2021 IEEE 4th International Conference on Multimedia Information Processing and Retrieval (MIPR) | 978-1-6654-1865-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/MIPR51284.2021.00009

Weakly-Supervised Damaged Building Localization and Assessment with Noise Regularization

Maria Presa-Reyes and Shu-Ching Chen
Knight Foundation School of Computing and Information Sciences
Florida International University, Miami, Florida 33199
Emails: {mpres029,chens}@cs.fiu.edu

***Abstract*—Not only does the destruction caused by natural disasters impair human lives, but it can also result in devastating damages to the community infrastructure and possibly cause the loss of historic structures as well as vital documents. Technological advances in remote sensing survey tools such as satellite images and aerial photographs have allowed emergency responders to rapidly and remotely conduct a comprehensive assessment of the damages caused by a disaster event. Most of the previously proposed research in the automatic identification and prediction of building damage assessments from optical remote sensing data depends on the availability of accurate geometric footprints of the affected area's structures. However, the available building footprints may rapidly become outdated as new infrastructures are built while old ones are demolished or renovated. We propose an end-to-end weakly-supervised damage assessment model where the assumption is that the building footprint is unknown during training.Instead, there is a rough estimate of the building's location and the level of damage it sustained. Ablation tests are conducted on both a large-scale satellite imagery set and a smaller set of aerial photographs prepared and curated by our team to demonstrate our proposed model's performance.**

***Keywords*-damage assessment; deep learning; convolutional neural networks**

I. Introduction

Damage assessment is a preliminary on-site survey of damages or failures caused by an accident or natural occurrences such as a hurricane, tsunami, or earthquake. These damage analyses are often conducted right after a disaster to document the degree of damage that can be replaced, repaired, or recovered. It can also allow estimating the time needed for repair, replacement, and rehabilitation. The first glimpse of the destruction caused by a disaster incident is made available by high-resolution satellite imagery or aerial photographs that allow experts to produce precise estimates of damage to the infrastructure without the need to be physically present on-site.

Technological developments in remote sensing survey instruments [1], [2], [3], such as satellite images and aerial photographs, have enabled emergency responders to perform a comprehensive damage assessment quickly and remotely [4]. After a disaster event, government agencies such as Federal Emergency Management Agency (FEMA) conduct a preliminary building damage assessment through (1) predictive modeling to estimate probable damages; and (2) visual using the imagery captured post-event to assess actual damages [5]. Nevertheless, manually identifying the impacted properties can be a slow and laborious job when disasters cover a wide area of the land.

Deep learning methods demonstrated great success in various research areas [6], [7], [8], [9], [10]. More specifically, the Convolutional Neural Network (CNN) is a well-known architecture that has achieved tremendous breakthroughs in image recognition and has been the preferred method for developing damage assessment models using optical remote sensing data. However, recently proposed approaches that apply deep learning for building damage assessment are limited due to their reliance on the availability of accurate geometries illustrating the footprint of the structure in the map. High-quality pre-disaster images are also expensive to produce and may rapidly become outdated as new structures are developed.

We proposed to apply weak supervision in the detection of the damaged building and the classification of the level of damage. The proposed work uniquely considers a scenario where a rough estimate of the damaged building's location and the degree of damage is the only data available to train the model. Such an approach will be valuable for rapidly identifying damaged buildings of interest and possibly expanding benchmark datasets for further studies without the need to review every structure in the image set. The main contributions of this study are summarized as follows:

- A novel fusion module is trained on the correspondent in-depth features at each spatial location extracted from the images captured before and after a disaster event.
- A weakly-supervised training approach is proposed, using noise regularization to train a robust model.
- Interpretable predictive results that serve as helpful visualizations in the form of heatmaps can be leveraged for rapid labeling jobs and further studies in building damage assessment.

The rest of the paper is organized as follows. Section II reviews the recent related works that apply deep learning methods and other statistical techniques to building localization and damage assessment. Section III introduces our

978-1-6654-1865-2/21/$31.00 ©2021 IEEE
DOI 10.1109/MIPR51284.2021.00009

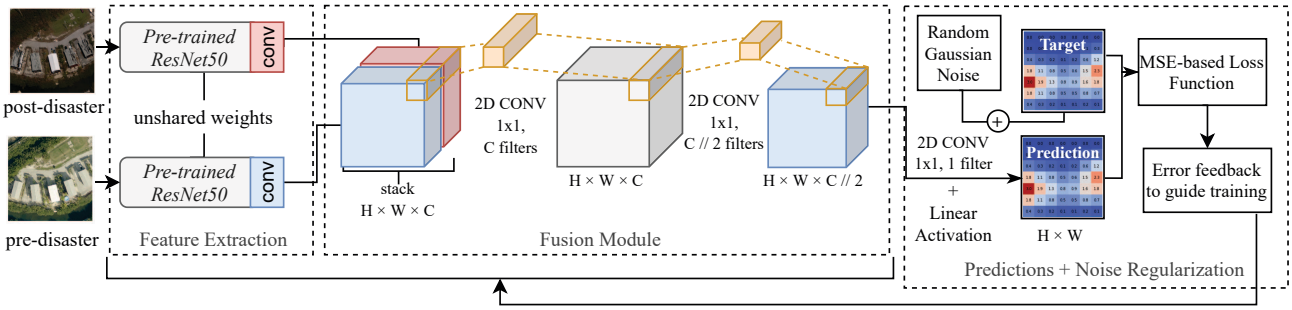

Figure 1: Proposed two-stream CNN architecture for the weakly-supervised damage assessment model applying the proposed fusion module to learn the deep feature correspondence at each feature location of the input image pair. Additive Gaussian noise is randomly applied to the target patches to help the model generalize better.

proposed approach. Section IV provides a brief discussion of the satellite images and aerial photographs in which the proposed approach is tested. Section V summarizes the experimental study with some discussions on the results. And finally, Section VI concludes the paper by providing suggestions on potential future work.

II. Related Work

Previously introduced building damaged scale classification research have been traditionally centered primarily on classifying damaged buildings into two categories (i.e., intact or destroyed) while also focusing on a single damage type due to lack of well-curated benchmark datasets. Nonetheless, more recent research works have explored a larger variety of damage levels and datasets such as xBD [11] have allowed researchers to develop more complex models that can address the challenges unique to this direction.

Currently proposed research works in building damage assessment can be summarized into two categories, instance-level damage classification [12], [13], [14] and pixel-level semantic segmentation [15], [16], [17]. Instance-level classification takes as input the image containing the overhead-view of the building and aims to predict the level of damage the building sustained. Semantic segmentation for pixel-level classification is often the preferred approach which aims to detect not only the building footprint but also predict the degree of damage at each pixel-level. Detecting the damage at each pixel-level has the advantage of obtaining more fine-grained results

Building damage assessment techniques from previous works rely on the availability of high-quality geometry of the building footprint. However, the available building footprints may rapidly become outdated as new infrastructures are built while old ones may get demolished. Pre-disaster images may become outdated as well, implying localization models may not correctly identify the location of the newly developed buildings to make the correct assessment.

III. Approach

An end-to-end convolutional neural network is developed in this work to automatically learn how to extract and fuse the characteristics from the images captured over an affected region before and after a disaster event. The assumption is to train a model under an scenario where data is very limited, and geometric building footprints are scarcely available. Given a pair of images, the proposed model's objective is to generate a two-dimensional predictive patch in a regression-based approach, where each cell from the patch will contain a value of the predicted level of damage—the higher the value, the greater the damage. The proposed approach allows the model to independently learn the specific patterns in the image that belong to a building without the apriori knowledge of the building's footprint.

A. Data Preprocessing

The damaged point values are encoded in an ordinal format starting with 1 as the lowest level of damage—the higher the value, the larger the damage. These values are first multiplied by 100 and placed into a grid on the location that corresponds to the location of the damaged building in the input image and generated the target label patch. A Gaussian kernel smooths out the point values in the grid, giving the damage more ground coverage of the surrounding area. Previously proposed works have demonstrated the importance of the region surrounding the damage structure when making predictions about the level of damage a building has sustained [14]. The Gaussian smoothing process is similar to the average filter, but make use of a special kernel representing a Gaussian curve form.

The formula of the Gaussian smoothing function G is shown in equation 1.

$$G(d_x, d_y, \sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{d_x^2 + d_y^2}{2\sigma^2}} \quad (1)$$

Where d_x is the horizontal axis distance from the origin, d_y is the vertical axis distance from the origin, and σ is the Gaussian distribution's standard deviation.

During training, additive zero-centered Gaussian noise is randomly applied to the target label patch in order to create synthetic perturbations in the data and mitigate over-fitting. These small perturbations are meant to aid in training a model that is robust to the noise often found in real-life data [18] such as Global Positioning System (GPS) location errors.

B. Framework Configuration

1) Feature Extraction: As shown in Figure 1, our proposed end-to-end framework is configured as a two-stream CNN network that implements a fusion module to learn from the correspondence between each image's feature vector. Each stream makes use of the ResNet50 architecture [19], pre-trained on ImageNet, to extract features from the last convolutional layer and obtain a feature vector that is correspondent to parts of the 2-D image. Both networks' pre-trained weights are entirely fine-tuned to the new damage-assessment datasets. Moreover, the weights between the networks are unshared, as this has been demonstrated by previous work [12] to allow each stream the flexibility to individually fine-tune their weights accordingly and achieve better results.

2) Fusion Module: The fusion module following the feature extraction step is trained to learn the deep correspondence between both feature vectors using the capabilities provided by the one-by-one (1×1) 2-D convolution technique [20]. The two-stream networks in the feature extraction phase take as input an image pair and from its last convolutional layer generates the feature vectors $F^{pre}, F^{post} \in \mathbb{R}^{H \times W \times D}$, where H, W, D are the height, width and total number of channels from the corresponding feature vector. The objective is to develop a fusion module, or function, such that $f : F^{pre}\ F^{post} \rightarrow P$, where $P \in \mathbb{R}^{H \times W}$ is the predictive output patch. The fusion model requires the vectors F^{pre} and F^{post} to first be stacked at the same spatial location (x, y) across channel d, namely:

$$F^{\text{stack}}_{x,y,2d} = F^{pre}_{x,y,d}, \qquad F^{\text{stack}}_{x,y,2d-1} = F^{post}_{x,y,d}, \tag{2}$$

where $\mathbf{F^{stack}} \in \mathbb{R}^{H \times W \times C}$ and $C = 2D$. It is also assumed $1 \leq x \leq H$, $1 \leq y \leq W$, $1 \leq d \leq D$. As shown in Figure 1, the 1×1 convolution that follows works as a coordinate-dependent transformation implemented to takes as input the stacked feature vectors and define the correspondence at each spatial location (x, y). This convolution approach leads to dimensionality reduction with its combination being mathematically equivalent to a multi-layer perceptron at each (x, y) [20].

While the first 1×1 convolution layer takes care of learning the deep correspondence between F^{pre} and F^{post}, the 1×1 convolutional layer that follows further refines the fused features and reduces the dimensions, while the last convolutional layer applies a single 1×1 convolutional layer followed by a Linear activation generates the final predictive patch P.

3) Predictive Results Post-processing: When running inference, the proposed model makes overlapping strides throughout the test dataset and outputs the predictive patches for each stride. These predictive patches are then merged together, with the maximum value of the overlapping cells calculated to generate the final predictive heatmap where the higher the values of the output grid cell, the higher the damage that was sustained by the building located in that specific area.

IV. Data

1) Satellite Imagery: As a benchmark to test our techniques, xBD [11] is a newly introduced large-scale dataset built for the advancement of building detection and damage assessment across various levels of damage and types of damage. xBD provides multi-band satellite pre- and post-event imagery with building polygons, damage type classification labels, ordinal damage level labels, and corresponding satellite metadata from a variety of disaster events. The dataset includes about 700,000 building annotations from 15 countries across more than $5,000\text{km}^2$ of imagery.

There is a broad variety of disaster events that can be found in the xBD dataset, including hurricanes, earthquakes, fires, etc. Within the xBD dataset, hurricane and flooding events are well represented but also indicate intra-class differences that need to be considered. For example, a large wind damage that is mirrored in their rooftop could have been sustained by other buildings. On the other hand, their rooftop may have left other buildings that experienced a significant flooding effect intact. Therefore, the water covering the building is a crucial clue to assessing this form of damage when identifying damage caused by floods. In addition, building damage, depending on the form and size of the building, is often visually varied.

The xBD data is delivered in the form of image tiles of size 1024×1024 that is padded with empty pixels bringing the size of the patch up to 1120×1120 in order to allow the corner buildings to be closer to the center of the cropped image patch in some cases while training the model and improving its predictive capabilities. While training the model, random 224×224 crops are made to these tiles to feed into the model. The centroids of the building's polygons are computed as the damage assessment points. Table I summarizes the number of damage point instances available from both datasets.

2) Aerial Photographs: The Irma dataset of aerial photographs represent one disaster event with a focus on the wind damage. This dataset was processed and curated by our team and is composed of aerial imagery and damage assessment labels from open sources with a focus on the damages caused by Hurricane Irma in 2017 at the Florida Keys [21]. The aerial imagery was collected in the affected

Table I: Data summary of the Irma data and xBD data

No.	Concepts	Irma	xBD
		# of damage instances	
1	Affected	1219	-
2	Minor Damage	2739	36860
3	Major Damage	1082	29904
4	Destroyed	577	31560

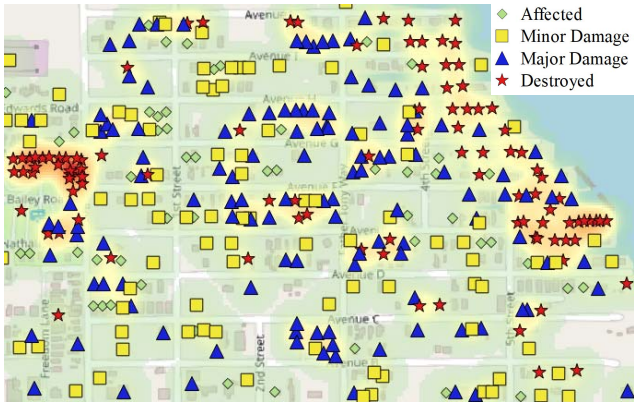

Figure 2: Irma damaged building location and damage level visualization from the Big Pine Key south area, one of the most affected regions after Hurricane Irma.

areas identified by FEMA and the National Weather Service during Hurricane Irma, where the Monroe county area located in South Florida received most of the substantial damages [22]. Hurricane Irma struck the Florida Keys as a category four storm with the maximum sustained wind speed of 132 `mph` and storm surge reaching up to 8 `feet` [21]. The eye of the storm made landfall over Cudjoe Key, and consequently, the Lower and Middle Keys received the highest impact.

The Irma data is composed of high resolution aerial photographs representing a Ground Sampling Distance (GSD) of 50 cm. Along-side the aerial photographs taken before and after the disaster event, the damage assessment labels were also gathered by our team from public sources, as well as combined and curated. Various damage assessment reports were combined from different data sources which include Monroe County's official public preliminary damage assessment report [23], Xian *et al.*'s assessment of the damages of more than 1600 residential buildings [24], and FEMA's Historical Damage Assessment Database [25].

Different from xBD, the Irma data includes the affected damage level, along with minor damage, major damage, and destroyed. As shown in Figure 2, damage labels from the Irma data are only the estimates of the damaged building's location and do not include the geometric information of the building footprint. The damage labels can be matched with building footprint geometries gathered from other data sources. Nonetheless, there is also the challenge of matching the location of the points to the right footprint.

The different damage levels are summarized as follows in accordance to FEMA's official guide in damage assessment [5]:

- *No damage*: the structure remains unchanged as seen from the birds' eye-view.
- *Affected*: the structure exhibits minimal effects, such as some missing shingles in the building rooftop but it is still habitable according to FEMA's standards.
- *Minor damage*: it constitutes of damages that do not necessarily affect the integrity of the structure but may make it inhabitable until repairs are done.
- *Major damage*: the structure sustained substantial damage and requires extensive repairs in order to make it habitable
- *Destroyed*: the structure is a total loss to the point where repair will not be feasible for recovery.

V. Experiments and Analyses

A. Experimental Setup

In this paper, two datasets are tested on the proposed methods. The Irma data samples are split into three non-overlapping areas for training, validation, and testing. Two of the lower and middle keys, Sugarloaf Key and Cudjoe Key, are selected as validation and testing set areas. The xBD satellite images are also split in a similar manner following the original xBD data split provided in train, test, and holdout splits in the 80/10/10% split ratio. In the proposed approach, the test set is used as a validation set to make an unbiased evaluation of our model while training and storing the best performing results. The holdout set is used to test the final already trained model.

Using the Mean Squared Error (MSE) formula, two loss functions are defined and assessed for the proposed approach—the patch loss $\mathcal{L}_{patch}$ function and the pixel loss $\mathcal{L}_{pixel}$ function. As demonstrated in equation 3, $\mathcal{L}_{patch}$ first calculates the losses for each grid cell (i, j) at each patch level P_m, where $(i, j) \in P_m \in \mathbb{R}^{N\times 2}$ and N is the total number of cells in the grid. These losses are then summed-up across a batch and divided by the total number patches M in the batch to obtain the average loss from each patch-level.

$$\mathcal{L}_{patch} = \frac{1}{M}\sum_{m=0}^{M}[\frac{\sum_{i,j\in P_m}(Y(i,j) - \hat{Y}(i,j))^2}{N}] \quad (3)$$

As an alternative, the pixel loss $\mathcal{L}_{pixel}$ function demonstrated in equation 4 calculates losses for the individual cells in the grid, treating each predicted cell as its own individual sample. The variable K represents total number of individual cells in each training batch, and is also assumed to be equal to $M \times N$.

$$\mathcal{L}_{pixel} = \frac{1}{K}\sum_{k=0}^{K}(Y_k - \hat{Y}_k)^2 \quad (4)$$

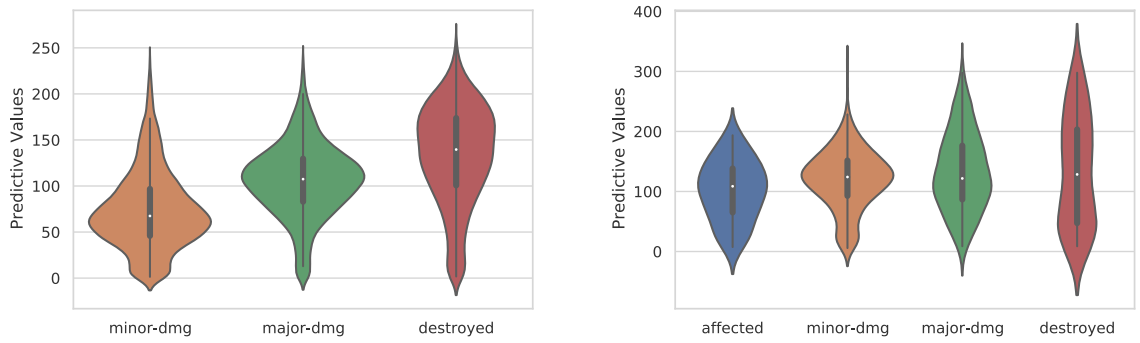

Figure 3: Violin plots of the sampled damaged points from predictions made on the xBD (left) and Irma (right) test split, grouped by the different levels of damage.

The performance improvements made by the proposed fusion module are also corroborated by comparing the results from the post-only model and pre-post fusion model configurations detailed as follows:

Post-only configuration: A single CNN model, based on the ResNet50 architecture and pre-trained on ImageNet is entirely fine-tuned to extract the features from the images taken after the disaster event. The original classification head is removed and replaced with a 1×1 2-D convolutional layer followed by a linear activation function to output the predictive patch.

Pre-post fusion configuration: This configuration makes use of the proposed fusion module highlighted in Figure 1, more details about this configuration have be found in Section III-B.

Each epoch, the model is trained on the entire training set using a small batch size of 32 samples to further regularize and improve the model's generalization capability [26]. As the model trains, it is evaluated at the end of each epoch on the validation set with the best performing model that has the lowest validation loss saved—models are trained for no longer than 100 epochs. In these experiments, the model's weights are optimized using the Stochastic Gradient Descent with an initial learning rate of $\eta = 0.001$. During training, the learning rate is multiplied by factor of 0.1 after there have been no improvements to the validation loss for 10 consecutive epochs.

B. Results and Discussion

For testing purposes, values for damaged location points in the testing set are sampled using bilinear interpolation [27] from the final predictive heatmaps' output predictive value cells to compare with the labeled assessment data. In bilinear interpolation, given the sampling point's location, the four cell centers from the input predictive patch that are nearest to the sampled cell's center are weighted based on the distances and then averaged. Table II and Table III

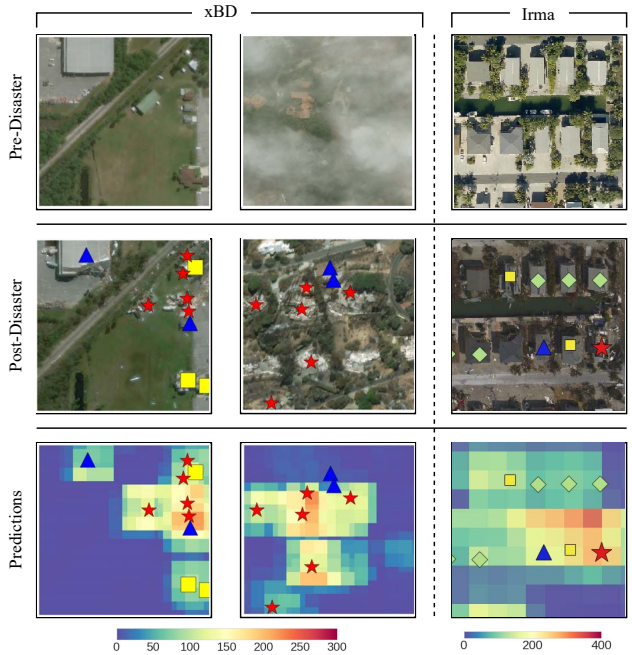

Figure 4: Qualitative results summary of the model's output predictive values on the test set for the xBD and the Irma data. The point locations for the damage buildings are overlaid on the post-disaster images and the model's predictive patch. The legend for the damage point labels is as follows: ◊ - affected, □ - minor-damaged, △ - major-damaged, and ⋆ - destroyed.

summarize the performance results for the xBD data and the Irma data test splits using well-known regression metrics MSE and Mean Absolute Error (MAE). These metrics are utilized to compare the performance among the model configurations described in previous sections. The closer the predictions made by the model are to the target damage

instances, the more confidence can be place into the model to more accurately detect the different damaged buildings.

The pre-post fusion model has consistently achieved the best performance. Moreover, losses calculated from each individual pixel-level may perform better on smaller datasets due to MSE's sensitivity to outliers, which is essential when detecting the minority classes. Noise regularization plays an essential role in augmenting the training data and helping to train a more robust model. As part of the ablation study, the proposed pre-post fusion model is tested with and without noise regularization, demonstrating the performance improvements that can be achieved by applying it. Although the proposed noise regularization hinders the performance of the smaller and more limited Irma data, it has proven to be an essential technique when working with larger datasets such as xBD.

Furthermore, the proposed patch-level loss function has also helped the model achieve better performance and reduce errors. Unlike the pixel-level loss function, the patch-level loss is not as sensitive to noise and places more weights on the surrounding regions.

Table II: Performance summary on the xBD data

Method	Loss Function	Noise Reg.	MSE	MAE
post-only	$\mathcal{L}_{pixel}$	✓	1.2259	0.9951
pre-post fusion			1.1867	0.9900
		✓	1.1285	0.9344
post-only	$\mathcal{L}_{patch}$	✓	1.2111	0.9828
pre-post fusion			1.1437	0.9436
		✓	**1.1282**	**0.9266**

Table III: Performance summary on the Irma data

Method	Loss Function	Noise Reg.	MSE	MAE
post-only	$\mathcal{L}_{pixel}$	✓	1.7227	1.4860
pre-post fusion			1.4714	1.2342
		✓	1.5663	1.3203
post-only	$\mathcal{L}_{patch}$	✓	1.5356	1.2722
pre-post fusion			**1.3893**	**1.1235**
		✓	1.5282	1.3039

Figure 3 demonstrates the violin plots of the best performing model configurations for the xBD and Irma data—its broader segments represent members of the population that are more inclined to take on the given value; the skinnier sections express a lower likelihood. It can be observed that the weakly-supervised model can find an evident pattern among different damage-levels—this is especially true with the results of the xBD data. Even if there is an offset between the target values and the predicted values, it is clear that the majority of the sampled predicted points are grouped on a segment following an ordinal paradigm—the higher the value from the predictive cell, the higher the level of damage sustained by the building located in that cell.

The model has successfully, and through implicit means, learned to detect different damage levels sustained by buildings. Figure 4 illustrates the qualitative results generated by the proposed model on the xBD and Irma datasets. These predictive outputs serve as the interpretable visualizations that can be leveraged for rapid labeling jobs and further studies in the weakly-supervised building damage assessment effort. Further refinement must be done to the predictive outputs to train a model that can further separate the individual buildings and detect the damages they sustained.

VI. Conclusion and Future Work

This paper describes a novel approach to identifying damaged buildings from remote sensing images and predicting the level of damage that the building sustained after a disaster event. The proposed model removes the dependency on the available high-quality building footprint geometries by assuming that only an estimate of the damaged building's location and damage level is available to train the model. Weak supervision, together with the inclusion of perturbations and noise regularization, are critical elements to develop a robust model to label noise, which can adapt better across multiple domains and various types of areas. The proposed model's predictive results can be further used to visualize, identify quickly, and flag damaged buildings, and conduct further studies and research. As part of the future work, we will continue to improve the model's performance and develop techniques to better handle the class imbalance in the data by assigning higher weights to the samples from minority classes during training.

Acknowledgment

For Shu-Ching Chen, this research is partially supported by NSF CNS-1952089.

References

[1] K. Zhang, J. Yan, and S.-C. Chen, "Automatic construction of building footprints from airborne lidar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2523–2533, 2006.

[2] J. Yan, K. Zhang, C. Zhang, S.-C. Chen, and G. Narasimhan, "Automatic construction of 3-D building model from airborne lidar data through 2-d snake algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 3–14, 2014.

[3] K. Zhang, S.-C. Chen, D. Whitman, M.-L. Shyu, J. Yan, and C. Zhang, "A progressive morphological filter for removing nonground measurements from airborne lidar data," *IEEE transactions on geoscience and remote sensing*, vol. 41, no. 4, pp. 872–882, 2003.

[4] L. Zheng, C. Shen, L. Tang, T. Li, S. Luis, S.-C. Chen, and V. Hristidis, "Using data mining techniques to address critical information exchange needs in disaster affected public-private networks," in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2010, pp. 125–134.

[5] Federal Emergency Management Agency (FEMA), "Damage assessment operations manual," 2016. [Online]. Available: https://www.fema.gov/media-library-data/1558541566358-30e29cac50605aae39af77f7e25a3ff0/Damage_Assessment_Manual_4-5-2016.pdf

[6] M. E. P. Reyes, S. Pouyanfar, H. C. Zheng, H.-Y. Ha, and S.-C. Chen, "Multimedia data management for disaster situation awareness," in *International Symposium on Sensor Networks, Systems and Security*. Springer, 2017, pp. 137–146.

[7] Y. Yan, M. Chen, M.-L. Shyu, and S.-C. Chen, "Deep learning for imbalanced multimedia data classification," in *2015 IEEE international symposium on multimedia (ISM)*. IEEE, 2015, pp. 483–488.

[8] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M.-L. Shyu, S.-C. Chen, and S. Iyengar, "A survey on deep learning: Algorithms, techniques, and applications," *ACM Computing Surveys*, vol. 51, no. 5, pp. 1–36, 2018.

[9] S. Pouyanfar, Y. Yang, S.-C. Chen, M.-L. Shyu, and S. Iyengar, "Multimedia big data analytics: A survey," *ACM computing surveys (CSUR)*, vol. 51, no. 1, pp. 1–34, 2018.

[10] S. Pouyanfar, Y. Tao, A. Mohan, H. Tian, A. S. Kaseb, K. Gauen, R. Dailey, S. Aghajanzadeh, Y.-H. Lu, S.-C. Chen *et al.*, "Dynamic sampling in convolutional neural networks for imbalanced data classification," in *2018 IEEE conference on multimedia information processing and retrieval (MIPR)*. IEEE, 2018, pp. 112–117.

[11] R. Gupta, B. Goodman, N. Patel, R. Hosfelt, S. Sajeev, E. Heim, J. Doshi, K. Lucas, H. Choset, and M. Gaston, "Creating xBD: A dataset for assessing building damage from satellite imagery," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 10–17.

[12] A. Fujita, K. Sakurada, T. Imaizumi, R. Ito, S. Hikosaka, and R. Nakamura, "Damage detection from aerial images via convolutional neural networks," in *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*. IEEE, 2017, pp. 5–8.

[13] T. Ci, Z. Liu, and Y. Wang, "Assessment of the degree of building damage caused by disaster using convolutional neural networks in combination with ordinal regression," *Remote Sensing*, vol. 11, no. 23, p. 2858, 2019.

[14] M. Presa-Reyes and S.-C. Chen, "Assessing building damage by learning the deep feature correspondence of before and after aerial images," in *IEEE Conference on Multimedia Information Processing and Retrieval*. IEEE, 2020, pp. 43–48.

[15] R. Gupta and M. Shah, "RescueNet: Joint building segmentation and damage assessment from satellite imagery," *arXiv preprint arXiv:2004.07312*, 2020.

[16] H. Hao, S. Baireddy, E. R. Bartusiak, L. Konz, K. LaTourette, M. Gribbons, M. Chan, M. L. Comer, and E. J. Delp, "An attention-based system for damage assessment using satellite imagery," *arXiv preprint arXiv:2004.06643*, 2020.

[17] Y. Shen, S. Zhu, T. Yang, and C. Chen, "Cross-directional feature fusion network for building damage assessment from satellite imagery," *arXiv preprint arXiv:2010.14014*, 2020.

[18] T. Poggio, V. Torre, and C. Koch, "Computational vision and regularization theory," *Readings in computer vision*, pp. 638–643, 1987.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[20] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv:1312.4400*, 2013.

[21] H. Wetherington, "Hurricane Irma recovery." [Online]. Available: https://www.monroecounty-fl.gov/726/Hurricane-Irma-Recovery

[22] National Oceanic and Atmospheric Administration (NOAA), "Hurricane IRMA: Emergency response imagery of the surrounding regions." [Online]. Available: https://storms.ngs.noaa.gov/storms/irma/download/metadata.html

[23] Monroe County - GIS, "Hurricane Irma - damage assessment." [Online]. Available: http://monroecounty-fl.maps.arcgis.com/apps/webappviewer/index.html?id=87fc14264bdf43fba1669413381dfe3a

[24] S. Xian, K. Feng, N. Lin, R. Marsooli, D. Chavas, J. Chen, and A. Hatzikyriakou, "Rapid assessment of damaged homes in the Florida Keys after Hurricane Irma," *arXiv preprint arXiv:1801.06596*, 2018.

[25] Federal Emergency Management Agency (FEMA), "Geoplatform.gov," 2020.

[26] D. Masters and C. Luschi, "Revisiting small batch training for deep neural networks," *arXiv preprint arXiv:1804.07612*, 2018.

[27] E. J. Kirkland, "Bilinear interpolation," in *Advanced Computing in Electron Microscopy*. Springer, 2010, pp. 261–263.