# MapAI: Precision in Building Segmentation

Sander Jyhne[1], Dr. Morten Goodwin [2], Dr. Per-Arne Andersen [3], Dr. Ivar Oveland [4], Dr. Alexander Salveson Nossum [5], Karianne Ormseth [6], Mathilde Ørstavik [7] and Andrew C. Flatman [8]

1. *The Norwegian Mapping Authority and Centre for Artificial Intelligence Research at University of Agder, Grimstad, Norway*
2. *Centre for Artificial Intelligence Research at University of Agder, Grimstad, Norway*
3. *Centre for Artificial Intelligence Research at University of Agder, Grimstad, Norway*
4. *The Norwegian Mapping Authority, Kristiansand, Norway*
5. *Norkart, Kristiansand, Norway*
6. *AI:Hub, Grimstad, Norway*
7. *Norkart, Oslo, Norway*
8. *The Danish Agency for Data Supply and Infrastructure*

## Abstract

*MapAI: Precision in Building Segmentation* is a competition arranged with the Norwegian Artificial Intelligence Research Consortium (NORA) [1] in collaboration with Centre for Artificial Intelligence Research at the University of Agder (CAIR)[2], the Norwegian Mapping Authority[3], AI:Hub[4], Norkart[5], and the Danish Agency for Data Supply and Infrastructure[6]. The competition will be held in the fall of 2022. It will be concluded at the *Northern Lights Deep Learning conference* focusing on the segmentation of buildings using aerial images and laser data. We propose two different tasks to segment buildings, where the first task can only utilize aerial images, while the second must use laser data (LiDAR) with or without aerial images. Furthermore, we use IoU and Boundary IoU [1] to properly evaluate the precision of the models, with the latter being an IoU measure that evaluates the results' boundaries. We provide the participants with a training dataset and keep a test dataset for evaluation.

**Keywords:** artificial intelligence; machine learning; deep learning; computer vision; remote sensing; semantic segmentation

## Introduction

Buildings are an essential component of information regarding population, policy-making, and city management [2]. Using computer vision technologies such as classification, object detection, and segmentation has proved helpful in several scenarios, such as urban planning and disaster recovery [2, 3]. Segmentation is the most precise method and can give detailed insights into the data as it highlights the area of interest.

Acquiring accurate segmentation masks of buildings is challenging since the training data derives from real-world photographs. As a result, the data often have varying quality, large class imbalance, and contains noise in different forms. The segmentation masks are affected by optical issues such as shadows, reflections, and perspectives. Additionally, trees, powerlines, or even other buildings may obstruct the visibility [4]. Furthermore, small buildings have proved to be more difficult to segment than larger ones as they are harder to detect, more prone to being obstructed, and often confused with other classes [5]. Lastly, different buildings are found in several diverse areas, ranging from rural to urban locations. The diversity poses a vital requirement for the model to generalize to the various combinations. These hardships motivate the competition and our evaluation method, detailed in Evaluation Methodology.

This competition follows mostly the same strategy as previous NORA competitions [6, 7].

## Dataset Details

For the competition, we provide the participants with a dataset containing aerial images, laser data, and ground truth masks for the buildings. We split the dataset into a training dataset and a test dataset. The training dataset

---

[1] https://nora.ai
[2] https://cair.uia.no
[3] https://kartverket.no
[4] https://aihub.no/
[5] https://www.norkart.no
[6] https://eng.sdfi.dk/

(a) Aerial image sample from the training dataset



(b) Lidar sample from the training dataset
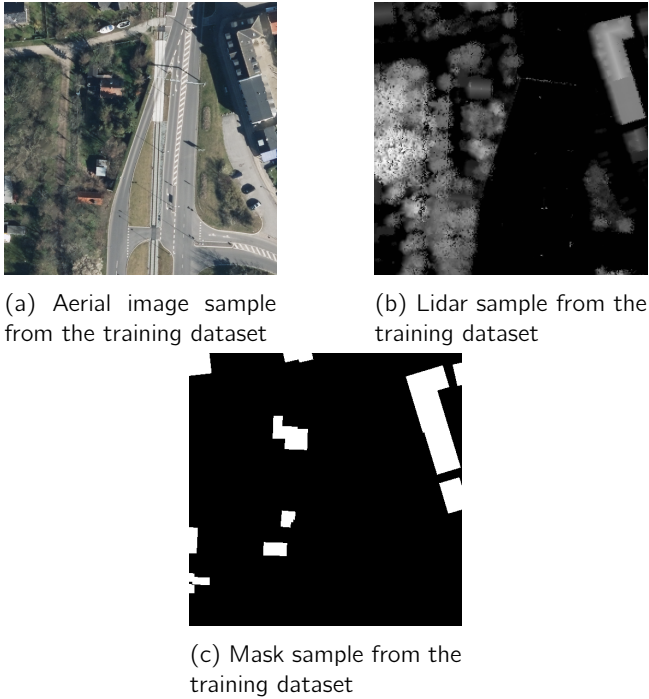


(c) Mask sample from the training dataset

Figure 1: Sample data from the training dataset

is released at the start of the competition, while the test dataset will be kept hidden until the competition is over. When the competition is complete, we will release the full dataset.

The training dataset consists of several different locations in Denmark. Area variability ensures a diverse dataset with several different environments and building types. The test dataset consists of seven locations in Norway, comprising urban and rural cities.

The data is derived from real-world data. As a result, there are cases where the buildings in the aerial image do not correspond to a ground truth mask. In addition, the ground truths in the test dataset are generated using a DTM, which skews the top of the buildings in images compared to the ground truths. The training dataset is generated using a DSM, which does not skew the top of the buildings. In Figure 1, we can see examples of the different datatypes present in the training dataset.

**Task Descriptions**
We present two subtasks; (1) an aerial image segmentation task and (2) laser data segmentation task. The participants are encouraged to submit for both tasks; however, it is not mandatory. Note that both tasks contribute equal weight to the final score, i.e., 50% each. You can only reach the maximum score by a separate submission for both tasks

*Task 1: Aerial Image Segmentation Task*
The aerial image segmentation task aims to solve the segmentation of buildings only using aerial images. Segmentation using only aerial images is helpful for

several scenarios, including disaster recovery in remote sensing images where laser data is unavailable. We ask the participants to develop machine learning models for generating accurate segmentation masks of buildings solely using aerial images.

*Task 2: Laser Data Segmentation Task*
The laser data segmentation task aims to solve the segmentation of buildings using laser data. Segmentation using laser data is helpful for urban planning or change detection scenarios, where precision is essential. We ask the participants to develop machine learning models for generating accurate segmentation masks of buildings using laser data with or without aerial images.

*Submission*
We have developed a new template for competition participation and submission. The contestants will have to fork a GitHub repository, create a folder for their team and develop their methods in the folder. When the participants are ready for submission, they push the changes to their fork and create a pull request to the original repository. Automatic Github actions will test their code and models and store their results. The GitHub repository and instructions are found at https://github.com/Sjyhne/mapai-competition.

**Evaluation Methodology**
We evaluate both tasks with the same metrics described in the following paragraphs. Overall there will be a first place and a second place winner of MapAI, determined by the sum of the score $S$ from both tasks.

We use a combination of Intersection-over-Union (IoU) and Boundary Intersection-over-Union (BIoU) to evaluate the submitted segmentation mask. The formula for the evaluation score $S$ is presented in Eq. 1.

$$S = \frac{BIoU + IoU}{2} \tag{1}$$

The Intersection-over-Union (IoU), also known as the Jaccard Index, measures the similarity between two samples, $G$ (ground truth) and $P$ (prediction), by dividing the intersecting area by the total area, seen in Eq. 2.

$$IoU = \frac{Intersection}{Union} = \frac{|G \cap P|}{|G| + |P| - |G \cap P|} \tag{2}$$

The Boundary Intersection-over-Union (BIoU) calculates the IoU of the boundary of the prediction and ground truth. A variable $d$ determines the width of the boundary used in the calculation. Fig 2 visualizes how we calculate the Boundary IoU, where $G$ and $G_d$ denote the ground truth and the edge of the ground truth with thickness $d$. Similarly to $G$, $P$ and $P_d$ denote the predicted mask and the edge of the predicted mask with thickness $d$.

The final score for the competition is the average of the score $S$ from both tasks.
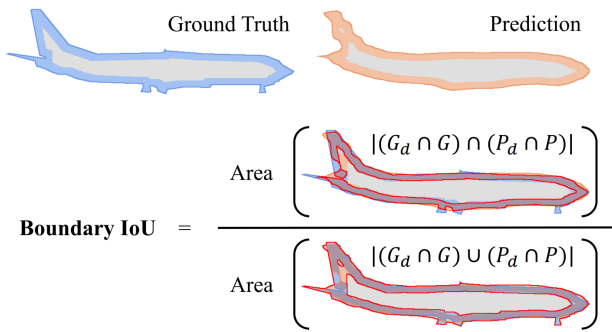
Figure 2: The Boundary Intersection-over-Union (BIoU) used to measure the accuracy of the segmentation boundary [1]

**Summary**

This paper presents the first version of *MapAI: Precision in Building Segmentation* task. The competition aims to advance the remote sensing image segmentation field by presenting two subtasks with an evaluation focused on small buildings and segmentation boundaries. We hope the competition motivates students to participate and contribute to advancing the field.

**References**

1. Cheng B, Girshick RB, Dollár P, Berg AC, and Kirillov A. Boundary IoU: Improving Object-Centric Image Segmentation Evaluation. CoRR 2021; abs/2103.16562. arXiv: 2103.16562. Available from: https://arxiv.org/abs/2103.16562

2. Grecea C, Bălă A, and Herban S. Cadastral requirements for urban administration, key component for an efficient town planning. Journal of Environmental Protection and Ecology 2013 Jan; 14:363–71

3. Sun W, Bocchini P, and Davison B. Applications of artificial intelligence for disaster management. Natural Hazards 2020 Sep; 103:2631–89. DOI: 10.1007/s11069-020-04124-3

4. Aletta S, Szabo G, Bertalan L, Varga Z, Enyedi P, and Szabo S. Building Extraction Using Orthophotos and Dense Point Cloud Derived from Visual Band Aerial Imagery Based on Machine Learning and Segmentation. Remote Sensing 2020 Jul; 12:2397. DOI: 10.3390/rs12152397

5. Kampffmeyer M, Salberg AB, and Jenssen R. Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks. *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2016 :680–8. DOI: 10.1109/CVPRW.2016.90

6. Hicks S, Jha D, Thambawita V, Halvorsen P, Singstad B, Gaur S, Pettersen K, Goodwin M, Parasa S, Lange T de, et al. MedAI: Transparency in Medical Image Segmentation. Nordic Machine Intelligence 2021; 1:1–4

7. Nordmo TAS, Kvalsvik O, Kvalsund SO, Hansen B, and Riegler MA. Fish AI: Sustainable Commercial Fishing. Nordic Machine Intelligence 2022; 2:1–3

8. 2D Semantic Labeling - Vaihingen data. Available from: https://www.isprs.org/education/benchmarks/UrbanSemLab/2d-sem-label-vaihingen.aspx

9. 2D Semantic Labeling Contest - Potsdam. Available from: https://www.isprs.org/education/benchmarks/UrbanSemLab/2d-sem-label-potsdam.aspx

10. Maggiori E, Tarabalka Y, Charpiat G, and Alliez P. Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE. 2017