WEI ZHAI
Department Urban and Regional Planning
University of Florida
Email: wei.zhai@ufl.edu

Wei Zhai is currently a third-year Ph.D. student in the Department of Urban and Regional Planning at the University of Florida (UF). He is also pursuing a dual master's degree in Electrical and Computer Engineering. Prior to UF, he was trained as a city planner in his undergraduate and master's majors in China. Given the background in urban planning, he is interested in developing and applying new approaches to solve urban and transportation problems in practice.

Zhai's previous research expertise has applied state-of-the-art machine learning techniques with geotagged big data to solve practical problems in planning and transportation. He is interested in incorporating Natural Language Processing techniques with spatial analyses. For instance, in his published work he has employed Place2vec to identify urban functional regions based on POIs (https://doi.org/10.1016/j.compenvurbsys.2018.11.008). In addition, he has developed an approach, Augmented Space-time-weighted Edit Distance, to extract the similarities between human activities by comparing the human activities to natural sentences (https://doi.org/10.1016/j.jtrangeo.2019.05.003).

As a research assistant at the International Center for Adaptation Planning and Design at UF, Zhai's current research focuses on developing the near real-time system for human-hazard interaction during a natural disaster by using geotagged social media data, images and human activity data. He is working on using deep-learning approaches to extract situational awareness information from geotagged Twitter data. He has also proposed a transfer learning-based model to quantify damage levels in geotagged images and videos, testing the performance by using Google Street View after a hurricane. To predict the real-time travel demand in high-resolution urban space during natural disasters, he is developing a deep learning-based sequence model by considering the local event data and real-time weather data.

Taking advantage of the opportunities offered by the Age of Artificial Intelligence, Zhai's ultimate research goal is to design automatic solutions for smart, sustainable, resilient cities. He will continue his research in employing spatial data, spatial analysis approaches, and machine learning techniques to analyze, monitor, and manage inclusive and resilient urban growth.

Detecting Disaster Damage based on Geotagged Images: A Case Study of Google Street View Images

Conventional damage assessment largely relies on remote sensing imagery because it has the advantage of large-scale damage evaluation. However, it cannot reveal the profile view of the streets to display damage from the perspective of the human eye. Recently, though there have been studies employing Google Street View (GSV) images to display and evaluate damage, they did not quantitatively measure the visual damage. To address this gap, we employed transfer learning based on the VGG 19 architecture and trained the labeled "damage" and "no damage" images. We conduct an empirical study in Mexico Beach, Florida, which was wiped out by Hurricane Michael in 2018. The result reveals the difference between the damage perceived by human vision system and the damage identified by the satellite image. This also demonstrates that previous methods on damage detection cannot completely reveal the circumstance perceived by people. The techniques in this study can be used as a near-real-time approach for FEMA or local response teams to quickly evaluate damage severity once images or videos are taken.

1. Methodologies

The proposed approach generally contains four parts. (1) From the social media platform and Google Images, we can crawl damage-related images and then label the dataset into two categories: "damage" and "no damage." (2) By fine-tuning the VGG 19 model (Simonyan and Zisserman, 2014), we can train the labeled training set and evaluate the model performance based on the test set. Specifically, in this work, two fine-tuned VGG 19 nets are proposed based on the Class Activation Map (CAM) and gradient-weighted class activation map (grad-CAM) methods. (3) To show the damaged area in the GSV image, CAM and grad-CAM are applied. Then, the damage severity can be calculated based on the intensity channel of the GSV image. (4) We validate the results of the detected VD from two aspects. First, we assess the detected VD in the GSV image by comparing it with manually delineated damage. Second, we show the spatial distribution of the detected damage and overlay it with remotely sensed imagery. Correlation analysis is applied for both evaluation steps.

2. Results

2.1 Validation of Visual Damage (VD) results

We trained the fine-tuned VGG 19 nets and tested the generalization ability in the test set. To localize the damage region in the image, CAM and grad-CAM are incorporated with VGG 19 nets. For the CAM-based architecture, the training accuracy in our experiment was 92.21±0.23, and the test accuracy was 81.89±0.35. For the grad-CAM-based architecture, the training accuracy in our experiment was 90.02±0.43, and the test accuracy was 78.12±0.55. We then applied the trained models to predict all the GSV images captured in Mexico Beach.

To show the effectiveness of the proposed approach and compare the performance between CAM and grad-CAM, we randomly selected 66 GSV images. Then, we manually delineated the damaged area in each image, and the human annotated image was defined as the delineated damage map (DDM). We used the percentage of the manually delineated area in the image as the damage value for the DDM. Fig. 1 shows examples of the original GSV image, the automatic damage map (ADM) derived from CAM, the ADM derived from grad-CAM, the DDM and the extraction of the delineated damage area. Fig. 2 indicates the accuracy of damage detected in a single image.

To examine whether the VD is related to remotely sensed damage (RSD), we overlaid the spatial distribution of VD with the RSD map. More specifically, the RSD was extracted based on the Hurricane

Michael Imagery provided by the National Oceanic and Atmospheric Administration (NOAA), which can be accessed via https://storms.ngs.noaa.gov/storms/michael/index.html#7. At each sampled site, we created five buffers with a radius from 10 m to 50 m. Then, we used the pixel number within the buffer to represent the RSD. The Pearson's correlation coefficients were applied to indicate the relationship between VD and RSD (Table 1). The result indicates that the correlation coefficient between DVL and RSD captured by 20 m buffered zones was the highest. In other words, damage within the 20 m buffered zone was perceived the most in GSV in this study area. Compared to the CAM, the correlation coefficient under grad-CAM was slightly weaker. In addition, the RSD in the 40 m and 50 m buffered zones were not significantly correlated with the VD.

2.2 Near-real-time Detection of the Disaster

The techniques introduced in this study can be used as a near-real-time approach for FEMA or local response teams to quickly evaluate damage severity once images or videos are taken. The images or videos can be taken by unmanned aerial vehicles (UAVs) or eyewitnesses. Once the deep learning model is trained, the method requires 1.2 seconds to process a single geotagged image on average and 5 minutes and 23 seconds to process this video. The video demonstrates that near-real-time detection of damage can be achieved if data can be rapidly collected during a disaster event.

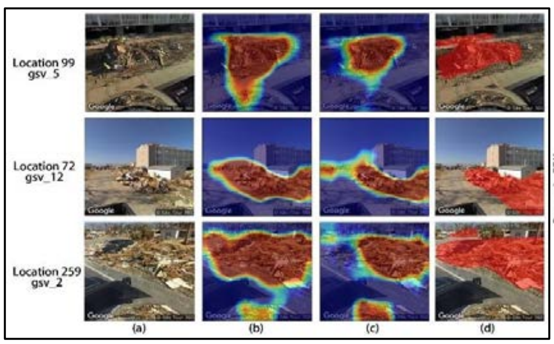

Fig.1 (a) Original image; (b) ADM derived from (c) ADM derived from grad-CAM; (d) DDM

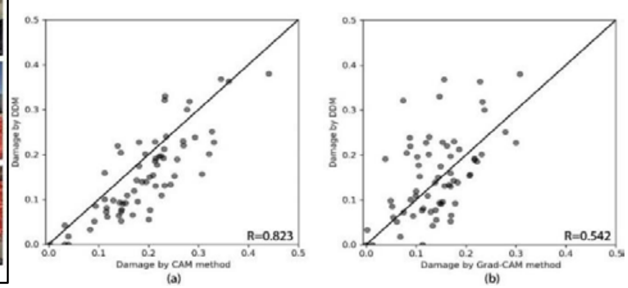

Fig.2 Correlations between ADM and DDM by CAM CAM and grad-CAM models

Table 1. Pearson's correlation coefficients between VD and RSD under different pitch parameters

Pitch Parameters	Buffer Distance	Pearson's correlation coefficients CAM VD and RSD	Pearson's correlation coefficients Grad-CAM VD and RSD
Pitch=-30°,0°,30°	10 m	0.152**	0.122**
	20 m	0.182**	0.121**
	30 m	0.130**	0.068*
	40 m	0.112**	0.044
	50 m	0.076*	0.010

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Reference

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409–1556*.