TRANSPORTATION POOLED FUND PROGRAM
QUARTERLY PROGRESS REPORT

Lead Agency (FHWA or State DOT): Kansas DOT

INSTRUCTIONS:
Lead Agency contacts should complete a quarterly progress report for each calendar quarter during which the projects are active. Please provide a project schedule status of the research activities tied to each task that is defined in the proposal; a percentage completion of each task; a concise discussion (2 or 3 sentences) of the current status, including accomplishments and problems encountered, if any. List all tasks, even if no work was done during this period.

Transportation Pooled Fund Program Project # *(i.e, SPR-2(XXX), SPR-3(XXX) or TPF-5(XXX)* **TPF-5(535)**	**Transportation Pooled Fund Program - Report Period:** ☐Quarter 1 (January 1 – March 31) ☐Quarter 2 (April 1 – June 30) ☐Quarter 3 (July 1 – September 30) ☒Quarter 4 (October 1 – December 31)	
TPF Study Number and Title: **TPF-5(535): Human-centered Steel Bridge Inspection enabled by Augmented Reality and Artificial Intelligence**		
Lead Agency Contact: **David Behzadpour**	**Lead Agency Phone Number:** **785-291-3847**	**Lead Agency E-Mail** David.Behzadpour@ks.gov
Lead Agency Project ID: Click or tap here to enter text.	**Other Project ID (i.e., contract #):** Click or tap here to enter text.	**Project Start Date:** 10/1/2024
Original Project Start Date: **10/1/2024**	**Original Project End Date:** **9/30/2027**	**If Extension has been requested, updated project End Date:** Click or tap to enter a date.

Project schedule status:

☒On schedule	☐On revised schedule	☐Ahead of schedule	☐ Behind schedule

Overall Project Statistics:

Total Project Budget	Total Funds Expended This Quarter	Percentage of Work Completed to Date
$600,000	0	2%

Project Description:

The main objective of this proposed research is to provide state DOTs practical tools for supporting human-centered steel bridge inspection with real-time defect (e.g., fatigue cracks and corrosion) detection, documentation, tracking, and decision making. The proposed research will not only bridge the gaps identified in the IDEA project, but also expand the existing capability by developing AI algorithms for crack and corrosion detection. In addition to AR headsets, the project will also develop AR-based inspection capability using tablet devices. The tablet device can be used to perform AR-based inspection directly in a similar way to the AR headset. It can also leverage Unmanned Aerial Vehicles (UAV) for remote image and video acquisition during inspections, enabling bridge inspections from a distance in a human-centered manner.

Progress this Quarter
(includes meetings, work plan status, contract status, significant progress, etc.):

1. **Project Kickoff Meeting**
   Project kickoff meeting was held on Oct. 14, 2024, via teleconference. Participants include Jian Li (U Kansas) – PI
   Caroline Bennett (U Kansas) – Co-PI, William Collins (U Kansas) – Co-PI, Luke Attard (U Kansas) – Graduate Student
   David Behzadpour (KDOT) – Technology Transfer Engineer, Sally Mayer (KDOT) –Assistant Bureau Chief, Research
   Mark Hurt (KDOT) – Bureau Chief, Structures and Geotech, Dan Wadley (KDOT) – Bureau Chief, Research
   David Snoke (NCDOT) – Bridge Inspection Program Manager, Shawn Hart (Caltrans) – Senior Bridge Engineer
   Justin Wilson (TxDOT) – Transportation Engineer Supervisor. The meeting minutes are attached.

2. **Task 1: CV and AI algorithms for crack and corrosion inspection**

   Deep learning models have been trained on a publicly available dataset for fatigue crack detection of steel bridges. This dataset was released by the 1st internation competition on structural health monitoring and contains 200 high-resolution images of steel fatigue cracks on welded joints. Two deep learning architectures are considered including YOLOv11 and DeepLabV3+. Preliminary results show that YOLOv11 model performed better than DeepLabV3+.

3. **Task 2: AR-based software for human-centered bridge inspection**
   Development has begun on a new HoloLens 2 app for infrastructure inspection using the latest version of Microsoft's Mixed Reality Tool Kit (MRTK) and Unity version 6. The application front end has been developed, and underlying systems are being created and tested. The use of Unity Sentis has been employed for machine learning (ML) inferencing to be done locally on the HoloLens without the need for a secondary device on a local network to make inferences and allowing inspections to be more efficient and require less user setup. Graduate students have collaborated to work out issues with the inference process producing incorrect outputs that do not resemble their input images. Functionality has been developed to utilize the HoloLens camera to take images for ML inferencing, and work has begun on anchoring inferences to their real life positions through the use of unity's projector components, which allow a 2D image (which is in this case the ML inference) to be projected on a 3D mesh, such as the spatial mesh generated by the HoloLens in real time based off of its surroundings. The software package will require holograms to maintain spatial persistence across sessions so that data gathered inspecting infrastructure can be reused when it is next inspected. Microsoft World Locking Tools have been identified as a promising solution to this problem and implementation of these tools has begun.

Anticipated work next quarter:

1. **Task 1: CV and AI algorithms for crack and corrosion inspection**
   A more comprehensive database for crack detection will be established which includes more diverse steel bridge images. The deep learning models will be further refined for enhanced performance.

2. **Task 2: AR-based software for human-centered bridge inspection**
   Over the next reporting period we plan to achieve improved hologram anchorage and persistence as well as native ML processing and long-term data collection and storage. We plan to refine our approach to image anchoring using Unity's

projector components and properly integrate World Locking Tools in order to allow anchorage to persist across sessions. ML processing solutions such as YoloSharp have been identified and will be used to achieve native ML processing.

Significant Results:

1. **Task 1: CV and AI algorithms for crack and corrosion inspection**

The dataset used is the IPC-SHM 2020 dataset, which contains 200 high-resolution images of steel fatigue cracks on welded joints, with a resolution of 4928x3264 pixels. Among these, 120 images are annotated with binary labels, where cracks are marked with a label of 1 and non-crack regions are labeled as 0. These 120 annotated images were used as the training set, while the remaining 80 images served as the test set.

To improve the model's ability to generalize and handle a variety of scenarios, data augmentation was applied to the training images. Instead of performing standard augmentation on individual images, a custom script was used to apply several transformations iteratively. The process begins by passing each original image through a for-loop 7 times, with each iteration applying a random transformation to the image. A random function determines which augmentation operation—such as horizontal flip, rotation, brightness adjustment, or contrast adjustment—should be applied, depending on whether the output of the random function exceeds a predefined threshold. This random selection ensures that each image undergoes different combinations of transformations, creating a more diverse set of images for training. After this process, the training dataset is effectively 7 times larger than the original, with varied versions of each image, exposing the model to different orientations, lighting conditions, and shapes of cracks.

Once the augmentation process is completed, the next step is the CREC cropping strategy, which was inspired by Hsu et al. (2022). The CREC cropping method is designed to increase the crack-to-background ratio in each image. After the augmentation, each image is cropped to a bounding box that tightly encapsulates the crack region, ensuring that only the crack and a small surrounding area are retained. This cropping process helps remove unnecessary background information, which could distract the model from learning the important features of the crack. The cropped images are then resized to meet the input requirements of the deep learning models and are normalized as per standard practices to ensure consistent training across all images.

Fig. 1 Image Preprocessing (Original->Horizontal Flip->Contrast Adjustment->Random Rotation->CREC)

The result of this preprocessing pipeline is a dataset that is more focused on the cracks themselves, with a significant reduction in background noise. To visually illustrate this, Figure 1 shows the progression of an image through each preprocessing step. Starting with the original image, it is then augmented through various transformations such as horizontal flips and rotations, followed by brightness and contrast adjustments. Finally, the image undergoes the CREC cropping step, where the crack is isolated from the surrounding background, providing the model with a clearer view of the crack region. This ensures that the model is trained on images that emphasize the features most relevant for crack detection. Through this comprehensive preprocessing approach, which combines augmentation and targeted cropping, the dataset is enriched and focused on the critical crack regions. The increased diversity in the images, coupled with the reduced background noise, helps improve the model’s ability to detect cracks accurately and generalize to different scenarios in real-world SHM applications.

Two advanced deep learning models—DeepLabv3+ and YOLOv11L-Seg—were implemented for steel fatigue crack recognition. Both models were chosen for their effectiveness in segmentation tasks, and their performance was evaluated using the mean Intersection-over-Union (mIoU) metric. The training losses of the two models are shown in Figure 2.

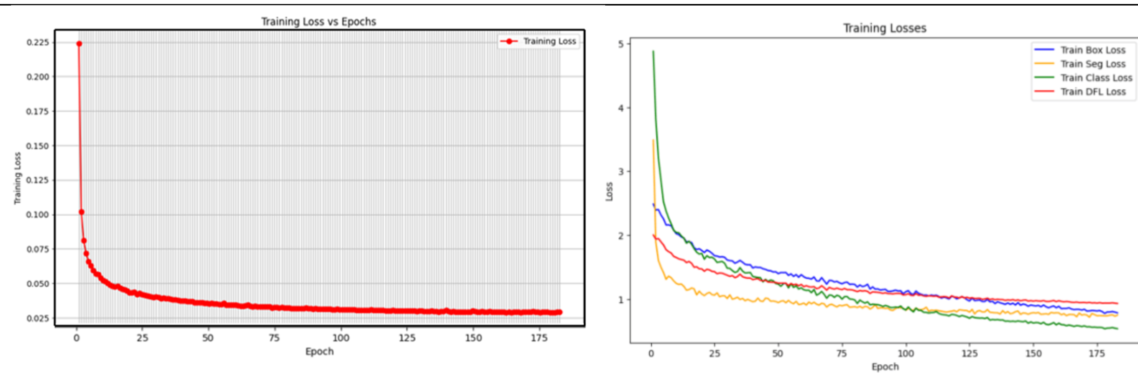

Fig. 2 Training loss for DeepLabv3+ (left) and YOLOv11(right)

To further refine the segmentation results and enhance model performance, post-processing techniques were applied to the predicted masks generated by the models. These steps were essential for addressing challenges posed by fine, fragmented crack structures and noisy backgrounds, common in steel fatigue crack detection tasks. The post-processing involved the application of morphological operations, specifically:

- Erosion: This operation was used to remove small, isolated false positives that often appear along the boundaries of predicted crack regions. By shrinking the predicted mask slightly, erosion refined the crack boundaries, ensuring cleaner and more accurate segmentation outputs.
- Dilation: Following erosion, dilation was applied to enhance the connectivity of true positive regions. This operation expanded the predicted mask slightly, filling in gaps and reconnecting fragmented areas within the crack predictions. The result was a more complete and cohesive representation of the crack structures.

To achieve optimal results, these operations were performed using a kernel size of 3, with two iterations of dilation and one iteration of erosion. This sequence ensured a balance between removing noise and preserving the integrity of the crack predictions.

Sample fatigue crack segmentation results are shown in Figure 3 for DeepLabV3+ and Figure 4 for YOLOv11. As can be seen, postprocessing improved the IoU for both models, and YOLOv11 achieved better performance than DeepLabV3+.

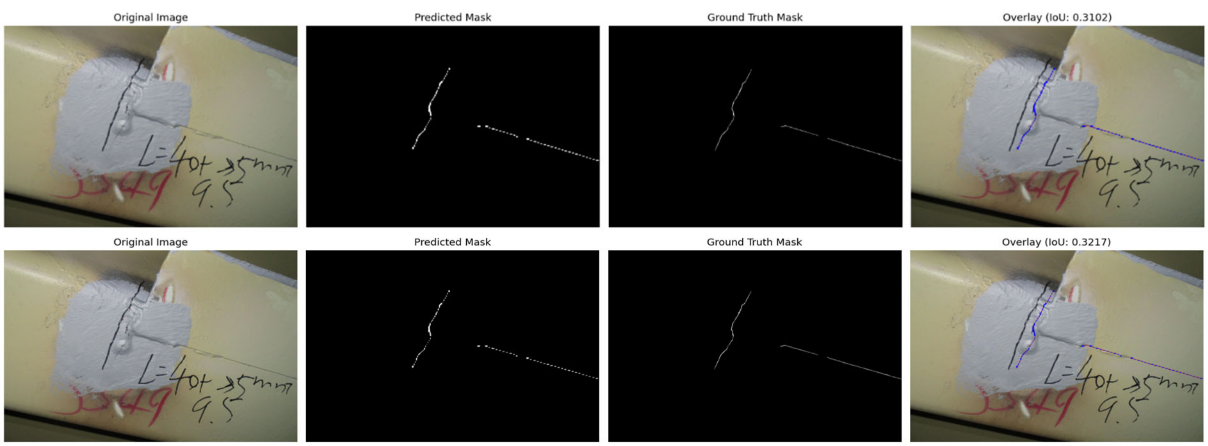

Fig. 3 Sample crack detection result with DeepLabV3+ (Top row: before postprocessing, bottom row: after postprocessing)

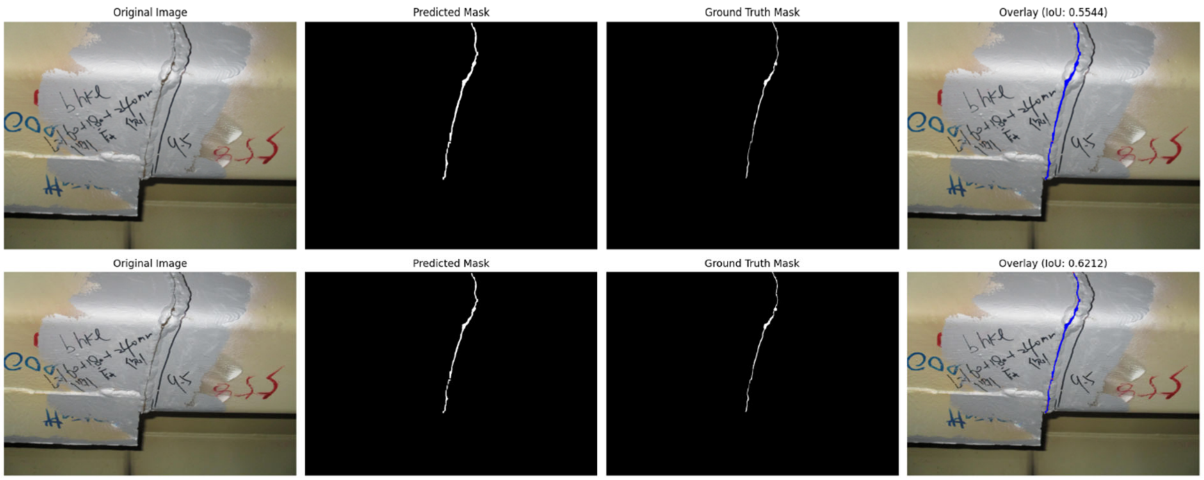

Fig. 4 Sample crack detection result with YOLOv11 (Top row: before postprocessing, bottom row: after postprocessing)

2. **Task 2: AR-based software for human-centered bridge inspection**
   The user interface for our AR platform on HoloLens has been developed. Figure 5 shows the application user interface. This interface will allow inspectors to interact seamlessly with the inspection tool through the headset. The featured buttons will allow the user to capture images for ML processing and storing, view the spatial mesh generated by the HoloLens, customize the view of the processed images as well as exit the program. Figure 6 features the debugging tool developed to allow for more efficient development, allowing error codes to be viewed from within the HoloLens as features are added.

Figure 5. HoloLens user interface.

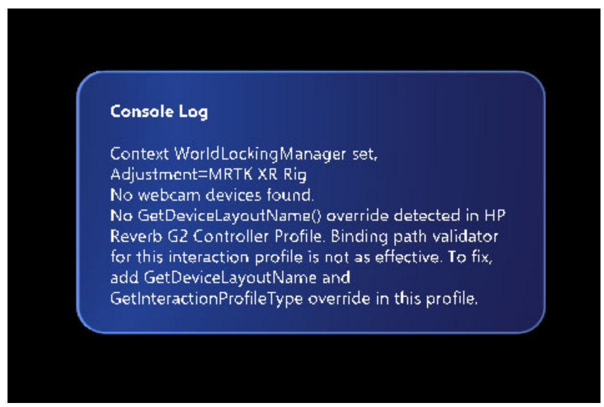

Figure 6. HoloLens development debug tool.

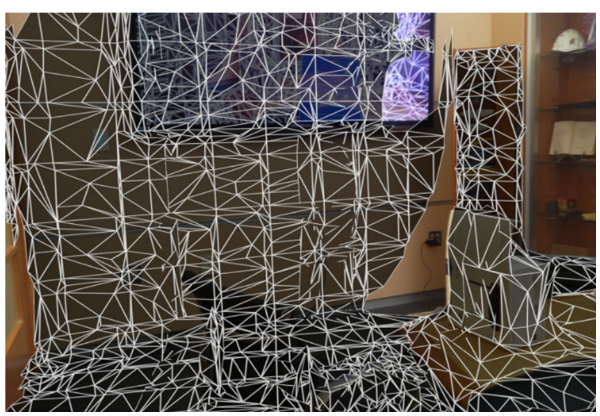
Figure 7. HoloLens generated spatial mesh.

HoloLens offers the feature of spatial mesh generation for apps in which it is necessary for virtual components of the app to interact with boundaries and objects present and around the user. This tool has been enabled as part of our software package and will be used in the process of anchoring processed images to their respective physical positions. Figure 7 shows the real-time spatial mesh generation enabled in our HoloLens application.

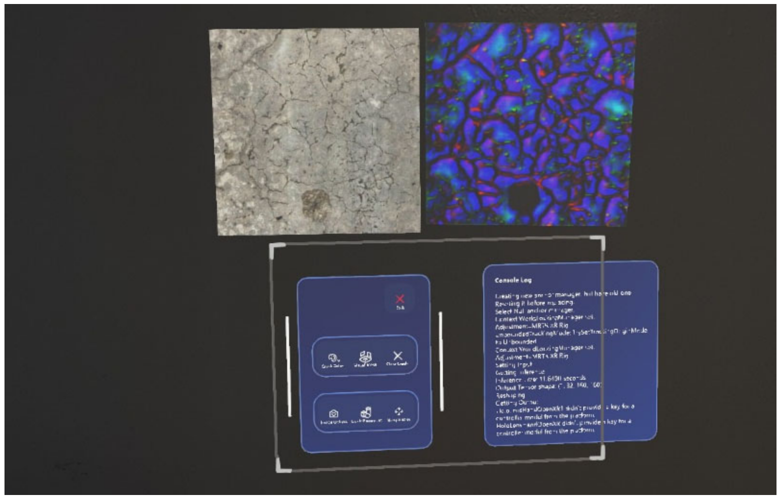

Figure 8. Deep Learning inference on a concrete crack, processed natively on the HoloLens.

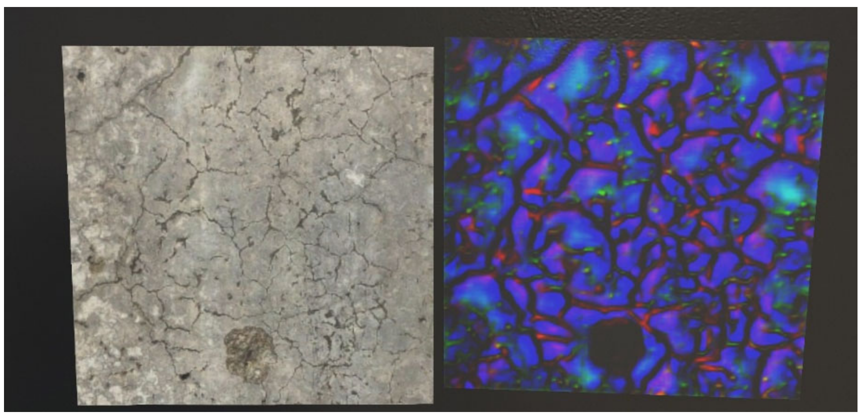

Figure 9. Image of crack before (left) and after (right) Deep Learning processing.

Issues were encountered with respect to moving the ML inference to run natively on the HoloLens. Our deep learning models were trained using Ultralitics YOLO, however these models are not natively compatible with Unity and as such would not run natively on the HoloLens. To solve this issue, the YOLO models were exported as ONNX models. Originally, we had planned to process the inferences using Unity's Barracuda package, built for processing ML inferences and tensors, however this package proved to be outdated, and could not properly process our models. It was then discovered that Unity has a more up-to-date package for these tasks named Sentis. Sentis was able to properly import our model and make an inference using it. The model outputs can be seen in figure 8 and figure 9. However, they are incorrect and did not match their input images, so we are now moving towards using YoloSharp, a third-party package for processing YOLO models in C Sharp, in order to process the ML inferencing.

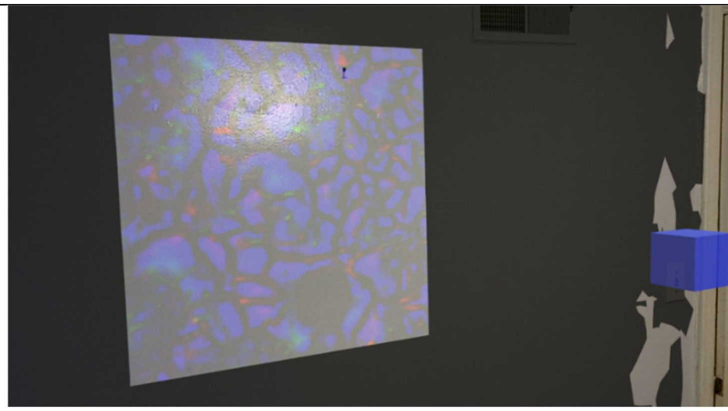

Figure 10. Deep Learning output anchored to a wall via projection from the blue cube.

Anchoring ML inferences of concrete damage to their real-life counterparts has been a major question of the AR software package all through development, and we believe that Unity's projector components will provide a viable solution. Unity's projector components are used in order to project a 2D image onto a 3D surface in the same way that a real projector would, wrapping around corners and edges of objects. We plan to use this feature to project the inference results onto the surfaces they are taken from as seen in figure 10.

Maintaining the position of holograms after inspection for future inspections is known as spatial persistence and is a big part of the AR software's requirements. To solve this, Microsoft's "World Locking Tools" have been identified as the most likely solution to maintain spatial persistence. We have chosen this software package as the best possible option for spatial persistence as it is designed with our specific requirements in mind and would allow for spatial anchors to be saved and loaded across sessions. Installation and implementation of this software package is ongoing.

Circumstance affecting project or budget. (Please describe any challenges encountered or anticipated that might affect the completion of the project within the time, scope and fiscal constraints set forth in the agreement, along with recommended solutions to those problems).

N/A

Potential Implementation:

Project Kickoff Meeting: TPF-5(535) Human-centered Steel Bridge Inspection enabled by Augmented Reality and Artificial Intelligence

Meeting Minutes

10/14/2024, 1:30-2:30 pm CDT

Present:	Jian Li (U Kansas) – PI
	Caroline Bennett (U Kansas) – Co-PI
	William Collins (U Kansas) – Co-PI
	Luke Attard (U Kansas) – Graduate Student
	David Behzadpour (KDOT) – Technology Transfer Engineer
	Sally Mayer (KDOT) –Assistant Bureau Chief, Research
	Mark Hurt (KDOT) – Bureau Chief, Structures and Geotech
	Dan Wadley (KDOT) – Bureau Chief, Research
	David Snoke (NCDOT) – Bridge Inspection Program Manager
	Shawn Hart (Caltrans) – Senior Bridge Engineer
	Justin Wilson (TxDOT) – Transportation Engineer Supervisor
Unable to attend:	Hoda Azari (FHWA) – Nondestructive Evaluation Research Program Manager
	Jorgomai Ceesay (FHWA) – Supervisory Civil Engineer (Structural)
Next meeting:	TBD, in about 3 months following submission of first quarterly report

1. Introductions

- J. Li began introductions of all meeting attendees, as noted above .

2. Project overview by J. Li (slides appended at end of minutes)

- J. Li provided brief overview of project
- Fatigue cracks and corrosion are huge issues related to steel bridge inspection and maintenance
- Current inspections rely on human vision, which can be labor-intensive, expensive, and unreliable.
- Previous research utilized augmented reality (AR) to superimpose virtual objects onto the physical world.
  - Brief overview of AR, including other applications and previous KU research
  - NCHRP IDEA 223 utilized AR to assist inspectors with fatigue crack inspection
    - Recognition of challenges related to image-based crack detection (false positives common with corrosion, degradation of surface coatings, etc.)
    - Project used video-based approach tracking movement of feature points
    - Other challenges addressed in previous work include camera motion compensation by tracking local distances
    - Application and lab-based validation of AR application of fatigue crack identification
    - Discussion of threshold value selection and corresponding feature point identification as well as feature point anchoring on structure as HoloLens moved relative to structure
    - Limitations identified in field implementation

      - Namely, no crack opening detected due to very small amount crack movement, although relative displacements of connect plate were captured
- Project Plan for TPF-5(535)
  - Inclusion of corrosion along with fatigue crack identification
  - Utilization of both video- and image-based tools
  - Two deployment mechanisms for AR inspections: AR Inspector (HoloLens headset or similar) and unmanned aerial vehicle (UAV) connected to tablet used by inspector
  - Project will conduct both lab and field testing
- Introduction of specific project tasks:
  - Task 1: Computer vision (CV) and artificial intelligence (AI) algorithms for crack and corrosion inspection
    - Continue to improve the video-based fatigue crack detection method
    - New AI-based algorithms for fatigue crack and corrosion detection using images
  - Task 2: AR-based software for human-centered bridge inspection
    - Focus on algorithm integration and software development to support project goals
    - Discussion of UAV platform used in project
      - Need to access live feed of UAV video
      - Additional sensors may be needed on UAVs
      - Multiple UAVs/platforms currently being explored
  - Task 3: Laboratory and Field Testing
- Project schedule introduced. Project started October 2024 and is planned for three years
- Communication plan for duration of project:
  - Quarterly reports submitted to FHWA TPF, followed by virtual meetings with group
  - Midterm project meeting planned in year 3 for on-site meeting
  - Virtual project meeting planned at completion of project

3. Questions and Discussion with group

- D. Snoke asked about the need to clean surface prior to inspection in the event of excessive corrosion
  - J. Li responded that this could occur if absolutely necessary, but hopefully the machine learning model will be trained to adequately identify corrosion without removing surface debris.
- D. Snoke also noted that NCDOT Aviation Department has lots of opinions on selection of UAVs, etc. and would be happy to provide information and give feedback.
- S. Hart also noted that DJI UAVs are no longer being used by Caltrans (can continue using DJI that they currently own but cannot purchase new UAVs). Caltrans currently uses Skydio UAVs, among others.

Recorded by: W. Collins

Reviewed by: J. Li