EVALUATION & RESULTS

4.1 Overview

In this chapter, we will discuss the metrics used to evaluate our model's performance, enabling us to predict how it will perform on unseen culvert video inspection data. We will also present and analyze the results obtained from these evaluations, providing insights into the model's accuracy and reliability in real-world applications.

4.2 Evaluation Metrics

Evaluating the performance of computer vision models, particularly in the domains of object detection and image classification, requires the use of well-established quantitative metrics. These metrics allow researchers and practitioners to assess how well models generalize to unseen data, how accurately they recognize objects or classify images, and how their predictions align with ground-truth annotations. The primary evaluation metrics we used in this study are mean Average Precision (mAP), precision, recall, F1-score, accuracy, and the confusion matrix.

4.2.1 Confusion Matrix

The confusion matrix provides a granular view of the classification performance by showing how predictions are distributed across the actual class labels. Each row corresponds to the true class, and each column corresponds to the predicted class [28]. It highlights where the model is making errors, such as misclassifying one class as another, and is particularly useful for evaluating multiclass classification models. Figure 6 illustrates a confusion matrix, which presents and summarizes the difference between the predicted and actual classes generated by a classification model.

		Predicted class	
		Positive	Negative
Actual Class	Positive	True Positive	Flase Positive
	Negative	False Negative	True Negative

Figure 6-Confusion matrix for binary classification

4.2.2 Accuracy

Accuracy is the simplest and most intuitive metric, defined as the proportion of correct predictions among the total number of predictions [28]. While widely used in image classification tasks, accuracy alone can be misleading in the presence of class imbalance or when evaluating object detection tasks with multiple classes and varied object sizes. In object detection, accuracy is less frequently used in isolation, as it does not account for localization quality or multiple instances per image.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Predictions} \qquad \text{Equation 1}$$

4.2.3 Precision and Recall

Precision measures how many of the predicted positive instances are actually correct, whereas recall measures how many of the actual positive instances the model was able to identify [29]. In object detection, precision and recall are often computed across multiple Intersection over Union (IoU) thresholds. A high precision indicates a low false positive rate, while a high recall reflects a low false negative rate. Balancing these two is crucial, especially in safety-critical applications where missed detections or false alarms have different implications.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad \text{Equation 2}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad \text{Equation 3}$$

4.2.4 F1-Score

The F1-score is the harmonic mean of precision and recall. It provides a single metric that balances both concerns, particularly useful when class distribution is skewed or when both false positives and false negatives are costly [30]. For image classification, a high F1-score indicates that the model is not only accurate but also robust in handling both positive and negative predictions.

$$F1\ Score = 2 \times \frac{Precision\ \times Recall}{Precision + Recall} \quad \textbf{Equation 4}$$

4.2.5 Mean Average Precision ([19])

To evaluate the performance of object detection models, we used mAP metric. MAP is a widely recognized performance measure in object detection tasks. It combines both precision and recall across different classes and thresholds, providing a comprehensive assessment of the model's accuracy [21]. The formula (Equation 5) for mAP involves calculating the Average Precision (AP) for each class and then taking the mean of these AP values. AP is the area under the Precision-Recall curve for a given class. It can be calculated by taking the precision at different recall levels (Equation 5).

$$\text{AP} = \sum_{n} (R_n - R_{n-1}) P_n \quad \textbf{Equation 5}$$

where P_n is the precision at the n-th threshold and R_n is the recall at the n-th threshold. MAP calculates the AP for each class and then averages these values to give an overall score, effectively summarizing the model's ability to correctly identify and localize defects in the culvert inspection images. Using mAP, we can ensure that our model not only detects defects accurately but also maintains a high level of reliability across various defect types.

$$\text{mAP} = \frac{1}{N}\sum_{i}^{N} \text{AP}_i \qquad \textbf{Equation 6}$$

where N is the number of classes and AP_i is the Average Precision for the i-th class. This metric is particularly useful in our context, as it helps ensure that our YOLOv8 model not only detects and localizes culvert defects accurately but also maintains consistent performance across different defect types, thereby validating the model's robustness and effectiveness in real-world applications.

MAP50, or Mean Average Precision at 50% Intersection over Union (IoU) threshold, is a specific metric used to evaluate the performance of object detection models. IoU is a measure used to quantify the accuracy of an object detector's predicted bounding box with respect to the ground truth bounding box. It is calculated as the area of overlap between the predicted bounding box and the ground truth divided by the area of their union. In the context of mAP50, the model's predictions are considered correct if the IoU between the predicted bounding box and the ground truth bounding box is at least 50%. This threshold is a common benchmark used to evaluate object detection models.

4.3 Results

To detect structural defects in culverts, we developed a YOLOv8 model using 34,390 labeled images. However, the initial results were unsatisfactory, as the model performed great in one class while overlooking others. This issue stemmed from the unbalanced nature of our dataset. Initially, we had 20 labels, but after annotating 77 videos, only ten labels were used. The distribution among these ten labels was also imbalanced. Figure 7 presents the normalized confusion matrix for this model, which indicates that the model's performance is quite low. As an example, Figure 8 shows a batch of a test set.

4.3.1 Classification

In this project, we developed and evaluated two types of image classification models. The first model type focused on binary classification, aiming to detect whether a given video frame contains a defect or not. To accomplish this, we experimented with four distinct image classification algorithms, which are summarized in Table 4.

Table 4-Binary Classification Results

#	Model	# Classes	Accuracy	Precision	Recall	F1 score
1	Yolov11	2	91%	90.3%	90.5%	90.4%
2	Resnet 50	2	83%	88%	89%	88%
3	VGG + XGBoost	2	85%	84%	84%	84%
4	ConvNeXt	2	74%	85%	67%	67%

Among the tested models, YOLOv11 consistently outperformed the others across all evaluation criteria. Specifically, YOLOv11 achieved the highest F1-score and accuracy, indicating its superior ability to correctly distinguish between defective and non-defective frames while maintaining a balanced trade-off between false positives and false negatives.

The performance advantage of YOLOv11 can be attributed to its robust feature extraction capabilities and optimized architecture, which proved especially effective in recognizing subtle defects within noisy or complex visual contexts. Given its reliability and efficiency, YOLOv11 was selected as the primary binary classifier for the first stage of the defect detection pipeline.

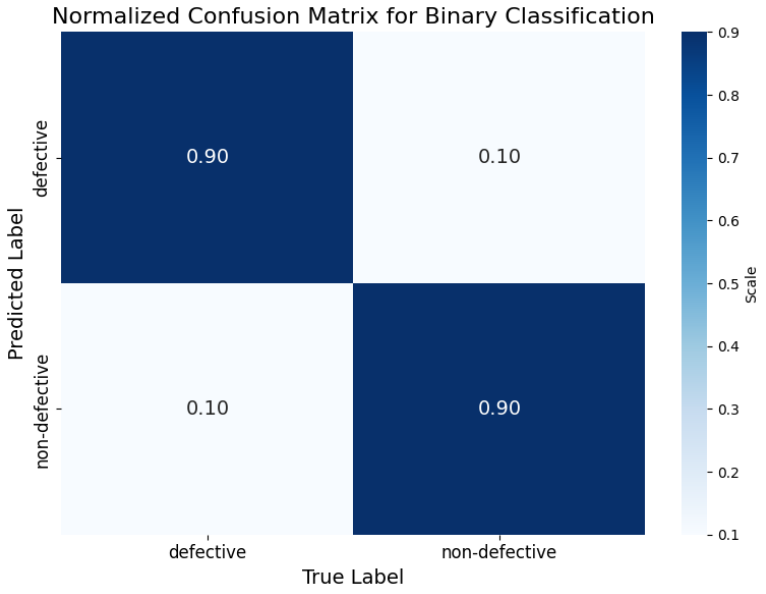

Figure 7-Confusion matrix of Yolo model for binary classification

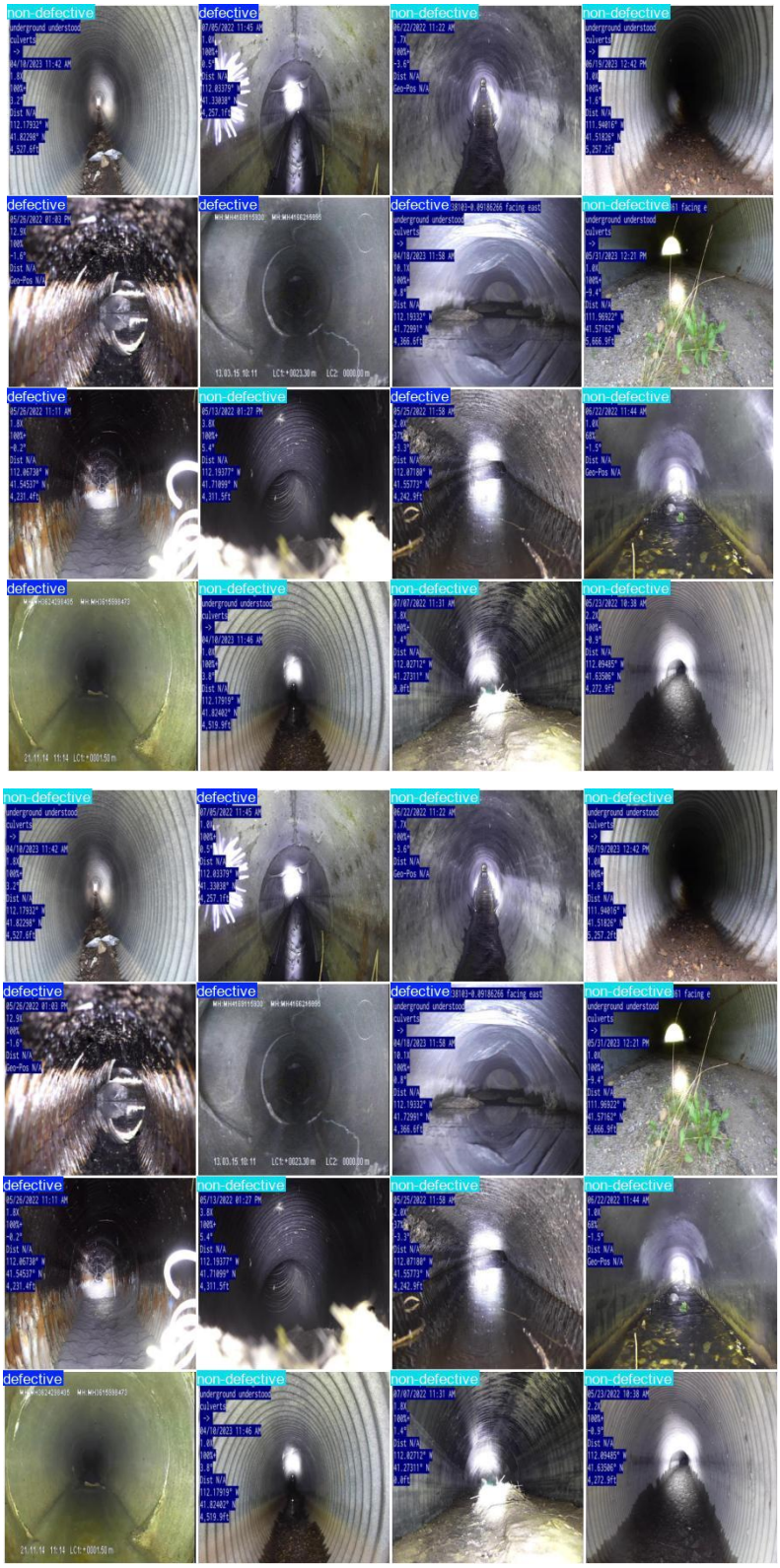

Figure 8-Examples of Yolo model predictions (bottom images) and true labels (top images) for a batch of testing set

In the next phase of this research, the focus shifted from binary classification to a more detailed analysis of pipe conditions through multiclass image classification models. The goal was to not only detect whether a frame was defective but also to classify the type of defect present in the image. To facilitate this, we identified and defined five major structural defect categories commonly observed in Utah's culvert inspection records

Each major defect category encompassed multiple specific defect classes, allowing for a more granular understanding of structural issues (Table 2). After defining the categories, we proceeded to annotate a large dataset of images, assigning each image to its appropriate class within a category. By training separate models per category, we aimed to improve model focus, reduce confusion between dissimilar classes, and ultimately enhance classification performance within each defect group.

To implement this multiclassification framework, we employed two state-of-the-art deep learning models: YOLOv11 and EfficientNet. These models were selected for their strong performance in prior image classification tasks and their architectural ability to generalize across diverse visual inputs. YOLOv11 provided fast and accurate real-time inference, while EfficientNet offered high accuracy with optimized computational efficiency through compound scaling.

The performance of both models was evaluated independently across each defect category using standard metrics, including accuracy, precision, recall, and F1-score. The results, summarized in YOLOv11, meanwhile, also demonstrated strong performance in Corrosion (96% accuracy) and Crack-Fracture (94% accuracy). It showed improved recall in several categories compared to EfficientNet but slightly more fluctuation in precision. Notably, in the Break/Hole/Collapse/Kink category, YOLOv11 achieved 81% accuracy with 66% precision and 76% recall, leading to a respectable F1-score of 68%. Overall, both models performed well, with EfficientNet exhibiting more consistent precision across categories and YOLOv11 showing higher recall, particularly for defect types where comprehensive detection is crucial. Findings demonstrate how each model performed within the context of the structural defect classification, offering insight into the strengths and weaknesses of each approach across different types of culvert anomalies. Due to class imbalance in the dataset, we merged the original five or three classes within each category into three or two classes to ensure more balanced representation and improve model performance.

EfficientNet achieved high accuracy across all categories, with its best performance in Corrosion (96% accuracy) and Crack-Fracture (94% accuracy). However, the model showed relatively lower performance in Deformation and Joints, both recording F1-scores around 61–63%, likely due to the visual similarity of these defects or fewer training samples. YOLOv11, on the other hand, demonstrated strong results in Corrosion as well (96% accuracy) and slightly better recall for Crack-Fracture than EfficientNet (81% vs. 68%). Interestingly, YOLOv11, meanwhile, also demonstrated strong performance in Corrosion (96% accuracy) and Crack-Fracture (94% accuracy). It showed improved recall in several categories compared to EfficientNet but slightly more fluctuation in precision. Notably, in the Break/Hole/Collapse/Kink category, YOLOv11 achieved 81% accuracy with 66% precision and 76% recall, leading to a respectable F1-score of 68%. Overall, both models performed well, with EfficientNet exhibiting more consistent precision across categories and YOLOv11 showing higher recall, particularly for defect types where comprehensive detection is crucial.

Table 5-Results of multiclassification models

Model	Defect	Accuracy	Precision	Recall	F1 Score
EfficientNet	Crack-Fracture	94%	79%	68%	72%
	Break-Hole-Collapse-Kink	83%	76%	74%	75%
	Corrosion	96%	84%	70%	73%
	Deformation	79%	61%	60%	61%
	Joints	77%	64%	62%	63%
YOLOv11	Crack-Fracture	94%	72%	81%	75%
	Break-Hole-Collapse-Kink	81%	66%	76%	68%
	Corrosion	96%	81%	87%	83%
	Deformation	78%	59%	60%	59%
	Joints	81%	61%	69%	64%

4.3.2 Object Detection

In the next phase of the project, we extended our analysis by developing an object detection model aimed at localizing and classifying structural defects within individual video frames. Unlike the classification models used in previous stages, which only provided global labels for entire images, object detection allows for precise identification of the location, extent, and type of each defect present in a frame. This level of granularity is essential for implementing Utah's culvert rating system, which relies on both the presence and severity of localized defects to assign condition scores.

To train the detection models, we manually annotated thousands of frames by drawing bounding boxes around visible structural defects and assigning each box a corresponding class label, as defined in Table 3. These annotations served as the ground truth for model training and evaluation. However, due to significant class imbalance in the original structural defect labels, where some defect types were vastly underrepresented, we applied a similar label merging strategy used in the multiclass classification phase. Specifically, we merged related classes into broader categories to ensure more consistent training signals and improve detection accuracy for underrepresented defect types.

For model development, we implemented and tested two object detection architectures: YOLOv11 and YOLOv12, both of which are advanced versions of the YOLO family known for their real-time inference speed and accuracy. These models were trained and evaluated on the annotated dataset using standard object detection metrics such as mAP. The detailed performance results of the YOLOv12 model across the merged defect categories are presented in Table 7. These findings provide valuable insight into the models' ability to both detect and distinguish structural defects, forming the foundation for automated culvert condition assessment.

Table 6-Results of object detection models

Model	defects	Precision	Recall	Model performance (mAP)
YOLOv12	Bre-Hol-Col-Kin-5	89.1%	68.9%	80.3%
	Joints-3	84%	77.5%	83.8%
	Joints-5	87%	87.9%	91.6%
	Corrosion-3	88.7%	47.1%	68.8%
	Corrosion-5	90.9%	66.7%	79.3%
	Deformation-3	82.4%	44.3%	63.3%
	Deformation-5	94%	71.9%	84.4%
	Crack-Fract-3	84.5%	60.8%	75.1%
	All	87.6%	65.6%	78.3%
YOLOv11	Bre-Hol-Col-Kin-5	89.5%	48.9%	69%
	Joints-3	86.2%	56.1%	71.8%
	Joints-5	89.9%	74.8%	82.9%
	Corrosion-3	96.8%	25.2%	61%
	Corrosion-5	83.3%	66.7%	78.4%
	Deformation-3	75%	11.4%	42.4%
	Deformation-5	94%	51.6%	73.1%
	Crack-Fract-3	87.9%	37.9%	63.7%
	All	87.8%	46.6%	67.8%

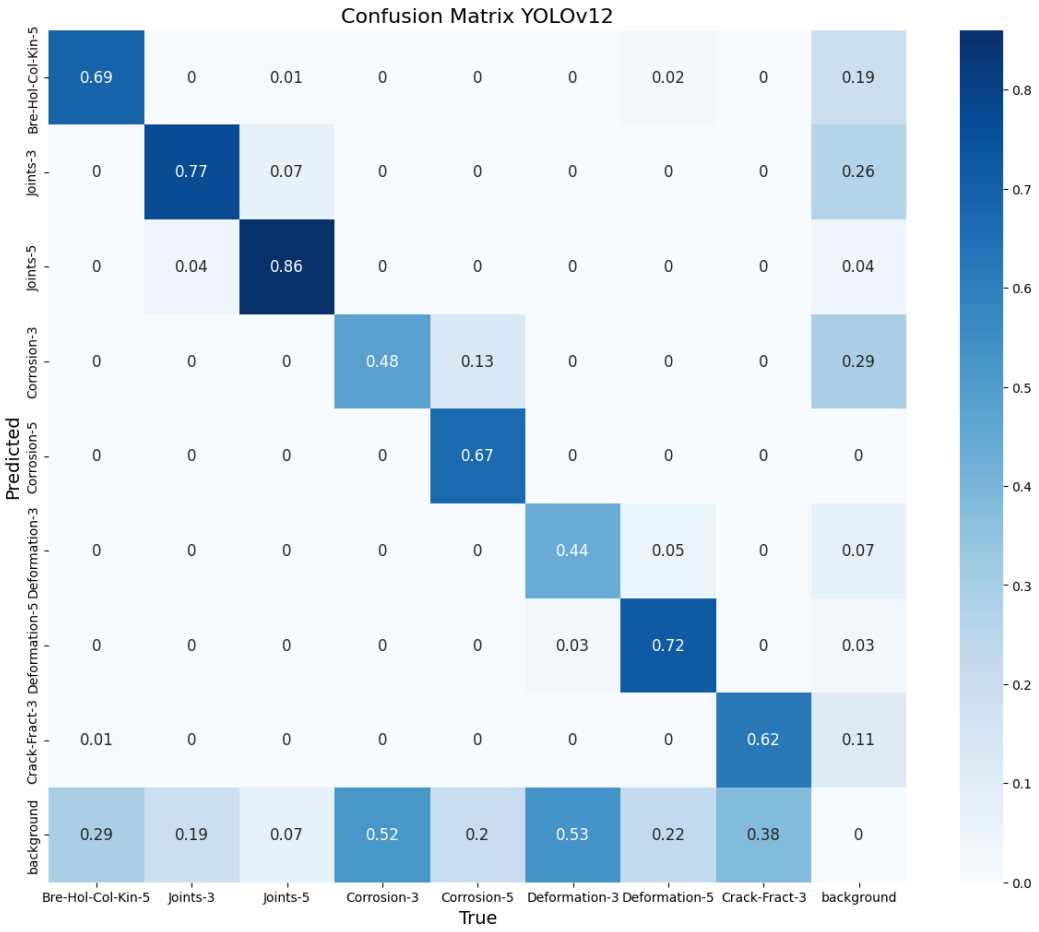

Figure 9-Normalized confusion matrix for the YOLOv12 model

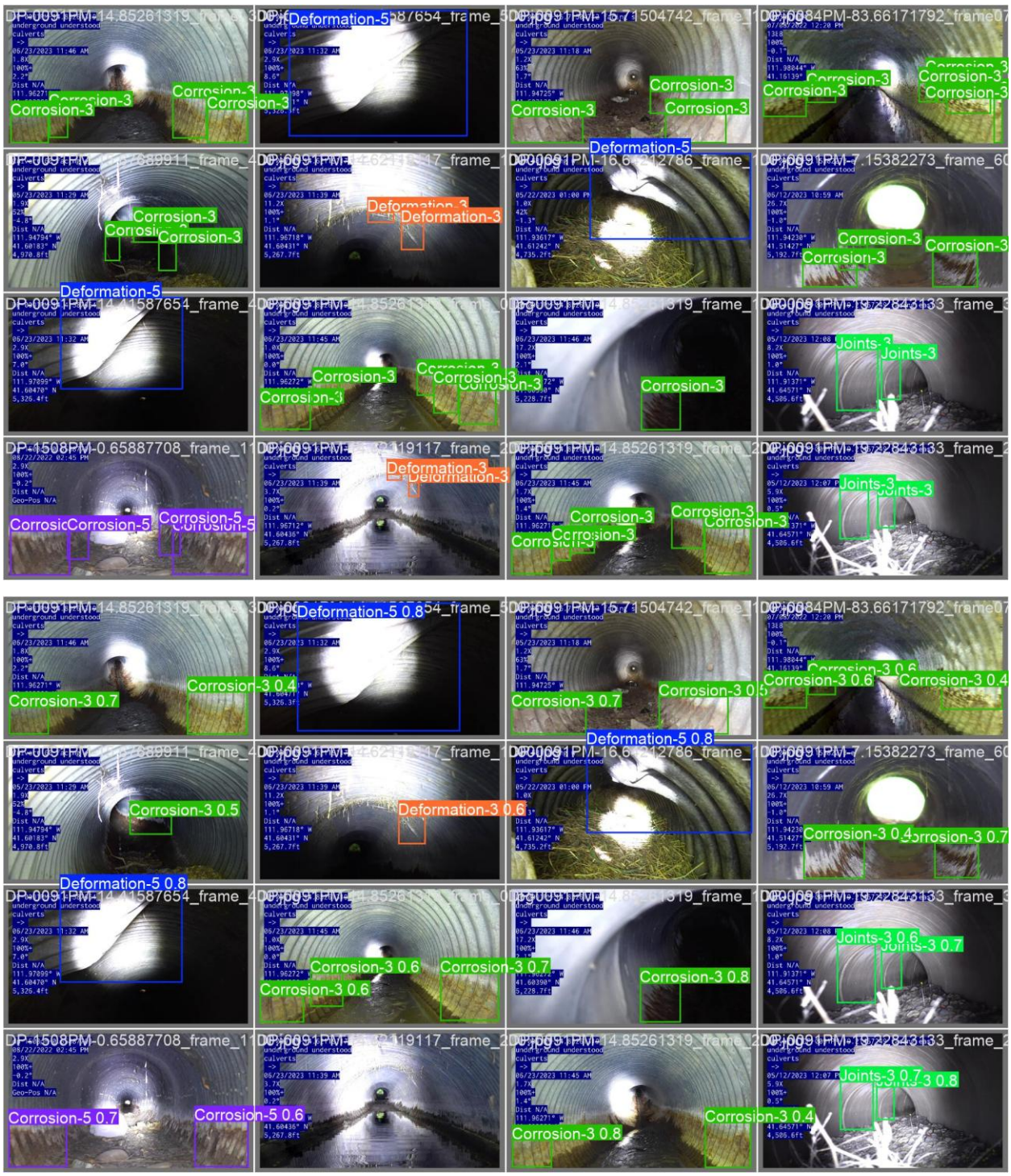

Figure 10-Examples of YOLOv12 model predictions (bottom images) and true labels (top images) for a batch of testing set

4.3.3 Graphical User Interface

To enable the practical use of the developed models by UDOT, we designed and implemented graphical user interfaces (GUIs) tailored to each type of model developed in this project. These GUIs are intended to provide a user-friendly experience for UDOT staff, allowing them to run complex video analysis tasks on culvert inspection footage without requiring programming knowledge. By simply launching the application on a laptop, a UDOT employee can import inspection videos and receive meaningful outputs, depending on the model selected for analysis.

Each GUI is specifically designed to align with the function and workflow of its corresponding model type, resulting in slightly different user experiences and outputs across the three interfaces. The first GUI is built for the binary classification model. In this interface, the user imports a culvert inspection video and selects an output directory. The model then automatically processes the video and filters out all frames identified as defective. These defective frames are saved in the designated output folder for further review. This tool acts as a rapid screening mechanism, allowing inspectors to focus on frames where defects are likely present.

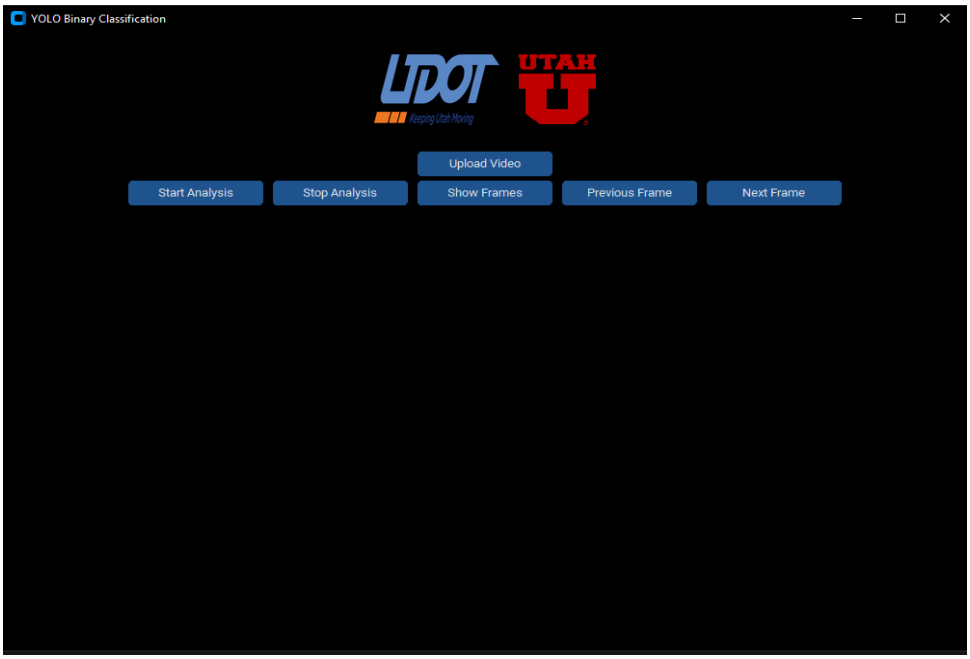

Figure 11-Binary classification-GUI

The second GUI is designed for the multiclass image classification models, which categorize detected defects into specific types. Upon launching the interface, the user inputs a culvert inspection video and designates a path for the output text file. Once the "Start Analysis" button is clicked, the system begins processing the video frame by frame. Each frame is classified by the trained multiclass model into one of several defect categories. The GUI employs a rule-based logic to track consistency in predictions: If the same defect label is detected in 15 consecutive frames, the GUI updates the predicted condition rating for the entire inspection video to reflect that defect type. If a more critical label is detected over 15 new consecutive frames at any point, the system updates the rating again to reflect the more severe condition. At the end of the analysis, the GUI generates a detailed text report named after the original video file. This report includes a summary of the inspection, the total number of frames associated with each defect label, their corresponding timestamps and frame numbers, and the final predicted rating for the video.

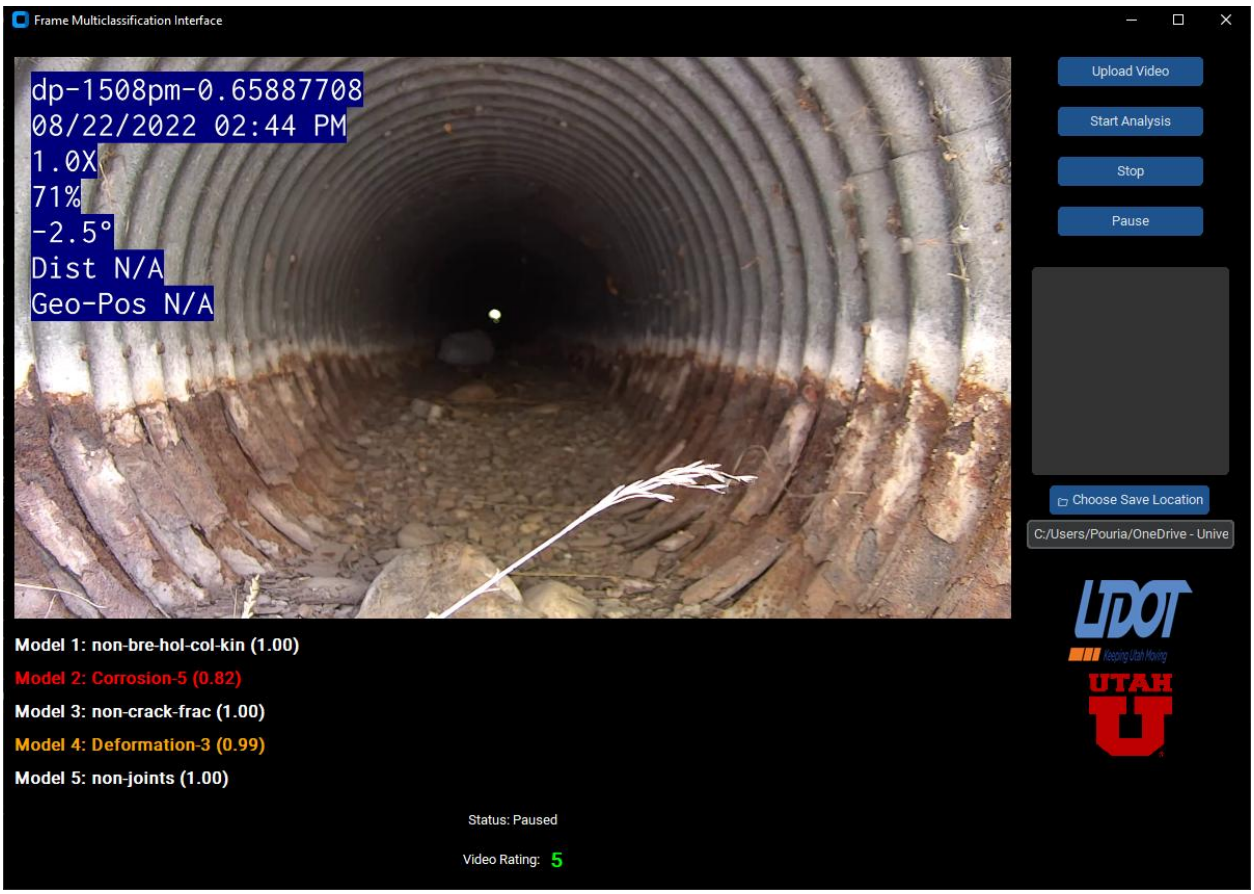

Figure 12-Multiclassification-GUI

The third GUI functions similarly to the second, but it is built around the object detection model. After importing the inspection video and specifying the output file location, the user starts the analysis process. The object detection model processes each video frame in real time, detecting and localizing structural defects by drawing bounding boxes around them. These boxes are labeled according to the predicted defect category. Similar to the multiclass classification GUI, if the model detects boxes of the same class in 15 consecutive frames, the video's rating is updated based on that defect type. If a more severe defect label is later detected over 15 consecutive frames, the system updates the rating to reflect the more critical condition. Once the analysis is complete, the GUI produces a comprehensive summary report, listing the detected defects, their bounding box coordinates, the timestamps and frame numbers where they occurred, and the final rating prediction for the entire inspection video.

Figure 13-Object detection-GUI

To evaluate the performance of both the Multiclassification-GUI and the Object Detection-GUI, we tested them on a set of 56 real-world culvert inspection videos provided by UDOT. These videos were previously assessed by inspectors, serving as a reliable benchmark for validation. Each GUI was used to independently analyze the videos and predict the overall condition of the culvert based on its trained model logic and output aggregation rules.

The results were highly promising, especially for the Object Detection-GUI, which correctly predicted the condition of 46 out of 56 videos, resulting in an accuracy of approximately 84%. This demonstrates the model's strong capability to localize and interpret structural defects with high reliability. In comparison, the Multiclassification-GUI achieved 42 correct predictions, corresponding to a 75% accuracy. While slightly less accurate, the multiclassification approach still provided valuable insights and maintained consistent performance across a variety of defect types.

Importantly, in most of the cases where the Object Detection-GUI failed to predict the correct condition, the model tended to be conservative, flagging defects where none were present according to human reviewers. This conservative bias may be preferable in certain safety-critical applications, as it prioritizes caution over risk. These misclassifications typically arose from subtle image features or lighting artifacts that resembled true defects, prompting the model to issue a higher severity rating than necessary.

Overall, the evaluation demonstrates the potential of these automated tools to streamline the inspection process, reduce subjectivity in condition assessments, and significantly save time, labor, and costs in culvert management. By automating defect recognition and condition rating from video footage, these GUIs can help prevent infrastructure failures and prioritize maintenance more effectively, especially in high-volume inspection workflows where manual review is impractical.

5.0 CONCLUSIONS

5.1 Summary

This chapter will discuss the conclusions obtained following the development of our computer vision models, showcasing their effectiveness. Additionally, we will discuss any limitations or challenges encountered during the research process, providing a comprehensive evaluation of the model's capabilities and areas for improvement.

To address the challenges associated with traditional culvert inspections in Utah, we proposed leveraging novel computer vision algorithms to enhance the interpretation of culvert inspection videos. This approach aims to significantly reduce the time and resources spent on manual interpretation, offering a more efficient and accurate alternative. To achieve this, we utilized the available culvert inspection videos and images from UDOT's database, labeled them using CVAT, and developed different computer vision models. This project introduced a multi-phase approach, including binary classification, multiclass image classification, and object detection models—each designed to progressively refine the detection and categorization of structural defects in inspection footage.

5.2 Findings

We started with a limited dataset of culvert inspection videos, which posed challenges for training reliable models. To address this, we expanded the dataset through additional video collection and data augmentation, enabling better model generalization and supporting the development of a multi-stage classification and detection framework.

The first model implemented was a binary classification model, designed to distinguish between defective and non-defective frames. Despite the limited initial data, the model achieved an impressive 91% accuracy, demonstrating its reliability in filtering out defective frames for further analysis. Building upon this, we developed a set of multiclass image classification models, each focused on categorizing specific types of structural defects. The best-performing model reached a 96% accuracy, while the least accurate achieved 77%, reflecting the varying difficulty across defect categories and class distribution within the dataset.

To gain spatial insight into the location and extent of defects, we developed an object detection model trained on annotated bounding box data. The final model achieved an average mAP of 78% across all defect classes. This performance indicates that the model not only classifies but also localizes defects with reasonable accuracy, providing detailed input for downstream culvert condition assessment.

The true effectiveness of the system was validated through testing on 56 real inspection videos. The object detection model correctly identified the condition of 84% of the culverts, confirming its practical viability for deployment. This result illustrates the potential of automated video analysis tools to support infrastructure inspection workflows, reduce subjectivity, and significantly improve efficiency in defect detection and rating.

Overall, the system demonstrates a successful progression from a small, constrained dataset to a fully functional, real-world application capable of assisting UDOT engineers in efficiently assessing culvert conditions. With further refinement and expansion, such tools hold strong promise for broader application across transportation infrastructure monitoring.

5.3 Limitations and Challenges

While the study produced robust results, several limitations and challenges were encountered:

- One of the primary challenges was the initial lack of sufficient labeled data. Although the dataset was expanded through additional data collection and augmentation, certain defect classes remained underrepresented. This class imbalance affected the training of both multiclass classification and object detection models, particularly in categories like Deformation, where performance metrics were lower due to fewer high-quality samples.
- Manual annotation of structural defects, especially bounding boxes in object detection, requires domain expertise and is prone to human error or inconsistency. Variation in how defects were labeled or categorized across frames may have impacted model training and evaluation. Inconsistent annotation boundaries or overlapping defect types also introduced noise into the dataset.
- Some structural defects, such as hairline cracks, subtle deformations, or joint displacements, are visually challenging to detect due to poor lighting, motion blur, or low resolution in inspection videos. These subtle visual cues often led to false negatives or incorrect classifications, especially in low-contrast or noisy frames.
- Users with limited hardware may experience slower inference times.

Despite these challenges, this research demonstrates the strong potential of computer vision and AI as transformative tools for automating culvert inspection in Utah. The success of the developed models, even under constrained conditions, highlights the feasibility of integrating AI-driven solutions into infrastructure maintenance workflows. Future work could explore the use of Vision-Language Models (VLMs), which may further simplify the process by enabling defect detection and condition assessment through natural language prompts, reducing the need for manual coding and huge datasets for training.

REFERENCES

[1] P. Mohammadi, A. Rashidi, and S. Asgari, “Improving Culvert Condition Prediction Models Using Federated Learning: The Case Study of Utah,” 2024.

[2] J. N. Meegoda, J. A. Kewalramani, and A. Saravanan, “Adapting 360-Degree Cameras for Culvert Inspection: Case Study,” *J. Pipeline Syst. Eng. Pract.*, vol. 10, no. 1, 2019, doi: 10.1061/(asce)ps.1949-1204.0000352.

[3] R. Rayhana, H. Yun, Z. Liu, and X. Kong, “Automated defect-detection system for water pipelines based on CCTV inspection videos of autonomous robotic platforms,” *IEEE/ASME Trans. Mechatronics*, vol. 29, no. 3, pp. 2021–2031, 2023.

[4] D. Youngblood and C. D. M. Smith, “Enhanced culvert inspections-best practices guidebook,” Minnesota. Dept. of Transportation. Research Services & Library, 2017.

[5] P. Mohammadi, A. Rashidi, and S. Asgari, “An Improved Hybrid XGBoost Model for Culvert Inspection Using Swarm Intelligence Algorithms,” *Computing in Civil Engineering 2023*. pp. 100–108, Jan. 25, 2024. doi: doi:10.1061/9780784485224.013.

[6] P. Mohammadi, A. Rashidi, and S. Asgari, “Simulating Federated Learning with Data Augmentation for Culvert Condition Prediction in Utah: A Case Study,” in *2024 Winter Simulation Conference (WSC)*, 2024, pp. 2337–2347.

[7] S. Grier, “Large culvert inspection procedures.” Purdue University, 2022.

[8] P. Mohammadi, S. Asgari, A. Rashidi, and R. Alder, “Culvert Inspection Framework Using Hybrid XGBoost and Risk-Based Prioritization: Utah Case Study,” *J. Constr. Eng. Manag.*, vol. 151, no. 6, p. 4025052, 2025.

[9] A. Hawari, M. Alamin, F. Alkadour, M. Elmasry, and T. Zayed, “Automated defect detection tool for closed circuit television (cctv) inspected sewer pipelines,” *Autom. Constr.*, vol. 89, pp. 99–109, 2018.

[10] X. Yin, Y. Chen, A. Bouferguene, H. Zaman, M. Al-Hussein, and L. Kurach, “A deep

learning-based framework for an automated defect detection system for sewer pipes," *Autom. Constr.*, vol. 109, p. 102967, 2020.

[11] S. S. Kumar, M. Wang, D. M. Abraham, M. R. Jahanshahi, T. Iseley, and J. C. P. Cheng, "Deep learning–based automated detection of sewer defects in CCTV videos," *J. Comput. Civ. Eng.*, vol. 34, no. 1, p. 4019047, 2020.

[12] X. Yin, T. Ma, A. Bouferguene, and M. Al-Hussein, "Automation for sewer pipe assessment: CCTV video interpretation algorithm and sewer pipe video assessment (SPVA) system development," *Autom. Constr.*, vol. 125, p. 103622, 2021.

[13] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, "Review of image classification algorithms based on convolutional neural networks," *Remote Sens.*, vol. 13, no. 22, p. 4712, 2021.

[14] M. Farhadmanesh, A. Rashidi, A. K. Subedi, and N. Marković, "A computer vision-based standalone system for automated operational data collection at non-towered airports," *IEEE Access*, 2024.

[15] S. Xu, M. Zhang, W. Song, H. Mei, Q. He, and A. Liotta, "A systematic review and analysis of deep learning-based underwater object detection," *Neurocomputing*, vol. 527, pp. 204–232, 2023.

[16] A. Hassandokht Mashhadi, A. Rashidi, and N. Marković, "A GAN-Augmented CNN Approach for Automated Roadside Safety Assessment of Rural Roadways," *J. Comput. Civ. Eng.*, vol. 38, no. 2, p. 4023043, 2024.

[17] M. Sohan, T. Sai Ram, and C. V. Rami Reddy, "A review on yolov8 and its advancements," in *International Conference on Data Intelligence and Cognitive Informatics*, 2024, pp. 529–545.

[18] R. Khanam and M. Hussain, "Yolov11: An overview of the key architectural enhancements," *arXiv Prepr. arXiv2410.17725*, 2024.

[19] Y. Tian, Q. Ye, and D. Doermann, "Yolov12: Attention-centric real-time object

detectors," *arXiv Prepr. arXiv2502.12524*, 2025.

[20] R. Sapkota, R. H. Cheppally, A. Sharda, and M. Karkee, "RF-DETR Object Detection vs YOLOv12: A Study of Transformer-based and CNN-based Architectures for Single-Class and Multi-Class Greenfruit Detection in Complex Orchard Environments Under Label Ambiguity," *arXiv Prepr. arXiv2504.13099*, 2025.

[21] L. Zhang, G. Ding, C. Li, and D. Li, "DCF-Yolov8: an improved algorithm for aggregating low-level features to detect agricultural pests and diseases," *Agronomy*, vol. 13, no. 8, p. 2012, 2023.

[22] P. Mohammadi, A. Rashidi, and S. Asgari, "Privacy-preserving culvert predictive models: A federated learning approach," *Adv. Eng. Informatics*, vol. 61, p. 102483, 2024.

[23] S. Yang, W. Xiao, M. Zhang, S. Guo, J. Zhao, and F. Shen, "Image data augmentation for deep learning: A survey," *arXiv Prepr. arXiv2204.08610*, 2022.

[24] A. Erfani, N. Shayesteh, and T. Adnan, "Data-augmented explainable AI for pavement roughness prediction," *Autom. Constr.*, vol. 176, p. 106307, 2025.

[25] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. big data*, vol. 6, no. 1, pp. 1–48, 2019.

[26] B. Dwyer, J. Nelson, and J. ea Solawetz, "Roboflow (version 1.0)[software]," *Comput. Vis.*, 2024.

[27] B. Sekachev *et al.*, "opencv/cvat: v1.1.0." Zenodo, Aug. 2020. doi: 10.5281/zenodo.4009388.

[28] P. Mohammadi, A. Rashidi, M. Malekzadeh, and S. Tiwari, "Evaluating various machine learning algorithms for automated inspection of culverts," *Eng. Anal. Bound. Elem.*, vol. 148, pp. 366–375, 2023, doi: https://doi.org/10.1016/j.enganabound.2023.01.007.

[29] A. H. Mashhadi, N. Markovic, and A. Rashidi, "Estimating Construction Work Zones Capacity Using Deep Neural Network," in *Construction Research Congress 2022*, 2022,

pp. 98–107.

[30] A. Hassandokht Mashhadi, A. Rashidi, J. C. Medina, and N. Marković, “A hybrid framework for predicting crash severity in construction work zones using knowledge distillation and conditional GANs,” *J. Comput. Civ. Eng.*, vol. 39, no. 2, p. 4025010, 2025.

APPENDIX A: UDOT's PIPE DEFECT RATING SHEETS

UDOT PLASTIC OR HDPE PIPE RATING SYSTEM

CATEGORY	MINOR DEFECTS		MODERATE DEFECTS		SIGNIFICANT DEFECTS		MAJOR DEFECTS		CRITICAL DEFECTS	
	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE
SHAPE	Minor bumps or bulges - no change in diameter - Area is less than 2 in. diameter	1	Bumps and bulges in pipe - greater than 2 in. diameter - no inside diameter lost	2	This refers to bulges or vertical deformation in pipe. No cracking or fractures present. ≤5% of inside diameter lost. * Minor wall flattening (≤5%).	3	This refers to bulges or vertical deformation in pipe. No cracking or fractures present. ≤5% to >10% of inside diameter lost. * Visible out of roundness (elliptical shape) with no cracks. * Significant wall flattening (>5% to ≤10%) or increased wall curvature.	4	This refers to bulges or vertical deformation in pipe. No cracking or fractures present. >10% of inside diameter lost. * Significant visible out of roundness (elliptical shape) >10% with no cracks. * Extreme wall flattening (>10%) with reversal of curvature (global bucking) and/or kinks. * A defect where the inward bulge is sharp crested taking shape of heart point or shark fin. * A sharp outward folding of pipe wall.	5
SURFACE DAMAGE	Blisters or degradation at single location - less than 6 in. diameter	1	Blisters at multiple locations - less than 10% of surface covered	2	Damage to surface due to erosion or wear, ≤10% wall thickness removed. * Ultraviolet degradation - based on amount of degradation shown, Minor amount. * Blisters on wall - < 25% of surface covered.	3	Damage to surface due to erosion or wear, >10% to ≤25% wall thickness removed. * Ultraviolet degradation - based on amount of degradation shown - Pipe ends showing discoloration. * Blisters on wall - ≥ 25% of surface covered.	4	Damage to surface due to erosion or wear, >25% wall thickness removed. * Ultraviolet degradation - based on amount of degradation shown - Degradation resulting of cracked or broken pipe walls.	5
LOCAL BUCKLING, SPLITS AND CRACKS	Crack that is perpendicular to flow direction. No opening between crack. One max per pipe section. Less than 1/4 of circumference.	1	Longitudinal crack ≤ 12 in. in length with or without water infiltration - no soil infiltration. * Crack that changes from perpendicular to longitudinal (or reverse). * Circumferential crack between 1/4 of diameter and 1/2 of diameter. * Initiation of local bucking indicated by rippling in wall.	2	Combination of Circumferential and Longitudinal cracks or multiple number of each in pipe section. * Water infiltration through circumferential cracks. * Two longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions) ≤ 12 in. in length. * Advanced and widespread local wall bucking indicated by extensive interior surface ripping.	3	Circumferential cracks ≥ 1/2 of pipe circumference. * Cracks/Fractures with significant soil migration or water infiltration. * Cracks/Fractures with vertical offset - pieces of pipe has moved. * Two longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions) > 12 in. in length. * Cracks with soil infiltration. * Pipe wall buckles inward locally. * Kinks through full wall thickness.	4	Broken Pipe - can see soil. * Broken Pipe - Can see void behind pipe. * Hole in pipe. * Collapsed Pipe. * Three or four longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions).	5
BARREL ALIGNMENT	Horizontal alignment shows small visible deviations (<5%) from installed conditions and does not affect joints or barrel. * Vertical alignment has minor sagging or heaving (<5%).	1	Vertical misalignment with sags < 10% with sediment accumulation	2	Change in alignment greater than (>) 5° and less than or equal to (≤) 10°. * Alignment deviations that affect condition of joints or barrel. * Vertical misalignment causing ponding or sediment accumulation at sags between 10% and 30% of diameter	3	Change in alignment greater than (>) 10°. * Alignment deviations that cause breakage at joints or barrel. * Vertical misalignment causing ponding or sediment accumulation at sags > 30% of diameter.	4	Changes in alignment that cause hole in pipe. * Changes in alignment causing blockage of pipe	5
JOINTS	Offset is visible at joint with no effect on pipe - not a quantifiable amount of offset	1	Offset is visible but less than 1 wall thickness.	2	Offset is greater than or equal to (≥) 1 pipe wall thickness but less than (<) 1.5 wall thickness - no distress visible. * Separation is up to 1 pipe wall thickness - no distress visible. * Exposed or missing gasket materials. * Infiltration/exfiltration or soil migration through joints - no structural damage. * Roots visible through joints - no structural damage.	3	Offset is greater than or equal to (≥) 1.5 pipe wall thickness. * Separation is greater than (>) 1 pipe wall thickness. * Possible exposed joint sealant. * Infiltration/exfiltration or soil migration through joints - visible structural damage. * Roots visible through joints - structural damage. * Joint distress directly causes distress to barrel/end section, roadway/shoulder, or embankment.	4	Offset joint where soil is showing	5

UDOT CORRUGATED METAL PIPE RATING SYSTEM

CATEGORY	MINOR DEFECTS		MODERATE DEFECTS		SIGNIFICANT DEFECTS		MAJOR DEFECTS		CRITICAL DEFECTS	
	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE
SURFACE DAMAGE	Single dent or bulge - no change in diameter - Area is less than 2 in. diameter	1	Multiple dents or bulges - Total area less than 4 inches diameter	2	Small dents or impact damage to pipe wall or end section with no wall breaches - area greater than 4 inches diameter.	3	Large dents or impact damage to pipe wall section with localized wall breaches, no more than one corrugation over circumferential length of 6 in.	4	Dents or damage that warrant engineering inspection. * Through-wall holes > 1 corrugation over a length of more than 6 in. allowing unimpeded soil infiltration.	5
CORROSION	Single area of freckeled rust	1	Isolated areas of freckled rust.	2	Freckled rust, corrosion of pipe wall material.	3	Corrosion of pipe material and widespread section has loss <10% of wall thickness. * Localized deep pitting. * Several holes (< 4 per square yard) less ≤ 1 in. diameter. * Penetration possible with hammer pick strike.	4	Widespread through wall penetration/corrosion. * Invert missing in localized section. * Holes > 1 in. diameter or holes grouped together > 4 per square yard.	5
ABRASION	Visible abrasion at single location less than 6 inches diameter	1	Visible abrasion of wall or coating at 2 locations with total affected area less than 12 inches diamter	2	Small or local abrasion of wall or coating at more than 2 locations or area gtreater than 12 inches diameter with no breaches in the coating exposing structural wall if signs of corrosion.	3	Widespread abrasion of protective coating with breaches exposing the pipe material and allowing through-wall penetration during inspection probing with pick.	4	Abrasion has worn holes in pipe.	5
SHAPE	Visible deformation. Isolated at single corrugation	1	Smooth curvature of barrel, deformation <5% of inside diameter.	2	Deformation of barrel ≥5% to 10% of inside diameter. * Minor wall flattening or bulges (≤5%).	3	Deformation of barrel ≥10% to 15% of inside diameter. * Visible out of roundness (elliptical shape) with no cracks. * Significant wall flattening (>5% to ≤10%) or increased wall curvature.	4	Deformation of barrel ≥15% of inside diameter. * Significant visible out of roundness (elliptical shape) >10% with no cracks. * Extreme wall flattening (>10%) with reversal of curvature (global bucking) and/or kinks. * A defect where the inward bulge is sharp crested taking shape of heart point or shark fin. * A sharp outward folding of pipe wall.	5
CRACKS BREAKS KINKS HOLES	Crack that is perpendicular to flow direction. No opening between crack. One max per pipe section. Less than 1/4 of circumference.	1	Longitudinal crack ≤ 12 in. in length with or without water infiltration - no soil infiltration. * Crack that changes from perpendicular to longitudinal (or reverse). * Circumferential crack between 1/4 of diameter and 1/2 of diameter. * Initiation of local bucking indicated by rippling in wall.	2	Combination of circumferential and longitudinal cracks or multiple number of each in pipe section. * Water infiltration through circumferential cracks. * Two longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions) ≤ 12 in. in length. * Advanced and widespread local wall bucking indicated by extensive interior surface ripping.	3	Circumferential cracks ≥ 1/2 of pipe circumference. * Cracks/Fractures with significant soil migration or water infiltration. * Cracks/Fractures with vertical offset - pieces of pipe has moved. * Two longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions) > 12 in. in length. * Cracks with soil infiltration. * Pipe wall buckles inward locally. * Kinks through full wall thickness.	4	Broken Pipe - can see soil. * Broken Pipe - can see void behind pipe. * Hole in pipe. * Collapsed Pipe. * Three or four longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions).	5
BARREL ALIGNMENT	Horizontal alignment shows small visible deviations (<5%) from installed conditions and does not affect joints or barrel. * Vertical alignment has minor sagging or heaving (<5%).	1	Vertical misalignment with sags < 10% with sediment accumulation	2	Change in alignment greater than (>) 5° and less than or equal to (≤) 10°. * Alignment deviations that affect condition of joints or barrel. * Vertical misalignment causing ponding or sediment accumulation at sags between 10% and 30% of diameter.	3	Change in alignment greater than (>) 10°. * Alignment deviations that cause breakage at joints or barrel. * Vertical misalignment causing ponding or sediment accumulation at sags > 30% of diameter.	4	Changes in alignment that cause hole in pipe. * Changes in alignment causing blockage of pipe	5
JOINTS	Offset is visible with no effect on pipe - not a quantifiable amount of offset	1	Offset is visible but less than 1 wall thickness.	2	Offset is greater than or equal to (≥) 1 pipe wall thickness but less than (<) 1.5 wall thickness - no distress visible. * Separation is up to 1 pipe wall thickness - no distress visible. * Exposed or missing gasket materials. * Infiltration/exfiltration or soil migration through joints - no structural damage. * Roots visible through joints - no structural damage.	3	Offset is greater than or equal to (≥) 1.5 pipe wall thickness. * Separation is greater than (>) 1 pipe wall thickness. * Possible exposed joint sealant. * Infiltration/exfiltration or soil migration through joints - visible structural damage. * Roots visible through joints - structural damage. * Joint distress directly causes distress to barrel/end section, roadway/shoulder, or embankment.	4	Offset joint where soil is showing	5
INFILTRATION EXFILTRATION	Signs of past infiltration (staining) at isolated location - no current infiltration	1	Signs of past infiltration (staining) at multiple locations -no current infiltration	2	Minor water infiltration through leak-resistant seams, but no soil infiltration.	3	Significant water infiltration and evidence of fine soils infiltrating through seams. * Evidence of piping due to exfiltration.	4	Course soil infiltration through seam openings. * Possible hollow sounds behind structure wall near seams indicating loss of backfill support.	5

UDOT CORRUGATED METAL PIPE RATING SYSTEM

CATEGORY	MINOR DEFECTS		MODERATE DEFECTS		SIGNIFICANT DEFECTS		MAJOR DEFECTS		CRITICAL DEFECTS	
	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE
SEAM ALIGNMENT	Seams minorly out of alignment - with no affect on pipe	1	Slight cocked seams without cusp effect, but does not affect cross section shape.	2	Cocked seams that it affects cross section shape. * Cusped effect with local wall bending.	3	Cocked seams severely affecting cross section shape. * Cusp effect with seam cracking. Seam capacity loss imminent.	4	Seam cracking causing failure or holes	5
SEAM BOLTS/ FASTENERS	Single missing bolt	1	<5% loose or missing bolts in any seam.	2	5% to 15% loose or missing bolts in any seam.	3	> 15% missing bolts in any seam.	4	> 50% missing bolts in any seam	5
SEAM BOLT HOLES	Cracking at single bolt hole	1	Minor yielding of steel and/or cracking/splitting < 1 in. long local to bolt holes. * Minor corrosion developing around bolt holes or on bolts.	2	Yielding of steel and/or cracking/splitting 1 in. up to 3 in. long local to bolt holes. * Corrosion with section loss around bolt holes or on bolts.	3	Significant yielding of steel at bolt holes. * Cracking/splitting >3 in. long local to bolt holes. * Corrosion with major section loss around bolt holes or on bolts.	4	Bolt holes corroded to level that no bolts can be replaced - over 50% of bolt holes	5

UDOT TIMBER PIPE RATING SYSTEM

CATEGORY	MINOR DEFECTS		MODERATE DEFECTS		SIGNIFICANT DEFECTS		MAJOR DEFECTS		CRITICAL DEFECTS	
	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE
CONNECTIONS AND MISSING MEMBERS	Single loose bolt or fastener	1	Two loose bolts or fasteners (not on single member)	2	Multiple loose bolts and fasteners. * Freckled rust (no pitting or section loss), rust staining (connection is functioning as designed).	3	Missing bolts, rivets or fasteners, broken welds. * Surface rusting with some pitting, pack rust without distortion (connection is functioning as designed).	4	Connection integrity in question, imminent collapse, missing members, collapsed section. * Missing bolts, rivets, or fasteners, broken welds causing movement in connection elements. * Heavy rusting with section loss, and/or pack rust causing distortion.	5
DECAY	Visible decay - no penetration	1	Visible decay - surface scraping of material only	2	Decay allowing probe penetration ≤10% of member cross section. * Localized hollow sounds.	3	Decay allowing probe penetration >10 % to ≤20% of member cross section, but is away from connections and tension of bending member. * Fruiting bodies.	4	Probe penetrates > 20% of cross section. * Probe penetrates > 10% of cross section near connections or in tension zone of bending member.	5
CHECKS AND SHAKES	Checks or shakes penetrating <5% of member thickness.	1	Checks or shakes penetrating 5% to 15% of member thickness, but away from connection and tension zones of bending members.	2	Checks or shakes penetrating 15% to 50% of member thickness, but away from connection and tension zones of bending members.	3	Checks or shakes penetrating >50% of member thickness. * Checks or shakes penetrating 5% to 10% of member thickness, at connection and tension zones of bending members.	4	Checks or shakes penetrating >10% of member thickness, at connection and tension zones of bending members.	5
SHAPE	Minor deflection visible, but not quantifiable	1	Smooth curvature of barrel, deformation <5% of inside diameter.	2	Deformation of barrel ≥5% to 10% of inside diameter. * Minor wall flattening or bulges (≤5%).	3	Deformation of barrel ≥10% to 15% of inside diameter. * Visible out of roundness (elliptical shape) with no cracks. * Significant wall flattening (>5% to ≤10%) or increased wall curvature.	4	Deformation of barrel ≥15% of inside diameter. * Significant visible out of roundness (elliptical shape) >10% with no cracks. * Extreme wall flattening (>10%) with reversal of curvature (global bucking) and/or kinks. * A defect where the inward bulge is sharp crested taking shape of heart point or shark fin. * A sharp outward folding of pipe wall.	5
STRUCTURAL CRACKS	Shrinkage cracks - not structural	1	Structural cracks have been arrested.	2	Structural cracking exists, but projects < 5% into member cross section.	3	Structural cracking ≥5% to 25% into member cross section.	4	Structural cracking ≥25% into member cross section.	5
DELAMINATION	Minor surface delamination at a single isolated location - less than 12 in diameter	1	Minor surface delamination at a single isolated location - less than 24in diameter	2	Delamination length less than the total member depth and away from connections and tension zones of bending members.	3	Delamination length ≥ total member depth and away from connections and tension zones of bending members.	4	Delamination near connections or in tension zones, imminent collapse of member or structure.	5
ABRASION/ IMPACT DAMAGE	Minor abrasion to surface from impacts - no damage	1	Minor abrasion damage due to impacts - no member section loss	2	Section loss < 10% of member cross section.	3	Section loss 10% to 20% of member cross section.	4	Section loss > 20% of member cross section.	5
DISTORTION	Minor observed sagging of single member - amount of sagging not quantifiable	1	Minor observed sagging of multiple non adjacent member - amount of sagging not quantifiable	2	Warping or sagging of single or few members not requiring mitigation or has been previously mitigated.	3	Warping or sagging causing distortion of cross sectional shape. * Crushing of members.	4	Significant distortion of cross sectional shape or widespread warping, crushing or sagging.	5

UDOT MASONRY PIPE RATING SYSTEM

CATEGORY	MINOR DEFECTS		MODERATE DEFECTS		SIGNIFICANT DEFECTS		MAJOR DEFECTS		CRITICAL DEFECTS	
	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE
MASONRY UNITS AND MOVEMENT	Minor stess or expansion cracking - surface cracking only	1	Cracking of individual units. * Surface weathering or spalling.	2	Split or cracked masonry units. * Large areas of moderate spalling, scaling or weathering. * Pronounced movement or dislocation of masonry units, but does not warrant engineering evaluation.	3	Widespread cracking, splitting, splitting, or crushing of masonry units, or missing units. * Large areas of heavy spalling, scaling or weathering. * Significant movement of individual units.	4	Holes through structure, units missing for entire cross section. * Visible movement or distortion of cross sectional shape, structure appears unstable.	5
MORTAR	Vegetation/roots sprouting between units, no widespread missing mortar.	1	Localized cracked or missing mortar (<10%). * Widespread areas of shallow mortar deterioration, possible minor water infiltration (no active flow) or exfiltration.	2	10% to 50% of mortar missing, no unit movement. * Extensive mortar deterioration, small flow but no fines, infiltration or exfiltration through joints.	3	>50% of mortar missing, no unit movement. * Large roots through joints (no unit movement).	4	Backfill infiltration. * Roadway voids. * Mortar missing or large roots with unit movement.	5
EFFLORESCENCE	Localized areas of efflorescence < 2 in^2.	1	Widespread areas of efflorescence without rust staining.	2	Heavy buildup of efflorescence with rust staining.	3	Exposed rebar	4	Broken or missing rebar	5

UDOT MANHOLES, CATCH BASINS, AND BURIED JUNCTION STRUCTURES RATING SYSTEM

CATEGORY	MINOR DEFECTS		MODERATE DEFECTS		SIGNIFICANT DEFECTS		MAJOR DEFECTS		CRITICAL DEFECTS	
	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE
MATERIAL DEGRADATION OF INSIDE SURFACE	Crack (crack is a line in pipe that has not shown opening or deformation) that is vertical. No opening between crack. One max per manhole.	1	Multiple cracking between 0.01 in. and 0.05 in. width horizontal to grade. Single crack around interior or exterior (if visible) of manhole. * Moisture on wall from seepage. * Grate, MH Cover, slightly off proper grade. * Localized spalls less than 1/2 in. depth and less than 6 in. diameter. No exposed rebar. * Ladder and attachments have surface corrosion or light pitting. * Efflorescence but no rust emanating from crack. * Single open crack (fracture) - vertical. * Missing brick in brick/masonry manhole in chimney, wall, or bench. No visible soil or void.	2	Split or cracked masonry units. * Missing mortar in brick or masonry manhole. * Slight discoloration of masonry units. * Spalling and/or delamination from 1/2 in. to 3/4 in. in depth and larger than 6 in. diameter. No exposed rebar. Some rust staining from spalled areas, structure stable. * Ladder and attachments have heavy corrosion, pitting on surface, minor loss of section. * Displaced structural elements, minor visible movement of masonry units. * Infiltration - no soils present. * Efflorescence and rust emanating from crack/fracture. * Exterior manhole cracking - are above grade. * Single open crack (fracture) - horizontal.	3	Widespread cracking, splitting, splitting, or crushing of masonry units, or missing units. * Significant movement of individual brick or masonry units. * Spalling with exposed or minor corrosion of rebar - rebar still intact. * Widespread spalling greater than 3/4 in. in depth or delamination. * Slabbing of concrete. * Ladder and attachments as heavy corrosion, pitting on surface, loss of section, not safe. * Multiple open cracks (fractures) on inside or outside of manhole. * Significant infiltration with soils. * Minor change in shape of masonry cross section. * Cracks/Fractures with significant soil migration or water infiltration. * Cracks/Fractures with vertical offset - pieces of pipe have moved. * Large areas of rust staining emanating from cracks/fractures. * Manhole frame and cover offset from manhole.	4	Holes in concrete manhole. * Visible movement or distortion of cross sectional shape, structure appears unstable. * Visible corrosion of rebar. * Major distortion in shape of masonry cross section. * Masonry units missing through structure wall. * Manhole frame or cover broken. * Holes in brick manhole with soil visible or void visible. * Hole in brick manhole in channel. * Collapsed manhole. * Offset joints in concrete manhole.	5
JOINT WITH PIPE	Cracking of mortar around pipe/manhole connection	1	Missing pieces of mortar around connection between pipe and manhole - no infiltration or distress	2	Small joint separation but no infiltration and no indication of distress. * Joint separation, offset, or rotation.	3	Indication of distress to pipe or structure wall.	4	Joint separations, offset, or rotation with significant backfill infiltration and pipe vertical offset with exposed backfill material.	5

UDOT CONCRETE PIPE RATING SYSTEM

CATEGORY	MINOR DEFECTS		MODERATE DEFECTS		SIGNIFICANT DEFECTS		MAJOR DEFECTS		CRITICAL DEFECTS	
	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE	DESCRIPTION	SCORE
CRACKS *(< 0.05 INCHES)* **FRACTURES** *(≥ 0.05 INCHES)*	Crack (not showing signs of opening or movement) that is perpendicular to flow direction. One max per pipe section.	1	Crack that extends along pipe longitudinally. Can be a single crack at a hinge point. * Crack that changes from perpendicular to longitudinal (or reverse). * Efflorescence but no rust emanating from crack. * Two longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions). * Fracture that is perpendicular to flow direction. One max per pipe section.	2	Combination of Circumferential and Longitudinal cracks or multiple number of each in pipe section. * Water infiltration through circumferential cracks. * Efflorescence and rust emanating from crack/fracture. * Fracture that extends along pipe. Described per pipe section. Can be a single fracture at a hinge point. * Three longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions). * Fracture that may start as longitudinal and change to circumferential or the reverse. Does not cross a joint. * Two longitudinal fractures located at hinge points (12, 3, 6, 9 o'clock positions).	3	Three or Four longitudinal cracks located at hinge points (12, 3, 6, 9 o'clock positions). * Cracks/Fractures with significant soil migration or water infiltration. * Cracks/Fractures with vertical offset - pieces of pipe have moved. * Large areas of rust staining emanating from cracks/fractures.	4	Broken Pipe - can see soil. * Broken Pipe - can see void behind pipe. * Hole in pipe. * Collapsed Pipe.	5
SLABBING **SPALLING** **DELAMINATION** **PATCHES**	Minor spalling of less than 1/2 in. depth and less than 2 in. diameter. No exposed rebar	1	Localized spalls less than 1/2 in. depth and less than 6 in. diameter. No exposed rebar.	2	Spalling and/or delamination from 1/2 in. to 3/4 in. in depth and larger than 6 in. diameter. No exposed rebar. Some rust staining from spalled areas, structure stable.	3	Patched areas that are delaminated or deteriorating. * Widespread spalling greater than 3/4 in. in depth or delamination. * Slabbing of concrete. * Spalling with exposed or corroded rebar.	4	Not Applicable	
DETERIORATION	Multiple plugged weep holes. * Slight damage to surface, minor wear.	1	Pipe cement material is eroded or worn to level that aggregate is showing - abrasion less than 0.25 in. deep over less than 20% of pipe surface cross section.	2	Moderate to severe scaling - pipe cement material is eroded or worn to level that aggregate is projecting above level of remaining cement mix. * Pipe cement material is eroded or worn to level that aggregate is showing - abrasion between 0.25 in. and 0.5 in. deep over less than 30% of pipe surface cross section.	3	Pipe cement material is eroded or worn to level that aggregate is missing at locations and there are pockets in the wall - rebar not exposed. * Impact damage with exposed rebar.	4	Pipe has deteriorated to level where the rebar has corroded but not broken. * Pipe has deteriorated to level where the rebar has corroded but not broken. * Pipe has deteriorated to level where the rebar has failed and broken such that pieces are sticking out of wall. * Complete invert deterioration and loss of pipe wall section.	5
BARREL ALIGNMENT	Horizontal alignment shows small visible deviations (<5%) from installed conditions and does not affect joints or barrel. * Vertical alignment has minor sagging or heaving (<5%).	1	Vertical misalignment with sags < 10% with sediment accumulation	2	Change in alignment greater than (>) 5° and less than or equal to (≤) 10°. * Alignment deviations that affect condition of joints or barrel. * Vertical misalignment causing ponding or sediment accumulation at sags between 10% and 30% of diameter.	3	Change in alignment greater than (>) 10°. * Alignment deviations that cause breakage at joints or barrel. * Vertical misalignment causing ponding or sediment accumulation at sags > 30% of diameter.	4	Changes in alignment that cause hole in pipe. * Changes in alignment causing blockage of pipe	5
JOINTS	Offset is visible at joint with minor joint material showing	1	Offset is visible but less than 1 wall thickness. * Moderate spall along edge of spigot end.	2	Offset is greater than or equal to (≥) 1 pipe wall thickness but less than (<) 1.5 wall thickness - no distress visible. * Separation is up to 1 pipe wall thickness - no distress visible. *Exposed or missing gasket materials. * Large spalls along edge of spigot end. * Infiltration/exfiltration or soil migration through joints - no structural damage. * Roots visible through joints - no structural damage.	3	Offset is greater than or equal to (≥) 1.5 pipe wall thickness . * Separation is greater than (>) 1 pipe wall thickness. * Possible exposed reinforcement or joint sealant. * Infiltration/exfiltration or soil migration through joints - visible structural damage. * Roots visible through joints - structural damage. * Joint distress directly causes distress to barrel/end section, roadway/shoulder, or embankment.	4	Offset joint where soil is showing	5