Unsupervised Semantic Segmentation with Pose Prior (USP)

Max Midwinter, Chul Min Yeum, Sriram Narasimhan

Chul Min Yeum

Assistant Professor

Civil and Environmental Engineering

University of Waterloo

Presentation at SHMII11
Aug 8, 2022

UNIVERSITY OF WATERLOO
FACULTY OF ENGINEERING

Computer Vision for Smart Structure

Current Visual Inspection Process



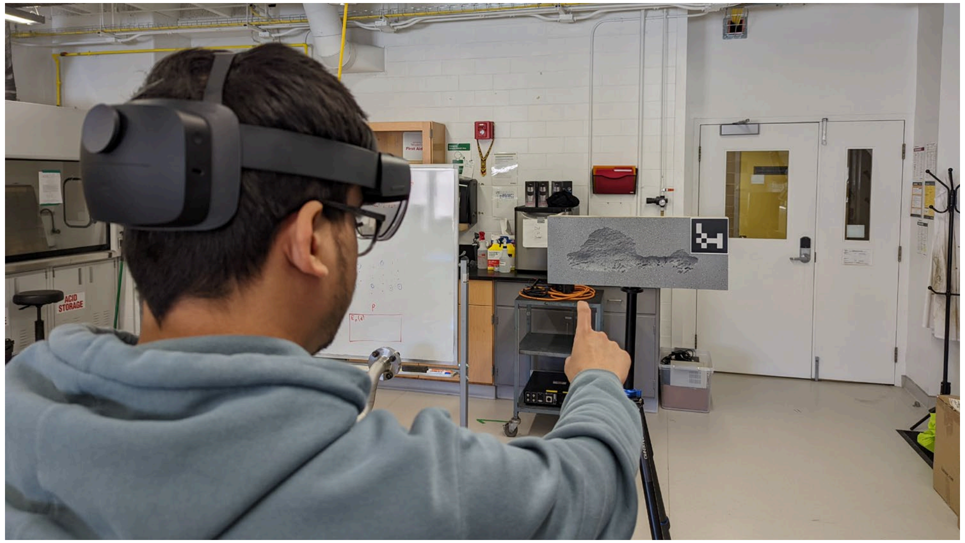

- Manual process for inspectors
- Use basic tools such as measuring tape.

Disadvantages:

- Time-consuming
- Inaccurate
- Expensive
- Inaccessible regions
- Inefficient
- Dangerous

Vision-based Inspections

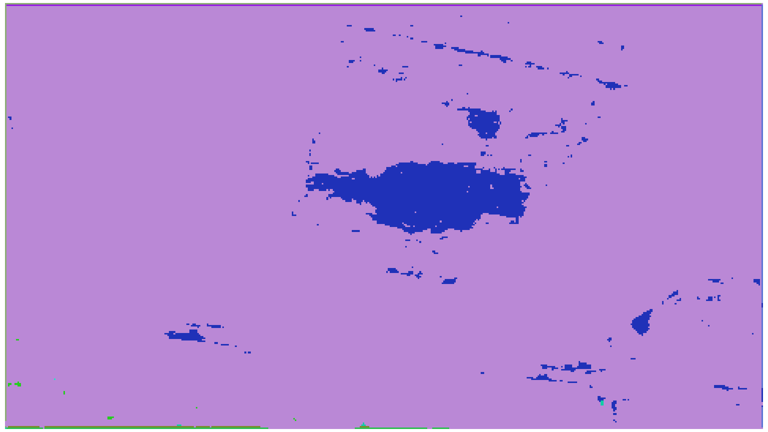
- Automated platforms: robots, drones, wearable glasses
- Use AI to automatically detect and localize defects in images.
- Measure defect sizes using 3D sensors (LiDAR, time-of-flight).



Vision-based Inspection Tasks: Quantification?

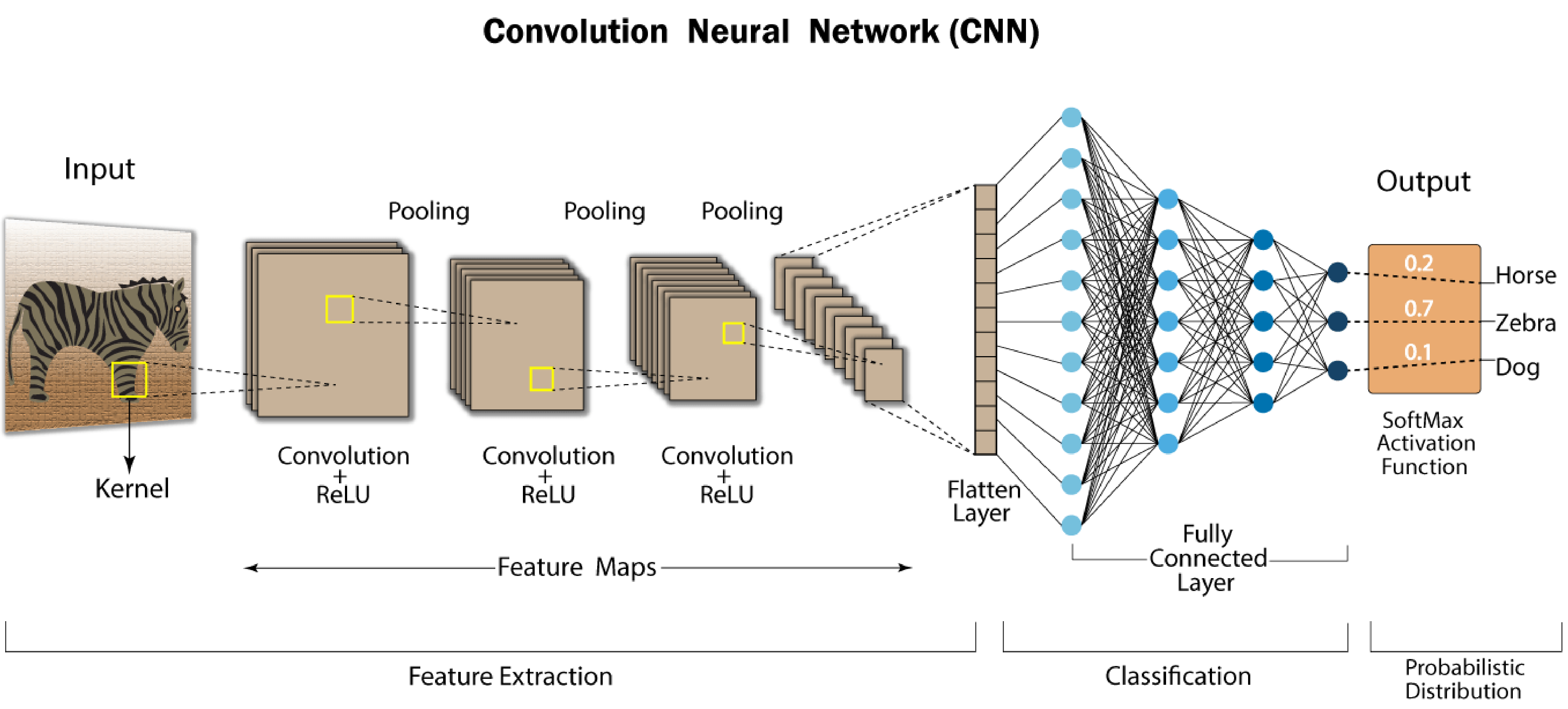
Classification

Detection



Segmentation

Convolutional Neural Network



Challenges

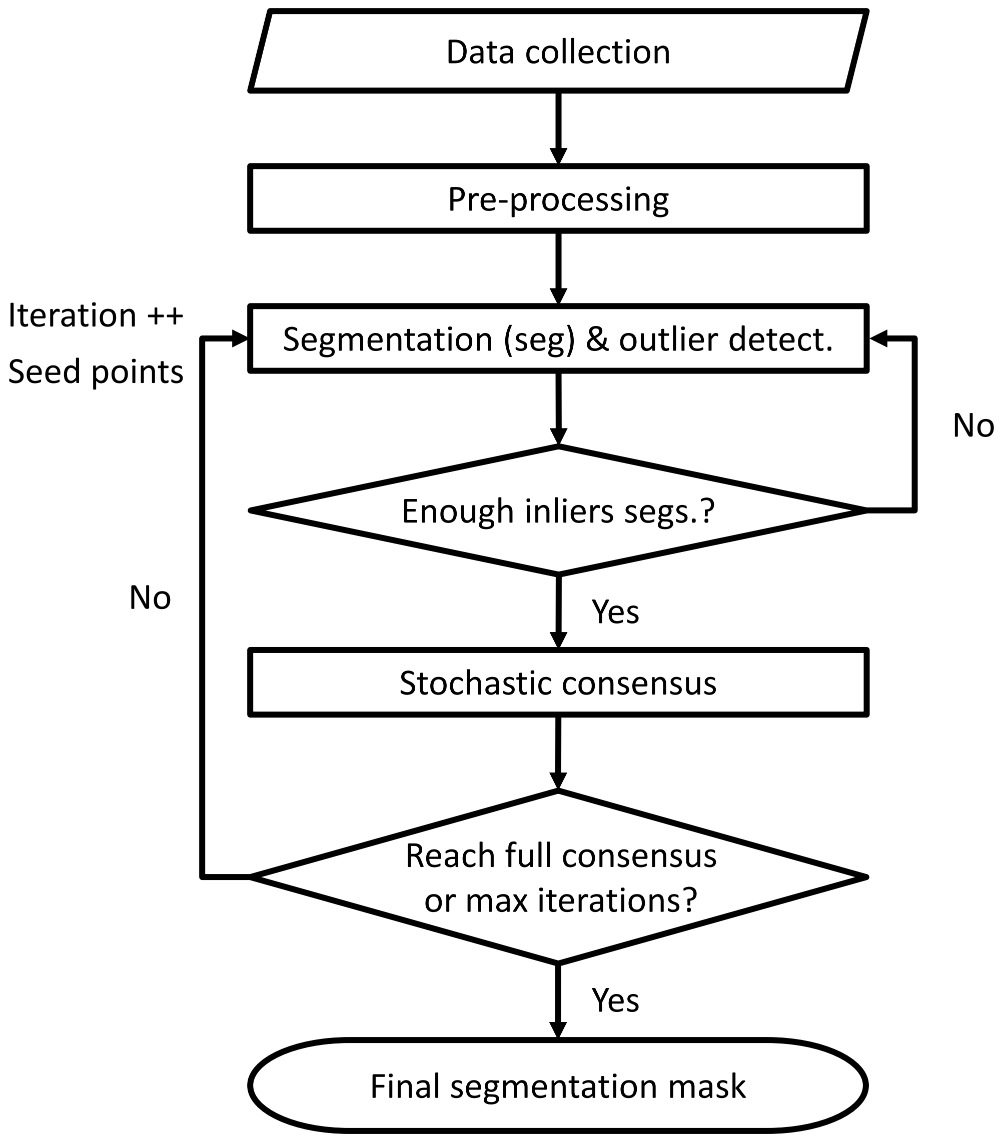
1) Creating a labeled segmentation datasets is time-consuming, tedious, and expensive for training supervised segmentation models.

2) In a typical inspection using a robotic platform (e.g., drones), videos are captured, and many frames contain the same ROI, captured from different camera positions and angles is challenging to specify unique segmentation ROI by weeding out incorrect or, merging suboptimal segmentation and selecting/integrating correct segmentation results.

Objectives

1. Propose unsupervised semantic segmentation with pose prior (USP), which best leverages novel developments in machine learning and capabilities of existing robotic inspection.
2. Utilizes **a series of images** (with pose) featuring the ROIs from various known camera poses to improve segmentation accuracy.
3. Provide accurate segmentation masks without using eliminates the need for expensive labeled data; meanwhile enforcing consistency between segmentation results from multiple frames.

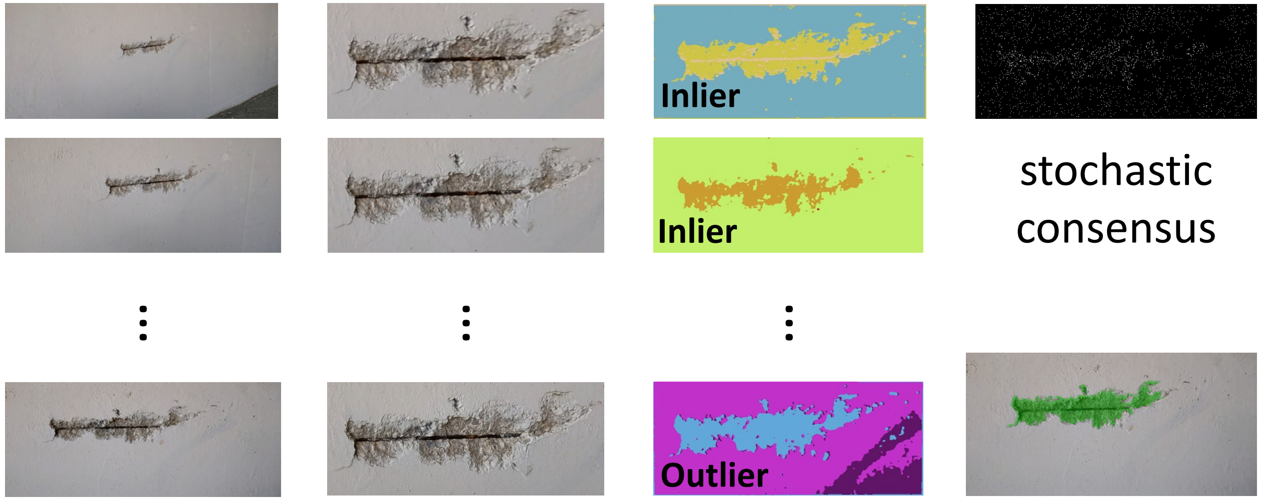
Overview of the USP



Inspection site & target ROI

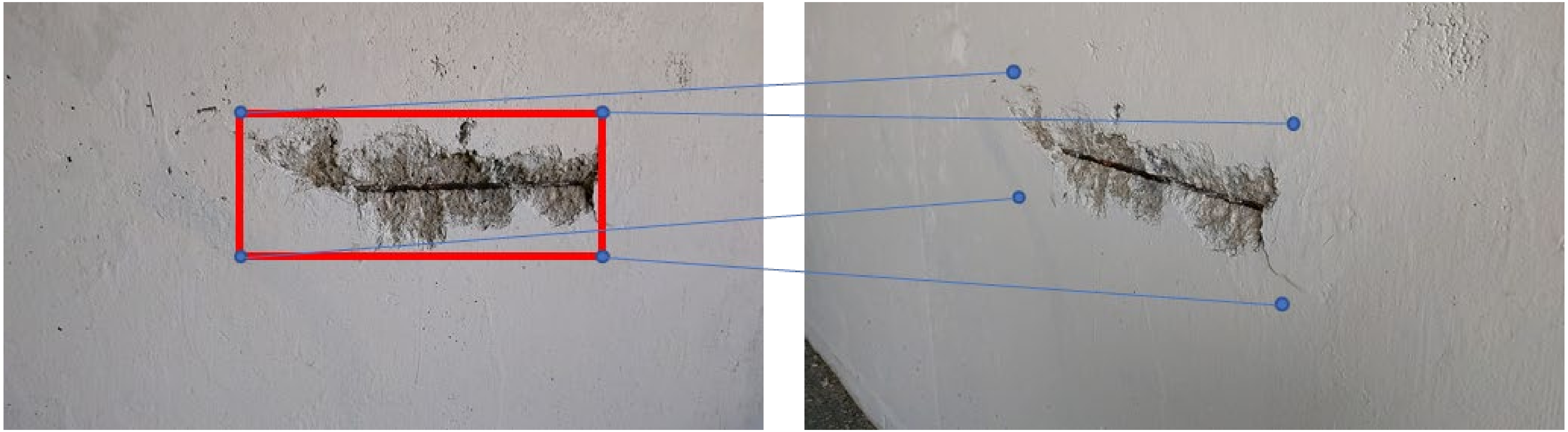
Pre-requisite

- Camera pose
- Bounding box detector (or manual selection)



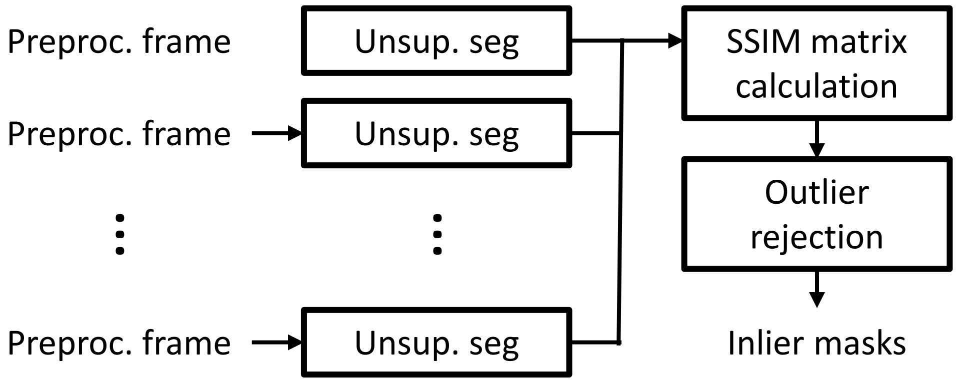
Data collection | Pre-pros. frames | Seg. & outlier rejection | Final seg. mask

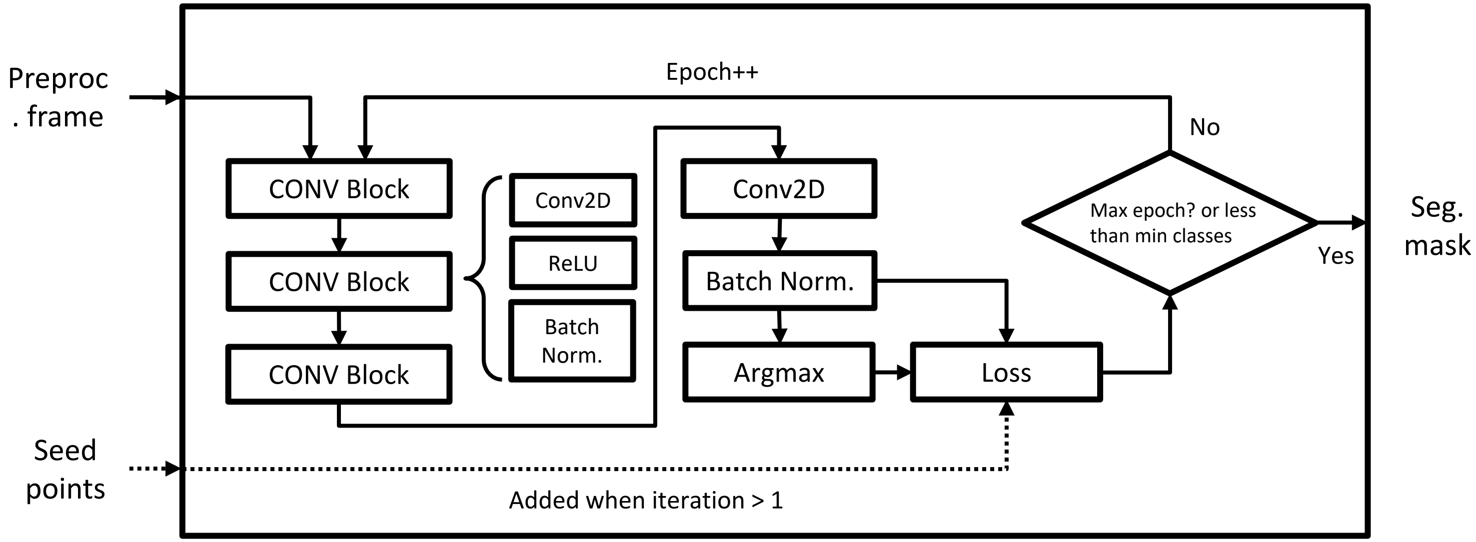
Pre-Processing: Homography Transformation



Reference frame **Other frames**

Unsupervised Segmentation (Architecture)





Kim, W., Kanezaki, A., & Tanaka, M. 2020. "Unsupervised learning of image segmentation based on differentiable feature clustering." IEEE Transactions on Image Processing, 29, 8055-8068.

Unsupervised Segmentation (Loss Function Design)

$$L = \underbrace{L_{\text{sim}}(\{\boldsymbol{r}'_n, c_n\})}_{\text{feature similarity}} + \mu \underbrace{L_{\text{con}}(\{\boldsymbol{r}'_n\})}_{\text{spatial continuity}} + \nu \underbrace{L_{\text{scr}}(\{\boldsymbol{r}'_n, s_n, u_n\})}_{\text{scribble information}}.$$

1) Feature similarity

Enhance the similarity of the similar features.

2) Spatial Discontinuity

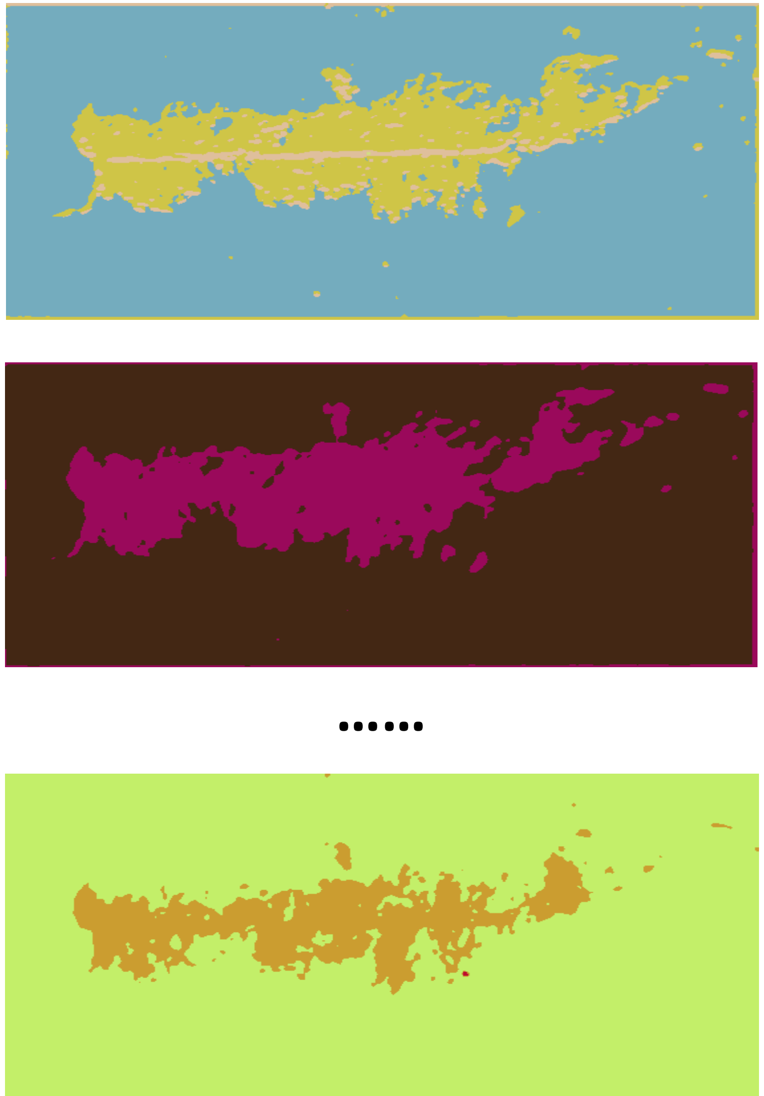
Group similar pixels into clusters: The constraint is introduced that favors the cluster labels to be the same as those of the neighboring pixels.

3) Scribbles (seed point) as a user input

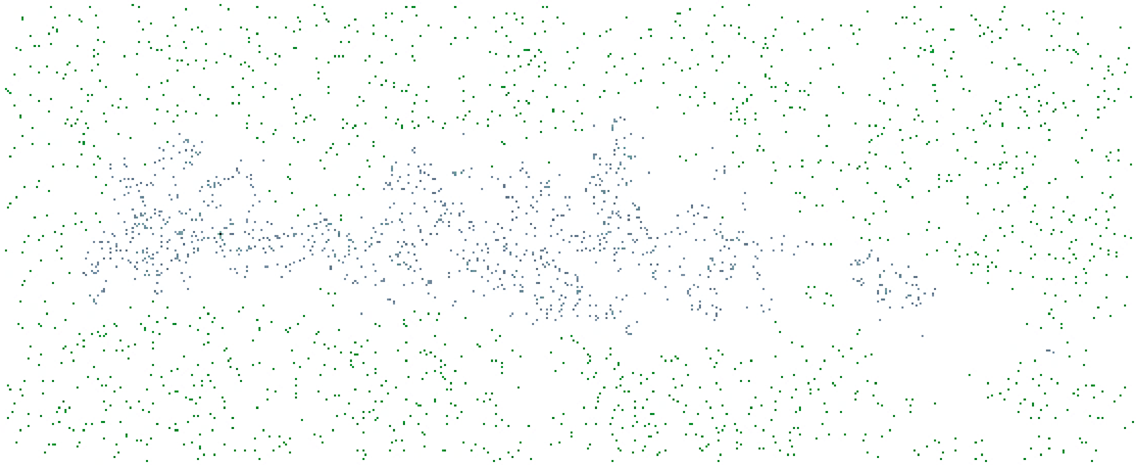
Incorporate multi-frames segmentation results.

Kim, W., Kanezaki, A., & Tanaka, M. 2020. “Unsupervised learning of image segmentation based on differentiable feature clustering.” IEEE Transactions on Image Processing, 29, 8055-8068.

Stochastic Consensus



Inlier Segmentation



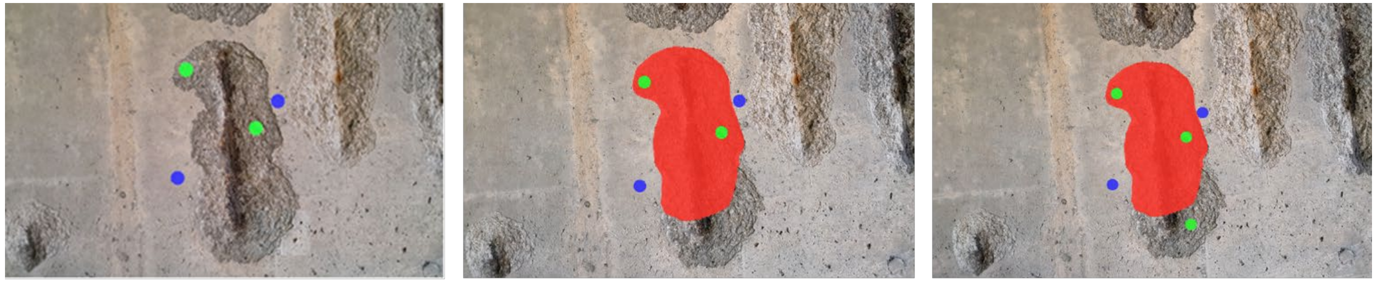
Seed point segmentation

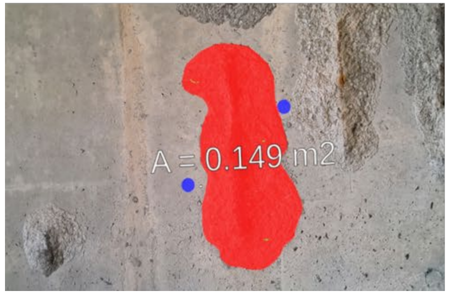
Experiment: Site

Overview of the test bridge

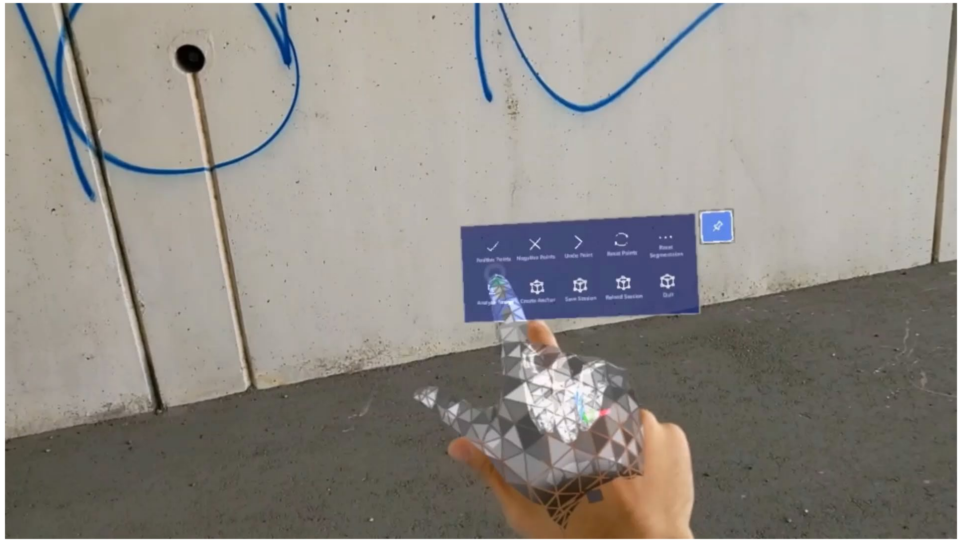
Previous Work: XRIV

- e**X**tended **R**eality **I**nspection and **V**isualization (XRIV).
- Localize and quantify defects.
- Interactive segmentation through MR.
- Limited Range: <5m.





Al-Sabbag, Z. A., Yeum, C. M., & Narasimhan, S. (2022). Interactive defect quantification through extended reality. *Advanced Engineering Informatics*, *51*, 101473.



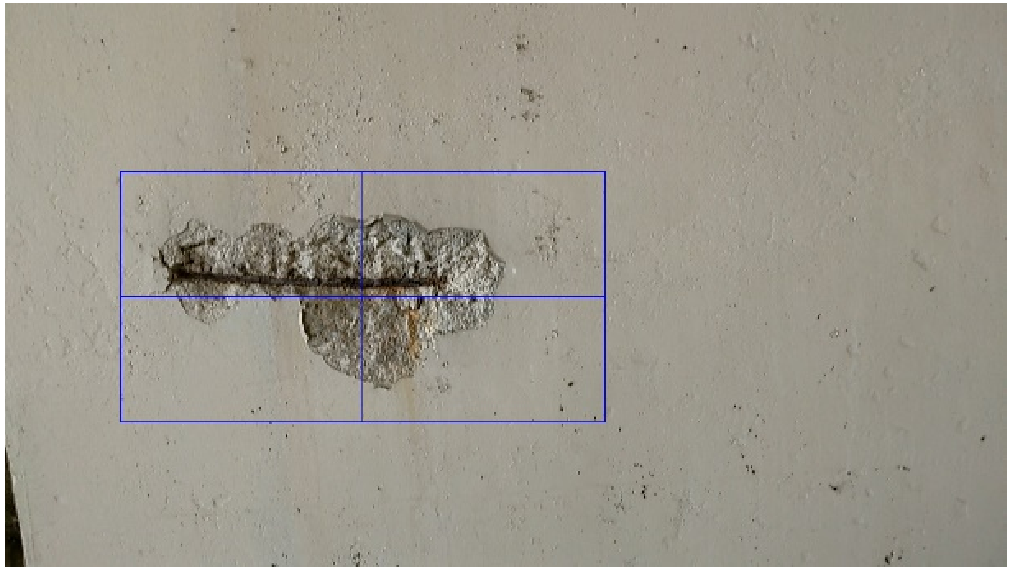
rames from Two Damage Locations

Damage 1

Damage 2

Bounding Box Detection from A Reference Frame (Manual)

Damage 1



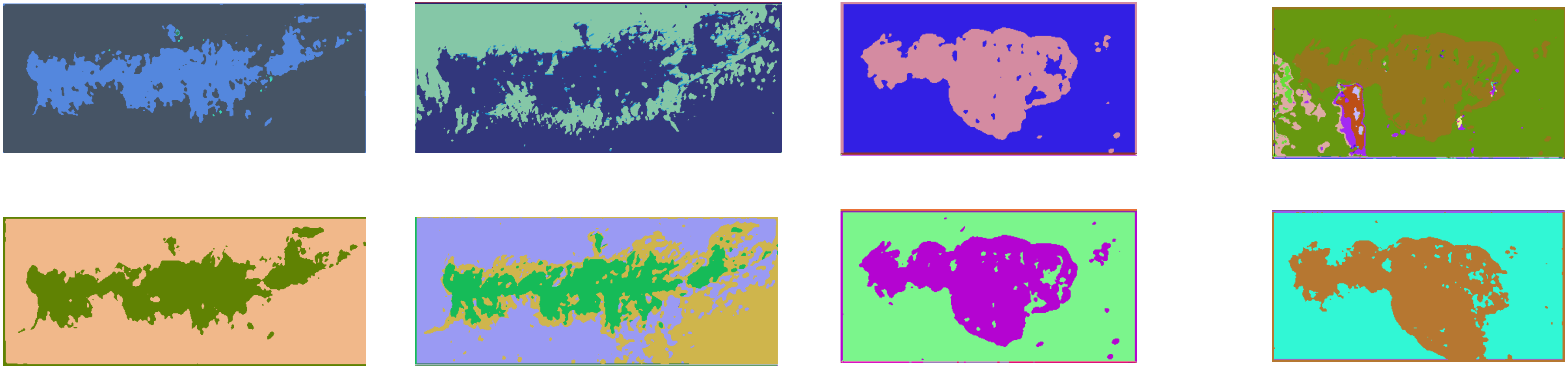
Damage 2

Preprocessed Frames

Damage 1

Damage 2

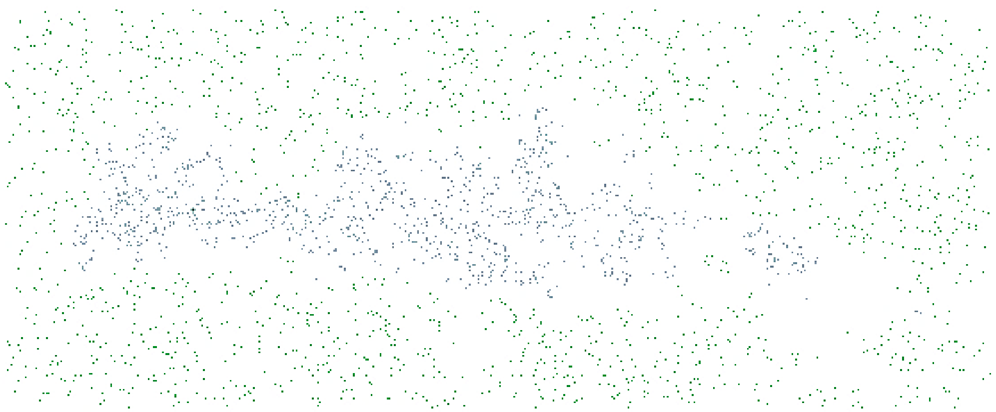
Raw Frames from Two Damage Locations



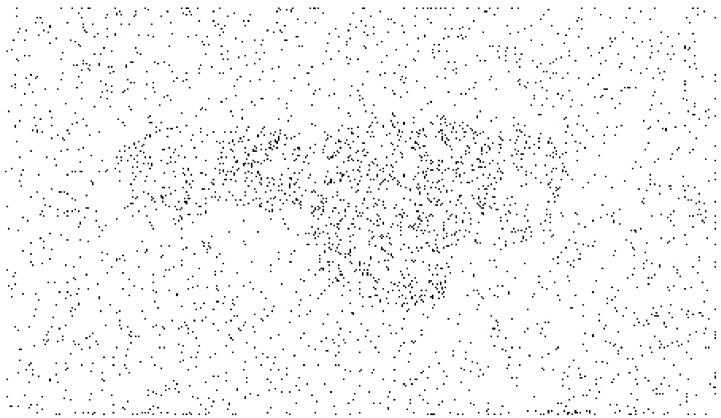
Inlier Outlier Inlier Outlier

Damage 1 Damage 2

Stochastic Consensus



Damage 1



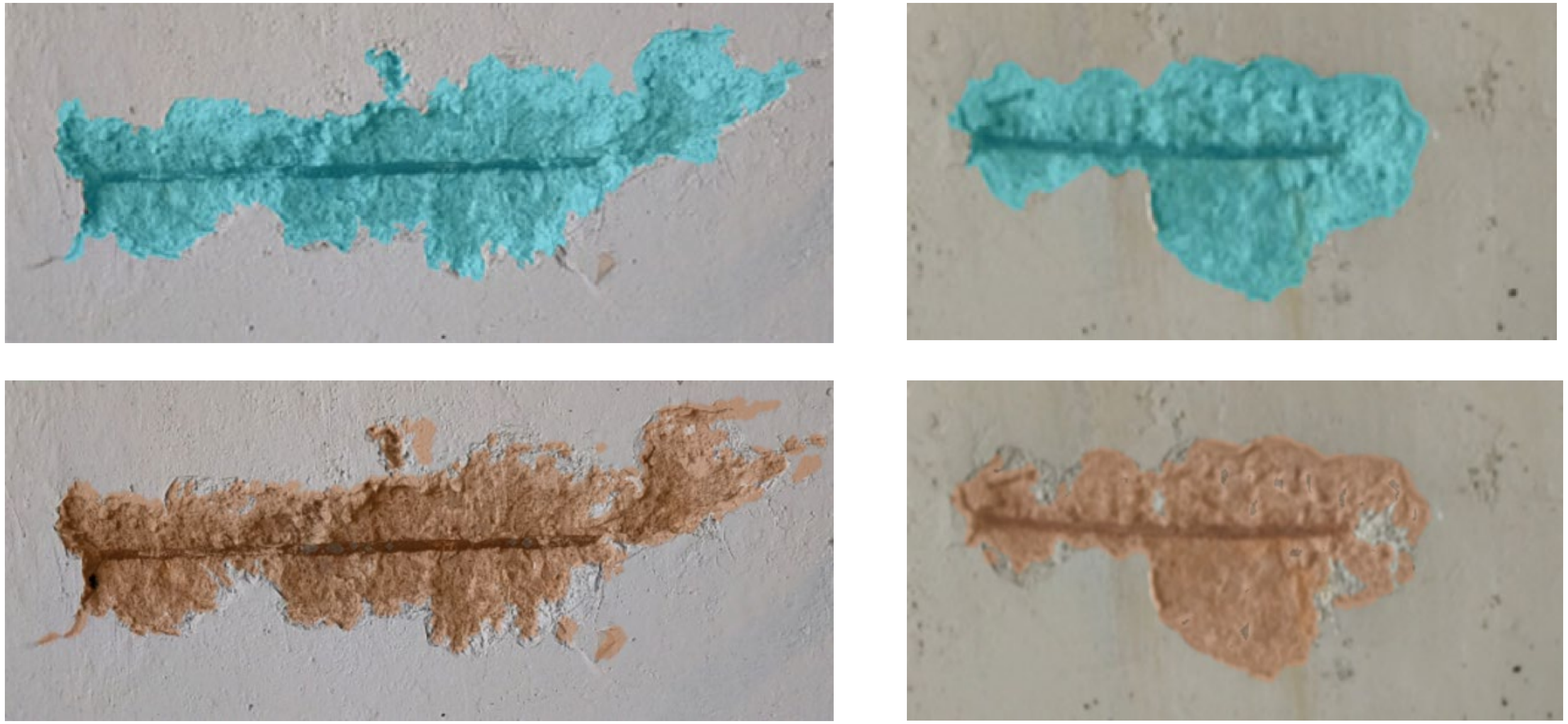
Damage 2

Segmentation Results

Damage 1 | Damage 2

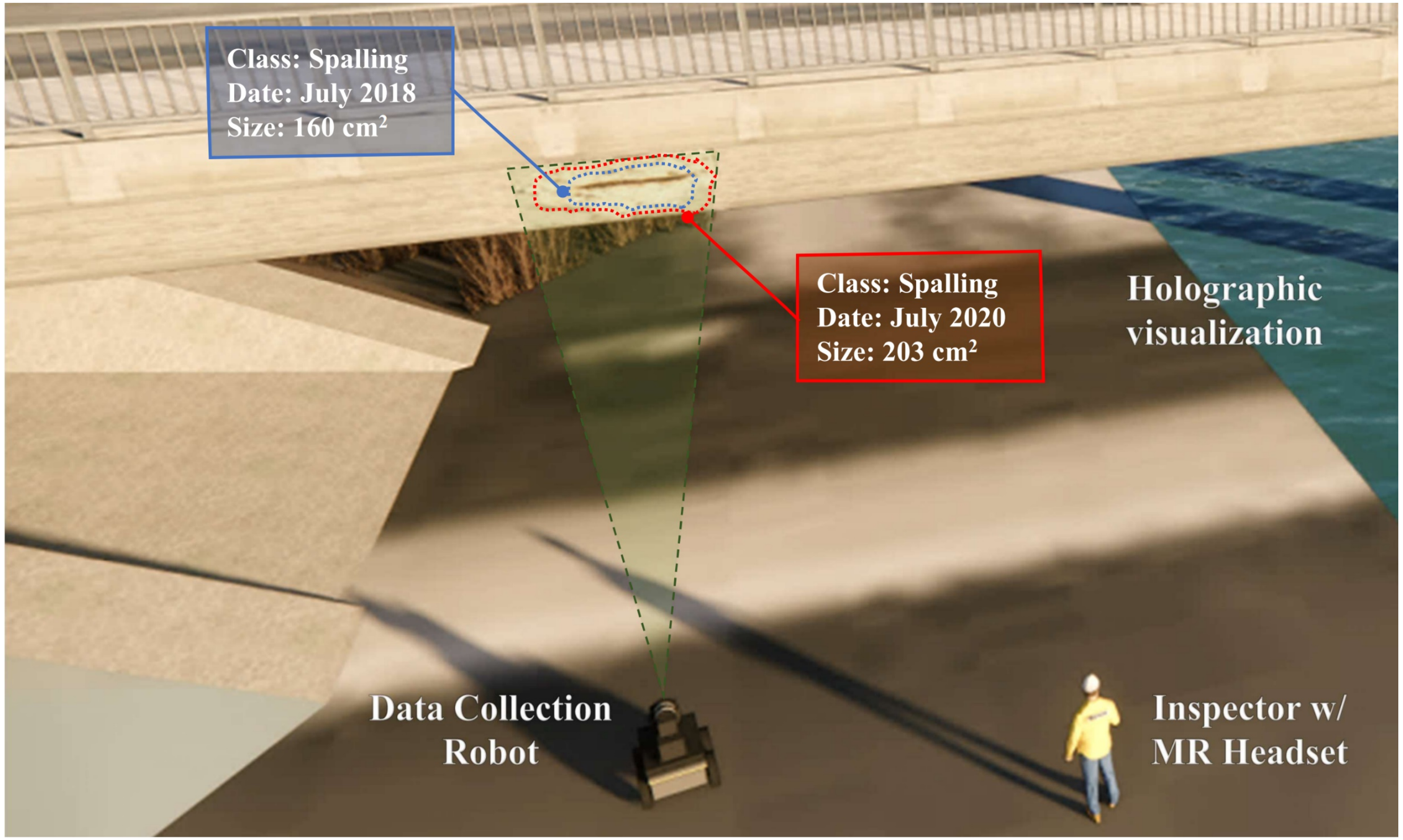
Ground truth

USP



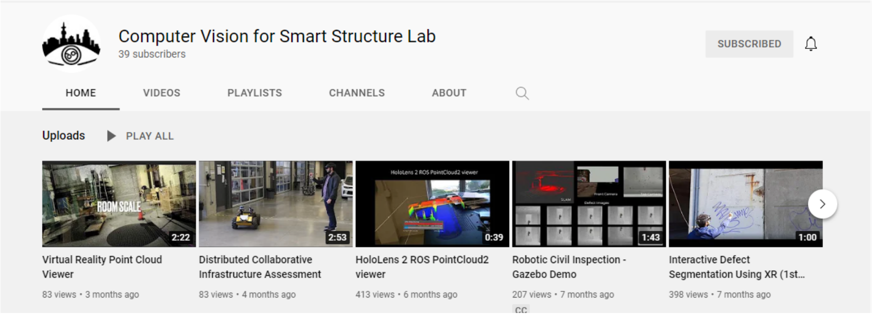
	IoU (%)
Defect 1	84.8
Defect 2	87.7

Human-Machine Collaborative Inspection (HMCI)



- **Robot** scans bridge using high-quality cameras and LiDAR to build a 3D map.
- **Remote server** analyze images to locate defects using deep learning.
- **Inspector** visualizes defect region on the real structure using MR.
- Inspector can re-visit structure and compare changes of defect region size over time.

Thank you! Any Question



https://cviss.net

UNIVERSITY OF WATERLOO
FACULTY OF ENGINEERING

ROGERS

Computer Vision for Smart Structure

Background