QUANTIFYING BUILDING DAMAGE USING SATELLITE IMAGERY

Amar Ananth, Madhav Jariwala, Sai Saathvik Domala

Luddy School of Informatics, Fall 2022

ABSTRACT

This project's goal is to use satellite imagery to understand building damage before and after disasters using the xBD Dataset[1]. Climate change, a hot topic in recent years with floods, tornadoes, hurricanes, and tsunamis occurring more frequently and with greater intensity as direct effects of global warming and is expected to cause more severe natural disasters in the near future. Highly-populated civilizations are becoming more and more susceptible to natural disasters. Local damage assessments must be completed promptly, completely, and accurately in order to build an effective and efficient adaption strategy. In order to identify areas impacted by disasters, we suggest a unique deep-learning-based method that compares pairs of pre-disaster and post-disaster satellite pictures of the buildings and applies them to damage prediction and classification. We show that the model can identify local damage with greater accuracy using less sophisticated and less granular ground truth data than those employed by earlier segmentation techniques. We hope to be able to provide a deep-learning framework for automating building segmentation and quantifying damage levels. This framework's applications would include emergency disaster response, vulnerability analysis, and threat demarcation.

1. INTRODUCTION

Automated and rapid damage assessment of infrastructure following disasters is critical for providing emergency response and resources. Though damage assessments done by the ground crew are more accurate, they are also time-consuming and require great physical effort. Aerial imagery has been able to capture buildings with extreme precision. Large-scale post-disaster building damage assessments can be aided by the near real-time and high-coverage information provided by satellite images. U-Net[2] with different combinations[3] and combinations of U-Net and transformers[4] have been widely used in building segmentation and building damage assessment. In this report, we would be focusing on using novel deep-learning techniques to perform building segmentation and building damage classification. We implement a Siamese Convolution Neural Network with ResNet as its base containing two identical sub-networks where one takes in the pre-disaster image while the other on the post-disaster image. Moreover, we perform a comparative study of results obtained through various ResNet architectures to classify the extent of building damage.

2. DATASET

We are using xBD, which is a large-scale dataset for the advancement of change detection and building damage assessment. The dataset includes post-disaster imagery with transposed polygons from before the disaster over the buildings, as well as damage classification labels. The dataset provides four damage categories - No Damage, Minor Damage, Major Damage, and Destroyed for each building. Over 45,000KM2 of polygon labeled pre and post-disaster imagery is included in the dataset. The xBD dataset is provided in the train, test and holdout splits in a 80/10/10% ratio, respectively.

Table 1. Class-wise representation of damage

Class Name	Description
No damage	Undisturbed. No sign of water, structural or burn damage
Minor damage	Building partially burnt, water surrounding structure, roof elements missing or visible cracks
Major damage	Partial wall or roof collapse, water surrounded
Destroyed	Completely collapsed, completely covered with water/mud or no longer present

Table 2. xBD data splits and annotation counts

Split	Images	Polygon
Train	18,336	632,228
Test	1,866	109,724
Holdout	1,866	108,784

3. RELATED WORK

This section discusses recent methods used for building damage classification using satellite images. There are various

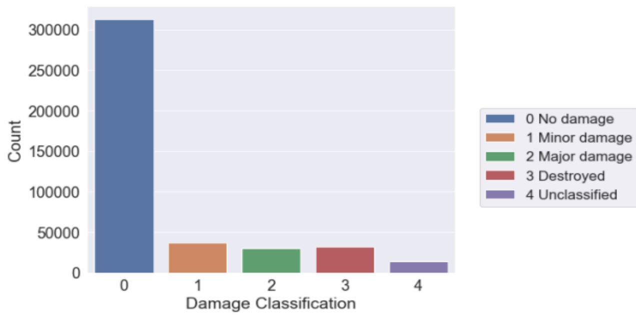

Fig. 1. Distribution of damage class labels

papers that use Siamese networks, CNN, and attention mechanisms to solve the problem of building damage classification. We would like to go through a few of them in detail. The first architecture that we explored was DAHiTrA[4]. This model uses hierarchical transformers to classify building damage based on satellite images in the aftermath of natural disasters. This architecture utilizes U-Net to establish a hierarchical feature-based difference block. After each Downsample block, there is a different block that calculates the difference between two transformer-encoded features.

Also, the recently released BDANet[3] where Self-attention is used to map features from spatial to the feature domain can be efficiently applied through transformers. This paper introduced cross-directional attention in a framework for building segmentation and damage assessment. This architecture also utilizes U-Net which provides a hierarchical structure.

There are numerous papers that utilize the U-Net's[2] capability to extract low-dimension features but very few use transformers. Our research is dependent on whether transformers can be useful for building segmentation instead of using an architecture similar to U-Net. And based on the results of the building segmentation, try doing the damage classification.

Based on a sizable dataset of xBD satellite images, this research builds a unique end-to-end benchmark model known as the pyramid pooling module semi-Siamese network (PPM-SSNet)[9]. By including residual blocks with dilated convolution and squeeze-and-excitation blocks into the network, the suggested model's high accuracy is made possible. By using concurrent learned attention mechanisms via a semi-Siamese network for end-to-end input and output purposes, the highly automated process of satellite imagery input and damage categorization result output is simultaneously achieved.

4. SEGMENTATION

4.1. Pre-processing

We have pre- and post-disaster images for each of the frames in our dataset. We also have bounding boxes for the buildings visible in the pre-disaster photographs in the form of JSON files. We were able to make a Black/White mask with this, which marks buildings with White and the Background with Black. Since we don't have bounding boxes for our post-disaster images, one of our segmentation goals is to develop a similar Black/White mask for them. After the post-disaster mask is constructed, we can extract the corresponding bounding boxes and compare them to the pre-disaster image to determine the degree of damage.

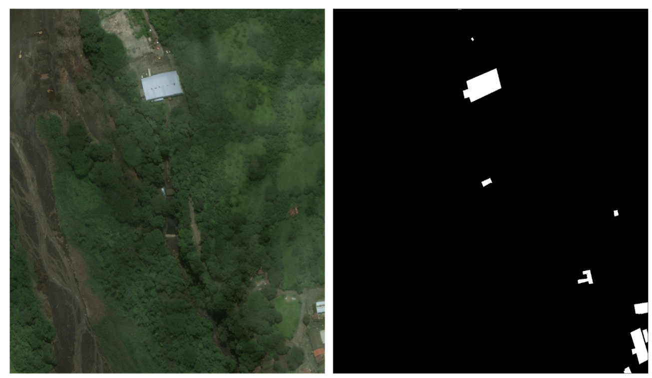

Fig. 2. Raw Satellite Image and B/W masked image

4.2. Semantic Segmentation

Every pixel of the masked image has been colored either Black (background) or White(building). An FC-Densenet[5] network is given both the raw and masked images as input, and when given a raw image, it learns to produce a segmented masked BW image. To locate areas of the post-disaster image that contain buildings, we use this segmented, masked BW image. The degree of damage will then be classified by comparing the post-disaster image areas to the pre-disaster image areas. Another segmentation model we implemented was SegFormer[6]. The model consists of a transformer encoder that is hierarchical and lightweight multi-layer perceptron as decoders to achieve great results on image segmentation. We used a pre-trained model from Hugging Face[7] to train this model. The input to this model is similar to the one above where we pass raw images and pass the black-and-white patch to calculate the loss function. The output of the model is our building segmentation which we, later on, use for building damage classification.

Upon testing these models, our results were not satisfactory to go ahead with the classification. Here are the results of the building segmentation model:

Fig. 3. Test Raw Satellite Image and B/W masked image

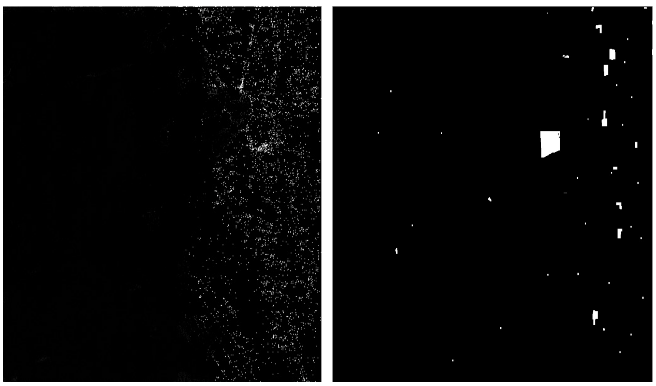

Fig. 4. Results : FC-DenseNet - Generated B/W Mask (Left), SegFormer - Generated B/W Mask (Right)

4.3. Alternative Approach

The transformers and CNN-based architectures that we experimented on were not able to extract the low-level image features accurately. Since our main goal was to classify the images based on the damage, we went ahead with the already provided meta-data to generate masks to extract each building separately into an image and pass pre and post-disaster images together into the Siamese[9] network architecture.

5. MODEL ARCHITECTURE

For this problem, the xBD dataset contains highly imbalanced data with an under-represented target class making the task at hand more difficult. To overcome this, we implemented Siamese Neural Networks (SNN) which is a group of neural network architectures that includes two or more sub-networks that are exactly identical. The term "identical" in this context refers to their setup, which includes the same parameters and weights. In order to compare feature vectors and identify commonalities between inputs, parameter updating is replicated across both sub-networks.

Here, we contrast between three simple yet powerful Deep Learning Architectures that have been used in Image classification previously and have given great results for various applications. The three architectures include the family of ResNets- ResNet18, ResNet34, and ResNet50. Although previous work that has been established in this problem uses advanced, state-of-the-art architectures, we are able to produce better results through these uncomplicated models using our unique approach to segmentation.

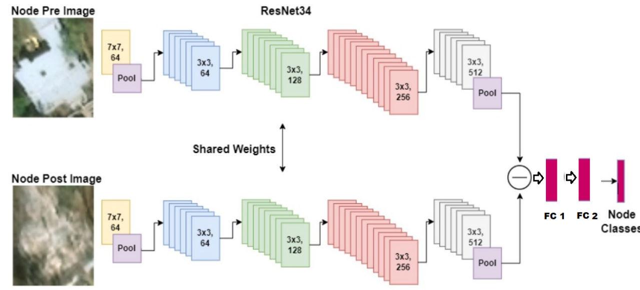

Fig. 5. Siamese CNN with ResNet34

layer name	output size	18-layer	34-layer	50-layer
conv1	112×112	7×7, 64, stride 2		
conv2_x	56×56	3×3 max pool, stride 2		
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax		
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9

Fig. 6. ResNet18, ResNet34, ResNet50 architecture

The batch size was given as 256 images per epoch with 2 hidden layers with 32 hidden units and 0.5 dropout rate after receiving the output from the ResNet architecture. The learning rate for these models is 0.0003 with each model running for 50 epochs. The loss function used here was negative log-likelihood loss with an Adam optimizer.

$$l(\theta) = -\Sigma_{i=1}^{n}\Big(y_i log\bar{y}_{\theta,i} + (1-y_i)log(1-\bar{y}_{\theta,i})\Big) \quad (1)$$

6. MODELING RESULTS

From Table 3, the accuracy for ResNet-18 is 0.76 and the Macro F-1 score is 0.610 which increases gradually as we add more layers. Adding more layers to the network would increase the parameters of the model, as well as the complexity to run them. We are considering the macro F-1 score because we are working with an imbalanced dataset and using the macro F-1 score does not take the number of samples into the consideration.

Table 3. ResNet Architecture Results

	Accuracy		**Macro F-1**	
	Train	**Test**	**Train**	**Test**
ResNet-18	0.766	0.471	0.610	0.354
ResNet-34	0.789	0.510	0.616	0.3706
ResNet-50	0.815	0.528	0.630	0.4009

7. CONCLUSION

After major natural catastrophes, the extent of building damage maps offers vital information for organizing successful rescue and relief operations. Identifying the level of damage may help concentrate resources into regions accordingly. Additionally, the sooner this information is obtained, the more likely it is to save trapped and injured people. Like most geographical phenomena, catastrophe damage spreads such that nearby buildings experience comparable patterns of destruction. In this study, we developed a Siamese network to overcome highly imbalanced data using various versions of ResNet architectures through a unique segmentation approach that is currently missed by existing methods. Other publications are trained on an image containing multiple buildings along with unnecessary structures using state-of-the-art models. Our paper leverages the use of simple yet powerful models by segmenting the buildings as separate images used in training from the pre-disaster and post-disaster building patches rather than the image as a whole. This way our model learns to classify damage only for the building which is our target neglecting the other patches in the image. We also concentrate on the cross-disaster generalization of our model, which would allow our approach to be applied to infer future catastrophes without training, speeding up the process.

8. FUTURE WORK

We have implemented both Segmentation and Classification for assessing damage in calamities. Although segmentation did not provide great results, we have utilized the metadata containing polygon coordinates to extract bounding boxes for conducting segmentation. One possibility in our future work would be to extend the semantic segmentation models to automate the segmentation and classification process. Ideally, we would like to train our model with more satellite data and have requested the dataset author to provide more images. Due to CPU and memory processing bottlenecks, we could not perform training for longer epochs which would be something to work on in the future if we could get access to faster computers. Another future possibility would be to train ResNet-101 and ResNet-152 to see how much better our models can perform. We would also like to perform a richer level of hyperparameter tuning.

In the future, we could build an end-to-end pipeline with an API where users can upload satellite images and we can predict the extent of damage and it would be great to deploy all this on a cloud platform.

Since we were unable to source data for the Ian Hurricane natural disaster, we would like to implement damage assessment on this as well. For any future calamities, we would like to build an automated system to predict the damage of buildings through plug-is form satellite images for faster reaction and response teams.

9. REFERENCES

[1] Ritwik Gupta, Richard Hosfelt, et al. xBD: A Dataset for Assessing Building Damage from Satellite Imagery

[2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation

[3] Yu Shen, Sijie Zhu, Taojiannan Yang, Chen Chen, et al. BDANet: Multiscale Convolutional Neural Network with Cross-directional Attention for Building Damage Assessment from Satellite Images

[4] Navjot Kaura, Cheng-Chun Leeb, Ali Mostafavib, Ali Mahdavi-Amiria, DAHiTrA: Large-scale building damage assessment using a novel hierarchical transformer architecture on satellite images

[5] Simon Jégou, Michal Drozdzal, David Vazquez, Adriana Romero, Yoshua Bengio : FC-DenseNet, The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation

[6] Enze Xie1, Wenhai Wang2, Zhiding Yu, et al. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers

[7] Hugging Face, The AI community building the future., https://huggingface.co/

[8] Gregory Koch, Richard Zemel, Ruslan Salakhutdinov, Siamese Neural Networks for One-shot Image Recognition

[9] Ali Ismail, Mariette Awad : Siamese, Towards Cross-Disaster Building Damage Assessment with Graph Convolutional Networks