FEW-SHOT FILTERING FOR THE DETECTION OF SPECIALIZED CHANGE IN REMOTE SENSING

Martin Hermann, Sudipan Saha & Xiaoxiang Zhu
Technical University of Munich, Germany
{martin.hermann,sudipan.saha,xiaoxiang.zhu}@tum.de

ABSTRACT

Identifying change in remote sensing observations - such as satellite images - is an important step for many applications. In practice, the aim is usually not to find all differences, but rather only specific types of change, such as urban development, deforestation, or even more specialized categories like road construction. However, often there are no large public data sets available for very fine-grained tasks or non-standard usecases that might, for example, occur in developing countries, and to collect the amount of training data needed for most supervised learning methods is very costly and often prohibitive.

For this reason, we formulate the problem of *few-shot filtering*, where we are provided with a relatively large change detection data set and a few instances of one particular type of change that we try to "filter out" of the learned changes. This would enable stakeholders to work with the available resources and adapt them to their individual needs, for a wide range of applications.

1 MOTIVATION AND OVERVIEW

Change Detection, that is, segmenting a pair of images of the same region but taken at two different points in time into *changed* and *unchanged* pixels, is a well-known task in remote sensing, with many applications in disaster assessment, urban planning, forest monitoring and other remote sensing domains (Shi et al., 2020). Usually, for these applications we are not interested in every change that occurred between the images, but limit our attention to certain categories such as building construction or destruction (Saha et al., 2020), deforestation (Bem et al., 2020) or flooding (Peng et al., 2019), and possibly even smaller subcategories.

For this reason, *supervised learning*, where we can exactly specify our interests via annotated samples, is a natural choice for these change detection tasks. However, the downside of these approaches is that they usually require large amounts of training data, which is expensive to produce. This might encourage research to focus on application domains where large data sets are already publicly available, which in turn often are annotated with rather general categories such as *urban change*.

This is a threat to more specialized tasks with limited focus, as they are likely to receive less attention and stay relatively underexplored. The situation is similar for stakeholders and regions that have a different distribution of changes: For example, what is interesting and uncommon in one area might be abundant in another and drown the information that is actually needed. As often, there is a danger that this happens particularly to developing countries, as we might assume that many data sets are created with the needs of developed countries in mind. (For example, Daudt et al. (2019) and Shi et al. (2021), while providing very valuable resources, are limited to regions in France and Hong Kong, respectively.)

Therefore, we propose an approach that works *top-down*: First, we learn to classify a broader type of change in a binary classification task, for which ample training data is available, and then try to *filter out* one particular subcategory via only a few examples, by which we enter the realm of *few-shot learning* (Wang et al., 2020). As few-shot learning deals with novel classes that the machine learning model has never seen before, whereas we in contrast try to specialize and split known classes, we propose the term *few-shot filtering* for this task. To achieve our goal, we combine methods from the deep clustering literature (Kanezaki, 2018) and the few-shot learning community (Chen et al., 2018), (Qi et al., 2018), (Gidaris & Komodakis, 2018).

2 Problem Statement

The (binary) change detection task is concerned with a pair of images $\mathbf{I}_1, \mathbf{I}_2 \in \mathbb{R}^{H \times W \times C}$, where $H \times W$ are the spatial dimensions and C denotes the number of channels (which might be larger than the usual 3 in standard computer vision and can also include, e.g., near infrared bands). It is assumed that $\mathbf{I}_1$ and $\mathbf{I}_2$ depict the same region, but are taken at different points in time, often multiple months or (such as, e.g., for applications in urban development) years apart. Our aim then is to derive a *change map* $\boldsymbol{C}_{I_1,I_2} \in \{0,1\}^{H \times W}$, i.e., a segmentation of the images into pixels that have changed in some meaningful way between $\mathbf{I}_1$ and $\mathbf{I}_2$ (denoted by a value of 1) and those that remain unchanged.

In supervised learning, we assume a training set of image pairs and their corresponding change map. This training set is usually limited to one change type $\mathcal{T}$, such as urban development or seasonal changes. A standard change detection task would now be to evaluate a model on a test set that is sufficiently similar to the train set. For the few-shot filtering problem however, we assume a *support set* that is very small (such as two or three small patches of a larger image) and contains examples of a change type $\mathcal{T}' \subset \mathcal{T}$. This is, e.g., the case for $\mathcal{T}$ denoting urban development, and $\mathcal{T}'$ then signifying road construction. Then, we evaluate on a *query set*, where we are now only interested in finding change of the new, restricted type $\mathcal{T}'$.

This is different from *semantic* or *multiclass change detection* (Daudt et al., 2019), (Mou et al., 2019), where we have $K > 1$ different categories of change and the aim is to find $\boldsymbol{C}_{I_1,I_2} \in \{0, 1, \ldots, K\}^{H \times W}$. In our setting, we still are interested only in binary classification, albeit restricted to one particular category $\mathcal{T}$. We could also consider our formulation as a variant of *weakly supervised* classification, with the broad labels of the train set acting as weak labels for the more specialized labels of the support and query sets. However, this view does not adequately highlight the few-shot nature of the task, where the support set in practice is only available after the training and consists just of a small amount of change instances.

Figure 1 illustrates this problem, where the images and the full change map are from the OSCD data set (Daudt et al., 2018), which is well known in remote sensing and uses freely accessible Sentinel-2 observations, whereas the filtered ground truths and few shot masks have been produced by hand by us.

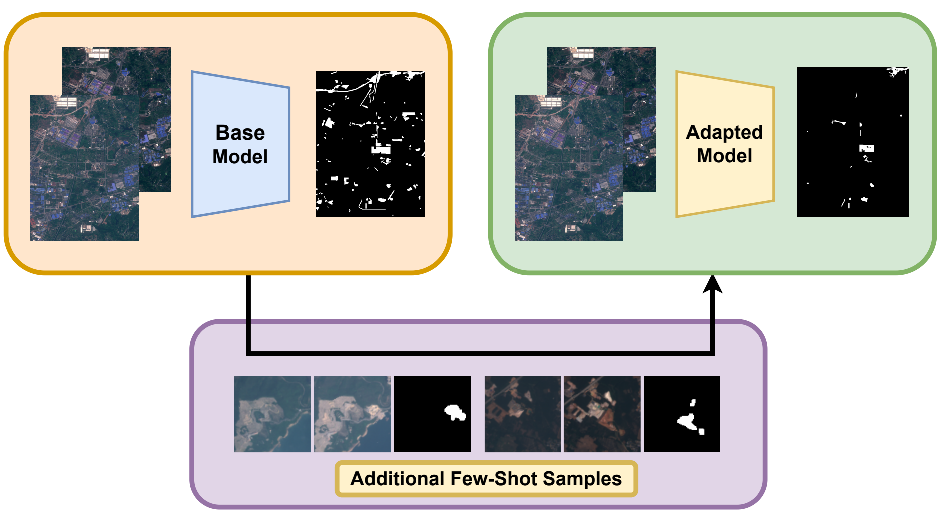

Figure 1: Illustration of the few-shot filtering process: The base model on the left has been trained on the full data and detects all change between the two satellite images. After we provide it with two additional observations of deforestation (from other areas), the adapted model on the left should now be able to *filter out* only deforestation instances.

REFERENCES

P. P. De Bem, O. A. de Carvalho Junior, R. Fontes Guimarães, and R. A. Trancoso Gomes. Change detection of deforestation in the brazilian amazon using landsat data and convolutional neural networks. *Remote Sensing*, 12(6):901, 2020.

W. Y. Chen, Y. C. Liu, Z. Kira, Y. C. F. Wang, and J. B. Huang. A closer look at few-shot classification. In *International Conference on Learning Representations*, 2018.

R.C. Daudt, B. Le Saux, A. Boulch, and Y. Gousseau. Urban change detection for multispectral earth observation using convolutional neural networks. In *International Geoscience and Remote Sensing Symposium*, pp. 2115–2118. IEEE, 2018.

R.C. Daudt, B. Le Saux, A. Boulch, and Y. Gousseau. Multitask learning for large-scale semantic change detection. *Computer Vision and Image Understanding*, 187:102783, 2019.

S. Gidaris and N. Komodakis. Dynamic few-shot visual learning without forgetting. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4367–4375, 2018.

A. Kanezaki. Unsupervised image segmentation by backpropagation. In *International Conference on Acoustics, Speech and Signal Processing*, pp. 1543–1547. IEEE, 2018.

L. Mou, L. Bruzzone, and X. Zhu. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2):924–935, 2019.

B. Peng, Z. Meng, Q. Huang, and C. Wang. Patch similarity convolutional neural network for urban flood extent mapping using bi-temporal satellite multispectral imagery. *Remote Sensing*, 11(21): 2492, 2019.

H. Qi, M. Brown, and D. G. Lowe. Low-shot learning with imprinted weights. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5822–5830, 2018.

S. Saha, F. Bovolo, and L. Bruzzone. Building change detection in vhr sar images via unsupervised deep transcoding. *IEEE Transactions on Geoscience and Remote Sensing*, 59(3):1917–1929, 2020.

Q. Shi, M. Liu, S. Li, X. Liu, F. Wang, and L. Zhang. A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

W. Shi, M. Zhang, R. Zhang, S. Chen, and Z. Zhan. Change detection based on artificial intelligence: State-of-the-art and challenges. *Remote Sensing*, 12(10):1688, 2020.

Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys (CSUR)*, 53(3):1–34, 2020.