CS 229 Final Report
StruX: Methods Comparison for Structural Damage Recognition and Classification

Tong Liu
tongliu@stanford.edu

Yitao Gao
yitaogao@stanford.edu

I. INTRODUCTION

Structural health monitoring (SHM) and rapid damage assessment after natural hazards and disasters have become an important focus in civil engineering. Nowadays, structural damage recognition using images is one of the important topics in vision-based SHM and structural reconnaissance, which greatly relies on human visual inspection and experience. It is time to implement the state-of-art DL technologies in civil engineering applications and evaluate its potential benefits.

The objective of the project is to classify the damage type(i.e.no/ flexural damage/ shear damage/ combined damage), spalling condition (spalling/ no spalling) and collapse mode (partial/ total/ no collapse) of the structures, which is mainly assessed by human experience traditionally.

The dataset of the structural damage image and label are provided by PEER Hub ImageNet (PHI), which including over 10k+ images of different structural types, damage states. The input of our algorithms is the structural damage images after standardization, Then we use different ConvNet models (VGG, ResNet, MobileNet) to output a predicted label of structural damage (damage type /damage level/ spalling condition/ collapse mode).

II. RELATED WORKS

The machine learning applications in the field of structural health monitoring have developed for years. Research scale of such damage is divided into local damage and global damage. Local damage, as indicated by its name, refers to the damage that occurs in part of the building structures, and the local damage detection mainly focuses on crack detection and spalling detection on a concrete structure.

Peng and Zhang implemented two ANN-based algorithms (backpropagation (BP) and delf-organizing Maps (SOM)) and their applications for the recognition of surface defects on images taken from bridges. The crack detection can not only be applied on the surface of bridges and buildings, but on the surface of highways as well.

O'Byrne and Pakrashi explored the crack detection on the surface of road and also explored the potential of low-cost facilities including DSLR and smartphone. Their experiment indicates that smartphones is a viable, low-cost method for executing quick assessments of road integrity. Masato uses deep neural network to detect the linear crack in tunnels. All of those researches indicate the flexibility of machine learning and computer vision.

The global damage refers to the damage visible

from the outside of structures. The classification of global includes classification of collapse mode and damage type. Gao uses feature extractor and fine-tuning to find the relative optimal model parameters and scope of application in the field of global damage classification. Cha uses region-based deep learning to detect the multiple damage type on the surface of concrete structure.

III. DATASET & FEATURE

Three different tasks are studied in our application, including classification of damage type, spalling condition and collapse mode. The dataset for each task is independent, and the size of each dataset is about 2500 images, which is relatively small. Furthermore, the dataset is divided into three parts: training set (2000 images), validation set (300 images) and test set (200 images).

The image is 224x224x3 RGB images. Images are standardized before importing into models, the value of pixels transform from 0-255 to 0-1. The output (labels) varies depend on the tasks. For damage type classification the number of class is 4, spalling condition is 2 and collapse mode is 3. Part of images are shown below.

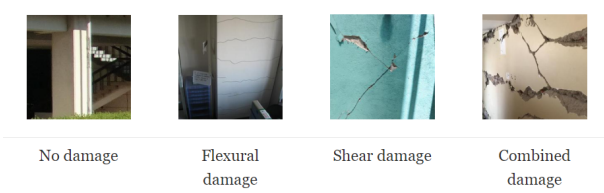

Fig. 1. Damage Type Classification

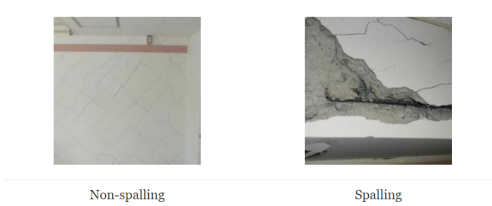

Fig. 2. Spalling Condition Classification

Fig. 3. Collapse Mode Classification

The feature used in the ConvNet models is the pixel values of images. After different Conv layers the labels are predicted.

IV. METHOD

In order to predict the label of structural damage, different models are implemented. The three baseline models are implemented including AlexNet, VGG-16 and ResNet50.

A. Baseline Models

The structure of AlexNet is shown as follows. The image firstly goes through a large convolutional layer and max pooling layer and repeat twice. Then the image goes through three smaller convolutional layers followed by a max pooling layer. Then flatten the image and connect the image with several full-connect layers. The activation functions are all ReLU.

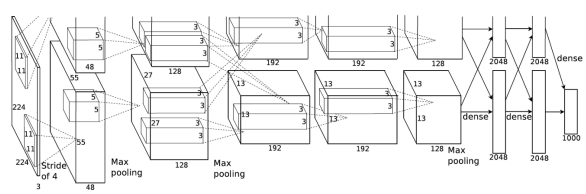

Fig. 4. AlexNet Model

Compared with AlexNet, the most significant difference in VGG-16 is that only 3x3 kernel and 2x2 maxpooling layer are used and the number of layer incrase dramatically. The structure of VGG-16 is shown below.

ResNet is a deeper neural network compared with AlexNet and VGG-16. There will be a degradation problem in a deep neural network

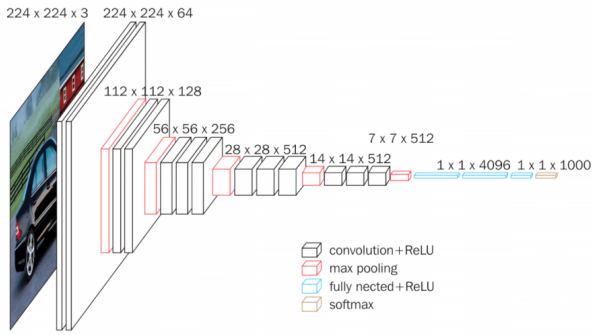

Fig. 5. VGG-16 Model

when increasing layer number, because gradient of the model will explode or vanish when the layer increase. As a result, an identity shortcut connection is used in ResNet, which directly let the data skip one or more layer.It is also called residue learning. Residue learning helps to increase the number of layer in CNN dramatically. The figure of residue learning is shown below.

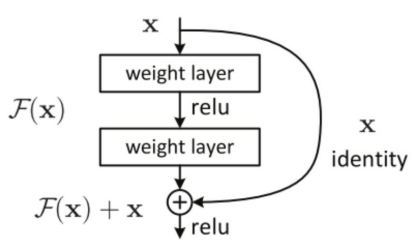

Fig. 6. Residue Learning

B. Baseline Results

The baseline results and training curves are shown as below.

TABLE I
BASELINE RESULT

Task	classes	Training Accuracy	Test Accuracy
Spalling Condition	2	94.5%	65.6%
Collapse Mode	3	97.5%	58.1%
Damage Type	4	94.2%	57.7%

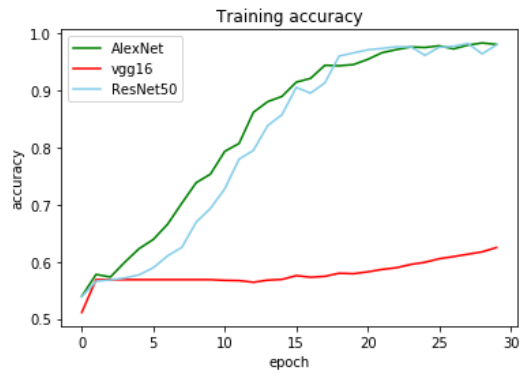

Fig. 7. Training Accuracy

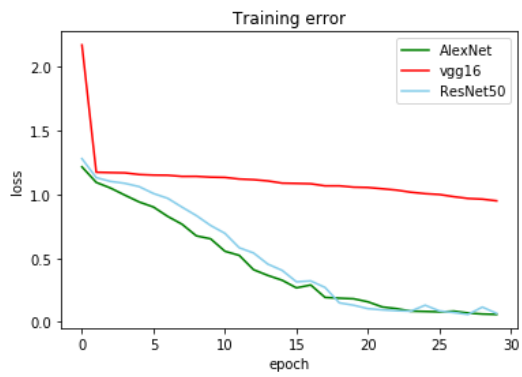

Fig. 8. Training Error

Fine-tuning are not used in baseline models to investigate the performance of naive models without any pre-trained parameters. From the training curve and prediction result, it can be perceived that for damage type prediction, the model tends to be over-fitting since the training accuracy is very high, while test accuracy is relatively low. To further improve the model, several methods can be implemented to result prediction including fine-tuning and regularization. Additionally, more models will trained to find the optimal models.

C. Model Improvements

Fine-tuning is implemented in the improvement of baseline model. Initially, pre-trained model parameters are obtained from ImageNet and assigned to corresponding layers. In this way,parameters in lower-level layers are frozen

in order to accelerate the training process and to extract the feature of images (pixel gradients, texture and color). Parameters in rest of the higher-level conv-blocks and fully-connected layers are unfrozen and kept updated in the training. The VGG16 with fine-tuning is shown below.

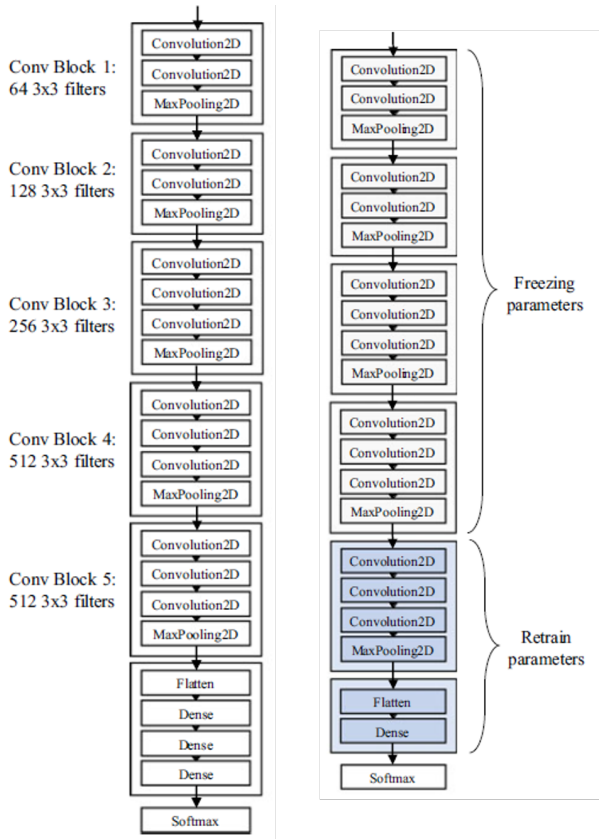

Fig. 9. Model with Fine-tuning

In addition to fine-tuning, several hyper-parameters are tuned to improve model performance as well. Four alternative models are tested including VGG19, MobileNet, MobileNetV2 and InceptionV3. The list of hyperparameters are shown below.

- Dropout Rate
- L1/L2 Regularization
- Learning Rate & Learning Rate Decay
- Number of Conv Blocks frozen & unfrozen
- SGD & RMSProp & Adam & Momentum

The improved model training curve and prediction results summary are shown below.

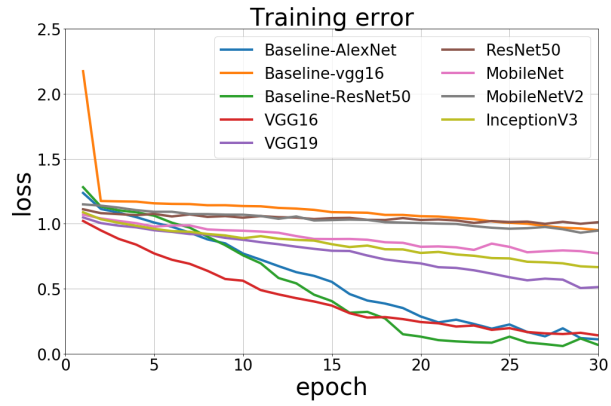

Fig. 10. Baseline Training Error

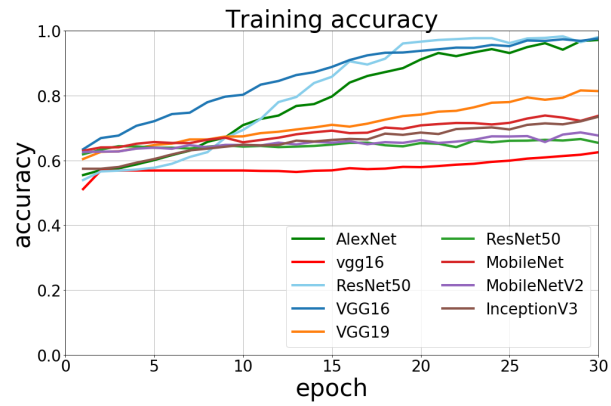

Fig. 11. Baseline Training Accuracy

	Task	Spalling Condition	Collapse Mode	Damage Type
Baseline Model	AlexNet	59.61%	53.30%	56.20%
	VGG16	65.63%	58.10%	57.70%
	ResNet50	57.94%	51.20%	52.10%
Improved Model	VGG16	84.76%	67.40%	64.11%
	VGG19	84.34%	66.50%	62.34%
	ResNet50	79.02%	62.30%	58.42%
	MobileNet	84.75%	68.20%	61.90%
	MobileNetV2	85.45%	69.11%	64.50%
	InceptionV3	80.50%	63.20%	57.14%
Paper		91.50%	—	68.80%

Fig. 12. Baseline Result Summary

The learning curves including training error and accuracy are shown in the figure. Models

without fine-tuning have large variance and issue of over-fitting, which can be found from the huge difference between training accuracy and test accuracy. The issue of overfitting is controlled after applying fine-tuning, dropout and L2 regularization. But bias issue occurs after fine-tuning, where accuracy of training set decrease dramatically.

From table, MobileNetV2 has highest accuracy among those models, which is close to the result from paper (Gao, 2018).

D. INTERMEDIATE OUTPUT

we have clear intuition about the input and output of Nerual Network, while we don't know the output in the intermediate layer, which is also important. The ideal convolutional neural network can extract low-level feature(shape, edge) in first several Conv blocks. Take damage type classification for instance, the intermediate output is shown below.

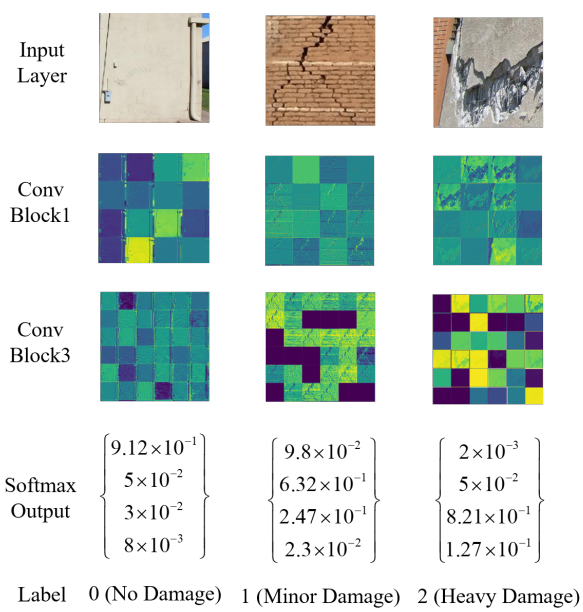

Fig. 13. Intermediate Output

As we go deeper in the layers, the activations become increasingly abstract and less visually interpretable. They begin to encode higher-level concepts such as single borders, corners and angles. Higher presentations carry increasingly less information about the visual contents of the image, and increasingly more information related to the class of the image.

That's the case we use fine-tuning with layer-frozen in neural network. If we train the model from draft without fine-tuning, even though the training accuracy is very high, while no shape and edges are detected in the layers, which lead to low accuracy in validation set and test set. Fine-tuning is a powerful tool to control variance and improve model performances.

V. CONCLUSIONS & FUTURE WORK

Deep convolutional Neural Network is implemented to classify structural damage. VGG16 and MobileNetV2 perform fast computation but has limited accuracies for multi-classes. The performance can be improved through fine-tuning and hyperparameter tuning.

Besides the three tasks, further recognition such as damage localization and quantification can be studied, which can also be combined with drone or satellite imagery in post-disaster structural recognition.

The prediction accuracy is not as high as expected. The reason is probably from the bad input image. Image segmentation is probably adopted in pre-processing.

REFERENCES

[1] Gao, Y. Q., Li, K. B., Mosalam, K. M., & Gnay, S. (2018, June). Deep Residual Network with Transfer Learning for Imagebased Structural Damage Recognition. In Eleventh US National Conference on Earthquake Engineering, Integrating Science, Engineering & Policy.

[2] Peng, J., Zhang, S., Peng, D., & Liang, K. (2018, October). Research on Bridge Crack Detection with Neural Network Based Image Processing Methods. In 2018 12th International Conference on Reliability, Maintainability, and Safety (ICRMS) (pp. 419-428). IEEE.

[3] O'Byrne, M., Pakrashi, V., Schoefs, F., & Ghosh, B. (2018, August). Damage Assessment of the Built Infrastructure using Smartphones. In Civil Engineering Research in Ireland 2018 (CERI2018), University College Dublin, Ireland, 29-30 August, 2018. University College Dublin.

[4] UKAI, M. (2019). Tunnel Lining Crack Detection Method by Means of Deep Learning. Quarterly Report of RTRI, 60(1), 33-39.

[5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[6] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[7] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[8] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

[9] Cha, Y. J., Choi, W., Suh, G., Mahmoudkhani, S., & Bykztrk, O. (2018). Autonomous structural visual inspection using regionbased deep learning for detecting multiple damage types. ComputerAided Civil and Infrastructure Engineering, 33(9), 731-747.

CONTRIBUTION

Tong & Yitao both train the models locally and Yitao also train the model in Google Cloud. Tong finished the final write-up and Yitao modified the report. Poster is finished by Tong & Yitao modified the poster. Tong generates the figure in final write-up with Photoshop.

CODE

The code is compressed in zipfile and uploaded to OneDrive.

The link is `https://1drv.ms/u/s!Ak_MpY2Hets4gZNCuryxYMVVj7HD5g`