PIXEL-WISE T-TEST: A NEW ALGORITHM FOR PERSISTENT BUILDING DAMAGE DETECTION IN SYNTHETIC APERTURE RADAR IMAGERY

Anonymous authors
Paper under double-blind review

ABSTRACT

This paper elaborates a new change detection algorithm, the Pixel-Wise T-Test (PWTT), developed to leverage the particular characteristics of war-induced building damage. The algorithm is deployed and tested in several cities damaged during the war in Ukraine. Despite being simple and lightweight, the algorithm produces results with accuracy statistics rival State of the Art methods that use deep learning and expensive high resolution imagery. Furthermore, the workflow is deployed entirely within the Google Earth Engine environment, allowing for the generation of near-real time damage maps that allow humanitarian practitioners to immediately get the count of damaged buildings in a user-specific area of interest.

1 INTRODUCTION

Satellite imagery is used extensively in the field of disaster response. The highest levels of accuracy are generally achieved using Convolutional Neural Networks (CNN) on very-high resolution (VHR) optical satellite imagery. However, such approaches are both financially and computationally expensive, and optical imagery is not consistently available due to cloud cover. The use of open-access Synthetic Aperture Radar (SAR) imagery mitigates some of these problems, and has been fruitfully applied to the detection of damage from natural disasters, but not war. This paper develops a new methodology that seeks to leverage the advantages of both approaches for the detection of building damage.

In an analysis of war-related destruction resulting from the Syrian war, Mueller et al. (2020) achieve an area under the curve of 0.92 by training a CNN on damage annotations carried out but the United Nations and high resolution satellite imagery. An important innovation of their research involves exploiting temporal persistence of war-related damage. Damage resulting from natural disasters results in one time period and reconstruction efforts are initiated as soon as possible. In contrast, conflict-induced damage occurs additively over time and reconstruction efforts are often delayed for months or even years by fighting. Leveraging the fact that damaged buildings in war zones tend to stay damaged, the addition of temporal smoothing more than doubles model precision for certain cases.

Despite the high accuracy achieved in Syria, there are a number of limitations to the application of this approach elsewhere. High resolution optical imagery from providers such as Maxar or Planet is expensive; monitoring large areas can become very costly, particularly if repeated observations are required. Clouds pose an additional challenge for optical imagery; while damage assessment for conflicts in the Middle East is not significantly hampered by cloud cover, certain areas in Ukraine are often shrouded in clouds. Of the 262 images of Mariupol taken by the PlanetScope constellation in 2022, only 52 are cloud-free. This not only entails a reduction in the quantity of data available for use, but may adversely affect the performance of a model that relies on consistent temporal lags since cloud cover varies seasonally.

The use of open-access Synthetic Aperture Radar imagery solves many of the problems associated with expense and coverage consistency. Considerable progress has been made in the application of interferometric coherence to identify buildings damaged by a single shock such as an earthquake (Kumar et al., 2022; Tiongson & Ramirez, 2022). However, because interferometry relies on phase information, it requires Single-Look Complex (SLC) data and computationally expensive workflows.

Several approaches to general change detection in SAR imagery have been elaborated in recent years, most notably Canty et al. (2019), who build upon the sequential complex Wishart-based algorithm developed by Conradsen et al. (2003) that results in the generation of a test statistic for each pixel. This algorithm is particularly well suited to the monitoring of areas with high levels of anthropogenic activity, such as mines, agricultural fields, ports, and airports. A key strength of this approach is that only relies on the amplitude of the backscatter included in the Sentinel-1 Ground-Range Detected (GRD) products. This enables the use of Google Earth Engine, allowing for planetary scale analysis in near real time. However, a stated limitation of this model is that it is purely data-driven and unsupervised, precluding the use of temporal information that leverages the persistence of war-induced damage.

The nature of war-related damage is that a feature such as an apartment complex will remain static, experience a significant one-time change (destruction), and then remain static again in its altered state. This differs in important ways from scenarios well suited for interferometry (a one-time change on a precise date) and Canty's sequential complex Wishart-based algorithm (frequent change).

2 METHODOLOGY

The approach taken in this paper fills the gaps outlined in the literature above by conducting a Pixel-Wise T-Test (PWTT) on Synthetic Aperture Radar imagery. The T-Test is one of the most common forms of hypothesis testing, and is widely used in clinical studies to determine whether a treatment has led to a significant change in the mean value of an outcome variable. The temporal persistence of building damage can be thought of in similar terms; using the onset of a conflict (or a destructive event) to separate the sample into a pre- and post- shock period, we would expect a pixel displaying damage to have a different mean value across a long post-shock period. We would also expect relatively little variance in pixel values in both the pre- and post- shock periods for damaged areas. The integration of pixel variance into the test also allows for the isolation of damage, rather than simple change as measured by algorithms such as Canty et al. (2019); airports and train stations are likely to have high pixel variance, meaning that these areas are unlikely to achieve high T-values.

2.1 PRE-PROCESSING

Sentinel-1 scenes acquired in Interferometric Wide-Swath mode is accessed in Google Earth Engine, spanning one full year before and after the onset of the war in Ukraine (24/02/2022). To ensure consistency in the look angle, Sentinel-1 imagery is first disaggregated by orbital pass and relative orbit number. Scenes are further separated by polarization, Vertical-Vertical (VV) and Vertical-Horizontal (VH). Finally, samples are split into pre- and and post war periods.

2.2 THE PIXEL-WISE T-TEST

Following pre-processing, a standard T-Test is carried out:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where x_1 is the pre-war sample and x_2 is the post-war sample. The final damage probability raster uses the absolute values of the T-statistic for each pixel. This is because the directionality of the change in backscatter resulting from the destruction of a building is unknown a-priori; depending on the angle of a building's orientation relative to the satellite, the transition from a standing building to rubble could either lead to an increase or decrease in backscatter. A gaussian filter with a radius of 20 meters is used to smooth the result.

After the generation of the damage probability raster, building-level inference is carried out using Microsoft Buildings, an open-access global building footprint dataset. The mean pixel value of the damage probability raster is calculated within each building footprint, and added as a feature. A cutoff value of the test statistic is chosen based on the degrees of freedom (the total number of images used minus 2), and according to the desired statistical significance threshold (e.g. T¿1.96 for large samples at the 95% confidence level).

3 RESULTS

The results of this analysis are visualized in the following interactive application. The tool updates in near-real time, ingesting new imagery as it becomes available. It displays the damage probability raster, and allows for the generation of an estimate of damaged buildings in a user-specified area of interest.

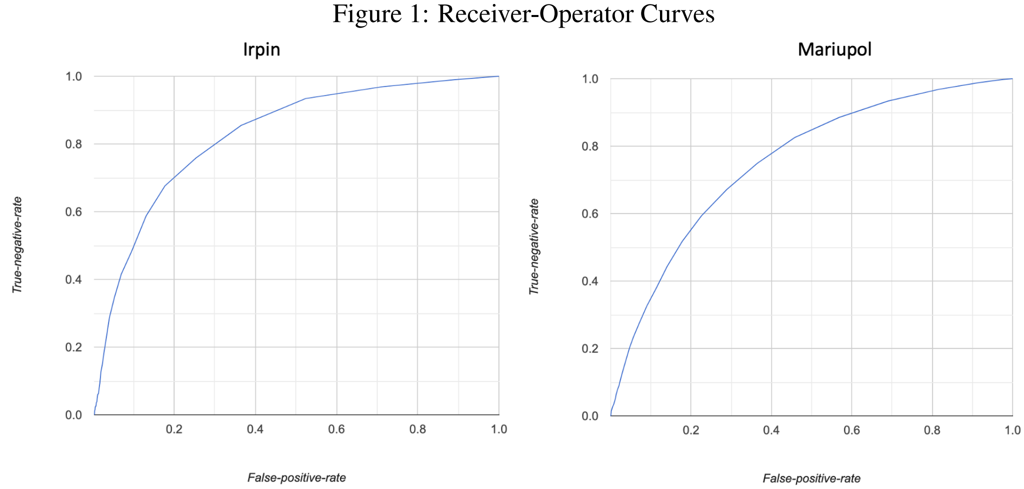
A dropdown menu in the tool allows users to conduct accuracy assessment using labeled data for Mariupol and Irpin. To assess the accuracy of this workflow, manually annotated building damage labels from UNOSAT are used. These data come in the form of the latitude and longitude of locations in which damage was identified on the basis of high resolution satellite imagery. Because these data only indicate the location of damaged points, undamaged areas must be quantified. Mueller et al. (2020) overcome this problem by drawing a 32-by-32 meter grid over the area and counting the number of damaged points within each grid cell. A latent variable that may bias accuracy statistics generated using this approach simply involves the number (or even the presence) of buildings in a grid cell.

Instead, this paper conducts accuracy assessment at the building level by spatially joining UNOSAT damage annotation points with building footprints. A building footprint is labeled as damaged if a damage annotation point falls within it, and labeled undamaged otherwise. This allows for the generation of more meaningful accuracy statistics, and a more stringent test of model accuracy than previous studies.

3.1 MARIUPOL AND IRPIN

The cities of Mariupol and Irpin suffered some of the most extensive damage from the war. Both endured heavy shelling in the first month of the conflict. The PWTT algorithm is run on both cities using a cutoff date of March 1st, 2022, yielding two datasets of building footprints with associated damage probability values. Receiver-Operator Curves for Mariupol and Irpin yield an area under the curve of 0.75 and 0.82, respectively:

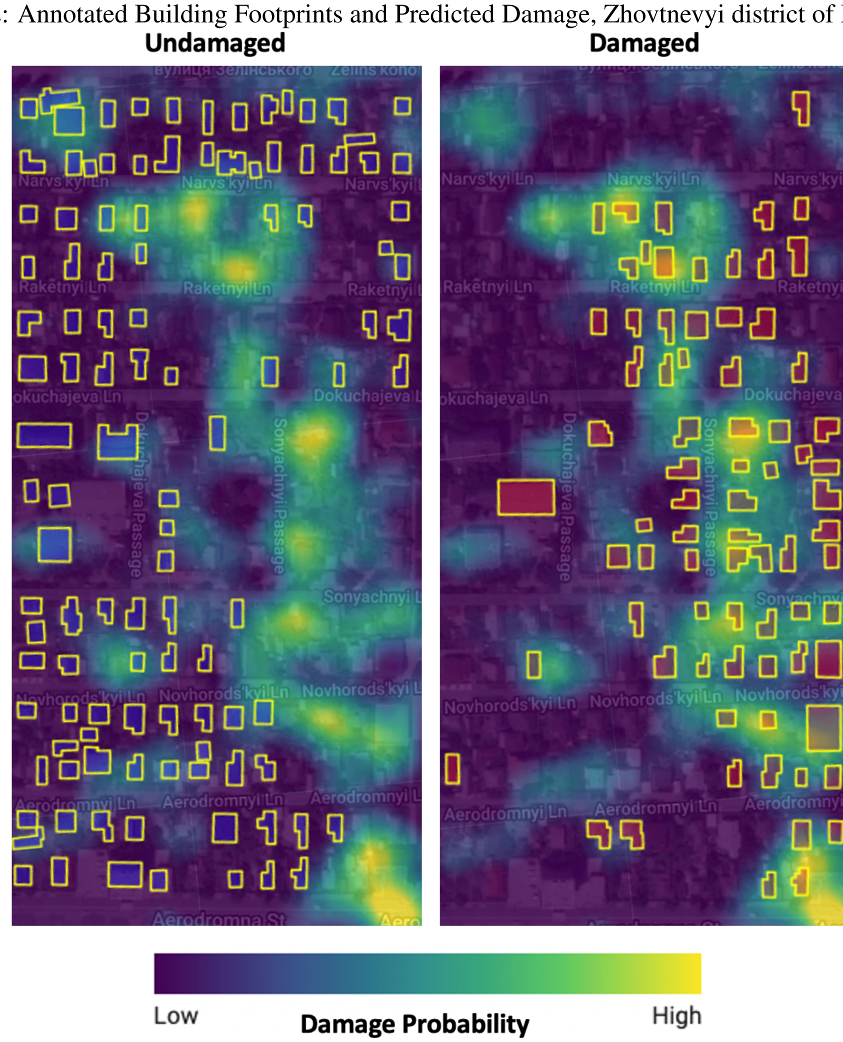
Figure 1: Receiver-Operator Curves



The application of deep learning to high resolution optical imagery results in an overall AUC of 0.90 and precision of 0.42 (Mueller et al., 2020). Though AUC for both Mariupol and Irpin falls short of this, precision is actually higher, at 0.62 for Mariupol and 0.58 for Irpin when using the optimal cutoff from the ROC curves. Depending on the nature of the task at hand, a more conservative model could be deployed by selecting a higher cutoff.

Not only are comparable (and by some measures, better) results achieved using open access imagery and a relatively simple algorithm, the accuracy assessment in this case is more stringent than in Mueller et al. (2020). Some of the building footprints used in the accuracy assessment are smaller than the size of a single Sentinel-1 pixel (10 meters by 10 meters). The figure below shows the damage probability raster for the Zhovtnevyi district of Mariupol. The panel on the right shows annotated building footprints labeled damaged by UNOSAT, and the panel on the left shows undamaged buildings.

Figure 2: Annotated Building Footprints and Predicted Damage, Zhovtnevyi district of Mariupol



Though the mean area of these footprints is 112 square meters—roughly the size of a single Sentinel-1 pixel— there is strong spatial alignment between the annotated footprints and the PWTT damage prediction.

A limitation of this analysis is that the data necessary for formal validation is only available in two cities. However, the interactive application also includes the locations of georeferenced conflict footage from the Centre for Information Resilience, which allows for ad-hoc ground-truthing of predicted damage. Figure 5 in the appendix shows geolocated photographs of a school in Kharkiv that was bombed, and high values on the damage probability raster.

3.2 Beirut and Turkey

An important advantage of this approach is that it is highly flexible, and can be adapted to different use cases. Though the previous two examples used a relatively even number of pre- and post-war images, this approach works well even when a relatively small number of post- destruction images are available. This enables its use in situations more akin to natural disasters, where the majority of destruction results from a single traumatic event such as an explosion on an earthquake.

To demonstrate the utility of the PWTT in the assessment of damage from individual events, this subsection performs a visual assessment of the algorithm's performance in two scenarios: the explosion in Beirut on August 8th, 2020, and the earthquake in Southern Turkey on February 6th, 2023. Both scenarios use 12 months of pre-event imagery. For Beirut, only one month of post-event imagery is included, and in Turkey, a single post-event image is included.

Figure 5 in the appendix displays two images. On the top is the damage probability raster generated using 12 months' worth of imagery prior to August 8th, 2020, and one month following the event. It shows a significant amount of damage radiating outwards from the epicenter of the explosion. The image on the bottom is a placebo test of the algorithm using the exact same parameters as the top image, but using the same date in the previous year. Though some areas display significant change, the quantity and magnitude thereof is much lower. Upon closer inspection, these areas are revealed to be places one might expect change: the crane loading terminal at the port of Beirut, and several construction sites.

Finally, Figure 5 displays the results from the use of a single post-event image in central Kahramanaras, which was devastated by an earthquake on February 6th, 2023. The left panel shows building footprints with an average T value exceeding 1.4, and the panel on the right shows a high resolution image of the same area taken the day after the earthquake.

4 Conclusion

This paper develops a novel algorithm for the detection of building damage using a Pixel-Wise T-Test on Synthetic Aperture Radar imagery. The pipeline can be deployed immediately following a disaster or the onset of a conflict, and can ingest new imagery as it becomes available, becoming more accurate over time. The use of SAR imagery means that the model can be deployed virtually anywhere on earth regardless of atmospheric conditions and cloud cover. The current approach achieves accuracy statistics that rival those of computationally and financially more expensive models, using only open-access imagery and relatively simple computation. In doing so, this methodology takes a significant step towards the generation of accurate, near-real time damage maps of building damage that can be put to use by humanitarian practitioners.

References

Morton J. Canty, Allan Aasbjerg Nielsen, Knut Conradsen, and Henning Skriver. Statistical analysis of changes in sentinel-1 time series on the google earth engine. *Remote Sensing*, 2019. doi: 10.3390/rs12010046.

Knut Conradsen, Allan Aasbjerg Nielsen, J. Schou, J. Schou, and Henning Skriver. A test statistic in the complex wishart distribution and its application to change detection in polarimetric sar data. *IEEE Transactions on Geoscience and Remote Sensing*, 2003. doi: 10.1109/tgrs.2002.808066.

Amit Kumar, Manisha Manisha, Preet Lal, Alisha Prasad, Pratyush Tripathy, and Purabi Saikia. Analysing urban damage and surface deformation based hazard-risk in kathmandu city occurred during nepal earthquake (2015) using sar interferometry. *Advances in Space Research*, 2022. doi: 10.1016/j.asr.2022.02.003.

Hannes Mueller, Andre Groeger, André Gröger, André Gröger, Jonathan Hersh, Andrea Matranga, Andrea Matranga, and Joan Serrat. Monitoring war destruction from space using machine learning. *arXiv: General Economics*, 2020. doi: 10.1073/pnas.2025400118.

Samuel Francisco Tiongson and Ryan Ramirez. Mapping of ground surface deformations and its associated damage using sar interferometry: a case study of the 2020 masbate earthquake. *E3S web of conferences*, 2022. doi: 10.1051/e3sconf/202234703014.

Frank D. W. Witmer. Remote sensing of violent conflict: eyes from above. *Journal of remote sensing*, 2015. doi: 10.1080/01431161.2015.1035412.

5 Appendix

Figure 3: Geolocated footage of a destroyed school in Kharkiv showing high PWTT damage probability

Figure 4: PWTT on the Port of Beirut

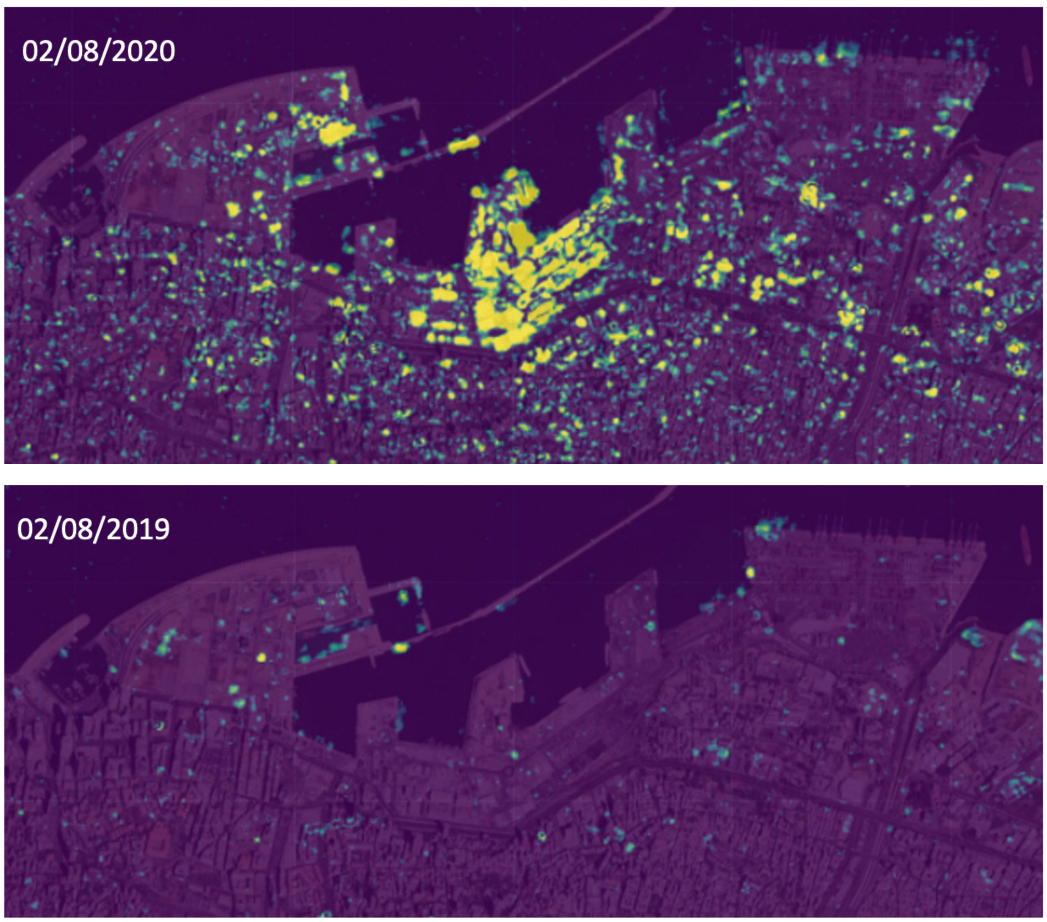


Figure 5: Predicted damaged buildings in Turkey and High-Resolution image