

DEEP FEW-SHOT LEARNING FOR BI-TEMPORAL BUILDING CHANGE DETECTION

M. Khoshboresh-Masouleh*, R. Shah-Hosseini

School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran –
 (m.khoshboresh, rshahosseini)@ut.ac.ir

KEYWORDS: Deep few-shot learning, Meta-learning, Change Detection, Building Extraction

ABSTRACT:

In real-world applications (e.g., change detection), annotating images is very expensive. To build effective deep learning models in these applications, deep few-shot learning methods have been developed and prove to be a robust approach in small training data. The study of building change detection from high spatial resolution satellite observations is important to research in remote sensing, photogrammetry, and computer vision nowadays, which can be widely used in a variety of real-world applications, such as map generation and updating. As manual high-resolution image interpretation is expensive and time-consuming, building change detection methods are of high interest. The interest in developing building change detection approaches from optical remote sensing images is rapidly increasing due to larger coverages, and lower costs of optical images. In this study, we focus on building change detection analysis on a small set of building changes from different regions that sit in several cities. In this paper, a new deep few-shot learning method is proposed for building change detection using Monte Carlo dropout and remote sensing observations. The setup is based on a small dataset, including bitemporal optical images labelled for building change detection.

1. INTRODUCTION

Current urban and rural space growth rates and obvious effects of building construction on different applications, such as building damage assessments, building change detection, population estimation, and urban planning continuous monitoring of building footprints are becoming ever more significant (Hardin et al., 2007; Hegazy and Kaloop, 2015; Khoshboresh-Masouleh et al., 2020; Masouleh and Shah-Hosseini, 2018). The interest in developing building change detection methods from bi-temporal optical images with a different platform (e.g., UAV, Airborne, and Spaceborne) is rapidly increasing due to larger land-covers, and lower costs of remote sensing data (Bayanlou and Khoshboresh-Masouleh, 2020; Khoshboresh Masouleh and Saradjian, 2019; Khoshboresh-Masouleh and Hasanlou, 2020). Despite rapid advances in computer vision and remote sensing, change detection is a major challenge in mapping. However, building change detection from an optical remote sensing image with high accuracy and precision needs considerable effort in developing robust methods. Fig. 1 shows different challenges of building change detection from optical images such as small change (cf. Fig. 1a), complex roofs (cf. Fig. 1b), multiscale change (cf. Fig. 1c), and no-change (cf. Fig. 1d).

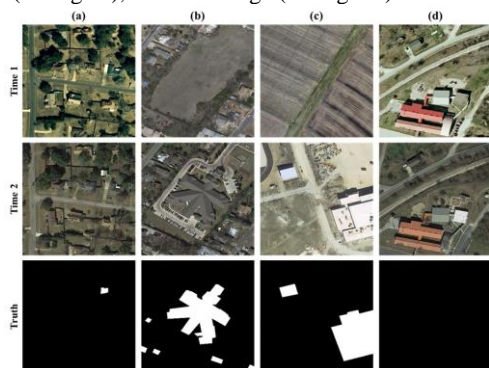


Figure 1. Example bi-temporal data from the LEVIR-CD dataset for building change detection

Owing to the advances of deep learning (LeCun et al., 2015) and the new datasets released in computer vision, remote sensing, and photogrammetry (Khoshboresh Masouleh and Saradjian, 2019), new insights have been presented in the field of building change detection for bi-temporal data (Chen and Shi, 2020; Hou et al., 2020; Khoshboresh Masouleh and Shah-Hosseini, 2019). Although some efforts have been devoted to the development of deep learning approaches, little attention has been devoted to deep few-shot learning for change detection. Table 1 presents the overview of the methods recently published in applied convolutional neural networks for building change detection from bi-temporal images, based on highlighting characteristics and precision with a focus on research contributions. Although the related approaches are robust, it's not still a good performance for building change detection, particularly, in robust building change detection from a small training dataset.

Method	Precision	Reference
Double CNN	50%	(Nemoto et al., 2017)
FC-EF	62%	(Daudt et al., 2018)
FC-Siam-conc	52%	(Daudt et al., 2018)
FC-Siam-diff	58%	(Daudt et al., 2018)
Multiscale CNN	63%	(Shunping Ji et al., 2019)

Table 1. Overview of recent building change detection methods using bi-temporal RGB images

Building change detection approaches for bi-temporal data need to be analyzed in different regions with uncertainty modelling. Therefore, despite the previous approaches that try to extract building change without uncertainty map from small regions with deep learning models and large-scale datasets, in this paper, we focus on building change detection analysis on a small set of building change from different areas that sit in several cities. In this study, to build generalized and effective building change detection models in remote sensing, deep few-shot learning methods have been developed and prove to be a robust method in small training data.

* Corresponding author

2. FEW-SHOT CNN

In real-world applications such as building change detection, annotating images is very expensive (Chen et al., 2020; Lin et al., 2020). This issue becomes more important in the tasks of pixel-wise classification, where dense labels are harder to annotate. The goal of few-shot learning in building change detection is to predict a binary map of a change class given a few pairs of support and query images containing the same change class and the binary truth maps for the support images (Masouleh and Sadeghian, 2019). Existing few-shot approaches generally learn from a handful of support images and then feed learned information into a parametric module for segmenting the query (Wang et al., 2019).

To build generalized deep learning models in building change detection, deep few-shot learning models have been developed and prove to be a robust tool in small labelled data. One efficient method is to fine-tune the pre-trained network. In this study, we use the MSCnet (Khoshboresh-Masouleh et al., 2020) as a specialized backbone to take advantage of its strong capacity for building detection. Unlike the previous methods that use the general CNN series as the backbone, we use the specialized model for building detection as the backbone.

Fig. 2 illustrates an overview of the Few-Shot CNN for building change detection. There are two same encoder block tied convolutional blocks for extracting features from t_1 and t_2 images. The proposed method learns the representation from the image pair and then tries to find the relationship between them. Using the residual blocks in convolutional neural networks, despite the improvement of accuracy, significantly increases the cost of calculations. The increased cost of calculations significantly affects the integration of the residual blocks with the convolutional blocks. Assuming that the convolutional block includes two convolutional filters, the input value passes through two filters and, then, is added to the initial value. As the theory suggests, the processed value is added to the initial value merely for preventing the reduction of the features created by the convolutional filters. This process makes it difficult to employ these layers in encoding-decoding networks. To promote the use of residual blocks in encoding-decoding networks, a new method called convolutional residual blocks based on the depth dropout method was used in this study (Khoshboresh-Masouleh and Shah-Hosseini, 2020).

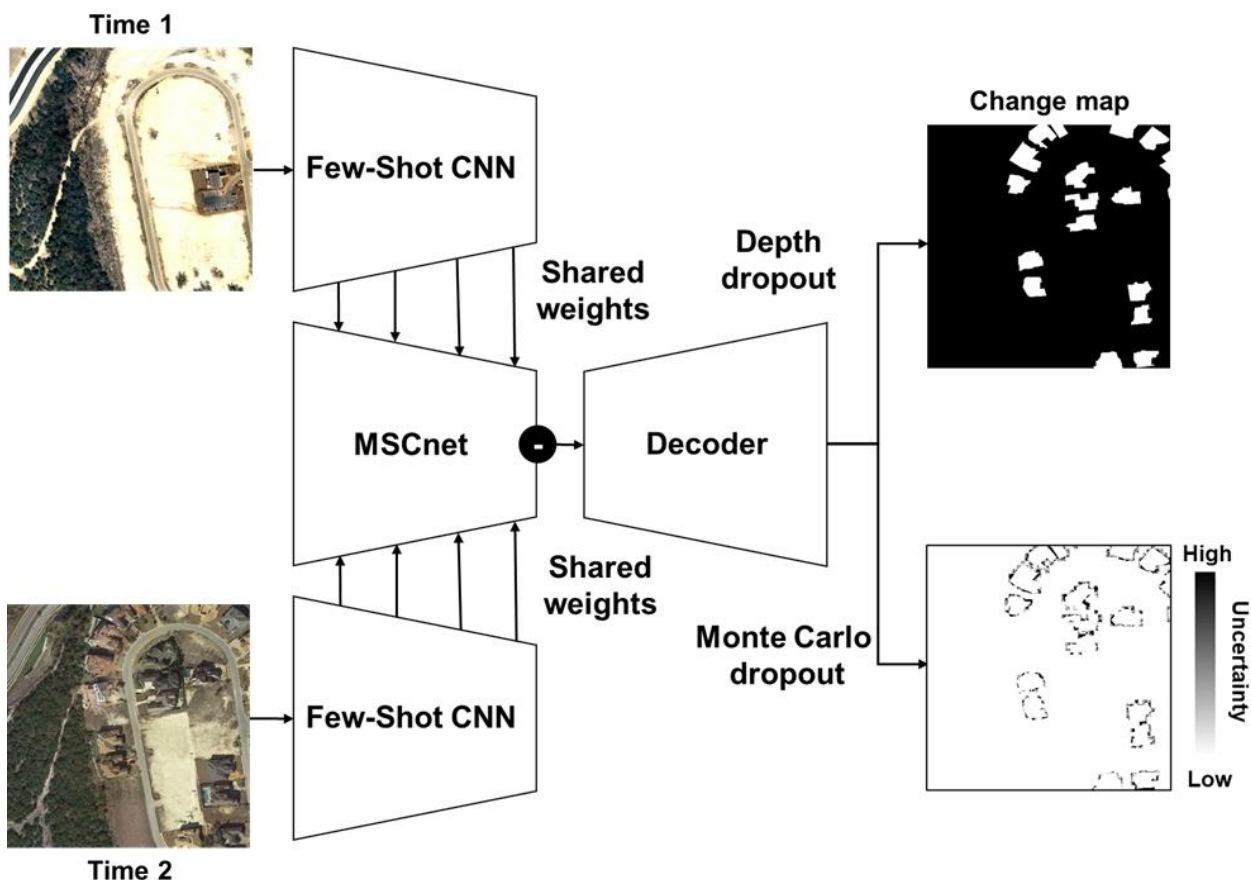


Figure 2. Few-Shot CNN. This example shows the performance of the Few-Shot CNN on the LEVIR-CD. The network takes bi-temporal RGB images as input and outputs a building change map

3. UNCERTAINTY MODELLING

Uncertainty modelling is a crucial step in the evaluation of the robustness for deep and few-shot learning models in remote sensing, especially when applied in risk-sensitive areas, such as building change detection (Cipolla et al., 2018; Khoshboresh Masouleh and Saradjian, 2019; Li and Alström, 2020). Knowing the confidence with which we can trust the building

change detection is important for decision-making in remote sensing (Khoshboresh Masouleh and Shah-Hosseini, 2020). There are some approaches to uncertainty modelling for deep learning proposed, but most of them need to sample several times, which is destructive to bi-temporal applications (Huang et al., 2018). In this paper, we focus on Monte Carlo dropout (Gal and Ghahramani, 2016) for uncertainty modelling in building change detection.

In this study, we use the Monte Carlo dropout as the aleatoric and epistemic uncertainty estimator for building change detection. In image segmentation, aleatoric is an uncertainty metric of the intrinsic, irreducible noise found in the image, usually associated with the image acquisition process. Epistemic is uncertainty over the actual values of a model's parameters arising from the finite size of the training datasets (Hüllermeier and Waegeman, 2020). As depicted in Fig. 2, total uncertainty was obtained by the Monte Carlo dropout.

4. RESULTS AND DISCUSSION

To evaluate the Few-Shot CNN, forty test images were selected from study areas. The features of these test images include small change, complex roofs, multiscale change, and no-change. To improve the training process of the MSCnet model, pretrained weights have been used in (Khoshboresh-Masouleh et al., 2020). Pretrained weights have been trained based on MSCnet model training with 2868 patches of very high-resolution images for building detection from different optical sensors. Table 2 shows the splitting statistics of the dataset. Few-Shot CNN was trained with ADAM optimizer (Kingma and Ba, 2017) using the default parameters and with a batch size of 64 for 250 epochs for building change detection. The generated dataset includes 190 bi-temporal RGB images that are

from the LEVIR-CD (Chen and Shi, 2020), with a size of 512×512 pixels and a spatial resolution from 0.5 m that cover different cities in the United States.

As shown in Fig. 3, four samples are selected from the test area for assessing the performance of the proposed method. The evaluation metric of intersection over union (IoU) (for change map) and entropy (for uncertainty map) are used to evaluate the performance of the proposed method. The average IoU and entropy for all scenes are about 92.4% and 0.12, respectively. The results using the LEVIR-CD dataset demonstrate the more reasonable accuracy achievement of the Few-Shot CNN to building change detection with small training data.

Dataset	Reference	Training	Test
IND	(Khoshboresh Masouleh and Saradjian, 2019)	300*	-
WHU-I	(S. Ji et al., 2019)	204*	-
Inria	(Maggiori et al., 2017)	200*	-
LEVIR-CD	(Chen and Shi, 2020)	150	40

Table 2. Dataset splitting statistics. * Denotes used only for the pretrained step

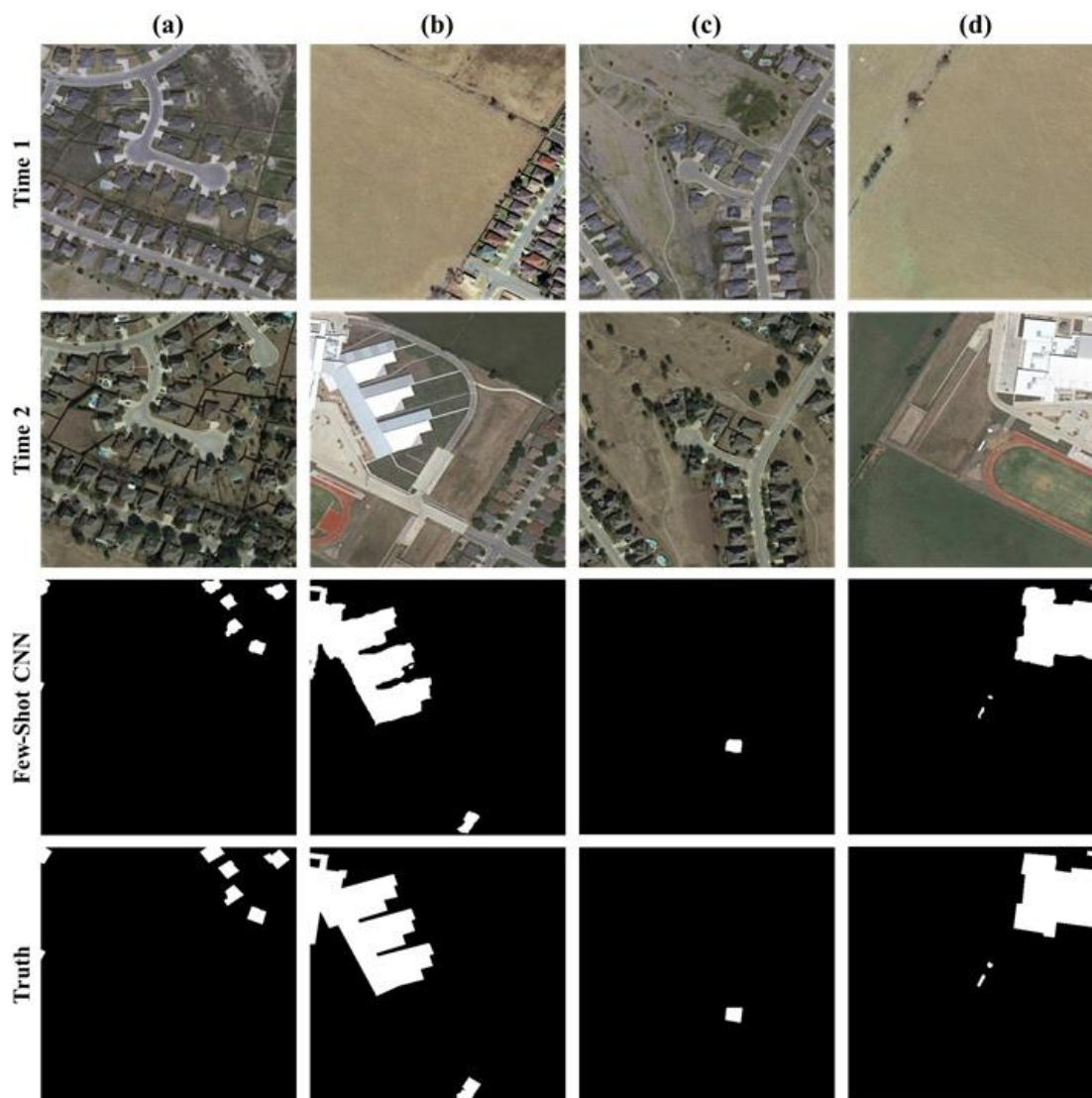


Figure 3. Building change detection results of the LEVIR-CD in four samples

5. CONCLUSIONS

In this study, using the fusion of Few-Shot CNN, Depth dropout, and Monte Carlo dropout, a novel method was presented and implemented for building change detection from bi-temporal RGB images. Then, the proposed network was tested on four remote sensing data sets with high spatial resolution and different challenges in mapping. The purpose of this study is to investigate the capabilities of this algorithm in the field of change detection. This is important because the access of some countries to high spatial resolution remote sensing images is far better and easier than that of other countries, and this approach can make the automatic mapping process possible with a relatively inexpensive remote sensing data source, with full coverage of different regions.

REFERENCES

- Bayanlou, M.R.R., Khoshboresh-Masouleh, M., 2020. SAMA-VTOL: A new unmanned aircraft system for remotely sensed data collection, in: SPIE Future Sensing Technologies. Presented at the SPIE Future Sensing Technologies, International Society for Optics and Photonics, p. 115250V. <https://doi.org/10.1117/12.2580533>
- Chen, H., Shi, Z., 2020. A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection. *Remote Sensing* 12, 1662. <https://doi.org/10.3390/rs12101662>
- Chen, Y., Zeng, X., Chen, X., Guo, W., 2020. A survey on automatic image annotation. *Appl Intell* 50, 3412–3428. <https://doi.org/10.1007/s10489-020-01696-2>
- Cipolla, R., Gal, Y., Kendall, A., 2018. Multi-task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Presented at the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Salt Lake City, UT, USA, pp. 7482–7491. <https://doi.org/10.1109/CVPR.2018.00781>
- Daudt, R.C., Saux, B.L., Boulch, A., 2018. Fully Convolutional Siamese Networks for Change Detection. [arXiv:1810.08462 \[cs\]](https://arxiv.org/abs/1810.08462).
- Gal, Y., Ghahramani, Z., 2016. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, in: International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 1050–1059.
- Hardin, P.J., Jackson, M.W., Otterstrom, S.M., 2007. Mapping, Measuring, and Modeling Urban Growth, in: Jensen, R.R., Gatrell, J.D., McLean, D. (Eds.), *Geo-Spatial Technologies in Urban Environments: Policy, Practice, and Pixels*. Springer, Berlin, Heidelberg, pp. 141–176. https://doi.org/10.1007/978-3-540-69417-5_8
- Hegazy, I.R., Kaloop, M.R., 2015. Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt. *International Journal of Sustainable Built Environment* 4, 117–124. <https://doi.org/10.1016/j.ijse.2015.02.005>
- Hou, B., Liu, Q., Wang, H., Wang, Y., 2020. From W-Net to CDGAN: Bi-temporal Change Detection via Deep Learning Techniques. *IEEE Trans. Geosci. Remote Sensing* 58, 1790–1802. <https://doi.org/10.1109/TGRS.2019.2948659>
- Huang, P.-Y., Hsu, W.-T., Chiu, C.-Y., Wu, T.-F., Sun, M., 2018. Efficient Uncertainty Estimation for Semantic Segmentation in Videos. [arXiv:1807.11037 \[cs\]](https://arxiv.org/abs/1807.11037).
- Hüllermeier, E., Waegeman, W., 2020. Aleatoric and Epistemic Uncertainty in Machine Learning: An Introduction to Concepts and Methods. [arXiv:1910.09457 \[cs, stat\]](https://arxiv.org/abs/1910.09457).
- Ji, Shunping, Shen, Y., Lu, M., Zhang, Y., 2019. Building Instance Change Detection from Large-Scale Aerial Images using Convolutional Neural Networks and Simulated Samples. *Remote Sensing* 11, 1343. <https://doi.org/10.3390/rs11111343>
- Ji, S., Wei, S., Lu, M., 2019. Fully Convolutional Networks for Multisource Building Extraction From an Open Aerial and Satellite Imagery Data Set. *IEEE Transactions on Geoscience and Remote Sensing* 57, 574–586. <https://doi.org/10.1109/TGRS.2018.2858817>
- Khoshboresh Masouleh, M., Saradjian, M.R., 2019. ROBUST BUILDING FOOTPRINT EXTRACTION FROM BIG MULTI-SENSOR DATA USING DEEP COMPETITION NETWORK, in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Presented at the ISPRS WG IV/3, WG I/8 & WG II/4-
ISPRS International GeoSpatial Conference 2019, Joint Conferences of 5th Sensors and Models in Photogrammetry and Remote Sensing (SMPR) and 3rd Geospatial Information Research (GI Research) (Volume XLII-4/W18) - 12–14 October 2019, Karaj, Iran, Copernicus GmbH, pp. 615–621. <https://doi.org/10.5194/isprs-archives-XLII-4-W18-615-2019>
- Khoshboresh Masouleh, M., Shah-Hosseini, R., 2020. A hybrid deep learning-based model for automatic car extraction from high-resolution airborne imagery. *Appl Geomat* 12, 107–119. <https://doi.org/10.1007/s12518-019-00285-4>
- Khoshboresh Masouleh, M., Shah-Hosseini, R., 2019. Development and evaluation of a deep learning model for real-time ground vehicle semantic segmentation from UAV-based thermal infrared imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 155, 172–186. <https://doi.org/10.1016/j.isprsjprs.2019.07.009>
- Khoshboresh-Masouleh, M., Alidoost, F., Arefi, H., 2020. Multiscale building segmentation based on deep learning for remote sensing RGB images from different sensors. *JARS* 14, 034503. <https://doi.org/10.1117/1.JRS.14.034503>
- Khoshboresh-Masouleh, M., Hasanlou, M., 2020. Improving hyperspectral sub-pixel target detection in multiple target signatures using a revised replacement signal model. *European Journal of Remote Sensing* 53, 316–330. <https://doi.org/10.1080/22797254.2020.1850179>
- Khoshboresh-Masouleh, M., Shah-Hosseini, R., 2020. A Deep Learning Method for Near-Real-Time Cloud and Cloud Shadow Segmentation from Gaofen-1 Images [WWW Document]. *Computational Intelligence and Neuroscience*. <https://doi.org/10.1155/2020/8811630>
- Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization. [arXiv:1412.6980 \[cs\]](https://arxiv.org/abs/1412.6980).
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>
- Li, B., Alstrøm, T.S., 2020. On uncertainty estimation in active learning for image segmentation. [arXiv:2007.06364 \[cs\]](https://arxiv.org/abs/2007.06364).

- Lin, H., Upchurch, P., Bala, K., 2020. Block Annotation: Better Image Annotation for Semantic Segmentation with Sub-Image Decomposition. arXiv:2002.06626 [cs, eess].
- Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P., 2017. Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark, in: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Presented at the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3226–3229. <https://doi.org/10.1109/IGARSS.2017.8127684>
- Masouleh, M.K., Sadeghian, S., 2019. Deep learning-based method for reconstructing three-dimensional building cadastre models from aerial images. JARS 13, 024508. <https://doi.org/10.1117/1.JRS.13.024508>
- Masouleh, M.K., Shah-Hosseini, R., 2018. Fusion of deep learning with adaptive bilateral filter for building outline extraction from remote sensing imagery. JARS 12, 046018. <https://doi.org/10.1117/1.JRS.12.046018>
- Nemoto, K., Hamaguchi, R., Sato, M., Fujita, A., Imaizumi, T., Hikosaka, S., 2017. Building change detection via a combination of CNNs using only RGB aerial imageries, in: Remote Sensing Technologies and Applications in Urban Environments II. Presented at the Remote Sensing Technologies and Applications in Urban Environments II, International Society for Optics and Photonics, p. 104310J. <https://doi.org/10.1117/12.2277912>
- Wang, K., Liew, J.H., Zou, Y., Zhou, D., Feng, J., 2019. PANet: Few-Shot Image Semantic Segmentation With Prototype Alignment, in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Presented at the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 9196–9205. <https://doi.org/10.1109/ICCV.2019.00929>