 DOI: 10.3217/978-3-85125-603-1-20

Detection of Bomb Craters in WWII Aerial Images*

Simon Brenner[1], Sebastian Zambanini[1] and Robert Sablatnig[1]

Abstract— **The analysis of aerial images from World War II surveillance flights allows a preliminary estimation of unexploded ordnance risk for large scale construction projects. To support this task, which is currently carried out manually, an automatic approach for the detection of bomb craters in such historical images was developed and evaluated.**

I. INTRODUCTION

Unexploded Ordnance (UXO) from World War II still poses a hazard for construction projects in Central Europe. Specialized companies provide a preliminary risk estimation by retrieving and interpreting aerial images from WWII surveillance flights over areas of interest. For this, such historical aerial images have to be georeferenced and searched for certain objects that indicate increased combat activity in the surveyed area, such as bomb craters, trenches or artillery stations. Currently, both the georeferencing and the search for warfare evidence are performed manually by specialists. Within the FFG-Bridge project DeVisOR (**De**tection and **Vis**ualization of unexploded **O**rdnance **R**isks), which was conducted from 2015 to 2017 in cooperation with the Luftbilddatenbank Dr. Carls GmbH (LBDB) as an industrial project partner, methods for the automation of the aforenamed tasks were developed.

This paper is concerned with the detection of warfare evidence in WWII aerial images, where we focused on the detection of bomb craters. First, they are the most frequent type of warfare-related objects found in the aerial images; second, they are the most direct evidence for the presence of UXO, as at least 10% of all bombs that were dropped in WWII are assumed to have not exploded [6]. We developed a machine learning approach based on Convolutional Neural Networks (CNNs) for the automatic detection of bomb craters. Furtheremore, the integration of the detector into the working environment of our industrial partner in the form of a plugin for the GIS *ArcMap* will be elaborated.

II. RELATED WORK

Merler et al. [6] developed an approach to automatically generate UXO risk maps via bomb crater detection in WWII aerial images. They created a set of *eigen-craters* from a principal component analysis of example craters and trained a classifier based on a variant of AdaBoost. The results are promising, but they are aiming at detecting clusters of craters. The dataset they were using for evaluation was not published.

More work has been done on the automatic detection of asteroid impact craters on extraterrestrial surfaces [3][7][5], lately also using CNNs [2]. Although the problem seems to be similar to ours, there are some major differences. First, asteroid impact craters exhibit a high variation in size; second, the source images are of better quality than the historical aerial images; and third, on other planets there are no trees or man made objects that can easily be confused with craters.

III. BOMB CRATER DATASET

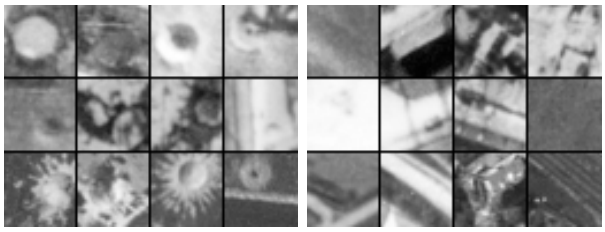
To our knowledge, no labelled dataset for crater detection in historical aerial images is currently publicly available. LBDB as an industrial partner provided a selection of their finished projects, covering both urban and rural areas. In these projects, the historical aerial images have been georeferenced and the bomb craters mapped by experts. The analysis is only performed within a defined region of interest (ROI). Usually, several overlapping historical images from different dates have been georeferenced in the ROI, to ease the determination of warfare evidence by the expert.

In total the provided projects contain about 10000 craters in world coordinates. After semi-automatically assigning the craters to the images in which they are visible, we ended up with roughly 20000 craters in image coordinates, along with their diameters. Equally, the ROI is mapped to the individual images; this is crucial for evaluation, as no ground truth data are available outside the ROI.

The crater positions exported in this format are very flexible and can be used in different ways to train and evaluate machine learning algorithms. For our approach, a CNN was trained for a binary classification problem on image patches (see Section IV). We therefore prepared our training data by extracting image patches of positive and negative examples. As the ground resolutions of the individual aerial images are approximately known, images patches with a fixed absolute size are extracted. A statistical assessment of crater sizes resulted in a mean crater diameter of 7.8m with a standard deviation of 2.0. Assuming a normal distribution, 95% of the crater diameters can be assumed to lie between 3.8m and 11.8m. In order to provide the classifier with some context, we set the patch size to 20x20m. The patches are extracted with a sliding window over the ROIs, with a stride of 1/4 the patch size. Patches containing a crater position are stored as positive examples, patches not containing a crater are randomly selected as negative samples or discarded, in order to match the number of positive samples. This approach was

*This work was supported by Austrian Research Promotion Agency (FFG) under project grant 850695

[1]Simon Brenner, Sebastian Zambanini and Robert Sablatnig are with Faculty of Informatics, Institute of Computer Aided Automation, Computer Vision Lab, TU Wien, 1040 Vienna, Austria sbrenner@cvl.tuwien.ac.at, zamba@cvl.tuwien.ac.at, sab@cvl.tuwien.ac.at



(a) positive (b) negative

Fig. 1: Examples of training patches

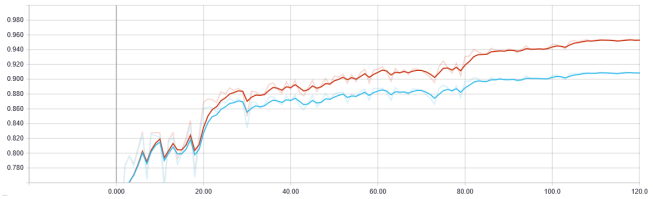


Fig. 2: Training of the CNN. Mean accuracy on the training set (red) and test set (blue) over training epochs.

preferred to the direct extraction of patches centered at the crater positions, as it already mimics the planned detection procedure and implicitly introduces translational variation within the examples. With this approach about 85000 patches per class were generated. Figure 1 shows some examples of extracted patches; note that both the positive and the negative examples are very heterogenous. The dataset was randomly divided into training and test set with a ratio of 4:1 and scaled to fit the CNNs input.

IV. THE CNN CLASSIFIER

First experiments with local binary patterns, Haar features and simple neural networks could not deliver satisfying results. We therefore employed the DenseNet CNN architecture [4], which emerged as state of the art in image classification in 2017. Using the training data described in Section III, we trained a 40 layer DenseNet with an input size of 32x32 pixels on a binary classification problem. The network was optimized using Nesterov Momentum. After experiments with learning rates we ended up at the learning curve depicted in Figure 2, plateauing at a mean accuracy of 0.91 on the test set after approximately 110 epochs.

Detection is performed by moving a sliding window of fixed ground size of 20x20m over the ROI, with a stride of 1/4 the window size, similarly to the generation of the training data. The sub-images covered by the sliding window are then scaled to the correct input size and processed by the CNN, which returns a confidence for that window containing a crater.

As craters only make up an average of 0.4% of the observed areas, classification by simply applying a threshold to the confidences leads to an unfeasible number of false positives; the precision of the raw approach on realistically distributed data was at 0.04. Raising the threshold does not solve the problem; it only lowers the recall without significantly improving the precision. The following section describes a post-processing approach that exploits spatial information to tackle the problem.

V. POSTPROCESSING

To alleviate the precision problems described in the previous section, spatial information and a-priori assumptions about bomb crater distributions are exploited to filter the CNN output. Specifically, the following ideas are employed:

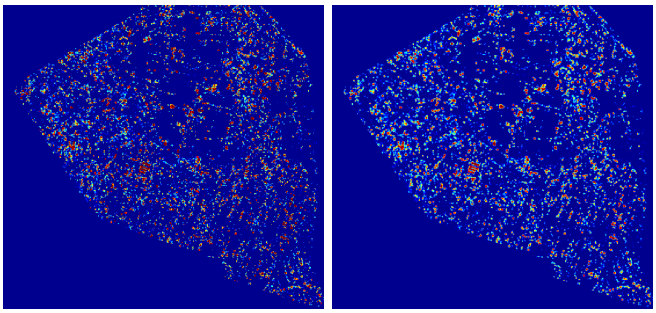
1) **Spatial proximity prior:** Bombs are typically dropped in clusters; 'lonely' detections are therefore more likely to be false positives.
2) **Non-cluster suppression:** As every individual crater is hit by the sliding window more than once, it is unlikely that a real crater is only detected in one of those window positions; therefore detections that are not part of a cluster are considered outliers.
3) **Non-maximum suppression:** This is a standard operation in object detection that reduces multiple detections of the same object to the one with the maximum confidence.

To practically apply the above named ideas, first the confidences computed for sliding window positions are stored in a 2-dimensional *confidence map*. Figure 3a visualizes an example of such a map. To introduce the spatial proximity prior, the map is convoluted with a gaussian kernel with $\sigma = 0.5$ (Figure 3b). This penalizes isolated areas of high confidences. The resulting filtered confidence map is thereafter thresholded to produce a binary *detection map* (Figure 3c). The threshold was set to 0.88, which maximizes the F1-score of the detector. For the non-cluster suppression, 8-connected components with less than 6 elements are removed from the detection image (Figure 3d). Finally, all detections which are not local maxima in the filtered confidence map are removed from the detection map, resulting in the final set of detections (Figure 3e). Figure 3f shows those detections superimposed on the original image.

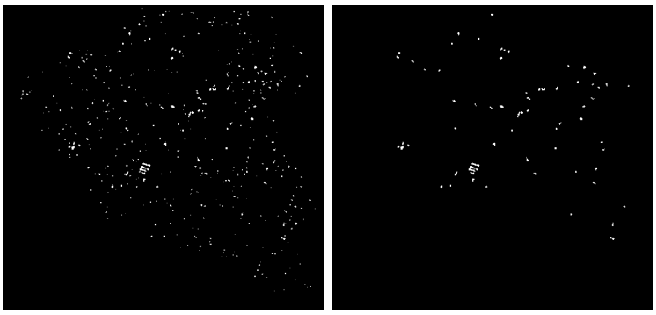
The described procedure converts the CNN outputs to individual crater positions for single images. However, as described in Section III, typically several overlapping aerial images are available for a given ROI. Therefore, detections from different images have to be merged to receive the complete set of craters for a ROI. In this stage, double detections from different images are eliminated. For that purpose, all craters are transformed to world coordinates and a neighborhood-based clustering is applied, where neighborhood is defined by a maximum euclidean distance, which was determined empirically. The clusters are then replaced by their centroids.

VI. RESULTS

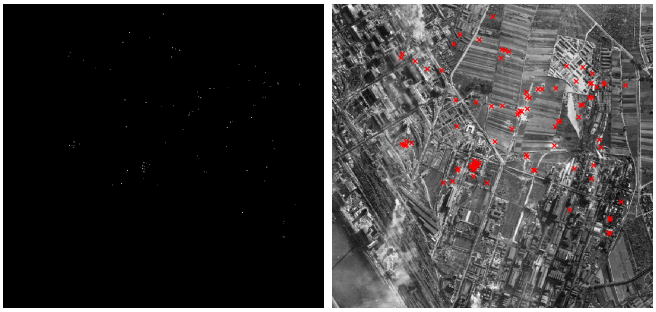
The CNN was evaluated on the test set of image patches as described in Section III. At the maximum accuracy of 0.91 at a decision threshold of 0.5 and a 1:1 ratio of positive and negative examples, patches with craters were detected



(a) confidence map (b) gauss filtered confidence map



(c) detection map (d) non cluster suppression



(e) non maximum suppression (f) final detections

Fig. 3: Visualization of postprocessing steps (example)

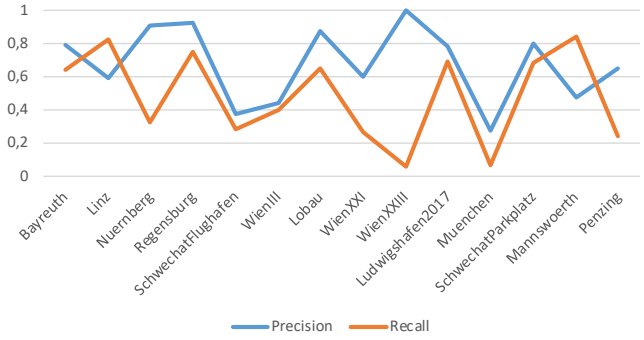


Fig. 4: Precision and recall for the different test projects

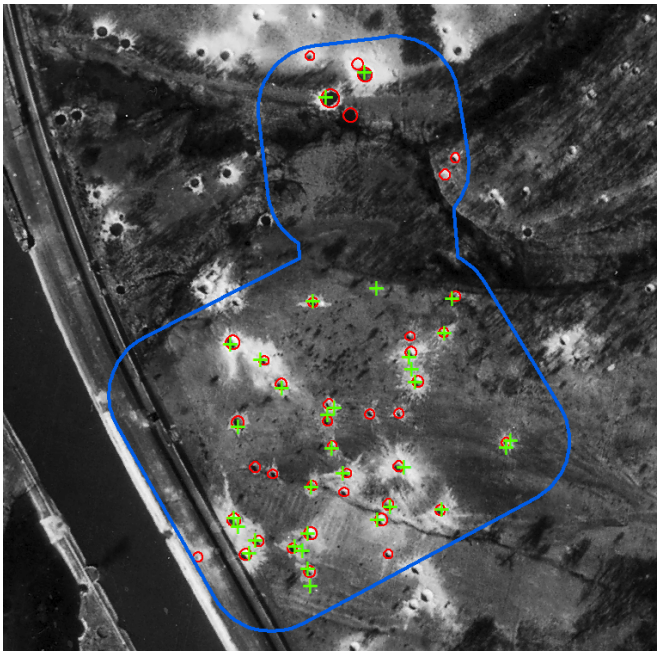


Fig. 5: A detetion result created by the ArcMap plugin. Green: our detected craters. Red: ground truth craters. Blue: ROI.

with precision of 0.907 and a recall of 0.913. In a realistic scenario with a ratio of approximately 250 negative examples on 1 positive example the precision drops to 0.04.

The end-to-end detection solution including the postprocessing steps described in Section V was evaluated on 14 full example projects provided by LBDB with approximately 2500 ground truth craters in total. A detection was regarded a true positive if its euclidean distance to the closest ground truth crater is smaller than the diameter of that crater. Figure 4 shows the precision and recall of the detector for the individual test projects. As shown in this figure, the results vary significantly from project to project. The average results, weighted by the number of ground truth craters present in a project, are a precision of 0.74 and a recall of 0.6.

VII. IMPLEMENTATION

As this work was developed in the course of an industry-oriented project, the described detector was implemented in a form that could easily be adapted into the worlflow of the company partner. To this end we extended the *ArcMap* plugin for the registration of historical aerial images, which was developed earlier in the project [1], with our crater detection approach. The plugin enables the user to automatically detect craters in the source images linked to an *ArcMap* project. These images are first processed by the dense CNN using the TensorFlow framework. The resulting confidences are then post-processed in the plugin code and the resulting detections transformed to the world coordinate system of the project. Finally, the craters from different images are merged. The detected craters are then present as features in the *ArcMap* project and can be refined and edited by the user. Figure 5 shows an example of our detection results.

VIII. CONCLUSIONS

In this project, an approach for the automatic detection of bomb craters in WWII aerial images was developed and implemented. While the performance of our solution does

not allow an application as a fully automated system, it can assist the human specialist in a semi-automated fashion, thus saving time and effort. Additionally, the automated detection can provide a second instance of inspection, in the sense that it may draw the specialist's attention to ambiguous points, thus reducing the chance of overlooking important evidence in large regions of interest.

As a further use case, our approach would allow to automatically generate risk estimations for large areas, using the intermediate confidence maps. This would allow the company to quickly provide potential customers with a first rough estimate of the UXO risk of a certain area, before performing a more detailed analysis.

REFERENCES

[1] S. Brenner, S. Zambanini, and R. Sablatnig, "Image registration and object detection for assessing unexploded ordnance risks a status report of the devisor project," in *Proceedings of the OAGM&ARW Joint Workshop*. Verlag der Technischen Universität Graz, 2017, pp. 109–110.

[2] J. P. Cohen, H. Z. Lo, T. Lu, and W. Ding, "Crater detection via convolutional neural networks," *CoRR*, vol. abs/1601.00978, 2016.

[3] W. Ding, T. F. Stepinski, Y. Mu, L. Bandeira, R. Ricardo, Y. Wu, Z. Lu, T. Cao, and X. Wu, "Subkilometer crater discovery with boosting and transfer learning," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 4, pp. 39:1–39:22, July 2011.

[4] G. Huang, Z. Liu, and K. Q. Weinberger, "Densely connected convolutional networks," *CoRR*, vol. abs/1608.06993, 2016.

[5] J. R. Kim, J.-P. Muller, S. van Gasselt, J. G. Morley, and G. Neukum, "Automated crater detection, a new tool for mars cartography and chronology," *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 10, pp. 1205–1217, 2005.

[6] S. Merler, C. Furlanello, and G. Jurman, "Machine learning on historic air photographs for mapping risk of unexploded bombs," in *Proceedings of the 13th International Conference on Image Analysis and Processing*, ser. ICIAP'05. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 735–742.

[7] P. G. Wetzler, R. Honda, B. L. Enke, W. J. Merline, C. R. Chapman, and M. C. Burl, "Learning to detect small impact craters," in *WACV/MOTION*. IEEE Computer Society, 2005, pp. 178–184.