Superpixel-wise Assessment of Building Damage from Aerial Images

Lukas Lucks, Dimitri Bulatov, Ulrich Thönnessen and Melanie Böge

Fraunhofer Institute of Optronics, System Technologies and Image Exploitation
Gutleuthausstr. 1, 76275 Ettlingen, Germany

Keywords: Damage Detection, Superpixels, Feature Extraction, Random Forest, Classification.

Abstract: Surveying buildings that are damaged by natural disasters, in particular, assessment of roof damage, is challenging, and it is costly to hire loss adjusters to complete the task. Thus, to make this process more feasible, we developed an automated approach for assessing roof damage from post-loss close-range aerial images and roof outlines. The original roof area is first delineated by aligning freely available building outlines. In the next step, each roof area is decomposed into superpixels that meet conditional segmentation criteria. Then, 52 spectral and textural features are extracted to classify each superpixel as damaged or undamaged using a Random Forest algorithm. In this way, the degree of roof damage can be evaluated and the damage grade can be computed automatically. The proposed approach was evaluated in trials with two datasets that differed significantly in terms of the architecture and degree of damage. With both datasets, an assessment accuracy of about 90% was attained on the superpixel level for roughly 800 buildings.

1 INTRODUCTION

According to the United States National Centers for Environmental Information (NCEI) and National Oceanic and Atmospheric Administration (NOAA), 2017 was the most expensive year of losses since 1980[1]. A total of 16 weather and climate disaster events caused US$ 306.2 billion of damage that had to be covered by insurance companies. In many cases, handling insurance claims and carrying out loss adjustment is sometimes more expensive than the loss itself. A loss adjuster has to be routed through a damaged city to reach each building in a portfolio. Depending on the severity of the natural disaster, the degree of urbanization, and the development of the area, it sometimes takes weeks to locate an insured building. Then, to identify the roof damage, the loss adjuster usually has to climb onto the roof; this tends to be a bottleneck in the damage assessment of a building.

However, policyholders need to be paid as soon as possible so that they can begin repairs. Hence, insurance companies are extremely interested in (1) reducing the time-consuming steps required and (2) automating the process as much as possible to predict losses faster. In this study, we focus on the damage caused by storms such as Hurricane Irma. Close-range aerial images of the roof structures are obtained in a high-throughput manner as many buildings can be imaged in a single flight. These aerial images are then analyzed to estimate the percentage of damaged roof area. In this way, although the opinions of loss adjusters cannot be replaced completely, automated localization of roof damage simplifies the process, eliminates the need for the time-consuming and dangerous task of climbing onto roofs, and accelerates the assessment of buildings with different degrees of damage. Moreover, this approach will provide better insight into the situation and help insurance companies set priorities accordingly.

Image-based detection of damage or other anomalies is a pattern recognition problem. Here, we leverage state-of-the-art tools for the analysis of patterns and textures. We present a detailed review of relevant literature and introduce the proposed workflow for automatic estimation of roof damage. The developed method is tested on two different portfolios after Hurricane Irma; the images in the datasets are post-loss (i.e., post-event) images and are often available in high resolution. An outline of the roof area is also required; building footprints are often available from cadastral offices or freely available databases such as Open Street Map (OSM). Because of registration errors in images and non-nadir views, the outlines have to be adjusted to roof polygons as a preprocessing step in the algorithm. Next, superpixel decompo-

---

[1] https://www.ncdc.noaa.gov/billions/overview, 03/12/2018

sition of the roof area is applied to reduce the computation time required for the subsequent classification step. To perform classification, three features are considered: unfiltered image channels, filter banks, and morphological profiles. Moreover, higher-level features are also assessed to extract the straight lines and repetitive texture patterns associated with roofs. The features are collected and classified superpixel-wise using the Random Forest classifier. The only user interaction is to select training regions to define references of damaged and undamaged building parts. Finally, each superpixel is assigned to either the damaged or undamaged class. These results can be used to localize roof damage and compute the grade of damage for an entire building.

2 RELATED WORK

An obvious approach is to determine the amount of damage from pre- and post-loss data. Several such approaches have been proposed. Block-wise damage assessment was conducted by considering differences in the average intensities and variances of images (Zhang et al., 2002). In another study (Tomowski et al., 2010) of a relatively small and rural region in Darfur, Sudan, four different cost functions were applied to four texture parameters; thus, in total, 16 responses were analyzed.

As machine-learning methods are implemented to integrate multiple features and identify appropriate thresholds user-determined thresholds have become less popular. Various studies have investigated morphological profiles, structural and radiometric features (Pesaresi et al., 2007) or correlation coefficients from a co-occurrence matrix (Rathje et al., 2005). Another study (Gueguen and Hamid, 2015) also generated data with annotated geo-referenced relevant changes on the ground that is now freely available and often used as training data for semi-supervised techniques. However, all images in this dataset must be re-sampled to a uniform scale, resulting in rather low resolution. To deal with certain radiometric properties in the images that are substantially different, different approaches have been used: assessing the damage object-wise rather than pixel-wise to analyze linear segments (Huyck et al., 2005) or measure certain metrics such as the normalized differential vegetation index (NDVI) (Gamba et al., 2007), segment properties (Im et al., 2008), etc.

Another study (Tu et al., 2017) used pre-event satellite images only to localize the buildings, projecting the building footprints onto a high-resolution post-event image. To correlate both buildings in sa-

tellite images and buildings in aerial images, the authors applied Support Vector Machines (SVMs) over the composed hue, saturation, value (HSV) indexes of the pixels and the 128 entries of the dense Scale-Invariant Feature Transform (SIFT) descriptor.

Moreover, (Fujita et al., 2017) applied Convolutional Neural Networks (CNNs) to analyze pairs of pre- and post-event color images if available or only post-event images where pre-event images were not available. The training data were annotated manually and all available images were stored within a dataset, which contained images of several buildings that were destroyed by flooding. However, it was impossible to assess the intermediate damage grades in this dataset using the proposed approach. Two important trends can be observed in (Tu et al., 2017) and (Fujita et al., 2017): first, using recent techniques, pre-event images are not necessary and, second, these techniques rely on high-dimensional spaces of features without explicit semantic meaning. A good example of these high-dimensional feature spaces is a CNN, which represent a universal framework that provides suitable solutions to a wide class of problems, including the assessment of roof damage from aerial images as considered herein (Fujita et al., 2017; Vetrivel et al., 2017; Cooner et al., 2016). However, CNNs crucially depend on a huge amount of training data for the affected regions, which cannot always be retrieved rapidly. Recently, significant progress has been achieved in the pixel-wise collection of results using CNNs (see (Maggiori et al., 2016) for a detailed discussion of these methods).

Many studies have shown a well-designed approach with a standard classifier can produce results similar to those obtained using CNNs (Fujita et al., 2017; Cooner et al., 2016) and emphasized the necessity of using three-dimensional (3D) features to improve the results (Vetrivel et al., 2017). For these reasons, we will postpone the implementation of CNNs to future work and, instead, pursue an alternative strategy.

The study of (Sirmacek and Unsalan, 2009) is one example for approaches that rely only on post-event images and model assumptions rather than on large training datasets. The model assumption is that shadows are missing around destroyed buildings. Thus, the challenge is to distinguish between buildings with and without shadowed regions. However, this method does not recognize whether roof tiles have been blown away by the wind and shadows can often be mistaken for other objects, necessitating relatively large datasets of features. In a subsequent study (Ma and Qin, 2012), spectral, spatial, and morphologic features were combined to achieve building-

wise detection rates around 90%. Other studies (Rasika et al., 2006; Gerke and Kerle, 2011) have focused on the detection of damage from images taken at oblique angles. The previously mentioned study of (Vetrivel et al., 2017) focused on localizing damage. In this approach, the image is subdivided into superpixels, which are approximately equally-sized image segments that ideally coincide with the image edges. Around every superpixel, a patch is formed and subjected to CNN-based evaluation. At the same time, 3D features resulting from the eigenvectors of the structure tensors of different radii are calculated in a point cloud and projected onto the superpixels. Moreover, to integrate features from different modalities, a multiple-kernel-learning framework (Vetrivel et al., 2017) was investigated. Based on the findings, the authors reported that it is very difficult to differentiate some complex textures from damage without considering 3D features.

Two main conclusions can be drawn from these prior studies: First, high-resolution aerial images can be used for precise localization of roof damage and, thus, can greatly support the loss evaluation that is traditionally carried out by loss adjusters. In addition, recently developed classification techniques can be used to conduct damage assessment sufficiently using only post-event images. However, due to variation between different natural disasters, it is difficult to determine a priori which type of destruction will occur (Dell'Acqua and Gamba, 2012). Hence, large and sophisticated pre-trained databases are of limited usefulness. To cope with this, our algorithm is designed to compensate for a small number of training examples by using a hand-tailored, purpose-based, and quickly extractable feature set.

3 DAMAGE DETECTION

A successful damage detection and other important components of the algorithm depend on several steps, including roof outlining supported by free geographic data, superpixel decomposition, selection of training data, and classification. The result can be transfered into a task-specific damage grading.

3.1 Generating Roof Outlines from Building Footprints (Alignment)

The analysis process requires roof outlines enclosing the areas of interest. This outline must fit the roof outline and must not include the surrounding ground. For many cities, freely available building footprints

exist, such as those from Open Street Map. However, in many cases, the building outlines do not satisfy these requirements. In the absence of precisely fitting roof outlines, freely available building outlines may be aligned to the actual roof outline as follows.

Let $\mathcal{J}$ denote the region of interest in an airborne RGB image, including a building and the surrounding area, and let $\mathcal{P}$ be the footprint of this building. We are looking for a transformation φ to align $\mathcal{P}$ with the roof of the building such that the corresponding edges of the transformed building outline $\mathcal{P}(\varphi)$ coincide with the roof outline in $\mathcal{J}$. In our case, φ is a two-dimensional translation but, in general, it may be necessary to use six- or even eight-dimensional vectors to represent affine or projective transformations. In addition, all pixels of the image patch in $\mathcal{J}$ are labeled as inside, outside, and border depending on their position according to $\mathcal{P}(\varphi)$; this label mask is denoted as $\mathcal{M}$ $(= \mathcal{M}(\mathcal{P}))$. Since the freely available building outlines and roofs overlap approximately, we assume that these data provide a sufficient starting point for the objective function.

The objective function achieves two purposes. First, a modification of the mutual-information function is applied to assess the dominant color, $\mathbf{f} \in \mathbb{R}^3$, sampled from a 3D histogram over the color values of all pixels in $\mathcal{J}$ labeled as inside according to $\mathcal{M}$. Second, we ensure that the norm of the image gradient is significantly higher at the border pixels than at pixels inside $\mathcal{P}(\varphi)$. Thus, the overall cost function can be expressed as follows:

$$E(\varphi) = \sum_{\mathbf{p} \in \mathcal{M}(\mathcal{P}(\varphi))} \{\alpha w_{\mathbf{f}}(\mathcal{M}(\mathbf{p})) \|\mathcal{J}_{\mathbf{f}}(\mathbf{p})\|$$
$$+ (1 - \alpha)\tilde{w}_{\nabla}(\mathcal{M}(\mathbf{p})) \|\nabla \mathcal{J}(\mathbf{p})\|\}, \quad (1)$$

where $\|\cdot\|$ is the L_1 norm, which is applied to achieve robustness with respect to outliers, $\tilde{\ }$ denotes Gaussian smoothing, and $\alpha = 0.5$ is a balance parameter, which may be adjusted depending on the distinctiveness of $\mathbf{f}$ (e.g., higher for red if most buildings in the area have red roofs and lower for green and gray since these colors are typically associated with vegetation and streets, respectively). Furthermore, $\mathcal{J}_{\mathbf{f}}(\mathbf{p}) = \mathcal{J}(\mathbf{p}) - \mathbf{f}$ is computed for each channel; $w_{\mathbf{f}}$ is equal to 1 inside $\mathcal{P}(\varphi)$ and 0 otherwise. Finally, w_{∇} is equal to -1 at the border of $\mathcal{P}$, $\varepsilon = 0.01$ inside, and 0 outside. Equation (1) can be minimized by applying the gradient-free Nelder-Mead method as in (Lagarias et al., 1998).

3.2 Superpixels

The damage grade is derived directly from the amount of damaged roof area. To locate damaged roof patches as precisely as possible, the roof area is subdivided

into small sub-areas. Each sub-area is classified as either *undamaged* or *damaged*.

A suitable choice for sub-areas is derived by a superpixel decomposition of the building mask. Superpixels (Jiang et al., 2015) are small image entities that include several image pixels and they typically group pixels of uniform color, texture, etc. Although superpixel computation can be time-consuming, the results are certainly worth the effort because the extracted features gain robustness and they are invariant to changes in the image resolution. Further, the computation of superpixels reduces the classification time since fewer entities need to be evaluated. Finally, as introduced in Sec. 3.3, high-level features can be incorporated as the use of superpixels implies that the data describe regions and neighborhood relations.

For superpixel decomposition, we follow the implementation of compact superpixels as described in (Veksler et al., 2010), which is based on the use of Graph Cuts (Boykov et al., 2001). This approach involves minimizing an energy function, which consists of a data term that prohibits a superpixel from leaving the area that was initially reserved for it and a smoothness term that ensures that the superpixels have compact borders. Although newer and faster methods exist, this tool was used here because the parameter settings had already been validated and the risk of under-segmentation is quite low. After segmentation, several quite fast post-processing steps are necessary: tiny superpixels are merged with their neighbors, any superpixels that lie outside of or on the border of the roof outline are removed, and topologically consistent and surjective labeling is enforced.

The key to detecting damaged roof areas lies in evaluating of a variety of properties (see Sec. 3.3) for each superpixel, which allows to classify them either as undamaged or damaged.

3.3 Features and Classification

Features summarize the properties of each superpixel, taking into account all varieties of roof types and damages such as blown-away shingles, collapsed areas, etc., and are used to differentiate between damaged and undamaged parts of a roof. The most obvious of these features are the unfiltered color channels of a red-green-blue (RGB) image and combinations thereof, such as the saturation, NDVI, and opponent gaussian color space (OGCS) (Geusebroek et al., 2001). Here, differential morphological profiles, which are very popular in remote sensing due to their invariance under changes in the contrast and shape of a characteristic region, were applied. The MR8 filter bank (Varma and Zisserman, 2005) offers

rotational invariance and facilitates the detection of edges and blobs. Entropy as an indicator for homogenity and therewith intactness of a superpixel is also evaluated. Each of these features is computed pixelwise but the obtained values are grouped to derive the mean and variance for each superpixel. The use of superpixels results in the inclusion of high-level features such as lines and recurring structures. For example, a long line on the roof indicates a proper edge, while short lines of differing orientations are characteristic of damage. Similarly, a regular texture pattern is a strong indicator of an undamaged superpixel even if the average gradient norm is high. To enhance the superpixels near long edges, straight line segments with minimum lengths of 1.5 m are computed in the image, rasterized and, finally, intersected with the superpixels. The number of line segments running through a superpixel is denoted as the lineness (Bulatov et al., 2011). To further characterize texture properties, a modified version of the HOG (histogram of oriented gradient) features was used (Dalal and Triggs, 2005). The gradient orientations weighted by their occurrence are collected modulo π, discretized into 0.1π steps, normalized, smoothed as in (Pohl et al., 2017), and sorted in ascending order. Finally, the cumulative distribution is evaluated. Only the first three entries of the histogram are used to highlight the superpixels with only a few dominant gradient orientations. The three buildings in Fig. 1 show selected features and their corresponding values (alarm rates) at the superpixels. The features clearly delineate the damaged sections of the roof from the undamaged ones. In addition, the roof edges and texture salience can be detected.

Altogether 52 features were considered to determine whether the roof area covered by one superpixel is damaged or not. To make this decision, a Random Forest classifier (Breiman, 2001) with suitable training data was used. The number of trees was estimated by adding trees to the Random Forest and analyzing the out-of-bag error. It was found that from about 40 trees the error did not significantly change (less than 0.5%) with further trees. Hence, the amount of trees for the classification was limited to that number. The advantages of the Random Forest algorithm are its robustness against redundant features and its efficiency during calculations because of parallelization. The Random Forest algorithm supplies a probability for each prediction. In our case, this output corresponds to the probability that a superpixel is damaged or not damaged. Hence, a superpixel with a probability of less than 0.5 is classified as undamaged and others are classified as damaged.

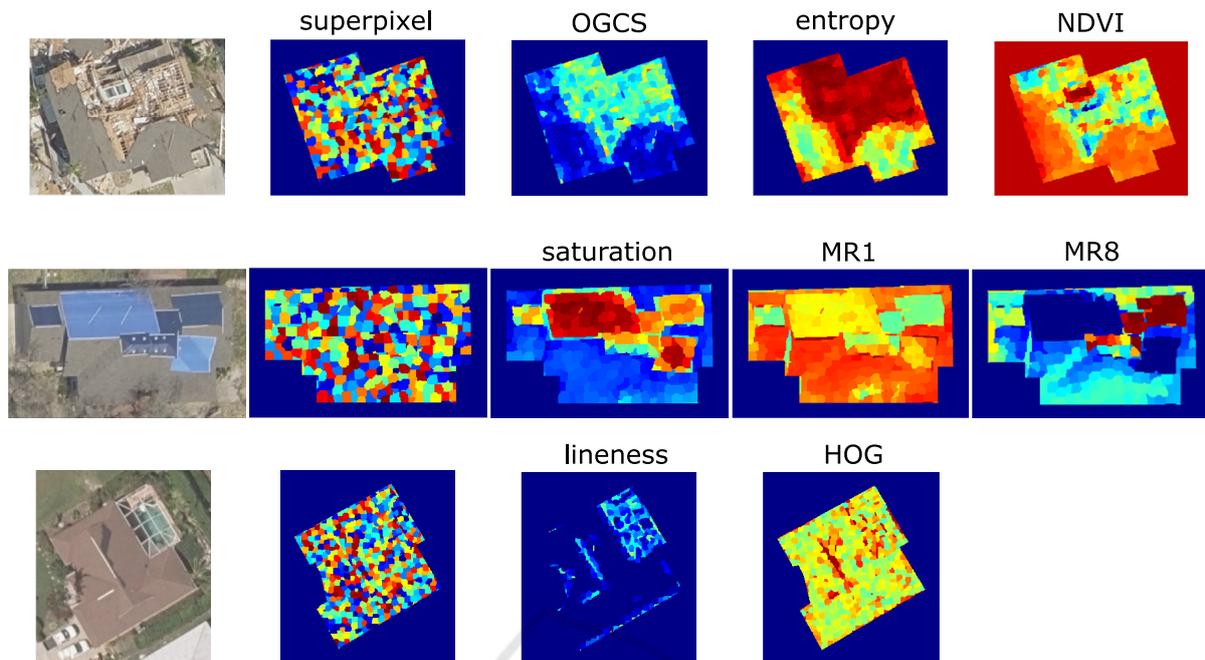Tests involving forward feature selection have

Figure 1: Alarm rates of selected features shown for three exemplary buildings. The first seven feature images clearly delineate the damaged parts of the roof from undamaged areas. The lineness and HOG reliably indicate the positions of roof edges and the texture salience.

shown that not all feature were necessary. However, in order not to perform too many modifications in our algorithm, we relied on the robustness of the classifier with respect to the redundant and irrelevant features (Genuer et al., 2010; Warnke and Bulatov, 2017).

3.4 Training Data

The selection of training data is an essential step for reliable classification. Hence, it is important to include appropriate representatives of as many situations as possible in the training set, including different types of damage, such as blown-away shingles and broken roofs, as well as different types of undamaged roofs, such as solar panels, areas of sound roof material, intact pools, air conditioners, and other superstructures (as illustrated in Fig. 2). The occlusion of roofs by trees is an ambiguous problem as it can range from branches covering the view of the roof or small branches scattered on the roof causing no damage to uprooted trees fallen on the roof and causing significant damage. As in other cases of ambiguousness, by way of precaution, it was decided to classify all regions covered by trees as damaged.

The training data must be chosen carefully for each class. For the undamaged class, as many intact roof entities as possible, such as solar panels, chimneys, air conditioners, and different kinds of roof material, should be included in the training data. For the damaged class, all types of damage should be included. It is also necessary to include heterogeneous structures with different degrees of damage and different types of roof structures. Moreover, incorrect assignments of undamaged roof areas as damaged (and vice versa) may disturb the classification.

The selection of training data is the only step in the workflow that requires user input. Here, we selected buildings that were distributed evenly across the dataset to represent a portfolio. For each building, superpixels covering wide parts of the roofs were labeled as either damaged or undamaged. As many types of structures as possible were selected from both classes to be included in the training data. The higher the inhomogeneity of roofs is (meaning roofs consisting of various materials or differing strongly in architecture), the more training data are necessary to distinguish between damage from different roof materials.

However, the training data do not have to be recreated for every dataset. In cases with similar images and roof types, the training data from different datasets may be transferable. We will consider this possibility in Sec. 4.

3.5 Damage Grading

Since the damage is localized by assigning each superpixel to one of the two classes, we are able to compute the corresponding areas of damaged and unda-

Figure 2: High-resolution images showing different types of damage.

maged roof. Thus, it is possible to determine the ratio between the total number of pixels in damaged super-pixels and the number of pixels assigned to the roof area, which reflects the damage grade of the building. Depending on the task or for map visualization (see Fig. 3), these values can further be decomposed into damage categories: intact, light damage, medium damage, and heavy damage.

4 RESULTS AND DISCUSSION

To demonstrate the usability of the proposed algorithm, we analyzed two different datasets: D1 with a ground resolution of 7.5 cm, which comprises 421 randomly chosen buildings in the city of Rockport, Texas, with homogeneous roof structures suffering heavy damage, and D2 with a ground resolution of 15 cm (sampled up to 5 cm/px), which comprises 416 buildings in Marco Island, Florida, with inhomogeneous roof structures and less damage. Both cities were impacted by Hurricane Irma in 2017. The relatively small number of buildings enabled a qualitative comparison with a manual reference set generated by experts. Every superpixel in each data set covered approximately 0.75×0.75 m^2.

Herein, we evaluated the performance and accuracy of the proposed procedure. Moreover, glass roofs represent a special case that is unique to roof damage detection in that the subjacent floor is visible in the aerial images; thus, the algorithm has to deal with both the floor and roof structure simultaneously; to test the behavior of the proposed method in this case, the pools attached to buildings in D2 were also evaluated. In addition, the superpixel-wise approach was compared with a pixel-wise approach, the transferability of the training data was investigated, and the potential for using additional classes was explored. Finally, due to unambiguous boundary values, we excluded from the evaluation all superpixels (or pixels)

that share a border with the roof outline; thus, we also evaluated the results of roof boundary registration.

4.1 Accuracy

Precisely fitting roof outlines were available for these datasets and were used to evaluate the accuracy. This enabled us to avoid propagating the systematic error associated with the alignment (see Sec. 3.1).

Representative results from D1 are shown in Fig. 4. The middle column shows the location maps of the damage resulting from the assignment of damaged or undamaged superpixels. The probability maps show the certainty of the resulting decisions: dark blue and dark red indicate classifications of undamaged and damaged roofs, respectively, with high certainty. Interestingly, the white patch of roof on the lower-left corner in the second example was correctly classified as damaged, although this patch does not differ visually from other intact white roof structures.

To assess the accuracy of the procedure, wide ranges of each building were labeled. A three-fold cross-validation was applied to all labeled examples to avoid bias due to the choice of training and testing data. The available portfolio is divided into training data and test data. Note that the partitioning in these two datasets has to be applied on a building-wise basis; some buildings are assigned to training set others to test set. It would also be possible to use a superpixel-wise partitioning randomly chosen from each building. But this holds the danger that the classifier would likely learn samples of each damage that is present and result in non-representative good results of damage detection due to an over-fitted classifier. The validation of all 54517 labeled samples in D1 revealed a testing accuracy of 89%. The corresponding confusion matrix is given in Table 1. Thus, it can be concluded that undamaged patches of roof can be recognized slightly better than damaged patches.

Figure 3: Visualization of damaged buildings on a map. Buildings in the portfolio are colored according to the damage ratio (green: intact, yellow: light damage, orange: medium damage, and red: heavy damage).
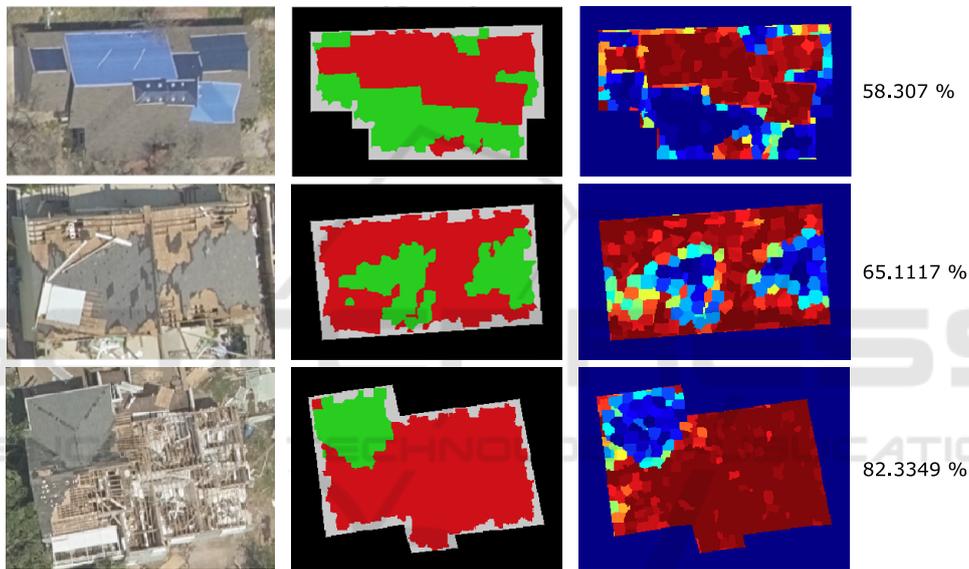

58.307 %

65.1117 %

82.3349 %

Figure 4: Selected results for D1: (left) orthophoto, (middle) locations of damage, and (right) probability map and damage grading. Dark red indicates regions that are most likely damaged while dark blue indicates regions that are most likely undamaged. The proportion of damaged roof area is given by the percentage values to the right.

The roofs in D2 are more diverse, featuring varied structures with different architectures and colors as well as elevated objects, than those in D1, which are generally homogeneous. In D2, the roofs are bordered by straight or curved lines and often include surfaces with many complex orientations. Due to this structural diversity, a larger amount of training data is required since distinguishing between roof structures and dam-age is more difficult. Examples of damage localized in D2 are shown in Fig. 5a. The three-fold cross-validation (see Tab. 1) revealed an overall testing accuracy of 91% although only 76% of all damaged roof patches that are present were recognized, which is approximately 10% less than that in D1. Here, only damage to the main roofs (no glass roofs) was considered. This lower recognition rate can be attributed to the lower resolution of the images and, more importantly, the greater inhomogeneity due to differing architecture and roof colors. Moreover, less damaged regions are present in the dataset. Hence, even a lower recognition rate can lead to a high overall accuracy.

To further evaluate our method, the results should be compared with other studies, keeping in mind the different settings and databases. Another approach based on neural networks (Vetrivel et al., 2017), the achieved overall accuracies are sligtly lower (89% or 91% referring to around 93%). However, the differences are small, even though the presented method needs less training data and no additional 3D information. Looking at methodes which only rely on pre and post event images (Thomas et al., 2014), it can be

217

Table 1: Confusion matrices for results of damage detection in D1 and D2.

D1		predicted			D2		predicted		
		undam.	dam.	prec. (%)			undam.	dam.	prec. (%)
reference	undam.	30672	2109	92.13	reference	undam.	31793	1564	95.31
	dam.	4069	17667	81.35		dam.	1906	4998	72.39
	acc. (%)	88.17	87.27	88.67		acc. (%)	94.34	76.17	91.38

(a) only main buildings　　　　　　　　　　　(b) with pools

Figure 5: Damage location maps for several examples in D2.

concluded that the performance are sligtly lower. The achieved accuracies range from around 68% to 91%, but it must be noted that the evaluation was done on the building level and not superpixel-wise, which makes the comparison difficult.

4.2 Glass Roofs

Swimming pools with glass roofs were attached to several buildings in D2 and presented a challenge for the proposed approach. A glass roof provides a nadir-view of the ground floor. Thus, the entity comprises both the ground floor and the structure of the glass roof. This algorithm was not designed for such simultaneous evaluation of roof and floor structures. As aspected, the rate of damage recognition was only approximately 61%. The detection of destroyed or completely missing glass structures is understandably difficult. Nevertheless, impressive results were attained in this trial: in many cases, the algorithm learned to differentiate between regular roof grids and destroyed roof grids so even damage to the glass roof could be identified in many cases as shown in Fig. 5b. The overall testing accuracy was 89%, which is comparable to that obtained without considering pools.

4.3 Comparison of Superpixel-wise and Pixel-wise Evaluation

We also tested a pixel-wise approach and compared it to the use of superpixels. The evaluation of 100

buildings in D1 required 12570000 pixels to be analyzed compared to 62600 superpixels. Thus, the number of superpixels was less than 1% of the number of pixels and greatly reduces the computational requirements. The evaluation of superpixels took only 6% of the computation time required to evaluate pixels but provided 10% better accuracy. One example of an imprecise pixel-wise result is shown in Fig. 6. While it was initially hypothesized that a pixel-wise localization would be more effective, the lower accuracy achieved in the pixel-wise evaluation can be attributed to the loss of context and neighborhood information using a pixel-wise approach.

4.4 Multiclass Evaluation

We further analyzed the potential for multi-class evaluation in which regions were classified into several classes that were more specific than damaged and undamaged. The damaged class was divided into damaged roofs and pools (glass roofs) and additional classes, such as solar panels, shadows, and air conditioners, were used for further differentiation in the undamaged class. However, this multi-class evaluation did not affect the accuracy of the proposed method.

4.5 Alignment

To test the performance of the building outline alignment introduced in Sec. 3.1, 100 buildings in D2 were manually selected and the necessary translation para-

superpixelwise   pixelwise

Figure 6: Comparison between superpixel-wise and pixel-wise evaluation often leads to a less precise result since information including neighborhoods and surroundings are omitted.

meters to align the available building outline with the roof were extracted. The algorithm was run on the same data and the resulting offsets between building outline and roof outline were compared. The initial root-mean-squared error (RMSE) was about 1.58 m and was reduced to 0.7 m by the alignment, which is close to the average size of one superpixel. Unfortunately, but understandably, the result differed for non-damaged and damaged buildings (0.62 and 0.87 m, respectively). The error in alignment can be attributed to poor fitting between the roof and the roof outline in terms of shape or size and insufficiencies in the convergence behavior of the optimization.

5 CONCLUSION AND FUTURE WORK

We presented a semi-automated approach for detection and localization of damaged roof patches in post-event aerial images. Images of sufficient resolution and precise outlines of the roofs are required to carry out the analysis. The roof area is divided into superpixels, which are then classified as either damaged or undamaged. In this way, the damage can be rated and localized. Providing training examples for the classifier constitutes the only interactive step in the algorithm. Because of a sophisticated choice of features, only a small number of training examples are required, in contrast to previous studies (Fujita et al., 2017), which rely on large data banks and perform classification using feature sets of deep neuronal architectures.

For future work, the hard choice of a threshold (in this case, it was set to 0.5) for classifying superpixels can be adjusted on the basis of Receiver Operated Characteristics (ROC) curves. The threshold may need to be adapted according to its surroundings. It is possible, though not probable, that an isolated superpixel remains undamaged but is surrounded by damaged superpixels. Therefore, corrections using Markov Random Fields could be useful.

Differentiating between types of roof damage is a complex challenge that should be investigated further. In this work, the detected anomalies summarize light (e.g., lost shingles) and heavy (e.g., collapsed parts of a roof) damage, with uprooted trees and branches overlapping a roof causing severe damage. However, in the current approach, collapsed roofs receive the same damage grade as those that have lost shingles. In future work, it would be useful to consider such types of damage by e.g. including 3D information and near-infrared measurements. Nevertheless, the results obtained by our procedure are extremely important for insurance companies; it enables them to make a first quick extrapolation of incurred loss and the sum insured associated with that to make the money available or to contact re-insurance companies.

Finally, even though the proposed approach was developed in the context of damage detection, the method and the provided insights of this paper were useful and could be empolyed for other related applications, such as roof analysis for installation of solar panels or roof window detection.

REFERENCES

Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239.

Breiman, L. (2001). Random Forests. *Machine learning*, 45(1):5–32.

Bulatov, D., Solbrig, P., Gross, H., Wernerus, P., Repasi, E., and Heipke, C. (2011). Context-based urban terrain reconstruction from UAV-videos for geoinformation applications. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3822:75–80.

Cooner, A. J., Shao, Y., and Campbell, J. B. (2016). Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 Haiti earthquake. *Remote Sensing*, 08-00868(10):1–17.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Proceedings of the*

IEEE Conference on Computer Vision and Pattern Recognition, volume 1, pages 886–893. IEEE.

Dell'Acqua, F. and Gamba, P. (2012). Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives. *Proceedings of the IEEE*, 100(10):2876–2890.

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

Gamba, P., Dell'Acqua, F., and Odasso, L. (2007). Object-oriented building damage analysis in VHR optical satellite images of the 2004 tsunami over Kalutara, Sri Lanka. In *Urban Remote Sensing Joint Event, 2007*, pages 1–5. IEEE.

Genuer, R., Poggi, J.-M., and Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14):2225–2236.

Gerke, M. and Kerle, N. (2011). Automatic structural seismic damage assessment with airborne oblique pictometry© imagery. *Photogrammetric Engineering & Remote Sensing*, 77(9):885–898.

Geusebroek, J.-M., Van den Boomgaard, R., Smeulders, A. W. M., and Geerts, H. (2001). Color invariance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12):1338–1350.

Gueguen, L. and Hamid, R. (2015). Large-scale damage detection using satellite imagery. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1321–1328.

Huyck, C. K., Adams, B. J., Cho, S., Chung, H.-C., and Eguchi, R. T. (2005). Towards rapid citywide damage mapping using neighborhood edge dissimilarities in very high-resolution optical satellite imagery Application to the 2003 Bam, Iran, earthquake. *Earthquake Spectra*, 21(S1):255–266.

Im, J., Jensen, J., and Tullis, J. (2008). Object-based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing*, 29(2):399–423.

Jiang, L., Lu, H., My, V. D., Koch, A., and Zell, A. (2015). Superpixel segmentation based gradient maps on RGB-D dataset. In *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1359–1364. IEEE.

Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E. (1998). Convergence properties of the Nelder–Mead Simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1):112–147.

Ma, J. and Qin, S. (2012). Automatic depicting algorithm of earthquake collapsed buildings with airborne high resolution image. In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, pages 939–942. IEEE.

Maggiori, E., Tarabalka, Y., Charpiat, G., and Alliez, P. (2016). Fully convolutional neural networks for remote sensing image classification. In *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, pages 5071–5074. IEEE.

Pesaresi, M., Gerhardinger, A., and Haag, F. (2007). Rapid damage assessment of built-up structures using VHR satellite data in tsunami-affected areas. *International Journal of Remote Sensing*, 28(13-14):3013–3036.

Pohl, M., Meidow, J., and Bulatov, D. (2017). Simplification of polygonal chains by enforcing few distinctive edge directions. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10270 LNCS:3–14.

Rasika, A., Kerle, N., and Heuel, S. (2006). Multi-scale texture and color segmentation of oblique airborne video data for damage classification. In *Proceedings of ISPRS Midterm Symposium 2006 Remote Sensing: From Pixels to Processes*, pages 08–11.

Rathje, E. M., Woo, K.-S., Crawford, M., and Neuenschwander, A. (2005). Earthquake damage identification using multi-temporal high-resolution optical satellite imagery. In *Proceedings of the IEEE on Geoscience and Remote Sensing Symposium*, volume 7, pages 5045–5048. IEEE.

Sirmacek, B. and Unsalan, C. (2009). Damaged building detection in aerial images using shadow information. In *4th International Conference on Recent Advances in Space Technologies*, pages 249–252. IEEE.

Thomas, J., Kareem, A., and Bowyer, K. W. (2014). Automated poststorm damage classification of low-rise building roofing systems using high-resolution aerial imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(7):3851–3861.

Tomowski, D., Klonus, S., Ehlers, M., Michel, U., and Reinartz, P. (2010). Change visualization through a texture-based analysis approach for disaster applications. In *ISPRS Proceedings on Advanced Remote Sensing Science*, volume XXXVIII, pages 263–269.

Tu, J., Li, D., Feng, W., Han, Q., and Sui, H. (2017). Detecting damaged building regions based on semantic scene change from multi-temporal high-resolution remote sensing images. *ISPRS International Journal of Geo-Information*, 6(5):131.

Varma, M. and Zisserman, A. (2005). A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1-2):61–81.

Veksler, O., Boykov, Y., and Mehrani, P. (2010). Superpixels and supervoxels in an energy optimization framework. *Proceeding on European Conference on Computer Vision*, pages 211–224.

Vetrivel, A., Gerke, M., Kerle, N., Nex, F., and Vosselman, G. (2017). Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images, and multiple-kernel-learning. *ISPRS Journal of Photogrammetry and Remote Sensing*.

Warnke, S. and Bulatov, D. (2017). Variable selection for road segmentation in aerial images. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42.

Zhang, J.-F., Xie, L.-L., and Tao, X.-X. (2002). Change detection of remote sensing image for earthquake-damaged buildings and its application in seismic disaster assessment. *Journal of Natural Disasters*, 11(2):59–64.