INSPECTING AND PRESERVING INFRASTRUCTURE THROUGH ROBOTIC EXPLORATION

PROJECT WD1

A TRAINING FRAMEWORK OF ROBOTIC OPERATION AND IMAGE ANALYSIS FOR DECISION-MAKING IN BRIDGE INSPECTION AND PRESERVATION

FACULTY: RUWEN QIN[1], GENDA CHEN[1], SUZANNA LONG[1], ZHAOZHENG YIN[2], SUSIL LOUIS[3]
STUDENTS: MUHAMMAD MONJURUL KARIM[1], TIANYI ZHAO[2]
[1] MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY
[2] STONY BROOK UNIVERSITY
[3] UNIVERSITY OF NEVADA, RENO

AUGUST 3-4, 2020



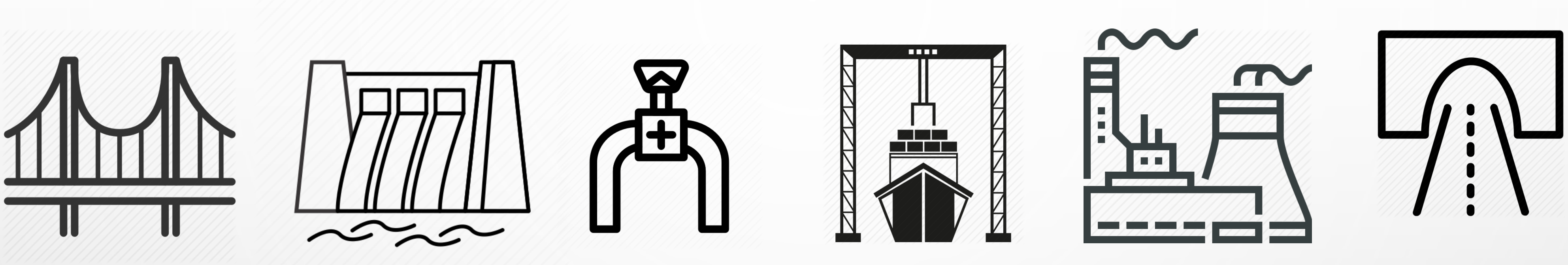
Outline

- Problem Statement
- The Approach
- Results
- Recommendations

Outline

- Problem Statement
- The Approach
- Results
- Recommendations

US Civil Infrastructure Systems



- Operate in harsh environment due to meteorological and electrochemical impacts
- Subject to multiple hazards such as earthquake, flood, hurricane, tornado, and tsunami
- Deteriorate rapidly and are approaching their designed life spans

Bridges

- Over 600,000 highway bridges in the National Bridge Inventory (NBI)
- 39% have exceeded their design life of 50 years
- 9% are structurally deficient and require significant repair
- National Bridge Inspection Standards require bridges to be inspected every two years to ensure that there are no cracks, rusting, or other damage
- Over hundreds of bridges need to be inspected every day

Problems of The Current Practice of Bridge Inspection

blocking traffic

data collection is time-consuming

difficulty of accessing elevated bridge structures

dangerous field activities

Bias of human visual inspection

Robotic Technology for Bridge Inspection

- To address the cost, accessibility, safety, and reliability concerns of the current inspection practice
- To move forward to data-driven asset management
- Robotic platforms equipped with sensors

BIRDS – Missouri S&T, Dr. Genda Chen

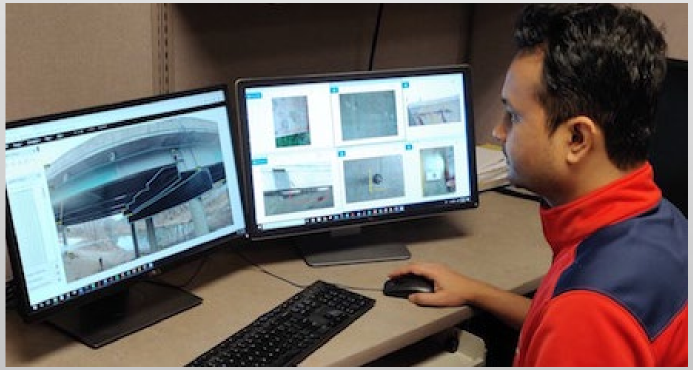
Climbing robot – University of Nevada, Reno, Dr. Hung La

Future Work at the Human-Technology Frontier

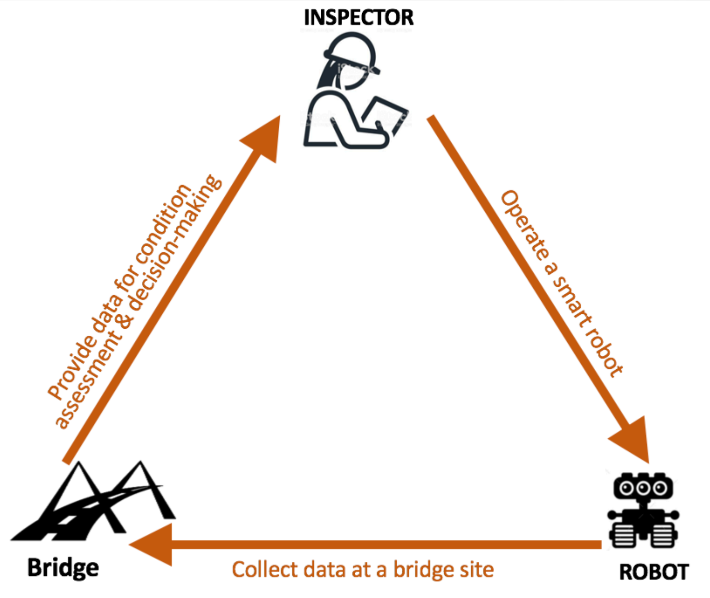
- Current practice – Inspectors visually check the bridge conditions at the site


https://studopedia.org/1-87707.html

- Future work – cooperative robot-inspector survey of bridges

- A paradigm shift of bridge inspection due to the advancement of new technologies

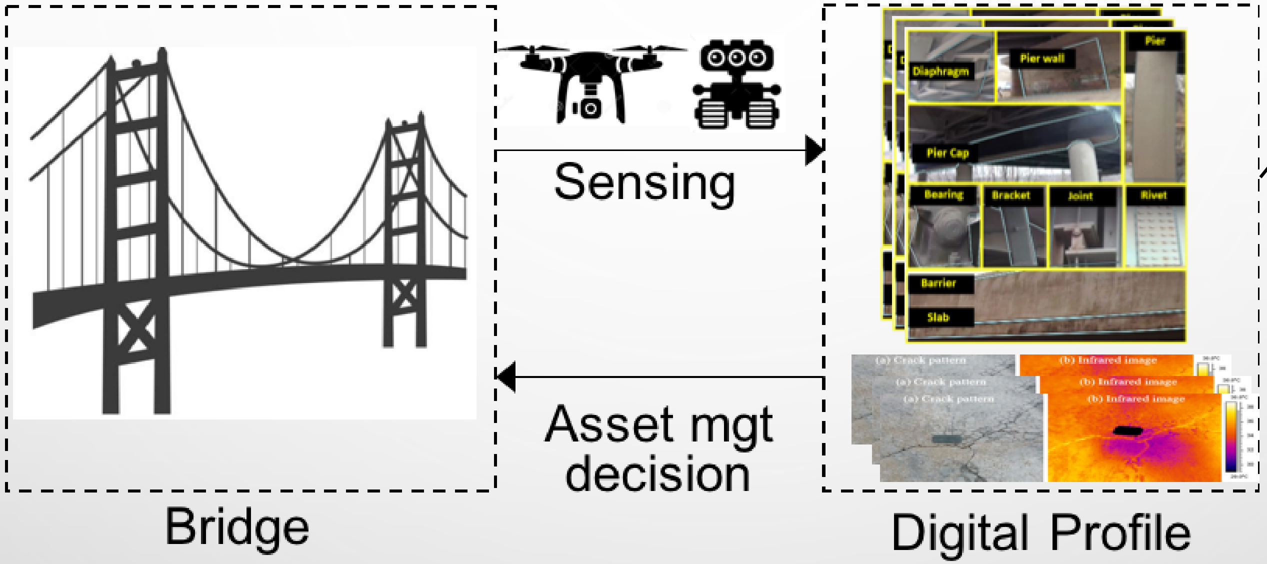


Ready for the Cooperative Robot-Inspector Survey Regime?

- Possible problems
  - Do robots “know” how to efficiently collect high quality data of bridges?
  - Have inspectors developed the skills of operating robots efficiently, safely, and with comfort?
  - Are inspectors efficient in analyzing the rapidly collected big complex inspection image data?
  - Is it an issue to let the robot stand by for a long time until the inspector finishes reviewing the data to move forward to the next step?
  - Is a data analytic tool developed for one bridge applicable to another bridge?
  - …
- Needs
  - <u>Robots</u>: maximize the utilization and efficiency of robots in assisting inspectors to collect data given the operational constraints.
  - <u>Workers</u>: make inspectors safe, comfortable, and efficient in operating robot and analyzing the big complex image data.
  - <u>Bridges</u>: create an accurate digital profile of every bridge, that can be updated over time, to support data-driven asset management.

A Cyber-Physical System for Asset Management

- Data-driven asset management for bridges requires evaluating defects at the structural element level



Inspection video data analytics for creating the digital profile of bridges

- Robots capture inspection video data with bridge elements mixed with cluttered background
- Digitally profile bridge conditions as a cyber system for managing the physical system (the bridge)
- Assessment results provide the decision support for preservation

Inspection Video Data Analytics

Inspection video data analytics for creating a digital profile of the bridge

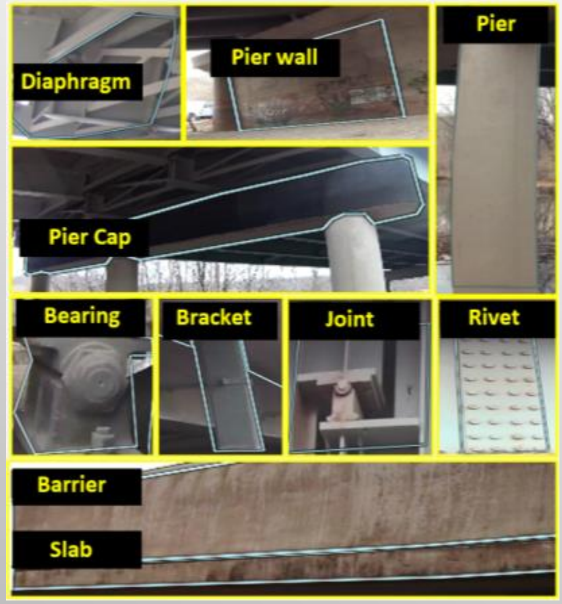
- Detect and segment bridge elements from inspection video data and sort them out

- Detect defects on each element type and accordingly classify defects for assessing the element condition

excellent good fair poor

Challenges in Inspection Video Data Analytics

- Big data collected in fast speed. A standard RGB camera collects 108K images per hour
- Complex: multiclass bridge elements mixed with cluttered background
- Variations: elements of 600,000 bridges exhibit large variations in videos



Bridge Elements
1) Barrier
2) Slab
3) Bearing
4) Pier
5) Pier cap
6) Pier wall
7) Rivet
8) Truss
9) Bracket
10) Joint

- Viewpoint change

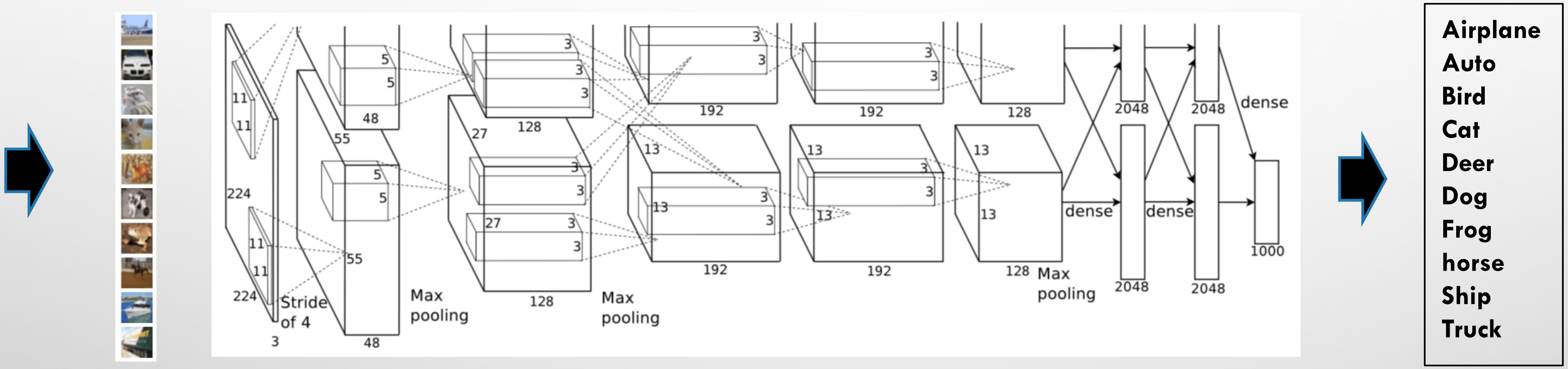
- Scale variation

- Lack of illumination

- Camera vibration
- …

Does Deep Learning Solve the Challenges?

- Deep learning
  - Powerful computational resources + large-scale annotated data
  - DL models such as convolutional Neural Networks (CNN) improve image analysis such as object detection and classification



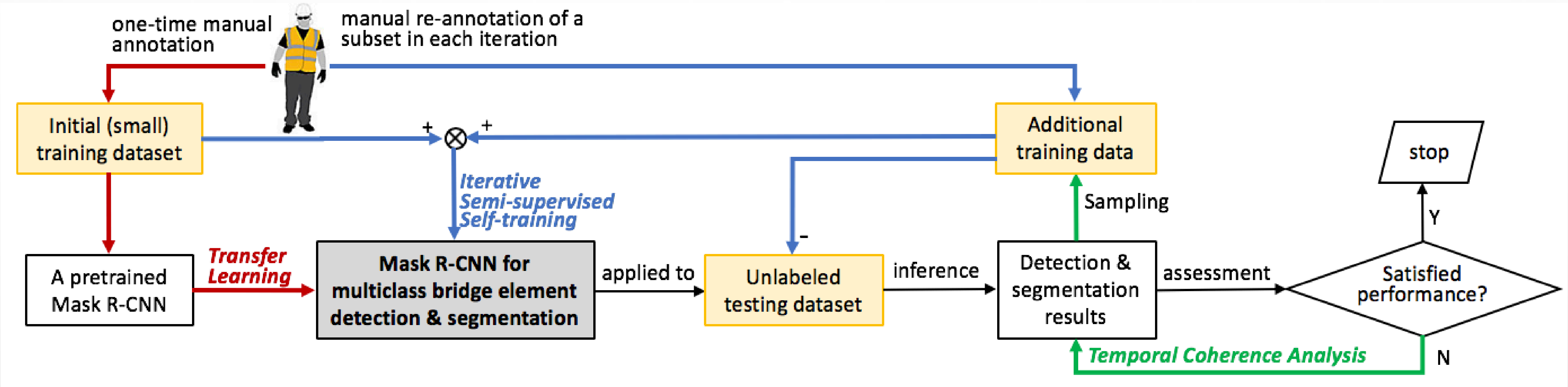
AlexNet in NIPS2012

- Heavily rely on intensive, tedious efforts of human to annotate training data
- Large-scale annotated datasets do not cover bridge elements or bridge conditions
  - ImageNet: 12.4 millions of labeled training samples
  - Microsoft COCO: 10 millions of labeled training samples
- High-quality labeled training data for domain-specific applications is difficult to acquire

Outline

- Problem Statement
- The Approach
- Results
- Recommendations

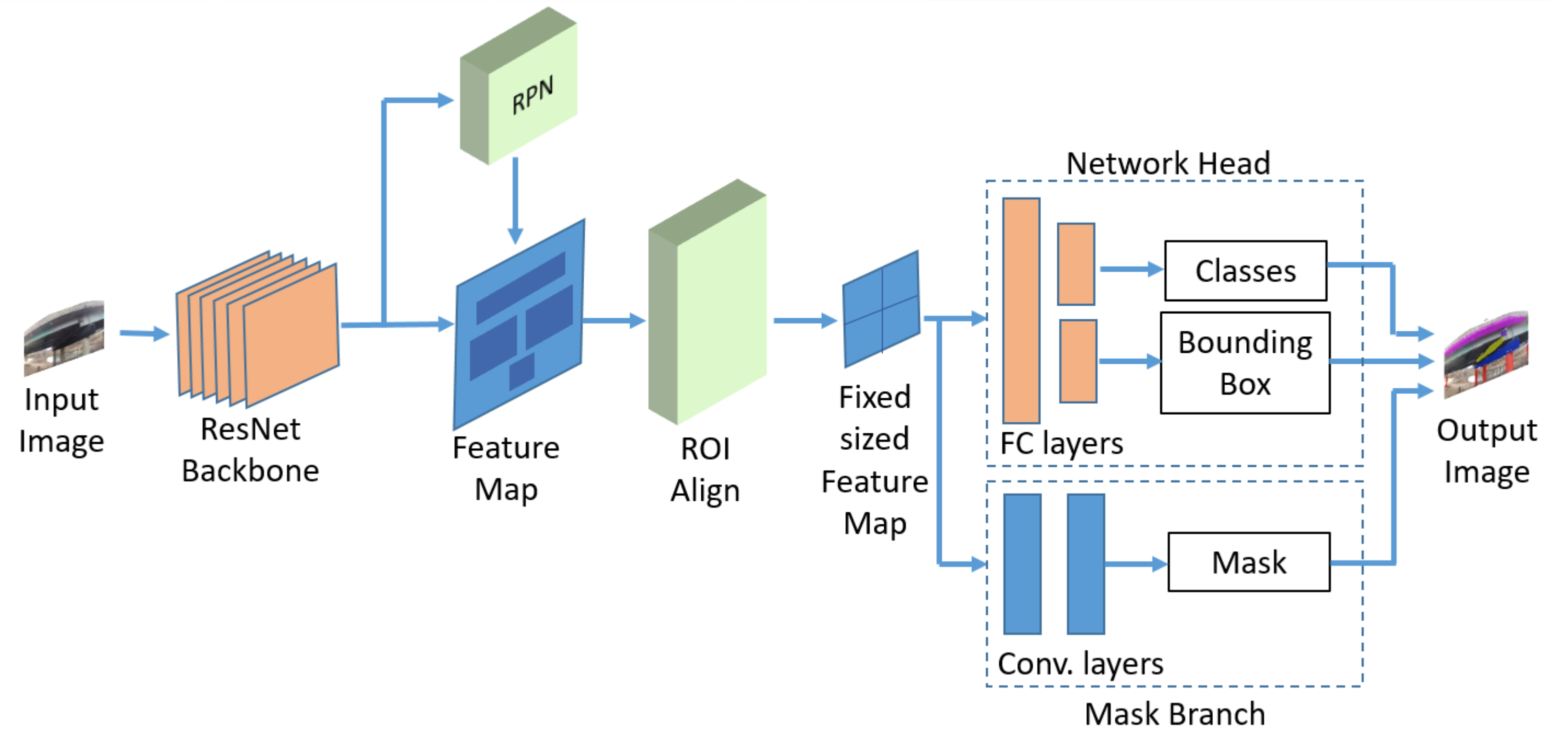
Engaging Inspectors to Develop the Video Data Analytics AI Tool



1. Transfer learning for initial adaption
2. Temporal coherence analysis for correcting false negative detections and selecting additional training data
3. Iterative semi-supervised self-training (IS^{3}T) to boost the performance

Transfer Learning

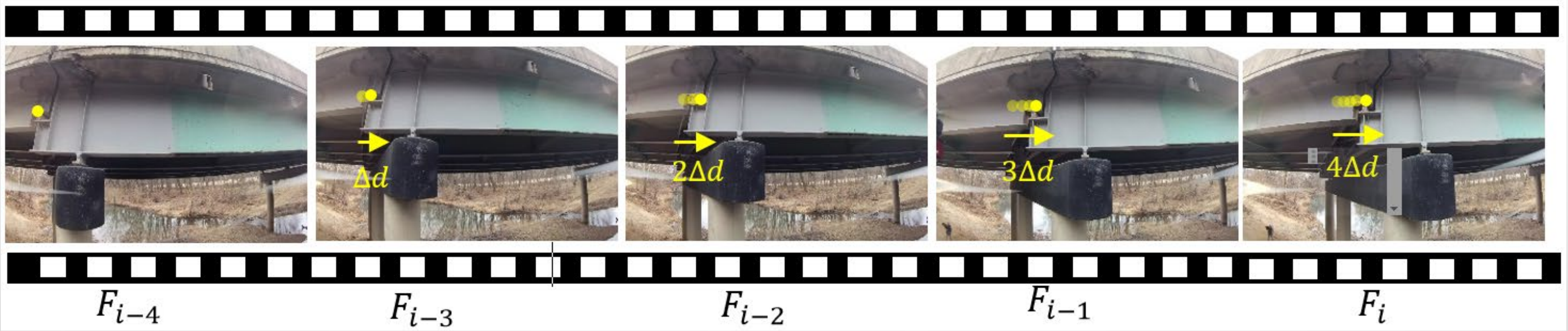
- A Mask R-CNN for segmenting multiclass bridge elements from video data
- Training this DL tool from the scratch requires a large volume of annotated data to achieve satisfied prediction accuracy



- A Mask R-CNN, pre-trained on Microsoft COCO, is transferred in for the task of multiclass bridge element detection and segmentation
- A small initial training dataset (T_0) annotated by inspectors: 40 images containing 482 labeled objects in 10 classes

Temporal Coherence Analysis

- An object in a frame is highly likely present in the neighboring frames within a range of displacement with similar confidence
- Example:

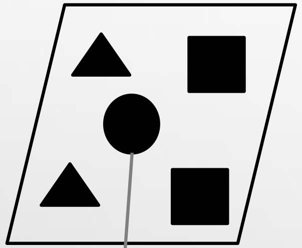


- Temporal coherence information: objects with high detection scores in preceding frames and their spatial locations
- Implemented in the post phase of Mask R-CNN
- Two-staged threshold: a detection with a score $S \in [t_l, t_u)$ indicates it may be a false negative detection

Temporal Coherence Analysis

Frame $i-(q+1)$

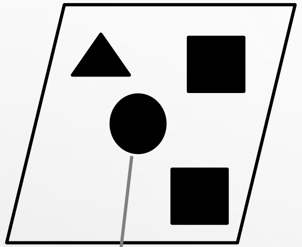
Set of detection $O_{i-(q+1)}$



A 'dot' is found within the distance $(q+1)\Delta d$ from $C_{i,j}$ with detection score $S_{i-(q+1),j'}$ and at the location $C_{i-(q+1),j'}$.

Frame $i-q$

Set of detection O_{i-q}



A 'dot' is found within the distance $q\Delta d$ from $C_{i,j}$ with detection score $S_{i-q,j''}$ and at the location $C_{i-q,j''}$.

$$S_{i,j} = \frac{1}{2}\left(S_{i-(q+1),j'} + S_{i-q,j''}\right)$$

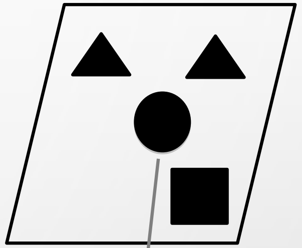
…

Search the same detection 'dot' within the surrounding area in a pair of preceding successive frames up to N times from the nearest pair to the farthest pair (i.e., $q = 1, \ldots, N-1$)

Recover this "false negative" detection by updating the detection score

Frame i

Set of detection O_i



The jth object is classified as 'dot' with detection score $S_{i,j}$ and at the location $C_{i,j}$.

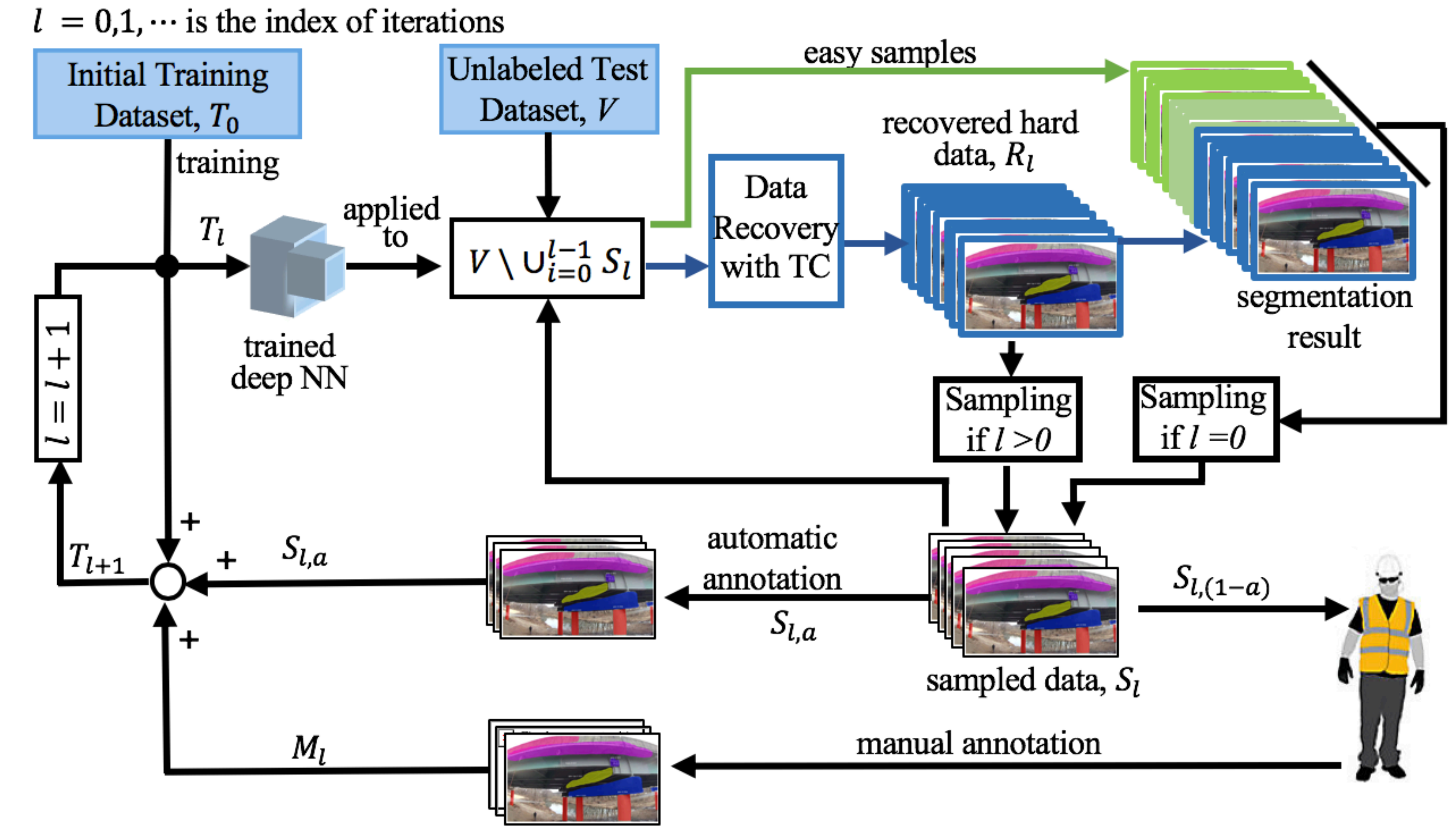
The detection score is lower than the upper threshold t_u of detection but greater than the lower threshold t_l, thus it may be a false negative detection.

Iterative Self-supervised Self-Training (IS^{3}T)

- Combine a few inspector-annotated high quality data with a set of data automatically annotated by the trained network to supplement the existing training dataset to retrain the network

$$T_{l+1} = T_l + M_l + S_{l,\alpha}$$

- The small dataset labeled by an experienced inspector is sampled from recovered hard data and thus can effectively boost the performance of the network



Outline

- Problem Statement
- The Approach
- Results
- Recommendations

The Solution Process

- After the Mask R-CNN is initially transferred, the IS^{3}T algorithm took three iterations to complete the refining process.
- Took 4.82 hours in total (3.58 hours for data annotation and 1.24 hours for training) to develop the AI tool.

	Transfer learning	Iterative S^{3}T		
Index of iteration, l	0	1	2	3
Training dataset (# images), T_l	40	48	85	118
Hard recovered data (# images), R_l	126	113	79	50
Incremental of training data (# images), S_l	8	37	33	
Automatically annotated data (# images), $S_{l,\alpha}$	0	26	26	
Manually annotated data (# images), M_l	8	11	7	

Testing Results

- Test dataset has 212 images of bridge A, which include 1,872 objects from the 10 classes
- Calculate precision, recall, and f1-score

$$\text{Precision} = \frac{\#\ \text{correct predictions}}{\#\ \text{predictions}};\ \text{Recall} = \frac{\#\ \text{correct predictions}}{\#\ \text{ground truth objects}};\ \text{f1} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$$

		Transfer learning	Iterative S^3T		
Index of iteration, l		0	1	2	3
IoU = 0.5	Precision (%)	80.3	81.7	90.7	91.8
	Recall (%)	74.4	90.3	90.1	93.6
	f1-Score (%)	77.2	85.8	90.4	92.7

- Transfer learning achieved 80.3% precision, 74.4% recall, and 77.2% f1
- IS^3T: ↑ 11.5% recall, ↑ 19.4% precision, and ↑ 15.5% f1, respectively

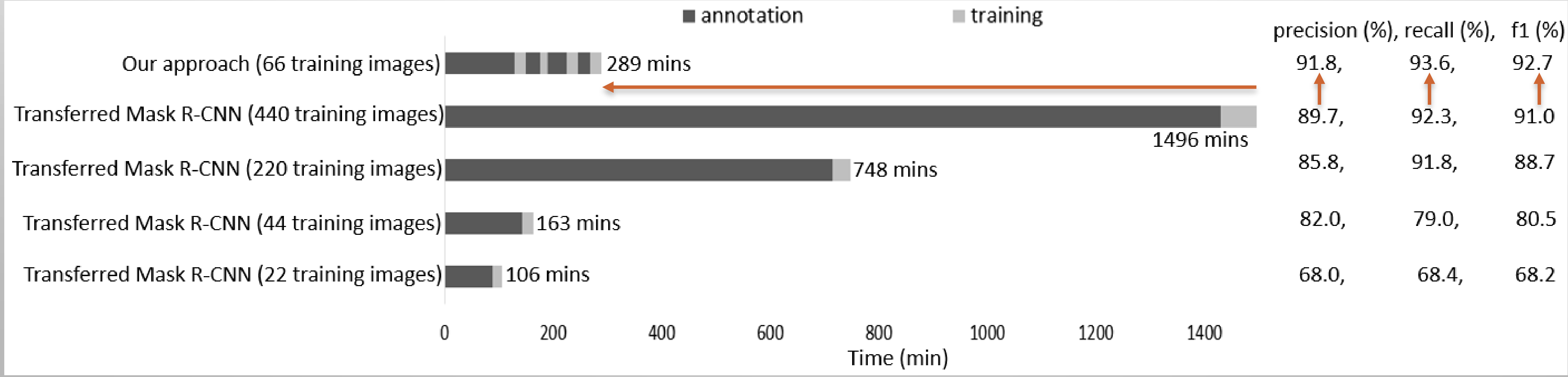
How Much Does Transfer Learning Help?

- Training from the scratch with a training dataset of 144 images
  - the training process has not been finished after 13.2 hours (600 epochs)
  - poor performance (32.3% precision, 18.3% recall, and 23.4% F1)
- Transfer learning with the initial training dataset of 40 images
  - the training process took only 20 minutes
  - reasonable performance (80.3%, 74.4%, and 77.2%)

Method	Training time (min)	Precision (%)	Recall (%)	f1 (%)
Training from scratch	792+	32.3	18.3	23.4
Transfer learning	20	80.3	74.4	77.2

Cost-Effectiveness Achieved by Engaging Inspectors

- Direct transfer learning
  - uses blind annotation. The model performance will increase at a rapidly increased cost of data annotation. But how many to annotate?
  - a performance of 89.7% precision, 92.3% recall, and 91.0% f1 took a large amount of tedious human annotation efforts (about 24 hours) plus 1.1 hours of training time to achieve.
- Proposed method
  - reduced the annotation time by 85% and took a comparable amount of time (only 6 more minutes) to train the model.
  - achieved a better performance (↑2.1% precision, ↑ 1.3% recall, ↑ 0.3% f1)



Helpfulness of the AI Tool for Inspection Image Data Analytics

- A small-scale example of detecting and segmenting bridge elements from 20 images

	Work time (min)	Accuracy (%)
w/o the developed AI tool	65	100
w/ the developed AI tool	0.27	93.7
change	↓99.5%	↓6.3%

- The time saving is tremendous, increasing in proportion to the size of real world task
- The gap of accuracy is diminishing as the job size increases due to human factors related issues (e.g., loss of attention, fatigue, etc.)

Outline

- Problem Statement
- The Approach
- Results
- Recommendations

Recommendations

1. No deep learning models are directly applicable to any tasks and, therefore, it is important to adapt models to different new tasks to achieve satisfied performance. Transfer learning and iterative boosting are useful methods that let the algorithms learn new features of a task from a small amount of data of the task, thus adapting to that task.
2. Keep-human-in-the loop is an important method to leverage human intelligence into the artificial intelligence algorithms. This should be achieved through a collaborative approach. That is, algorithms provide humans with its performance so that humans can figure out the weakness and provide inputs (e.g., additional training dataset annotated by humans) to the algorithms for improvement.
3. Model adaptability and the collaboration between artificial intelligence and human experts were integrated together as a solution for developing assistive intelligence that takes care of time-consuming, boring tasks and let humans focus on knowledge-intensive tasks. This will be a new style of work for future bridge professionals.

Thank you!

Q&A

qinr@mst.edu