ISPRS Journal of Photogrammetry and Remote Sensing xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

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

Anand Vetrivel [a,*], Markus Gerke [b], Norman Kerle [a], Francesco Nex [a], George Vosselman [a]

[a] University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), P.O. Box 217, 7500 AE Enschede, The Netherlands
[b] Technical University of Braunschweig, Institute of Geodesy and Photogrammetry, Pockelsstraβe 3, 38106 Braunschweig, Germany

ARTICLE INFO

Article history:
Received 6 December 2016
Accepted 1 March 2017
Available online xxxx

Keywords:
Oblique images
UAV
3D point cloud features
CNN features
Multiple-kernel-learning
Transfer learning
Model transferability
Structural damage detections

ABSTRACT

Oblique aerial images offer views of both building roofs and façades, and thus have been recognized as a potential source to detect severe building damages caused by destructive disaster events such as earthquakes. Therefore, they represent an important source of information for first responders or other stakeholders involved in the post-disaster response process. Several automated methods based on supervised learning have already been demonstrated for damage detection using oblique airborne images. However, they often do not generalize well when data from new unseen sites need to be processed, hampering their practical use. Reasons for this limitation include image and scene characteristics, though the most prominent one relates to the image features being used for training the classifier. Recently features based on deep learning approaches, such as convolutional neural networks (CNNs), have been shown to be more effective than conventional hand-crafted features, and have become the state-of-the-art in many domains, including remote sensing. Moreover, often oblique images are captured with high block overlap, facilitating the generation of dense 3D point clouds – an ideal source to derive geometric characteristics. We hypothesized that the use of CNN features, either independently or in combination with 3D point cloud features, would yield improved performance in damage detection. To this end we used CNN and 3D features, both independently and in combination, using images from manned and unmanned aerial platforms over several geographic locations that vary significantly in terms of image and scene characteristics. A multiple-kernel-learning framework, an effective way for integrating features from different modalities, was used for combining the two sets of features for classification. The results are encouraging: while CNN features produced an average classification accuracy of about 91%, the integration of 3D point cloud features led to an additional improvement of about 3% (i.e. an average classification accuracy of 94%). The significance of 3D point cloud features becomes more evident in the model transferability scenario (i.e., training and testing samples from different sites that vary slightly in the aforementioned characteristics), where the integration of CNN and 3D point cloud features significantly improved the model transferability accuracy up to a maximum of 7% compared with the accuracy achieved by CNN features alone. Overall, an average accuracy of 85% was achieved for the model transferability scenario across all experiments. Our main conclusion is that such an approach qualifies for practical use.

© 2017 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction and related works

Automated detection of severe building damages is crucial in the coordination of fast response actions after any destructive disaster event such as earthquakes. Remote sensing technology has been recognized as a suitable source to provide timely data for automated detection of damaged buildings for larger areas (Dell'Acqua and Gamba, 2012; Dong and Shan, 2013). In particular, multi-view oblique images from manned aircraft and unmanned aerial vehicles (UAV) have been recognized as most suitable (Fernandez Galarreta et al., 2015; Gerke and Kerle, 2011; Kerle and Hoffman, 2013). This is because these images capture both roofs and façades with very high spatial resolution, facilitating a

* Corresponding author.
E-mail addresses: a.vetrivel@utwente.nl (A. Vetrivel), m.gerke@tu-bs.de (M. Gerke), n.kerle@utwente.nl (N. Kerle), f.nex@utwente.nl (F. Nex), george.vosselman@utwente.nl (G. Vosselman).

http://dx.doi.org/10.1016/j.isprsjprs.2017.03.001
0924-2716/© 2017 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

Please cite this article in press as: Vetrivel, A., et al. 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 J. Photogram. Remote Sensing (2017), http://dx.doi.org/10.1016/j.isprsjprs.2017.03.001

holistic and detailed view of the building for damage assessment (Fernandez Galarreta et al., 2015). Several studies have demonstrated automated detection of damaged buildings from the above mentioned image types, where the heavily damaged buildings are identified by recognizing externally visible damage evidences such as spalling, debris, rubble piles and broken elements, which are the strong indicators of severe structural damage (Dong and Shan, 2013; Vetrivel et al., 2015a). These damage evidences alone are not sufficient to infer the actual damage state of the building, as it requires additional information such as damages to internal building elements (e.g., columns and beams), which can rarely be directly inferred from images. Even though the information that can be derived from the images is limited, it is typically sufficient for satisfying the requirements of the stakeholders involved in search and rescue processes (Dong and Shan, 2013). Furthermore, the information can be used to plan for subsequent detailed assessments, for example, identifying hotspots that require immediate attention, and prioritizing the locations for field inspection. Towards this, numerous automated methods have been proposed for detection of aforementioned visual damage evidences from very high resolution images (Dong and Shan, 2013; Ma et al., 2016). These methods are largely based on two approaches: (1) comparison of pre- and post-event data, and (2) damage detection based on mono-temporal post-event data alone. The methods based on supervised learning strategies have been demonstrated to be effective for damage detection, particularly for the mono-temporal approach (Gerke and Kerle, 2011; Vetrivel et al., 2015a). However, it is still challenging to adopt them for practical use. This is because the methods based on a supervised learning approach often do not generalize enough for them to be transferred to similar remote sensing data from unseen geographic locations, and Vetrivel et al. (2015a) discussed several reasons. One of the major factors is the poor generalization capability of the features and their representation used for constructing the supervised model, which is briefly described below:

(1) Numerous image features have been examined for damage detection, and often the texture features such as Histogram of oriented Gradients (HoG) and Gabor features have been reported as effective (Samadzadegan and Rastiveisi, 2008; Tu et al., 2016; Vetrivel et al., 2015a). Apart from feature selection, the choice of the feature representation strategy is also crucial, which is evident from the recent study by Vetrivel et al. (2016b), where the performance of the above mentioned texture features was found to be improved when represented using a Visual Bag of Words (BoW) framework (Ferraz et al., 2014). Though the BoW representation improved the accuracy, problems related to generalization still exist. For example, Vetrivel et al. (2016b) examined the generalization capability of three different texture features: speeded up robust features (SURF), HoG features and Gabor features in a BoW framework for damage detection, using very high resolution images from different geographic locations (e.g., Italy, Haiti, India, etc.). They reported that the performance of the features is moderately inconsistent for datasets from different places, i.e. particular features perform better for specific datasets. The difference between the accuracies produced by these features for different datasets was reported to be 3–4%. The same set of features in a similar experimental setting as reported in Vetrivel et al. (2016b) was examined by Tu et al. (2016) for another study area for damage detection. However, they reported contradictory findings: the difference in accuracy produced by different features was found to be higher ($\sim$10%), though there is no obvious explanation for this difference in results. Thus, identifying the generalized features for building a supervised classifier for damage detection is still challenging.
(2) Additionally, all aforementioned features which have been reported as being efficient for damage detection are based on gradient orientation distribution patterns. These features are adopted for the damage detection process based on the assumption that structurally deformed regions often result in non-uniform radiometric distributions when compared to regions of undamaged man-made structural elements. For example, Fig. 1a depicts the rudimentary gradient orientation pattern derived for damaged and undamaged image regions. However this assumption often fails in urban areas possessing complex texture (Vetrivel et al., 2016a). For example, consider Fig. 1b where the building elements possess complex textures that look similar to the radiometric pattern of damaged regions. In such areas, the reported texture features would fail, thereby hindering the automated assessment.

Overall, the previously reported features are found to be inadequate to create a strong generalized supervised model for damage detection, and a feature descriptor robust to aforementioned limitations is highly desirable.

Recently, the features from deep learning approaches such as Convolutional Neural Networks (CNNs) have been reported as being superior to conventional hand-crafted features, including the ones used in earlier state-of-the-art BoW framework for image classification in many applications including remote sensing (Hu et al., 2015; Karpathy et al., 2014; Sherrah, 2016; Szegedy et al., 2015; Zhou et al., 2015; Zuo et al., 2014). For example, several participants in the ISPRS urban scene classification challenge have achieved state-of-the-art accuracy for the ISPRS Vaihingen and Potsdam benchmark data sets using CNN features, outperforming all previously reported methods based on hand-crafted features (cf. ISPRS-Benchmark, 2016). Hence, we anticipate that CNN features would outperform the hand-crafted features in a damage detection application as well. This is examined in this paper.

In the real world, man-made structural elements are complex and they often possess irregular radiometric patterns due to several reasons other than damage, including radiometric degradation of elements due to aging, or presence of dirt (cf. Fig. 2). In such cases, our assumptions about damaged regions based on image-radiometric patterns may fail. In this scenario, the use of 3D geometric information could be of help to differentiate between the unusual radiometric pattern due to geometric deformation (damage) and other reasons. In general, 3D point clouds are an ideal source to infer geometric characteristics of structural elements. For example, Khoshelham et al. (2013) demonstrated the potential of 3D point cloud features derived from post-event LiDAR point clouds for building damage detection. The oblique-view aerial images from manned- and unmanned aerial vehicles which have previously been identified as effective for damage detection are usually captured with high block overlap, facilitating the generation of 3D point clouds (Nex and Remondino, 2014). We assume that the integrated use of 3D features from photogrammetric point clouds and CNN features from images would yield improved results. However, it is well known that the direct integration of features, i.e. stacking of features from different sources, possibly possessing different modalities, into a single feature vector for supervised classification is inefficient (Bucak et al., 2014; Gu et al., 2015). Alternatively, integrating features from different sources using a Multiple-Kernel-Learning (MKL) approach associated with a kernel-based classifier such as SVM has been reported to be effective and it is being commonly used (Bucak et al., 2014; Gu et al., 2015). In addition to feature subsets integration, the MKL also could be used to evaluate the significance of each feature

Please cite this article in press as: Vetrivel, A., et al. 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 J. Photogram. Remote Sensing (2017), http://dx.doi.org/10.1016/j.isprsjprs.2017.03.001

Download English Version:

https://daneshyari.com/en/article/6949149

Download Persian Version:

https://daneshyari.com/article/6949149

Daneshyari.com