AI-AR for remote visual bridge inspection by drone

Lapointe, Jean-François; Sekkati, Hicham; Allili, Mohand Saïd; Hebbache, Loucif; Amirkhani, Dariush; Hammouche, Nadir

Publisher’s version / Version de l'éditeur:

Proceedings of the 11th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII-11), pp. 462-465, 2022-08-12

AI-AR for Remote Visual Bridge Inspection by Drone

Jean-François Lapointe[1], Hicham Sekkati[1], Mohand Saïd Allili[2], Loucif Hebbache[2], Dariush Amirkhani[2], Nadir Hammouche[2]
[1]Human-Computer Interaction Team, Digital Technologies Research Centre, National Research Council of Canada
[2]Department of Computer Science and Engineering, University of Quebec in Outaouais, Gatineau, Quebec, Canada, J8X 3X7
Email : [jean-francois.lapointe, hicham sekkati]@nrc-cnrc.gc.ca, [mohandsaid.allili, hebl08, amid01, hamn14]@uqo.ca

ABSTRACT: Regular inspections of bridges and overpasses are required to ensure the safety of those infrastructures. This paper presents a system developed to do the remote visual inspection of concrete bridges and overpasses by combining drone, artificial intelligence and augmented reality technologies with the goal to assist bridge inspectors in conducting a remote visual inspection. It also reports on the results of a first field trial of this complex human-machine system. Deep learning networks can be trained to detect defects on independent images but the user interface must be adapted to track defects in a stream of images at a speed consistent with human vision to allow inspection. This first field trial allows for an initial estimate of a speed up factor of ten (10x) to conduct visual inspections. The developed system could lead to significant cost reduction to inspect the bridges and overpasses as well as significant reduction of downtimes for these transportation infrastructures.

KEY WORDS: SHMII-11; VISUAL INSPECTION; REMOTE INSPECTION; BRIDGE INSPECTION; DRONE; ARTIFICIAL INTELLIGENCE; AUGMENTED REALITY.

1 INTRODUCTION

The recent tragedy caused by the collapse of the Morandi bridge in Italy [1] illustrates the importance to conduct regular inspections of such transportation infrastructures [2]. Up to recently, these inspections were done mainly manually but the recent use of drones, also known as unmanned aircraft vehicles (UAVs), unmanned aircraft systems (UAS) or remotely piloted aircraft systems (RPAS), which capture aerial images of the bridges, allows to speed up the visual inspection of the infrastructures and make it easier to access these images. Reducing the inspection time can have great benefits such as reducing the down time of the infrastructures as well as potentially reducing inspection costs.

Unmanned aerial vehicles were used by Kim et al. [3] to inspect and detect cracks in the concrete bridges. To do this, a camera mounted on a drone was used to acquire images, and cracks in the bridge were detected in the images using convolutional neural networks (R-CNN). Benz et al. [4] used unmanned aircraft systems to collect data to inspect the bridge. A convolutional neural network called TERNAUSNET was then used for cracks detection. Kumar et al. [5] used YOLO-v3 to detect in real-time cracks and spalls in images acquired by UAV system.

In this paper, we present the results of a first field trial of a system that combines artificial intelligence (AI) and augmented reality (AR) to drones for the visual inspection of bridges and overpasses.

2 SYSTEM DESCRIPTION

2.1 Concept

The concept of our system is illustrated in Figure 1. The system combines several hardware and software components in an open-loop fashion that continuously interact the pilot/expert with the drone and a computing station to perform the inspection.

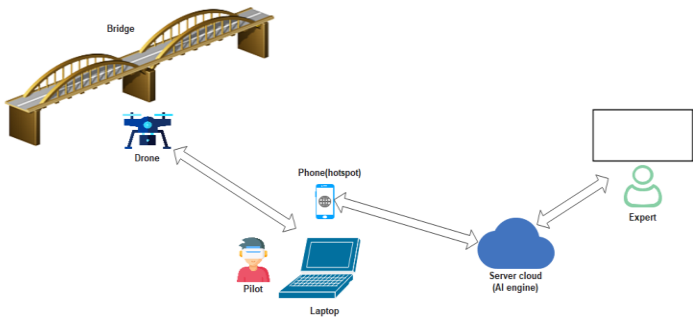

Figure 1. Concept of the system.

2.2 System architecture

The software architecture of the system is illustrated in Figure 2. The system has a modular architecture that can run on multiple platforms.

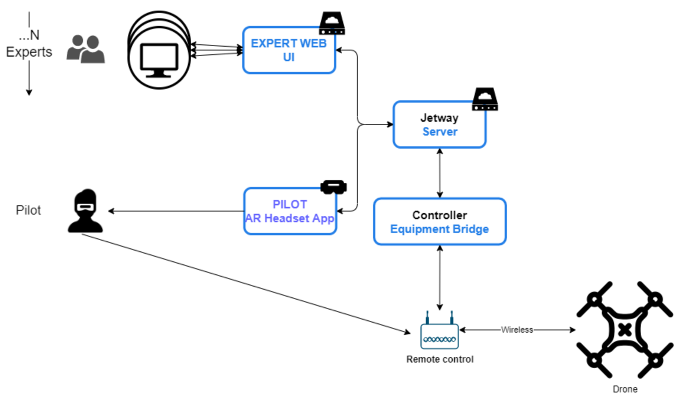

Figure 2. System architecture.

2.3 System hardware devices

The drone used for the trial was a microdrone (i.e. with a mass under 250 g), namely the DJI Mini 2 microdrone. The controller ran on an Android tablet, namely a Samsung Galaxy Tab S7+. The phone used for the hotspot (Wi-Fi) connection was a Samsung Galaxy S8+.

2.4 AI Module

The AI module, by using images acquired by the onboard camera, detects and segments the defects on the structure in the images and classifies them according to five types that are crack, spallation, efflorescence, exposed bar and corrosion. For this purpose, several recent deep learning models have been tested on the CODEBRIM dataset [12] to choose the most appropriate one.

Table 1 shows the comparison of defect detection results obtained with tested models based on the mAP (Mean Average Precision) metric and inference time; The mAP represents the mean value of the AP of all classes, and the AP is the area under the precision-recall curve; The precision and recall metrics are defined as follows:

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

Where TP, FP, and FN represent the number of true positives, false positives, and false negatives, respectively. To classify the detections as TP or FP, the IOU (Intersection Over Union) threshold is set to t = 0.5. The inference time metric represents the amount of time that an object detector takes to process one image (frame) of the test images (video).
It was observed that the YOLOX model outperforms the other models in terms of accuracy and speed, which is why we opted to use it for the defect detection. The YOLOX model is an improved version of YOLO's popular model. It improves both speed and accuracy. Recently it captured the first place at CVPR 2021's Automatic Driving Workshop with its Streaming Perception Challenge. Since CODEBRIM is a multi-label dataset, we used YOLOX as a multi-label detector which assigns more than one label for each bounding box. Figure 3 shows some sample results of YOLOX detection, where overlaid bounding boxes represent defects with multiple labels.

Table 1. Comparison of defect detection results.

Method	mAP@0.5	Inference time (s)
YOLOR[6]	0.86	0.05
SSD[7]	0.41	0.06
Faster RCNN [8]	0.36	0.14
DETR[9]	0.44	0.08
YOLOX[10]	**0.91**	**0.04**

Figure 3. Detection samples by YOLOX.

Another task of the AI module, that is still development, consists of semantic segmentation. We treated each class separately (due to the problem of the overlapping defects); we created 250 masks for the CODEBRIM dataset, and we used U-Net model as binary pixel classier for each class. Figure 4 illustrates some early images with defect segmentation using U-Net.

Figure 4. Early images of defect segmentation using U-Net.

2.5 AR Modules

The system uses two AR modules. The first one (see Figure 7) highlights the defects detected by the AI module on the web interface used by the remote experts (the inspectors). This AR web interface also allows one of the inspectors at a time to control the gimbal of the camera remotely. In addition, the web interface displays a map depicting the position of the drone on earth as well as aircraft information such as the latitude and longitude coordinates, the altitude, the orientation (pitch, roll, yaw) and allows to enable or disable the augmented reality. Finally, it allows to request and pass the control of the web interface to other remote inspectors in the case where there are many.

The second AR module (see Figure 5) is designed to provide a better piloting experience for the drone pilot. This module is built on the top of the Mixed Reality Toolkit (MRTK). Although still in development, it will be deployed on the Microsoft HoloLens 2 headset. This module will allow the pilot to visualize simultaneously both the real field of view and the video images captured by the drone during the flight in order to optimize his/her situational awareness. This aims at solving the *head down time* problem where the pilot constantly alternate his/her gaze between having a line-of-sight with the drone and looking at his/her remote control to capture information such as the battery level [11].

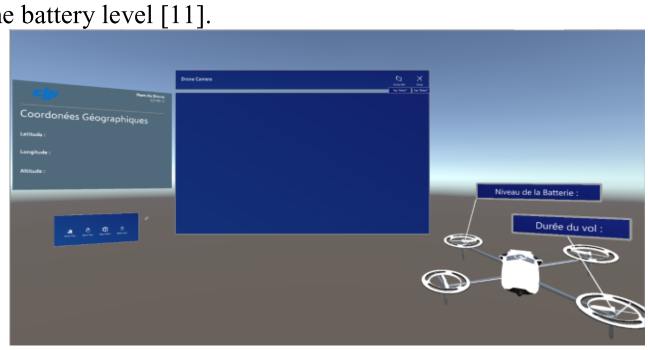

Figure 5. Simulated view of the pilot's AR environment.

3 FIELD TRIAL

The first field trial of the system was held in June 2022. The goals of this trial flight were to assess the technical maturity of the system, put in practice the user guide and to gather the lessons learned in order to refine the system for the next flight.

3.1 Timeline of the field trial

Once on the site (see Figure 6), the field trial took about an hour to complete. It allowed the research team to scan the bridge along its two sides with a drone for live remote inspection as well as two other flights to take ultra-high (4k) resolution video sequences of these sides for documentation and reuse to test new, enhanced, AI modules. It is important to note that although several people were on site to observe the trial, the experiment could have been conducted by a single person.

Figure 6. Aerial photo from the field trial site.

3.2 Results

Figure 7 below illustrates the view provided to the three remote volunteer inspectors that participated to the trial.

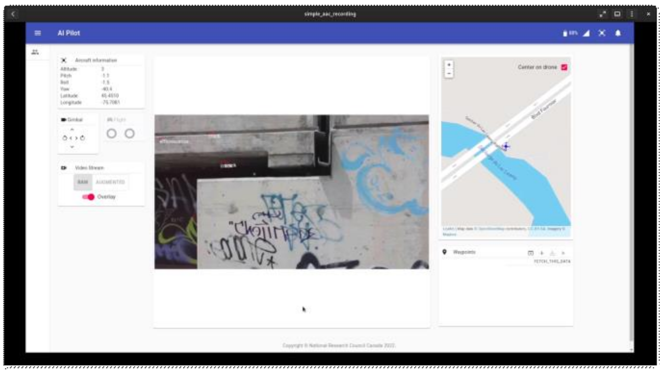

Figure 7. View provided to the remote inspectors through their web user interface.

It is possible to see on this figure that despite a relatively good bridge condition, several defects have been detected, namely efflorescence and cracks. It is important to note that the AI module successfully ignored the graffiti. We used the MS Teams application to provide audio connection among all the participants (pilot + remote inspectors).

3.3 Lessons learned

Since we were essentially able to complete the trial, the first and greatest lesson learned relates to completion time. As mentioned earlier, the whole inspection took about an hour to complete onsite. Our estimate is that it is 10 times faster than a regular visual inspection thus revealing the great efficiency of our system.

As a second lesson learned, the field trial revealed that the bandwidth provided by our smartphone connection was not sufficient to get the same results as those obtained in laboratory where the Wi-Fi connection provides a bandwidth of up to 180 Mbps. As a result, the video displayed on the web interface of the remote inspectors had a low frame rate (around 1 frame/s) to the point of reducing the user experience and effectiveness of the system. Also, since the onsite minimal system requirements were to use a smartphone, a tablet, and the drone's remote controller. It is important to keep these devices close to each other given that the tablet was connected by wire to the remote controller and that the smartphone was wirelessly providing a hotspot (Wi-Fi) connection to the tablet, with a practical range of about 15 m. In future trials, the bandwidth available onsite should be assessed first in order to predict the quality of the results.

A third lesson learned here was that the defects revealed by the inspection system were not always consistent from an image to another. For instance, a given defect could be detected on one image but not on another, thus giving the impression of inconsistent results. These inconsistencies could be the results of several factors including various lighting conditions and different viewpoints.

A fourth lesson learned was that a good audio connection should be provided between the drone's pilot and the remote inspectors to coordinate the flight and, if needed, revisit more closely some parts of the bridge. In this case, we used MS Teams to stream the audio connection but the low bandwidth connection affected the quality of the audio.

A fifth lesson learned was that the augmented reality images provided to the remote inspectors could take more space on their scree during the flight to allow a better visualization of the images and the detected defects. As such, information on the aircraft and control of its camera by the remote inspectors is of secondary importance and could be called on an as-needed basis thus taking less display space. Alternatively, the use of large screens is also an option.

4 DISCUSSION

This first field trial allowed practicing the deployment of this system and testing its different components such as the computing platform, the network connection as well as determining the task completion time.

We think that the addition of an accurate 3D model of the bridge would allow anchoring the detected defects on a model that could be used for later visualization and review of the evolution of defects in time. As such, we are investigating the idea of estimating depth by using deep learning methods with both mono and stereo image sequences. Moreover, since it is important to have the same resolution between the training and testing images, the AI module can be enhanced to get automatically the drone at the right position with regard to the bridge concrete surface. This will help avoiding false positives/negatives that can be caused by the mismatching of the distribution of training and testing images, and thus enhancing the inspection accuracy.

In terms of efficiency, this first field trial let us think the performance improvement we can expect using this system to visually inspect a bridge would be of an order of magnitudes (10x) improvement over traditional methods. This means that this could reduce the downtime of these transportation infrastructures by a factor of 10, which would be a significant improvement. Further field trials and a task analysis of the bridge inspection are now needed to confirm this number.

5 CONCLUSION

The field trial of this human-in-the-loop bridge inspection system, allowed to realize several great lessons. First, a good onsite network connection is required for good system performance and user experience for the remote inspectors. Second, due to various factors, the output of the AI module varies from an image to the other and thus results in a video stream where the defects highlighted often vary from an image to another even though the defects are visible on several images. This behavior must be corrected to give enough time for the inspectors to take note of each defect detected by the AI module.

Overall we expect our system to allow for a 10 times reduction in visual inspection time. Although further field trials and a task analysis of bridge inspection are needed to confirm this initial estimate. This represents promising results in the ability to reduce inspection cost and transportation infrastructure downtime.

ACKNOWLEDGMENTS

This project was supported in part by collaborative research funding from the National Research Council of Canada's Artificial Intelligence for Logistics Program.

REFERENCES

[1] G.M. Calvi, M. Moratti, G. J. O'Reilly, N. Scattarreggia, R. Monteiro, and D. Malomo. *Once upon a time in Italy: the tale of the Morandi bridge*. Structural Engineering International. 29(2), 2019, pp. 198–217. https://www.tandfonline.com/doi/abs/10.1080/10168664.2018.1558033

[2] JF. Lapointe, M.S. Allili, L. Belliveau, L. Hebbache, D. Amirkhani, H. Sekkati. *AI-AR for bridge inspection by drone*. In: Chen, J.Y.C., Fragomeni, G. (eds) Virtual, Augmented and Mixed Reality: Applications in Education, Aviation and Industry. HCII 2022. Lecture Notes in Computer Science, vol 13318. Springer, Cham, 2022, pp. 302-313. https://doi.org/10.1007/978-3-031-06015-1_21

[3] I.-H. Kim, H. Jeon, S.-C. Baek, W.-H. Hong and H.-J. Jung. *Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle*. Sensors 18 (6), 2018, 1881, 14 pages. https://doi.org/10.3390/s18061881

[4] C. Benz, P. Debus, H. K. Ha and V. Rodehorst. *Crack segmentation on UAS-based imagery using transfer learning*. In 2019 International Conference on Image and Vision Computing New Zealand (IVCNZ), IEEE, 2019, pp. 1–6. https://doi.org/10.1109/IVCNZ48456.2019.8960998

[5] P. Kumar, S. Batchu, N. Swamy S. and S. R. Kota. *Real-time concrete damage detection using deep learning for high rise structures*. In IEEE Access, vol. 9, pp. 112312-112331, 2021, https://doi.org/10.1109/ACCESS.2021.3102647

[6] CY. Wang, IH Yeh, HY Liao. *You only learn one representation: Unified network for multiple tasks*. arXiv preprint arXiv:2105.04206. 2021 May 10.

[7] W. Liu, D. Anguelov, D. Erhan, C. Szegedi, S. Reed, CY Fu and A. C. Berg. *SSD: Single shot multibox detector*. European conference on computer vision (ECCV 2016). Springer, Cham, 2016, pp. 21-37.

[8] S. Ren, K. He, R. Girshick and J. Sun. *Faster R-CNN: towards real-time object detection with region proposal networks*. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6), pp. 1137-1149, 1 June 2017, https://doi.org/10.1109/TPAMI.2016.2577031

[9] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov and S. Zagoruyko. *End-to-end object detection with transformers*. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, JM. (eds) Computer Vision – ECCV 2020. Lecture Notes in Computer Science(), vol 12346. Springer, Cham, 2020. https://doi.org/10.1007/978-3-030-58452-8_13

[10] Z. Ge, S. Liu, F. Wang, Z. Li and J. Sunet. *YOLOX: Exceeding YOLO series in 2021*. arXiv preprint arXiv:2107.08430 (2021).

[11] M. Safi, J. Chung and P. Pradhan. *Review of augmented reality in aerospace industry*. Aircraft Engineering and Aerospace Technology 91(9), 1187–1194 (2019). https://doi.org/10. 1108/AEAT-09-2018-0241

M. Mundt, S. Majumder, S. Murali, P. Panetsos and V. Ramesh. *Meta-learning convolutional neural architectures for multi-target concrete defect classification with the COncrete DEfect BRidge IMage dataset*. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019), pp. 11196– 11205 (2019). https://doi.org/10.1109/CVPR.2019.01145