

Extreme Weather Event Management and Climate Adaptation via Supervised Computer Vision-based Algorithms

Thomas Y. Chen
ECMWF ML Workshop

Background



Computer Vision, Satellite Imagery, and Building Damage Assessment: An Introduction

- Natural Disasters
 - 60,000 Deaths a Year
 - Immense infrastructure damage and economic loss
 - Increasing in frequency and intensity due to climate change
- Satellite Imagery
 - Quick and efficient, aids in the allocation of resources
 - Analyzed with deep learning based approaches to classify building damage



Previous Works

- Image Classification
 - Classical approaches, deep-learning techniques
- Computer Vision for Satellite Imagery
 - Marine ecology, weather forecasting, spread of disease
 - Agriculture, urban road damage
 - Change detection (multi-temporal fusion)



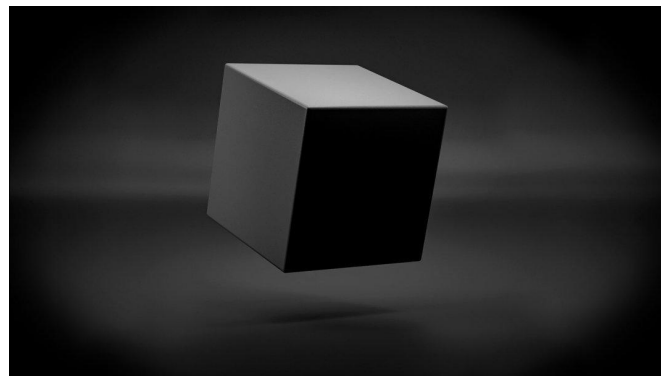
Previous Works

- Building Damage Assessment
 - Semantic building segmentation
 - Cross-region transfer learning
 - Semi-supervised approaches
 - xBD: most comprehensive dataset
- Disaster Relief: Social Media (NLP vs. CV)
- What do we contribute?
 - Interpretability
 - Quantitative and Qualitative



Interpretability

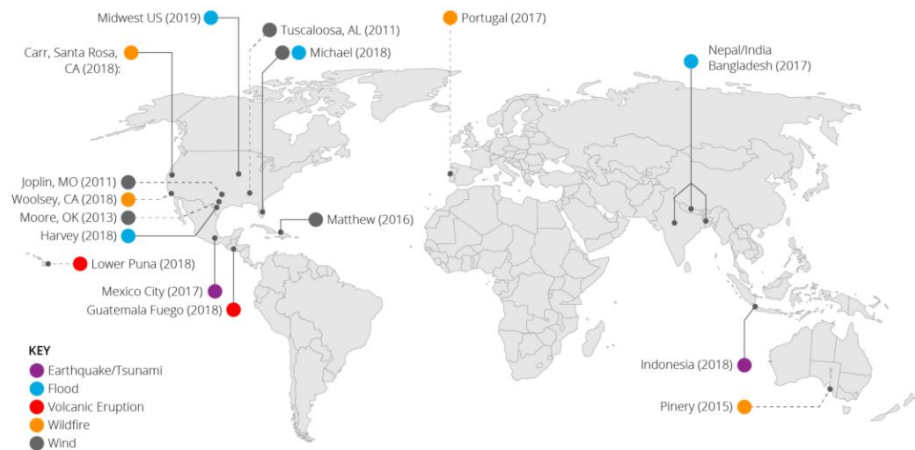
(Def) the degree to which a human can understand the cause of a decision of a machine learning algorithm



Research Process

- Dataset analysis
- Develop a baseline model to classify building damage based on the post-disaster image only
- Develop improvements to the baseline model to classify building damage based on other aspects of the image, namely the pre-disaster image and the disaster type
- Compare the results
- **Understand exactly what these networks are learning (leading to more interpretable models)**

xBD Dataset



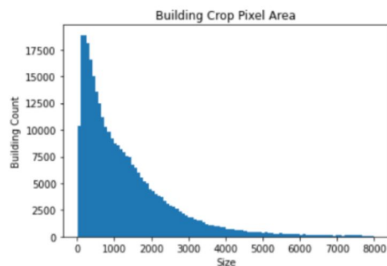
Score	Label	Visual Description of the Structure
0	No damage	Undisturbed. No sign of water, structural damage, shingle damage, or burn marks.
1	Minor damage	Building partially burnt, water surrounding the structure, volcanic flow nearby, roof elements missing, or visible cracks.
2	Major damage	Partial wall or roof collapse, encroaching volcanic flow, or the structure is surrounded by water or mud.
3	Destroyed	Structure is scorched, completely collapsed, partially or completely covered with water or mud, or no longer present.



Source: www.xview2.org

Preprocessing

- Creating building crops for per-building analysis, using labeled building polygons provided
- Discarding small/unclear buildings
- Other cleaning mechanisms
- Train on equally distributed dataset (equal number of crops for each category)



(a) 64 pixels



(b) 6232 pixels

Baseline model

- Based on the post-disaster image only
- ResNet18 (CNN architecture) - pre-trained on ImageNet data
- Cross-entropy loss
- Trained on 12,800 building crops
- Adam optimizer
- Learning rate of 0.001
- 100 epochs
- NVIDIA Tesla K80 GPU

Baseline model

- ResNet18 (CNN architecture) - pre-trained on ImageNet data
- Cross-entropy loss
- Trained on 12,800 building crops
- Adam optimizer
- Learning rate of 0.001
- 100 epochs
- NVIDIA Tesla K80 GPU

Layer Name	Output Size	ResNet-18
conv1	$112 \times 112 \times 64$	$7 \times 7, 64, \text{stride } 2$
conv2_x	$56 \times 56 \times 64$	$3 \times 3 \text{ max pool, stride } 2$ $\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3_x	$28 \times 28 \times 128$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4_x	$14 \times 14 \times 256$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5_x	$7 \times 7 \times 512$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
average pool	$1 \times 1 \times 512$	$7 \times 7 \text{ average pool}$
fully connected	1000	$512 \times 1000 \text{ fully connections}$
softmax	1000	

Baseline model

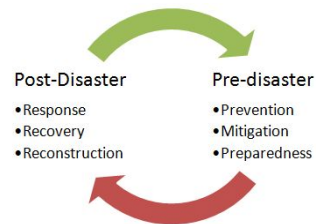
- ResNet18 (CNN architecture) - pre-trained on ImageNet data
- Cross-entropy loss
- Trained on 12,800 building crops
- Adam optimizer
- Learning rate of 0.001
- 100 epochs
- NVIDIA Tesla K80 GPU

$$-\sum_{c=1}^4 y_{o,c} \log(p_{o,c})$$

Layer Name	Output Size	ResNet-18
conv1	$112 \times 112 \times 64$	$7 \times 7, 64, \text{stride } 2$
conv2_x	$56 \times 56 \times 64$	$3 \times 3 \text{ max pool, stride } 2$
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3_x	$28 \times 28 \times 128$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4_x	$14 \times 14 \times 256$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5_x	$7 \times 7 \times 512$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
average pool	$1 \times 1 \times 512$	$7 \times 7 \text{ average pool}$
fully connected	1000	$512 \times 1000 \text{ fully connections}$
softmax	1000	

Improvements

- New types of input: pre-disaster image and disaster type
- Different loss functions:
 - Ordinal Cross-entropy loss
 - Mean squared error
- Other aspects remain the same



Results: Accuracy comparison

Table 1: Comparison of the Validation Accuracy on 9 Different Models

Model Accuracy on Validation Set with Chosen Loss (100 epochs)			
Model Input	Loss Function		
	Mean Squared Error	Cross-Entropy Loss	Ordinal Cross-Entropy Loss
Post-Disaster Image Only	45.3%	59.5%	64.2%
Pre-Disaster, Post-Disaster Images	50.2%	68.3%	71.2%
Pre-Disaster, Post-Disaster Images, Disaster Type	49.7%	72.7%	74.6%

Table 1. Comparison of accuracy on the validation set for nine different models. Unsurprisingly, the models trained on pre-disaster image, post-disaster image, and disaster type (all three modalities) performed the most accurately. Additionally, the models that utilized ordinal cross-entropy loss as their loss function achieved the best results.

Discussion

- Accuracy increases between three models: post-disaster image only, pre-and-post-disaster images, and pre-and-post disaster image plus disaster type
- Reasons for non-optimal accuracy
- Ordinal cross-entropy loss is the best criterion
- Contributes to the study of interpretability in deep learning models that classify building damage

Qualitative Interpretability

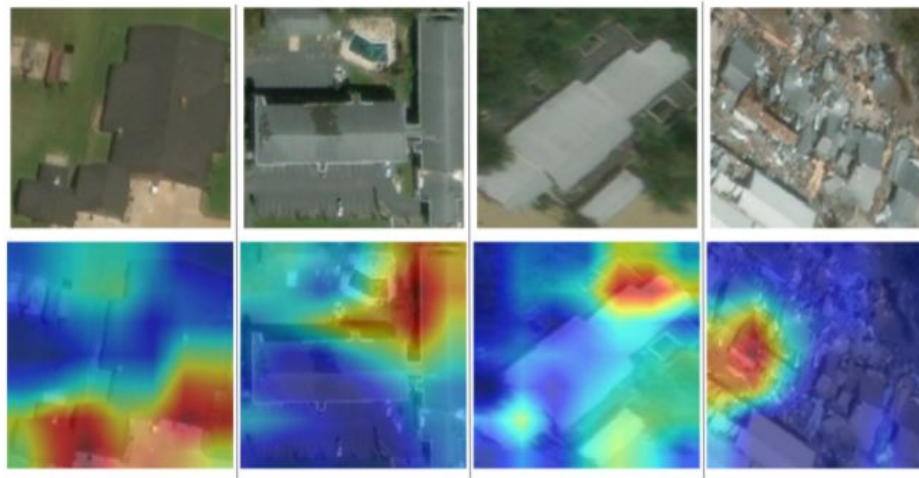


Figure 1: Gradient class activation maps [20] depict which parts of the building crop lead the baseline model to predict a certain classification. On the top are the original images (crops) and on the bottom are the corresponding gradient class activation maps. The images included are only post-disaster images. From left to right: (1) A building with label "no damage," after flooding in the Midwestern United States, (2) A building with label "minor damage," after Hurricane Michael, (3) A building with label "major damage," after Hurricane Harvey, and (4) A building with label "destroyed," after Hurricane Michael.

Conclusion

- We find that inputting different combinations of information does indeed improve model performance.
- Our study leads the way for more effective and efficient damage assessment in the event of a disaster.
- Climate change
- AI/ML is key

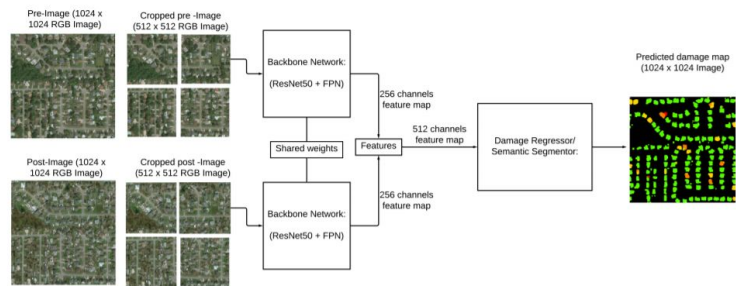


Future Work

- Other modalities of input
 - Neighboring buildings
- Different combination methods of the pre-disaster image and post-disaster image
- Qualitative interpretability - deployment
- Cleaner dataset, more distinct differences between major damage and minor-damage, for instance.



Incidents detection map. Here we illustrate a map that could be created with our Incidents Model. By detecting and geo-locating incidents in tweets, relevant maps can be created to help inform in response efforts.



Thanks for listening. Any questions?

