

東京大学大学院新領域創成科学研究科

複雑理工学専攻

令和 4 年度

修士論文

(要旨)

**Toward Faster and Accurate Post-disaster Damage Assessment:  
Development of End-to-End Building Damage Detection  
Framework with Super-resolution Architecture**

2022 年 9 月 2 日提出

指導教員 吉川 一郎 教授

富 宣超

**Toward Faster and Accurate Post-disaster Damage Assessment:  
Development of End-to-End Building Damage Detection  
Framework with Super-resolution Architecture**

**Xuanchao Fu**

47-206131

Complexity Science and Engineering

Completed: September 2022

Keywords: *Remote Sensing, Deep Learning, Super-resolution, Semantic segmentation,  
Building damage, Natural disasters*

# Toward Faster and Accurate Post-disaster Damage Assessment: Development of End-to-End Building Damage Detection Framework with Super-resolution Architecture

東京大学大学院 新領域創成科学研究科 複雑理工学専攻

Yoshikawa Lab. Fu Xuanchao

## 1. INTRODUCTION

The damage level of buildings in various natural disasters typically has a strong spatial correlation with human casualties and economic losses, and thus there is an urgent need for high-accurate building damage detection (BDD) approach with low response time to support post-disaster rescue and response. In recent years, BDD approaches based on high-resolution (HR) remote sensing images (e.g., less than 1m) and deep learning techniques have been extensively researched and applied with increasing of the remote sensing satellite performance, especially the convolutional neural network (CNN)-based models have even performed well in BDD practice after various severe disasters in recent years and has become a standard structure of BDD approaches.

As another aspect of BDD, in recent years, increasingly small/nanosatellites have been launched (more than 1,000 satellites per year), and the constellation of satellites has been realized. In such a situation, continuous monitoring of hazard damage can be done, while the faster decision for planning the next observation based on earlier satellite observation must be important. The human decision may be a bottleneck for such operation, and thus it is important to establish an automated process in a ground station, such as the Kashiwa-Campus site, which can provide “human-like” results and replace human planning. This system is referred to as Smart Ground Station (Smart GS).

However, since previous CNN-based BDD approaches rely heavily on high-resolution remote sensing satellites with long revisit times (typically, > 24 hours), the images that SmartGS can acquire at high frequency are mainly low-resolution (LR) images taken by small/nanosatellites, it is less practical to be integrated into Smart GS. Therefore, incorporating super-resolution architecture into the BDD approaches to perform accurately and detail BDD even in the case when only low-resolution images can be used after a disaster is an effective solution

Based on Super-Resolution Generative Adversarial Network (SRGAN) and U-Net convolutional network, an efficient and novel BDD framework is proposed in this research for obtaining high-accuracy BDD results from HR pre-disaster images and LR post-disaster images. We trained the framework using two disasters from the xBD dataset and tested three different structures. The results show that our End-to-End training framework performs significantly better than the results of the two-stage training framework. We also discussed the reasons for its advantages, and its

generalizability and expandability, through small-scale experiments.

## 2. DATA AND METHOD

xBD dataset is a widely known natural disaster building damage dataset, which provides paired pre/post-disaster HR (GSD = 0.5m) three-band RGB optical images, it is well suited to simulate the “without post-disaster HR images” scenario, which we assumed (Gupta et al., 2019). For better quantitative evaluation of the performance of the SR/BDD models for tsunami disaster cases, we selected images of two disasters from the dataset, the 2018 Sunda Strait tsunami and 2018 Sulawesi earthquake and tsunami, and reclassified the building damages levels from the original four types (no damage, minor damage, major damage, destroy) to two types (no damage, damaged or destroyed).

In data processing, 1024 size images are randomly cropped to get 256 size images and save the images with building. Then, the images with high cloud coverage are automatically removed using the pre-trained EfficientNetV2 image classification model. For super-resolution training, we produced realistic LR post-disaster images (1/2, 1/4 and 1/8 the size of HR) by Gaussian blurring and Bicubic as much as possible. The format of the processed dataset is shown in Fig. 1 (blue: no damage building, yellow: damaged or destroyed building).

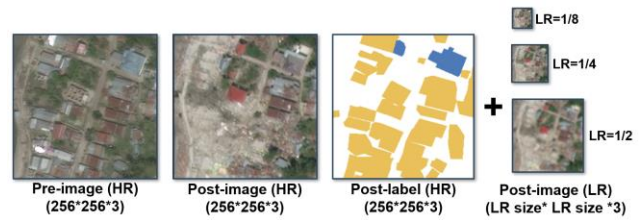


Figure 1

As illustrated in Fig. 2, our proposed framework consists of the SR and Seg modules. The SR module at the first part is used to transform the LR post-disaster images into SR post-disaster images of the same resolution as the HR images, while the Seg module at the second part is used to evaluate the damage of buildings from the paired HR pre-disaster images and SR post-disaster images. Based on this overall structure, we tested with three BDD frameworks with different structures and training process in this research to find an appropriate structure for our purpose.

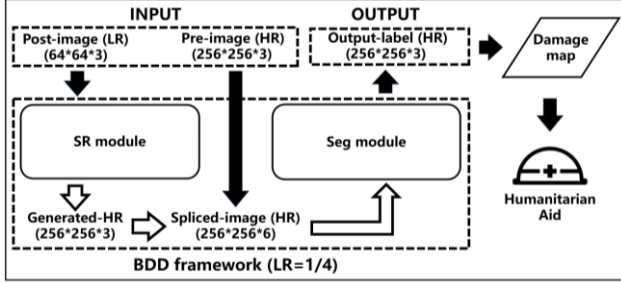


Figure 2

We construct a double-modal model based on the U-Net (Ronneberger et al., 2015), we call it as BDD-U-Net as the Seg module in the framework. This structure can concatenate features of both pre-disaster and post-disaster images extracted from each convolutional layer. And the models used as SR modules in this paper are divided into two types: SRGAN and BDD-SRGAN. Both models have the same structure of generator and discriminator as the original paper of SRGAN (Ledig et al., 2017), with the difference that SRGAN uses the original loss function, while the loss function of BDD-SRGAN is modified. The original SRGAN uses one of the feature maps of the VGG network pre-trained on ImageNet for calculating the perceptual loss, while in BDD-SRGAN, the above VGG network is replaced with the last feature map in the encoder of the post-disaster image in the pre-trained or in-training BDD-U-Net. For End-to-End training, BDD-U-Net can use the output of BDD-SRGAN as post-HR.

Based on the above BDD framework structure and modules, the structure and training process of the three frameworks we proposed (SR-B, BSR-B, EEBSR-B) and the two frameworks that served as the control group (HR-B, BIC-B) are shown in Table 1. We chose the Overall F1 score as the main evaluation criterion for the performance of the BDD framework and tested the SR module by PSNR/SSIM and visual effects to evaluate the SR performance.

Framework	SR/Seg module	Training
HR-B	HR images/BDD-U-Net	Seg only
BIC-B	Bicubic/BDD-U-Net	Seg only
SR-B	SRGAN/BDD-U-Net	independently
BSR-B	BDD-SRGAN/BDD-U-Net	Seg→SR
EEBSR-B	BDD-SRGAN/BDD-U-Net	Seg→SR (Each Batch)

Table 1

### 3. RESULTS AND DISCUSSIONS

We compared the training results of all five frameworks at LR sizes of 1/2, 1/4 and 1/8. Their Overall F1 scores results are shown in Figure 3. The results show that the EEBSR-B outperforms other frameworks in terms of Overall F1 scores and approaches HR-B at all resolutions, which can also be

observed in Figure 4. On the other hand, by observing the results of PSNR/SSIM, we also found that the PSNR and SSIM/EEBSR-B did not always outperform the two-stage training frameworks. This can be explained that the EEBSR-B optimized to obtain more accurate BDD results rather than obtain accurate SR.

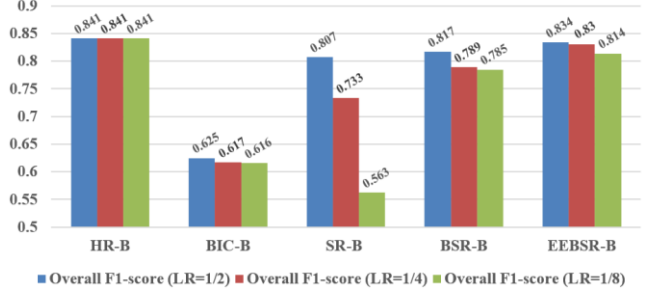


Figure 3



Figure 4

As a discussion of the above results, we imported the SAM (Woo et al., 2018) into the SR module, and by analyzing its output attention maps, we believe that one reason for the advantage of EEBSR-B could be its more focus on the building boundaries. We also selected two other disasters (wildfire and tsunami) from xBD and trained the HR-B, BIC-B, and EEBSR-B frameworks in the same way as above, and the results show that EEBSR-B still achieves performance close to HR-B in the other disaster datasets, demonstrating a degree of generalizability. Moreover, we have demonstrated through comparative experiments that by replacing the basic structure of the SR module with the more advanced ESRGAN (Wang et al., 2018), we can improve the performance of EEBSR-B by a small margin, demonstrating a degree of expandability.

### 4. SUMMARY

This study validates that obtaining high accuracy BDD results without HR post-disaster images is practically feasible and has the potential to be of great help for future post-disaster response efforts. There are reasons to believe that designing BDD frameworks that perform better in terms of accuracy and timeliness with integration into the Smart GS system is of great significance and is an important topic for future works.