The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Sparse Damage Per-pixel Prognosis Indices via Semantic Segmentation

Takato Yasuno*1

*1 Research Institute for Infrastructure Paradigm Shift (RIIPS)

Efficient inspection and accurate prognosis are required for civil infrastructures with more than 30 years since completion. If we can detect damaged photos automatically per-pixels from the record of the inspection record and countermeasure classification of drone inspection vision, then it is possible that countermeasure information can be provided more flexibly, whether we need to repair and how large the expose of damage interest. A piece of damage photo is often sparse as long as it is not zoomed around damage, exactly the range where the detection target is photographed, is at most only one percent. In this paper, we propose three damage detection methods of transfer learning which enables semantic segmentation in an image with low pixels using damaged photos of drone inspection. Furthermore, we propose prognosis indices to make a decision repair-priority such as the counts index of pop-outs region and the per-pixel area counts index of each pop-out based on morphology image processing. In fact, we show the results applied this method using the 40 drone inspection images whose size is 6,000 x 4,000 on an infrastructure, where each image is partitioned into 400 crops, so the total number of input images is 16,000 for training deep neural network. Finally, future tasks of damage detection modeling are mentioned (211words).

1. Introduction

Deterioration of civil engineering structures is progressing in recent years, including a large number of concrete structures. Improving efficiency of scheduled inspections is a pressing issue, since the cost of inspections comprises a large proportion of maintenance costs for local governments, which are also experiencing manpower shortage for technical personnel. There are often opportunities to apply deep learning as a method for improving efficiency of inspections on social infrastructure and studies have been conducted on this issue. Dam general inspection is required for dam once every 30 years and as a result, images of damage have been accumulating (Ministry of Land, Infrastructure, 2013). If it were possible to utilize images of damage that are attached to inspection reports, data from scheduled inspections from past years can be input for the purpose of deterioration learning. If it could be possible to automatically calculate numerical scores for the extent of damage based on images of damage, this would be useful in deciding whether any repairs work should be performed and for setting the order of priority among candidates for repairs. There are past studies on detecting cracks in concrete on bridges, structures, plants, etc.

Especially, for dam structural health monitoring, it is important to prognosis pop-outs owing to be greater impact on the health of dam embankment. In area of low quality aggregates, as a result of the water absorption of the concrete, the soft stone having a high water absorption rate becomes saturated. When the freezing temperature is reached, pressure due to volume expansion occurs. However, the detection model for pop-out is only at its incipient stages, so it would be difficult to claim that this is an established means for concrete damage deterioration learning and prediction. This paper proposes a practical method applies semantic segmentation (segmentation) of concrete damage using images of damage from drone-base inspections. Results are shown from actually applying this method on sparse images of damage, focusing on images of pop-outs among images of damage to dam embankment. Finally, references will be made to issues of damage detection modeling as well.

Contact: Takato Yasuno, RIIPS on 5-20-8, Asakusabashi, Taito-ku, Tokyo, 111-8648, tk-yasuno@yachiyo-eng.co.jp

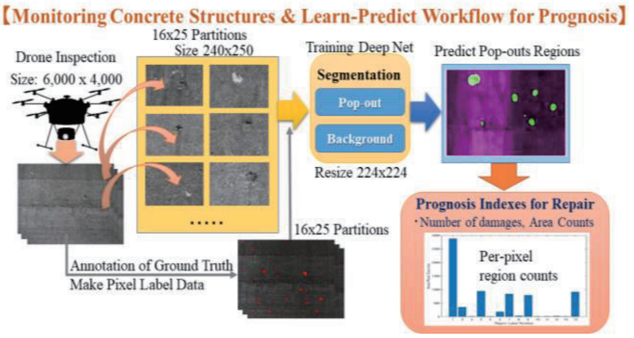

Figure 1: Monitoring concrete structures from drone-base inspection to train segmentation networks and damage prediction for prognosis indices.

2. Related Studies and Damage Images

2.1 Damage detection studies for civil infrastructures

Since 2002, there has been an accumulation of studies (Wu, 2002) (Chun, 2015) on resolving damage detection using neural networks (ANN) for the purpose of continuous surveillance of bridges. Many instances of damage detection modeling for machine learning have been conducted over the past 15 years, including the ANN, as well as the PCA, SVM, GA and other such solution methods (Gordan, 2017). Since the potential of convolutional neural networks (CNN) to exhibit high degrees of accuracy in classifying one million images into 1,000 classes was reported in 2012 (AlexNet, 2012), there has been active reporting of studies on solution methods of the CNN, which provides solutions with greater accuracy than conventional methods for label categorization of overall images, object detection and semantic segmentation at the pixel level. There have been a number of studies conducted on damage classification of at the whole-image level for cracks and corrosion of road pavement, structures and bridges, for detection of damage to civil engineering structures (Gopalakrishnan, 2018) (Ricard, 2018), as well as damage segmentation at the pixel level (Hoskere, 2017). A report was made on a study that applied deep CNN to conduct four classes of damage segmentation, namely no damage, only separation, exposure of rebar (with and without rust), using 734 images of damage (Guillamon, 2018). The breakdown of the damage classes, however, indicated a distribution biased to the third class, for which there were 510 images, and as such, distortion in the training images cannot be denied. Dimensions of the images of damage were widely varied, being 640 x 480, 1,024 x 768, and 1,600 x 1,200. The potential for learning with the index that represents the degree of matching between prediction and reality, mIoU (class mean IoU) to the level of 0.6 to 0.8 was indicated by using some types of CNN models for fully convolutional networks (FCN) in entering images of such diverse dimensions. The use of the damage detection modeling that utilizes solution method of CNN, however, has just been started and as such, it would be difficult to claim that this is an established general-purpose method for damage detection in management of infrastructure. This paper proposes a practical method for damage segmentation with considerations for sparse characteristics of damage images from drone-base inspections. Furthermore, using the output of prediction RGB-images by the trained semantic segmentation, we propose two morphological indices such as the number of identified damages and the per-pixel counts of each damage region for prognosis to make a decision repair-priority.

Table 1: Comparison of the per-pixel counts between the target pop-outs region and the background region.

Example consisting of 40 damage drone inspection images	Total number of pixels per damage image	The number of pixels per image	Percentage per image
Background	954,339,801	23,858,495	99.4%
Damage to region of interest (ROI)	5,660,199	141,505	**0.6%**
Total per image	960,000,000	24,000,000	100.0%

2.2 Characteristics of Damage Images

This paper provides a practical observation on characteristics of images of damage, using 40 images of damage in which pop-out has been captured through drone-base inspection of dam embankment, whose size is 6,000 x 4,000. While generality cannot be guaranteed with these characteristics, they are considered to lead the way for utilizing images of damage. Characteristics of general conditions and damage for pop-out is as follows (CERI, 2016). Pop-out is a crater-like indentation generated by destruction due to the expansion of aggregate particles on the concrete surface. These are often observed in aggregates with high water absorption and in poor quality. Pop out is the meaning of “jumps out suddenly”. In the case of low-quality aggregate, as a result of the water absorption of the concrete, the soft stone having a high water absorption rate becomes saturated. At this time, when the freezing temperature is reached, pressure due to volume expansion occurs, the surface portion peels off, and then a crater-like hole is formed.

Table 1 shows the summary value for the damage area (region of interest: ROI) subject to detection, as well as other regions in the background, counted at pixel level. No advance manipulation was conducted on images to unify photographing distance and picture quality. The number of pixels per image was 24 million pixels. The proportion of these that include targeted damage was only 0.6%. The first characteristic of damage image is the sparsity of the area comprised of ROI.

3. Per-pixel Learning and Prognosis Indices

3.1 Damage segmentation for prediction

The FCN-Alex (Long, 2015), as well as the SegNet-VGG16 (Badrinarayanan, 2016) are compared where appropriate, as a method for learning transfers of semantic segmentation. The solution method used in this paper by itself does not present any innovation but the extremely sparse proportion of detection target ROI on any given image is a characteristic and the intention was to derive a practical method that can be applied to images of damage with sparse pixel labels. The FCN-Alex is a transfer learning of AlexNet and the CNN is implemented to the deepest layer, making it a deep neural net (DNN) of 23 layers in depth. Learning is possible with relatively short calculation time and prediction output for exhaustive detection of targeted damage can be achieved. SegNet-VGG16 is a method of transfer learning used to identify objects for automatic driving and a DNN with depth of 91 layers.

This paper applies the four deep neural networks described above to images of damage to compare calculation execution time, accuracy and prediction output image. There is a problem of no improvements being evident with loss functions when the SGDM is used in the optimization method for hyper parameters, as gradients of the detection target are eliminated due to the sparse characteristic of the damage image. In order to overcome this issue, the gradient of the detection target is captured with good sensitivity and the previously updated quantities are deleted where appropriate, and the RMSProp, which has a characteristic formula for error function that eliminates the amount of change in gradients of detection targets by taking square root of the amount of change in gradient, is adopted (Hinton, 2012) (Mukkamala, 2017). The weighting factor for the updating amount was set to 0.99. The learning coefficient for the overall model was set to 1E-5 and the minibatch was set to 32.

3.2 Morphological indices for prognosis

The word morphology commonly denotes a branch of biology that deals with the form and structure of animals and plants. We use the same word here in the context of mathematical morphology as a toll for extracting image components that are useful in the representation of region shape. We are interested also in morphological techniques for pre- or post-processing, such as morphological filtering, thinning, and pruning. In image-processing applications, dilation and erosion are used most often in various combinations (Serra 1992; Gonzalez 2008). This paper proposes some prognosis indices to make a decision repair-priority such as the counts index of pop-outs region and the per-pixel area counts index of each pop-out based on morphological image processing, such as dilation and erosion operation.

On the prediction of pop-out damage segmentation, there are some extremely small size of pop-outs, so that we may overlook them. Also, the shape of pop-outs are not always like circle, but these are complex shape or partially connected with various size of pop-outs. This paper proposes two practical morphological operation, dilation and erosion, in terms of the union (or intersection) of an image with a translated shape called structuring element. At first, we translate the prediction RGB-image of pop-out segmentation into a binary mask image with pop-out foreground (1-valued pixel, white color) and with background (0-valued pixel, black color). Dilation is an operation that "grows" or "thickens" objects in the extremely small images of pop-outs. This growing is controlled by a shape referred to as a structuring element, such as linear, disk, octagon etc. This paper proposes the disk-shaped structuring element with radius r=3. Further, erosion "shrinks" or "thins" objects in a binary image like complex shape and partially connected with various size of pop-outs. This paper proposes these morphological operations applied to the masked prediction of pop-outs images. By these operation, it is possible to avoid overlooking the small pop-outs, and we can extract the complex shape or partially connected pop-outs, in order to count the number of pop-out and the each region pixel size more accurately and efficiently.

4. Applied Results

4.1 Training results

The input data was 40 images whose size is 6,000 x 4,000 from drone-base inspections of dam embankment. In order to bring them closer with the input size of deep pre-trained network, we partitioned each original image into 25 x 16 equal 400 crops whose size was 250 x 240. The usage rate of the training and test data was set to Train: Test = 99:1. The transition of loss function in the learning process applied to the pop-out segmentation is shown in Figure 2. The calculation conditions are 490 cycles per epoch for a total of 24,500 repeated calculations in 50 epochs. The loss value of the FCN-AlexNet is transitioning at a lower level than SegNet-VGG16. This FCN models, however, have large dispersion of loss values and their disadvantage is that they make for unstable learning processes. The loss function of the SegNet-VGG16 does not offer minimum values, but up and down fluctuations remain small early on, which can be interpreted to offer superior stability for the learning process.

Table 2: Comparison of indices for pop-out segmentation models.

DNN model	Time calculation	Mean mIoU	Weighted wIoU
FCN-AlexNet	466min.	0.5811	0.9861
SegNet-VGG16	832min.	**0.5967**	0.9856

Table 2 shows the calculation time, accuracy, mean-IoU and weighted-IoU index of respective segmentation model. The FCN-AlexNet offers a relatively short calculation time of 466 minutes. This net achieved the index such as mIoU = 0.5811, and wIoU = 0.9861. Meanwhile, the SegNet-VGG16 offers about two times calculation time compared with FCN-AlexNet, and indicates the score of mIoU of 0.5967 and wIoU of 0.9856. While each weighted wIoUs have almost no difference, but regarding the score of the mean mIoU the SegNet-VGG16 is superior with the FCN-AlexNet to select better pop-out predictor.

4.2 Prediction results

Figure 2 shows an output RGB-image of predictions for a test image whose size is 600 x 400, using the trained SegNet-base predictor of pop-out segmentation. Here, the region of prediction are shown in green color. In contrast, the region of background are shown in magenta color. Figure 3 shows the translated binary mask with pop-out foreground (1-valued) and with background (0-valued). We operated the morphological operations applied to the masked prediction of pop-outs images. Further, we compute the centroid of each pop-out region and set the pop-out number in order to represent the counts index accurately, here the total count of pop-outs is 14. Figure 4 shows the per-pixel counts of each region based on the morphological image pre-processing. Figure 5 shows the bar chart that we can visualize the volume indices of pop-outs and it is possible to compare the largest size, middle size, and extremely small size of pop-outs in order to make a decision of repair-priority for infrastructure manager.

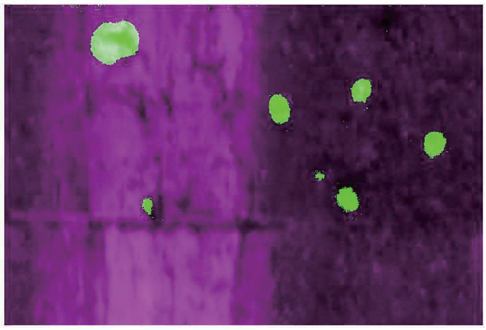

Figure 2: Trained SegNet-base prediction of pop-outs (RGB image) Here, green indicates prediction, magenta denotes background.

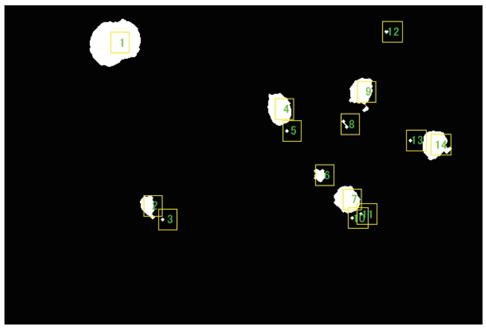

Figure 3: Counts index of identified pop-out centroid based on morphological operations with dilation and erosion.

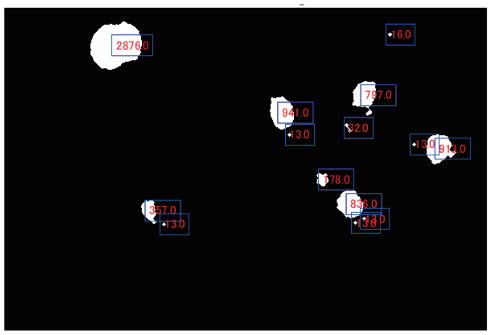

Figure 4: Per-pixel counts of each pop-out prediction region based on morphological image processing.

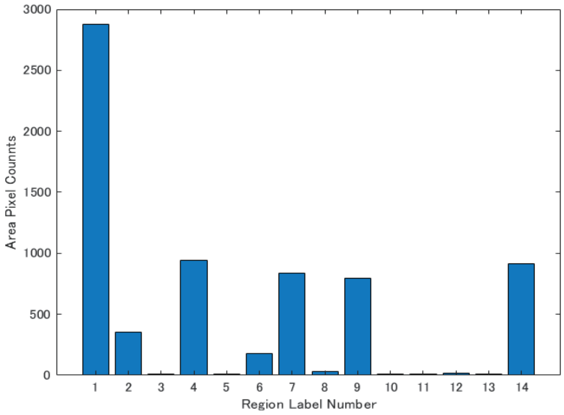

Figure 5: Area pixel count index of each pop-out for prognosis measure indices to make a decision regarding repair-priority.

5. Conclusion

5.1 Concluding remarks

This paper proposed a method for detecting pop-out by semantic segmentation, using images of damage obtained from drone-base inspections. In fact, we show the results applied this method using the 40 drone inspection images at a dam embankment, where each image is partitioned into 400 crops, so the total number of input images is 16,000 for training deep neural network. Based on transfer learning, per-pixel higher accurate prediction is possible, even to sparse damage images whose pop-out ratio per-pixel is only one percent compared with the background. The SegNet-VGG16 exhibited the better accuracy and achieved class mean mIoU index of 59.67% and weighted index wIoU of 98.56%. Furthermore, we demonstrated to compute some morphological indices, such as the counts index of identified pop-outs centroid and the per-pixel area counts index of each pop-out region for prognosis to make a decision repair-priority more accurately and efficiently.

5.2 Future works

The scope of this paper was the segmentation of pop-out for prognosis, using images from drone-base inspection of dam embankment. Monitoring various damages for standard dam inspection prescribes concrete crack, scaling, pop-outs, water leak, efflorescence etc (Ministry of Land, Infrastructure, 2013). In contrast, creation of dataset for training and prediction of various segmentation models for being predictive diagnosis before occurred pop-outs, such as “crack” and “scaling” are the issue for health monitoring. Infrastructure manager administrates a lot of aging structures other than dam well. Learning of damage segmentation models using a diverse range of images for a wide variety of other infrastructures will be the issue for the future. Predictor of damage segmentation intelligence created from scratch, i.e. U-Net by data mining accumulated images is also a challenging issue. Furthermore, 3-dimension segmentation is more useful for volume counting to measure the depth of damage.

[Acknowledgments] We would like to express our gratitude to obtain practical information about Deep Learning and Image Processing Toolbox for training and prediction from Mr. Shinichi Kuramoto and Mr. Takuji Fukumoto.

References

[Ministry of Land, Infrastructure 2013] Water Management Land Conservation Bureau River Environment Division : Dam General Inspection Procedure Commentary, 2013.

[CERI 2016] Public Works Research Institute CERI : Survey and Countermeasure Guidance on the Structure Suspected of Frost Damage (draft), 2016.

[Wu 2002] Wu, Z., Xu, B. Yokoyama, K. : Decentralized Parametric Damage Detection Based on Neural Networks, Comput. Civ. Infrastruct. Eng., 17, pp.175-184, 2002.

[Chun 2015] Chun, P., Yamashita, et al. : Bridge Damage Severity Quantification using Multipoint Acceleration Measurement and Artificial Neural Networks, 2015.

[Gordan 2017] Gordan, M., Razak, H.A. et al. : Recent Development in Damage Identification of Structures using Data Mining, Latin American Journal of Solids and Structures, pp.2373-2401.

[AlexNet 2012] Krizhevsky, A. et al. : ImageNet Classification with Deep Convolutional Neural Networks, NIPS, 2012.

[Gopalakrishnan 2018] Gopalakrishnan, K., Gholami, H. et al. : Crack Damage Detection in Unmanned Aerial Vehicle Images of Civil Infrastructure using Pre-trained Deep Learning Model, International Journal for Traffic and Transport Engineering, 8(1), pp.1-14, 2018.

[Ricard 2018] Ricard, W., Silva, L. et al. : Conclete Cracks Detection based on Deep Learning Image Classification, MDPI Proceedings, 2, 489, pp.1-6, 2018.

[Hoskere 2017] Hoskere, V., Narazaki, Y. et al : Vision-based Structural Inspection using Multiscale Deep Convolutional Neural Networks, 3rd Huixan International Forum on Earthquke Engineering for Young Researchers, 2017.

[Guillanmon 2018] Guillamon, J.R. : Bridge Structural Damage Segmentation using Fully Convolutional Networks, Universitat Politecnica de Catalunya, 2018.

[Long 2015] J. Long, E. Shelhamer, T. Darrell : Fully Convolutional Networks for Semantic Segmentation, CVPR, pp3431-3440, 2015.

[Badrinarayanan 2016] V. Badrinarayanan, A. Kendall, et al., SegNet: Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, ArXiv:1511.00561v3, 2016.

[Hinton 2012] G. Hinton, N. Srivastava, K. Swersky : Lecture 6d – A Separate, Additive Learning Rate for Each Connection, Slides Lecture Neural Networks for Machine Learning, 2012.

[Mukkamala 2017] M.C. Mukkalmala et al.: Variants of RMSProp Adagrad with Logarithmic Regret Bounds, 2017.

[Serra 1992] Serra,J., Vincent,L.: An Overview of Morphological Filtering, Circuits, Systems and Signal Processing, 11(1), pp47-108, 1992.

[Gonzalez 2008] Gonzalez, R.C., Woods, R.E. : Digital Image Processing, 3rd ed., Prentice Hall, 2008.

[Yasuno 2018] T. Yasuno: Infra Machine Learning for Predictive Maintenance via Classification Models, 32th Journal of Society for Artificial Intelligence, 3Z1-04, 2018.

[Yasuno 2019] T. Yasuno, M. Amakata, J. Fujii, Y. Shimamoto : Color-base Damage Feature Enhanced Support Vector Classifier for Monitoring Quake Image, IAPR: CCIW, 2019.

(2019.Feb.15)