Building Thinking Machines: Enhancing the Built Environment through Artificial Intelligence

Mark Sarkisian, PE, SE, LEED BD+C, Partner
Alessandro Beghini, PhD, PE, SE, Associate Director
Samantha Walker, PE
Abhinav Bindal
Skidmore, Owings & Merrill LLP
San Francisco, CA

Anthony Sarkis
Diffgram
Santa Clara, CA

Abstract

Artificial intelligence is currently being applied across a wide range of fields, from self-driving car technology to art and music generation. This paper explores different ways in which the built environment can be enhanced through artificial intelligence, from seismic assessment to construction verification and infrastructure health monitoring. A variety of computer vision approaches have been used for these purposes, including novel combinations of established machine learning, optical character recognition and geometric methods.

This paper presents results from a machine learning model that was trained to identify and categorize post-earthquake building damage from photographs. Using this tool, photos could be directly uploaded to an earthquake database to identify and catalog building damage in real time, rather than manually cataloging the photos in the weeks and months following the event. Eventually, this tool could allow individuals to obtain an immediate initial assessment of the condition of their building following an earthquake, rather than waiting weeks or months for an in-person inspection. A machine learning model that was trained to identify the early formation of hinges in timbrel vaults subjected to lateral loads is also presented.

To help automate the construction verification process, machine learning models were trained to identify structural components from site photos and "read" shop drawings. These tools could be used to support construction inspection efforts and to reduce shop drawing review times from several days to a matter of hours, with an engineer reviewing a summary output from a machine learning program. A case study is also presented in which machine learning was used to monitor building health by identifying and measuring check formation and growth in glulam girders from a project under construction in California. Finally, this paper will discuss future applications of machine learning to enhance structural design.

Introduction

The concept of artificial intelligence (AI) has existed for decades, but the world is only now on the verge of an AI revolution due to the following factors: (1) access to large amounts of digital data and (2) access to powerful and affordable cloud computing power.

In parallel, the need for a technology revolution in the architecture, engineering and construction (AEC) industry is imminent. The current world population of 7.6 billion is expected to reach 8.6 billion in 2030 (United Nations, 2017). Infrastructure demands will increase drastically and projects will need to be delivered more quickly to meet this rapid population growth. Design and construction processes need to be made exponentially more efficient and sustainable in order to meet these demands in a responsible manner.

There are many opportunities to enhance the built environment through artificial intelligence. Incorporating AI into construction verification and infrastructure health monitoring could improve quality control and decrease construction time. AI-enabled infrastructure health monitoring tools could be used to improve maintenance and safety practices by evaluating the deterioration of both non-structural components, such as exterior walls, and structural components. Photo recognition-based systems can be applied to

continuously evaluate and monitor bridges and highlight potential maintenance and safety issues in advance. AI damage assessment tools could make recovery and rebuilding efforts more efficient by rapidly and automatically evaluating structural integrity following major natural or man-made events. In addition to practical applications, AI can also be incorporated into current design processes to optimize for multiple parameters and develop new design options that may not have been thought of otherwise.

In light of these opportunities and the impending global transformations described above, Skidmore, Owings and Merrill (SOM) and software developer Anthony Sarkis have been experimenting with applying a variety of computer vision approaches to monitor and evaluate the built environment. This paper presents results from (1) construction verification proof-of-concept studies, (2) a health monitoring case study for checking in glulam girders, (3) post-earthquake building damage assessment and (4) hinge detection and collapse identification in timbrel vaults subjected to lateral loads.

1. Construction Verification

To advance efforts in applying machine learning to the construction industry, SOM and software developer Anthony Sarkis performed several proof-of-concept studies utilizing photo recognition for construction verification. The intention of these studies was to provide an initial viability assessment of these applications with minimal data gathering, annotation and computational effort. Therefore, small datasets on the order of tens of images were used to train the models.

All models were trained using the Tensorflow library (Abadi, et al., 2016). Depending on the element in question, two different photo recognition techniques were employed: (1) object detection and (2) semantic segmentation. In object detection, a bounding box is drawn around each class identified in the image. In sematic segmentation, every pixel in the image is associated with a class.

A machine learning model was first trained to recognize rebar from site photos using object detection based on Faster RCNN (Ren, et al., 2015) and Resnet-101 (He, et al., 2016). The model was initialized with the pretrained weights from the MS COCO dataset. However, as shown in Figure 1 below, the density of rebar is typically such that multiple, overlapping bounding box regions are identified in the photo, rendering the photo recognition results meaningless.

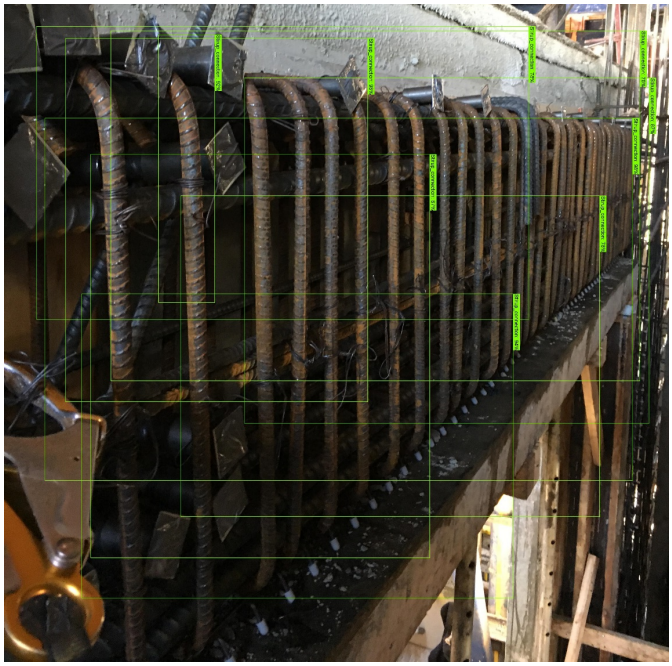

Figure 1: Rebar identification through object detection.

To increase precision in the photo recognition results, a second machine learning model was trained using semantic segmentation. Two different algorithms were used: (1) an approach based on VGG 16 with skip connections and (2) DeepLabV2 (Chen, et al., 2016) with ResNet101. As shown in both Figure 2 and Figure 3, the second algorithm achieved significantly better results than the first in terms of detecting the exact region of the main reinforcing bars shown in the foreground of the images.

Figure 2: Rebar identification through semantic segmentation using VGG 16.

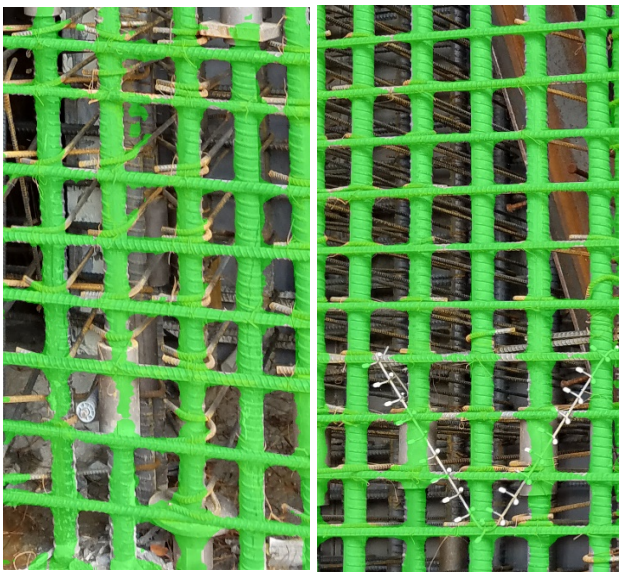

Figure 3: Rebar identification through semantic segmentation using DeepLabV2.

The semantic segmentation technique can be combined with classical computer vision techniques to differentiate between different types of rebar, such as vertical bars, horizontal bars, ties and stirrups. Combining these techniques with geometric methods, the size and spacing of the rebar can also be estimated.

A second machine learning model was trained to identify gusset plates from photographs using object detection. The model was able to identify gusset plates with a high degree of confidence from photographs taken at various angles. For example, in Figure 4 the model detected not only the two gusset plates in the foreground facing the camera, but also the gusset plate in the background at an angle to the camera. This success of this model may be due to the unique shape of gusset plates, which makes them clearly distinguishable from other elements on a construction site.

Figure 4: Identification of gusset plates from site photos using object detection.

In order to compare objects identified from site photos to the construction documents, a machine learning model was also trained to identify gusset plates from shop drawings. Figure 5 shows a gusset plate being detected both from a site photo and a shop drawing using object detection. The machine learning model was able to identify gusset plates in shop drawings to a high degree of confidence. This performance is likely due to both the distinguishable shape of gusset plates and the consistency and precision of steel shop drawings.

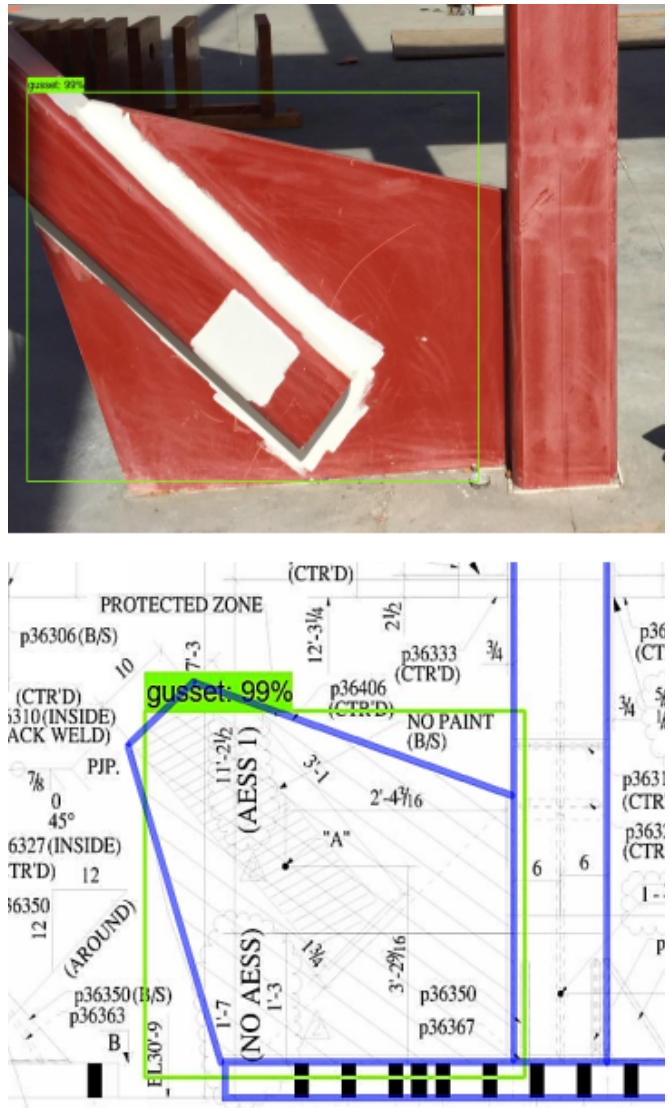

Figure 5: Identification of a gusset plate from both a site photo and shop drawings (blue outline added manually to highlight the steel shapes).

Methods to extract useful information from shop drawings were further explored using machine learning. Since a large portion of the information provided in shop drawings is in textual form, some shop drawings were processed through Google's Cloud Vision API optical character recognition platform (Google, 2018). As shown in Figure 6, the text was successfully extracted from the drawing, however the different text blocks are meaningless without any indication of the drawing element to which they refer.

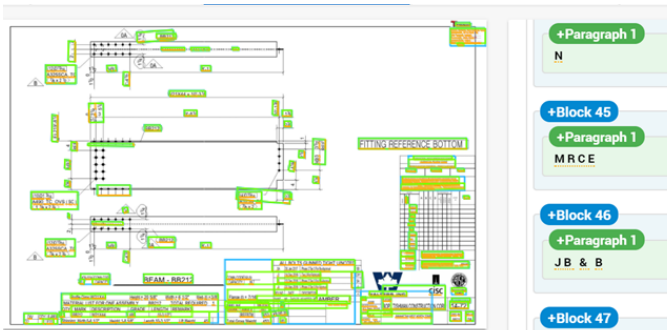

Figure 6: A shop drawing processed through Google's Cloud Vision API *(Google, 2018)*.

To provide better context to the extracted text blocks, object detection can first be applied to the shop drawing to divide it into components and sub-components. Figure 7 shows the results of a machine learning model that was trained to recognize the different component views shown in steel beam shop drawings.

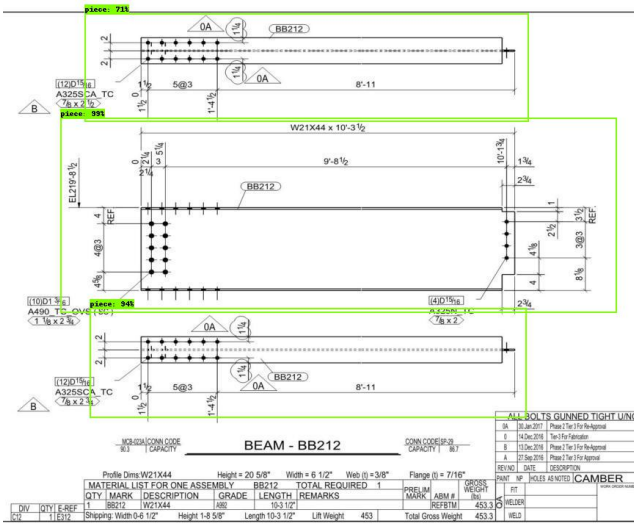

Figure 7: Different component views identified in a steel beam shop drawing using object detection.

In Figure 8, a machine learning model was trained to recognize subcomponents for each element, such as dimension ticks, weld symbols and bolt groupings. By subdividing shop drawings into manageable components and subcomponents through object detection and using optical character recognition to extract their associated text blocks, machine learning models can be trained to "read" shop drawings.

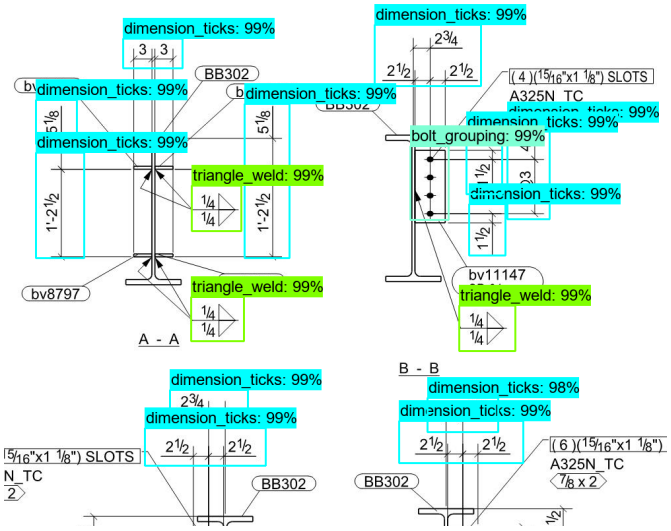

Figure 8: Different subcomponents identified in a shop drawing using object detection.

Using object detection, another machine learning model was trained to recognize shear studs from site photos. Because shear studs are small and are typically counted over a large area, a digital single-lens reflex (DSLR) camera was used to take high-resolution photographs of large areas containing shear studs. However, machine learning models are very limited in terms of resolution. It is computationally intractable to feed a full DSLR image into a typical machine learning model on today's hardware. To get around this limitation, the images were divided into nine sub images and then each sub-image was processed independently. Afterwards, the photos were recombined. This allowed the onsite user to take a single photograph, while enabling the system to process the images at an effective resolution. A sample of a processed sub-image is shown in Figure 9 below.

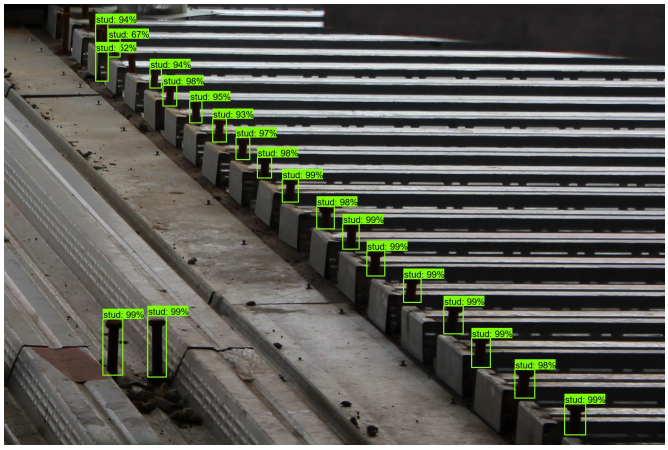

Figure 9: Shear studs identified from a sub-image.

A sample of a recombined photo is shown in Figure 10 below. Although some studs in far corners of the photo were not identified, the lack of false positives in these results indicate that a larger training set could produce a more robust model.

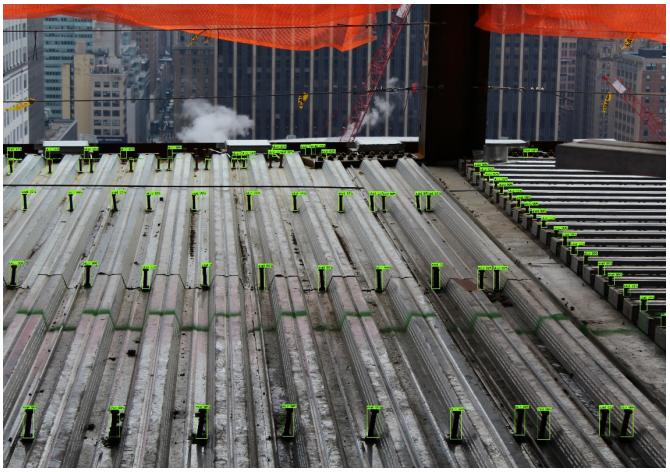

Figure 10: Shear stud object detection results shown in a recombined photo.

In another application, a machine learning model was trained to recognize bolts from site photos. Figure 11 illustrates that not all the bolts were successfully identified, potentially due to the angle and lighting conditions of the photo. However, no false positives were identified in the photo. Therefore, a larger and more varied training set could lead to improved results. Issues with identifying bolts in perspective could also be controlled by limiting the machine learning model to head-on photos of bolt groups.

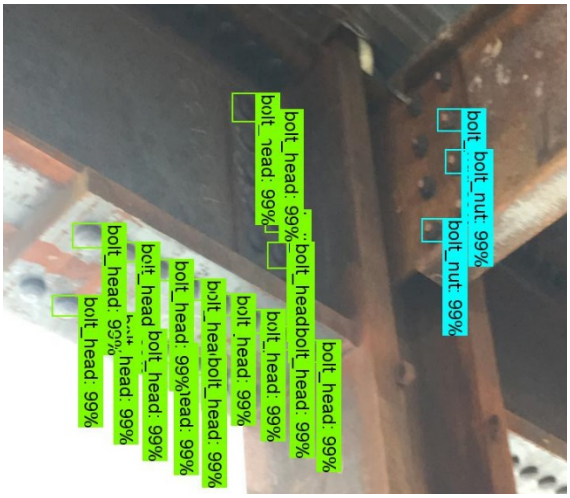

Figure 11: Bolts identified from a site photo using object detection.

Tests were also performed to detect welds using semantic segmentation, as shown in Figure 12 below. However, because welds by nature blend into the adjacent steel members, they are not easily visible in photographs and are therefore difficult to detect through photo recognition.

Figure 12: Weld detection using semantic segmentation.

2. Health Monitoring Case Study: Checking in Glulam Girders

In a project under construction in Southern California, significant separations were noted in certain glulam girders a few weeks after they were installed. An example of these separations is shown in Figure 13 below, with the green tape indicating the locations of some notable separations. An investigation deduced that these separations were partly due to abnormal fabrication methods and party due to major temperature and humidity differences between the fabrication environment and the site environment in which they were installed. The glulam girders with significant separations were removed and replaced with new glulam girders.

Figure 13: Significant separations noted in glulam girders in a project under construction in Southern California.

Prior to their replacement, SOM and Sarkis took several images of these girders and used them in a proof-of-concept study to train a machine learning model to recognize the separations in the images using object detection. The object detection technique was then combined with classical computer vision to determine the dimensions of the separations. The measurements were calibrated based on the depth of the glulam laminations. An example of these results is shown in Figure 14 below.

Figure 14: Separation identification and dimension measurement in a glulam girder using object detection and classical computer vision methods.

Due to the success of the proof-of-concept study, SOM and Sarkis were commissioned to devise a Glulam Monitoring System (GMS) to monitor a few glulam girders at select locations for check formation and growth.

SOM selected four locations at which to monitor glulam girders; two exhibiting some checking and two at critical locations, but not exhibiting checking at the time the locations were selected. Each location will be monitored with a DSLR camera, which will be programmed to take a picture once per day using a timed remote shutter release. These locations will be monitored for one year after building systems commissioning, which is approximately how long will take for the wood members to become acclimatized to the new environment.

The images can be uploaded to a platform being developed by Sarkis to analyze the images and obtain results regarding the number and size of the checks identified in the images. A screenshot showing a preview of this platform is presented in Figure 15.

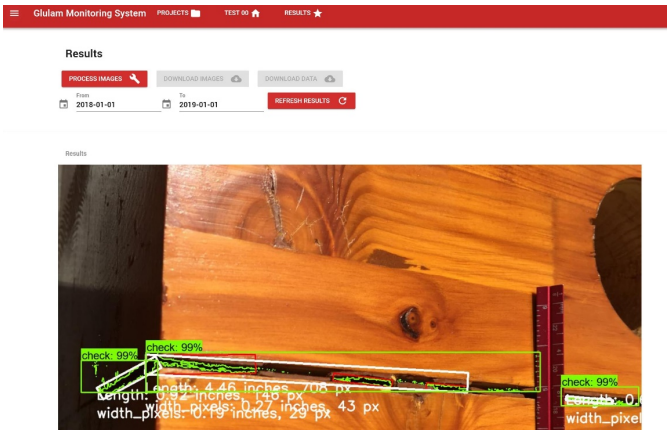

Figure 15: Screenshot showing a preview of the GMS platform for glulam check identification and measurement for images.

SOM designed a custom camera box system in which to house the cameras and auxiliary equipment, while being seamlessly integrated into the building surroundings. These boxes are shown in Figure 16. The boxes were fabricated out of plywood. They were designed to fit between and be secured to parallel joists. An adjustable base was designed to tweak the camera angle for each location being monitored. A hole was cut in the front of the box to provide the camera lens with an unobstructed view of the girder being monitored.

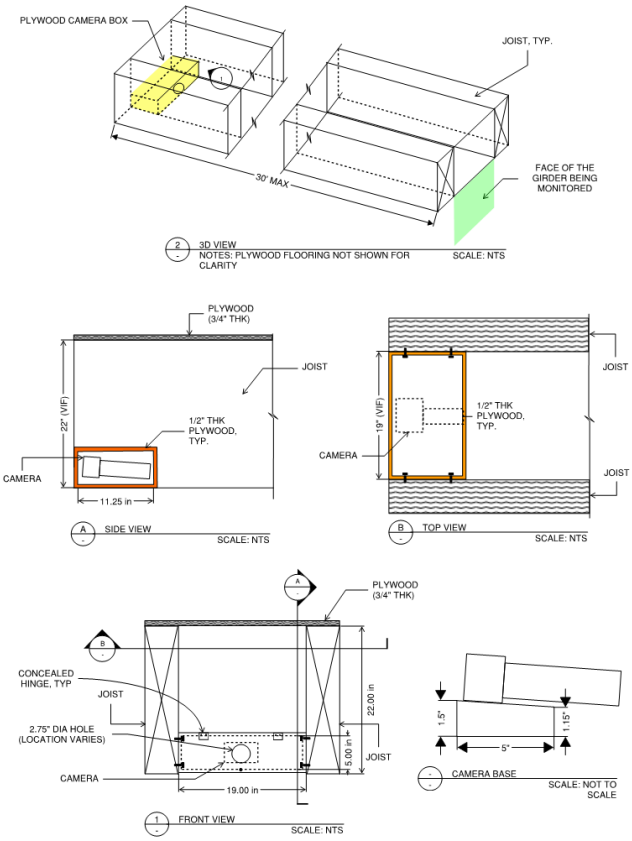

Figure 16: GMS camera box sketches.

The camera boxes are first being installed at various locations throughout the site to gather approximately 1000 images of checking in glulam girders to be used in the initial training set. Images will be taken once per hour at each location to obtain images with various lighting conditions. Figure 17 and Figure 18 on the following page show photos of camera boxes installed at the temporary locations.

Figure 17: Camera boxes being installed at their temporary locations.

Figure 18: Close-up view of one of the camera boxes installed at a temporary location.

Figure 19 below shows examples of the training set images that are being gathered to train the machine learning model for check identification using object detection.

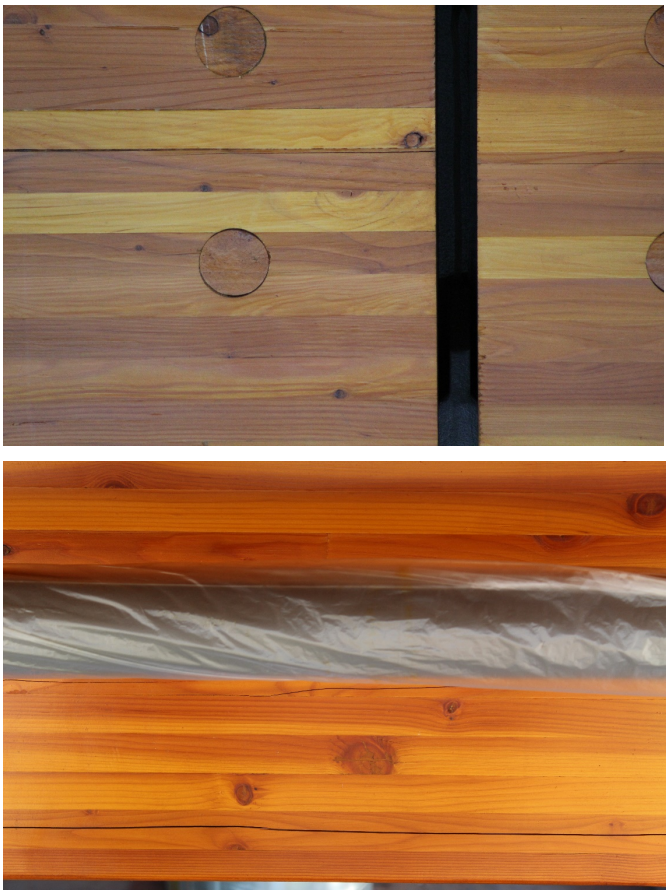

Figure 19: Examples of training set images for check detection.

Once the machine learning model has been trained and tested, the camera boxes will be installed at their final locations. WiFi-enabled SD cards will be used to obtain the images remotely from the cameras. If needed, an external battery source will be connected to the camera to keep it powered without the need for human intervention once the building is open to the public.

3. Post-Earthquake Building Damage Identification

One of the greatest economical losses associated with earthquake events is due to downtime from repair and rebuilding. Home and business owners of damaged buildings can wait months for an inspection to be completed by an expert. In some cases, they cannot occupy their home or operate their business until this inspection is complete. Photo recognition technology can be employed to expedite this evaluation process and mitigate this lost time. A comprehensive and vetted machine learning-based application could provide users with immediate feedback regarding the level of damage of their building, whether or not it can be occupied and whether or not it requires major repairs.

In addition, in the aftermath of a major earthquake, local and international reconnaissance teams upload hundreds of images to various databases, not all of which can be rapidly evaluated by experts. Integrating machine learning into these databases could make the photo cataloguing process not only more uniform and efficient, but also more comprehensive.

Following the 7.1 magnitude Central Mexico earthquake on Tuesday, September 19th, 2017, SOM sent a team of engineers to Mexico City to contribute to post-disaster recovery efforts by documenting building damage and providing technical support to local structural reconnaissance efforts. The preliminary observations from this reconnaissance mission are documented in Diaz et al. (2017).

The SOM team uploaded their in-field photos to the Earthquake Engineering Research Institute (EERI) Virtual Clearinghouse. The different locations these photos were taken are shown in the map in Figure 20.

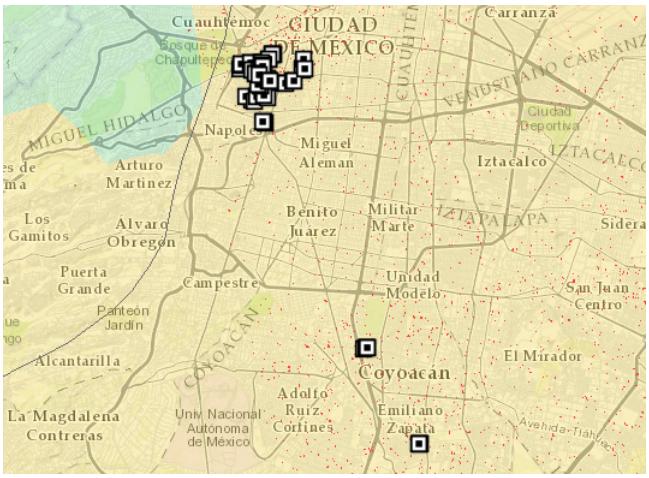

Figure 20: Enlarged map showing the location of data uploaded by the SOM reconnaissance team to the EERI Virtual Clearinghouse (EERI, 2017).

In an effort to enhance post-earthquake reconnaissance capabilities, SOM and Sarkis performed tests using machine learning to identify and classify building damage from the photos taken by the SOM reconnaissance team. Some results from these tests are shown in Figure 21 below.

Figure 21: Examples of building damage identification and classification using machine learning.

(a) (b)

Figure 22: Comparison of results from training set images from machine learning models trained using (a) 3,000 iterations and (b) 10,000 iterations.

An object-detection based machine learning model was trained using approximately 40 annotated images in the training set. The model was trained to differentiate between the following eight classes: (1) nonstructural light damage, (2) nonstructural medium damage, (3) nonstructural heavy damage, (4) nonstructural severe damage, (5) structural light damage, (6) structural medium damage, (7) structural heavy damage and (8) structural severe damage. The model was based on a Faster RCNN (Ren, et al., 2015) approach using Resnet-101 (He, et al., 2016). It was initialized with the pretrained weights from the MS COCO dataset.

One model was trained using 3,000 iterations, while the other was trained using 10,000 iterations. Figure 22 shows a comparison of the results from the two models using the same training image. The model trained to 10,000 iterations most likely overfit the data. This means that it very accurately represents the training data, at the expense of being as useful to novel data. Using a similar method with a training set of at least 1000 images may yield a model that generalizes well to new data.

While the training set used for these initial proof of concept studies was relatively small, the promising results demonstrate the potential to train a robust machine learning model using a larger training set.

Figure 23: Training set images showing detection of progressive hinge formation and arch collapse.

4. Timbrel Vault Hinge Detection and Collapse Identification

In 2017, SOM was invited to participate in an annual timbrel vault workshop in Madrid, Spain organized by architect Julio Jesus Palomino Anguí and hosted by the University of Alcalá. SOM was asked to provide insight on the seismic behavior of masonry arches and on methods to enhance their resistance to seismic loading. A reinforcement system using fiber grid material was studied through numerical analysis and tested on a shake table at the workshop. The fiber grid reinforcement proved to significantly improve the lateral load resistance of masonry arches. The findings of the 2017 workshop are detailed in Sarkisian et al. (2018).

SOM and Sarkis performed tests using a few images from the 2017 workshop footage to train a machine learning model to detect hinge formation in masonry arches subjected to lateral loading on a shake table. The training set results are shown in Figure 23 and Figure 24.

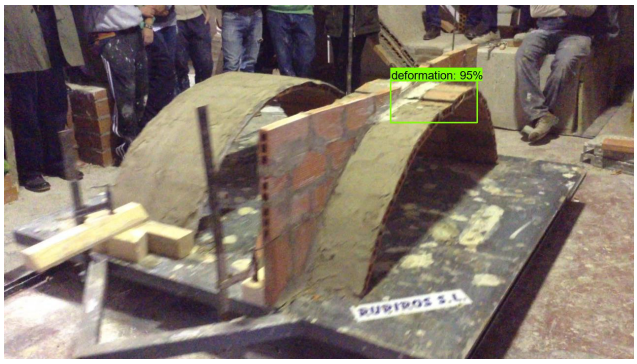

Figure 24: Training set image showing hinge detection in a ribbed arch.

During the 2018 workshop, three iPhones were installed around the shake table to capture slow-motion videos of the shake table tests. The objective of capturing this footage was to train a robust model for live hinge detection at the 2019 timbrel vault workshop. Two iPhones were positioned on either side of the arches and one was hung from above the arches. This camera placement was selected to eliminate potential object detection complications associated with images in perspective and to provide a clear, unobstructed

view of each arch by capturing head-on shots of the arches from the side and top views. Slow-motion videos recorded at 240 frames-per-second were used to gather more data for the training set and to capture the moment of hinge formation more precisely. In total, over 3,000 images, including footage from the 2017 and 2018 workshops, were collected. Only a small portion were annotated and used as a training set for this study.

The machine learning model was able to detect partial and full arch collapse from the side view photos, however it had difficulty identifying the hinges in this view to a confidence greater than 50 percent. In Figure 25, partial collapse of the arch in the foreground is correctly identified to a confidence of 96 percent, however a confidence of greater than 50 percent was not achieved for the visible hinge in the same arch. In Figure 26, full collapse of the arch in the foreground is correctly identified to a confidence of 99 percent, however a confidence of greater than 50 percent was not achieved for the visible hinges in the arch in the background.

Figure 25: Side view of arch showing partial collapse detection. A confidence greater than 50 percent was not achieved for the visible hinge.

Figure 26: Side view of arch showing full collapse detection. A confidence greater than 50 percent was not achieved for the visible hinges in the arch in the background.

The machine learning model was able to detect middle hinges in the top view photos, however it had difficulty identifying hinges on the side of the arches to a confidence greater than 50 percent. In Figure 27 and Figure 28, the middle hinge is correctly identified in the leftmost arch, however there also appears to be a hinge in the same arch at the top left of the photo which is not identified in either photo. In Figure 28, collapse is also detected in the leftmost arch. Although it is hard to confirm this behavior from the photo, the presence of two hinges corroborates the collapse assessment.

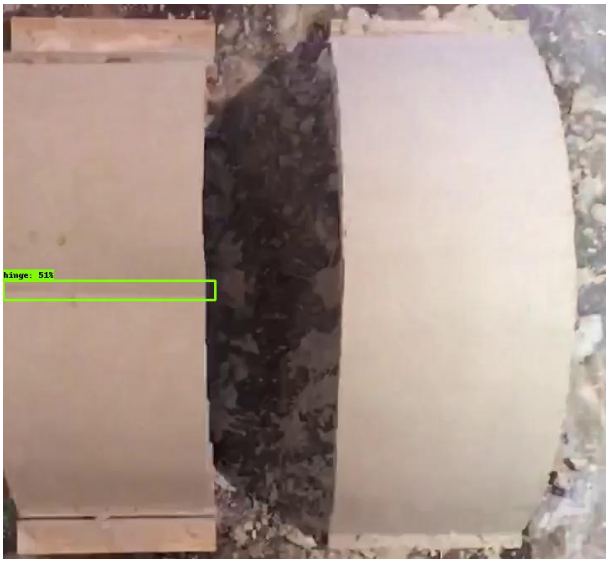

Figure 27: Top view of arches showing middle hinge detection.

Figure 28: Top view of arches showing middle hinge and collapse detection.

Since different views were found to be better for detecting different classes, a machine learning model can be trained using photos shot from various angles to build a more robust system for real-time hinge and collapse detection during next year's workshop. Footage shot by other workshop participants can be included in the training set to increase the quantity and variability of the data.

Conclusion and Future Work

Different methods of incorporating AI into the AEC industry were presented in this paper, from construction verification, to infrastructure health monitoring, to building damage and structural performance assessment. While these studies are helpful in demonstrating the possibilities of applying AI to enhance the built environment, there is significant room to further develop these models and make them more robust by increasing the quantity and variety of the images used in the training set. In addition, AI is a rapidly evolving industry. Different computer vision techniques and more efficient deep learning algorithms are constantly being developed and can be incorporated into these systems as well.

SOM is also investigating applying AI to expand their Environmental Analysis Tool™ (EA Tool™). SOM's EA Tool™ estimates the carbon footprint of a building taking into consideration initial construction, service life, repair after hazardous events and deconstruction. Applying machine learning to past project data, the tool could be expanded to reliably predict parameters such as material quantities, schedule, cost and carbon footprint depending on a few key inputs such as building location, use, number of stories and structural system. For new projects, the tool could also suggest structural systems to optimize certain parameters, such as material quantities and carbon footprint. This data could also be combined with a photo recognition system such that taking a photograph of a building would provide the user with an estimation of the structural system, material quantities, construction time and carbon footprint.

SOM is also studying potential applications of AI in design. Using data from SOM's WT 260 wind tunnel in Chicago, SOM is training a machine learning model to recognize how different tall building shapes impact wind effects. Once tested and vetted, this tool could help designers quickly evaluate numerous massing options before testing a select few in the wind tunnel. The tool could also be used to propose new massing options and be combined with other models to optimize shapes for multiple parameters such as wind effects, solar energy and material quantities.

References

Abadi, M. et al., 2016, *Tensorflow: a system for large-scale machine learning,* Operating Systems Design and Implimentation (OSDI), Savannah, Goergia.

Chen, L.-C. et al., 2016, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *Institute of Electrical and Electronics Engineers (IEEE) Transactions on Pattern Analysis and Machine Intelligence,* Volume 40, pp. 834-848.

Diaz, A., Murren, P. & Walker, S., 2017, *Preliminary Observations in the Aftermath of the September 19, 2017 Puebla-Morelos Earthquake,* Skidmore, Owings & Merrill LLP, San Francisco, California.

Earthquake Engineering Research Institute (EERI), 2017, *Puebla, Mexico Earthquake Clearinghouse,* [Online] Available at: www.eqclearinghouse.org/2017–09–19–puebla–mexico/
[Accessed September 2017].

Google, 2018, *Google Cloud Vision,* [Online] Available at: https://cloud.google.com/vision/
[Accessed 15 August 2018].

He, K., Zhang, X., Ren, S. & Sun, J., 2016, *Deep residual learning for image recognition,* Proceedings of the Institute of Electrical and Electronic Engineers (IEEE) Conference on Computer Vision and Pattern Recognition, pp. 770-778, Las Vegas, Nevada.

Ren, S., Kaiming, H., Ross, G. & Jian, S., 2015, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems,* pp. 91-99.

Sarkisian, M. et al., 2018, *Reinforcement of Masonry Vaults for Seismic Loading,* Structures Congress 2018, Fort Worth, Texas.

United Nations, Department of Economic and Social Affairs, Population Division, 2017, *World Population Prospects: The 2017 Revision, Key Findings and Advance,* Working Paper No. ESA/P/WP/248, New York, New York.