Editorial

Editorial for Special Issue: "Remote Sensing based Building Extraction II"

Jiaojiao Tian [1,*], Qin Yan [2], Mohammad Awrangjeb [3], Beril Kallfelz (Sirmacek) [4] and Nusret Demir [5]

1 Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Wessling, Germany
2 Chinese Academy of Surveying and Mapping, Beijing 100830, China
3 Institute for Integrated and Intelligent Systems, Griffith University, Nathan, QLD 4111, Australia
4 Independent Scientist, 7553 Hengelo, The Netherlands
5 Department of Space Science and Technologies, Remote Sensing Division, Akdeniz University, 07058 Antalya, Turkey
* Correspondence: jiaojiao.tian@dlr.de

1. Introduction

Accurate building extraction from remotely sensed images is essential for topographic mapping, urban planning, disaster management, navigation, and many other applications [1]. The easily available very-high resolution 2D/3D dataset and the rapid development of image processing techniques, especially the convolutional neural networks (CNN) and deep learning techniques have further boosted the research on building-extraction-related topics. Especially in recent years, many research institutes and associations have provided open-source datasets and annotated training data to meet the demand for advanced artificial intelligence models, which brings new opportunities to develop advanced approaches for building extraction and monitoring. Hence, there are higher expectations of the efficiency, accuracy, and robustness of building extraction approaches. They should also meet the demand of processing large datasets at the city, national, and global levels. Moreover, challenges remain on transform learning and dealing with imperfect training data, as well as unexpected objects in urban scenes such as trees, clouds, and shadows.

As a follow-on Special Issue of "Remote Sensing based Building Extraction", this Special Issue "Remote Sensing based Building Extraction II" has further collected the cutting-edge approaches for automatic building segmentation [1–4], vectorization [5,6], and regularization [7], dense matching [8], 3D reconstruction [9–11], and road detection [12]. The proposed methods fall into two main categories depending on the use of the input data sources: 2D building extraction and 3D reconstruction/segmentation.

2. 2D Building Extraction

Deep learning (DL) shows remarkable performance in extracting buildings from high-resolution remote sensing images. How to improve the performance of DL methods, especially the perception of spatial information, is worth further study. Paper [2] proposed a building extraction network (B-FGC-Net) with a feature highlighting, global awareness, and cross-level information fusion to achieve improved profitability of accurate extraction and information integration for both small- and large-scale buildings. Focusing on the promotion of the robustness of the interactive segmentation, Shu et al. [1] propose one Progress Guidance Representation Net (PGR-Net) to utilize the distance of newly added clicks to the boundary of the previous segmentation mask as an indication of the interactive segmentation progress, and this information is employed with the previous segmentation mask and positive and negative clicks to form a progress guidance map. This progress guidance map is then fed into a CNN with the original RGB image.

Furthermore, they propose an iterative training strategy for the training of the network and adopt an adaptive zoom-in technique during the inference stage for further performance promotion. Farmland constitutes an important resource for human survival and development. With complex ground features and scattered distribution, building extraction from farmland remains a challenging topic. To this end, Paper [3] proposes an attention-enhanced U-Net for building extraction from farmland based on Google and WorldView-2 remote sensing images. First, a Resnet unit is adopted as the infrastructure of the U-Net network encoding part, then the spatial and channel attention mechanism module is introduced between the Resnet unit and the maximum pool and the multi-scale fusion module is added to improve the U-Net network. Second, the buildings are extracted from WorldView-2 and Google images through farmland boundary constraints. Third, boundary optimization and fusion processing are carried out to further refine the building extraction results. In order to investigate the photovoltaic potential of urban buildings, Paper [4] proposed a pseudo-label-guided self-supervised learning (PGSSL) semantic segmentation network structure to extract building information from high-resolution remote sensing images. The pseudo-label-guided learning method allows the feature results extracted by the pretext task to be more applicable to the target task and ultimately improves segmentation accuracy.

To further close the gap between airborne images and vector representation, Van den Broeck and Goedemé [5] propose a fully automated end-to-end workflow for large-scale roof-part polygon extraction from UHR orthoimagery (0.03 m GSD). Their workflow comprised three steps: (1) An multitask fully convolutional network (FCN) was utilized for the semantic segmentation of roof-part objects and edges; (2) A bottom–up clustering algorithm was used, given the predicted roof-part edges, to derive individual roof-part clusters, where the predicted roof-part object area distinguish roof from non-roof; and (3) The roof-part clusters were vectorized and simplified into polygons. The methodology is trained and tested on a challenging dataset comprising of UHR aerial RGB orthoimagery (0.03 m GSD) and LiDAR-derived digital elevation models (DEMs) (0.25 m GSD) of three Belgian urban areas (including the famous touristic city of Bruges). Li et al. [6] explore the idea of combining three deep learning models, each model performing specific tasks, for automated extraction of building footprint polygons from very high-resolution aerial imagery. Their approach uses the U-Net, Cascade R-CNN, and Cascade CNN models to obtain building segmentation maps, building bounding boxes, and building corners, respectively, thus allowing for the direct production of building maps in a vector format. A polygon construction strategy based on Delaunay triangulation is designed to integrate the outputs from the deep learning models effectively, as well as to generate high-quality vector data. To solve the problem of edge discontinuity and incompleteness generated by semantic edge detection, Xia et al. [7] propose a multitask learning Dense D-LinkNet (DDLNet), which adopts full-scale skip connections and edge guidance module to ensure the effective combination of low-level information and high-level information.

3. 3D Reconstruction/Segmentation

The use of 3D building models is essential and provides realistic data for spatial and environmental analysis for various applications such as creating digital, generating simulations to predict and prepare for future scenarios, and creating various urban analytical processes, especially those that consider environmental impact, which is a growing global concern. To obtain a precise 3D model with lower cost, dense stereo matching has been studied persistently in the field of computer vision, remote sensing, and photogrammetry. Along with the development of deep learning, the Guided Aggregation Network (GA-Net) achieves state-of-the-art performance via the proposed Semi-Global Guided Aggregation layers and reduces the use of costly 3D convolutional layers. To solve the problem of GA-Net requiring large GPU memory consumption, Xia et al. [8] propose an efficient end-to-end network GA-Net-Pyramid for dense matching a pyramid architecture to modify the model. Starting from a downsampled stereo input, the

disparity is estimated and continuously refined through the pyramid levels. Thus, the disparity search is only applied for a small size of stereo pair and then confined within a short residual range for minor correction, leading to highly reduced memory usage and runtime. Manual modelling of urban buildings is very time-consuming and costly. Due to the complexity of the dense urban regions, research oriented toward the automatic reconstruction of buildings is still an open topic. In the manuscript titled "Parameter-Free Half-Spaces Based 3D Building Reconstruction Using Ground and Segmented Building Points from Airborne LiDAR Data with 2D Outlines"[9], the authors propose a new half-spaces based algorithm for building reconstruction from airborne laser point clouds. In contrast to the related algorithms, which divide 2D outlines of buildings into smaller parts and then process them while taking only convex shapes into account, the proposed algorithm performs reconstruction without division, while also considering concave parts of the rooftops. The method works in two stages, where the input data is processed first to obtain the definition of the base model of each building and the corresponding half-spaces. In the second stage, a building shape is generated by performing 3D Boolean operations over the analysed half-spaces.

A major challenge of large-scale building reconstruction from airborne LiDAR point clouds is the reconstruction of missing vertical walls. Paper [10] provided a fully automatic building reconstruction approach to infer vertical walls based on the connection between planar segments of both roofs and walls. The reconstruction model is obtained by using an extended hypothesis-and-selection-based polygonal surface reconstruction framework. Experimental results demonstrated that the proposed method is superior to the state-of-the-art methods in terms of reconstruction accuracy and robustness. The study also generated a new dataset consisting of the point clouds and 3D models of 20k real-world buildings which can stimulate research in urban reconstruction and the use of 3D city models in urban applications. To further refine the extracted building boundaries, Hui et al. [11] propose a multi-constraints graph segmentation method for building extraction from airborne LiDAR data and achieve satisfactory results. The graph structure is generated using the three-dimensional spatial features of points. To reduce computational cost the point-based building extraction is transformed into an object-based building extraction and geometric morphological features are computed for each segmented object. Finally, a multi-scale progressively growing optimisation method is employed to recover the omitted building parts.

Besides buildings, digital maps of road networks are a vital part of digital cities and intelligent transportation. This study [12] provided a comprehensive review of road extraction based on various remote sensing data sources. It is divided into three parts. Part 1 provides an overview of the existing data acquisition techniques for road extraction, including data acquisition methods, typical sensors, application status, and prospects. Part 2 underlines the main road extraction methods based on four data sources. Part 3 presents the combined application of multisource data for road extraction. It can provide a comprehensive reference for research on existing road extraction technologies.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shu, Z.; Hu, X.; Dai, H.; Progress Guidance Representation for Robust Interactive Extraction of Buildings from Remotely Sensed Images. *Remote Sens.* **2021**, *13*, 5111. https://doi.org/10.3390/rs13245111.
2. Wang, Y.; Zeng, X.; Liao, X.; Zhuang, D. B-FGC-Net: A Building Extraction Network from High Resolution Remote Sensing Imagery. *Remote Sens.* **2022**, *14*, 269. https://doi.org/10.3390/rs14020269.
3. Li, C.; Fu, L.; Zhu, Q.; Zhu, J.; Fang, Z.; Xie, Y.; Guo, Y.; Gong, Y. Attention Enhanced U-Net for Building Extraction from Farmland Based on Google and WorldView-2 Remote Sensing Images. *Remote Sens.* **2021**, *13*, 4411. https://doi.org/10.3390/rs13214411.
4. Chen, D.-Y.; Peng, L.; Zhang, W.-Y.; Wang, Y.-D.; Yang, L.-N. Research on Self-Supervised Building Information Extraction with High-Resolution Remote Sensing Images for Photovoltaic Potential Evaluation. *Remote Sens.* **2022**, *14*, 5350. https://doi.org/10.3390/rs14215350.
5. Van den Broeck, W.A.J.; Goedemé, T. Combining Deep Semantic Edge and Object Segmentation for Large-Scale Roof-Part Polygon Extraction from Ultrahigh-Resolution Aerial Imagery. *Remote Sens.* **2022**, *14*, 4722. https://doi.org/10.3390/rs14194722.
6. Li, Z.; Xin, Q.; Sun, Y.; Cao, M. A Deep Learning-Based Framework for Automated Extraction of Building Footprint Polygons from Very High-Resolution Aerial Imagery. *Remote Sens.* **2021**, *13*, 3630. https://doi.org/10.3390/rs13183630.
7. Xia, L.; Zhang, J.; Zhang, X.; Yang, H.; Xu, M. Precise Extraction of Buildings from High-Resolution Remote-Sensing Images Based on Semantic Edges and Segmentation. *Remote Sens.* **2021**, *13*, 3083. https://doi.org/10.3390/rs13163083.
8. Xia, Y.; d'Angelo, P.; Fraundorfer, F.; Tian, J.; Fuentes Reyes, M.; Reinartz, P. GA-Net-Pyramid: An Efficient End-to-End Network for Dense Matching. *Remote Sens.* **2022**, *14*, 1942. https://doi.org/10.3390/rs14081942.
9. Bizjak, M.; Žalik, B.; Lukač, N. Parameter-Free Half-Spaces Based 3D Building Reconstruction Using Ground and Segmented Building Points from Airborne LiDAR Data with 2D Outlines. *Remote Sens.* **2021**, *13*, 4430. https://doi.org/10.3390/rs13214430.
10. Huang, J.; Stoter, J.; Peters, R.; Nan, L. City3D: Large-Scale Building Reconstruction from Airborne LiDAR Point Clouds. *Remote Sens.* **2022**, *14*, 2254. https://doi.org/10.3390/rs14092254.
11. Hui, Z.; Li, Z.; Cheng, P.; Ziggah, Y.Y.; Fan, J. Building Extraction from Airborne LiDAR Data Based on Multi-Constraints Graph Segmentation. *Remote Sens.* **2021**, *13*, 3766. https://doi.org/10.3390/rs13183766.
12. Jia, J.; Sun, H.; Jiang, C.; Karila, K.; Karjalainen, M.; Ahokas, E.; Khoramshahi, E.; Hu, P.; Chen, C.; Xue, T.; et al. Review on Active and Passive Remote Sensing Techniques for Road Extraction. *Remote Sens.* **2021**, *13*, 4235. https://doi.org/10.3390/rs13214235.