BOISE STATE UNIVERSITY

Integrated APC-GAN and AttuNet framework for automated pavement crack pixel-level segmentation: A new solution to small training datasets

Comprehensive Exam
by
Tianjie Zhang

Committee Members

Yang Lu, Ph.D.
Tim Andersen, Ph.D.
Eric Henderson, Ph.D.
Mahmood Mamivand, Ph.D.

November 11, 2022

Emphasis Area: Data Science

- ❑ Background and Motivation
- ❑ Literature Review
- ❑ Methodology
- ❑ Results

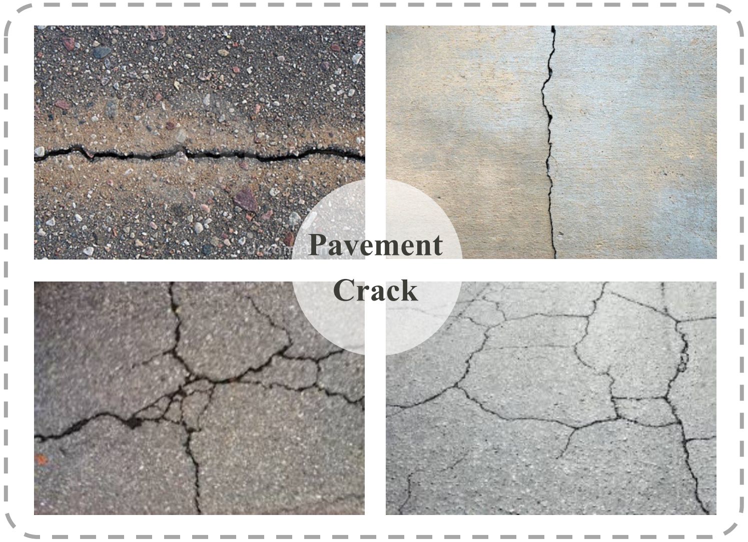
Background

Crack has become one of the primary defects in **pavement[1].**

- Seriously affects the service life of road
- People's comfortable when driving

Traditional inspection[2]:

- Time-consuming
- Labor-intensive
- Dangerous
- Affect Traffic Flow



From: Deep Crack

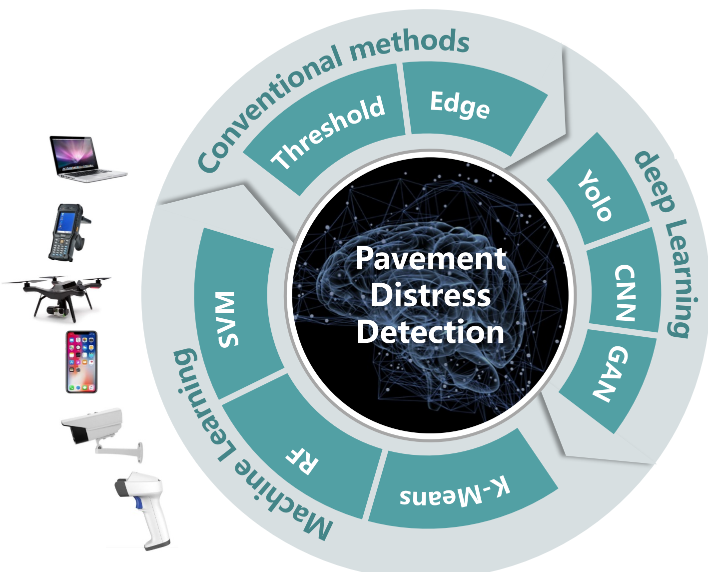
From:https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpp/15049/007.cfm

1. Wang, W., M. Wang, H. Li, H. Zhao, K. Wang, C. He, J. Wang, S. Zheng, and J. Chen. Pavement crack image acquisition methods and crack extraction algorithms: A review. Journal of Traffic and Transportation Engineering (English Edition), Vol. 6, No. 6, 2019, pp. 535-556.
2. Yu, Y., M. Rashidi, B. Samali, A. M. Yousefi, and W. Wang. Multi-image-feature-based hierarchical concrete crack identification framework using optimized SVM multi-classifiers and D–S fusion algorithm for bridge structures. Remote Sensing, Vol. 13, No. 2, 2021, p. 240.

Background

Intelligent pavement detection and identification becomes a significant task.

- Image processing
- Machine learning
- **Deep learning**



Deep Learning: needs a lot of data to train

The training dataset is **lacking or uneven**, making it insufficient to train an accurate segmentation model.

Collecting road images are costive.

From: https://www.surveylegend.com/customer-insight/open-ended-survey-questions/

Introduction

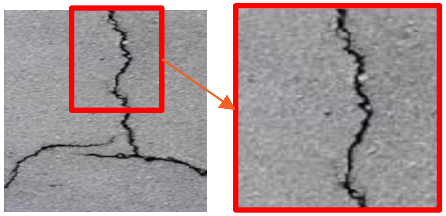
Data augmentation

Network Structure

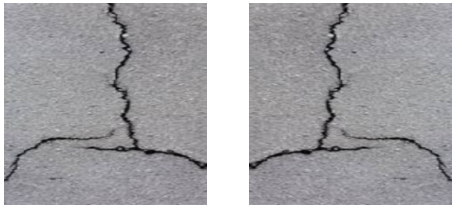
Traditional image augmentation → GAN → DCGAN

- Crop
- Flip
- Adjust brightness
- contrast
- Blur

…



Crop



Flip

Mazzini, D., P. Napoletano, F. Piccoli, and R. Schettini. A novel approach to data augmentation for pavement distress segmentation. Computers in Industry, Vol. 121, 2020, p. 103225.

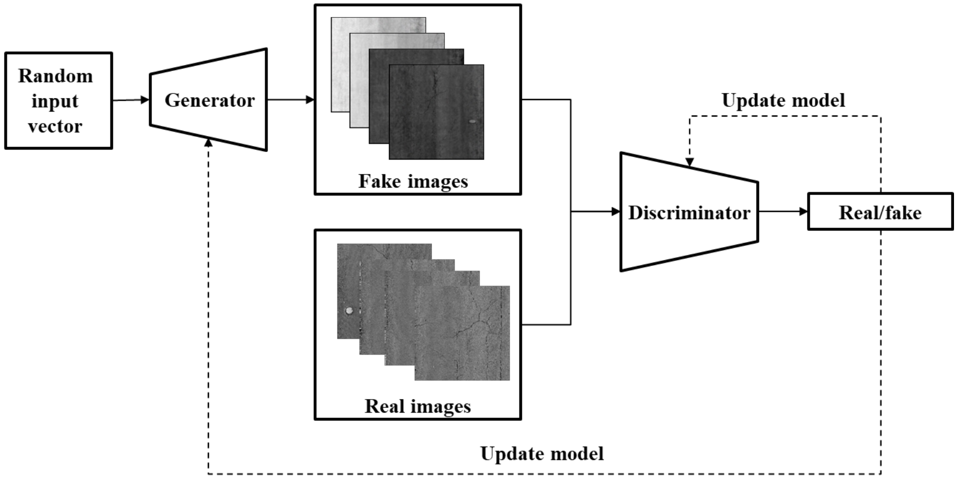
Introduction

Data augmentation

Network Structure

Traditional image augmentation → GAN → DCGAN

In 2014, Goodfellow[1] proposed the concept of **generative adversarial networks** (GANs), which can produce real-like images through a battle between a generator and discriminator.



1. Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. Advances in neural information processing systems, Vol. 27, 2014.

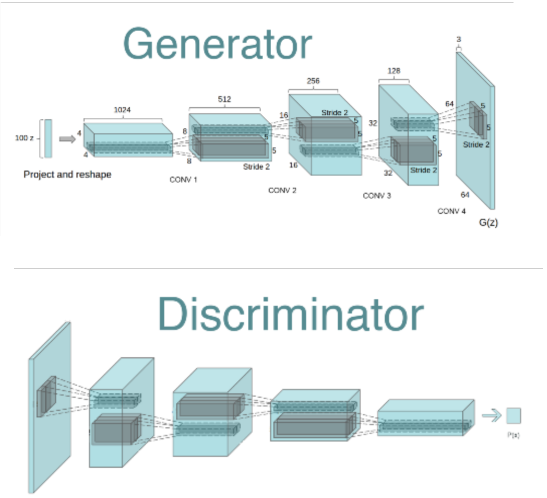
Introduction

Data augmentation | Network Structure

Traditional image augmentation → GAN → DCGAN

Alec Radford[1] proposed deep convolutional generative adversarial networks (**DCGANs**) based on the conception of GAN, and it showed good representations of images. His main work is replacing the FNN with CNN. Drawbacks:

- Only work on small revolution images
- Discriminator studies too fast



From google images

1. Radford, A., L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

Introduction

Data augmentation | Network Structure

ResNet → U-Net → Attention module

ResNet was proposed by Kaiming He[1] in 2015 by learning residual functions with reference to the layer input.

Bang, S et al.[2] used ResNet-152 to increase the layers in structure in purpose of increasing the accuracy of cracks identification.

1. He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
2. Bang, S., S. Park, H. Kim, and H. Kim. Encoder–decoder network for pixel-level road crack detection in black-box images. Computer-Aided Civil and Infrastructure Engineering, Vol. 34, No. 8, 2019, pp. 713-727

Introduction

Data augmentation

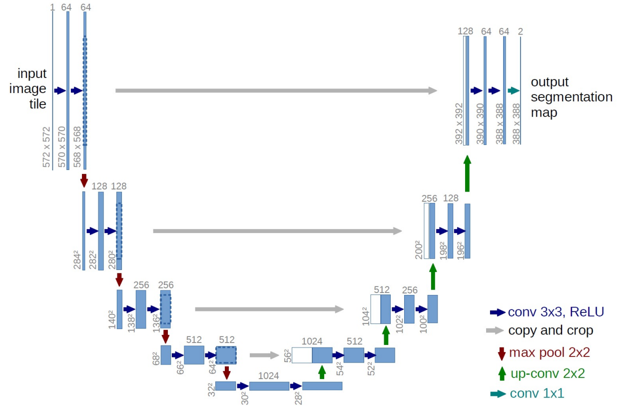
Network Structure

ResNet

U-Net

Attention module

The main idea of U-Net[1] is to replace pooling operators by upsampling operators to increase the output resolution. Also, it combines the high-resolution features with the upsampled output to learn more precise information based on small dataset.



From google images

1. Ronneberger, O., P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation.In International Conference on Medical image computing and computer-assisted intervention, Springer, 2015. pp. 234-241.

Introduction

Data augmentation

Network Structure

ResNet → U-Net → Attention module

The attention module[1] is popular in Nature Language Process (NLP) and now it is starting to be applied in the computer vision area. It can improve the sensitivity and efficiency of network to get rid of large amount data[2].

1. Wan, H., L. Gao, M. Su, Q. Sun, and L. Huang. Attention-based convolutional neural network for pavement crack detection. Advances in Materials Science and Engineering, Vol. 2021, 2021.
2. 21. Oktay, O., J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, and B. Kainz. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999, 2018.

Motivation

- In order to solve the problem that the pavement crack training dataset is lack and uneven, while we want to use deep learning methods to do automatic crack detection.

This paper proposed a framework for pavement crack segmentation which can work on a very small dataset.

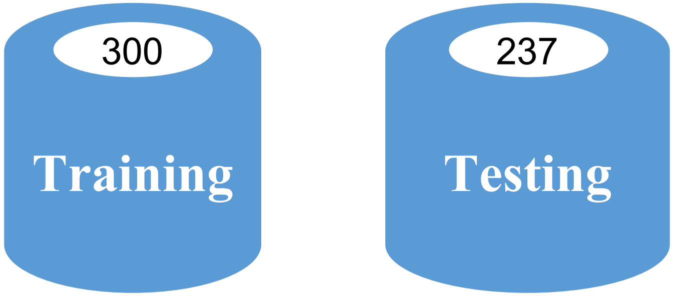
It contains an automated pavement crack generative adversarial network (APC-GAN) and a new pixel-level pavement crack segmentation network, AttuNet.

Data

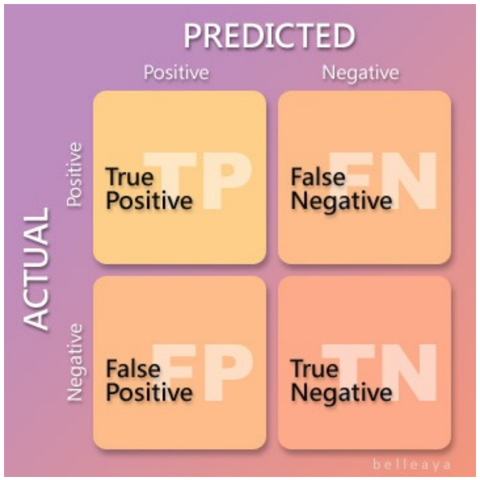
A **DeepCrack** dataset was utilized in this work to test the performance of the CNN networks. The DeepCrack dataset is an open-source dataset published in GitHub (https://github.com/yhlleo/DeepCrack).

This dataset consists of 537 RGB crack images with manually annotated segmentations. The image has a resolution of 544 * 384 pixels.
The images are divided into two subsets:

- 300 images for training
- 237 images for testing



Evaluation metrics



From google images

$$P = \frac{TP}{TP + FP}$$

Precision (P) can measure how accurate the predictions are.

$$R = \frac{TP}{TP + FN}$$

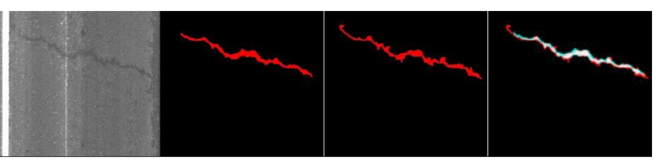
Recall (R) suggests the level of sensitivity for prediction results.

$$F_1 = \frac{2PR}{P + R}$$

F1 is defined based on the harmonic average of Precision and Recall.

$$mIoU = \frac{1}{2}\sum_{i,j}^{k=2} \frac{p_{ii}}{p_{ij} + p_{ji} - p_{ii}}$$

The mIoU is the most important index showing the segment ability of the method as it counts and compares each pixel. A higher value in mIoU means more pixels is classified accurately.



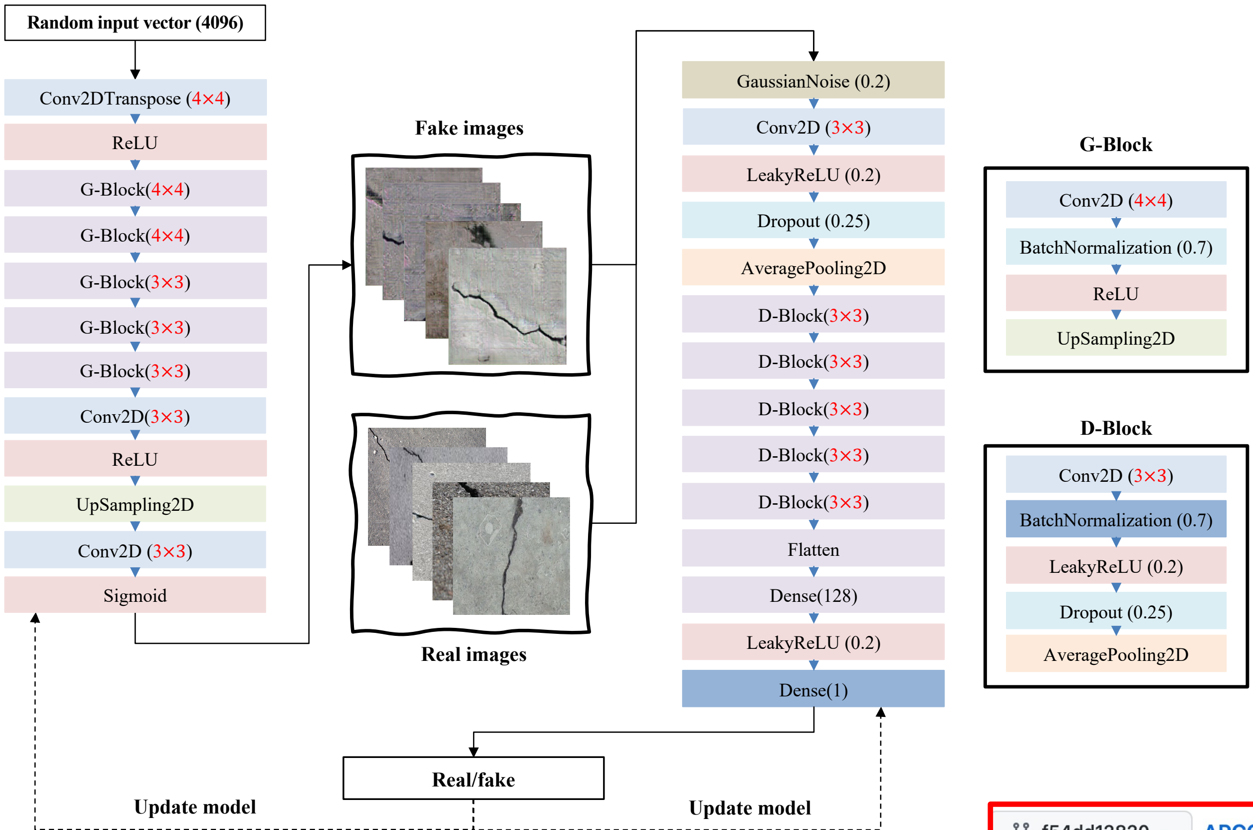
Raw image | Prediction | Ground Truth | IoU

$$mPA = \frac{1}{2}\sum_{i}^{k=2} \frac{p_{ii}}{t_i}$$

PA is a semantic segmentation metric that denotes the percentage of pixels that are accurately classified in the image.

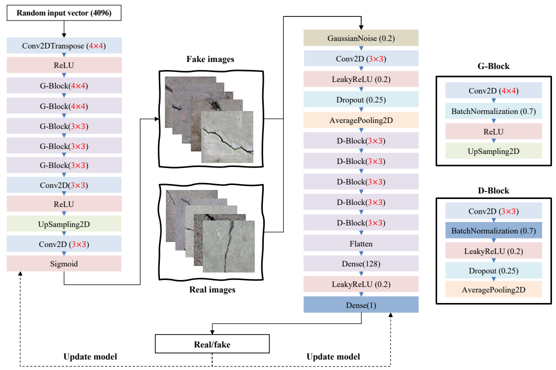
1. Ai, D., G. Jiang, L. S. Kei, and C. Li. Automatic pixel-level pavement crack detection using information of multi-scale neighborhoods. Ieee Access, Vol. 6, 2018, pp. 24452-24463.
2. Kaddah, W., M. Elbouz, Y. Ouerhani, V. Baltazart, M. Desthieux, and A. Alfalou. Optimized minimal path selection (OMPS) method for automatic and unsupervised crack segmentation within two-dimensional pavement images. The Visual Computer, Vol. 35, No. 9, 2019, pp. 1293-1309.

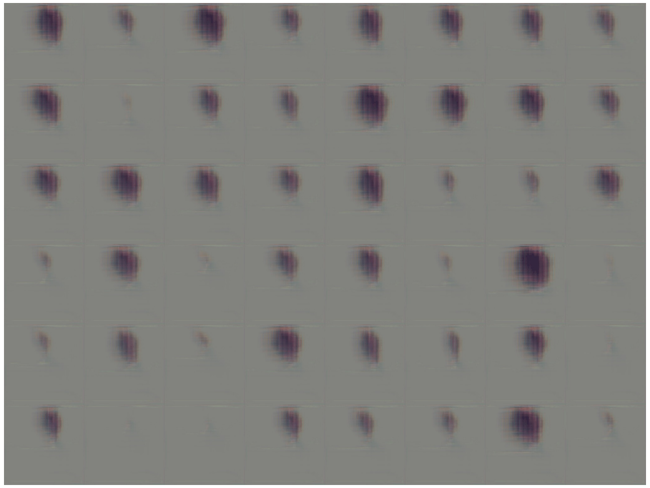
APC-GAN structure



- Structure of APC-GAN

APC-GAN





The improvements of the APC-GAN compared to the DCGAN:

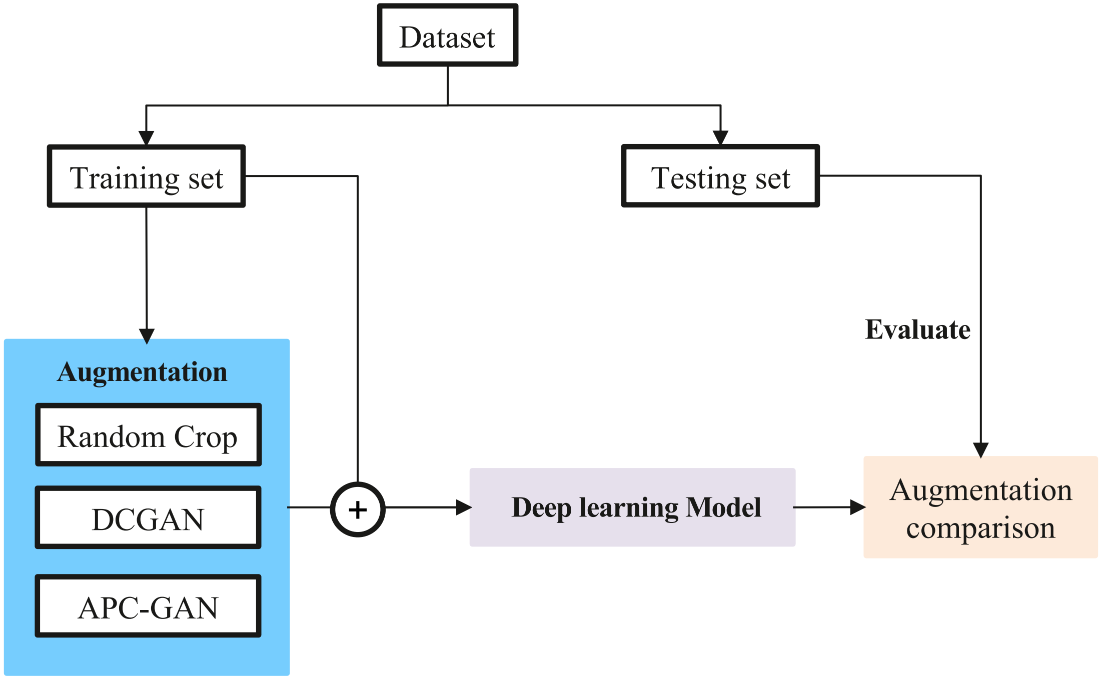
- **Large kernel size** is used. The kernel size is increased to 4×4 in generator and to 3×3 in discriminator.

- The **number** of convolutional layers is increased. More layers can help capture additional information which can eventually add sharpness to the final images produced.

- A **batch normalization layer** is followed by the convolutional layer.

- A **Gaussian noise layer** is added as the first layer of the discriminator to prevent the discriminator from studying too quick.

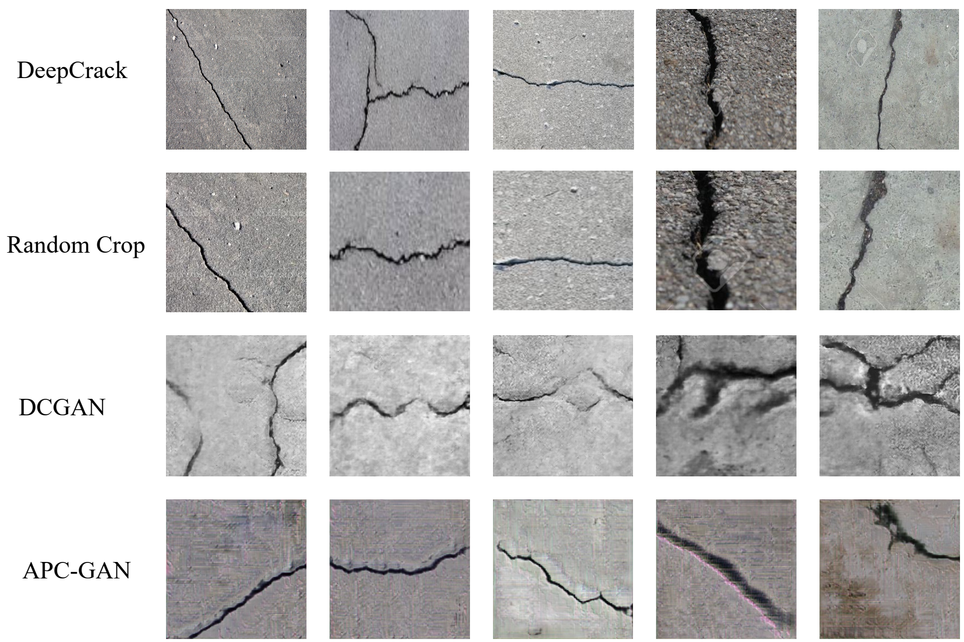
Validation process



- Procedure to evaluate the capacity of APC-GAN

APC-GAN

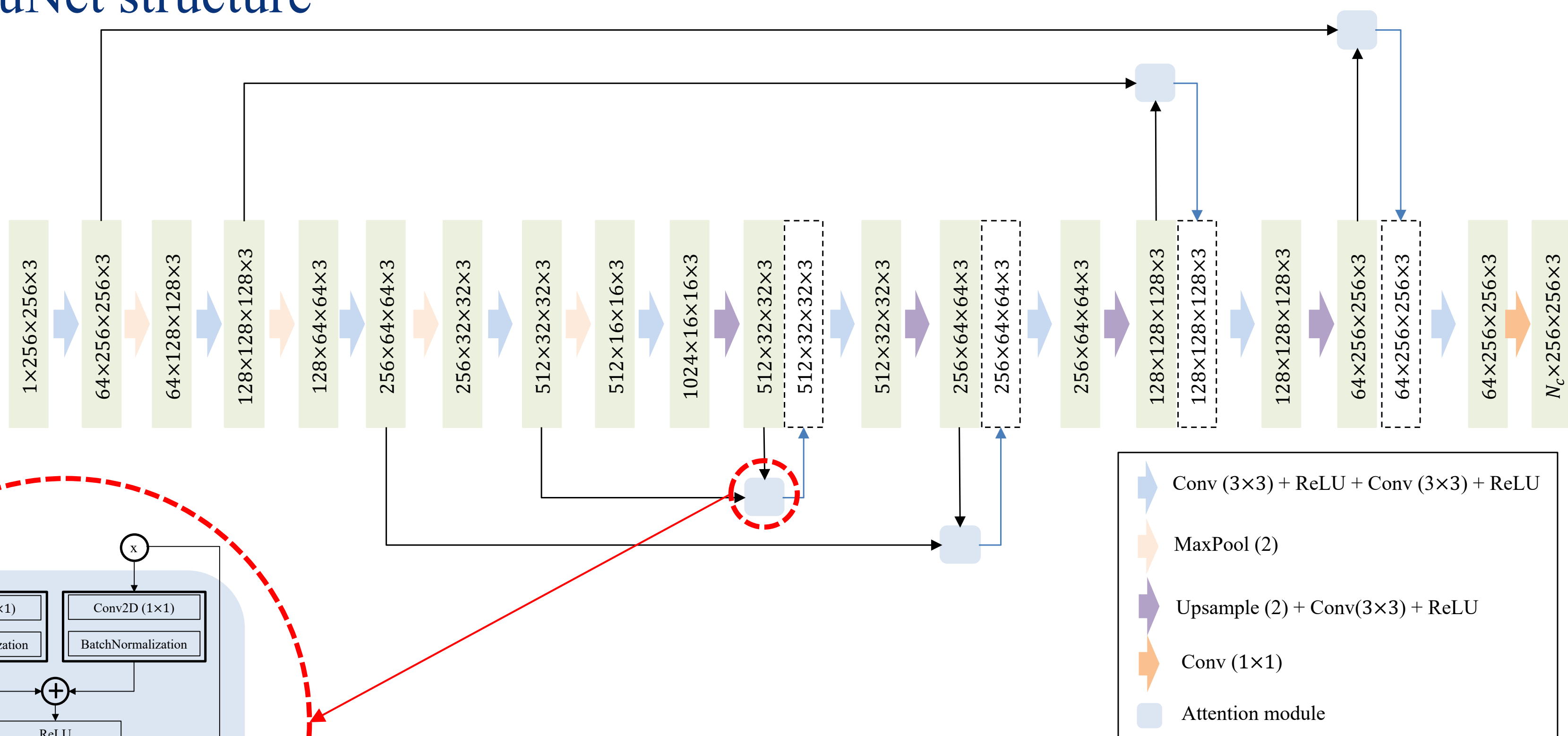
❖ Augmentation results from training set, Random Crop, DCGAN and APC-GAN

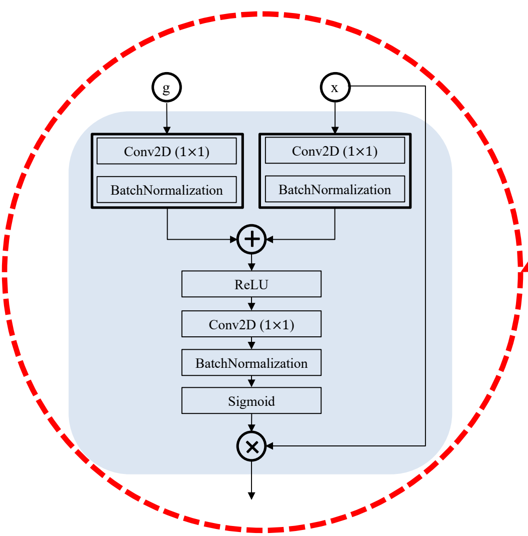


❖ Evaluation results

Data	**Augmentation**	P	R	F1	**mIoU**	mPA
DeepCrack	None	0.950	0.839	0.892	0.812	0.839
	APC-GAN	0.947	**0.868**	**0.906**	**0.836**	**0.868**
	DCGAN	0.949	0.851	0.897	0.822	0.851
	Random Crop	0.950	0.856	0.900	0.827	0.856

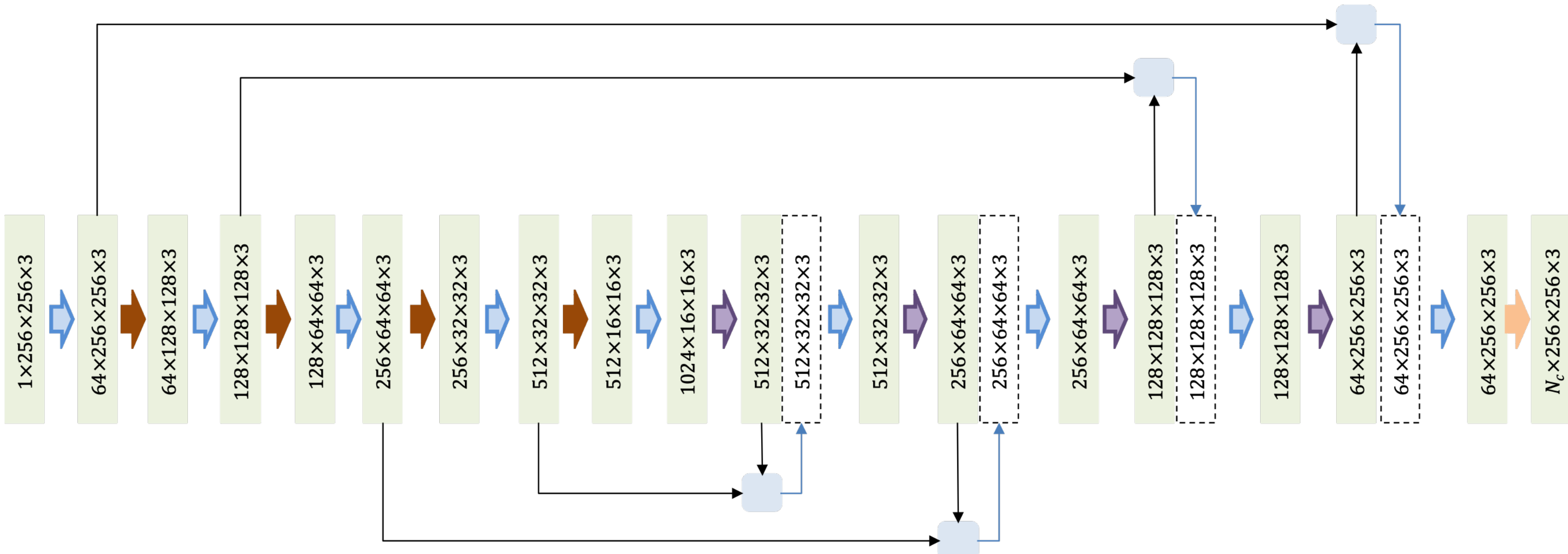
AttuNet structure



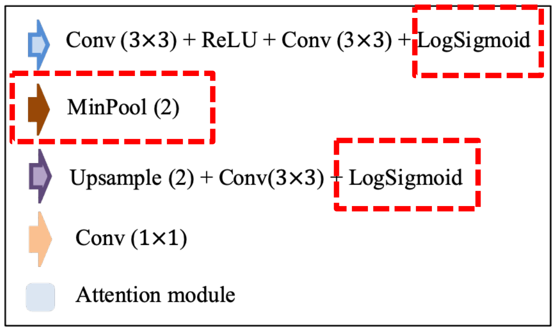


- This structure is modified from U-Net.
- Four attention modules are introduced to connect different layers. Each attention module has two inputs, one from the current layer and the other from a previous layer. Then, the output from attention module would be concatenated with the current layer.

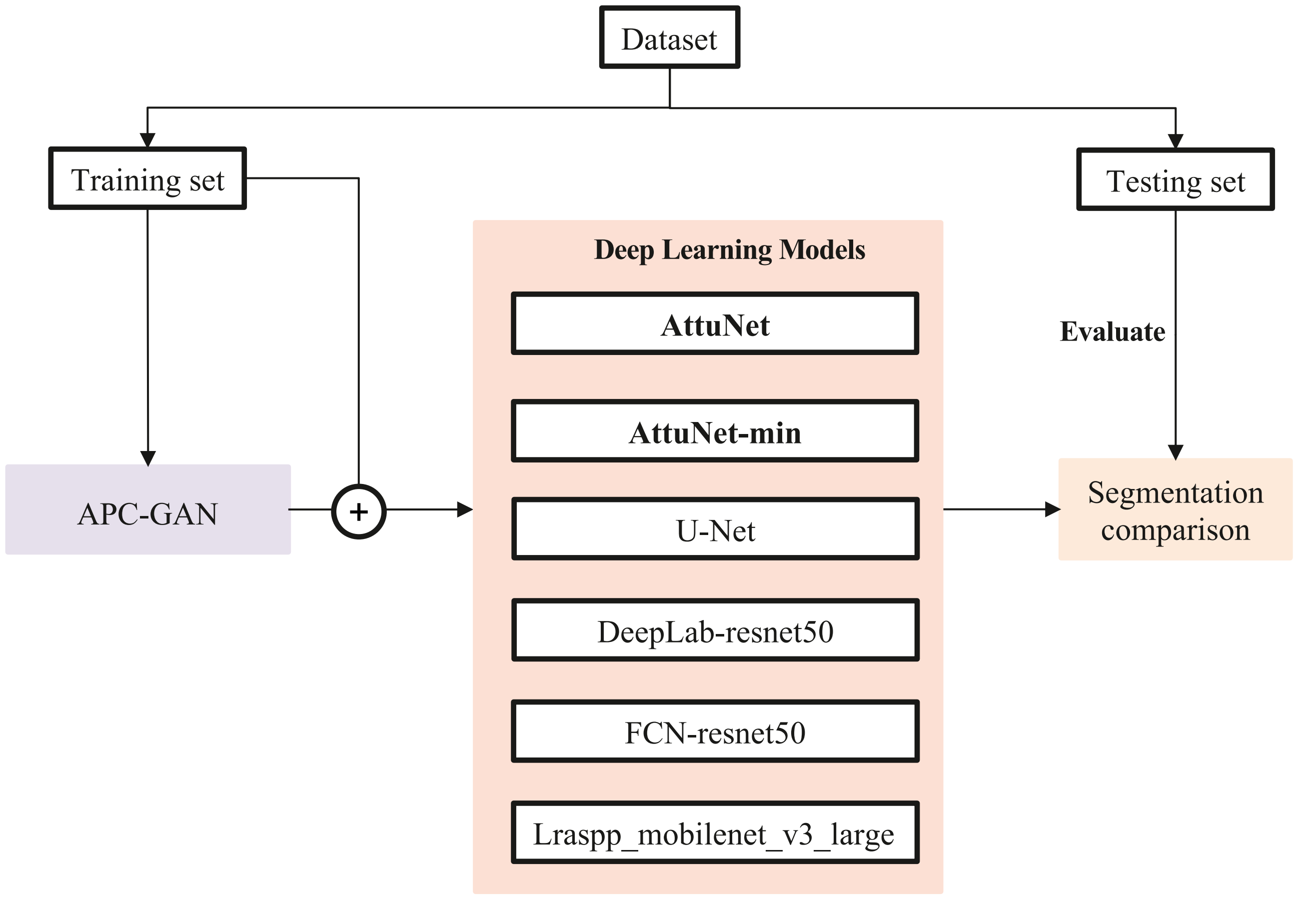
AttuNet-min



- ✓ The **max pooling layer** is replaced by the **min pooling layer**.
- ✓ The ReLU activation function is replaced by the LogSigmoid function.

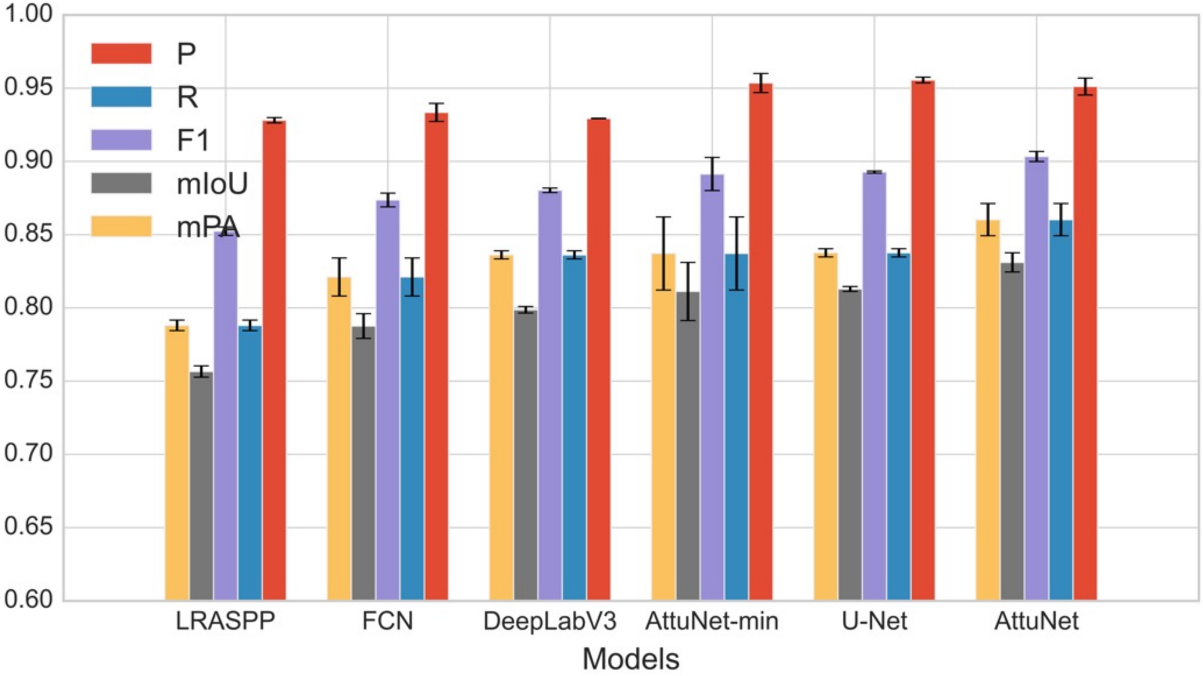


Validation process of AttuNet



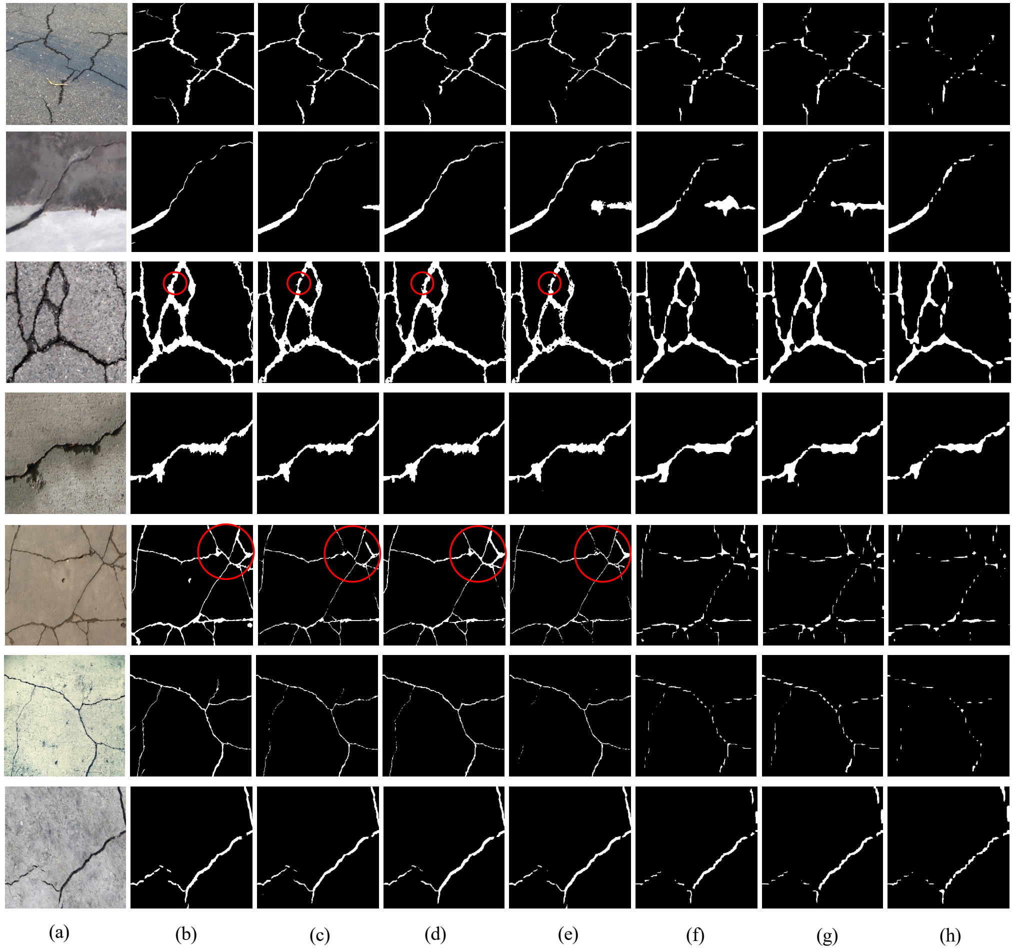
- ❖ Procedure to evaluate the capacity of AttuNet and AttuNet-min

Results



- The models in the Figure are **ranked by the mIoU** from lowest to highest.
- It shows that the **AttuNet** gets the highest mIoU (**0.831**) among all the models. It also gets the highest value in mPA, followed by **AttuNet-min** and U-Net, which means that the AttuNet has the highest performance.

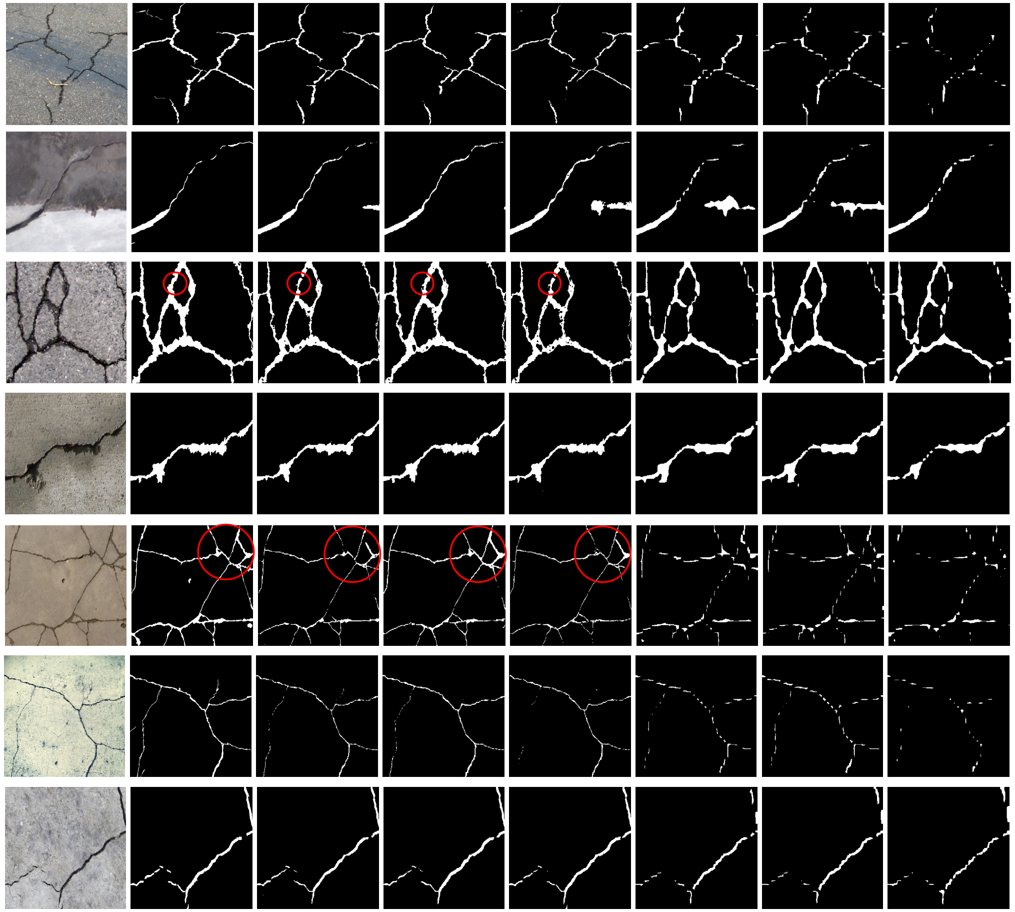
Results



- Several samples with cracks in various scenes and their segmentation results using different methods:

(a) Original image
(b) Ground Truth
(c) **AttuNet**
(d) **AttuNet_min**
(e) U-Net
(f) FCN_resnet50
(g) Deeplabv3_resnet50
(h) LRASPP_mobilenet_v3_large

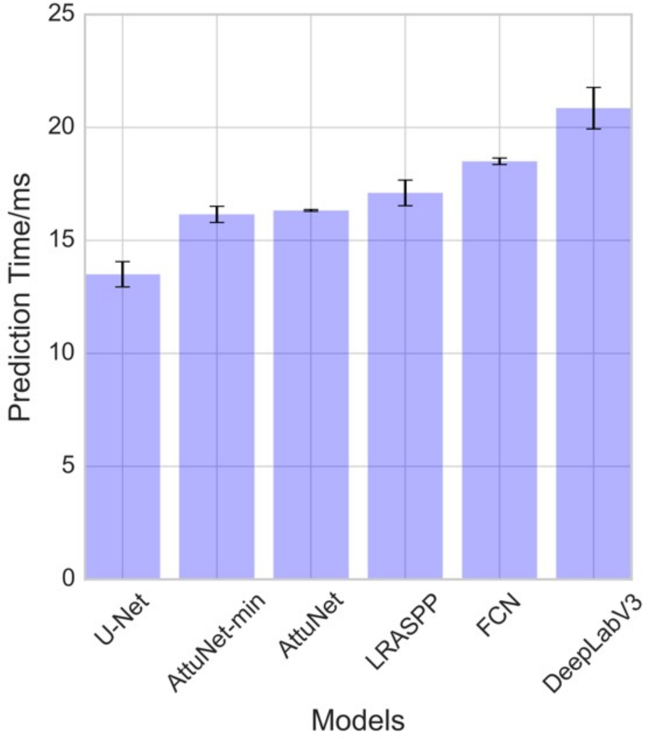
Results


(a) (b) (c) (d) (e) (f) (g) (h)

- The results from AttuNet and AttuNet_min are more complete and accurate than the segmented results from FCN-resnet50, DeepLab-resnet50 and LRASPP_mobilenet_v3_large.

- The segmented part in the red circle shows that the cracks are segmented more entirely by AttuNet_min than by AttuNet and U-Net. It shows that the AttuNet-min has a good performance in the continuous of the cracks as the segmentation image is much closer to the ground truth.

Results



- If we want to use the segmentation model in real-time crack detection work, the **prediction time per image** is an important factor.
- A faster prediction time of one model means this model is more suitable for real-time jobs.
- As we can see, the U-Net model consumes the least time while the DeepLabv3 consumes the largest time. The mean prediction time of AttuNet and AttuNet-min is 16.32 ms and 16.15 ms, respectively, which is relatively low.
- Perform better than LRASPP, FCN and DeepLabV3.

Conclusion

- The lack of training dataset is a very common situation in road maintenance as the cost of obtaining a large number of pavement top-view images and labeling these manually is very high. However, a small training dataset may cause the neural network overfitting and bad model performance in robustness. Therefore, this work proposed a novel pixel-level crack segmentation strategy for this problem.

- The proposed framework consists of an image augmentation method (**APC-GAN**) and a deep learning-based structure (**AttuNet**), which can accurately segment images with a small dataset.

- In this work, only a total of 300 color images are used for training. The performance of APC-GAN is evaluated and it shows a better ability to produce sharper contrast and more diverse images compared to DCGAN and Random Crop.

- The proposed AttuNet model combines the attention module with the CNN network. It gets the highest performance among the classic CNN models including U-Net, DeepLabv3, FCN and LRASPP. Compared with AttuNet, the AttuNet-min achieved a more continuous segmentation result by applying the min pooling layer and LogSigmoid activation function.

Publications

1. **Zhang, T.;** Rahman, M.A.; Peterson, A.; Lu, Y. Novel Damage Index-Based Rapid Evaluation of Civil Infrastructure Subsurface Defects Using Thermography Analytics. Infrastructures 2022, 7, 55. https://doi.org/10.3390/infrastructures7040055

2. Behzadian A, Muturi TW, **Zhang T**, Kim H, Mullins A, Lu Y, Owor NJ, Adu-Gyamfi Y, Buttlar W, Hamed M, Aboah A. The 1st Data Science for Pavements Challenge. arXiv preprint arXiv:2206.04874. 2022 Jun 10.

Submitted manuscripts

1. **Tianjie Zhang**, Donglei Wang, Amanda Mullins, Yang Lu. Benchmark Study of Convolutional Neural Networks for Pavement Cracks Classification. International journal of pavement engineering, under review.

2. **Tianjie Zhang**, Donglei Wang, Amanda Mullins, Yang Lu. Integrated APC-GAN and AttuNet framework for automated pavement crack pixel-level segmentation: A new solution to small training datasets. IEEE Transactions on Intelligent Transportation Systems. Under review.

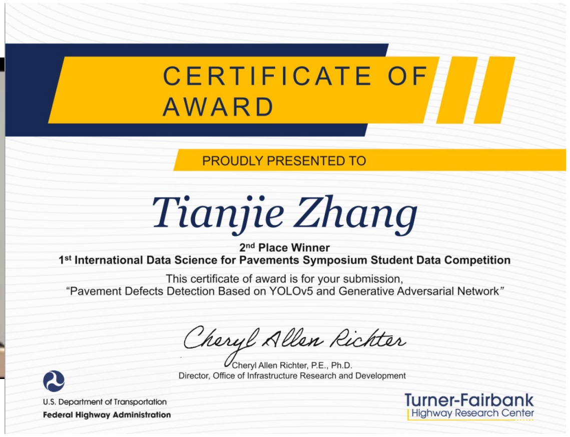
3. **Tianjie Zhang**, Donglei Wang, Yang Lu. RheologyNet: A New Solution to Predict the Rheology Properties of The Cementitious Materials. Cement and concrete research. Under review.

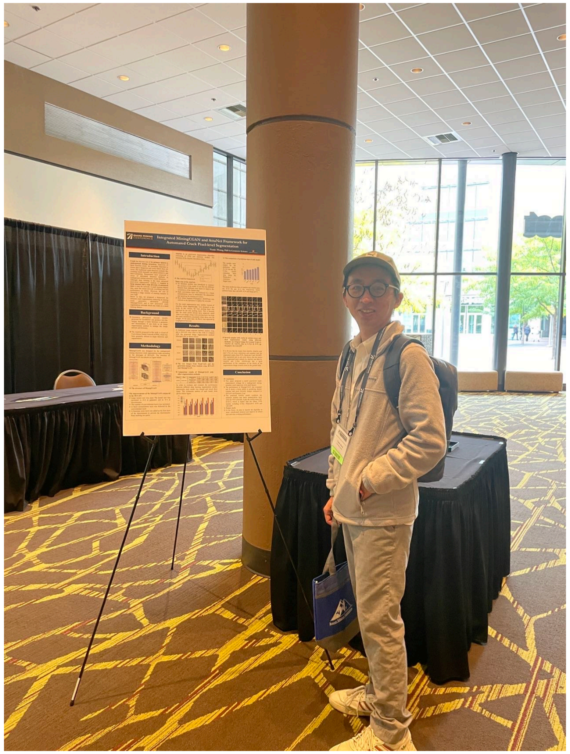
Competition and Conference

DSPS
Data Science for Pavements Symposium 2022

U.S. Department of Transportation
Federal Highway Administration

IDAHO MINING CONFERENCE





https://pavementdatascience.com/home

https://www.boisestate.edu/computing/2022/05/10/bsu-student-team-places-2nd-in-fhwa-student-data-competition/

https://idahominingconference.org/

Thank you!

Questions?

Tianjie Zhang

Student ID: 114178873

Email: tjzhang@u.boisestate.edu