ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-4/W4-2024
8th International Conference on Smart Data and Smart Cities (SDSC), 4–7 June 2024, Athens, Greece

Building extraction in urban and rural areas with aerial and LiDAR DSM

Emilia Hattula, Lingli Zhu, Jere Raninen

National Land Survey of Finland (NLS) firstname.lastname@nls.fi

Keywords: Deep Learning; UNet; Building Extraction; Digital Surface Model; Digital Elevation Model

Abstract

Automatizing the extraction of different objects from remote sensing data with deep learning methods has been a popular research topic. Buildings have been one of those popular objects to be extracted. Not only does the selection of neural network affect the results and accuracy of extracted buildings, but also the selection of different types of data for the task. Digital surface models (DSMs) are increasingly used in remote sensing and their demand has increased. Retrieving height information from surface models has proved helpful for accurate extraction of buildings. In this study was investigated, if the use of light detection and ranging (LiDAR) DSMs and DEMs with 25 cm pixel resolution will lead to more accurate building extraction results in comparison to the use of aerial DSMs. Results with UNet models trained with building vector labels, DSMs, DEMs and true orthophotos from multiple areas of Finland, were produced with different data combinations from two Finnish cities, Savonlinna and Pudasjärvi, to see, which combination would lead to the most accurate building detection results. Results were evaluated partly by visual inspection, and partly by quantitative assessment. Based on the tests carried out, combining the information from true orthophotos with LiDAR DSMs and 25 cm DEMs provided the most accurate results. In forest area, using LiDAR data increased the accuracy of building detection. However, in urban area, due to missing buildings from LiDAR data, its advantages were compromised. We suggest that the use of both imagery and LiDAR data should be the optimal solution.

1. Introduction

The use of artificial intelligence (AI) has been increasing fast during the last years and automatizing the extraction of different objects from remote sensing data with deep learning methods for reducing the amount of manual labour needed has been a popular research topic (Luo et al., 2021, Romero et al., 2016). Studying the extraction of buildings with the help of deep learning and convolutional neural networks (CNN), which have proved to be successful in different extraction tasks (Yang et al., 2018, Shao et al., 2020), has been one of these popular topics to investigate in the field of remote sensing. Not only does the selection of neural network affect the results and accuracy of extracted buildings, but also the selection of different types of data for the task. Digital surface models (DSMs) are increasingly used in remote sensing and their demand has increased, as there is an expanding market for their applications (Cornelis Stal and Goossens, 2013). Retrieving height information from DSMs and digital elevation models (DEMs) has proved helpful for accurate extraction of buildings (Guan et al., 2013, Maltezos et al., 2017, Gilani et al., 2016).

In this study has been investigated, if the use of light detection and ranging (LiDAR) DSMs and DEMs having originally 25 cm pixel resolution will lead to more accurate building extraction results in comparison to the use of aerial DSMs and 2 m DEM resampled into the 25 cm pixel resolution. Results with UNet models trained with building vector labels, DSMs, DEMs and true orthophotos from multiple areas of Finland, were produced with different data combinations to see, which data combination would lead to the best results. UNet is a neural network developed in 2015 by Ronneberger et al. (Ronneberger et al., 2015). The performance on the different DSM and DEM types were tested in 2 Finnish cities, Savonlinna and Pudasjärvi. Results were evaluated mainly by visual inspection, but in the forest areas also quantitative assessment was used.

2. Material

Data from multiple areas and types of Finland collected by the National Land Survey of Finland (NLS) were exploited in this study. True orthophotos had the pixel resolution of either 30 cm or 25 cm depending on the test area. Similarly, the aerial DSM was of the pixel resolution of either 30 cm or 25 cm, as well as the LiDAR DSM. The creation method for the LiDAR DSM was the last and only pulse and it was produced from five points per m ^{2} laser scanning data. The aerial DSM was produced together with the true orthophotos and five points per m ^{2} laser scanning data was used as help if there were problems in the coverage of the aerial images. There were two types of DEMs tested, both produced from LiDAR data: originally 2 m pixel resolution resampled into 30 cm or 25 cm pixel resolution depending on the test area used, and DEM originally of 25 cm pixel resolution. 2 m DEMs were based on LiDAR data with a density of half points per m ^{2}. It is publicly available from the NLS. 25 cm DEMs were made from five points per m ^{2} LiDAR data and were produced according to the needs or request. All of the data types were utilized in the format of GeoTIFF, with the coordinate system of ETRS89-TM35FIN. The true orthophoto data contained the information x, y, R, G, B and Near Infrared (NIR) that was not utilized.

The first test area for LiDAR DSM was the Savonlinna area. Data for the Savonlinna area had a pixel resolution of 30 cm. True orthophotos for the area were collected during the year 2021. Total of 25,72 km^2 data was used for testing the LiDAR DSM for the area. For different kinds of tests, the area was divided into 8.08 km^2 training area, 0.72 km^2 validation area, and 2.16 km^2 test area, as well as two forest test areas of size 5.76km ^{2} and 9km ^{2}, which are seen in Figure 1. For the Savonlinna area tests, only 2 m DEM resampled into 30 cm pixel resolution was used, as well as true orthophotos, LiDAR DSM and aerial DSM. More accurate, originally 30 cm DEM was not available for the test area.

https://doi.org/10.5194/isprs-annals-X-4-W4-2024-73-2024

(**a**) (**b**)

Figure 1. Two forest test areas from Savonlinna. (**a**) Forest test area 1 of size 5.76 km^2. (**b**) Forest test area 2 of size 9 km^2.

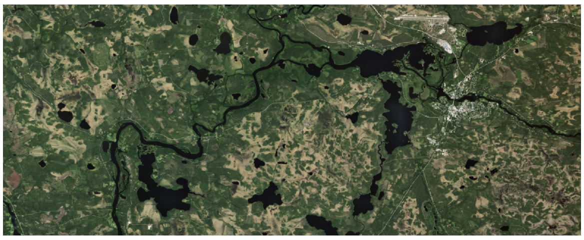


Figure 2. Pudasjärvi area for testing LiDAR and aerial DSM data, as well as DEM originally of 25 cm pixel resolution and 2 m DEM resampled into 25 cm pixel resolution.

The second test area was the Pudasjärvi area. The data had a pixel resolution of 25 cm. This area used also the originally 25 cm DEM for the tests. The Pudasjärvi test area included mapsheets S4333, as well as S5111A and B, covering total of 360 km^2. The data for Pudasjärvi area were collected during the year 2022. The Pudasjärvi test area is seen in Figure 2.

For testing the different DSM and DEM types in the Savonlinna and Pudasjärvi area, UNet models trained with data from Finland were used for extracting buildings. The model used to test Pudasjärvi area was trained with both 30 cm and 25 cm pixel resolution data. First training datasets given to the model had 30 cm pixel resolution and later the model had been finetuned with new, 25 cm pixel resolution datasets. All of the training datasets used 2 m DEM resampled into either 30 cm or 25 cm pixel resolution depending on the training dataset. The models used for producing outputs for Savonlinna area were trained with less datasets only having the pixel resolution of 30 cm, as 25 cm pixel resolution data hadn't been available for training at the time. Summary of the different datasets is seen in Table 1.

Dataset	**Location(s)**	**Area(s)**	**Pixel Resolution(s)**	**Purpose**
D1	Finland, Savonlinna, urban city area	2.16 km^2	30 cm	Testing
D2	Finland, Savonlinna, forest area	5.76 km^2	30 cm	Testing
D3	Finland, Savonlinna, forest area	9 km^2	30 cm	Testing
D4	Finland, Savonlinna, urban city area	8.08 km^2	30 cm	Training
D5	Finland, Savonlinna, urban city area	0.72 km^2	30 cm	Validation
D6	Finland, Pudasjärvi, rural area	25.38 km^2	25 cm	Testing

Table 1. Datasets used for testing LiDAR and aerial DSMs, and also 25 cm DEM and 2 m DEM resampled into 25 cm pixel resolution in the Pudasjärvi area.

3. Methods

Building extraction results with UNet models trained with different datasets of Finland were produced from test areas with LiDAR DSMs and aerial DSMs for evaluating which DSM type would lead to most accurate building detection results. In addition, more accurate 25 cm pixel resolution DEM was tested and compared with 2 m pixel resolution DEM resampled into 25 cm pixel resolution in Pudasjärvi area. The architecture of the UNets trained is seen in Figure 3.

3.1 Savonlinna LiDAR DSM test 1

2 UNet models were trained with Savonlinna data. Training set covered an area of 8.08 km^2 of Savonlinna, validation set 0.72 km^2 and test set 2.16 km^2. Model 1 used DSM produced from laser scanning data for training, validation, and testing. Model 2 used DSM made from aerial images for training, validation, and testing.

3.2 Savonlinna LiDAR DSM test 2

ATMU UNet model, at the time trained with 21 different training sets of Finland (30 cm pixel resolution true orthophotos, aerial DSM and 2 m DEM resampled into 30 cm pixel resolution), was used to produce outputs for Savonlinna test area. Two different outputs were produced; for the first output, the aerial DSM was used, and for the second output, LiDAR DSM. The results were compared.

3.3 Savonlinna LiDAR DSM test 3

ATMU UNet model, as well as model 1 and model 2 from Savonlinna DSM test 1 were tested with two forest test areas from Savonlinna, covering areas of 5.76 km^2 and 9 km^2. Model 1 produced outputs with LiDAR DSM, model 2 with aerial DSM, and the ATMU UNet model produced outputs with both LiDAR and aerial DSM. These four outputs were then compared.

3.4 Savonlinna LiDAR DSM test 4

In the fourth Savonlinna LiDAR DSM test, some LiDAR DSM data was added to the ATMU model's training data. The model originally had two Savonlinna training areas, their aerial DSM was replaced with LiDAR DSM, which was total of 8.08 km^2 of training data. Other training datasets still used aerial DSM. This new model was then used for producing outputs for the two Savonlinna forest test areas, both with LiDAR DSM and aerial DSM, and the final results were compared.

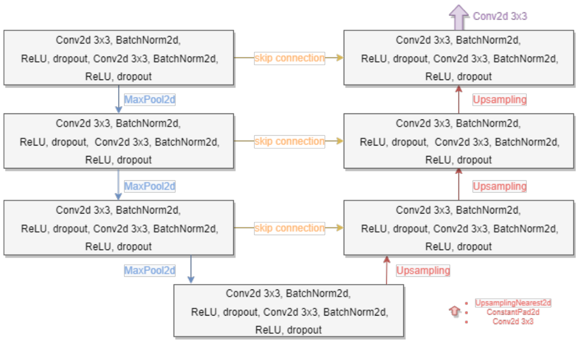


Figure 3. Implemented UNet structure used for model training (Hattula et al., 2023). Original UNet was developed by Ronneberger et al. in 2015 (Ronneberger et al., 2015).

3.5 Pudasjärvi LiDAR DSM test 1

For Pudasjärvi LiDAR test, ATMU model trained with 31 areas was used. Three different outputs were produced and compared for the Pudasjärvi test area. First output was produced with aerial DSM and 2 m DEM resampled into 25 cm pixel resolution. The second output was produced with LiDAR DSM and DSM originally of 25 cm pixel resolution. The last output was produced with LiDAR DSM and 2 m DEM resampled into 25 cm pixel resolution. This was done for finding out which combination of data would result in best building detection performance.

3.6 Pudasjärvi LiDAR DSM test 2

Pudasjärvi test area received a new version of LiDAR DSM, where some water areas with incorrect heights were completely removed. This DSM was tested with the ATMU model together with 25 cm pixel resolution DEM and compared to the previous results from the Pudasjärvi LiDAR DSM test 1 produced with the previous LiDAR DSM and 25 cm DEM.

3.7 Evaluation

Different kinds of evaluation methods were used in the different tests. In Savonlinna LiDAR DSM test 1 and 4 where model training was conducted, the main evaluation metric for the training and validation used was F1-score. The F1-score emphasizes the effect of correctly labeled building pixels and is defined through precision and recall, as seen in Equation (1).

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \tag{1}$$

For observing the model performance differences with the LiDAR DSM and aerial DSM, visual inspection was used in all of the tests. For Savonlinna LiDAR DSM tests 3 and 4, also the number of correctly detected, falsely detected, and missed buildings were counted and inspected.

4. Results

Some buildings were missing from the LiDAR data from both Savonlinna and Pudasjärvi areas, most likely due to rooftop materials and reflections, which led them being almost always missed from the building detection results with the LiDAR DSM. In addition, in Savonlinna area, LiDAR DSM and true orthophotos were from different years, which in some cases

(a) (b) (c)

Figure 4. Using LiDAR DSM for prediction often produced more accurate building shape. (**a**) A building in Savonlinna area. (**b**) Building detection result with aerial DSM. (**c**) Building detection result with LiDAR DSM.

(a) (b) (c)

Figure 5. Using LiDAR DSM for prediction helped in the detection of buildings partly covered by trees. (**a**) A couple buildings in Savonlinna area. (**b**) Building detection result with aerial DSM. (**c**) Building detection result with LiDAR DSM.

can also explain the missing buildings, if they have been demolished or build between the years. In an ideal case, the data should be from the same year. Also, both areas had some false heights in the water areas leading to false detections, for which those water areas were tested to be removed completely from the LiDAR DSM in Pudasjärvi LiDAR DSM test 2.

4.1 Savonlinna LiDAR DSM test 1

Model 1 achieved validation F1-score of 0.89998, and model 2 validation F1-score of 0.90390. With the DSM produced from laser scanning data, the model 1 produced less false detections with bridges and shore areas. With the LiDAR DSM also the detected building edges were cleaner and more accurate. An example is seen in Figure 4. Buildings partly covered by shadows and near trees were better detected with the help of LiDAR DSM in the Savonlinna city area. An example is seen in Figure 5. Some buildings missing from the LiDAR DSM were not detected. An example of a results with both DSM types in harbour/bridge area is seen in Figure 6.

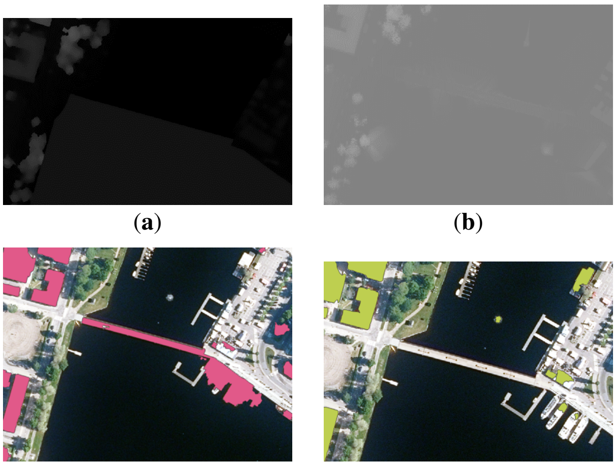
4.2 Savonlinna LiDAR DSM test 2

Less false detections in the harbor areas were produced with LiDAR DSM. Some buildings missing from the LiDAR DSM were not detected. Building outlines were a bit clearer when using LiDAR DSM. In addition, trees and shadows did not impact the building detection result with the LiDAR DSM as much as it did with the aerial DSM. In one case, window in the roof of a building had led to a hole in the building in the LiDAR DSM, leading to a false hole in the detected building. This situation is seen in Figure 7.

4.3 Savonlinna LiDAR DSM test 3

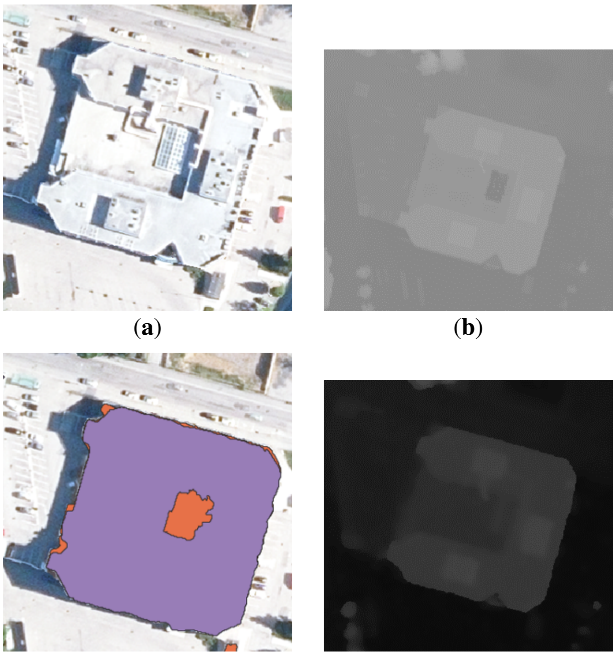
The forest area in Savonlinna had some false heights in the lakes of the LiDAR DSM, causing false detections in the water. An example of this kind of situation is seen in Figure 8. Some hay bales were falsely detected as buildings with the LiDAR DSM, as in the data they looked quite similar with the buildings. There were some cases, where the Savonlinna test model trained with LiDAR DSM was the only model detecting the buildings covered by some shadows and trees with the help of

https://doi.org/10.5194/isprs-annals-X-4-W4-2024-73-2024 | © Author(s) 2024. CC BY 4.0 License.



(**a**) (**b**) (**c**) (**d**)

Figure 6. Savonlinna bridge and harbor area. (**a**) Aerial DSM. (**b**) LiDAR DSM. (**c**) Building detection result with the aerial DSM. There's many false detections in the harbor area and the bridge. (**d**) Building detection result with the LiDAR DSM. Notably less false detections in the harbor area and the bridge.



(**a**) (**b**) (**c**) (**d**)

Figure 7. A building from Savonlinna area with glass construction on the rooftop. (**a**) True orthophoto of the building. (**b**) LiDAR DSM. (**c**) Building detection results with the LiDAR DSM and aerial DSM. Purple polygon represents the detection result with the LiDAR DSM and the orange polygon the result with the aerial DSM. A hole can be noted in the purple polygon in the location of the glass construction on the rooftop. (**d**) Aerial DSM.

the LiDAR DSM. In some cases, only the ATMU model found certain buildings when using the LiDAR DSM. The test model trained with Savonlinna area and LiDAR DSM performed the best in this test. Results are seen in Table 2.

Model	**Test area**	**Number of Detections**	**False Detections**
Model 1	Savonlinna, forest area 1	158	5
Model 2	Savonlinna, forest area 1	155	4
ATMU model with LiDAR DSM	Savonlinna, forest area 1	147	1
ATMU model with aerial DSM	Savonlinna, forest area 1	152	3
Model 1	Savonlinna, forest area 2	75	3
Model 2	Savonlinna, forest area 2	72	2
ATMU model with LiDAR DSM	Savonlinna, forest area 2	73	1
ATMU model with aerial DSM	Savonlinna, forest area 2	72	0

Table 2. The detected buildings and false detections of each model used in the two Savonlinna forest test areas.

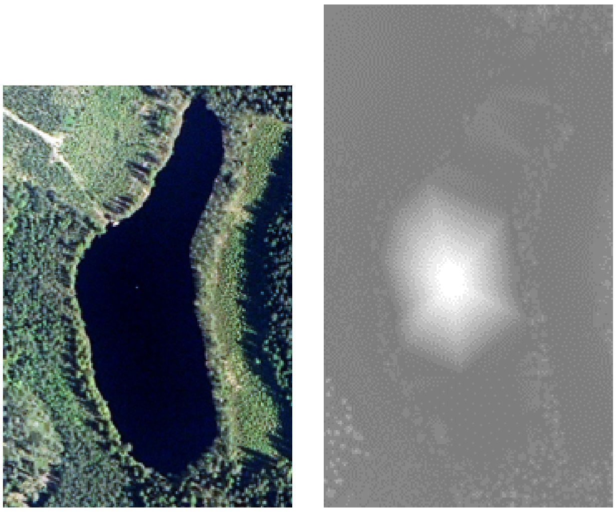
4.4 Savonlinna LiDAR DSM test 4

The ATMU model's performance with the Savonlinna forest areas and city area was inspected after replacing the model's Savonlinna training areas' aerial DSM with LiDAR DSM and retraining the model. After retraining, the model found some buildings partly covered by trees with the LiDAR DSM, but not with the aerial DSM. In general, the predictions with LiDAR DSM were completer and more accurate in comparison to the predictions with aerial DSM. False detections were produced to the harbor areas with both types of data, but considerably less with the LiDAR DSM. In some cases, with the aerial DSM multiple buildings were detected as one building, while with the LiDAR DSM they were detected as separate. An example of this kind of case is seen in Figure 9. Results from the forest areas with different models are seen in Table 3.

4.5 Pudasjärvi LiDAR DSM test 1

As in previous tests, it was noticed that also in Pudasjärvi area, the model found the shape of buildings better when LiDAR DSM was used. There were some false detections, when either LiDAR DSM or the more accurate 25 cm pixel resolution DEM were not used. An example where a couple buildings were detected only with the combination of LiDAR DSM and 25 cm DEM is seen in Figure 10b. False heights in the water areas in the LiDAR DSM caused false detections (Figure 10a). Some buildings were missed with the LiDAR DSM if they were missing from it (Figure 11). In some cases, the true orthophotos helped the detection and even if the building was missing from the LiDAR data, the building was detected. There were noticeably less false detections in the forest areas of Pudasjärvi when LiDAR DSM and 25 cm pixel resolution DEM was used,

https://doi.org/10.5194/isprs-annals-X-4-W4-2024-73-2024 | © Author(s) 2024. CC BY 4.0 License.



(**a**) (**b**)

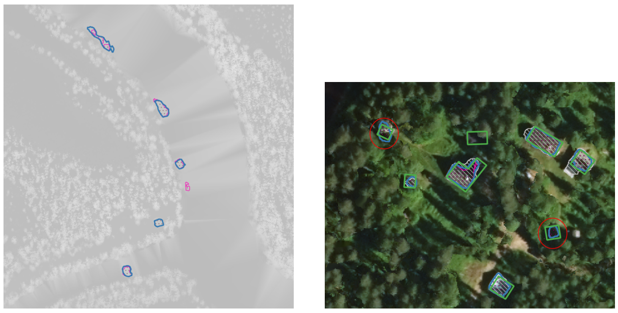
Figure 8. Some water areas in the LiDAR data had false heights seeming higher than the ground surrounding them. (**a**) A lake in the Savonlinna area seen in a true orthophoto. (**b**) The LiDAR DSM of the lake.

(**a**) (**b**)

Figure 9. In some cases, making prediction with aerial DSM produced one building for many buildings, while making prediction with the LiDAR DSM produced separate buildings. (**a**) Buildings in Savonlinna area. (**b**) Making prediction with aerial DSM produced one building (yellow), while making prediction with LiDAR DSM produced separate buildings (purple).

DSM type used for prediction	**Test area**	**Number of Detections**	**False Detections**
LiDAR DSM	Savonlinna, Forest area 1	155	0
Aerial DSM	Savonlinna, forest area 1	158	2
LiDAR DSM	Savonlinna, forest area 2	69	1*
Aerial DSM	Savonlinna, forest area 2	67	0

Table 3. The detected buildings and false detections in the two Savonlinna forest test areas after the model has been trained with some LiDAR DSM data in addition to the old aerial DSM datasets. *The false detection was due to faulty height in the LiDAR DSM's water area.



(**a**) (**b**)

Figure 10. Utilizing LiDAR DSM in Pudasjärvi area for building extraction. (**a**) Some of the water areas in Pudasjärvi area had faulty heights and seemed to be higher than the ground surrounding them. This caused false detections in the water areas, when LiDAR DSM was used. (**b**) Buildings detected only if LiDAR DSM and 25 cm DEM were used. Green building vectors from topographic database were used to help detecting where buildings are located.

proving out to be the best combination of data for the building detection task and the accuracy of detected building outlines.

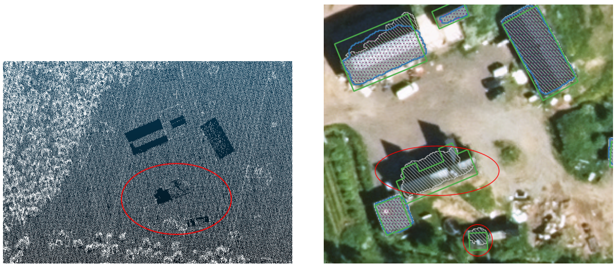
4.6 Pudasjärvi LiDAR DSM test 2

The water areas having false height were completely removed from the LiDAR DSM, replaced with holes in the data. An example is seen in Figure 12. This change in the LiDAR DSM removed the false detections from the water and shore areas completely when using LiDAR DSM for building detection, proving out to be a considerable improvement that solved one of the largest noticed problems with the use of LiDAR DSMs in these tests.

5. Discussion

Based on the multiple tests conducted in this study, building extraction results with UNet models trained with different datasets led to the best results if LiDAR DSM was used instead of aerial DSM. In addition, if 25 cm pixel resolution DEM was available for replacing the 2 m pixel resolution DEM resampled into 25 cm pixel resolution, it enhanced the building detection result further. The most notable improvements that these data types

https://doi.org/10.5194/isprs-annals-X-4-W4-2024-73-2024 | © Author(s) 2024. CC BY 4.0 License.



(a) (b)

Figure 11. Utilizing LiDAR and aerial DSM in Pudasjärvi area for building extraction. (**a**) Some buildings were partly or completely missing from the LiDAR DSM due to, for example, rooftop material. (**b**) Some buildings were detected only with aerial DSM due to missing building points in the LiDAR DSM. Detected buildings using LiDAR DSM can be seen with blue polygons, and detected buildings using aerial DSM can be seen with white polygons. The buildings surrounded with red had some points missing from the LiDAR DSM and were detected only with aerial DSM.

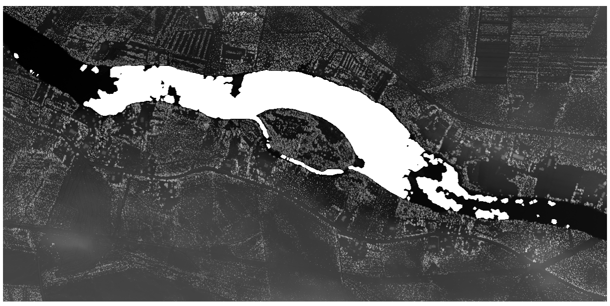


Figure 12. Pudasjärvi water areas with faulty heights were removed from the LiDAR DSM, which removed the false detections from water areas when utilizing the LiDAR DSM.

provided was the decreased false detections in the forest, harbor, and bridge areas, as well as the enhanced detection of buildings partly covered by shadows or trees. The detected shape of buildings was also enhanced in comparison to the ones produced with aerial DSM.

The biggest problems encountered with the use of the LiDAR DSM instead of aerial DSM were the false heights of some water areas leading to false detections, missing buildings from the LiDAR data due to, for example, rooftop material, plants on the roof, or reflections, and the availability of the LiDAR data, as for some areas the LiDAR data was not available from the same year as the true orthophotos. For the false heights in water areas, removing the data from the areas turned out to be a solution for the problem, and the false detections disappeared. Not only counting on the LiDAR data for building detection, but also on true orthophotos helps to some degree with the missing buildings of the LiDAR data, as the deep learning model might be able to still detect the buildings with the help of the true orthophotos. If only LiDAR data was used, those buildings would always end up missing from the model prediction.

6. Conclusions

In this paper, several different remote data types were investigated and it was evaluated, which data combination would lead to the most accurate building extraction results with UNet deep learning model for semantic segmentation. The model performance on LiDAR and aerial DSM data was evaluated on the Savonlinna, urban and forest, and Pudasjärvi, rural, areas. In addition, in the Pudasjärvi area, the effect of DEM originally of 25 cm pixel resolution was compared to DEM resampled from 2 m pixel resolution to 25 cm pixel resolution. The results were mainly evaluated visually, but in the Savonlinna forest test areas also quantitative assessment was used. It was noticed that the use of LiDAR DSM together with 25 cm DEM, if available, lead to less false detections in the forest, harbour, and bridge areas, as well as to more accurate extracted building shapes.

The use of LiDAR DSM had some problems, including some false heights in the data leading to false detections and missing buildings due to, for example, rooftop material. In addition, the availability of LiDAR data from the same year as the true orthophotos might be a problem in some cases, and should be taken into account. The benefits from LiDAR data vary in different environment. In forest area, using LiDAR data increased the accuracy of building detection. However, in urban area, due to the missing buildings from LiDAR data, its advantages were compromised. We suggest that the use of both imagery and LiDAR data should be the optimal solution.

Though the use of LiDAR data turned out to be helpful and leading to better predictions on forest areas, buildings covered by trees remain problematic for the deep learning model, and the topic should be investigated further. Moving towards the use of 3-dimensional (3D) data in the future might offer insights for the problem, as new solutions for 3D data are constantly developed (Li et al., 2020, Bello et al., 2020).

7. Acknowledgements

The work was supported by the National Land Survey's AI4TDB project (Artificial Intelligence for Topographic Database Accuracy Enhancement), which was funded by the Ministry of Agriculture and Forestry in Finland for the period of 1.1.2023-31.12.2023.

The authors also wish to thank CSC - IT Center for Science, Finland (urn:nbn:fi:research-infras-2016072531) and the Open Geospatial Information Infrastructure for Research (Geoportti, urn:nbn:fi:research-infras-2016072513) for computational resources and support.

References

Bello, S. A., Yu, S., Wang, C., Adam, J. M., Li, J., 2020. Review: Deep Learning on 3D Point Clouds. *Remote Sensing*, 12(11). https://www.mdpi.com/2072-4292/12/11/1729.

Cornelis Stal, Frederik Tack, P. D. M. A. D. W., Goossens, R., 2013. Airborne photogrammetry and lidar for DSM extraction and 3D change detection over an urban area – a comparative study. *International Journal of Remote Sensing*, 34(4), 1087-1110. https://doi.org/10.1080/01431161.2012.717183.

Gilani, S. A. N., Awrangjeb, M., Lu, G., 2016. An Automatic Building Extraction and Regularisation Technique Using LiDAR Point Cloud Data and Orthoimage. *Remote Sensing*, 8(3). https://www.mdpi.com/2072-4292/8/3/258.

Guan, H., Li, J., Chapman, M., Deng, F., Ji, Z., Yang, X., 2013. Integration of orthoimagery and lidar data for object-based urban thematic mapping using random forests. *International Journal of Remote Sensing*, 34(14), 5166-5186.

Hattula, E., Zhu, L., Raninen, J., Oksanen, J., Hyyppä, J., 2023. Advantages of Using Transfer Learning Technology with a Quantative Measurement. *Remote Sensing*, 15(17). https://www.mdpi.com/2072-4292/15/17/4278.

Li, D., Shen, X., Yu, Y., Guan, H., Li, J., Zhang, G., Li, D., 2020. Building Extraction from Airborne Multi-Spectral LiDAR Point Clouds Based on Graph Geometric Moments Convolutional Neural Networks. *Remote Sensing*, 12(19). https://www.mdpi.com/2072-4292/12/19/3186.

Luo, L., Li, P., Yan, X., 2021. Deep Learning-Based Building Extraction from Remote Sensing Images: A Comprehensive Review. *Energies*, 14(23). https://www.mdpi.com/1996-1073/14/23/7982.

Maltezos, E., Doulamis, N., Doulamis, A., Ioannidis, C., 2017. Deep convolutional neural networks for building extraction from orthoimages and dense image matching point clouds. *Journal of Applied Remote Sensing*, 11(4), 1 – 22. https://doi.org/10.1117/1.JRS.11.042620.

Romero, A., Gatta, C., Camps-Valls, G., 2016. Unsupervised Deep Feature Extraction for Remote Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 54(3), 1349-1362.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation.

Shao, Z., Tang, P., Wang, Z., Saleem, N., Yam, S., Sommai, C., 2020. BRRNet: A Fully Convolutional Neural Network for Automatic Building Extraction From High-Resolution Remote Sensing Images. *Remote Sensing*, 12(6).

Yang, H. L., Yuan, J., Lunga, D., Laverdiere, M., Rose, A., Bhaduri, B., 2018. Building Extraction at Scale Using Convolutional Neural Network: Mapping of the United States. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(8), 2600-2614.