



Deep-Learning-Based Edge Detection for Improving Building Footprint Extraction from Satellite Images [†]

Nima Ahmadian ^{1,*}, Amin Sedaghat ¹, Nazila Mohammadi ¹ and Mohammad Aghdami-Nia ²

¹ Department of Geomatics Engineering, Faculty of Civil Engineering, University of Tabriz, Tabriz 5166616471, Iran; a.sedaghat@tabrizu.ac.ir (A.S.); n.mohammadi@tabrizu.ac.ir (N.M.)

² School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 1417935840, Iran; m.aghdaminia@gmail.com

* Correspondence: n.ahmadian1998@gmail.com

[†] Presented at the 5th International Electronic Conference on Remote Sensing, 7–21 November 2023; Available online: <https://ecrs2023.sciforum.net/>.

Abstract: Buildings are objects of great importance that need to be observed continuously. Satellite and aerial images provide valuable resources nowadays for building footprint extraction. Since these images cover large areas, manually detecting buildings will be a time-consuming task. Recent studies have proven the capability of deep learning algorithms in building footprint extraction automatically. But these algorithms need vast amounts of data for training and they may not perform well under the low-data conditions. Digital surface models provide height information, which helps discriminate buildings from their surrounding objects. However, they may suffer from noises, especially on the edges of buildings, which may result in low boundary resolution. In this research, we aim to address this problem by using edge bands detected by a deep learning model alongside the digital surface models to improve the building footprint extraction when training data are low. Since satellite images have complex backgrounds, using conventional edge detection methods like Canny or Sobel filter will produce a lot of noisy edges, which can deteriorate the model performance. For this purpose, first, we train a U-Net model for building edge detection with the WHU dataset and fine-tune the model with our target training dataset, which contains a low quantity of satellite images. Then, the building edges of the target test images are predicted using this fine-tuned U-Net and concatenated with our RGB-DSM test images to form 5-band RGB-DSM-Edge images. Finally, we train a U-Net with 5-band training images of our target dataset, which contain precise building edges in their fifth band. Then, we use this model for building footprint extraction from 5-band test images, which contain building edges in their fifth band that are predicted by a deep learning model in the first stage. We compared the results of our proposed method with 4-band RGB-DSM and 3-band RGB images. Our method obtained 82.88% in IoU and 90.45% in F1-score metrics, which indicates that, by using edge bands alongside the digital surface models, the performance of the model improved 2.57% and 1.59% in IoU and F1-score metrics, respectively. Also, the predictions made by 5-band images have sharper building boundaries than RGB-DSM images.

Keywords: deep learning; building footprint extraction; edge detection; U-Net; digital surface model



Citation: Ahmadian, N.; Sedaghat, A.; Mohammadi, N.; Aghdami-Nia, M. Deep-Learning-Based Edge Detection for Improving Building Footprint Extraction from Satellite Images. *Environ. Sci. Proc.* **2024**, *29*, 61. <https://doi.org/10.3390/ECRS2023-16615>

Academic Editor: Riccardo Buccolieri

Published: 20 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Automatic building footprint extraction from remote-sensing imagery has various applications in urban planning, 3D modelling and disaster management [1]. Due to the advanced technology in acquisition of high-resolution satellite images, there are valuable resources for building footprint extraction nowadays [2]. Satellite images cover a vast amount of areas and contain complex backgrounds and rich information [3]. Since manually extracting building footprints from satellite images is a laborious and challenging task, automatic approaches should be considered in this case [4].

With recent developments in data science and artificial intelligence, deep learning algorithms are used widely in remote sensing [5,6]. Deep learning algorithms are capable of extracting features from satellite images automatically and using this information to solve problems [7]. Recently, much research has focused on extracting building footprints from remote-sensing images with deep learning models [3]. Deep convolutional neural networks and fully convolutional networks are used frequently for this task [8].

In a study, Aryal et al. proposed two scale-robust fully convolutional networks by focusing on multi-scale feature utilization and domain-shift minimization [9]. Yu et al. proposed a convolutional neural network called ConvBNet, which uses deep supervision in training with weighted and mask cross-entropy losses to ensure stable convergence [10]. In another study, Ji et al. proposed Siamese U-Net to improve the classification of larger buildings [11]. Ma et al. proposed GMEDN by focusing on using global and local features and mining multi-scale information, which has a local and global encoder with a distilling decoder [12].

Also, in some studies, LiDAR point clouds or digital surface models (DSMs) are used alongside RGB images to improve the accuracy of the building footprint extraction task. In a study, Yu et al. proposed MA-FCN and used digital surface models with RGB images to extract buildings from aerial images, which resulted in better predictions [13].

Although many studies proposed deep learning models for building footprint extraction from remote-sensing imagery, most of these models need a considerable amount of training data, which may not be available all the time. Moreover, noisy DSMs may lead to noisy building edges. To address this problem, in this study, we propose using building edge bands detected by a deep learning model alongside RGB images and digital surface models to improve the results of building footprint extraction from satellite images that contain a small amount of training data.

2. Methodology

In this study, our goal is to improve the accuracy of building segmentation maps in low training data conditions by using building edge bands alongside RGB-DSM images. To address this problem, we use the U-Net [14] model to detect building edges from satellite images. In this section, first, we present a brief review of U-Net model. Then, the advantages of deep-learning-based building edge detection over traditional edge detection methods will be discussed.

2.1. U-Net

U-Net is a fully convolutional network that has an encoder–decoder structure with skip connections between them. In U-Net structure, convolutional blocks are used to extract features from input data, pooling layers are used in the encoder to pass the output of each convolutional block to the next block by reducing the dimensions of the output by half and up-convolution layers are used to increase the dimensions of the output by two and pass it to the next convolutional block in the decoder part. Also, skip connections used in the U-Net model help the model retrieve spatial information from early stages of the model and reduce information loss. The structure of the U-Net model is shown in Figure 1.

2.2. Edge Detection Methods

Satellite images are rich in information and have complex backgrounds. Since our aim is to improve the accuracy of building segmentation maps by using edge bands, these edge bands should only contain the edges of buildings. By using conventional edge detection methods like Canny or Sobel filter, there will be complex edges detected from satellite images that do not contain building edges exclusively.

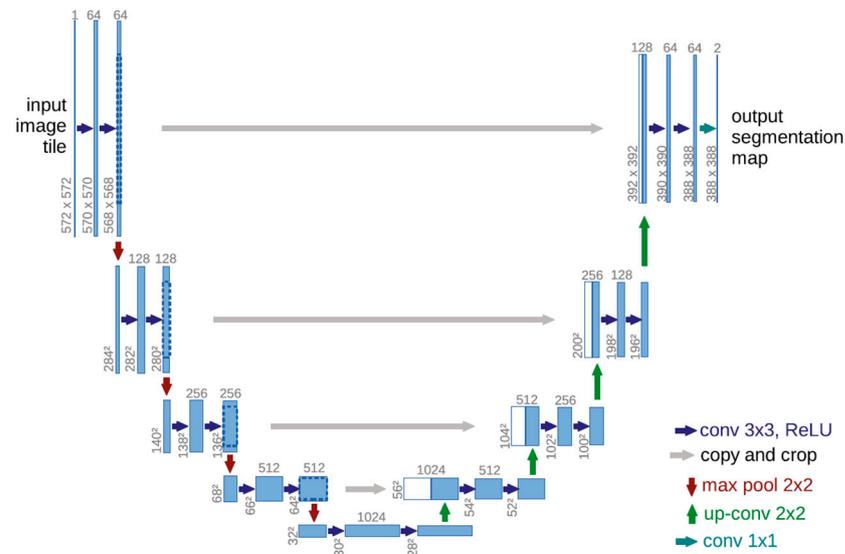


Figure 1. U-Net structure [14].

To address this problem, we use a deep learning model to detect building edges from remote-sensing images. For this purpose, we need a training dataset that contains images with corresponding binary building edge labels. First, we create binary building edge labels for the training dataset by applying Canny filter to the binary building masks of WHU and our target satellite datasets. Then, the U-Net model is trained with the WHU dataset and it is fine-tuned with our target satellite dataset. By using this strategy, we can use U-Net to exclusively detect building edges from test images, which do not contain the edges of other objects like trees or roads existing in the images.

In order to evaluate the impact of edge bands in building footprint extraction, we create 5-band RGB-DSM-Edge training and test images from the satellite data. The training images contain precise building edges in their fifth band since they are created from applying the Canny filter to the binary building masks of the training data. On the other hand, the fifth band of the test images contains building edges that are predicted by the U-Net model.

Finally, we train U-Net with RGB, RGB-DSM and RGB-DSM-Edge satellite images and compare their results with each other in order to evaluate the impact of using edge bands alongside RGB images and DSMs in the building footprint extraction task. Edge bands can help the model be aware of the building edges, which can lead to more complete segmentation maps with sharper boundaries for buildings.

3. Datasets

In this research, we used two datasets: WHU dataset and IEEE Data Fusion Contest 2019 dataset [15]. WHU dataset consists of aerial images with buildings of various shapes, sizes and colors. IEEE Data Fusion Contest 2019 dataset consists of satellite images and DSMs. For both of these datasets, binary building edge labels are created by applying a canny filter to the binary building footprint labels. An example of these datasets with building footprint and building edge labels is shown in Figure 2.

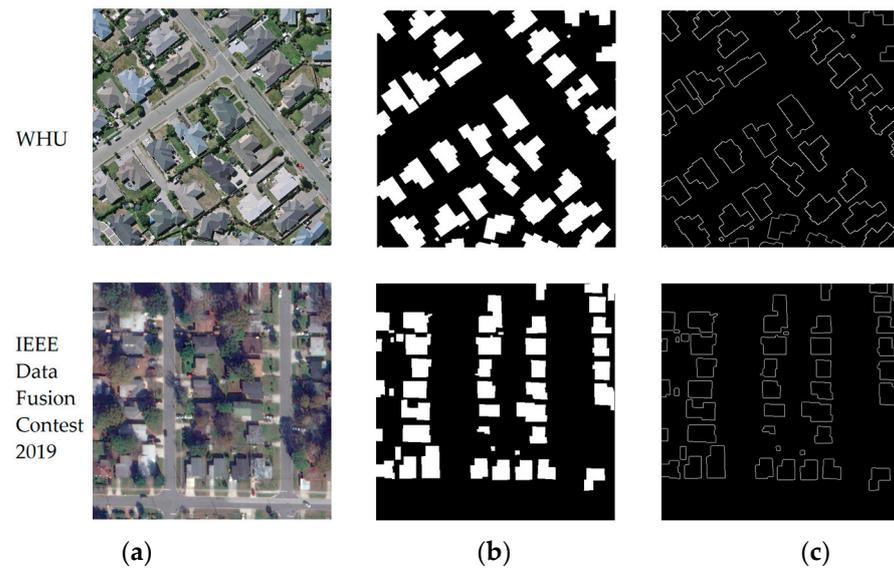


Figure 2. (a) Images, (b) binary building footprint masks and (c) binary building edge masks.

4. Results

In this section, the results of building edge detection and building footprint extraction with RGB-DSM-Edge images will be discussed. The results of building edge detection with U-Net are compared with Canny and HED [16] edge detection methods. Also, we used Mask-RCNN, MA-FCN and U-Net models for comparison by using RGB and RGB-DSM images in the building footprint extraction task. The quantitative results of the mentioned models are shown in Table 1.

Table 1. Results of mentioned models with RGB, RGB-DSM and RGB-DSM-Edge images.

Model	Bands	Accuracy	Precision	Recall	F1-Score	IoU
Mask-RCNN	RGB	96.22%	92.51%	85.69%	88.80%	79.91%
MA-FCN	RGB	95.83%	86.62%	86.38%	86.36%	76.60%
U-Net	RGB	96.16%	87.48%	90.20%	88.55%	79.78%
MA-FCN	RGB-DSM	96.13%	92.20%	83.87%	87.26%	78.07%
U-Net	RGB-DSM	96.40%	91.93%	86.24%	88.86%	80.31%
U-Net	RGB-DSM-Edge	96.73%	89.55%	91.66%	90.45%	82.88%

As shown in Figure 3, Canny and HED edge detection results contain a lot of noise, which cannot be used for building footprint extraction improvement with deep learning models. On the other hand, the edge detection result of the U-Net, which was trained with building edge labels, produced building edges that can be used alongside RGB-DSM images to improve the results of the building footprint extraction task.

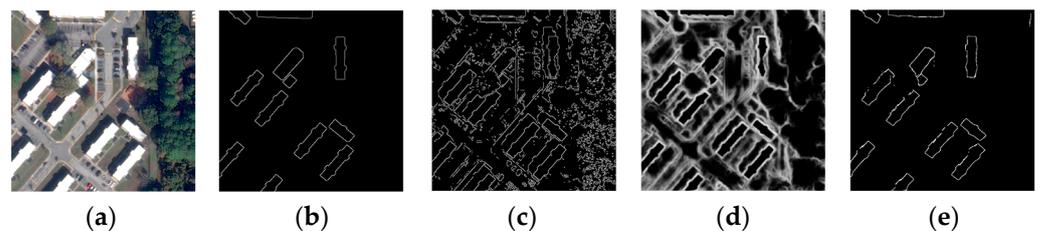


Figure 3. Building edge detection results: (a) satellite image, (b) ground truth building edge label, (c) Canny edge detection, (d) HED edge detection and (e) U-Net edge detection.

As shown in Table 1, RGB-DSM images improved the results of U-Net and MA-FCN models compared to the RGB images. Moreover, using edge bands alongside RGB images and DSMs improved the results and outperformed all other models in F1-score and IoU metrics. Since Mask-RCNN uses Region Proposal Network, ROI Align and ResNet architecture, it performs better in low data conditions, which helps it to detect more true positives and leads to better results in RGB images and the precision metric. Although edge bands helped the U-Net model to perform better than other cases, since edge detection U-Net was trained with a low quantity of training data, edge bands detected by U-Net are not that accurate, which may lead to lower true positives and lower precision. Our proposed method achieved 90.45% and 82.88% in F1-score and IoU metrics, respectively, which indicates the better quality of segmentation maps created by the U-Net model using RGB-DSM-Edge images. These results indicate the effectiveness of using edge bands alongside RGB-DSM images in datasets with a low quantity of training images.

In Figure 4, predictions made by the mentioned models are shown with test images and ground-truth binary labels. It is clear that DSMs improved the quality of segmentation maps produced by MA-FCN and U-Net models compared to RGB images but there is still room for improvement, especially in the number and boundaries of detected buildings. Edge bands addressed these problems effectively, since the segmentation maps produced by RGB-DSM-Edge images are more complete, especially in the third image. Also, detected buildings by RGB-DSM-Edge images have sharper building boundaries, which indicates the effectiveness of using edge bands in producing sharper building boundaries.

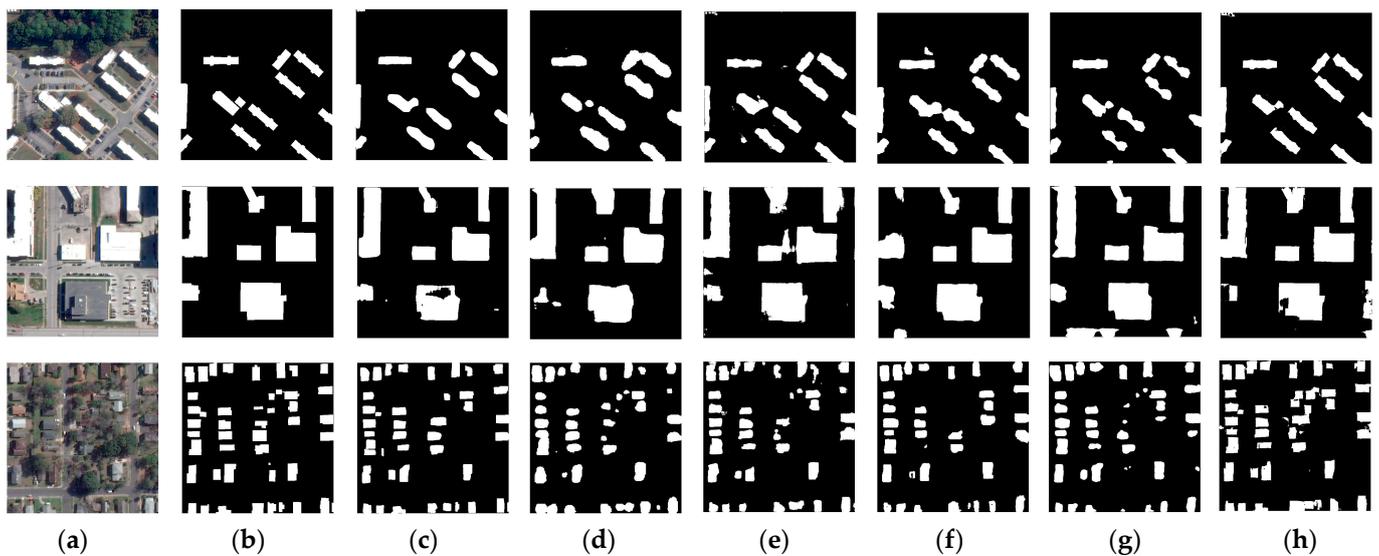


Figure 4. IEEE Data Fusion Contest 2019 Dataset: (a) test images, (b) ground truth labels, (c) Mask-RCNN RGB predictions, (d) MA-FCN RGB predictions, (e) U-Net RGB predictions, (f) MA-FCN RGB-DSM predictions, (g) U-Net RGB-DSM predictions and (h) U-Net RGB-DSM-Edge predictions.

5. Conclusions

In this study, we aimed to improve the results of building footprint extraction from satellite images with deep learning models both quantitatively and qualitatively. Since satellite images have complex backgrounds with various objects, traditional and state-of-the-art edge detection methods are not capable of detecting building edges exclusively. For this purpose, we proposed preparing building edge labels to train a U-Net model for the building edge detection task. Then, these edge bands were attached to RGB-DSM images to create RGB-DSM-Edge images, which were used for building footprint extraction with U-Net. We compared the results of our proposed method with other deep learning models with RGB and RGB-DSM images. The U-Net model trained with RGB-DSM-Edge images outperformed Mask-RCNN with RGB images and also U-Net and MA-FCN models with

both RGB and RGB-DSM images. Our proposed method reached 90.45% and 82.88% in F1-score and IoU metrics, respectively. Also, the segmentation maps produced by RGB-DSM-Edge images contain more complete detected buildings with sharper boundaries.

Author Contributions: Conceptualization, N.A.; methodology, N.A., A.S. and N.M.; validation, N.A. and A.S.; writing-original draft preparation, N.A. and M.A.-N.; writing-review and editing, N.A., A.S. and N.M.; visualization, N.A. and M.A.-N.; supervision and project administration, A.S. and N.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The images can be accessed from the following link: <https://iee-dataport.org/open-access/data-fusion-contest-2019-dfc2019> (accessed on 29 February 2024).

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

References

1. Sakeena, M.; Stumpe, E.; Despotovic, M.; Koch, D.; Zeppelzauer, M. On the Robustness and Generalization Ability of Building Footprint Extraction on the Example of SegNet and Mask R-CNN. *Remote Sens.* **2023**, *15*, 2135. [CrossRef]
2. Hosseinpour, H.; Samadzadegan, F.; Javan, F.D. CMGFNet: A deep cross-modal gated fusion network for building extraction from very high-resolution remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2022**, *184*, 96–115. [CrossRef]
3. Luo, L.; Li, P.; Yan, X. Deep Learning-Based Building Extraction from Remote Sensing Images: A Comprehensive Review. *Energies* **2021**, *14*, 7982. [CrossRef]
4. 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. [CrossRef]
5. Abbasi, M.; Shah-Hosseini, R.; Aghdami-Nia, M. Sentinel-1 Polarization Comparison for Flood Segmentation Using Deep Learning. *Proceedings* **2023**, *87*, 14. [CrossRef]
6. Jovhari, N.; Farhadi, N.; Sedaghat, A.; Mohammadi, N. Performance Evaluation of Learning-Based Methods for Multispectral Satellite Image Matching. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *X-4/W1-2022*, 335–341. [CrossRef]
7. Aghdami-Nia, M.; Shah-Hosseini, R.; Salmani, M. Effect of Transferring Pre-Trained Weights on a Siamese Change Detection Network. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**; *X-4/W1-2022*, 19–24. [CrossRef]
8. Shao, Z.; Tang, P.; Wang, Z.; Saleem, N.; Yam, S.; Sommai, C. BRRNet: A Fully Convolutional Neural Network for Automatic Building Extraction from High-Resolution Remote Sensing Images. *Remote Sens.* **2020**, *12*, 1050. [CrossRef]
9. Aryal, J.; Neupane, B. Multi-Scale Feature Map Aggregation and Supervised Domain Adaptation of Fully Convolutional Networks for Urban Building Footprint Extraction. *Remote Sens.* **2023**, *15*, 488. [CrossRef]
10. Yu, T.; Tang, P.; Zhao, B.; Bai, S.; Gou, P.; Liao, J.; Jin, C. ConvBNet: A Convolutional Network for Building Footprint Extraction. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 1–5. [CrossRef]
11. Ji, S.; Wei, S.; Lu, M. Fully Convolutional Networks for Multisource Building Extraction from an Open Aerial and Satellite Imagery Data Set. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 574–586. [CrossRef]
12. Ma, J.; Wu, L.; Tang, X.; Liu, F.; Zhang, X.; Jiao, L. Building Extraction of Aerial Images by a Global and Multi-Scale Encoder-Decoder Network. *Remote Sens.* **2020**, *12*, 2350. [CrossRef]
13. Yu, D.; Ji, S.; Liu, J.; Wei, S. Automatic 3D building reconstruction from multi-view aerial images with deep learning. *ISPRS J. Photogramm. Remote Sens.* **2021**, *171*, 155–170. [CrossRef]
14. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.
15. Le Saux, B.; Yokoya, N.; Hänsch, R.; Brown, M. *Data Fusion Contest 2019 (DFC2019)*; IEEE Dataport; IEEE: Piscataway, NJ, USA, 2019. [CrossRef]
16. Xie, S.; Tu, Z. Holistically-nested edge detection. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.