Vision-Based Detection of Bridge Damage Captured by Unmanned Aerial Vehicles

Final Report
September 2025

Principal Investigator: Paolo Stegagno

Electrical Computer and Biomedical Engineering
University of Rhode Island

Authors
Paolo Stegagno, Mayrai Gindy, Abdeltawab Hendawi, Stephen Licht

Sponsored By
Transportation Infrastructure Durability Center
List other Sponsors if applicable (i.e. MaineDOT)

TIDC

Transportation Infrastructure Durability Center
AT THE UNIVERSITY OF MAINE

A report from
University of Rhode Island
Department of Electrical Computer and Biomedical Engineering
2 E Alumni Ave FCAE 360
Phone: 401.874.5814
Website: https://web.uri.edu/cypher/meet/dr-paolo-stegagno/

About the Transportation Infrastructure Durability Center

The Transportation Infrastructure Durability Center (TIDC) is the 2018 US DOT Region 1 (New England) University Transportation Center (UTC) located at the University of Maine Advanced Structures and Composites Center. TIDC's research focuses on efforts to improve the durability and extend the life of transportation infrastructure in New England and beyond through an integrated collaboration of universities, state DOTs, and industry. The TIDC is comprised of six New England universities, the University of Maine (lead), the University of Connecticut, the University of Massachusetts Lowell, the University of Rhode Island, the University of Vermont, and Western New England University.

U.S. Department of Transportation (US DOT) Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgements

Funding for this research is provided by the Transportation Infrastructure Durability Center at the University of Maine under grant 69A3551847101 from the U.S. Department of Transportation's University Transportation Centers Program. [Include any acknowledgements for other contributors (i.e. your university or contributing DOTs/industry partners) here.]

Technical Report Documentation Page

1. Report No.	**2. Government Accession No.**	**3. Recipient Catalog No.**
4 Title and Subtitle Vision-Based Detection of Bridge Damage Captured by Unmanned Aerial Vehicles		**5 Report Date** September 2025 **6 Performing Organization Code**
7. Author(s) Paolo Stegagno https://orcid.org/0000-0003-1241-5732 Mayrai Gindy https://orcid.org/0000-0001-9322-4625 Abdeltawab Hendawi https://orcid.org/0000-0003-3385-3379 Stephen Licht https://orcid.org/0000-0002-9887-7681		**8 Performing Organization Report No.**
9 Performing Organization Name and Address University of Rhode Island, Department of Electrical, Computer and Biomedical Engineering		**10 Work Unit No. (TRAIS)** **11 Contract or Grant No.**
12 Sponsoring Agency Name and Address Transportation Infrastructure Durability Center (TIDC)		**13 Type of Report and Period Covered** Final Report; September 2022 – September 2025 **14 Sponsoring Agency Code**
15 Supplementary Notes		

16 Abstract

Bridge inspection is a vital component of any bridge management strategy of a state DOT. Significant funds are allocated to keeping the over 600,000 bridges in the U.S. safe. A routine bridge inspection is the most common type of inspection and is often performed from deck, ground or water levels or from permanent access structures, if available. Visual inspection is the predominant approach used in a routine inspection. With visual inspection, only basic tools for cleaning, probing, sounding, measuring, and visual aids are used. However, according to research, there can be significant variation in the condition ratings assigned to a structure simply based on visual inspection [1]. The use of unmanned aerial vehicles (UAVs) has recently been explored for the use of bridge inspections [2]. UAVs equipped with high resolution or infrared cameras can be used to scan a bridge taking hundreds of images and essentially building a navigable 3D model of the bridge. Recent advances in machine learning may be employed to automatically identify different types of bridge damage [3]. This research project will evaluate the effectiveness of using more autonomous methods for the collection and analysis of bridge deck images for the purpose of identifying the type and extent of damage in concrete decks.

17 Key Words UAV, bridge inspection, deep learning, concrete damage, 3D modeling		**18 Distribution Statement** No restrictions. This document is available to the public through	
19 Security Classification (of this report) Unclassified	**20 Security Classification (of this page)** Unclassified	**21 No. of pages**	**22 Price**

Form DOT F 1700.7 (8-72)

Contents

List of Figures

List of Tables

No Tables included in this report.

List of Key Terms

- ***UAV (Unmanned Aerial Vehicle)****: A drone or aircraft operated without a human pilot onboard, used for aerial data collection in bridge inspections.*
- ***RGB Camera****: A camera that captures images in red, green, and blue channels, used for high-resolution visual data acquisition.*
- ***YOLOv8 (You Only Look Once, version 8)****: A real-time object detection algorithm based on convolutional neural networks, used to identify damage types in bridge imagery.*
- ***CNN (Convolutional Neural Network)****: A class of deep learning models particularly effective for image recognition and classification tasks.*
- ***Image Annotation****: The process of labeling images with metadata (e.g., bounding boxes and damage types) to train machine learning models.*
- ***Efflorescence****: A type of concrete damage characterized by white, powdery deposits on the surface due to water migration and salt crystallization.*
- ***Spalling****: The breaking, chipping, or flaking of concrete surfaces, often due to freeze-thaw cycles or corrosion of embedded steel.*
- ***Delamination****: Separation of concrete layers, typically caused by corrosion or poor bonding, which may not be visible on the surface.*
- ***3D Modeling****: The process of creating three-dimensional representations of bridge structures using photogrammetry software.*
- ***Photogrammetry****: A technique for generating 3D models from overlapping 2D images, used to visualize bridge geometry and damage.*
- ***Image Stitching****: Combining multiple overlapping images into a single, continuous mosaic to visualize large concrete surfaces.*
- ***SLAM (Simultaneous Localization and Mapping)****: A method used by autonomous systems to map an environment while tracking their own location within it.*
- ***GPS-Denied Navigation****: Operating UAVs in environments where GPS signals are unreliable or unavailable, often using alternative sensors like cameras and sonar.*

- ***MATLAB Image Labeler***: *A tool used for manually annotating images to create labeled datasets for machine learning applications.*
- ***Agisoft Photoscan / Metashape***: *Commercial photogrammetry software used to generate 3D models from drone imagery.*
- ***Regard3D***: *An open-source photogrammetry software alternative to Agisoft, used for creating 3D models from images.*

Abstract

Bridge inspection is a vital component of infrastructure management, yet traditional visual methods are limited by accessibility, subjectivity, and cost. This project explored the integration of unmanned aerial vehicles (UAVs) and deep learning to automate the detection of structural damage in concrete bridge decks. A commercial drone was configured for underbridge flights and used to collect high-resolution imagery across multiple campaigns. Damage classes—spalling, spalling with exposed rebar, delamination, efflorescence, and cracking—were selected in collaboration with Steere Engineering. A custom image annotation workflow was developed, initially as a web-based app, but ultimately implemented in MATLAB due to deployment constraints. The annotated images were used to train a YOLOv8-based convolutional neural network. Results showed high accuracy for spalling, medium accuracy for efflorescence, while cracking and delamination remained challenging due to limited data and visual ambiguity. The project also explored 3D modeling using Agisoft Photoscan and Regard3D. Based on feedback from stakeholder Steere Engineering, we ultimately shifted toward stitched image mosaics of concrete surfaces to better visualize damage. A final meeting with Steere Engineering confirmed the value and interest from stakeholders into the proposed research, in particular the trained model and the mosaiced imagery, laying the groundwork for future deployment and refinement.

Chapter 1: Introduction and Background

1.1 Project Motivation

Routine bridge inspections are essential for maintaining structural safety, yet they often rely on manual visual assessments that are time-consuming, costly, and prone to subjectivity [1]. Accessing hard-to-reach areas, such as the underside of bridge decks, poses additional challenges. Recent advancements in UAV technology [2] and deep learning [3] offer promising alternatives for automating and enhancing the accuracy of bridge inspections [4-5].

1.2 Research, Objectives, and Tasks

The primary objective of this project was to develop a UAV-based system for collecting bridge imagery and apply deep learning techniques to automatically identify structural damage. Specific goals included:

- Reviewing neural network architectures and selecting YOLOv8 [12] for damage detection
- Configuring a commercial UAV for underbridge image collection
- Developing an image annotation workflow
- Training a YOLOv8-based neural network for multi-class damage detection
- Evaluating model performance across damage types
- Exploring 3D modeling and stitched image mosaics
- Engaging stakeholders for feedback and validation

1.3 Report Overview

This report presents the methodology used to collect and process bridge imagery, the development and training of the damage detection model, and the results of its evaluation. It concludes with recommendations for future work and potential applications in transportation infrastructure monitoring.

Chapter 2: Methodology

2.1 Materials

UAV Platform: A commercial drone equipped for underbridge flights and high-resolution image capture.
Cameras: RGB cameras mounted for upward-facing imaging.
Software Tools:

- MATLAB ImageLabeler for manual annotation
- Python for data formatting and model training
- YOLOv8 framework for object detection
- Agisoft Photoscan [15] and Regard3D [16] for 3D modeling
- Custom web-based annotation app (not deployed due to server restrictions)

2.2 Test Setup & Process

Drone Configuration: Completed by May 2024, enabling stable flight and image capture under bridges. The selected drone was a DJI Matrice 400 with DJI Guidance system [13]. The Guidance System was necessary to replace unreliable GPS signals using cameras and sonar. A MapIR Survey3 camera was installed on the top of the drone to collect the needed images. A picture of the drone setup is shown in Figure 1.

Figure 1: Drone setup with the guidance system and the camera.

Bridge Selection: The bridges selected for the study are Bridges Nos. 072551 and 072571, located on I-295 over West Natick Road and the Washington Secondary Trail. Both bridges are easily accessible and have four spans, of which only two above traffic. Therefore, imagery of the concrete ceiling of six spans were collected and used in the study. Two images of the selected bridge are provided in Figure 2.

Figure 2: Two images of the bridges selected for this study.

Image Collection: Image collection was replicated at different dates, times of day, and flight angles to capture different lighting conditions and apparent defect size. The initial collection was performed in June 2024, with following data collections in March 2025, and April 2025. The images were collected following a lawn-mower path under the bridge in manual flight at low speed to ensure overlap between consecutive images. Some data collection was executed emulating the actual drone flight with the camera only. During data collection, the camera collected one image every second, resulting in a large number of collected images.

Annotation App Development and Workflow: An annotation program was initially developed as a web-based app and initial in-house testing provided good results. Steere Engineering attempted to use the app but access issues due to internal departmental server restrictions prevented collaboration. Eventually, MATLAB ImageLabeler [14] feature was used to manually label defects. A Python program was developed to convert label files generated from MATLAB to a format usable within the PyTorch [17] framework.

Damage Classes: Five concrete damage classes were selected in consultation with Steere Engineering - spalling, spalling with exposed rebar, delamination, efflorescence, and cracking [6]. Limited data was available for delamination and cracking due to absence of such damage from the bridge selected for this study.

Model Training: A CNN based network [8-9], YOLOv8 was trained to detect the five classes of bridge deck damage using a dataset of 185 annotated training images and 65 validation images. The annotations were created manually using MATLAB's ImageLabeler and converted to a YOLO-compatible format in Python. Preprocessing included resizing the original high-resolution images (3000 × 4000 pixels) to 640 × 640 pixels and normalizing pixel values to improve training convergence. The model architecture featured a CSPDarknet backbone for feature extraction, a PANet neck for multi-scale feature fusion, and an anchor-free detection head. Training was performed over 100 epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 16. The loss function combined box loss, classification loss, and distribution-focused localization (DFL) loss. The training process showed consistent improvement across metrics, with convergence observed after approximately 70 epochs. Detailed evaluation results, including precision-recall curves, F1-confidence analysis, and confusion matrices, are presented in Chapter 3.

3D Modeling: Initial models were created using Agisoft Photoscan with legacy images (Jan–May 2024). Later, Regard3D was used with newly collected images (May 2025). While point clouds were promising, full 3D models were less effective. This is attributed to scarce presence of features in non-damaged concrete and metal sections.

Image Stitching: Based on stakeholder feedback, the focus shifted from 3D modeling to stitching concrete sections for damage visualization. A custom stitching program was developed in Python. The program used feature extraction and matching among consecutive images collected in flight to compute the relative translation and align the images. The identified damage was subsequently overlaid onto the stitched images to provide spatial context and enhance the visualization of defect locations.

Chapter 3: Results and Discussion

In this chapter we discuss the results of the automatic damage detection module (3.1), the 3D modeling via photogrammetry (3.2), and the stakeholder engagement (3.3).

3.1 Model Performance

Training Metrics: The YOLOv8 model showed consistent improvement during model training across 100 epochs, with decreasing loss values and stable precision and recall.
Detection Accuracy:

- **Spalling with Rebar**: High accuracy with optimal performance at 0.85 confidence threshold
- **Spalling**: High accuracy with good performance at 0.71 confidence threshold
- **Efflorescence**: Lower accuracy due to class imbalance and visual ambiguity; often limited resolution and training data prevented identification and the Efflorescence damage was confused with background. Larger efflorescence damages were mostly correctly identified.
- **Delamination Cracking:** Moderate performance; limited training data affected results

The **Confusion Matrix** presented in Figure 2 highlighted strong performance for spalling with and without exposed rebar, and background confusion for efflorescence. It was not possible to train and evaluate delamination and cracking classes due to insufficient instances of such defects on the test case bridges.

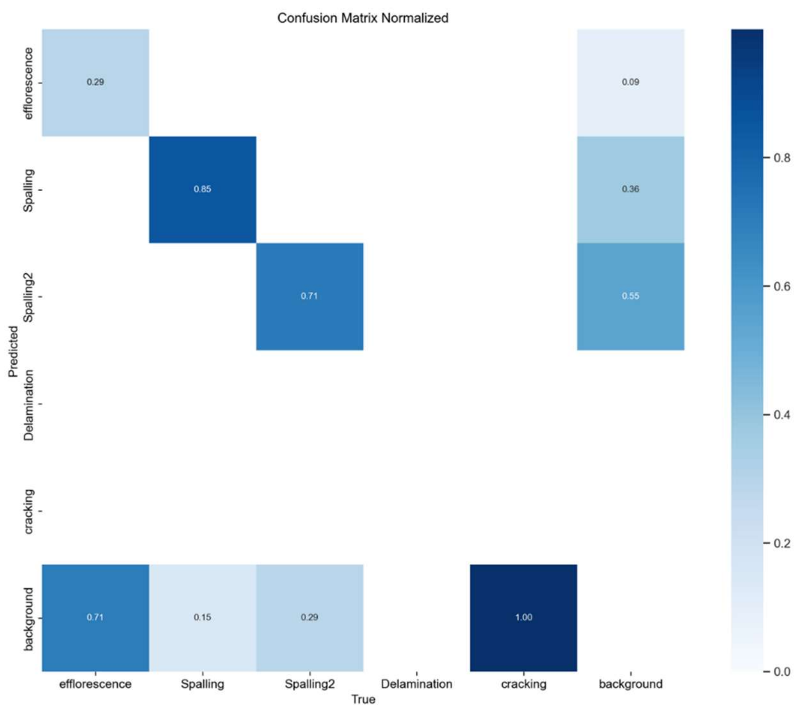


Figure 3: Confusion matrix of the YOLO-v8 trained network

Visual Validation: Bounding boxes aligned well with ground truth for most defects, though efflorescence was occasionally missed. Some examples of images with identified and missed damage are provided in Figure 3.

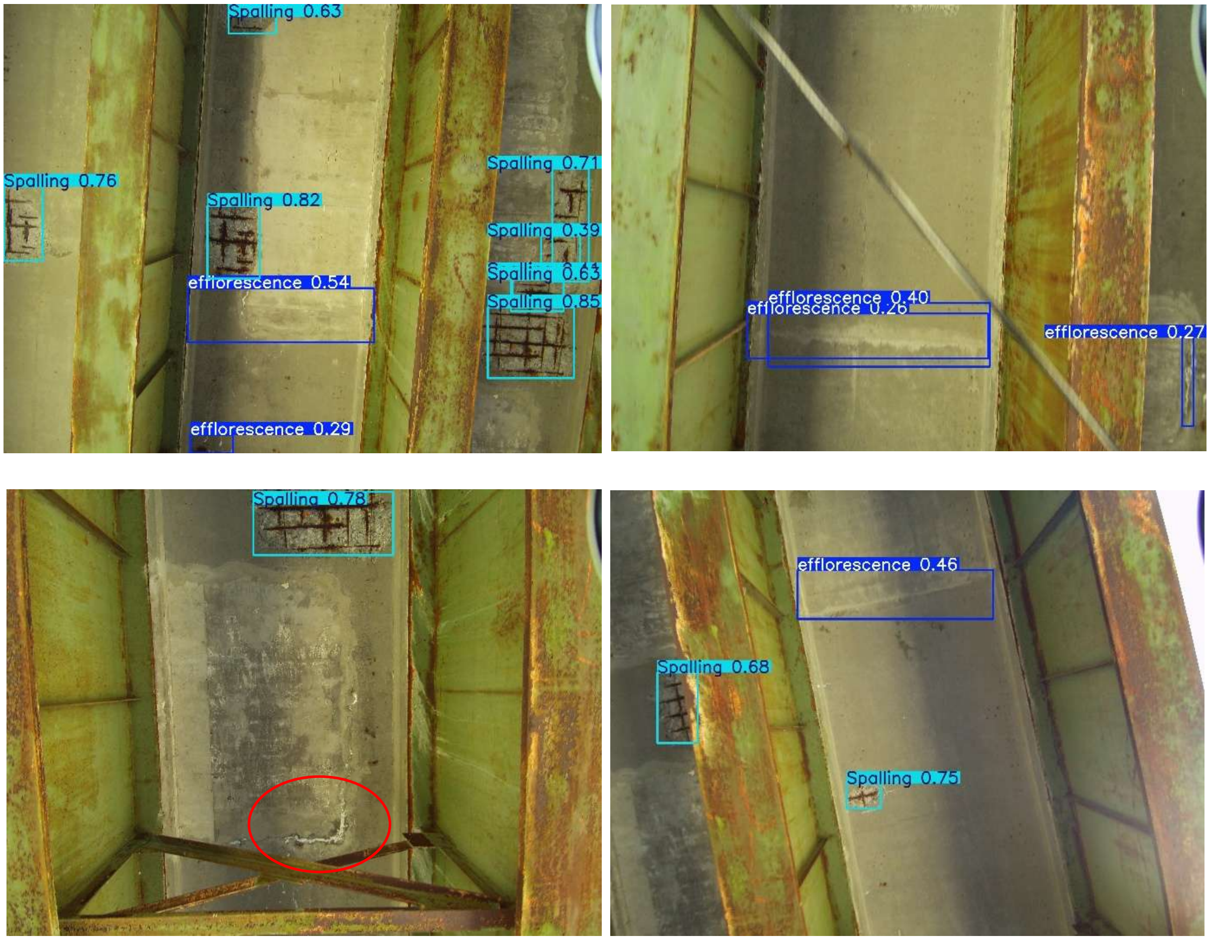
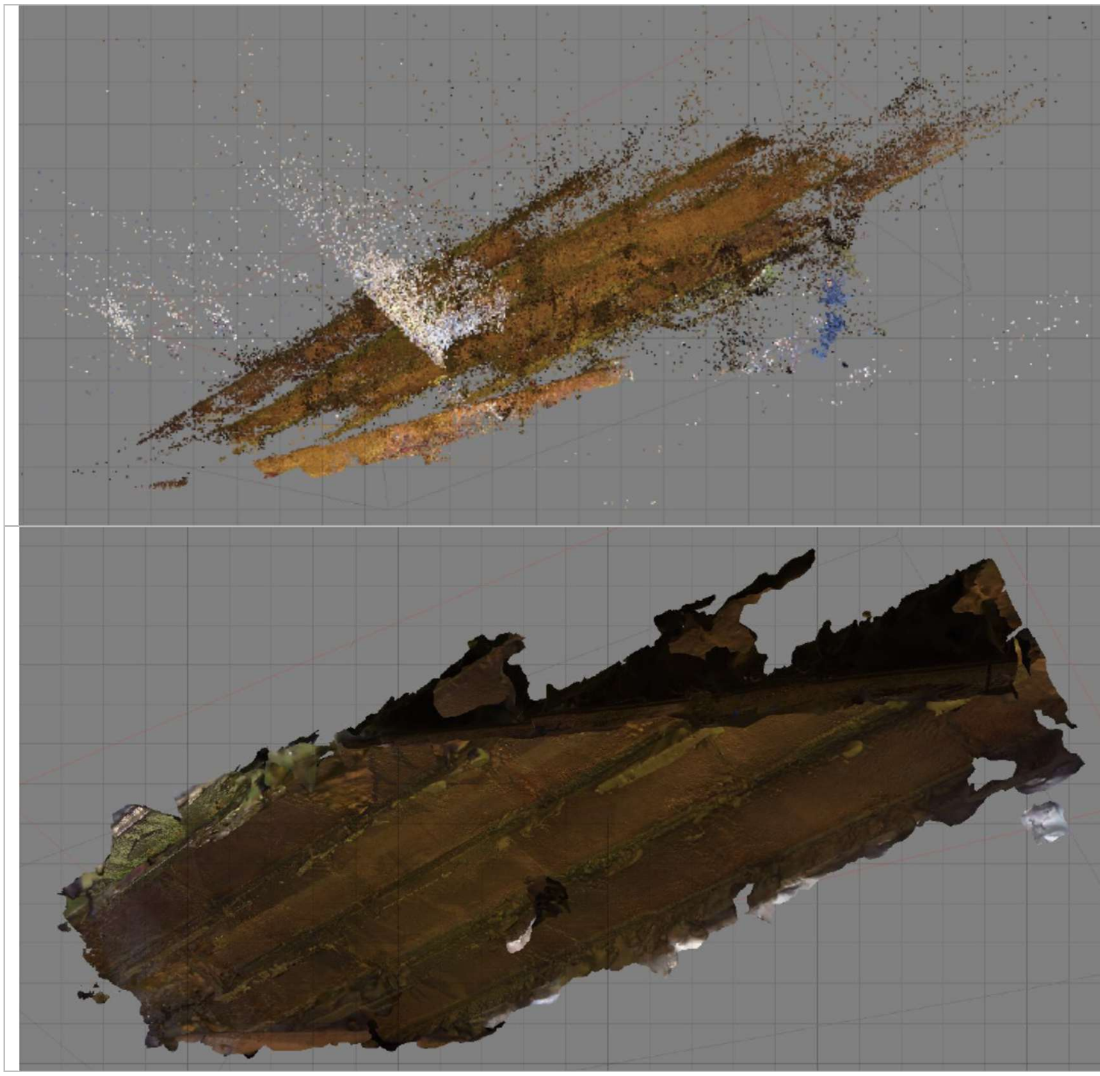


Figure 4: Examples of validation Batch labels identified by the YOLO-v8 trained network. Circled in red a missed Efflorescence defect.

3.2 3D Modeling

Agisoft Photoscan: Two undergraduate students explored photogrammetry using Agisoft PhotoScan Professional to generate 3D models of bridge structures from legacy drone imagery. Despite limitations in image quality and coverage, they successfully aligned and processed over 900 images, producing point clouds and mesh models that demonstrated the potential of photogrammetry for bridge inspection. Their work was presented as a poster at the 2024 URI E-Week. The pointcloud presented in Figure 4a showed a large number of outlier points and noise. A 3D mesh (Figure 4b) was subsequently created from the pointcloud. Upon visual inspection, it was determined that the quality of the 3D model was not sufficient to provide a good visualization of the damage. Metal sections were not modeled properly due to lack of features, bumps were identifiable in flat concrete surfaces, and the texture did not contain sufficient details to show bridge damage.

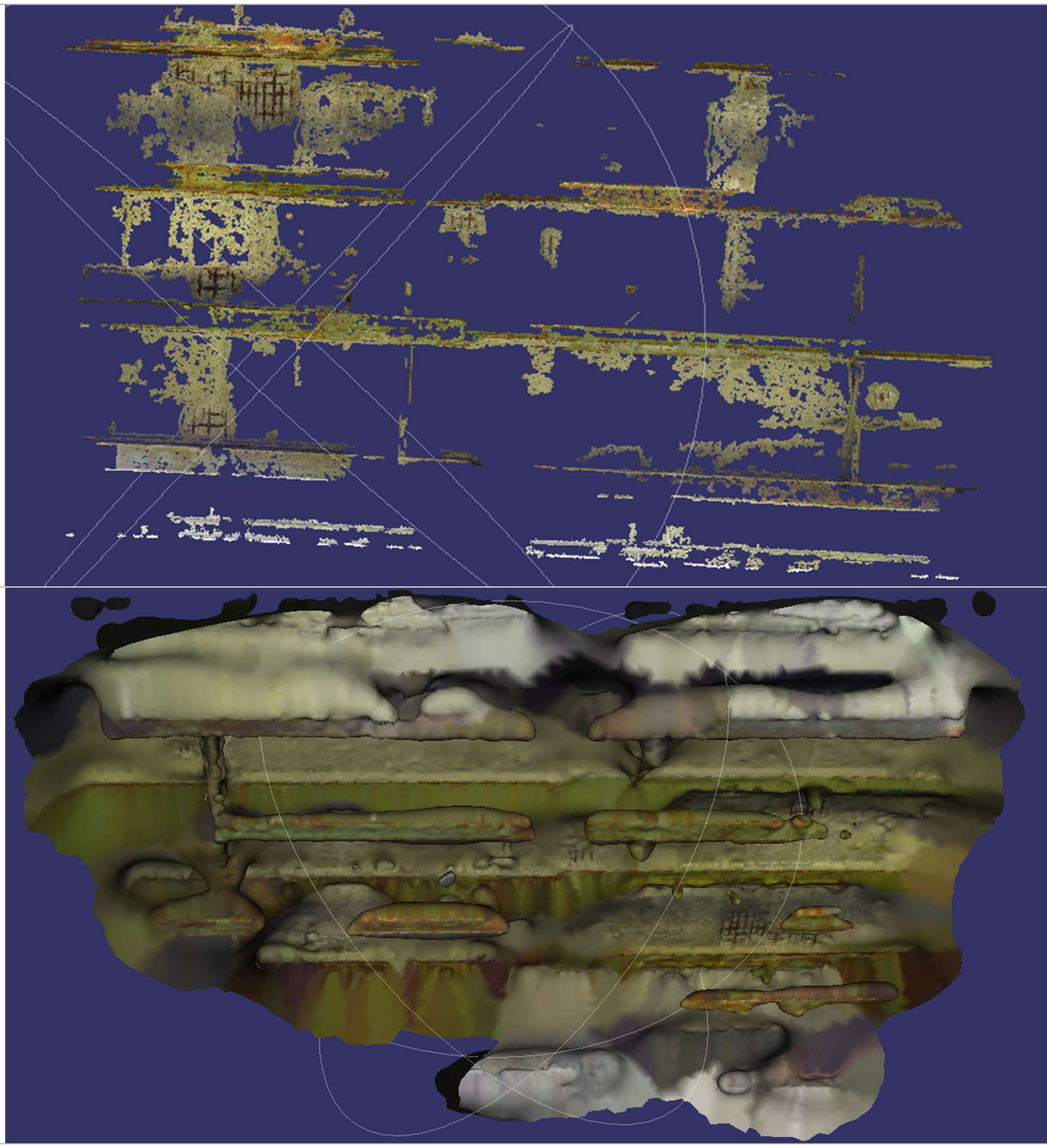


(a)

(b)

Figure 5:(a) Pointcloud and (b) 3D mesh generated with Agisoft PhotosScan.

Regard3D, a free alternative to costly Agisoft PhotoScan was considered and evaluated with imagery collected at the bridge selected in this study. Examples of the generated 3D model are provided in Figure 6. In general, Regard3D produced a clean pointcloud (Figure 6a) without outliers. Within the pointcloud, it was also possible to identify damage as spalling with exposed rebar and efflorescence. Unfortunately, the quality of the produced 3D mesh (Figure 6b) was insufficient to the scope of this project. In particular, flat concrete surfaces did not present many features and therefore were distorted in the mesh generation.



(a)

(b)

Figure 6: (a) Pointcloud and (b) 3D mesh generated with Regard3d

Image Stitching: In order to provide a contextualization of the damage within the bridge structure an image stitching of the collected images was developed and applied. The ideation of the stitching algorithm was also developed in response to direct requests from industry partner Steere Engineering to lay the foundation to further studies on the numerical quantification of the area affected by different damages. Among the collected data, five sequences of images were identified suitable for stitching and experimented on. The results of the stitching of two image sequences are presented in Figure 7. The first image is the product of the stitching of 10 images collected along a longitudinal flight along a single girder bay. The image can be used to map an entire girder bay. The success of the stitching process demonstrates that a concrete deck contains enough features for this task. On the other hand, the cross frames clearly constitute an impediment to proper data visualization and may hide some damage. The second image is the result of the stitching of 8 images on a flight across adjacent girder bays. This sequence was tested to see if girders were an obstruction of proper mapping of the concrete surface. The resulting image show that while it is possible to stitch cross girder imagery, the representation of the metal girders is affected by artifacts due to the different depth.

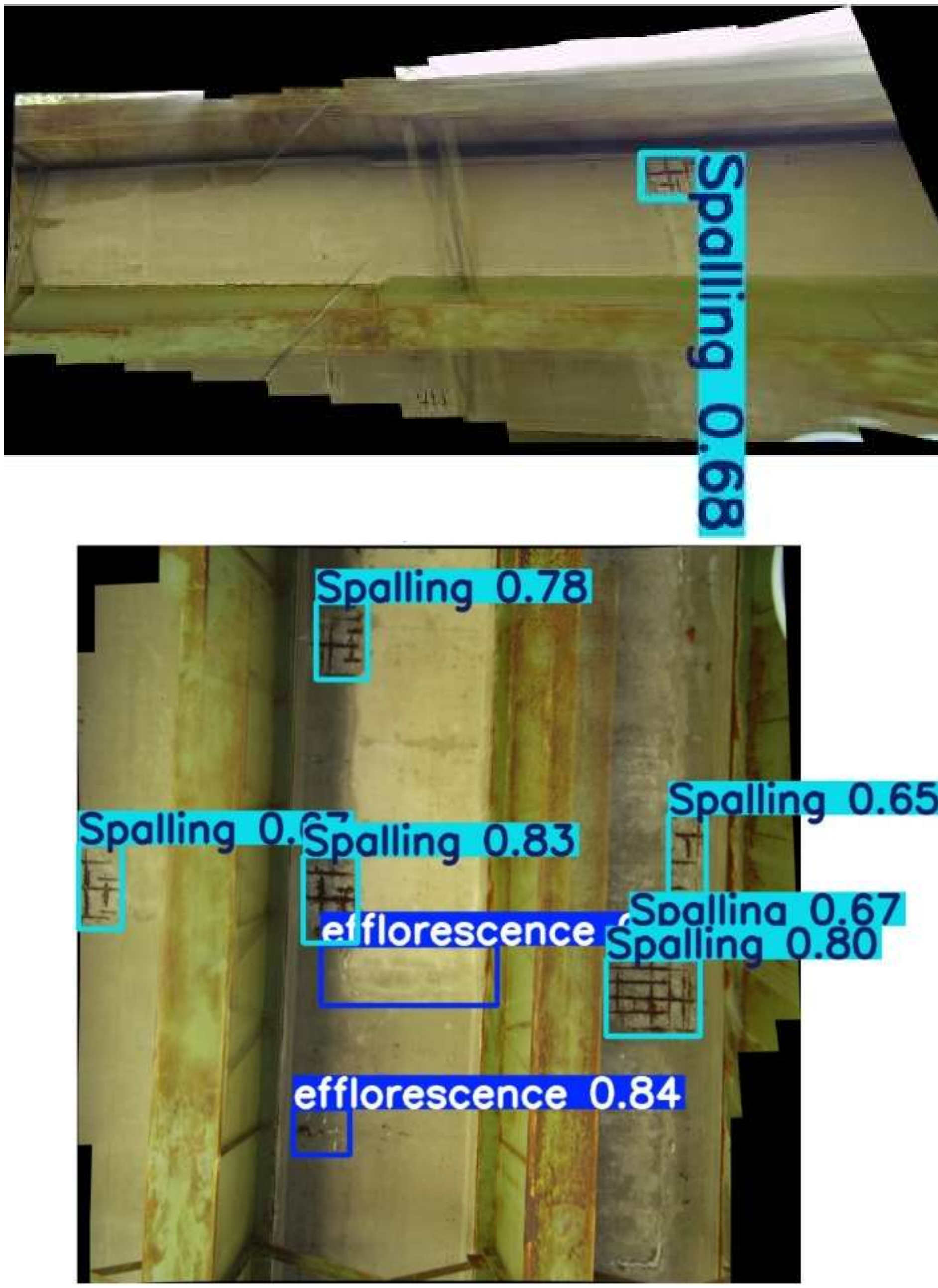


Figure 7: Two examples of stitched images.

3.3 Stakeholder Engagement

Steere Engineering provided valuable input throughout the project. Their involvement included consultation on the selection of relevant concrete damage classes and participation in a one-day workshop and final presentation, where project outcomes were discussed and feedback was collected. These collaborative activities supported the practical relevance of the research and aligned with the proposed stakeholder engagement goals.

Chapter 4: Additional Contributions

4.1 Educational Contributions

During the reporting period, three transportation-related graduate courses were offered by faculty affiliated with the UTC:

- **Management of Highway Bridges** (Spring 2024) - taught by Merve Iplikcioglu Kirtan (Steere Engineering)
- **Special Problems** (Fall 2024 and Spring 2025) - taught by Mayrai Gindy (Co-PI)

These courses supported graduate-level education in transportation infrastructure and provided opportunities for students to engage with emerging technologies and research themes related to bridge inspection and structural health monitoring. In particular, two Civil Engineering graduate students enrolled in **Special Problems** offered by co-PI Gindy actively contributed to the project development by interfacing with the Electrical Engineering Team to develop the tools for marking the drone images, providing expertise and time marking collected images, and participating in image collection in Spring 2025.

4.2 Exploratory Drone Development

As part of a parallel student-led initiative, a custom drone platform was developed to explore the feasibility of fully automated bridge inspections. The system integrates a PX4 flight controller, an onboard NVIDIA Jetson module, and a camera for real-time image acquisition and processing. While this effort remained exploratory and was not fully integrated into the main project workflow, it provided valuable insights into hardware integration and onboard SLAM capabilities, laying the groundwork for future autonomous inspection systems.

4.3 Dissemination

Project outcomes were disseminated through multiple presentations and outreach activities across academic, professional, and stakeholder venues:

- **October 2023** – Presentation at the *2023 Rhode Island Transportation Forum*: "Vision-Based Detection of Bridge Damage Captured by Unmanned Aerial Vehicles" (Paolo Stegagno)
- **May 2024** – Poster presentation at *URI College of Engineering E-Week*: "Using Photogrammetry and Drones to Inspect Bridges" by undergraduate students Arpi Donoyan and Ethan White
- **October 2024** – Presentation at the *2024 Rhode Island Transportation Forum*: "Drone Bridge Inspection – Current Efforts and Future Perspectives" (Paolo Stegagno)
- **June 2025** – One-day workshop and final project presentation hosted at Steere Engineering, including stakeholder feedback and discussion of results
- **August 2025** – Poster presentation at the 2025 TIDC conference, "Object-Guided Sequential Image Stitching Using YOLO and ORB Matching" (Xiangjun Meng)
- **October 2025** – Scheduled presentation at the *2025 Rhode Island Transportation Forum*: "Vision-Based Detection of Bridge Damage Captured by Unmanned Aerial Vehicles" (Paolo Stegagno)

These activities supported the dissemination of research findings, fostered stakeholder engagement, and contributed to the broader conversation on innovative approaches to bridge inspection and infrastructure monitoring.

Chapter 5: Conclusions and Recommendations

This project successfully demonstrated the integration of unmanned aerial vehicles (UAVs) and deep learning for the detection of bridge deck damage. The trained YOLOv8 model achieved high accuracy for identifying spalling and spalling with exposed rebar, validating the approach for practical applications. However, challenges remain in detecting more subtle defects such as efflorescence, cracking, and delamination, which require improved dataset balance, enhanced preprocessing, and more representative training samples.

The initial exploration of 3D modeling using photogrammetry software revealed limitations due to image quality and feature sparsity in non-damaged concrete surfaces. Based on stakeholder feedback, the project pivoted toward stitched image mosaics, which proved more effective for visualizing damage across large concrete areas. These mosaics were well received by Steere Engineering and offer a practical tool for inspection workflows.

Stakeholder engagement was a key component of the project, with Steere Engineering providing technical input on damage classification, participating in a one-day workshop, and offering feedback on the developed tools. This collaboration helped ensure the relevance of the research to real-world inspection practices.

The project continues to evolve. Two undergraduate students are currently working to improve the image stitching pipeline by developing a semantic segmentation network to isolate concrete surfaces and exclude structural elements such as girders and cross frames. This will enable more accurate mapping of damaged areas and support automated quantification of defect size and affected surface area.

Future work will focus on expanding the image database to include multiple bridges, increasing variability in damage types and environmental conditions. Additionally, improvements to the drone setup are underway to enable reliable data collection over water. The current visual localization system is challenged by reflective non-solid surfaces, and a new SLAM-based approach is being developed to use the bridge structure itself for localization. This advancement will allow for broader deployment across diverse bridge environments and support rapid scaling of the dataset.

In summary, the project has laid a strong foundation for UAV-based bridge inspection using computer vision and machine learning. Continued development, stakeholder collaboration, and technical refinement will be essential to transition this research into a deployable and scalable inspection solution.

References

[1] Federal Highway Administration. *Reliability of Visual Inspection for Highway Bridges*. FHWA-RD-01-020, 2001.

[2] Dorafshan, S., & Maguire, M. (2018). *Bridge inspection: human performance, unmanned aerial systems and automation*. Journal of Civil Structural Health Monitoring, 8(3), 443–476.

[3] LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. Nature, 521, 436–444.

[4] Rhode Island Department of Transportation. *Unmanned Aerial Vehicles in Transportation Applications*. RIDOT-RTD-16-1A, 2016.

[5] Rhode Island Department of Transportation. *Autonomous Airborne Thermal Imaging of Bridge Deck Structures*. RIDOT-RTD-17-1A, 2017.

[6] American Association of State Highway and Transportation Officials (AASHTO). *Manual for Bridge Element Inspection*, 2019.

[7] Jeong, E., Seo, J., & Waker, J. (2020). *Literature Review and Technical Survey on Bridge Inspection Using Unmanned Aerial Vehicles*. Journal of Performance of Constructed Facilities, 34(6).

[8] Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). *Deep learning-based crack damage detection using convolutional neural networks*. Computer-Aided Civil Infrastructure Engineering, 32(5), 361–378.

[9] Chen, F., & Jahanshahi, M. R. (2018). *NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion*. IEEE Transactions on Industrial Electronics, 65(5), 4392–4400.

[10] Kim, H., Ahn, E., Shin, M., & Sim, S.-H. (2018). *Crack and noncrack classification from concrete surface images using machine learning*. Structural Health Monitoring, 18(3), 725–738.

[11] Jahanshahi, M. R., Masri, S. F., Padgett, C. W., & Sukhatme, G. S. (2013). *An innovative methodology for detection and quantification of cracks through incorporation of depth perception*. Machine Vision and Applications, 24(2), 227–241.

[12] R. Varghese and S. M., "YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness," *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, Chennai, India, 2024, pp. 1-6.

[13] DJI. *Guidance User Manual V1.6*. Shenzhen DJI Technology Co., Ltd., 2015. https://www.dji.com/global/guidance [dji.com]

[14] The MathWorks, Inc. (2024). *Image Labeler App – Label images for computer vision applications*. Natick, Massachusetts: The MathWorks, Inc. https://www.mathworks.com/help/vision/ref/imagelabeler-app.html

[15] Agisoft LLC. (2024). *Agisoft Metashape User Manual*. https://www.agisoft.com

[16] Regard3D. (2024). *Regard3D: Free and Open Source Structure-from-Motion Program*. https://www.regard3d.org

[17] Paszke, A. et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. 10.48550/arXiv.1912.01703.

Appendix A: YOLOv8 Training and Evaluation Report

The following technical report was authored by a graduate student involved in the project. It documents the training setup, model architecture, evaluation metrics, and performance analysis of the YOLOv8-based damage detection system developed for this research.

Title: Performance Analysis of YOLOv8 for Multi-Class Damage Detection

Objective: The goal of this project was to develop and evaluate a CNN-based architecture with enhancements for efficiency and accuracy for identifying structure damage types such as efflorescence, spalling, delamination, and cracking.

1. Introduction

Structural integrity assessments play a crucial role in ensuring the safety and longevity of crucial infrastructure. This project aims to develop an automated defect detection system using You Only Look Once version 8 (YOLOv8) to identify common structural defects, such as efflorescence, spalling, and cracking, from vision images of an old bridge concrete structures. The system leverages state-of-the-art object detection techniques to provide accurate and efficient assessments, suitable for real-world applications.

2. Dataset Description

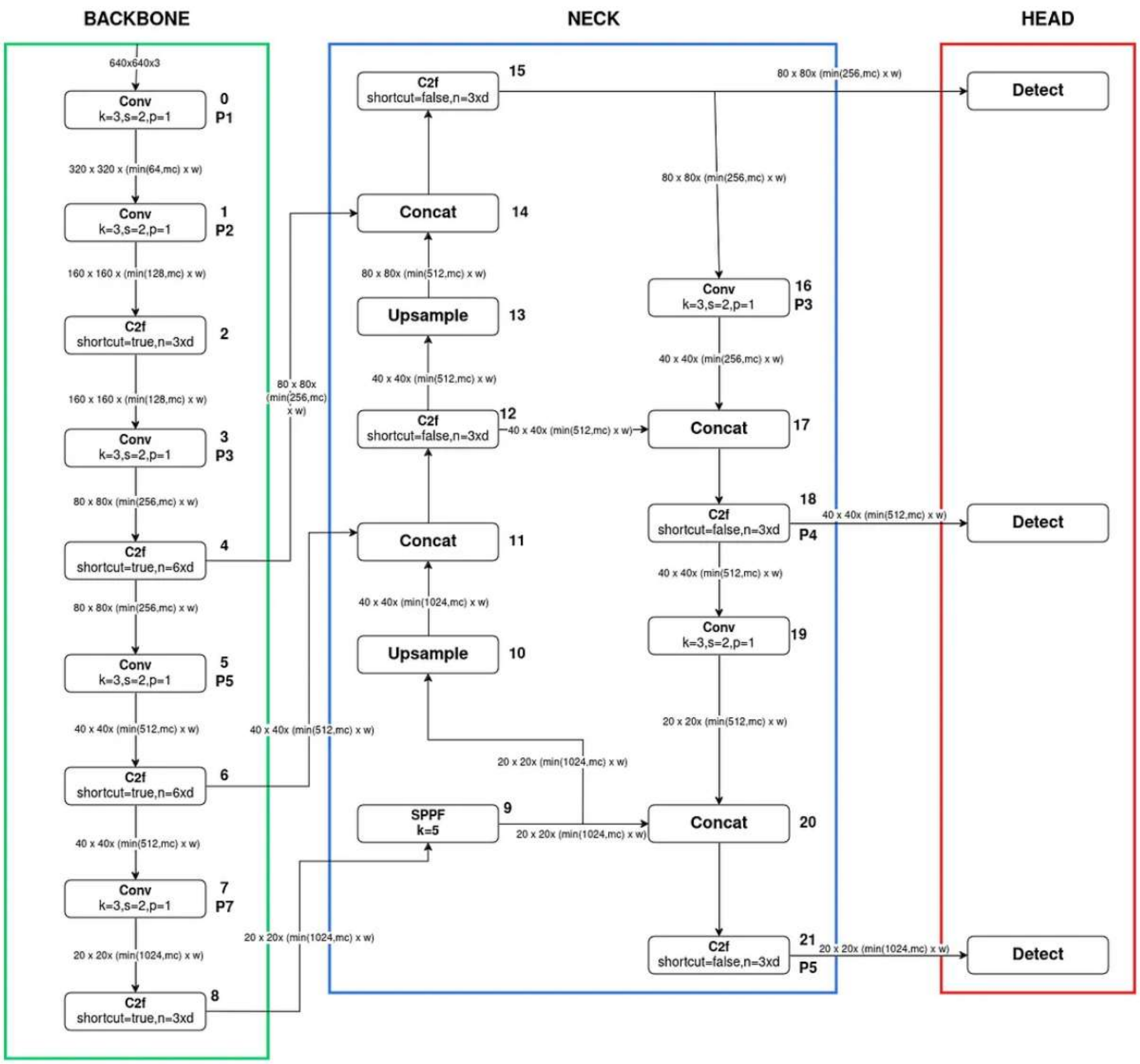
The dataset comprises high-resolution images of concrete structures captured under diverse conditions, offering a variety of real-world scenarios for structural defect detection. It includes 185 training images and 65 validation images across five classes: Efflorescence, Spalling, Spalling2, Delamination, and Cracking. The dataset presents several challenges, such as variability in lighting, shadowed areas, overexposed regions, class imbalance with more samples of spalling compared to efflorescence, and the presence of occlusions (e.g., human interference) and environmental noise (e.g., glare), which increase detection difficulty. Bounding box annotations were created using MATLAB's ImageLabeler tool and converted to a YOLO-compatible data format in Python for seamless training.

2.1 Preprocessing Steps:

Bounding boxes for defects were manually annotated using MATLAB ImageLabeler to ensure precise localization of each defect within the images. The annotated images and labels were then split into training (80%) and validation (20%) datasets using Python. To optimize the data for model training, the original high-resolution images (3000 × 4000 pixels) were resized to 640 × 640 pixels, and pixel values were normalized to improve convergence during the training process.

3. Model Architecture

The model is built on YOLOv8, a CNN-based object detection framework optimized for real-time performance. It features a CSPDarknet backbone for efficient feature extraction, utilizing cross-stage partial connections to enhance gradient flow and reduce redundancy. The neck employs a Path Aggregation Network (PANet) to merge features at multiple scales, improving small-object detection, while the head outputs bounding boxes, class probabilities, and confidence scores through an anchor-free detection mechanism. The model processes 640 × 640 input images and generates outputs including bounding boxes, class probabilities, and object confidence for detected defects. The YOLOv8 family offers lightweight and larger variants, providing flexibility for resource-constrained scenarios.



4. Training and Evaluation

4.1 Training Setup:

The model was trained using the Adam optimizer with an initial learning rate of 0.001 and a batch size of 16 for 100 epochs. The loss function consisted of three components: box loss for bounding box regression, classification loss for predicting correct classes, and DFL loss for distribution-focused localization. During training, the model demonstrated consistent improvement across metrics, with steadily decreasing loss values and validation metrics indicating convergence after 70 epochs.

4.2 F1-Confidence Curve:

The F1-confidence curve shows the model's performance across different confidence thresholds. For defects like spalling, the curve peaks at high confidence thresholds (~0.8), indicating precise and confident detections. However, for efflorescence, the curve has a lower peak and flattens at lower confidence thresholds, reflecting the challenge of distinguishing this class from the background.

F1-score is the harmonic mean of precision and recall:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

It balances **precision** (how many detected defects are true positives) and **recall** (how many true defects are detected) into a single metric.

The *Figure 1. F1-Confidencee Curve* shows different confidence thresholds, where: 1) low thresholds allow more predictions, increasing recall but potentially lowering precision. 2) High thresholds filter uncertain predictions, improving precision but reducing recall.

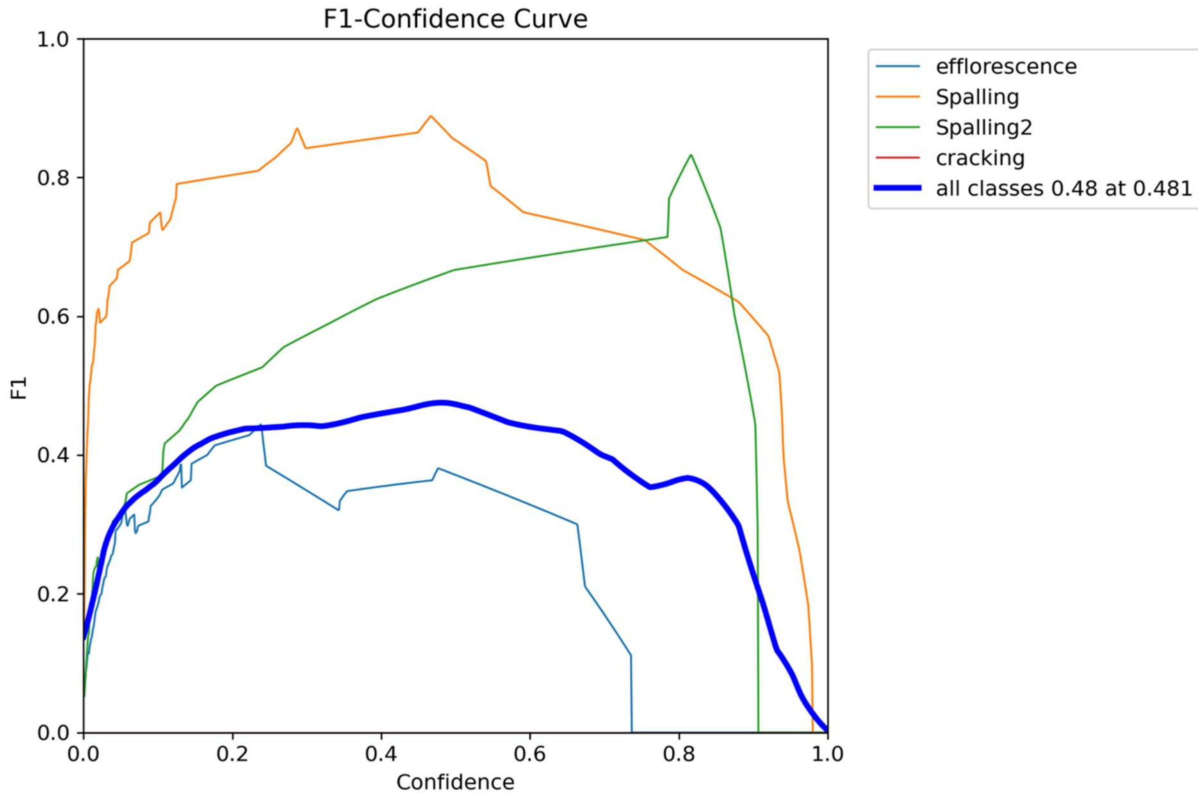


Figure 1. F1-Confidencee Curve

4.2 Key Observations from the F1 Curve:

For classes spalling, the curve peaks at a high confidence threshold (e.g., 0.8), indicating: 1) the model is confident and accurate in detecting spalling defects. 2) There's a good balance between precision and recall at higher thresholds.

For classes efflorescence, the curve is flatter and peaks at a lower threshold (e.g., 0.5), suggesting: 1) the model struggles to confidently distinguish efflorescence from other classes or the background. 2) Precision and recall do not simultaneously improve significantly at higher thresholds.

4.3 Class-Specific Behavior:

- Efflorescence:

  F1-scores remain relatively low across thresholds due to lower precision, caused by false positives with the background, and lower recall, resulting from missed true positives. The flatness of the curve indicates difficulty in identifying a threshold where the model achieves a strong balance between precision and recall.
- Spalling:

The high and narrow peak of the F1-score curve suggests that the model performs well within a specific confidence range, achieving both high precision and recall, indicating that it has successfully learned distinct and reliable features for spalling detection.

- Cracking:

  F1-scores are nearly perfect across thresholds, reflecting the model's ease in detecting cracks, likely due to their clear visual distinction.

4.4 Insights from the Curve

The optimal threshold for each class corresponds to the peak of the F1 curve, with spalling achieving a good balance between precision and recall at a threshold of 0.8, minimizing false positives and false negatives, while efflorescence performs better at a lower threshold of 0.5, maximizing recall at the expense of some precision. At low thresholds, the model predicts nearly everything as a defect, resulting in high recall but low precision and a drop in F1-scores. Conversely, at high thresholds, the model becomes overly conservative, achieving high precision but low recall, leading to many missed true defects.

4.5 Precision-Recall Curve

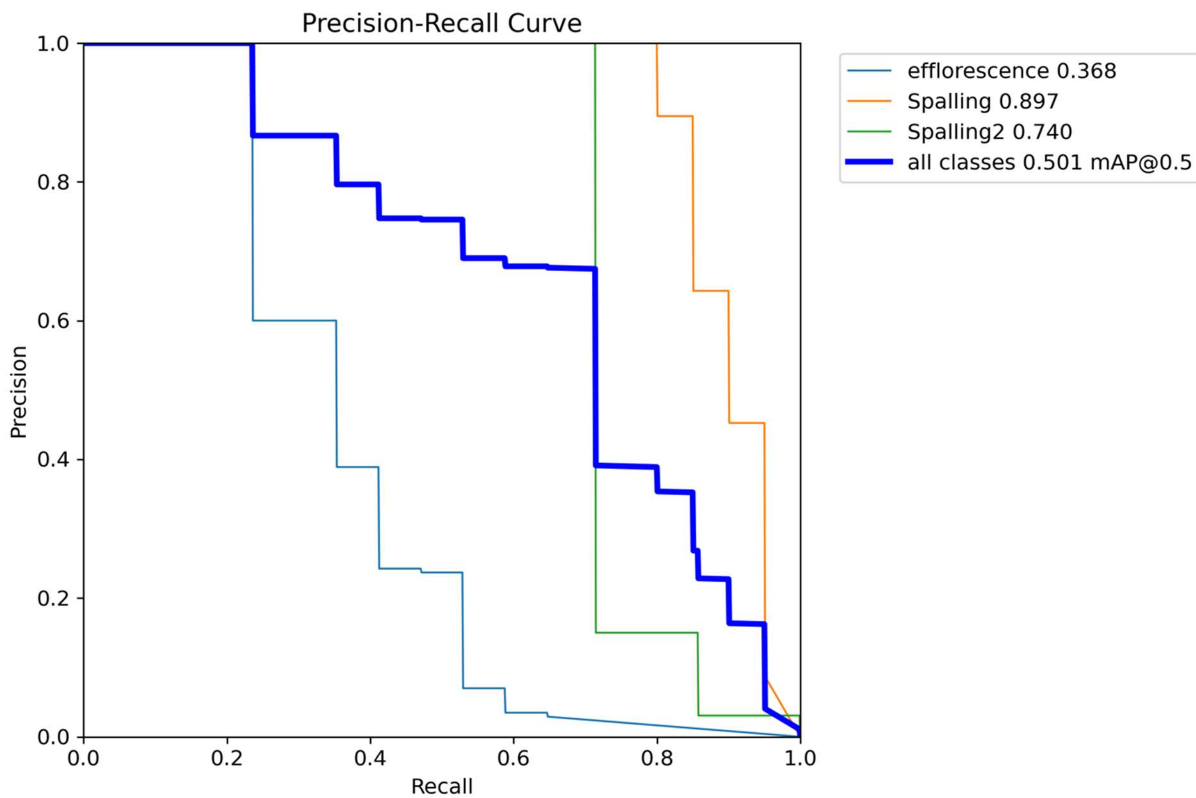


Figure 2. Precision-Recall Curve

The *Figure2. Precision-Recall Curve* highlights the trade-off between precision and recall. Classes like cracking achieve near-perfect scores, with precision and recall consistently high. In contrast, efflorescence shows a sharper drop-off, likely due to subtle visual features and variability in appearance. Future efforts should focus on improving recall for challenging classes like efflorescence.

4.6 Training Metrics & Validation Results:

The *Figure 3 training metrics* indicates the training and validation performance of a model over 100 epochs. The model improves steadily across all aspects. The box loss decreases consistently over epochs, reflecting enhanced accuracy in predicting bounding box coordinates, with reduced fluctuations indicating stable training. The classification loss drops rapidly during the initial epochs and slows as training progresses, showing the model's ability to quickly and effectively classify defects, with the curve smoothing toward convergence. Similarly, the distribution-focused localization (DFL) loss decreases steadily, highlighting the model's improved capability to refine localization predictions, with reduced fluctuations signaling robust training dynamics. Precision starts high and stabilizes with minor fluctuations, suggesting that the model maintains a good proportion of true positives among all predictions. Recall increases significantly during the early epochs before plateauing, indicating that the model improves at identifying true defects but may still miss some instances.

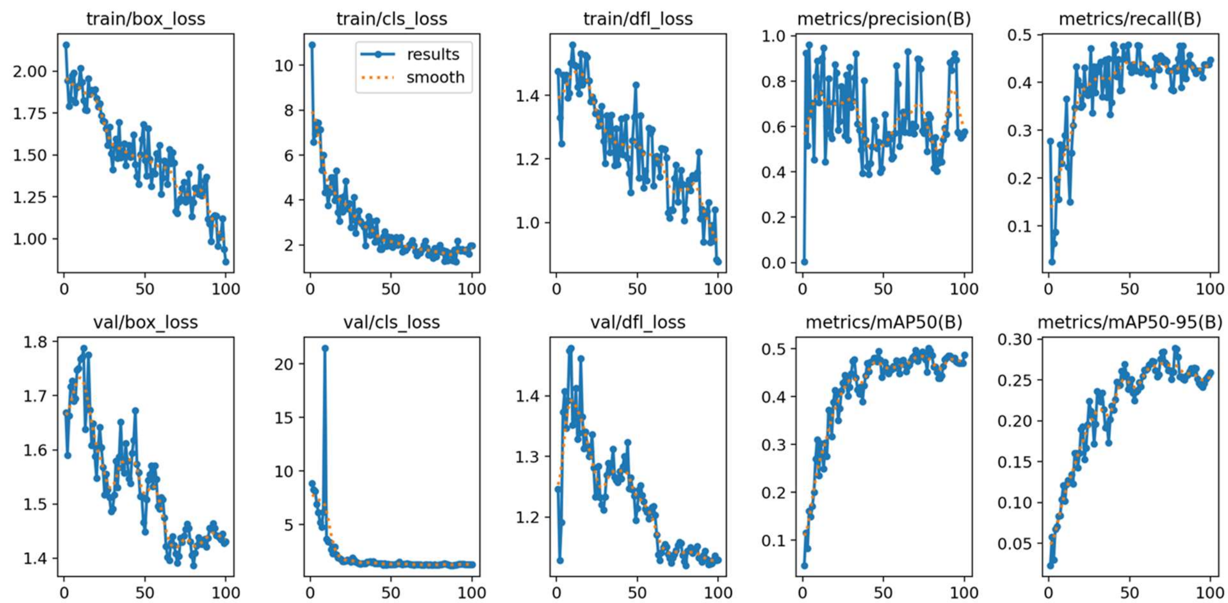


Figure 3 Training Metrics

Ground Truth vs. Predictions:

Visualizations of validation samples show accurate detections for most defects, with bounding boxes aligning well with ground truth. For example, in *Figure 4 Illustration of YOLO-v8 visualization samples: (a) validation_batch_labels vs. (b) validation_batch_pred,* spalling and cracking are detected with high confidence, while smaller or less distinct defects like efflorescence are occasionally missed.


2021_0101_015135_102.JPG
2021_0101_015050_078.JPG
efflorescence 0.7
Spalling 0.9
Spalling2 0.3
2021_0101_015349_191
2021_0101_014917_022.JPG
Spalling 0.9
Spalling2 0.8
efflorescence 0.4
2021_0101_014941_038.JPG
2021_0101_015108_087.JPG
2021_0101_014918_023.JPG
Spalling 0.9
Spalling2 0.4

(a)

(b)

Fig 4. Illustration of YOLO-v8 visualization samples: (a) validation_batch_labels; and (b) validation_batch_labels

4.7 Confusion Matrix

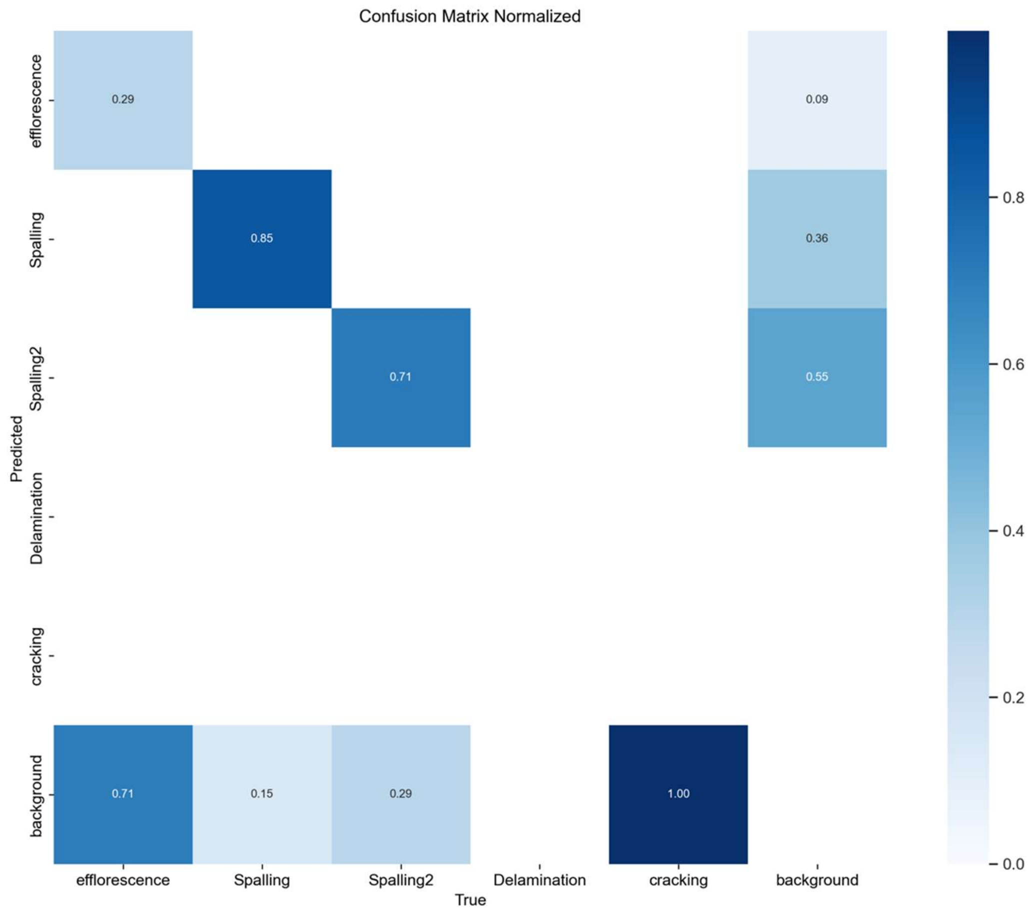


Figure 5. Normalized Confusion Matrix

The *Figure 5. Normalized Confusion Matrix* illustrates the model's performance in predicting different defect classes. High precision is evident for the "cracking" class, with a perfect score of 1.0, indicating the model accurately identifies cracks without confusion. The "spalling" and "spalling2" classes show relatively strong predictions with normalized values of 0.85 and 0.71, respectively, but there is some misclassification between these classes, likely due to their similar features. The "efflorescence" class has a lower accuracy of 0.29, with noticeable confusion with the background class (0.71), indicating challenges in distinguishing efflorescence from non-defective areas. The background class itself is predicted accurately in most cases, but its overlap with efflorescence suggests room for improvement. Overall, the matrix highlights the model's strengths in detecting distinct classes like cracking while indicating areas of improvement, particularly in differentiating between efflorescence and background.

5. Observations and Challenges:

Class Imbalance: The dataset contains significantly more examples of spalling compared to efflorescence, affecting model performance for the latter.

6. TimeLine

The timeline schedule of this project is shown in Table 1. At the end of each year, a comprehensive final report summarizing research findings and a specification for developing and implementing will be delivered. In this year, the general task can be categorized as follows: (1) **Task 1**: Collect Dataset. Collect high-resolution images of concrete structures under various conditions.; (2) **Task 2**: Annotation. Manually annotate bounding boxes for defects using

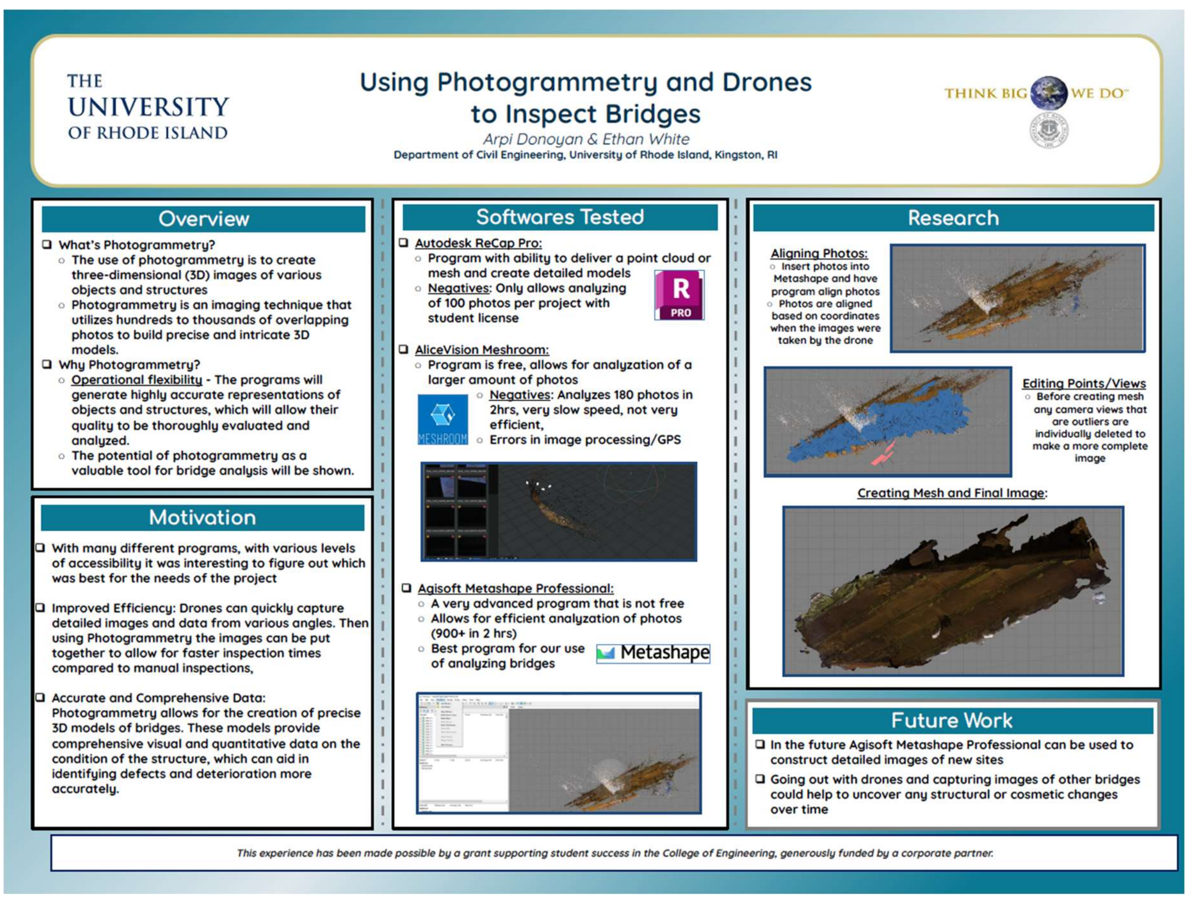
MATLAB ImageLabeler. (3) **Task 3**: Training & Validation. Train YOLOv8 on the dataset, analyze performance metrics, and validate results. (4) **Task 4**: Report Writing. Prepare a detailed report including results, analysis, and recommendations.

Table 1. Project schedule

Year	2024											
Months	1	2	3	4	5	6	7	8	9	10	11	12
Task 1												
Task 2												
Task 3												
Task 4												

Appendix B: Undergraduate Student Poster – URI E-Week 2024

The following poster was created by undergraduate students Arpi Donoyan and Ethan White and presented at the University of Rhode Island College of Engineering E-Week in May 2024. It summarizes their exploration of photogrammetry tools for drone-based bridge inspection, including software comparisons and workflow development.



Appendix C: Graduate Student Poster – TIDC Student Poster Contest 2025

The following poster and abstract were submitted by graduate student Xiangjun Meng and presented at the 2025 TIDC Student Poster Contest. The work introduces an object-guided image stitching framework for automated monitoring of spalling defects in bridge infrastructure using YOLO-based detection and ORB feature matching.

Poster Title: *Object-Guided Sequential Image Stitching Using YOLO and ORB Matching*

Author: Xiangjun Meng

Affiliation: Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island

Advisor: Paolo Stegagno

Event: TIDC Student Poster Contest, August 2025

Abstract:

This research presents an object-guided image stitching framework for the automated monitoring of spalling defects in bridge infrastructure. The proposed method addresses a key challenge in structural health monitoring [1]: capturing the spatial and temporal evolution of localized surface damage using sequential imagery. Traditional inspection methods often rely on manual documentation, which can be labor-intensive, inconsistent, and error-prone [2][3]. To overcome these limitations, we integrate object detection and feature-based matching to construct panoramic representations of structural surfaces over time. As illustrated in ***Figure 1***, a deep learning model (YOLO11), enhanced with modules such as SPPF, C2PSA, and C3k, is trained on a dataset of 498 annotated high-resolution images to detect spalling regions across sequential frames. Within each detected region, ORB features are extracted and matched between consecutive images to estimate object-specific motion flow [4]. By accumulating the relative displacement of matched keypoints, the method aligns and stitches multiple frames into a single panoramic image. The model demonstrates strong performance in detection precision (86-90%) and matching concrete defects across multiple images. Unlike traditional keypoint-based approaches [5], our object-aware pipeline maintains high spatial accuracy in defect localization, even in the presence of illumination changes and texture variation. The resulting panoramas preserve structural detail and provide an intuitive visual summary of damage evolution. This work contributes a novel integration of deep learning and classical computer vision to the field of bridge infrastructure monitoring. For the academic community, it opens new pathways for vision-based structural assessment. For public infrastructure stakeholders, it provides a reliable and efficient tool to support data-driven maintenance strategies and enhance safety in the built environment.

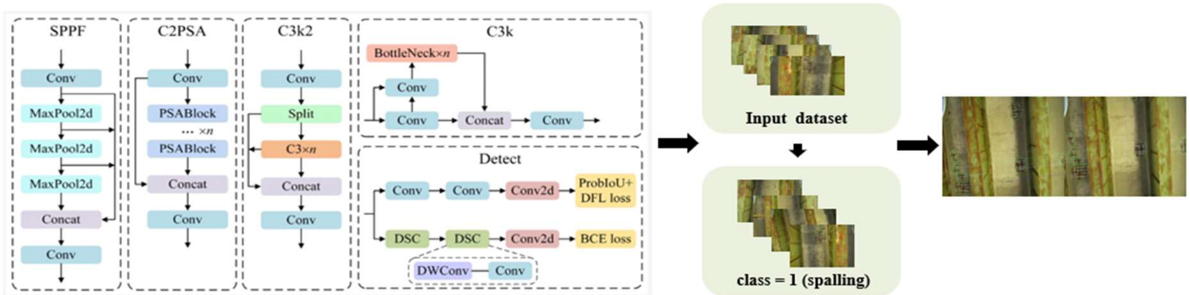


Figure1.*Overview of the proposed spalling detection and stitching framework. The input dataset consists of time-sequential surface images. A YOLO-based backbone enhanced with SPPF, C2PSA, and C3k modules is used for defect detection, targeting class = 1 (spalling). Feature extraction and matching are performed within detected regions across frames, followed by motion-based image stitching to generate a panoramic view of surface damage evolution.*

Acknowledgements: Funding for this research is provided by the Transportation Infrastructure Durability Center at the University of Maine under grant 69A3551847101 from the U.S. Department of Transportation's University Transportation Centers Program.
[provide any additional funding acknowledgements here i.e. your university name, any DOT support, etc.]

References

[1] Xu, W., Li, X., Ji, Y., Li, S. and Cui, C., 2024, BD-YOLOv8s: enhancing bridge defect detection with multidimensional attention and precision reconstruction. *Scientific Reports*, *14*(1), p.18673.

[2] Su, H., Kamanda, D.B., Han, T., Guo, C., Li, R., Liu, Z., Su, F. and Shang, L., 2024. Enhanced YOLO v3 for precise detection of apparent damage on bridges amidst complex backgrounds. *Scientific Reports*, *14*(1), p.8627.

[3] Li, T., Liu, G. and Tan, S., 2024. Superficial Defect Detection for Concrete Bridges Using YOLOv8 with Attention Mechanism and Deformation Convolution. *Applied Sciences*, *14*(13), p.5497.

[4] Mahjoubi, S., Barhemat, R., Guo, P., Meng, W. and Bao, Y., 2021. Prediction and multi-objective optimization of mechanical, economical, and environmental properties for strain-hardening cementitious composites (SHCC) based on automated machine learning and metaheuristic algorithms. *Journal of Cleaner Production*, *329*, p.129665.

[5] Guo, P., Meng, W. and Bao, Y., 2021. Automatic identification and quantification of dense microcracks in high-performance fiber-reinforced cementitious composites through deep learning-based computer vision. *Cement and Concrete Research*, *148*, p.106532.

Object-Guided Sequential Image Stitching Using YOLO and ORB Matching

Xiangjun Meng, Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island Advisor: Paolo Stegagno

Introduction

- Automate the detection of spalling defects across time sequence images using a trained YOLO11 model.
- Track the camera spatial movement by extracting and matching ORB features across consecutive frames.
- Estimate relative motion flow between images based on matched object regions.
- Stitch multiple frames into a single panoramic image that provides a comprehensive view of the bridge deck.
- Support visual inspection and structural monitoring by producing an aligned overview of surface degradation.

System Model

System Overview: YOLO11 + ORB-Based Image Stitching Framework

YOLO11 as a one-stage object detector that predicts spalling class bounding boxes and class probabilities directly from full images in a single forward pass.

Formulating the problem as Deep Learning: conversion

From normalized outputs to image coordinates:

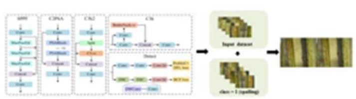
$$x_{\mathrm{abs}} = \sigma(\hat{x}) \cdot w_{\mathrm{cell}} + x_{\mathrm{cell\ origin}}$$

$$w_{\mathrm{abs}} = e^{\hat{w}} \cdot w_{\mathrm{anchor}}$$

Prediction Output per Image Grid Cell:

For each grid cell, YOLO11 predicts:

- (x, y) : center of the bounding box
- (w, h) : width and height
- c: object confidence score
- $p_1, p_2, \ldots, p_c$: class probabilities (e.g., spalling = class 1)



Object-Aware Motion-Based Image Stitching System

1) **Object Detection (YOLO11):** Each image frame is passed through a pretrained YOLO11 model to detect spalling regions (class = 1). For each detection, we obtain a bounding box:

$$\mathrm{bbox}_i = (x_i^{(1)}, y_i^{(1)}, x_i^{(2)}, y_i^{(2)})$$

2) **Feature Extraction and Matching (ORB + BFMatcher):** For each detected spalling object, ORB is used to extract keypoints $\{P_i\}$ and descriptors. Then match key points between adjacent frames.

3) **Motion Flow Estimation:** Accumulating the horizontal displacement across the image sequence to obtain the stitched position of each image

4) **Image Stitching on Global Canvas**

$$\mathbf{p}_k^A \leftrightarrow \mathbf{p}_k^B, \quad k = 1, 2, \ldots, N$$

Results

YOLO11 Training Dataset



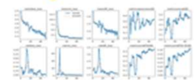
A total of 498 annotated images were utilized, comprising high-resolution frames capturing surface conditions of bridge structures over time.

- Training Set: 300 images
- Testing Set: 198 images
- Good variety of spalling imagery, but no cracking, delamination, and limited efflorescence

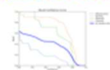
Associated Logos and Contacts
Go Here (Sponsors, Partners, etc.)

Feature Matching Result

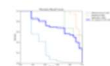
Training Cureve



PR Cureve & P Cureve



Confusion Matrix



Conclusion

An object-aware stitching framework that combines YOLO11 detection and ORB-based feature matching to reconstruct panoramic views of the bridge deck.

- **Reliable deep learning model to automatically identify spalling defects in bridge concrete elements (90%).**
- **Promising initial stitching results based on feature extraction.**
- **Future Work: improve stitching results by introducing semantic segmentation and image reprojection**

Acknowledgements: Funding for this research is provided by the Transportation Infrastructure Durability Center at the University of Maine under grant 69A3551847101 from the U.S. Department of Transportation's University Transportation Centers Program.

June 2024 – www.tidc-utc.org

TIDC

Transportation Infrastructure Durability Center
AT THE UNIVERSITY OF MAINE

35 Flagstaff Road
Orono, Maine 04469
tidc@maine.edu
207.581.4376

www.tidc-utc.org