Large-Scale Damage Detection Using Satellite Imagery

Lionel Gueguen[1], Raffay Hamid[1]
[1]DigitalGlobe Inc., 12076 Grant Street, Thornton, Colorado, USA.

Each year, hundreds of catastrophic events impact vulnerable areas around the world. Assessing the extent of damage caused by these crises is crucial in the timely allocation of resources to help the affected populations. Since disaster-locations are usually not readily accessible, the use of satellite imagery has emerged as a valuable source of information for estimating the impact of catastrophic events.

However, currently these assessments are mostly done by analyzing the pre- and post-event images of distressed areas by human photo-interpreters, making it a labor-intensive and expensive process. It is therefore important to scale-up damage detection to larger areas accurately and efficiently. Our work is a step towards solving this problem.

In the following, we summarize some of the key challenges that need to be addressed in this regard, and how our work contributes towards them:

1- Comprehensive Data-Set: Thus far, there has been a lack of comprehensive labeled data-set that could be used to explore automatic damage detection at scale. To this end, we present a benchmark data-set of 86 pairs of pre- and post-event satellite imagery of distressed areas covering 4,665 KM^2 with the associated ground truth of damaged regions acquired by expert interpreters. This data-set was collected by using the satellites of DigitalGlobe Inc. Our data-set covers 12 different regions from around the world, and spans a wide range of terrains and climates, with a variety of damage types (see Figure 1). To the best of our knowledge, the size and variability of our data makes our work the most thorough analysis of automatic damage detection ever published thus far.

2- Appropriate Feature Choice: The scale of our problem naturally presents an accuracy-efficiency tradeoff for the features being considered. To this end, we introduce the use of trees-of-shapes features [3] in the bag-of-visual-words model [1] that focuses more on the shape characteristics of a scene, as opposed to its edge attributes as done by other popular descriptors *e.g.*, SIFT [2] (see Figure 2 for tree-of-shapes illustration). Our results show that this difference proves to be quite important to detect damaged areas accurately. We present a thorough empirical analysis for the effectiveness of our scheme, and compare it to multiple alternatives. Figure 3 shows the ROC and EER obtained using multiple feature-sets used in different learning settings. Figure 4) illustrates some example damages our framework was able to automatically detect.

3- Label Acquisition Cost: Given the high skill-set required from the photo-interpreters to assess the damage accurately, acquiring reliable ground-truth labels is particularly challenging for our problem. This high label acquisition cost makes it important to explore the various learning paradigms that could utilize the labeled data effectively. To this end, we present a thorough comparison of different learning strategies, including supervised, unsupervised and semi-supervised methods. Our results suggest the use of semi-supervised learning as a good trade-off between the label-acquisition cost and the detection accuracy. We present a user-study of the photo-interpretation efficiency provided by our framework, and report a ten-fold speed-up compared to an exhaustive manual inspection, at a minimal loss in detection accuracy.

[1] Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *IEEE CVPR*, volume 2, pages 2169–2178, 2006.

[2] David G Lowe. Object recognition from local scale-invariant features. In *IEEE ICCV*, 1999.

[3] G.-S. Xia, J. Delon, and Y. Gousseau. Shape-based invariant texture indexing. *IJCV*, 88(3):382–403, 2010.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

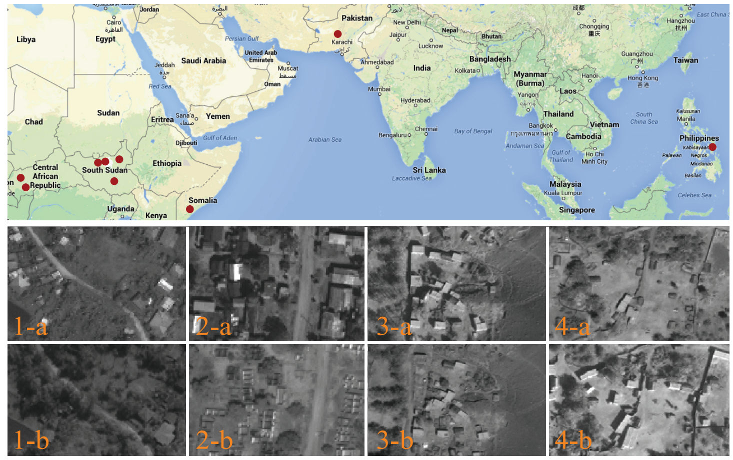


Figure 1: (Top) Example areas of interest shown with red dots. (Bottom) Different instances of types of changes considered in our data-set.

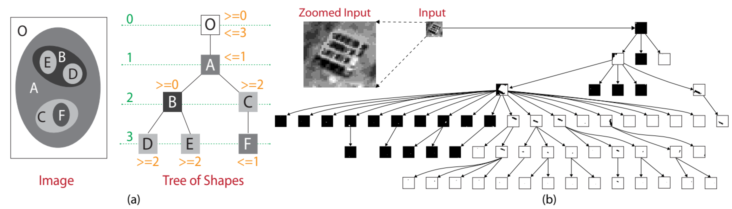
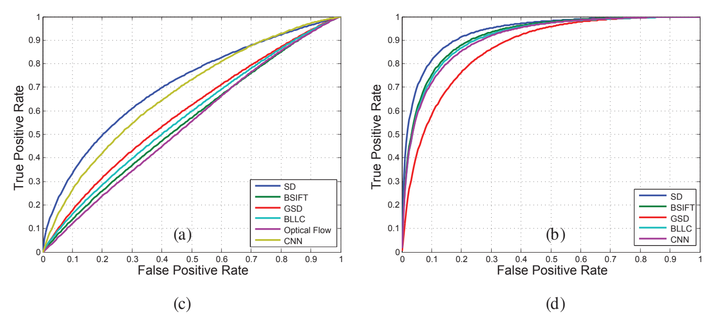
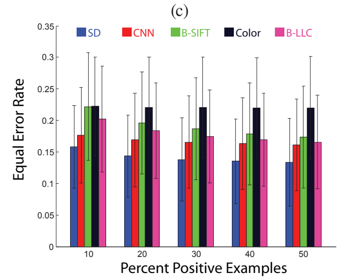


Figure 2: (**a**)- Illustration of tree of shapes. (**b**)- Example image of a roofless building decomposed into its tree of shapes.





Equal Error Rate (EER)

Feature Type	Unsupervised	Supervised
SD	0.339±0.088	0.123±0.068
CNN	0.360±0.120	0.161±0.072
B-SIFT	0.462±0.049	0.149±0.071
B-LLC	0.447±0.057	0.156±0.074
GSD	0.429±0.072	0.205±0.077
O-Flow	0.473±0.029	-

Figure 3: Average ROCs for multiple feature-sets used in (a) unsupervised, and (b) supervised setting. EER as (c) a function of training size in supervised setting, and (d) for fixed training size in different learning settings.

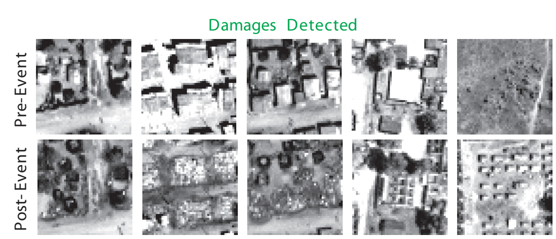


Figure 4: Example damaged areas detected by our framework.