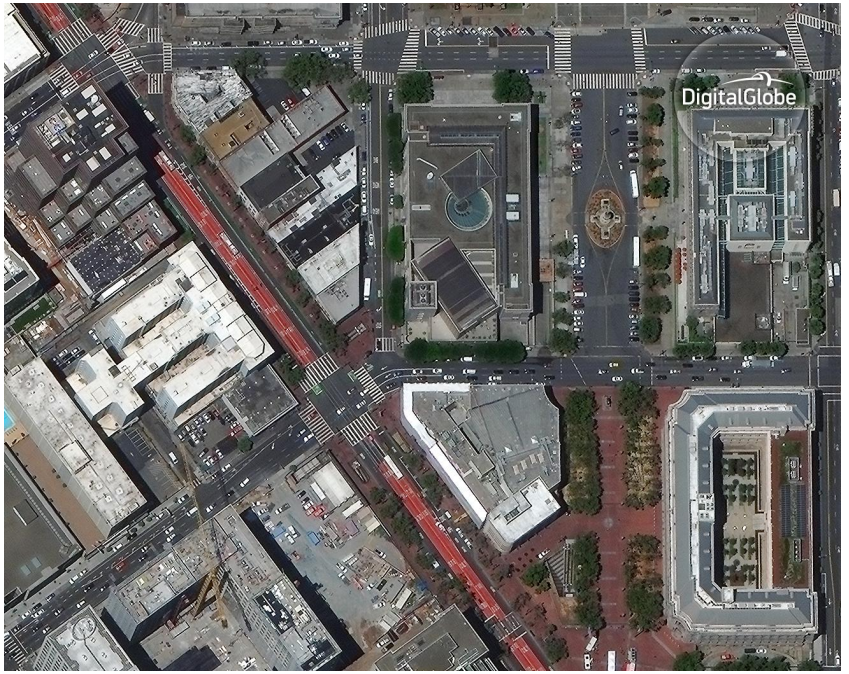
Benchmark Dataset for Automatic Damaged Building Detection from Post-Hurricane Remotely Sensed Imagery

Presentation by Jennifer Grannen

Introduction

- Current approaches
  - Reliant on human visual inspection
  - Labor intensive - inefficient and expensive
  - Not much focus on wind damage
  - Many approaches not generalizable
- Not many easily formatted datasets to train damage detection algorithms
- **Goal for this paper:** create scalable framework for benchmark dataset of hurricane damaged buildings
  - Developed for training object detection models



Data Development Process

- Obtain data with damaged buildings
- Join damage annotations with images to build dataset
- Create building footprint envelopes on images and merge with damage annotations

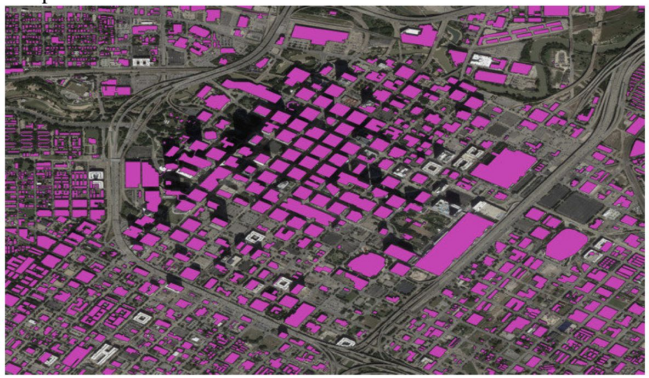
Data Criteria

- Verifiability
- Represent damage in visually apparent way
- Comprehensiveness in area of study
- Functionality
    - How easily can it be used for training?

Data Sources

- Annotation Data
  - TOMNOD project - crowdsourced annotation on DigitalGlobe satellite imagery
  - FEMA data
- Imagery Datasets
  - DigitalGlobe open sourced imagery - satellite images
  - NOAA aerial imagery survey - aerial images
- Raster Imagery with Building Footprints
  - Microsoft national building footprint dataset
  - Oak Ridge National Laboratory dataset
  - Parcel data from specific counties used to fill in gaps

(a) A sample image from Oak Ridge National Laboratory building footprints.



(b) A sample image from Microsoft's building footprints.

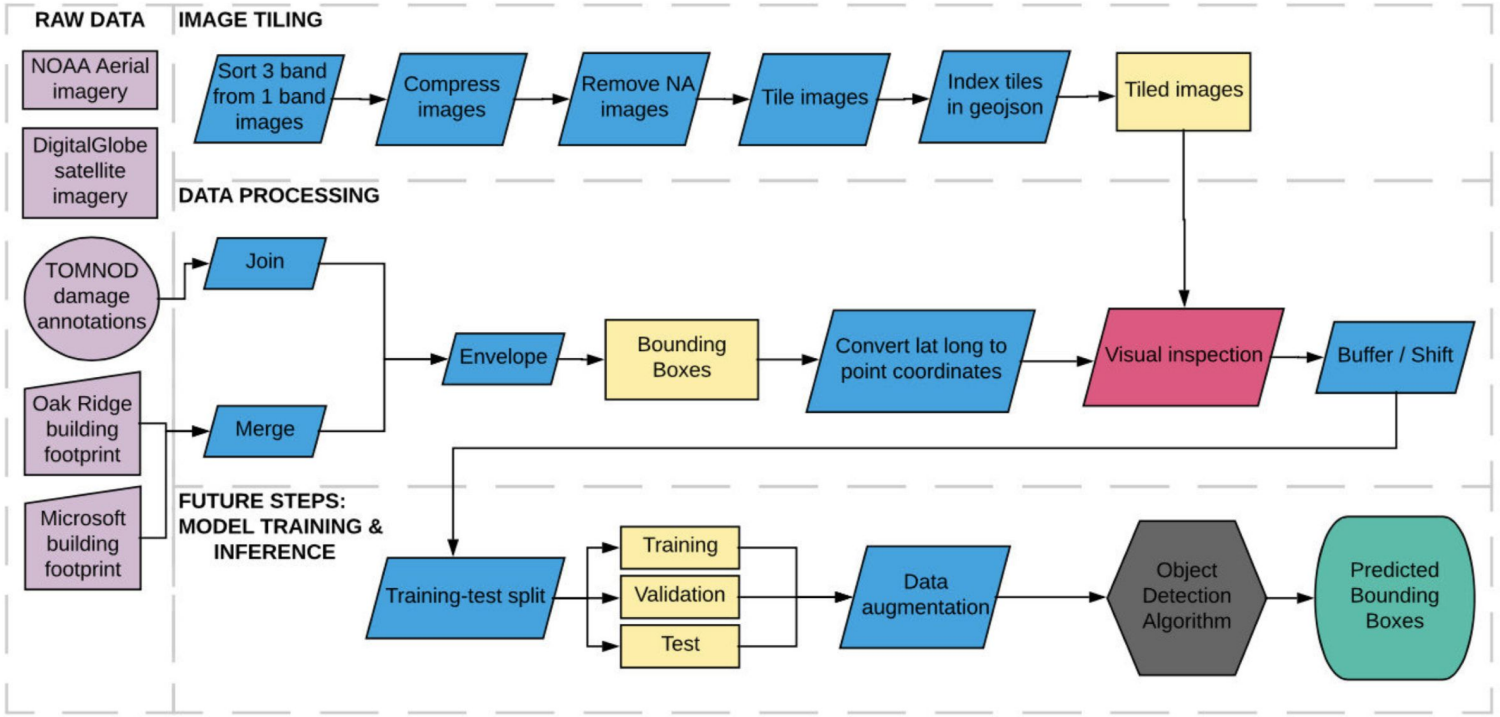
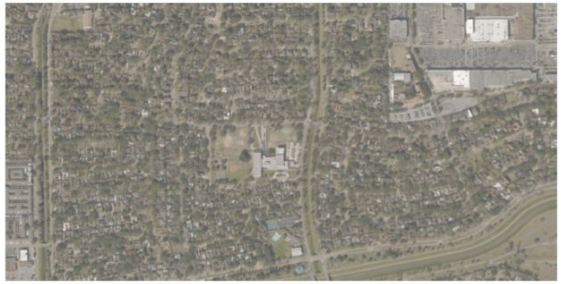


Fig. 1: **Benchmark Dataset Preparation Process.** In the above diagram we describe the steps of creating a benchmark dataset: the first row indicates the preprocessing steps which are required to convert the large raw datasets to a more manageable tiled format; the second row describes how the damage annotation vector data is joined with the raster data to obtain corresponding bounding boxes; the last row illustrates a traditional workflow that a machine learning practitioner will take to train object detection algorithms on the resulting benchmark dataset.

Conclusion

- Selected FEMA labels with NOAA images and parcel data building footprints
- This benchmark dataset is publicly available



(a) FEMA points

(b) Identified affected parcels

(c) Identified individual structures

(d) Bounding boxes from identified structures