Integrated APC-GAN and AttuNet Framework for Automated Crack Pixel-level Segmentation

Tianjie Zhang, Donglei Wang, Amanda Mullins and Yang Lu*

BOISE STATE UNIVERSITY

Introduction

Crack has become one of the **primary defects** in **Pavements**. Inspection of these cracks are incredibly important to ensure that a dangerous and costly failure does not occur. The traditional crack detection methods like counting cracks manually are labour-intensive and time-consuming. Researchers have proposed a series of automated detection methods in pavement cracks detection.

Although deep learning is the most advanced pixel-level segmentation method, it **requires a large amount and a wide diversity of annotated data** to train the network. A small training dataset may cause the neural network overfitting and bad performance in robust. However, the cost of obtaining a large number of training samples is very high.

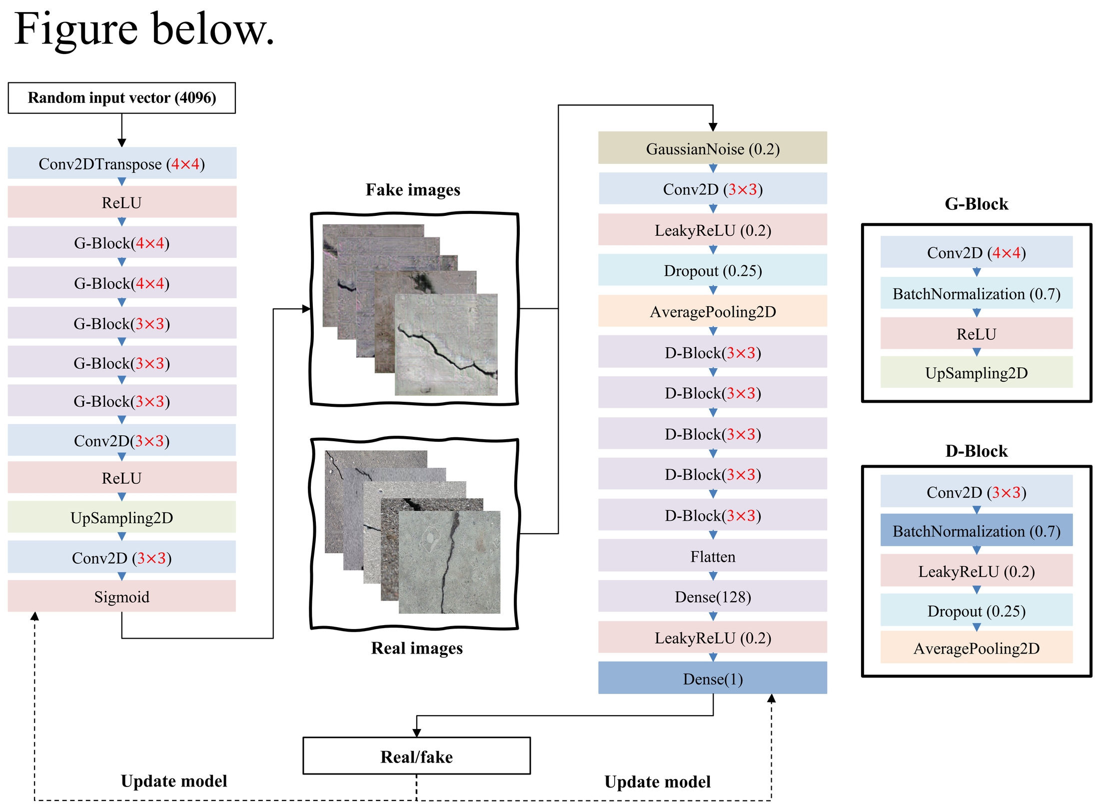
Based on this, we propose a framework for pavement crack segmentation containing an automatic pavement crack generative adversarial network (APC-GAN) and a new pixel-level crack segmentation network which can work on a small training dataset.

Background

- ❖ Generative adversarial network (**GAN**) proposed by Goodfellow can produce real-like images through a battle between a generator and discriminator. It can work as an image augmentation method to enlarge the image amount and diversity.
- ❖ The AttuNet proposed in this work is a kind of Convolution Neural Network (**CNN**), which has been gradually utilized in target detection and segmentation.

Methodology

APC-GAN is designed for the pavement crack segmentation tasks according to the shortages of DCGAN. The structure of APC-GAN is shown in Figure below.

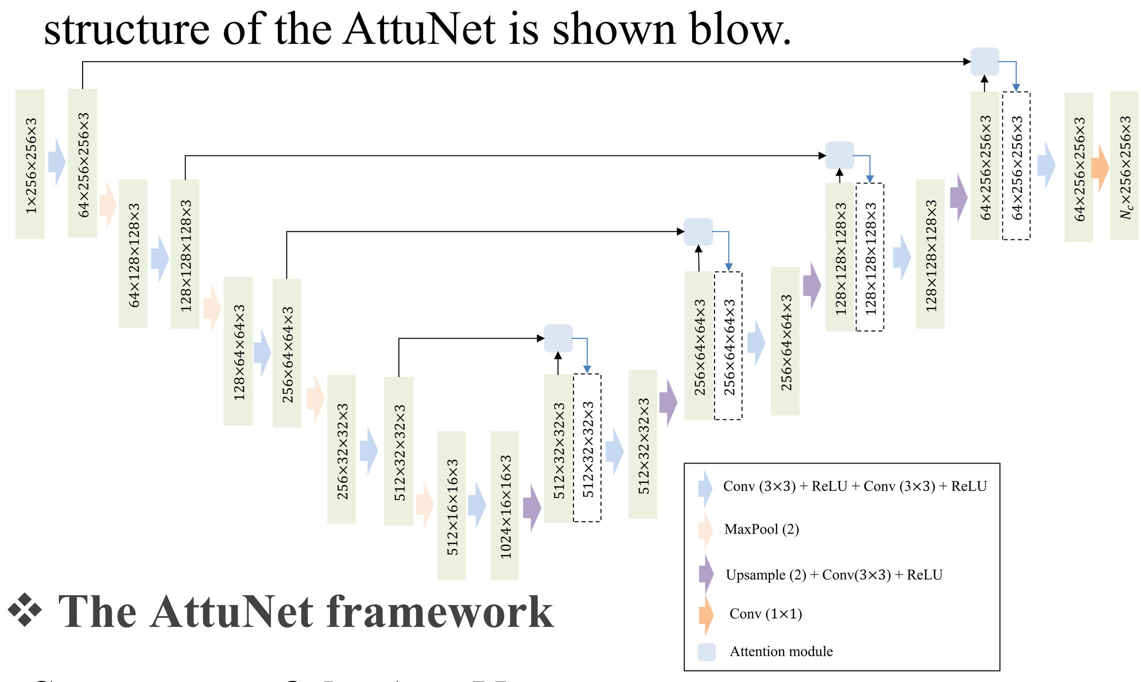


❖ The structure of APC-GAN

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.
- ➢ 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.

AttuNet is an pixel-level segmentation algorithm consisting of CNN and Attention module. The structure of the AttuNet is shown blow.



❖ The AttuNet framework

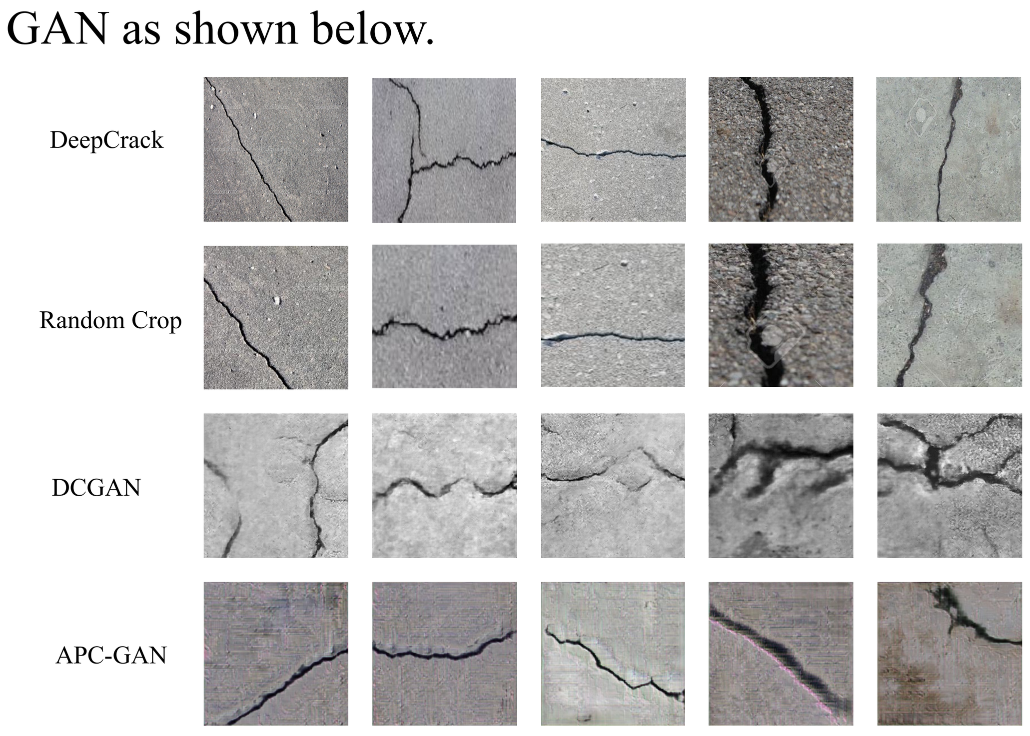
Some cons of the AttuNet:

- ➢ An attention module is introduced to increase the accuracy of the model. It can fuse the different features from different scale layers to improve the consistency of the feature map
- ➢ Each convolution layer is followed by a batch normalization layer to standardize the parameters and speed up the training procedure.
- ➢ Root Mean Squared Propagation (RMSProp) is utilized as the optimizer to update the network.

For the crack segmentation task, another version of AttuNet, called **AttuNet-min**, is designed in this work. In this version, the max pooling layer is replaced by the min pooling layer. This is because that the crack pixels always have relatively small value in an image, using a min pooling layer can keep the crack information accurately when down size the images. At the same time, the ReLU is replaced by LogSigmoid function .

Results

A traditional image augment method, random crop, and a DCGAN are used in this work to compare with the proposed image augmentation method: APC-GAN as shown below.

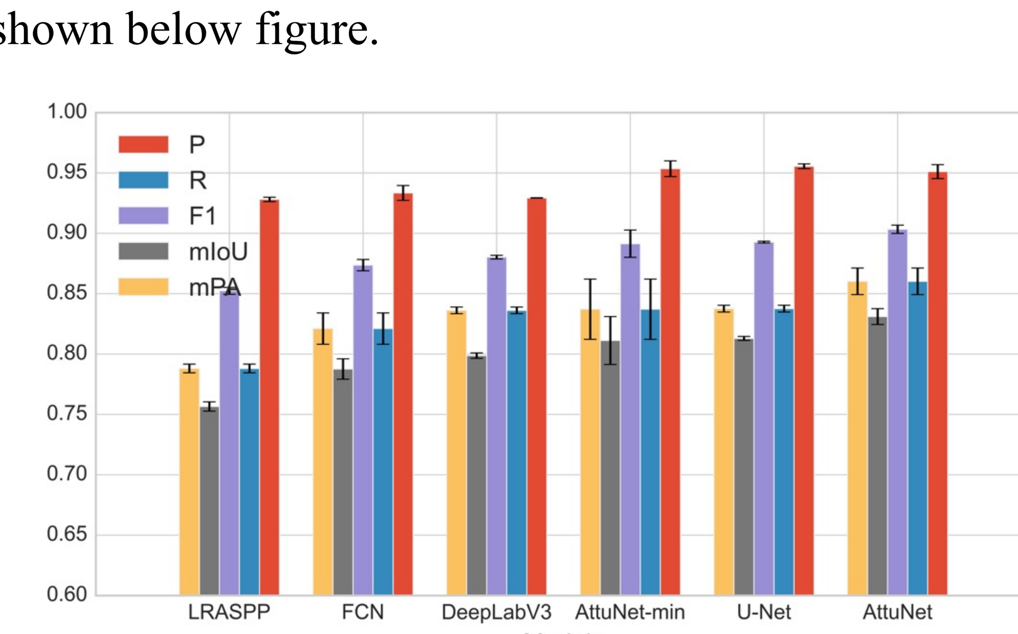


❖ **The raw images from DeepCrack and the generated images from random crop, DCGAN and APC-GAN.**

❖ **Comparison results of APC-GAN with other augmentation methods.**

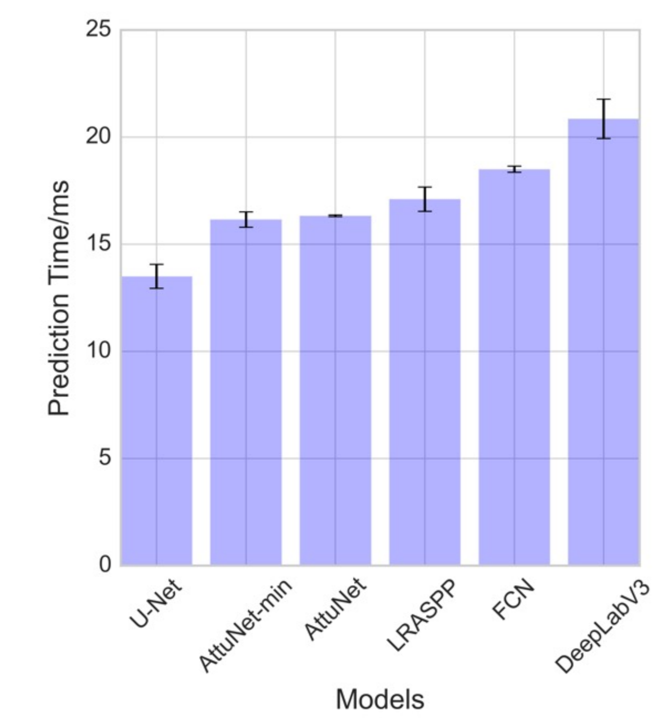
Data	**Augmentation**	P	R	F1	**mIoU**	mPA
DeepCrack	None	**0.950**	0.839	0.892	0.812	0.839
	APC-GAN	0.947	**0.868**	**0.906**	**0.836**	**0.868**
	DCGAN	0.949	0.851	0.897	0.822	0.851
	Random Crop	**0.950**	0.856	0.900	0.827	0.856

In order to evaluate and compare the performance between different deep learning segmentation models statistically, each model was trained and tested three times. The evaluation metrics of each model were shown below figure.

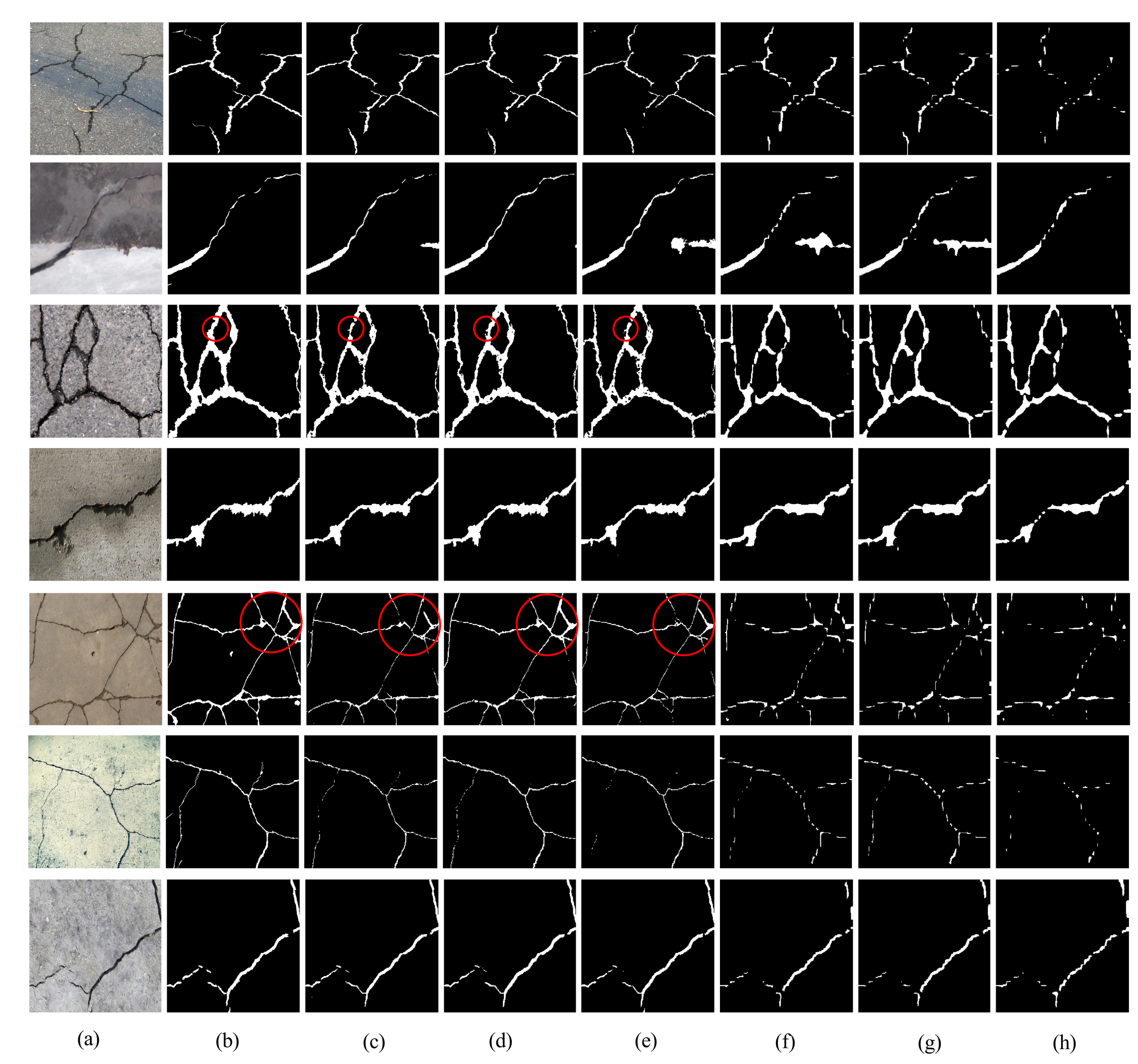


❖ **The comparison of prediction time of each model**

The models are ranked by the prediction time. As we can see, the U-Net model consumes least time while the DeepLabv3 consumes the largest time.



The mean prediction time of AttuNet and AttuNet-min are 16.32 Ms and 16.15 Ms, respectively, which perform better than LRASPP, FCN and DeepLabV3.



❖ **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 (g) Deeplabv3 (h) LRASPP**

It is obvious that the results from AttuNet, AttuNet_min and U-Net are more continuous and complete than the segmented results from FCN, DeepLabv3 and LRASPP. The segmented part in red circle shows that the cracks were segmented more entirely by AttuNet_min than by AttuNet and U-Net. AttuNet-min has a good performance in the continuous of the cracks as the segmentation image is much closer to the ground truth.

Conclusion

❖ **Conclusion**

- ➢ This paper proposed a novel pixel-level crack segmentation strategy for pavement crack inspection with small dataset.
- ➢ The performance of APC-GAN is evaluated and it shows a better ability in producing sharper contrast and more diversity images compared to DCGAN and Random Crop.
- ➢ The proposed AttuNet model combines the attention module and batch normalization layer with the CNN network. It gets the testing mean IoU (0.831), which is higher than the classic CNN models including U-Net, DeepLabv3, FCN and LRASPP.
- ➢ **Future work**
- ➢ In the future, we plan to transfer the algorithm to more complicated situation and improve its robust and real-time working performance.