International Cartographic Association
Association Cartographique Internationale

Building footprint extraction using National Agriculture Imagery Program imagery and 3D Elevation Program lidar data: a deep learning approach

Jung kuan Liu [a, *], Rongjun Qin [b], Samantha Arundel [a]

[a] *U.S. Geological Survey, Center of Excellence for Geospatial Information Science, Rolla, Missouri, United States, jliu@usgs.gov, Samantha Arundel, sarundel@usgs.gov*

[b] *Department of Civil, Environmental and Geodetic Engineering, The Ohio State University, Columbus, Ohio, United States qin324@osu.edu*

* Corresponding author

Keywords: building footprints, NAIP, 3DEP, lidar, deep learning

Abstract:

Accurate building footprint extraction is vital for numerous geospatial applications, including population estimation, urban planning, infrastructure development, disaster management, and other fields (Che et al. 2024; Vincent and Varalakshmi 2024). Building footprints and building height can be extracted from high-resolution optical imagery based on the texture and shadow details they provide (Liasis and Stavrou 2016; Cao and Huang 2023; Vincent and Varalakshmi 2024). However, the quality of derived building footprints and heights is lower in densely built areas where shadows overlap. In recent years, light detection and ranging (lidar) has been shown to be reliable data for extracting building footprints and estimating building height (Park and Guldmann 2019; Rottmann et al. 2022; Karsli et al. 2024). This approach also comes with challenges in terms of geometric irregularities, noisy points, point density, occlusion in dense urban areas, and accuracy (Karsli et al. 2024). A possible solution is integrating optical imagery and lidar data to fill data gaps and improve feature visibility.

The most common approaches to building footprint extraction can be classified into three categories: traditional manual interpretation or rule-based, machine learning (ML), and deep learning (DL) methods. Although the accuracy is expected to be high for traditional manual interpretation compared to other approaches, it is labour-intensive and time-consuming. Furthermore, semi-automated, traditional rule-based methods or classical image processing techniques struggle in complex environments such as dense urban areas or rural regions with occlusion from trees, clouds, or adjacent structures. The variability in building sizes, shapes, and appearances in different geographic contexts (Vincent and Varalakshmi 2024) often hinders these traditional manual interpretation or rule-based methods. Schlosser et al. (2020) used ML techniques such as random forests and support vector machines to classify building pixels from high-resolution imagery. These methods rely heavily on feature engineering, limiting them in capturing complex patterns across diverse geographies (Karsli et al. 2024). DL models such as convolutional neural networks (CNNs) are essential for building footprint extraction (Luo et al. 2021) because these can automatically learn complex spatial patterns and features from multi-source data, overcoming the limitations of traditional rule-based and ML approaches. By integrating spectral and elevation information, deep learning models improve accuracy, adaptability, and scalability, enabling robust extraction even in challenging environments with occlusions, shadows, and diverse architectural styles (Dabove et al. 2024).

In this study, we propose a novel approach for automated extraction of building footprints by combining U.S. Department of Agriculture (USDA) National Agricultural Imagery Program (NAIP) aerial imagery with U.S. Geological Survey (USGS) 3D Elevation Program (3DEP) lidar data, leveraging the power of deep learning. Our approach integrates multispectral NAIP imagery and 3DEP lidar elevation data to enhance the accuracy of building footprint extraction. NAIP provides high-resolution aerial imagery, while 3DEP offers precise elevation data, differentiating buildings from spectrally similar features like trees. The complementary nature of these datasets allows for improved structural identification by combining spectral and height-based characteristics.

To achieve this, we use a deep CNN that simultaneously processes NAIP imagery and lidar-derived digital elevation models (DEMs), ensuring robust building footprint detection. To refine building height estimation, we reclassified 3DEP lidar data using a custom-developed toolkit (Liu et al. 2024). The datasets undergo co-registration and resolution alignment before model training, ensuring consistency. Additionally, lidar data are processed to generate DEM and normalized digital surface model (nDSM) representations, facilitating object height differentiation. Given the high cost of manually labeled training data, we utilize SpatialAPI 20 shared data (Shi et al. 2020) from the 2nd ACM SIGSPATIAL International Workshop on Geospatial Data Access and Processing APIs to validate our approach.

https://doi.org/10.5194/ica-abs-10-177-2025

We hypothesize that integrating NAIP imagery and 3DEP lidar data will improve building footprint extraction accuracy compared to models relying on a single data source. To test this, we compare our extracted footprints with Microsoft's Global Building Footprints (GBF) dataset, which contains over 1.4 billion footprints derived from high-resolution satellite imagery using DL techniques. Initial results indicate that our method achieves higher precision and Intersection-over-Union (IoU) scores, particularly in areas with dense vegetation where spectral information alone is insufficient. Further evaluation includes benchmarking against state-of-the-art models based on either high-resolution imagery or lidar data alone, as well as traditional ML techniques. By assessing performance through precision, recall, F1 score, and IoU, we aim to quantify the benefits of incorporating both spectral and elevation data for improved footprint extraction.

This research presents a novel approach to building footprint extraction using a DL framework that integrates multispectral NAIP imagery and 3DEP lidar data. The dual-input CNN architecture allows for the effective fusion of spectral and elevation features, with the potential for superior performance compared to existing methods. The model has demonstrated robustness across diverse landscapes, enabling large-scale building footprint extraction essential for urban planning, disaster management, and geospatial analysis (Alaoui et al. 2022).

Disclaimer: Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Alaoui, H.M., Radoine, H., Chenal, J., Hajji, H. and Yakubu, H., 2022. Deep building footprint extraction for urban risk assessment – remote sensing and deep learning based approach. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-4/W3-2022, 83–86, doi.org/10.5194/isprs-archives-XLVIII-4-W3-2022-83-2022.

Cao, Y. and Huang, X., 2021. A deep learning method for building height estimation using high-resolution multi-view imagery over urban area: a case study of 42 Chinese cities. *Remote Sensing of Environment*, 264, 112590, doi.org/10.1016/j.rse.2021.112590.

Che, Y., Li, X., Liu, X., Wang, Y., Liao, W., et al., 2024. 3D-GloBFP: the first global three-dimensional building footprint dataset. *Earth System Science Data*, 16, 5357–5374, doi.org/10.5194/essd-16-5357-2024.

Dabove, P., Daud, M. and Olivotto, L. 2024. Revolutionizing urban mapping: deep learning and data fusion strategies for accurate building footprint segmentation. *Sci Rep*, 14, 13510, doi.org/10.1038/s41598-024-64231-0.

Karsli, B., Yilmazturk, F., Bahadir, M., Karsli, F. and Ozdemir, E., 2024. Automatic building footprint extraction from photogrammetric and lidar point clouds using a novel improved-Octree approach. *Journal of Building Engineering*, 82, doi.org/10.1016/j.jobe.2023.108281.

Liasis, G. and Stavrou, S., 2016. Satellite image analysis for shadow detection and building height estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119, 437–450, doi.org/10.1016/j.isprsjprs.2016.07.006.

Liu, J. L., Qin, R. and Song, S., 2024. Automated deep learning-based point cloud classification on USGS 3DEP lidar data using a transformer. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS),* Athens, Greece, doi.org/10.1109/IGARSS53475.2024.10641055.

Luo, L., Li, P. and Yan, X., 2021. Deep learning-based building extraction from remote sensing images: A comprehensive review. *Energies*, 14(23):7982, doi.org/10.3390/en14237982.

Park, Y. and Guldmann, J.-M., 2019. Creating 3D city models with building footprints and LIDAR point cloud classification: A machine learning approach. *Computers, Environment and Urban Systems,* 75, 76–89, doi.org/10.1016/j.compenvurbsys.2019.01.004.

Rottmann, P., Haunert, J.-H. and Dehbi, Y., 2022. Automatic building footprint extraction from 3D laserscans. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, X-4/W2-2022, 233–240, doi.org/10.5194/isprs-annals-X-4-W2-2022-233-2022.

Schlosser, A. D., Szabó, G., Bertalan, L., Varga, Z., Enyedi, P. and Szabó, S., 2020. Building extraction using orthophotos and dense point cloud derived from visual band aerial imagery based on machine learning and segmentation. *Remote Sensing*. 12(15):2397, doi.org/10.3390/rs12152397.

Shi, Y., Chen, X. and Zhang, T., 2020. Cloud-based deep learning on AWS Open Data Registry: Automatic building and road extraction from satellite and LiDAR. *ACM SIGSPATIAL 2020 International Workshop on Geospatial Data Access and Processing APIs*, doi.org/10.1145/3423452.3430693.

Vincent, J. M. and Varalakshmi, P., 2024. Extraction of building footprint using MASK-RCNN for high resolution aerial imagery. *Environmental Research Communications*, 6, 075015, doi.org/10.1088/2515-7620/ad5b3d.

Abstracts of the International Cartographic Association, 10, 177, 2025
32nd International Cartographic Conference (ICC 2025), 17–22 August 2025, Vancouver, Canada.
https://doi.org/10.5194/ica-abs-10-177-2025