Creating a Dataset to Boost Civil Engineering Deep Learning Research and Application

Murad Al Qurishee[1], Weidong Wu[2], Babatunde Atolagbe[3], Joseph Owino[4], Ignatius Fomunung[5], Mbakisya Onyango[5]

[1]Graduate Transportation Associate, Tennessee Department of Transportation, Nashville, Tennessee, USA
[2]University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA
[3]Transportation Engineer, Maryland Department of Transportation, Baltimore, Maryland, USA
[4]UC Foundation Professor and Department Head, Department of Civil & Chemical Engineering, The University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA
[5]Department of Civil & Chemical Engineering, The University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA
Email: mgp216@mocs.utc.edu, Weidong-Wu@utc.edu, zxc259@mocs.utc.edu, JosephOwino@utc.edu, Ignatius-Fomunung@utc.edu, MbakisyaOnyango@utc.edu

Abstract

With cutting edge deep learning breakthrough, numerous innovations in many fields including civil engineering are stimulated. However, a fundamental issue that civil engineering research community currently facing is lack of a publicly available, free, quality-controlled and human-annotated large dataset that supports and drives civil engineering deep learning research and applications on such as intelligent transportation including connected vehicle, structural health monitoring, and bridge inspection. This paper is a general discussion about demanding needs and construction of a long-anticipated dataset for researchers and engineers in civil engineering and beyond for providing critical training, testing and benchmarking data. The establishment of such a free dataset will remove a major hurdle and boost deep learning research in civil engineering and we hope this work will urge researchers, engineers, government agencies and even computer scientists to work together to start building such datasets. A framework has been developed for the proposed database. Also, some pilot study databases were developed for concrete crack detection, pavement crack detection using normal and infrared thermography, as well as pedestrian and bicyclist detection. A convolution neural network model called Faster RCNN was deployed to check the detection accuracy and a 98% detection accuracy of the proposed datasets was obtained.

Keywords

Dataset, Deep Learning, AI, Civil Engineering, ITS, Transportation

1. Introduction

With the breakthrough of deep learning due to advances of hardware such as GPU and Google cloud TPU chip [1], available large datasets as ImageNet [2] and benchmarks and algorithm improvements such as better activation functions, better weight-initialization schemes and better optimization schemes [3], numerous innovations have been unveiled and researchers from various communities are being excited. Civil engineering researchers have applied this novel computer technology to damage detection of structure [4] [5], concrete crack [6] [7], bridge structural components [8], pavement [9], tunnel [10], transmission tower [11] and roof [12]. Some of the most significant commercial products built on deep learning are Apple HomePod®, Amazon Echo®, and Google Home®.

One big force that drives the advance in computer vision, machine learning and deep learning is publicly available large and high-quality datasets that stand behind various competitions hosted by Kaggle platform that challenges data science and Artificial Intelligence. One example of such competition is Image Large Scale Visual Recognition Challenge (ILSVRC) that is hosted annually in order to challenge computer algorithms for object localization/detection from images and videos [13].

Data has been playing an irreplaceable role behind recent explosions and boom in artificial intelligence (AI). The key element or foundation of deep learning application and research is the dataset that supports training and testing deep neural networks and learning skills. In computer vision and machine learning (ML) research community, there are openly available datasets like ImageNet [2] that supported the pioneering deep learning publication AlexNet [14] and spawned the current AI boom, MS COCO [15] for objection, segmentation and captioning, CIFAR-10 [16], a labelled subset of 80 million tiny images that support ML [17], and MNIST [18] for handwritten digits. It took tremendous efforts, time and investment to build such large-scale datasets. For example, MS COCO researchers utilized more than 70,000 worker hours [15] to label and annotate millions of images and object instances. The data is so important to AI and deep learning, that it was reported very recently in June 2018 that the President's most senior technology advisor claimed that the White House may consider releasing some government data to push AI research [19].

However, the open dataset is of limited value to civil engineering research community. Most of the research papers in civil engineering fields used their proprietary datasets which are generally quite small compared to the above general-purpose datasets. While the success of deep learning owes to availability of large-scale labeled data [20], it would be less advantageous if only small datasets were used. Sun *et al.* [20] concluded that the performance of vision tasks increases logarithmically if the size of dataset gets large. If research data could be widely shared, a larger dataset would be created with fewer efforts. But researchers are reluctant to share their data due to legal issues and many other barriers

[21].

Various techniques have been proposed to solve the issue of small datasets for training and testing deep neural network. The creator of Keras deep learning library Chollet [22] proposed a few simple techniques to address small dataset for image classification problem, which includes data augmentation, and transfer learning [23]. Nevertheless, transfer learning only works for deep learning when the model features extracted from a large dataset from pre-training are general. In this sense, use for example a pre-trained ConvNet on ImageNet for transfer learning of crack and damage of concrete is still questionable and investigation must be done to properly evaluate the feasibility of using this technique for civil engineering applications purpose.

In the rest of the paper, a reader will expect the discussion of the application of AI in Civil Engineering research and how important a good database makes the AI techniques more efficient for the users. In the following section, a dataset structure is proposed and discussed how it can improve the traditional database limitations. At the end, some proposed databases were developed and tested using convolution neural network (CNN) models.

2. Civil Engineering Research and Application Needs

Modern computer technologies are making civil engineering, one of the oldest engineering disciplines, smarter and more intelligent. Worldwide cities and governments are launching smart cities [24] initiative. Intelligent Infrastructure can be one of the key characteristics that supports smart city initiatives. Intelligent infrastructure addresses intelligent transportation [25] [26], smart buildings and structures [27] [28], smart bridges and tunnels [29], smart pavement monitoring system [30], and etc.

The USDOT's Intelligent Transportation Systems (ITS) ITS Strategies Plan 2015-2019 outlined the goal of "Realizing Connected Vehicle (CV) Implementation" and "Advancing Automation" as the primary technological drivers of ITS [31]. Connected vehicle applications build an interoperable wireless communication network through dedicated short-range communications (DSRC) [32] or possibly on the deploying 5G telecom network in the near future, which collects vehicles, infrastructure (as traffic lights), and wireless devices (as cell phones) to prevent vehicle crash, improve mobility by reducing delay and congestion and benefit environment by cutting emission. In September 2016, US DOT selected New York City, Tampa-Hillsborough and Wyoming to launch its Connected Vehicle Pilot Deployment Program, which serves an initial effort to deploy, and test the cutting-edge CV technology [33]. Connected vehicle research has still been its early stage, which relies and demands sizable datasets that can serve benchmarking data for developing critical algorithms. Valuable, publicly available, selected and well-organized data related will be always appreciated in order to help advance fundamental knowledge and provide strong and timely support to connected vehicle research.

Allocation of Federal transportation funds and transportation infrastructure management and planning requires traffic monitoring, vehicle count, and classification. Very affordable video devices and State-of-the-art deep neural network based computer techniques/algorithms such as YOLO [34], and Faster R-CNN [35] can provide simple, manageable, cost-effective, real-time solution to any type of traffic counting and classification problem. A large dataset with images of all types of motorized and nonmotorized vehicles, pedestrians and bicyclists will ensure a well-trained deep neural network for detecting classifying and counting traffic.

Structural health monitoring [36] to buildings, bridges, tunnels, and any other civil infrastructure will provide a real-time and preventative strategy to identify and monitor potential damage to a structure. Application of wireless sensor network and wireless smart sensors [37] is the recent trends and future of civil infrastructure health monitoring. Trained deep neural network can be a strong candidate to automatically process the collected ambient vibrations, wind, strain, displacement data for structural damage detection and condition health assessment. However, training a deep neural network calls for a reliable large dataset. Other potential application of deep neural network can be dam and nuclear power plant concrete structure health monitoring, which may safeguard welfare and lives of hundreds of thousands of citizens.

Another critical civil engineering need is bridge inspection automation. With more than 56,000 or 9.1% structurally deficient bridges in the US, more and more bridges may require even shorter inspection interval than the basic 2-year requirement, which means more inspection and maintenance efforts, and higher costs and more dangerous works. With affordable Unmanned Aerial Vehicle (UAV) and deep learning computer technology, it is a trend to partially or completely replace vision inspection. Many civil engineering researchers including the authors of this article and computer scientists are working enthusiastically towards automation of bridge inspections. Recently we proposed a framework of coupling UAV and deep learning for civil infrastructure condition assessment. One of the major challenges civil engineering community face in applying deep learning in bridge inspection is shortage of an image dataset that has good representation of all bridge components to be inspected [38].

Deep learning has also been attempted to solve time-series based real-world applications. Researchers recently applied this computer technology to predict and forecast traffic flow [39] [40] in order to study traffic congestion and delay, and estimate building energy consumption [41] [42] to eventually develop smart grids. The traffic data used in these studies for training and testing are obtained from open data portal provided by State DOTs. And the electric power consumption data was collected from an individual residential customer. Not many publications have been found in these research areas partly due to insufficient quality data.

In summary, deep learning can transform many aspects of conventional civil engineering: structural health monitoring of important infrastructure in damage

detection automation, bridge inspection, intelligent transportation and connected vehicle, road condition assessment, traffic counting and planning, traffic flow prediction and building energy consumption estimation. We anticipate more and more exciting and meaningful applications of deep learning in civil engineering with large dataset and advanced deep neural network architectures available soon.
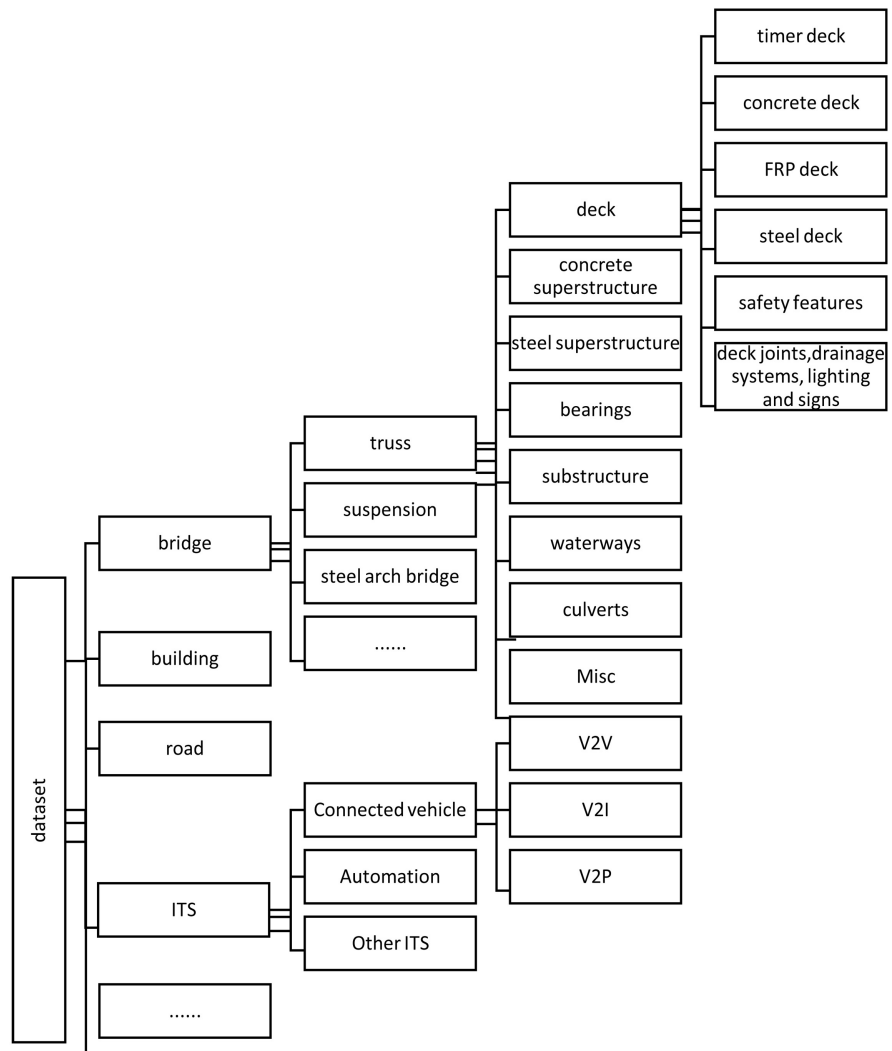
3. The Proposed Dataset

This section presents the difference of the proposed dataset from the existing ones, what should be included in the proposed dataset, how it can possibly be organized, the data collection methods and finally tools to build such a dataset.

3.1. Difference from Existing Datasets

The image datasets ImageNet [2] built by researchers from Princeton University and MS COCO [15] created by Microsoft provide good examples on how such datasets can be constructed, organized and made available to the civil engineering community and beyond. We can generally follow in the footprints of ImageNet and MS COCO. Nevertheless, ImageNet currently offers only images while the proposed dataset expects to include time-series data such as traffic flow, building energy consumption, connected vehicles data as well. The time-series data such as connected vehicle data [43] including basic safety message information as position, motion, vehicle size, road coefficient of friction, lights status and so on can be huge, which makes collect and store data very expensive. Furthermore, dataset as ImageNet hosts mostly daily life objects other than discipline specific targets. The needs for civil engineering applications are far beyond recognition and classification of daily life objects. Workers to label such images may need to receive Safety Inspection of In-Service Bridges training provided by FHWA National Highway Institute in order to produce quality labels for images to be used for bridge inspection. In addition, data in the proposed dataset rely on not only internet and contribution from researchers, but also government agencies. Even though government open data site data.gov offers tremendous amount of data, finding the right data for state-of-the-art in civil engineering deep learning research is without much luck.

3.2. Data Structure

The popular ImageNet was created with a hierarchical structure according to WordNet [44] that is a large lexical database of English words grouped into synonym set or synsets. Compared to ImageNet, the proposed dataset is more discipline (civil engineering) specific, it would be more appropriate to organize the dataset referring to government published documents, national standards or widely accepted classifications in the discipline. As seen in **Figure 1** that depicts the hierarchical structure of the proposed dataset, subtrees of class bridge are organized following National Bridge Inspection Standards and Bridge Inspector's

Figure 1. The hierarchical structure of the proposed dataset.

Reference Manual [45].

3.3. Data Collection

Data should have good representation and cover a broad range of research and application to serve the cutting-edge deep learning research in civil engineering. Labelling of images can be a daunting task, which may be completed in multiple ways such as web-based annotation LabelMe [46], Amazon Mechanical Turk (AMT) and even with a computer game [47].

Humans who receive special trainings may need to label some types of data such as images for highway bridge inspection.

The following is a list of various methods that may be used to collect data for the proposed dataset.

• Data mining online resources

Internet based data collection [48] is a relatively easy and cheap way to collect a large thus more representative data. Both ImageNet and MS COCO collect

images from internet. It is not a challenging task to use python script to automatically extract and scrape data. However, data collected from internet may be satisfying for daily life applications such as detection and recognition of cats and dogs, they may not always meet the demands of specific civil engineering research and applications. Another concern of using internet-based data is piracy and copyright protection. Fortunately, non-commercial use of data for education, and research is generally allowed. For example, researchers and educators may download images acquired from web by ImageNet for non-commercial and/or educational purchases under certain conditions and terms.

- Image data may be obtained by querying several image search engines such as Google Images, Bing Images, and Flickr. For example, MS COCO collected non-iconic images from Flickr. Kaggle can also be a good source to identify good data.
- Request data from DOT and local state DOTs and other government agencies.

This could be the best approach to obtain high quality data. The Freedom of Information Act (FOIA) [49] is a Federal law that gives individuals the right to access to any US federal agency records unless the agencies the release is prohibited by law or protected by nine exemptions, which means we may not necessarily be able to request all the data of interest.

There are tremendous amount of data and we first need to identify and decide what are the most valuable data to be requested that can be potentially used for deep learning studies. One example of valuable data can be images taken by bridge inspectors owned by state DOTs. These images may serve high quality data for supervised training and testing deep neural network for bridge inspection purpose.

- Collect and archive as traffic and data from publicly available open data portal provided by State DOTs and other agencies. Table 1 shows a few examples of publicly available free traffic data portal provided by state DOTs. The proposed dataset website should extract data from those data portals, organize and classify for deep learning traffic flow research use, which may need heavy involvement of data cleansing to better serve the deep learning research needs.
- Promote and encourage share of data.

High quality data is the heart of any research work and excellent data builds the best possible foundation for deep neural network related publications. However, individual author may have their small datasets and look for even larger dataset for their use. "Take one, return one" (a researcher may download data if they contribute) may encourage share of data among peer researchers.

- Launch competitions based on the proposed dataset to help advance and develop better algorithms for civil engineering deep learning research and application.

Kaggle sets an example of providing predictive modeling competition platform to solve a wide variety of problems in different fields of computer science,

Table 1. Selected traffic data portals provided by state DOTs [7] [38] [50]-[63].

Data Portal	States and Agencies	Comments
PeMS [64]	CA	Real-time data by 40,000 detectors across CA. Archived historical data with a wide variety of information available; free
NYDOT Real Time Traffic Information	NY City	Real-time traffic cameras; free
Smart Way [65]	TN DOT	four transportation management centers, 517 cameras, 174 message signs, 1015 roadway detection systems and 49 video detection systems in the four largest cities; real-time traffic cameras
TADA	GA DOT	Dynamic mapping interface and other data formats to present data collected
ITS Public Data Hub [66]	US DOT	ITS data

computer vision, medicine etc. With emerging technologies such as connected vehicle, numerous challenges have been remaining and competitions based on the proposed datasets will lead to better solutions, algorithms and produce high quality journal and conference publications, eventually accelerate AI and deep learning in civil engineering.

3.4. Construction and Maintenance

We can build an online dataset using the open-source data portal platform CKAN that allows easy data storage, distribution and share. CKAN is being used by public institutions [67] and government data catalogues, such as Data.gov and HealthData.gov in the US, data.gov.uk in the UK, and many others [68]. Construction and maintenance of such a dataset will need support from research grants, and donations.

4. Building and Testing a Pilot Study Database

A small dataset for concrete crack detection was built with 1499 concrete-crack images and 589 concrete-not-crack images. Figure 2 shows the sprite image of the proposed database for concrete crack detection. For bicycle detection and counting, a database was created with 988 test images and 4822 train images. The bicycle images were taken from Google using a special data scrapping software tool and the images were labeled and annotated using LabelImg software. Moreover, a dataset was created for pavement crack detection using a 336-test image and 2284 train. The images were collected using a hand-held mobile phone and a drone. A total of 11 categories of flexible pavement crack images and 7 types of rigid pavement crack images were included in this dataset. The database images were annotated and labeled using LabelImg software with more than 50 hours of manual labors. Additionally, an infrared thermography dataset
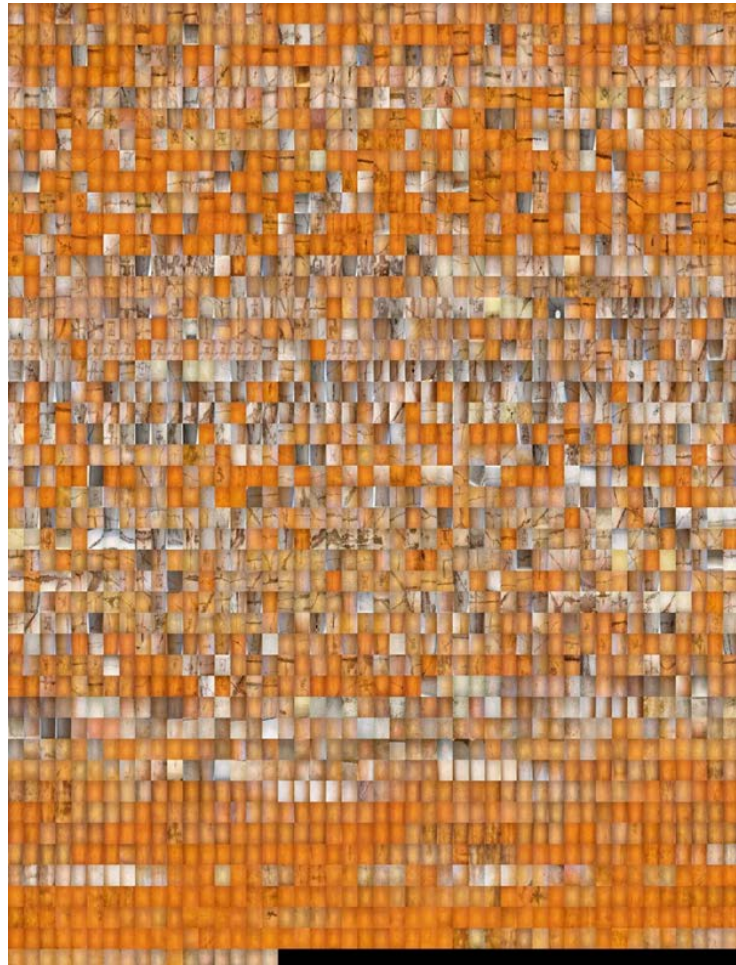
Figure 2. Sprite image of 1642 images of crack and not crack database.

was created with 24 test and 84 train infrared thermography images **Figure 3** shows the sprite images of the pavement crack detection images with infrared thermography images. The databases were uploaded in a Google share drive (http://bit.ly/2ujAhMd).

All the developed databases were tested using a deep learning convolution neural network model called Faster RCNN. The description of the test parameters and the procedures are out of the scope of this paper. The reader can find more depth in-formation of the model selection and training in this article [7]. However, the Faster RCNN model successfully detects the pavement crack using normal images (**Figure 4**) and infrared thermography crack images (**Figure 5**), as well as the pedestrian and bicyclist images (**Figure 6**). All the test images from the proposed pilot study databases show a 98% confidence level which means the database annotation and labeling are in the right direction.

5. Conclusion and Discussion

Big dataset is like fuel to engine that delivers power to the civil engineering AI research plane. A publicly available, free and labelled dataset is to address a
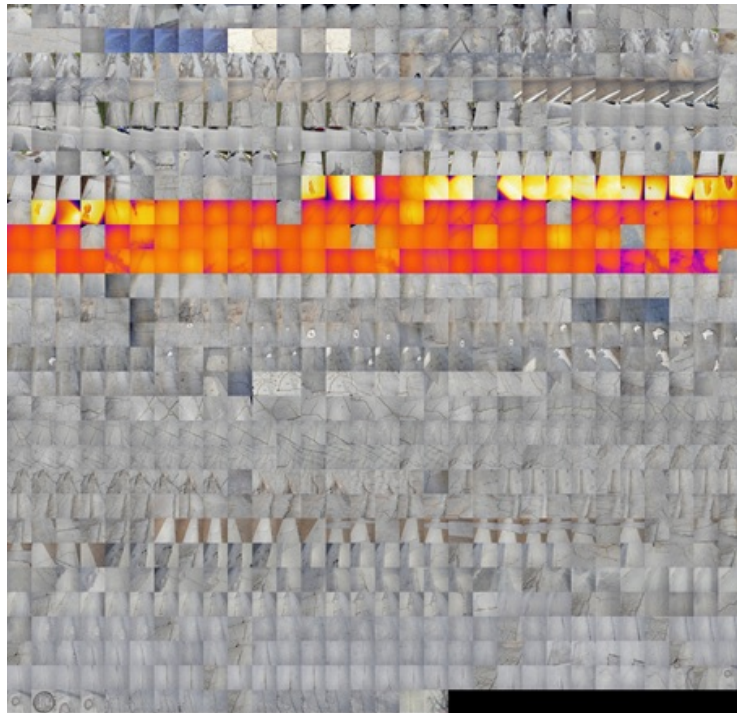
Figure 3. Sprite images of 2728 image data for pavement crack detection including infrared and normal images.
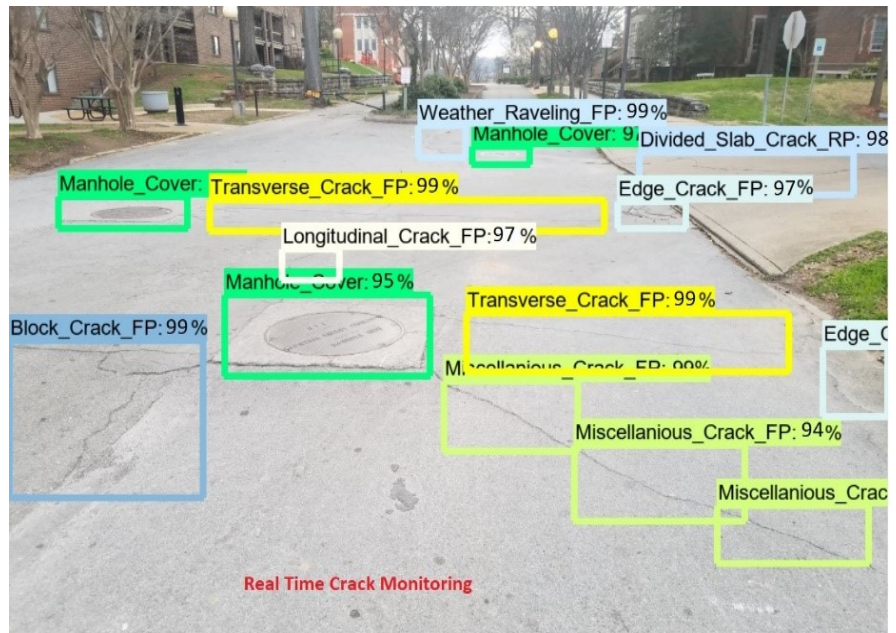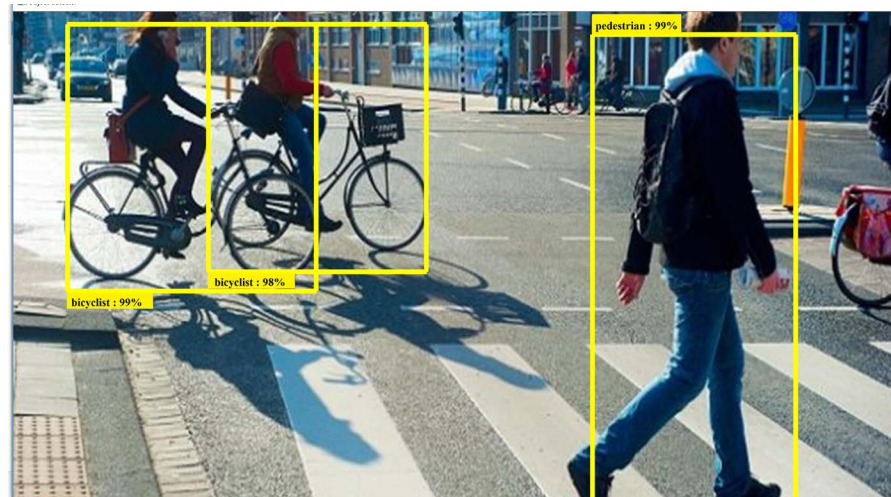

Figure 4. Pavement crack detection using Faster RCNN with 98% confidence.

fundamental issue of advancing deep learning research and application in civil engineering and beyond: high quality data. It would be a tremendous help to civil researchers to build their innovative and cutting-edge works in intelligent transportation, connected vehicle, structural health monitoring, bridge inspection, and more real-life applications on the proposed dataset. Our pilot study
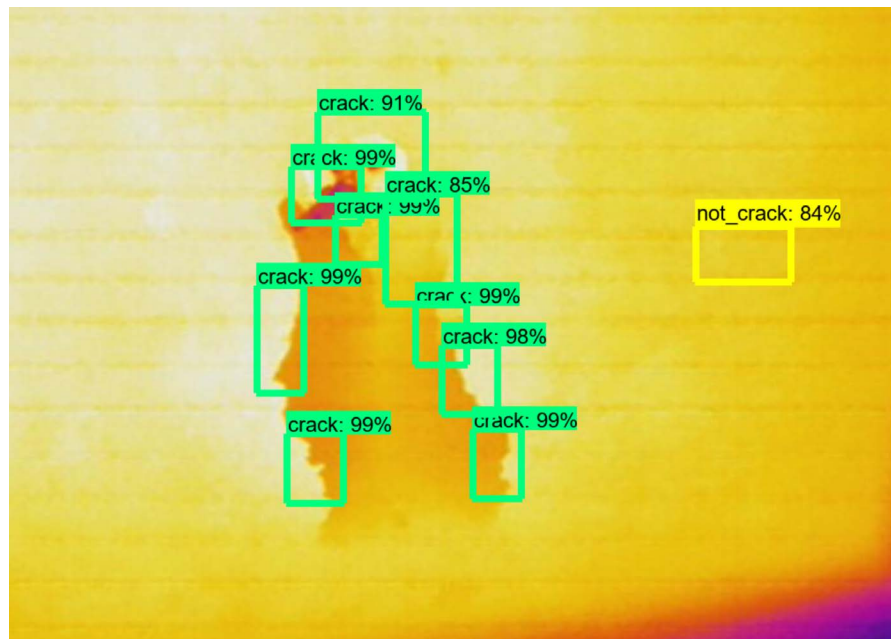
Figure 5. Concrete crack detection using faster RCNN with 98% confidence.

Figure 6. Pedestrian and bicyclist detection using faster RCNN with 98% confidence.

shows some of the proposed datasets for concrete crack detections, pavement crack detections as well as pedestrian and bicyclist detections with 98% confidence level.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Sato, K., Young, C. and Patterson, D. (2017) An In-Depth Look at Google's First Tensor Processing Unit (TPU). Google Cloud Big Data and Machine Learning Blog,

12.

[2] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L. (2009) ImageNet: A Large-Scale Hierarchical Image Database. 2009 *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, 20-25 June 2009, 248-255. https://doi.org/10.1109/CVPR.2009.5206848

[3] Chollet, F. (2017) Deep Learning with Python. Manning Publications Co., New York.

[4] Lin, Y.Z., Nie, Z.H. and Ma, H.W. (2017) Structural Damage Detection with Automatic Feature-Extraction through Deep Learning. *Computer-Aided Civil and Infrastructure Engineering*, **32**, 1025-1046. https://doi.org/10.1111/mice.12313

[5] Gulgec, N.S., Takáč, M. and Pakzad, S.N. (2017) Structural Damage Detection Using Convolutional Neural Networks. In: *Model Validation and Uncertainty Quantification*, Volume 3, Springer, New York, 331-337.
https://doi.org/10.1007/978-3-319-54858-6_33

[6] Cha, Y.J., Choi, W. and Büyüköztürk, O. (2017) Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Computer-Aided Civil and Infrastructure Engineering*, **32**, 361-378. https://doi.org/10.1111/mice.12263

[7] Qurishee, M.A. (2019) Low-Cost Deep Learning UAV and Raspberry Pi Solution to Real Time Pavement Condition Assessment.

[8] Cha, Y.J., Choi, W., Suh, G., Mahmoudkhani, S. and Büyüköztürk, O. (2017) Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types. *Computer-Aided Civil and Infrastructure Engineering*, **33**, 731-747. https://doi.org/10.1111/mice.12334

[9] Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T. and Omata, H. (2018) Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone. arXiv Preprint arXiv:1801.09454.

[10] Makantasis, K., Protopapadakis, E., Doulamis, A., Doulamis, N. and Loupos, C. (2015) Deep Convolutional Neural Networks for Efficient Vision Based Tunnel Inspection. 2015 *IEEE International Conference on Intelligent Computer Communication and Processing* (*ICCP*), Cluj-Napoca, Romania, 3-5 September 2015, 335-342. https://doi.org/10.1109/ICCP.2015.7312681

[11] Varghese, A., Gubbi, J., Sharma, H. and Balamuralidhar, P. (2017) Power Infrastructure Monitoring and Damage Detection Using Drone Captured Images. 2017 *International Joint Conference on Neural Networks* (*IJCNN*), Anchorage, AK, 14-19 May 2017, 1681-1687. https://doi.org/10.1109/IJCNN.2017.7966053

[12] Wang, F., Kerekes, J.P., Xu, Z. and Wang, Y. (2018) Residential Roof Condition Assessment System Using Deep Learning. *Journal of Applied Remote Sensing*, **12**, Article ID: 016040. https://doi.org/10.1117/1.JRS.12.016040

[13] Russakovsky, O., *et al*. (2015) ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, **115**, 211-252.
https://doi.org/10.1007/s11263-015-0816-y

[14] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) Imagenet Classification with Deep Convolutional Neural Networks. In: *Advances in Neural Information Processing Systems*, Springer, New York, 1097-1105.

[15] Lin, T.-Y., *et al*. (2014) Microsoft Coco: Common Objects in Context. In: *European Conference on Computer Vision*, Springer, New York, 740-755.
https://doi.org/10.1007/978-3-319-10602-1_48

[16] Krizhevsky, A., Nair, V. and Hinton, G. (2014) The CIFAR-10 Dataset.

https://www.cs.toronto.edu/~kriz/cifar.html

[17] Torralba, A., Fergus, R. and Freeman, W.T. (2008) 80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30**, 1958-1970.
https://doi.org/10.1109/TPAMI.2008.128

[18] Y. LeCun, C. Cortes, and C. Burges (2010) MNIST Handwritten Digit Database. AT & T Labs. Volume 2. http://yann.lecun.com/exdb/mnist

[19] Knight, W. (2018) The White House Promises to Release Government Data to Fuel the AI Boom.
https://www.technologyreview.com/s/611331/the-white-house-promises-to-release-government-data-to-fuel-the-ai-boom/

[20] Sun, C., Shrivastava, A., Singh, S. and Gupta, A. (2017) Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. 2017 *IEEE International Conference on Computer Vision* (*ICCV*), Venice, Italy, 22-29 October 2017, 843-852.
https://doi.org/10.1109/ICCV.2017.97

[21] Downey, A.S. and Olson, S. (2013) Sharing Clinical Research Data: Workshop Summary. National Academies Press, Washington DC.

[22] Chollet, F. (2016) Building Powerful Image Classification Models Using Very Little Data. Volume 13.

[23] Yosinski, J., Clune, J., Bengio, Y. and Lipson, H. (2014) How Transferable Are Features in Deep Neural Networks? In: *Advances in Neural Information Processing Systems*, Springer, New York, 3320-3328.

[24] Chourabi, H., *et al.* (2012) Understanding Smart Cities: An Integrative Framework. 2012 45*th Hawaii International Conference on System Science* (*HICSS*), Maui, HI, 4-7 January 2012, 2289-2297. https://doi.org/10.1109/HICSS.2012.615

[25] Figueiredo, L., Jesus, I., Machado, J.T., Ferreira, J.R. and De Carvalho, J.M. (2001) Towards the Development of Intelligent Transportation Systems. *Intelligent Transportation Systems*, 2001. *Proceedings*, Oakland, CA, 25-29 August 2001, 1206-1211. https://doi.org/10.1109/ITSC.2001.948835

[26] Dimitrakopoulos, G. and Demestichas, P. (2010) Intelligent Transportation Systems. *IEEE Vehicular Technology Magazine*, **5**, 77-84.
https://doi.org/10.1109/MVT.2009.935537

[27] Adeli, H. (2008) Smart Structures and Building Automation in the 21st Century. *International Symposium on Automation in Construction*, **25**, 5-10.
https://doi.org/10.3846/isarc.20080626.5

[28] Udd, E. (1993) Fiber Optic Smart Structures. *Society of Photo-Optical Instrumentation Engineers* (*SPIE*) *Conference Series*, Volume 10266.
https://doi.org/10.1117/12.145196

[29] Stajano, F., Hoult, N., Wassell, I., Bennett, P., Middleton, C. and Soga, K. (2010) Smart Bridges, Smart Tunnels: Transforming Wireless Sensor Networks from Research Prototypes into Robust Engineering Infrastructure. *Ad Hoc Networks*, **8**, 872-888. https://doi.org/10.1016/j.adhoc.2010.04.002

[30] Lajnef, N., Chatti, K., Chakrabartty, S., Rhimi, M. and Sarkar, P. (2013) Smart Pavement Monitoring System. United States Federal Highway Administration.

[31] Barbaresso, J., Cordahi, G., Garcia, D., Hill, C., Jendzejec, A. and Wright, K. (2014) USDOT's Intelligent Transportation Systems (ITS) ITS Strategic Plan 2015-2019.
https://doi.org/10.1109/JPROC.2011.2132790

[32] Kenney, J.B. (2011) Dedicated Short-Range Communications (DSRC) Standards in

the United States. *Proceedings of the IEEE*, **99**, 1162-1182.

[33] Hartman, K.K. Connected Vehicle Pilots. https://www.its.dot.gov/pilots/index.htm

[34] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779-788. https://doi.org/10.1109/CVPR.2016.91

[35] Ren, S., He, K., Girshick, R. and Sun, J. (2015) Faster r-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In: *Advances in Neural Information Processing Systems*, Springer, New York, 91-99.

[36] Farrar, C.R. and Worden, K. (2007) An Introduction to Structural Health Monitoring. *Philosophical Transactions of the Royal Society of London A*: *Mathematical, Physical and Engineering Sciences*, **365**, 303-315. https://doi.org/10.1098/rsta.2006.1928

[37] Kim, S., *et al*. (2007) Health Monitoring of Civil Infrastructures Using Wireless Sensor Networks. In: *Proceedings of the* 6*th International Conference on Information Processing in Sensor Networks*, ACM, New York, 254-263. https://doi.org/10.1145/1236360.1236395

[38] Wu, W., Qurishee, M.A., Owino, J., Fomunung, I., Onyango, M. and Atolagbe, B. (2018) Coupling Deep Learning and UAV for Infrastructure Condition Assessment Automation. 2018 *IEEE International Smart Cities Conference* (*ISC*2), Kansas City, MO, 16-19 September 2018, 1-7. https://doi.org/10.1109/ISC2.2018.8656971

[39] Polson, N.G. and Sokolov, V.O. (2017) Deep Learning for Short-Term Traffic Flow Prediction. *Transportation Research Part C: Emerging Technologies*, **79**, 1-17. https://doi.org/10.1016/j.trc.2017.02.024

[40] Lv, Y., Duan, Y., Kang, W., Li, Z. and Wang, F.-Y. (2015) Traffic Flow Prediction with Big Data: A Deep Learning Approach. *IEEE Transactions on Intelligent Transportation Systems*, **16**, 865-873.

[41] Mocanu, E., Nguyen, P.H., Gibescu, M. and Kling, W.L. (2016) Deep Learning for Estimating Building Energy Consumption. *Sustainable Energy, Grids and Networks*, **6**, 91-99. https://doi.org/10.1016/j.segan.2016.02.005

[42] Li, C., Ding, Z., Zhao, D., Yi, J. and Zhang, G. (2017) Building Energy Consumption Prediction: An Extreme Deep Learning Approach. *Energies*, **10**, 1525. https://doi.org/10.3390/en10101525

[43] NHTS Administration (2017) Federal Motor Vehicle Safety Standards; V2V Communications. *Federal Register*, **82**, 3854-4019.

[44] Miller, G.A. (1995) WordNet: A Lexical Database for English. *Communications of the ACM*, **38**, 39-41. https://doi.org/10.1145/219717.219748

[45] Ryan, T.W., Hartle, R.A., Mann, J.E. and Danovich, L.J. (2006) Bridge Inspector's Reference Manual. Report No. FHWA NHI, 03-001.

[46] Russell, B.C., Torralba, A., Murphy, K.P. and Freeman, W.T. (2008) LabelMe: A Database and Web-Based Tool for Image Annotation. *International Journal of Computer Vision*, **77**, 157-173. https://doi.org/10.1007/s11263-007-0090-8

[47] Von Ahn, L. and Dabbish, L. (2004) Labeling Images with a Computer Game. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, New York, 319-326. https://doi.org/10.1145/985692.985733

[48] Benfield, J.A. and Szlemko, W.J. (2006) Internet-Based Data Collection: Promises and Realities. *Journal of Research Practice*, **2**, 1.

[49] Kramer, V.H. and Weinberg, D.B. (1974) The Freedom of Information Act. *The*

Georgetown Law Journal, **63**, 49.

[50] Caltrans. Performance Measurement System (PeMS). http://pems.dot.ca.gov

[51] Dot, T. Smartway. https://smartway.tn.gov/traffic/

[52] GDOT. The Georgia Department of Transportation's Traffic Analysis and Data Application. https://gdottrafficdata.drakewell.com/publicmultinodemap.asp

[53] UDO Transportation. ITS Public Data Hub. https://www.its.dot.gov/data/

[54] Qurishee, M., Iqbal, I., Islam, M. and Islam, M. (2016) Use of Slag as Coarse Aggregate and Its Effect on Mechanical Properties of Concrete. *Proceedings of International Conference on Advances in Civil Engineering*, **3**, 475-479.

[55] Al Qurishee, M. (2017) Application of Geosynthetics in Pavement Design.

[56] Al Qurishee, M. and Fomunung, I. (2000) Smart Materials in Smart Structural Systems.

[57] Hasnat, A., Qurishee, M., Iqbal, I., Zaman, M. and Wahid, M. (2018) Effectiveness of Using Slag as Coarse Aggregate and Study of Its Impact on Mechanical Properties of Concrete.

[58] Atolagbe, B. (2019) Automatic Mesh Representation of Urban Environments.

[59] Islam, M.A. (2018) Intergrading Connected Vehicle Data into the Transportation Performance Measurement Process. The University of Alabama, Birmingham.

[60] Islam, M.A. (2019) A Literature Review on Freeway Traffic Incidents and Their Impact on Traffic Operations. *Journal of Transportation Technologies*, **9**, 504-516. https://doi.org/10.4236/jtts.2019.94032

[61] Islam, M.A., Sisiopiku, V.P., Ramadan, O.E. and Hadi, M. (2019) A Framework for Performance-Based Traffic Operations Using Connected Vehicle Data. *Simulation* (*NGSIM*), **6**, No. 8.

[62] Al Qurishee, M., Wu, W., Atolagbe, B., El Said, S. and Ghasemi, A. (2019) Non-Destructive Test Application in Civil Infrastructure.

[63] Al Qurishee, M., Wu, W., Atolagbe, B., El Said, S., Ghasemi, A. and Tareq, S.M. (2008) Wireless Sensor Network and Its Application in Civil Infrastructure.

[64] Cdot (2020) California Department of Transportation. http://pems.dot.ca.gov/

[65] Tdot (2020) SmartWay Traffic. https://smartway.tn.gov/Traffic?position=-85.97076699999997,35.88416963357008,7&features=incidents,traffic

[66] Usdot (2020) ITS DataHub. https://www.its.dot.gov/data/

[67] Winn, J. (2013) Open Data and the Academy: An Evaluation of CKAN for Research Data Management.

[68] OK International. https://ckan.org/