

Title: Deep Convolutional Neural Networks for Comprehensive Structural Health Monitoring and Damage Detection

Authors : Bing Zha
Yongsheng Bai
Alper Yilmaz
Halil Sezen

ABSTRACT

Vision-based structural health monitoring (SHM) has become an important approach to recognize and evaluate structural damage after natural disasters. Deep convolutional neural networks (CNNs) have recently attained a breakthrough in computer vision field, in particular for image classification task. In this article, we adopted deep residual neural network (ResNet) whose residual representations and shortcut connections mechanism has gained significant performance in various computer vision tasks. In addition, we applied transfer learning due to a relatively small number of training images. To test our approach, we used the dataset from the 2018 PEER Hub ImageNet Challenge distributed by Pacific Earthquake Engineering Research Center. This challenge proposed eight structural damage detection tasks: scene classification, damage check, spalling condition, material type, collapse check, component type, damage level and damage type which can be categorized as binary and multi-class (3 or 4 classes) classification problems. Our experiments with eight different tasks showed that reliable classification can be obtained for some tasks. Corresponding above eight tasks, classification accuracy varied from 63.1% to 99.4%. Our approach has attained third place for overall tasks in this challenge. Through the individual observation of training dataset, it is found that there are a large number of confusing images. Therefore, it is believed that the accuracy will be improved after making a precise training data.

INTRODUCTION

In these years, researchers prefer to apply knowledge and technologies of Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) into Civil Engineering (CE) and Structural Engineering (SE), especially into Structural Damage Detection (SDD) and Structural Health Monitoring (SHM) [1]. Some researchers also combined knowledge of Computer Vision and Structural Engineering, in particular on damage evaluation and health monitoring of building structures, and developed novel approaches for SDD and SHM [2] [3]. These approaches aim to automatically identify and classify them for experts and administrators from images and videos. The cost and time of classifying structural damage in post-event buildings will drop dramatically with their application.

Gong et al. employed the random forest (RF), support vector machine (SVM) and K nearest neighbor (kNN) to classify the damage of individual building with high-resolution SAR images and building footprint maps after Wenchuan Earthquake (magnitude 8.8) in China in 2008 [4]. Meanwhile, for satellite images obtained before and after the Hurricane 'Charley' struck Punta Gorda in 2004, Radhika et al. used artificial neural network (ANN) and SVM to detect the edges of the destroyed roofs of coastline buildings and classify damage [5]. In [6], Ho et al. applied image-based damage detection and machine learning on cable-stayed bridges to identify the defect of cable

Bing Zha, Alper Yilmaz, Photogrammetric Computer Vision Lab, Department of Civil, Environmental & Geodetic Engineering, The Ohio State University, Columbus, OH 43210, U.S.A. Email: zha.44, yilmaz.15@osu.edu;

Yongsheng Bai, Halil Sezen, Structural Engineering, Department of Civil, Environmental & Geodetic Engineering, The Ohio State University, Columbus, OH 43210, U.S.A. Email: bai.426, sezen.1@osu.edu.

surface by three cameras moving along the cables. As convolutional neural networks (CNNs) and deep learning (DL) outperform alternative techniques in Computer Vision (CV), researchers implemented them into SDD and SHM problems [2] [3].

BACKGROUND

Since 2012, great progress has been made on convolutional neural networks (CNN) in automatically classifying objects in images through the competition, which is called the Large Scale Visual Recognition Challenge (ILSVRC) in ImageNet. In these series of the challenge, 1000 classes of objects can be classified from 14 million images. The algorithms with deep CNNs, such as AlexNet, ZF Net, GoogleNet, VGGNet and ResNet, have been widely used and achieved the same accuracy as humans [7] [8]. Researchers implemented this state-of-the-art architecture into SDD and SHM, and aimed to gain the level of damage recognition as high as ILSVRC.

Yeum et al. applied AlexNet for automatically detecting damage in post-event structures with a large volume of images to identify two types of damage, such as buildings collapse and spalling of building components [3]. Gao et al. were inspired by the ImageNet and have built the Structural ImageNet for collecting pictures with building structural failures, and training CNN for structural damage detection. In order to overcome the limitation on a small number of sampling data in Structural ImageNet, the authors employed deep transfer learning (TL) to classify structural damage. Their tests show that although the proposed method gained a low accuracy in some tasks, such as damage type, they could do as well as human beings in other tasks [2]. In [9], the authors categorized structural damage by ResNet23 and identified 7 classes of the damage by 19-VGG networks with 1695 images cut from 339 original ones for about 250 buildings. Ali et al. also adopted Faster R-CNN (Faster Region Convolutional Neural Networks) in the task to detect damage and defects in historical masonry buildings based on high resolution images [10].

Cha et al. used Faster R-CNN to identify five classes of structural damage: concrete crack, steel corrosion with two levels (medium and high), bolt corrosion, and steel delamination [11]. And the authors in [12] propose another CNN for autonomous detection of pavement cracks. In addition, Kong and Li describe an application that detects cracks in a steel girder with a video stream [13]. Furthermore, the authors in [14] explain the different effects when they use two algorithms of CNNs (VGG16 and ZF Net) in detecting metallic corrosion from visual inspection, and Gibbons et al. implemented a regional CNN to locate the damaged bolt heads with a bounding box regression [15]. Chen et al. applied VGG16 into automatically detecting damaged fasteners on catenary support devices in various climates and regions [16].

METHODOLOGY

In this paper, two key components in deep learning are used for network modeling and training. One is the residual neural network which has proved very beneficial in image recognition tasks, and another is transfer learning technique which is a popular approach when a model trained on limited dataset. There are similarities between our

approach and the approach given in [2]. In this paper we employed a more advanced network and conducted experiments on comprehensive detection tasks. The details of these two methods are described as follows.

Residual Neural Network (ResNet)

Before the advent of residual neural network, there were several popular networks used to achieve image recognition tasks, such as AlexNet [17], VGG [18] and GoogleNet [19]. AlexNet and VGG are simply stacking convolutional and pooling layers so as to fulfill the complicated feature representation. Although they could integrate hierarchical features and led to a series of breakthroughs, a natural question arise: is it better as the number of layers increase? It is found that in practice when deeper networks starts converging, a degradation problem has been exposed: with the network depth increasing, accuracy gets saturated and then degrades rapidly. The residual neural network [8] is thus designed to solve the problem of such deep layers.

The core contribution of ResNet lies in its residual learning compared to the traditional plain stack of layers which can directly cause notorious problem of gradients vanishing and exploding. Furthermore, the ResNet is built upon each layer which learns an incremental function $F(x)$ according to the equation: $y = F(x) + x$. Instead of learning a direct mapping of $x \rightarrow y$ as done in many other CNNs, this residual learning redefine the mapping function which can be realized by skip connections. This structure connects components of different layers with an identity mapping. The authors' hypothesis is that it is easier to optimize the residual mapping than to optimize the original, unreferenced mapping [8]. It showed in many recognition tasks this structure can achieve higher accuracy than previous network structures. There are also a few papers providing explanation of the exact reason behind favourable results [20] [21]. We present two ways to explain the differences. First, from a mathematical viewpoint, if the identity mapping is optimal, we can easily set the residuals as zero than fit an identity mapping by stack of non-linear layers. It is pretty easy to come up with a solution like $F(x) = 0$ rather than $F(x) = x$. Another explanation from intuition perspective such that the whole hierarchical feature combinations can be further optimized. Fewer feature compositions in different layers may better represent object.

We adopted the ResNet152 architecture as training network. The number 152 here denotes the total number of convolutional layers and fully connected layers. There are three different sizes of convolution, also called kernel or filter: 1×1 , 3×3 and 7×7 . In practice, selection of convolution size is still a tricky task, therefore, several empirically indicated sizes of convolution are usually utilized. To this end, we utilized the original convolution combination of ResNet.

Transfer Learning

Big training dataset plays a vital role in deep learning technology. However, it is not trivial to obtain a large amount of training data. To remediate this problem, transfer learning is employed where a model developed for a task is reused as the starting point for a model on a second task. Therefore, with transfer learning, we basically need to exploit what has been learned in one task to improve generalization in another. Simply,

we fine tune the weights that a network has learned using a large amount of data at task A to a new Task B.

Transfer Learning has a long history of research and techniques to tackle different scenarios which basically can be categorized into two cases: 1) insufficient labeled training data to train our network from scratch, 2) an existing pre-trained neural network on a similar task. Considering the second case, it is reasonable to harness the technique of transfer learning to accomplish structural detection and recognition tasks.

As in Figure 1, the pre-train model is trained on ImageNet dataset which contains more than 10,000 categories [22]. In the convolutional layers block, neural networks prone to extract low-level features such as edges and curves in earlier layers and high-level task-specific features in the deeper layers. It is acknowledged that some levels of features can be reused in other similar tasks including the tasks of structural health monitoring. For transfer learning process, we fixed convolutional layers as feature extractor, remove the last fully connected layer replacing with new layer and trained the network using the new dataset.

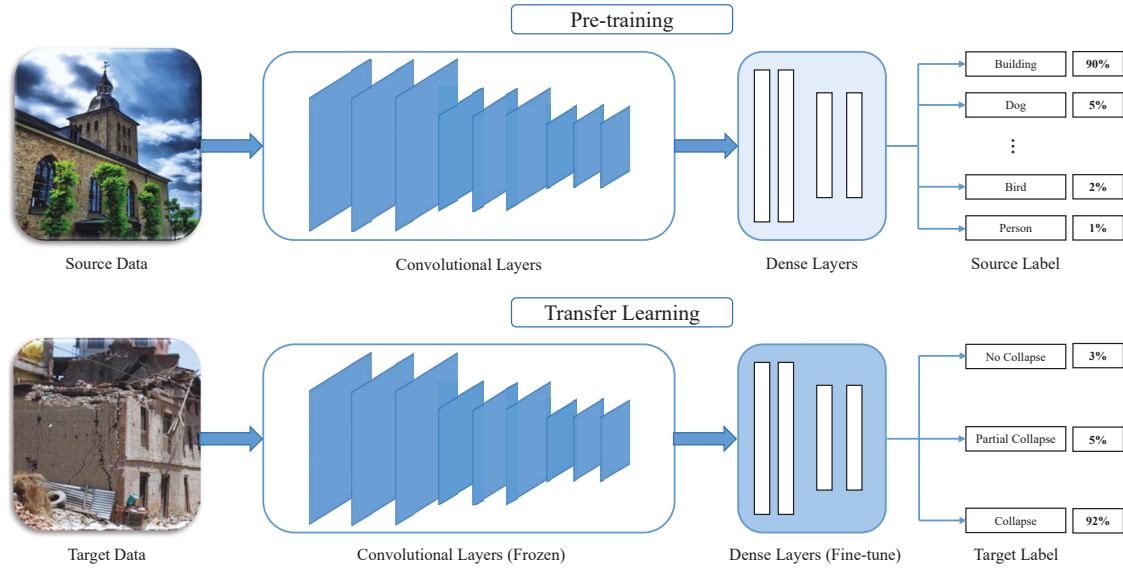


Figure 1. The general transfer learning framework (source data and target data are from ImageNet dataset and PEER Hub Challenge dataset, respectively).

EXPERIMENTS

Since the motivation of this paper is from PEER Hub ImageNet Challenge, the part of dataset can be seen in the challenge website and results has been submitted to Kaggle platform [23]. In addition, we used related images in [24] and tested them on eight trained models respectively.

Dataset and Training Settings

The dataset in PEER Hub ImageNet Challenge are used to measure the performance of designed model. PEER Hub ImageNet Challenge is a competition on multi-purpose classification, by which we can identify structural damage types, materials, and component types. There are eight detection and classification tasks for vision-based structural damage detection: 1) Scene level: pixel levels will help to detect cracks, and object and structural levels focus on identify spalling of components and collapse of buildings; 2) Damage or undamage state; 3) Spalling or Non-spalling; 4) Material type: steel and others; 5) Collapse mode: this task will distinguish global collapse, partial collapse and non-collapse of structures; 6) Component type: there are four types, including beams, columns, walls and others; 7) Damage level: no damage, minor damage, moderate damage and heavy damage; 8) Damage type: 4 types of structural member failure, such as no damage, flexural damage, shear damage and combined damage. All the images are RGB channel at 224×224 resolution. For each task, we first divide data into testing data consisting 20 percent of whole data and training data for the rest of them, then the training data is further partitioned into "real training data" which corresponds to 80 percent of training data and validation data. The reason of partition in such way is to validate model performance in every epoch using validation data and choose the best one at last. The basic image statistics can be seen in the table I.

Once dataset and network architecture are prepared, the remaining task is to train the designed neural network using the dataset. Most deep learning methods adopt stochastic gradient descent (SGD) for training [25]. In this paper, this method is also being used with learning rate 0.001 and momentum 0.9. The loss function is cross-entropy for classification problem. To maximize GPU usage, we use 40 min-batches. The total number of epochs is set as 100. The whole program is run under Windows system and tested on NVIDIA GeForce GTX 1080.

Evaluation of Results

We evaluate our result of task classification based on the confusion matrix, where the diagonal elements denote true positive and N denotes the total number of tested images. For this end, the metric called accuracy represents the percentage of the correctly classified images and is defined as:

$$Accuracy = \frac{\sum (TruePositive)}{N}$$

From the original challenge, the tasks are divided into three levels of difficulty. Damage check, material type and spalling condition belong to easy tasks. It shows from the Table I the test accuracy of material type is highest and hence indicates that the network has learned feature representation, which can distinguish this type of structural damage. However, the task of spalling condition and damage check obtained 79.6% and 81.9% test accuracy which are greatly lower than the other task. After checking the data of this two tasks, we found that the data is noisy which can be the key reason causing poor performance. Therefore, it is believed that the performance can be improved after cleaning training data from noise. The scene classification, collapse check and component type are at the medium level. The task of scene classification obtains more than 90.0% test accuracy. This task also contains the most number of training images among all

the data. The remaining two tasks plus the most difficult level tasks, damage level and damage type obtain similar test accuracy. The common characteristic for these tasks is that the classes between each task is similar and ambiguous which poses a significant challenge to the neural network. Overall, the test accuracy for eight structural damage tasks can basically be considered satisfactory for engineers to obtain a basic knowledge of the damage.

TABLE I. CLASSIFICATION RESULTS ON TESTING DATA

Detection Tasks	Number of Classes	Image Statistics			Test Accuracy (ResNet152)
		Train	Validation	Test	
Scene Classification	3	13939	3485	4356	93.8%
Damage Check	2	4730	1183	1479	81.9%
Spalling Condition	2	2635	659	824	79.6%
Material Type	2	3470	867	1085	99.5%
Collapse Check	3	2105	527	658	63.1%
Component Type	4	2104	526	658	71.7%
Damage Level	4	2105	527	658	67.8%
Damage Type	4	2105	527	658	67.5%

SUPPLEMENTARY TEST AND DISCUSSION

Besides the first experiment where we only used the provided data, we conducted our test experiment on structural damage images collected from other sources and want to further test our models, despite the fact that the number of images is very small. We manually label them and compared the results after they are classified with our models. As indicated in Table II, all the test accuracy are above 60%. Even though the accuracy is lower than the case in the first experiment, we think it can still be acceptable for these collected images. We also observed the scale of structures and components in images can affect the accuracy, and perhaps more data for training will improve the performance of these models.

TABLE II. CLASSIFICATION RESULTS ON REAL-WORLD DATA

Detection Tasks	Number of Classes	Number of Images	Test Accuracy (ResNet152)
Scene Classification	3	65	76.9%
Damage Check	2	65	80.0%
Spalling Condition	2	65	69.2%
Material Type	2	65	66.2%
Collapse Check	3	65	64.6%
Component Type	4	65	84.6%
Damage Level	4	65	63.1%
Damage Type	4	65	60.0%

CONCLUSIONS

In this paper, we presented residual neural network (ResNet) with transfer learning to detect and classify structural damage and gained better performance than VGG network. Compared to simply deep neural networks, ResNet can solve the gradient propagation in “deep” layers due to its residual learning mechanism and easily extend from natural images to man-made objects area. We conducted two groups of experiments to test the model performance, including the data separated from the original challenge and our own collection data. The results have demonstrated that the test accuracy is satisfactory for most of tasks. Although the current ResNet can obtain acceptable results, there is still large room to improve. The original dataset provided for the challenge is noisy. Therefore, we believe that when the more advanced neural network is designed in deep learning field and the larger data are provided as training, it is possible for researchers in structural engineering to combine their domain knowledge to obtain results with higher accuracy.

REFERENCES

1. Salehi, H. and R. Burgueno. 2018. “Emerging artificial intelligence methods in structural engineering,” *Engineering structures*, 171:170–189.
2. Gao, Y. and K. M. Mosalam. 2018. “Deep transfer learning for image-based structural damage recognition,” *Computer-Aided Civil and Infrastructure Engineering*, 33(9):748–768.
3. Yeum, C. M., S. J. Dyke, and J. Ramirez. 2018. “Visual data classification in post-event building reconnaissance,” *Engineering Structures*, 155:16–24.
4. Gong, L., C. Wang, F. Wu, J. Zhang, H. Zhang, and Q. Li. 2016. “Earthquake-induced building damage detection with post-event sub-meter VHR TerraSAR-X staring spotlight imagery,” *Remote Sensing*, 8(11):887.
5. Radhika, S., Y. Tamura, and M. Matsui. 2015. “Cyclone damage detection on building structures from pre-and post-satellite images using wavelet based pattern recognition,” *Journal of Wind Engineering and Industrial Aerodynamics*, 136:23–33.
6. Ho, H.-N., K.-D. Kim, Y.-S. Park, and J.-J. Lee. 2013. “An efficient image-based damage detection for cable surface in cable-stayed bridges,” *Ndt & E International*, 58:18–23.
7. Russakovsky, O., J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. 2015. “Imagenet large scale visual recognition challenge,” *International journal of computer vision*, 115(3):211–252.
8. He, K., X. Zhang, S. Ren, and J. Sun. 2016. “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.
9. Hoskere, V., Y. Narazaki, T. Hoang, and B. Spencer Jr. 2018. “Vision-based Structural Inspection using Multiscale Deep Convolutional Neural Networks,” *arXiv preprint arXiv:1805.01055*.
10. LuqmanAli, W. K. and K. Chaiyasarn. 2019. “Damage Detection and Localization in Masonry Structure using Faster Region Convolutional Networks,” *International Journal*, 17(59):98–105.
11. Cha, Y.-J., W. Choi, G. Suh, S. Mahmoudkhani, and O. Büyüköztürk. 2018. “Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types,” *Computer-Aided Civil and Infrastructure Engineering*, 33(9):731–747.
12. Zhang, A., K. C. Wang, B. Li, E. Yang, X. Dai, Y. Peng, Y. Fei, Y. Liu, J. Q. Li, and C. Chen. 2017. “Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network,” *Computer-Aided Civil and Infrastructure Engineering*, 32(10):805–819.

13. Kong, X. and J. Li. 2018. "Automated fatigue crack identification through motion tracking in a video stream," in *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018*, International Society for Optics and Photonics, vol. 10598, p. 105980V.
14. Atha, D. J. and M. R. Jahanshahi. 2018. "Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection," *Structural Health Monitoring*, 17(5):1110–1128.
15. Gibbons, T. J., S. Pierce, K. Worden, and I. Antoniadou. 2018. "Convolutional neural networks for the detection of damaged fasteners in engineering structures," in *Proceedings of the 9th European workshop on structural health monitoring (EWSHM 2019)*, NDT. net.
16. Chen, J., Z. Liu, H. Wang, A. Núñez, and Z. Han. 2018. "Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network," *IEEE Transactions on Instrumentation and Measurement*, 67(2):257–269.
17. Krizhevsky, A., I. Sutskever, and G. E. Hinton. 2012. "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105.
18. Liu, S. and W. Deng. 2015. "Very deep convolutional neural network based image classification using small training sample size," in *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, ISSN 2327-0985, pp. 730–734, doi:10.1109/ACPR.2015.7486599.
19. Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. 2015. "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9.
20. Veit, A., M. J. Wilber, and S. Belongie. 2016. "Residual networks behave like ensembles of relatively shallow networks," in *Advances in neural information processing systems*, pp. 550–558.
21. He, K., X. Zhang, S. Ren, and J. Sun. 2016. "Identity mappings in deep residual networks," in *European conference on computer vision*, Springer, pp. 630–645.
22. Deng, J., W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. 2009. "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, Ieee, pp. 248–255.
23. 2010, "Kaggle: Your Home for Data Science," <https://www.kaggle.com/>.
24. Li, K. 2017. *Collapse Experiments and Assessment of Masonry Wall Buildings*, Ph.D. thesis, The Ohio State University.
25. LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, 86(11):2278–2324.



CONTRIBUTING AUTHOR COPYRIGHT RELEASE FORM

As author of the chapter/contribution titled Deep Convolutional Neural Networks for Comprehensive Structural Health Monitoring and Damage Detection, to appear in the *Proceedings of Structural Health Monitoring 2019*, I hereby agree to the following:

1. To grant to DEStech Publications, Inc., 439 North Duke Street, Lancaster, PA, 17602, copyright of the above named chapter/contribution (for U.S. Government employees to the extent transferable), in print, electronic, and online formats. However, the undersigned reserve the following:
 - a. All proprietary rights other than copyright, such as patent rights.
 - b. The right to use all or part of this article in future works.

DEStech Publications thereby retains full and exclusive right to publish, market, and sell this material in any and all editions, in the English language or otherwise.

1 I warrant to DEStech Publications, Inc., that I am the (an) author of the above-named chapter/contribution and that I am the (a) copyright holder of the above-named chapter/contribution granted to DEStech Publications, Inc.

2 I warrant that, where necessary and required, I have obtained written permission for the use of any and all copyrighted materials used in the above-named chapter/contribution. I understand that I am responsible for all costs of gaining written permission for use of copyrighted materials.

3 I agree to assume full liability to DEStech Publications, Inc. and its licensee, and to hold DEStech Publications, Inc. harmless for any claim or suit filed against DEStech Publications, Inc. for violation of copyrighted material used in the above-named contribution.

Please sign and date this form and retain a copy for your records. Please include original form with your chapter/paper.

Thank you for your cooperation.

Please print name: Bing Zha, Yongsheng Bai, Alper Yilmaz, Hail Sezen

Signed: Bing Zha Dated: 04/29/2019

439 NORTH DUKE STREET • LANCASTER, PENNSYLVANIA 17602-4967, U.S.A. Toll Free:
(866) 401-4337 • Tel: (717) 290-1660 • Fax: (717) 509-6100 E-mail: info@destechpub.com •
Internet address: www.destechpub.com