# Automated Region-of-Interest Localization and Classification for Visual Assessment

**Chul Min Yeum**[1], Jongseong Choi[2], and Shirley J. Dyke[2,1]

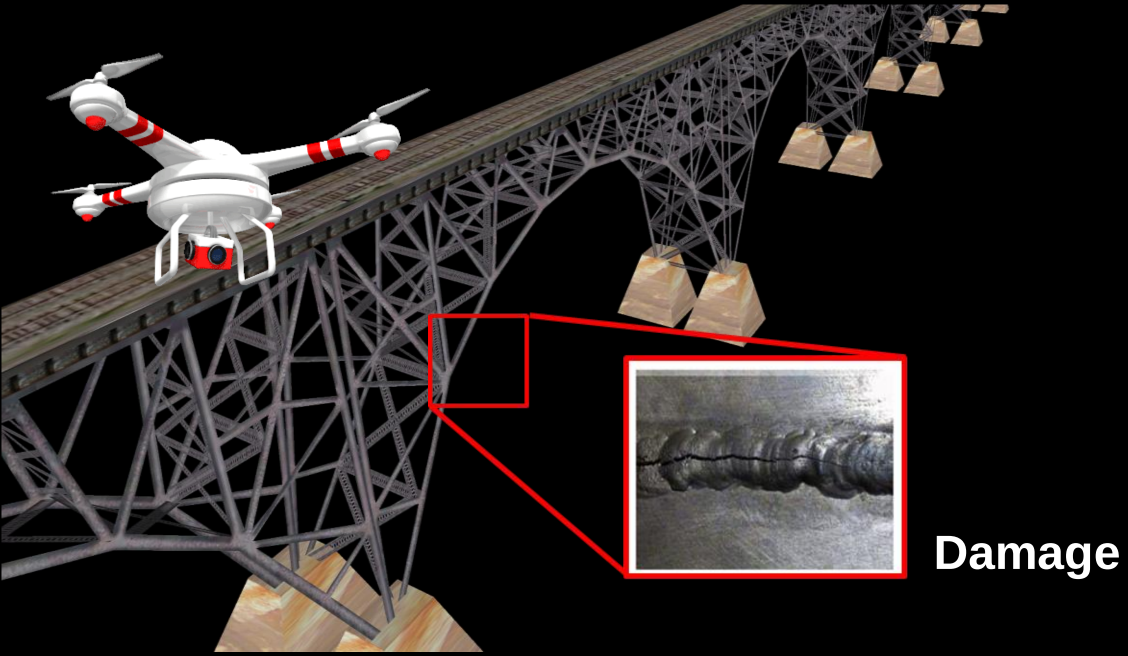
[1] Postdoctoral Researcher, Lyles School of Civil Engineering, Purdue University, USA

[2] Graduate Student, School of Mechanical Engineering, Purdue University, USA

[3] Professor, School of Mechanical and Civil Engineering, Purdue University, USA

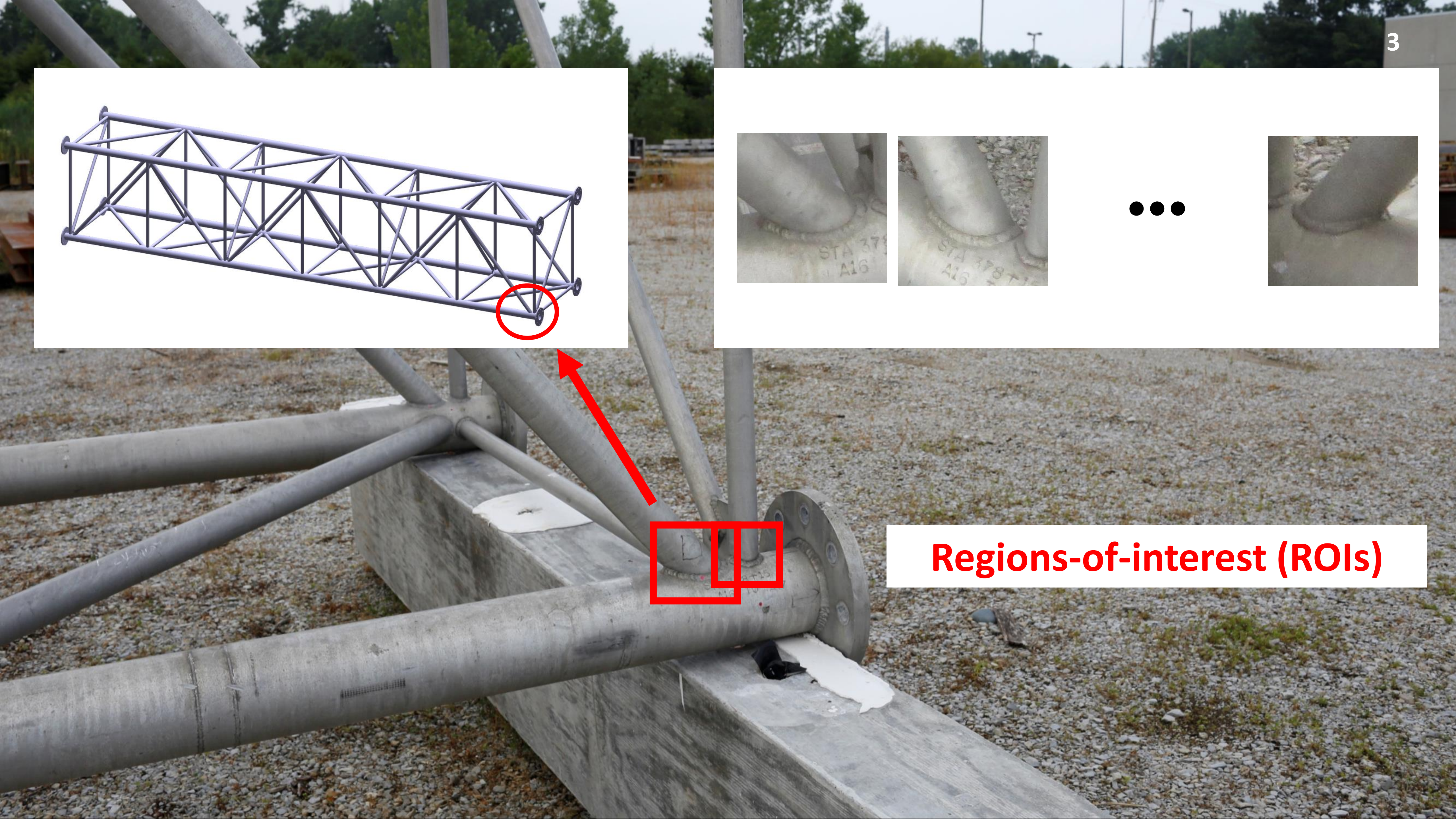
# Opportunity






Automated visual Inspection using drones

A large volume of images collected from drones


Regions-of-interest (ROIs)

## Objective

Develop a technique that can automatically localize and classify the **Regions-Of-Interest** (ROI) on each of the collected images so as to process and analyze only highly relevant and localized image areas for visual inspection or damage detection.

## Advantage

Develop an enabling technique to facilitate successful application of **existing damage detection techniques** on large volumes of actual images in an efficient and reliable way. The key is to avoid unnecessary processing of the large portion that are irrelevant and complex.

**Localization and classification**

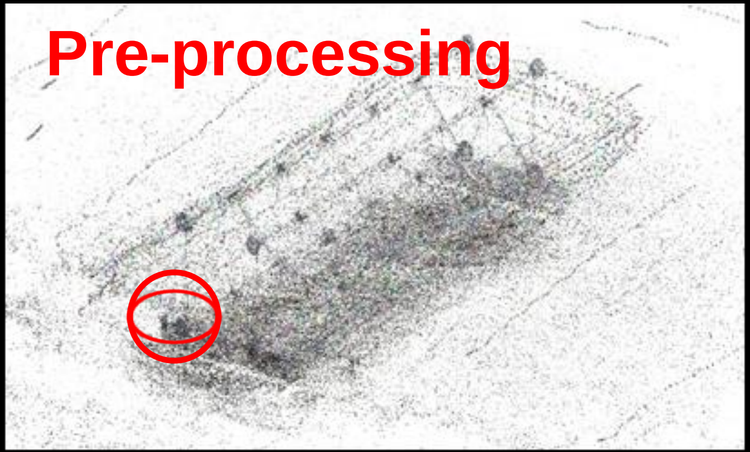
**ROIs**

**Human based visual inspection**

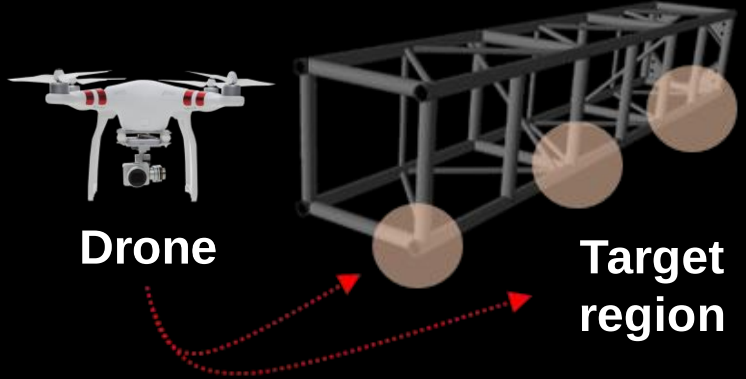
**Autonomous damage detection**

# Overview of the Technical Steps

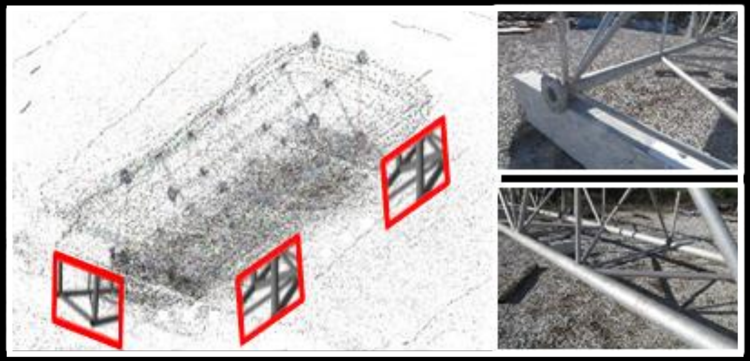




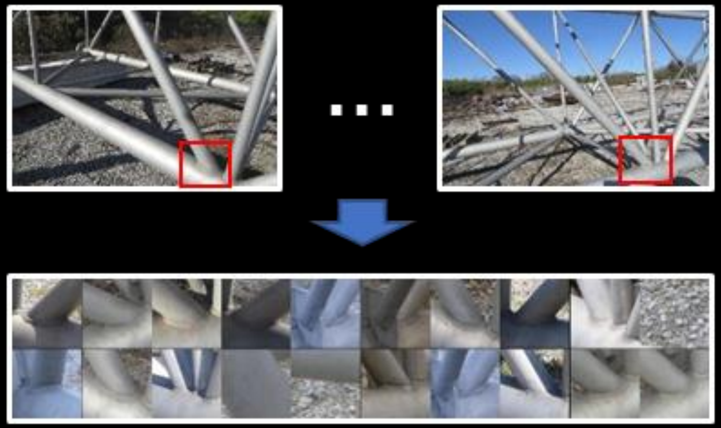

(a) Baseline model construction



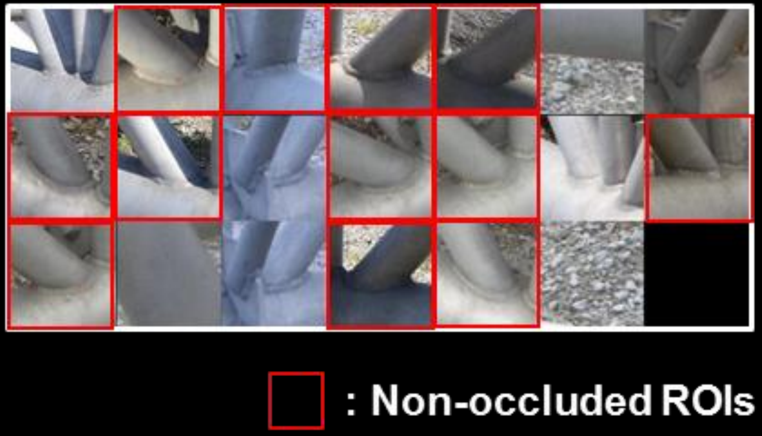

(b) Step 1: Image collection



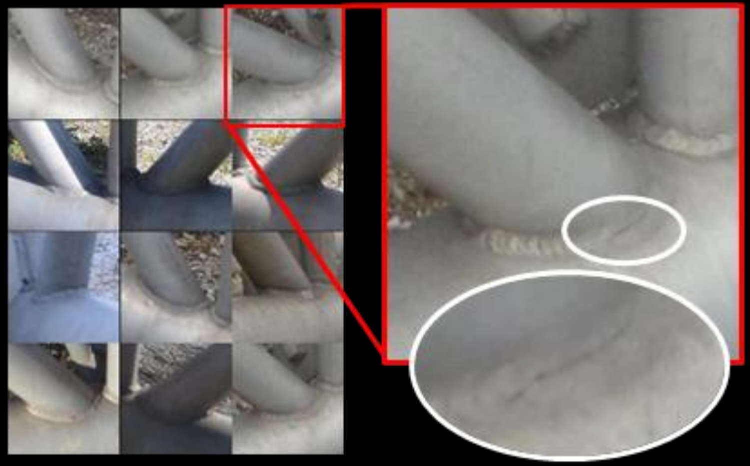
(c) Step 2: Image registration




(d) Step 3: ROI localization




(e) Step 4: ROI classification

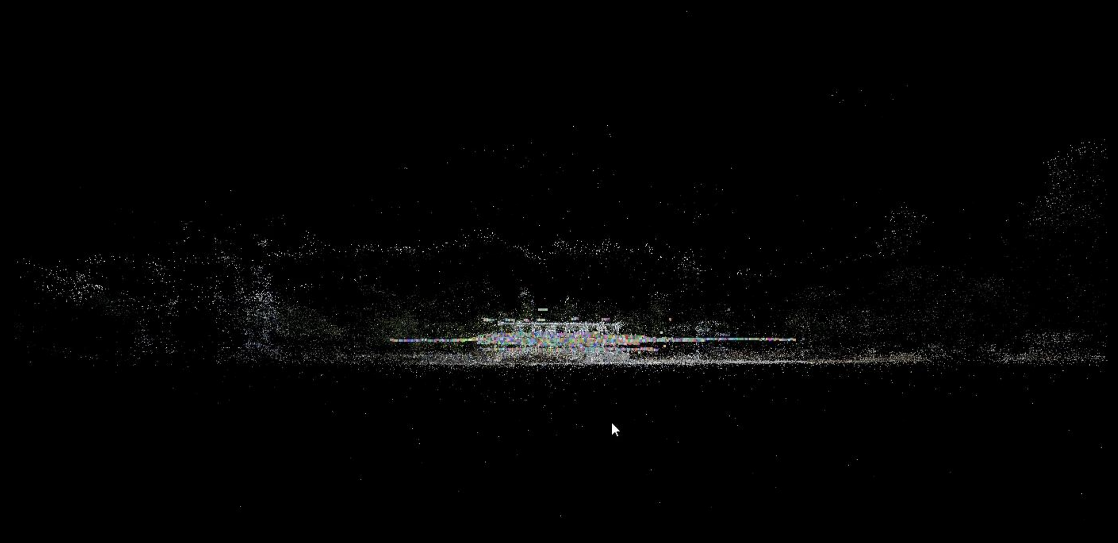


(f) Step 5: Damage detection

# What is Structure from Motion (SfM)?

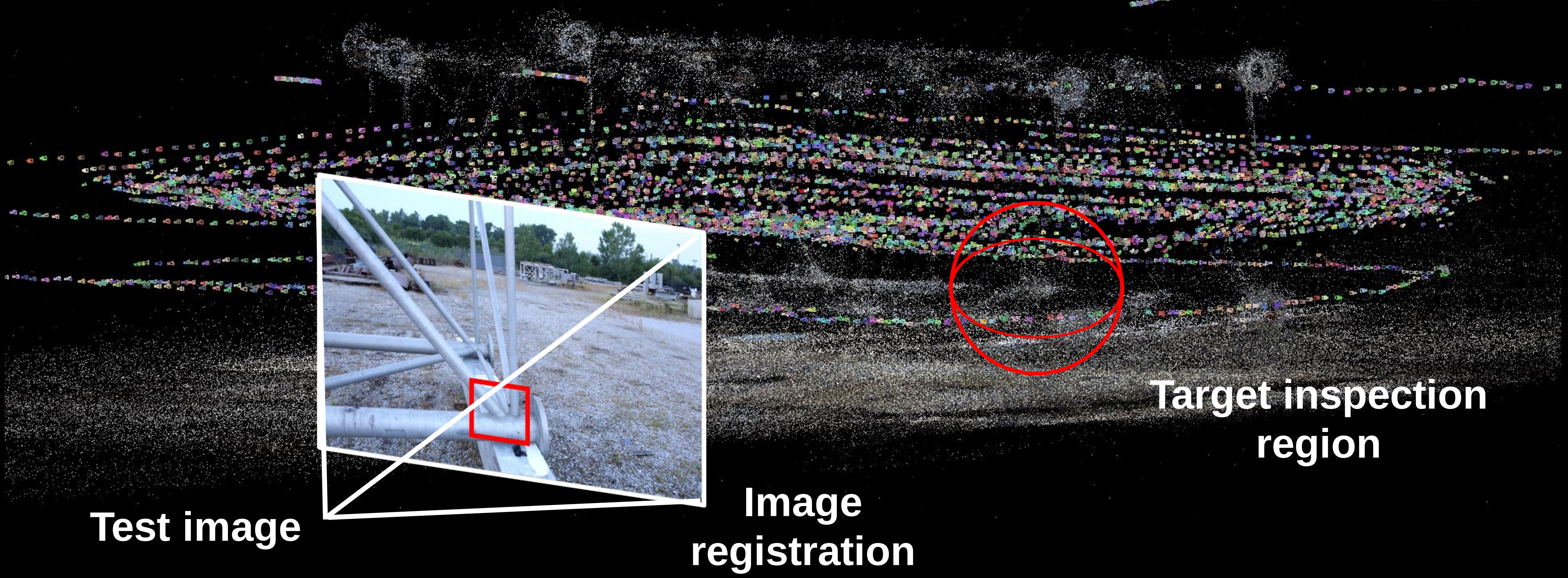


Pictures

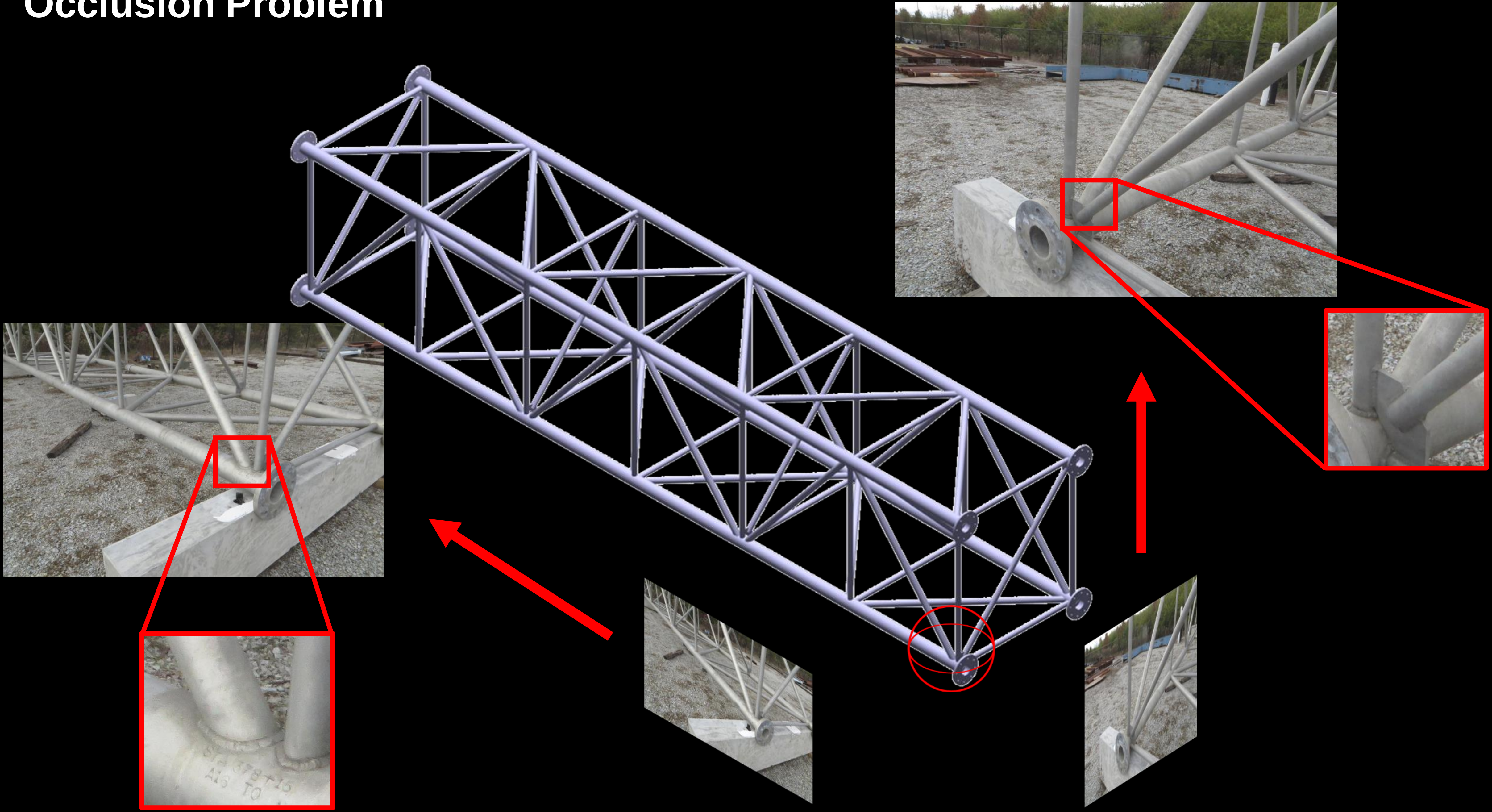


Scene structure & Camera locations and parameters

# ROI Localization using Geometric Relationships



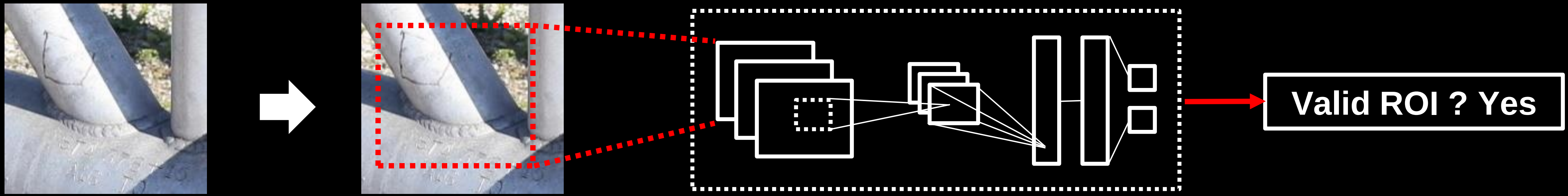
# Occlusion Problem



# ROI Classification using Convolutional Neural Network (CNN)

Collection of test images

Occlusion




Input image

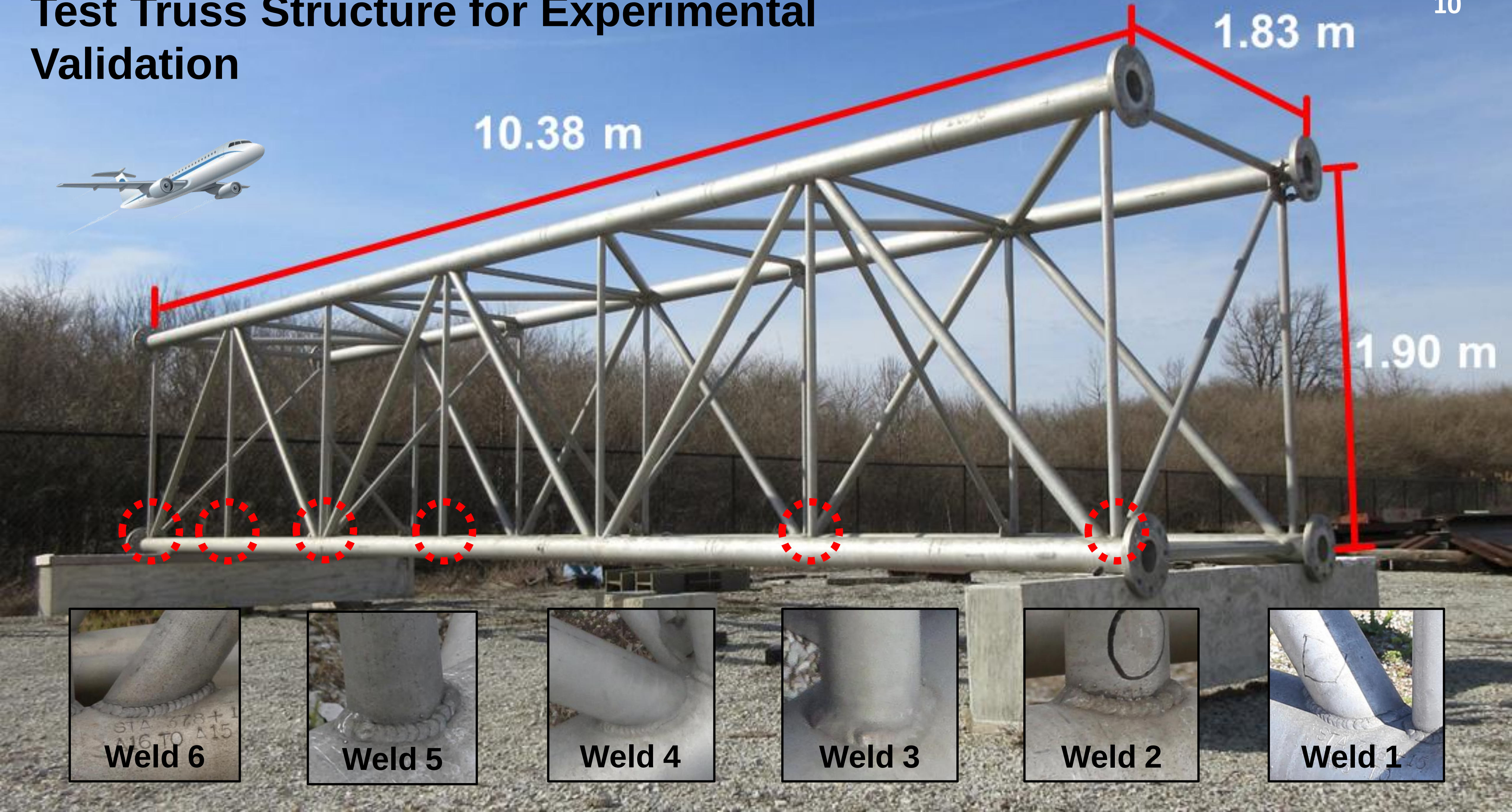
Data augmentation

Compute CNN features

Classification

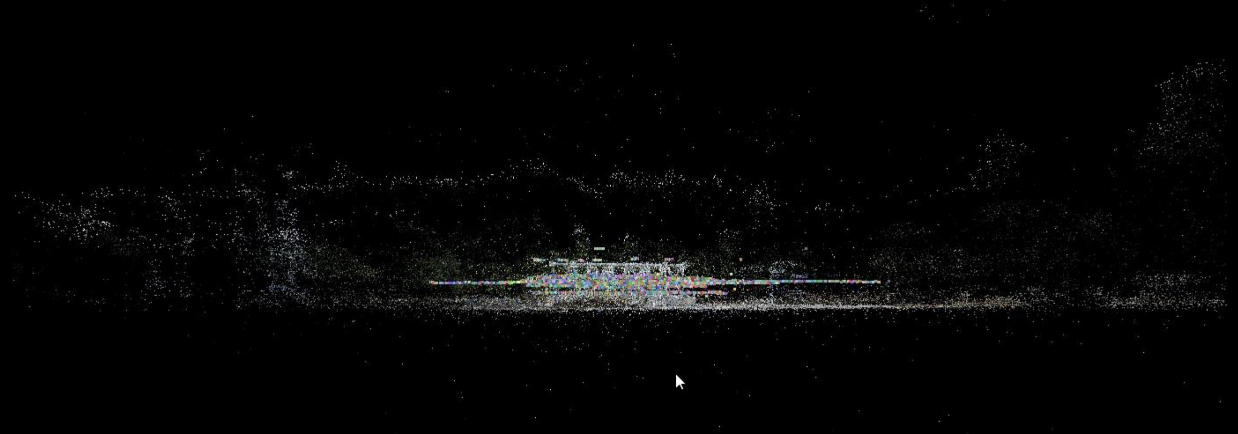
Training of binary occlusion classifier using convolutional neural network (CNN)

# Test Truss Structure for Experimental Validation



# Baseline Model Construction (Pre-processing)





A total of **5,321** images are collected from the test structure during **five** months and **11 different days** under different time window in a day and/or weather conditions.

# Training a Binary Occlusion Classifier (Pre-processing)

Positive

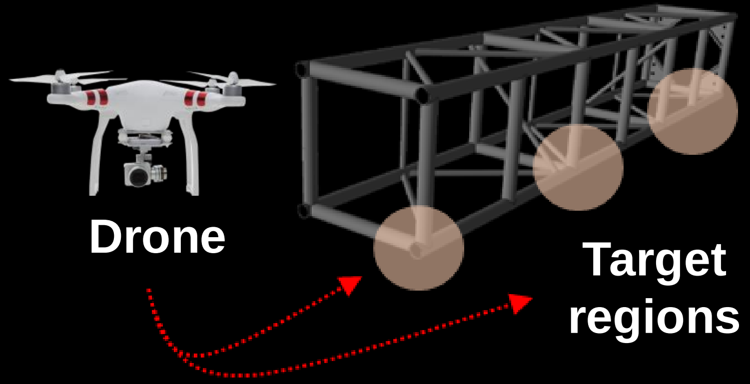
Negative

**If the ROI is <u>positive</u>, the entire weld line on the ROI is visible**

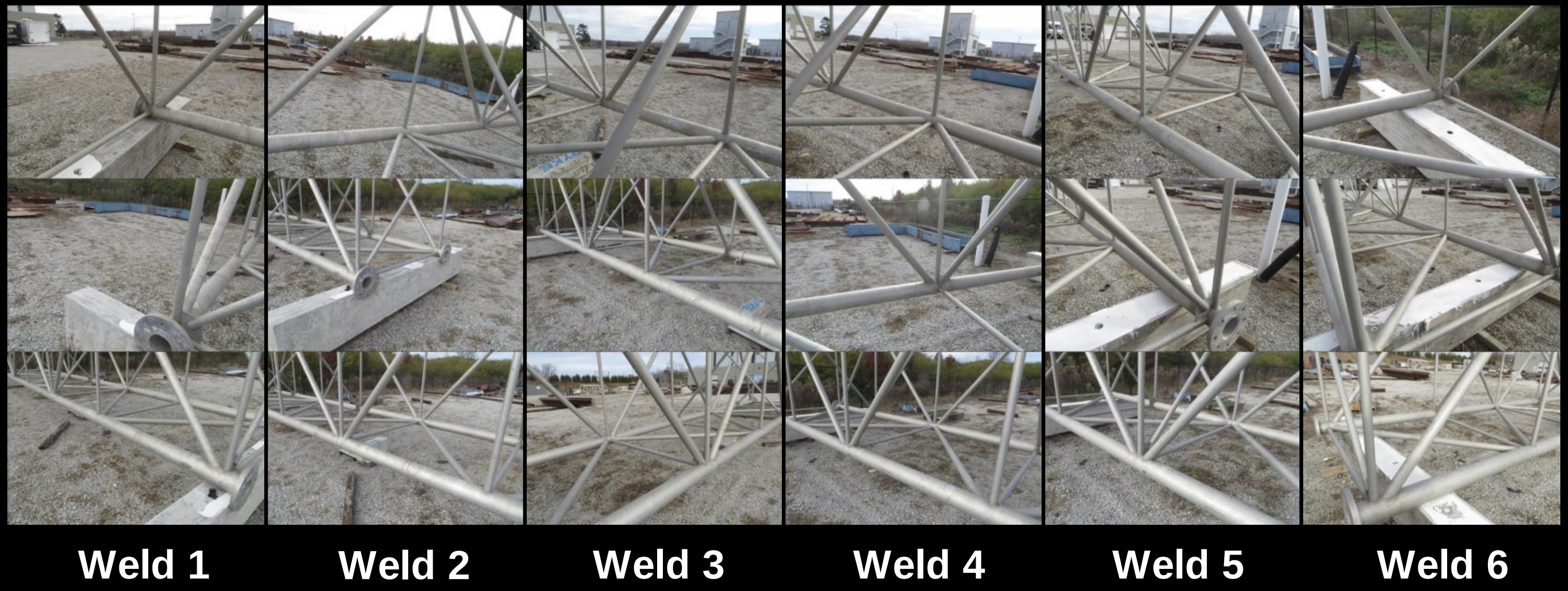
**<u>Configuration</u>**

- **CNN architecture : Alexnet for binary class.**
- **CNN framework (library) : MatCovnet (in  MATLAB)**
- **# of pos. and neg. images : 3,353/ 945 images**

# Image Collection and Registration




(b) Step 1: Image collection




| | Weld 1 | Weld 2 | Weld 3 | Weld 4 | Weld 5 | Weld 6 |
|---|---|---|---|---|---|---|
| # of images | 119 | 77 | 88 | 84 | 60 | 55 |

# ROI Localization



Weld 1 Weld 2 Weld 3 Weld 4 Weld 5 Weld 6

# Samples of Localized ROIs from Weld 1, 3, and 6

## Weld 1

## Weld 3

## Weld 6

# Results of the ROI Localization

| | Weld 1 | Weld 2 | Weld 3 | Weld 4 | Weld 5 | Weld 6 |
|---|---|---|---|---|---|---|
| # of images | 119 | 77 | 88 | 84 | 60 | 55 |
| # of localized ROIs | 104 | 51 | 54 | 70 | 45 | 47 |

Too small (insufficient resolution)

Not visible

# Samples of Localized and Classified ROIs from Weld 1, 3, and 6

## Weld 1

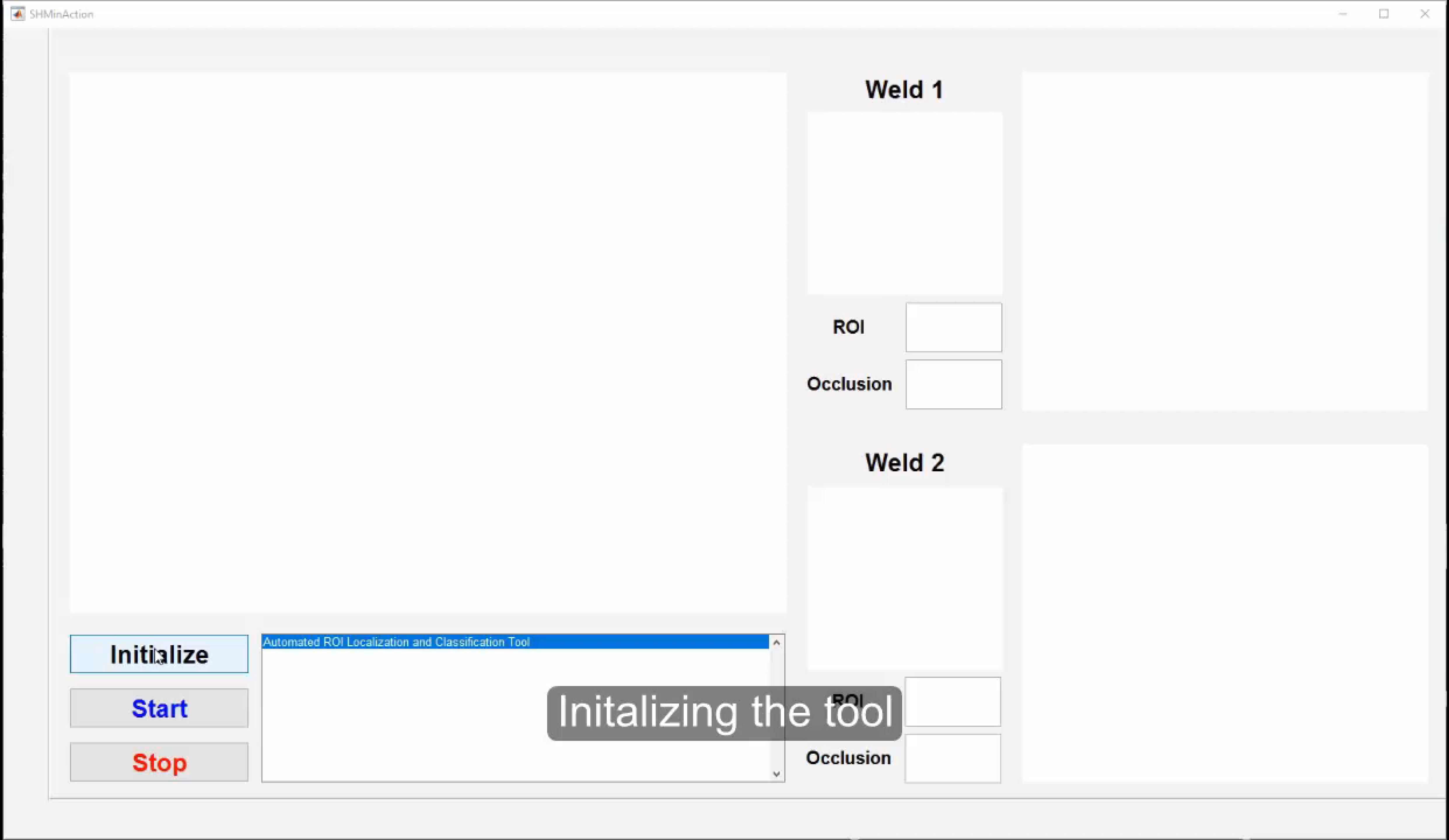
## Weld 3

## Weld 6

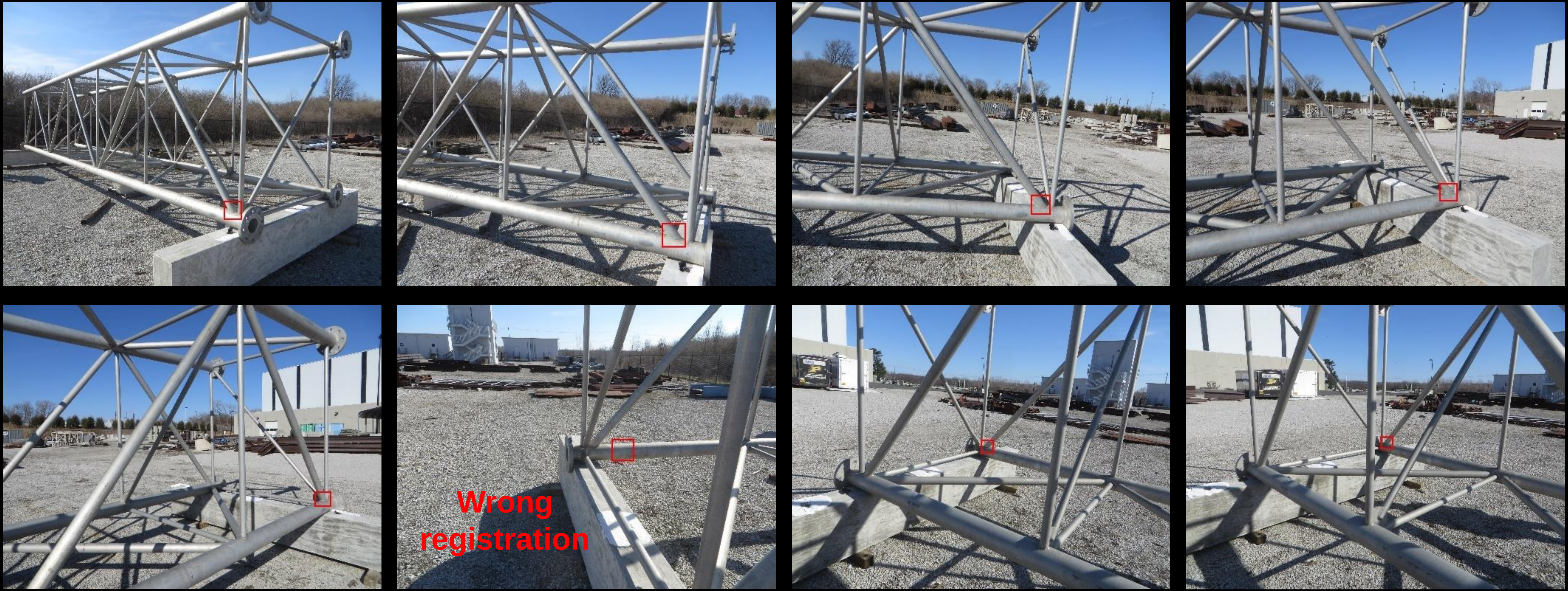
# Results of the ROI Localization and Classification

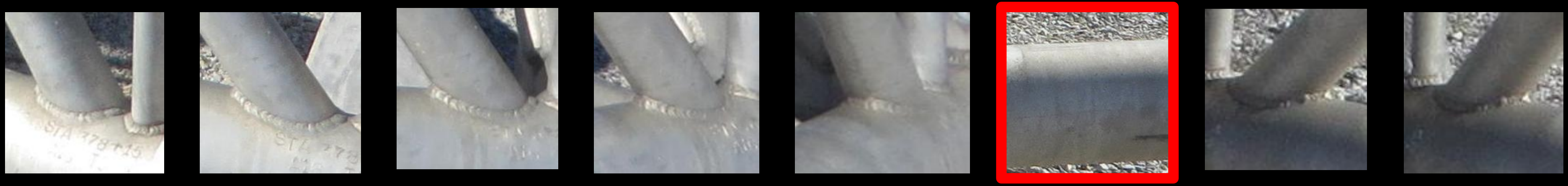
| | Weld 1 | Weld 2 | Weld 3 | Weld 4 | Weld 5 | Weld 6 |
|---|---|---|---|---|---|---|
| # of images | 119 | 77 | 88 | 84 | 60 | 55 |
| # of localized ROIs | 104 | 51 | 54 | 70 | 45 | 47 |
| # of classified ROIs (positive/negative) | 69/35 | 49/2 | 48/6 | 47/23 | 44/1 | 33/14 |
| Precision | 92.75% | 100% | 97.92% | 85.11% | 100% | 90.91% |

# Application of On-board Image Analysis


SHMinAction
Weld 1
ROI
Occlusion
Weld 2
ROI
Occlusion
Initialize
Start
Stop
Automated ROI Localization and Classification Tool
Initalizing the tool

# Test Images Collected Four Months Later



Detected as negative

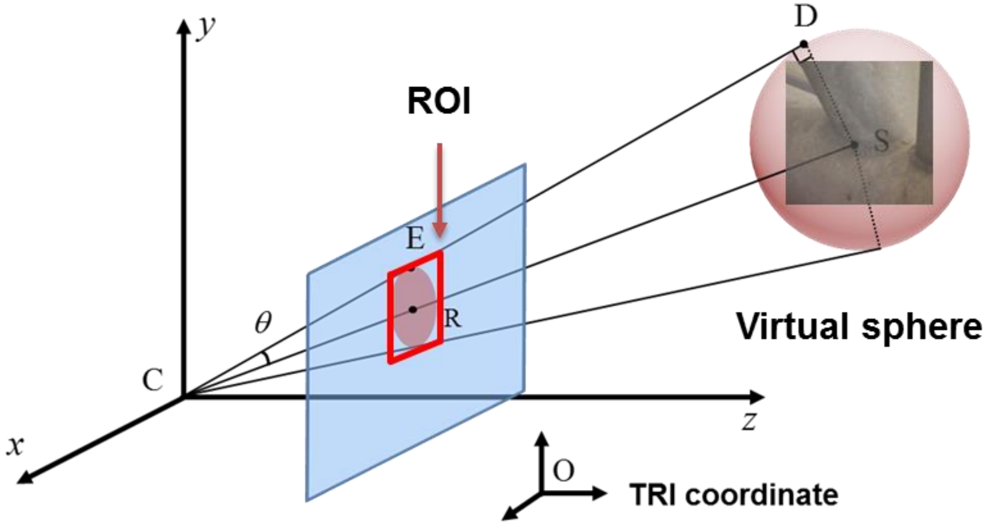
**Source code and data: https://github.com/chulminy**

# Backup Slides

# Backup Slide 1

**Constraints 1:** Bounding boxes should be entirely visible on the image

**Constraints 2:** Bounding boxes should be large enough to obtain useful ROIs

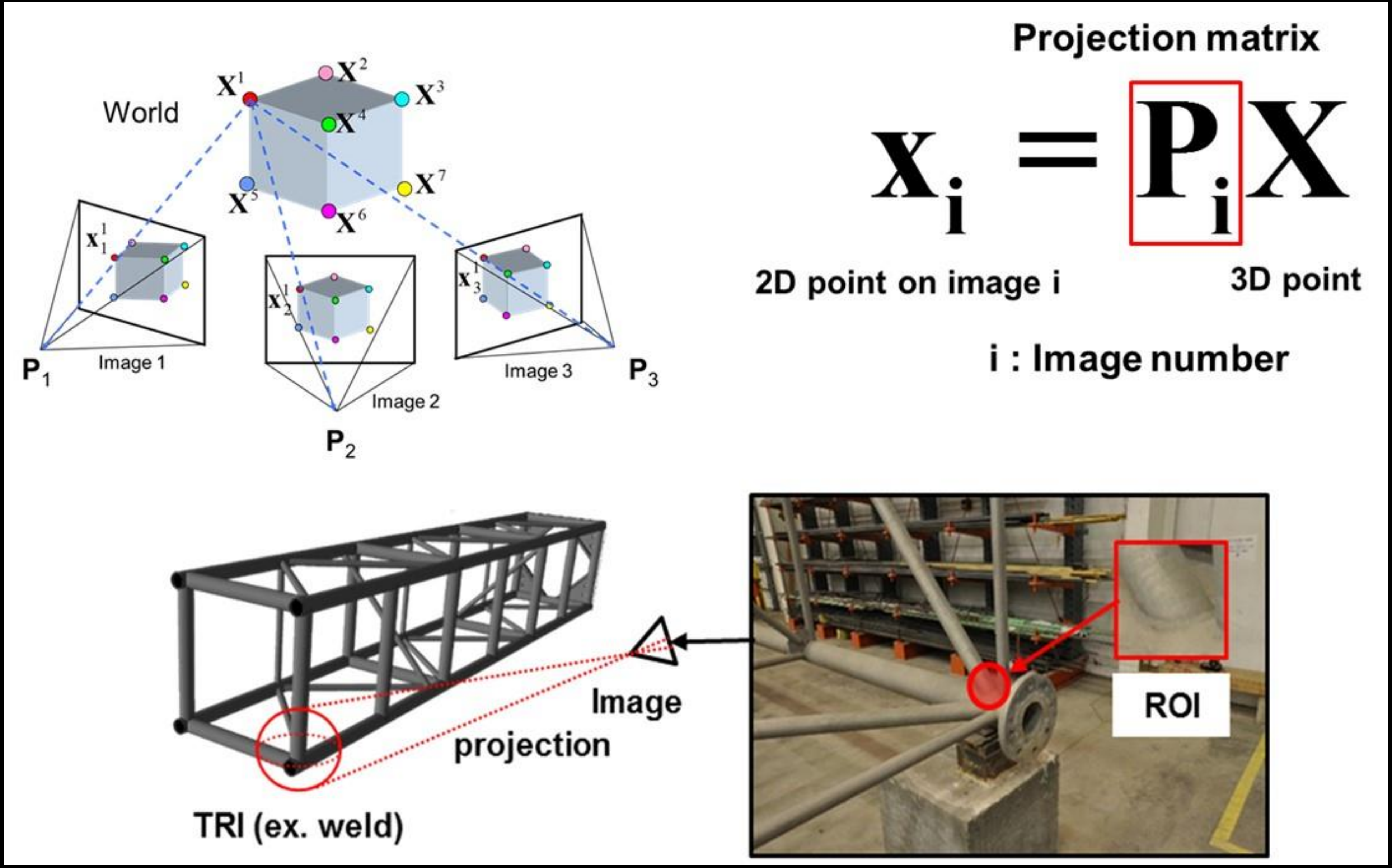


## Backup Slide 1



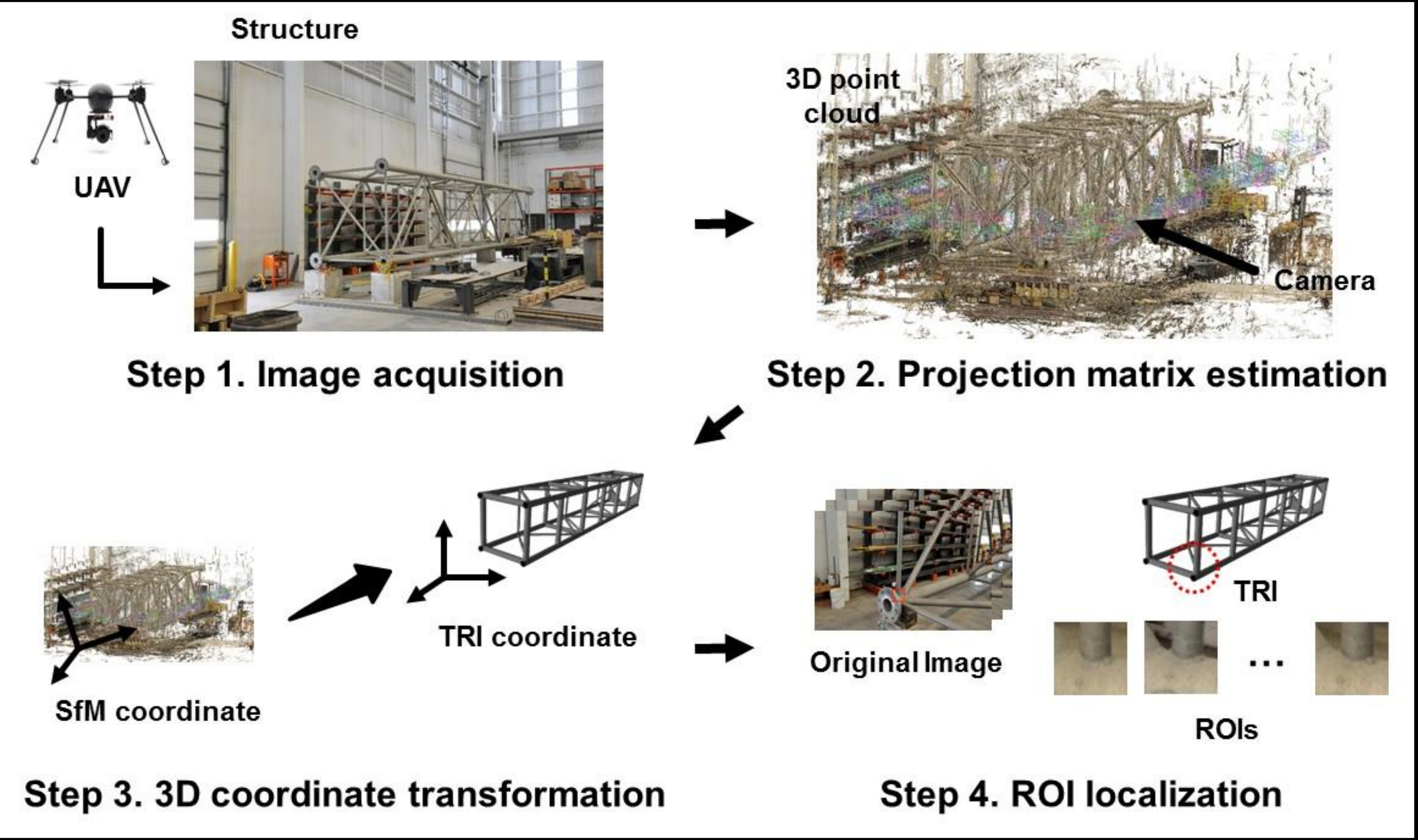
In this study, ROI classification successfully attains a relatively high accuracy. We obtain rates of 89.73% (743/828 images) true-positive (true classification of non-occluded ROIs) and 91.83% (225/245 images) true-negative, respectively. The precision is 97.37%, defined as the number of true-positives over the total number of positives.

# Backup Slide 2



# Backup Slide 3

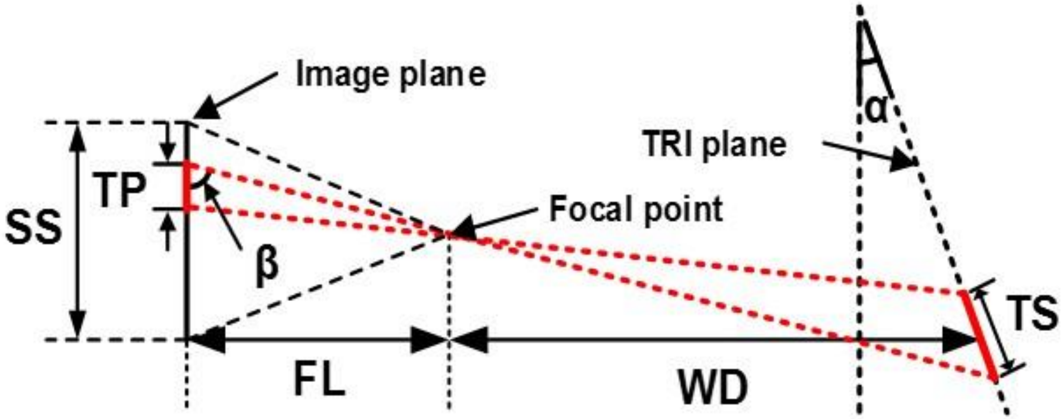




# Backup Slide 4



## 1. Working distance




**Example**

- SR = 4,288 px (Sensor resolution-Width)
- SS = 23.6 mm (Sensor size)
- TS = 63.5 x 2 mm (TRI size – diameter)
- TP = 127 px ( the min. size of the ROIs)
- FL = 18 mm (focal length)
- α = 0 ~ pi/3
- β = 0.92 ~ pi/2

**WD = 2,200 mm**

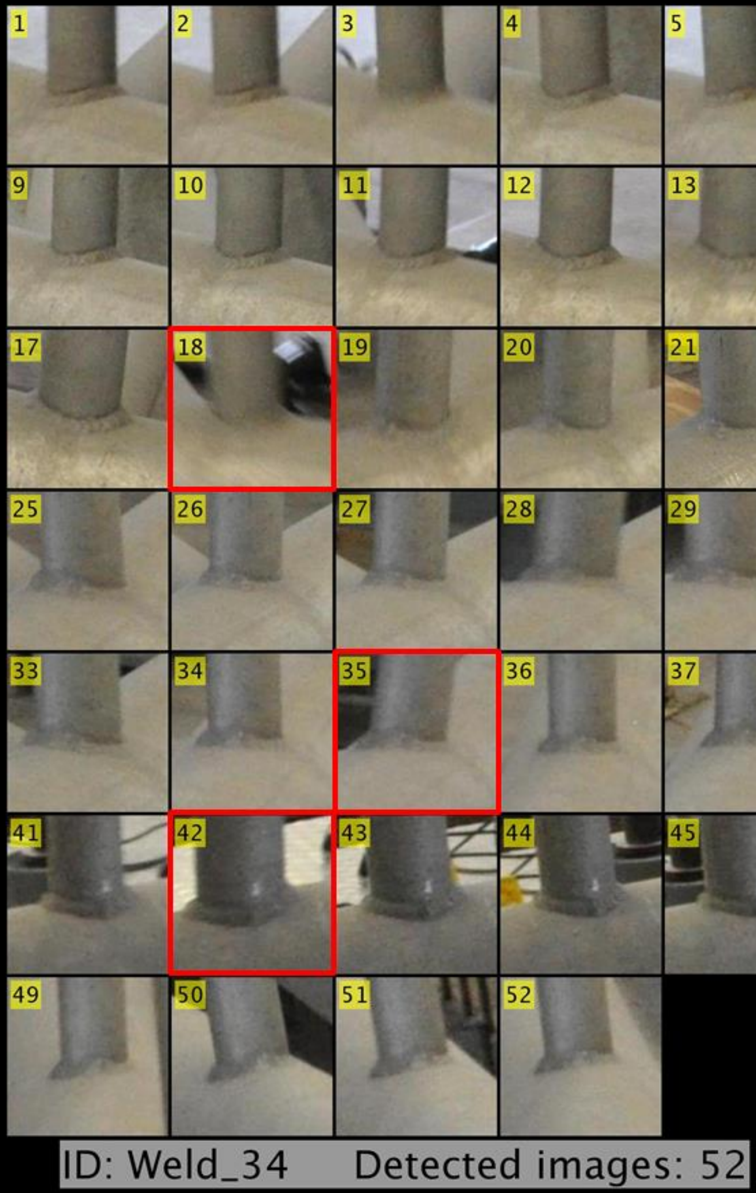
## 2. Motion blur

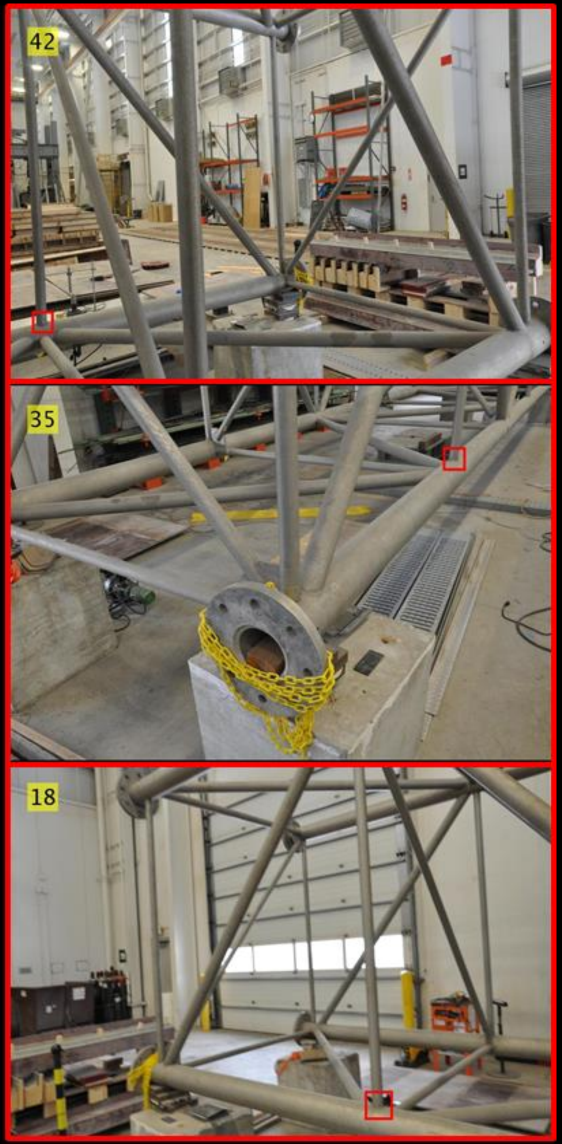
- Flying speed
- Light condition
- Shutter speed
- Vibration on the platform

## 3. Occlusion

# Backup Slide 5



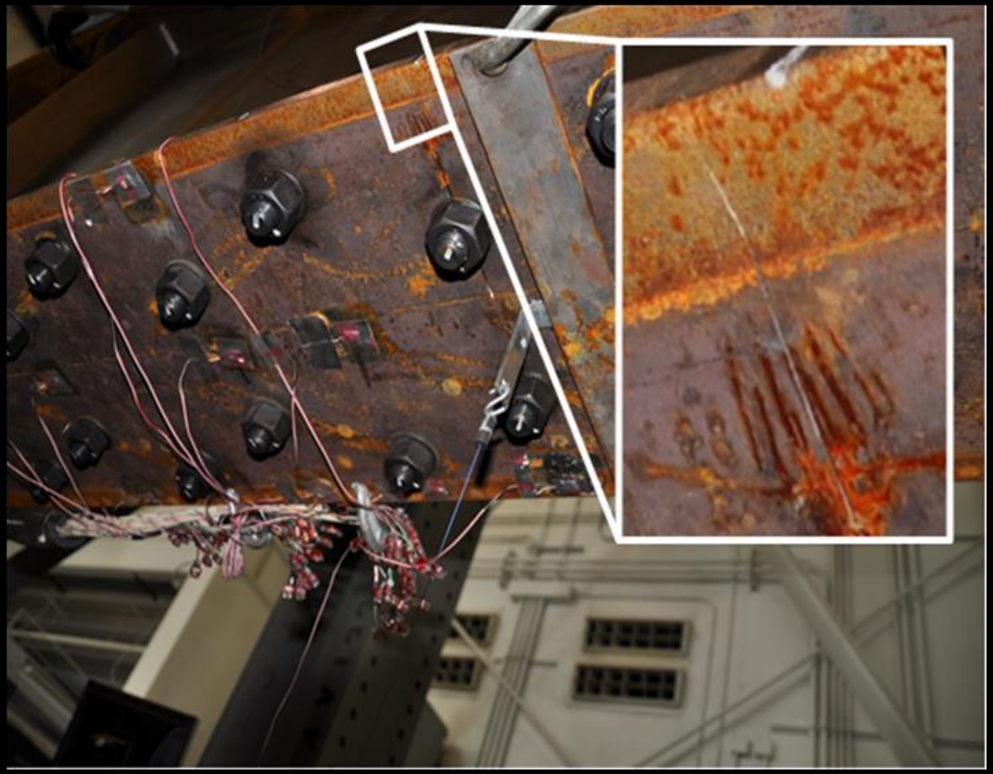


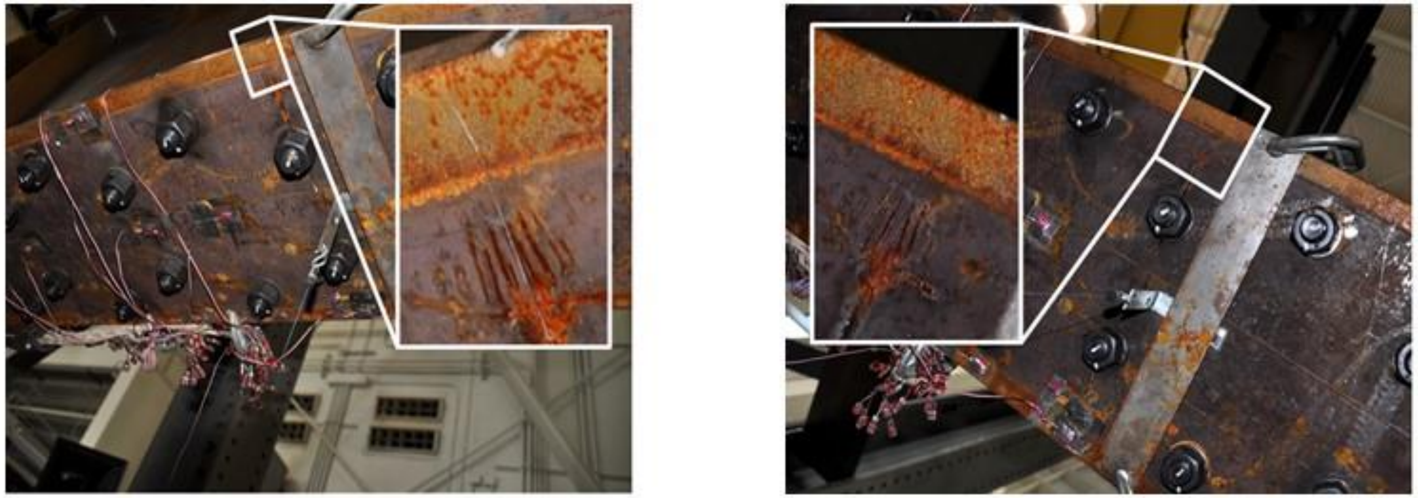
# Backup Slide 6



Fatigue testing on a steel girder (courtesy of Mattew H. Hebdon)



Non-crack area



Images of a fatigue crack from different viewpoints

- Many false-positive alarms and misdetections

  → **Detection of damage-sensitive areas**

- Visibility depending on viewpoints

  → **Use of many different viewpoints of object images**

# Backup Slide 6

We train a single binary classifier that is then applied to all welded connections. This approach is possible because the visual appearances of the welded connections are quite similar to each other, and considerably different than the occluded ones in Fig. 6(b). However, if the appearance of the TRIs were visually dissimilar, and common visual features were not shared with each other, multiple classifiers would need to be trained individually for each type of TRI.