Building foot-print extraction using deep learning at Lantmäteriet

Adam Hedkvist

Geographical information, Lantmäteriet

adam.hedkvist@lm.se

Table of contents

Machine Learning Algorithm Cheat Sheet

This cheat sheet helps you choose the best machine learning algorithm for your predictive analytics solution.
Your decision is driven by both the nature of your data and the goal you want to achieve with your data.

Microsoft Azure

What do you want to do?

Extract information from text

Text Analytics

Derives high-quality information from text

Answers questions like: What info is in this text?

- **Latent Dirichlet Allocation** ← Unsupervised topic modeling, group texts that are similar
- **Extract N-Gram Features from Text** ← Creates a dictionary of n-grams from a column of free text
- **Feature Hashing** ← Converts text data to integer encoded features using the Vowpal Wabbit library
- **Preprocess Text** ← Performs cleaning operations on text, like removal of stop-words, case normalization
- **Word2Vector** ← Converts words to values for use in NLP tasks, like recommender, named entity recognition, machine translation

Predict between several categories

Multiclass Classification

Answers complex questions with multiple possible answers

Answers questions like: Is this A or B or C or D?

- **Multiclass Logistic Regression** ← Fast training times, linear model
- **Multiclass Neural Network** ← Accuracy, long training times
- **Multiclass Decision Forest** ← Accuracy, fast training times
- **One-vs-All Multiclass** ← Depends on the two-class classifier
- **One-vs-One Multiclass** ← Depends on binary classifier, less sensitive to an imbalanced dataset with larger complexity
- **Multiclass Boosted Decision Tree** ← Non-parametric, fast training times and scalable

Predict between two categories

Two-Class Classification

Answers simple two-choice questions, like yes or no, true or false

Answers questions like: Is this A or B?

- **Two-Class Support Vector Machine** ← Under 100 features, linear model
- **Two-Class Averaged Perceptron** ← Fast training, linear model
- **Two-Class Decision Forest** ← Accurate, fast training
- **Two-Class Logistic Regression** ← Fast training, linear model
- **Two-Class Boosted Decision Tree** ← Accurate, fast training, large memory footprint
- **Two-Class Neural Network** ← Accurate, long training times

Generate recommendations

Recommenders

Predicts what someone will be interested in

Answers the question: What will they be interested in?

- **Use the Train Wide & Deep Recommender module** ← Hybrid recommender, both collaborative filtering and content-based approach
- **SVD Recommender** ← Collaborative filtering, better performance with lower cost by reducing dimensionality

Predict values

Regression

Makes forecasts by estimating the relationship between values

Answers questions like: How much or how many?

- **Fast Forest Quantile Regression** ← Predicts a distribution
- **Poisson Regression** ← Predicts event counts
- **Linear Regression** ← Fast training, linear model
- **Bayesian Linear Regression** ← Linear model, small data sets
- **Decision Forest Regression** ← Accurate, fast training times
- **Neural Network Regression** ← Accurate, long training times
- **Boosted Decision Tree Regression** ← Accurate, fast training times, large memory footprint

Discover structure

Clustering

Separates similar data points into intuitive groups

Answers questions like: How is this organized?

- **K-Means** ← Unsupervised learning

Classify images

Image Classification

Classifies images with popular networks

Answers questions like: What does this image represent?

- **ResNet** ← Modern deep learning neural network
- **DenseNet** ← Modern deep learning neural network

Find unusual occurrences

Anomaly Detection

Identifies and predicts rare or unusual data points

Answers the question: Is this weird?

- **One Class SVM** ← Under 100 features, aggressive boundary
- **PCA-Based Anomaly Detection** ← Fast training times

Machine Learning Algorithm Cheat Sheet

This cheat sheet helps you choose the best machine learning algorithm for your predictive analytics solution.
Your decision is driven by both the nature of your data and the goal you want to achieve with your data.

Microsoft Azure

What do you want to do?

Extract information from text

Text Analytics

Derives high-quality information from text

Answers questions like: What info is in this text?

- **Latent Dirichlet Allocation** ← Unsupervised topic modeling, group texts that are similar
- **Extract N-Gram Features from Text** ← Creates a dictionary of n-grams from a column of free text
- **Feature Hashing** ← Converts text data to integer encoded features using the Vowpal Wabbit library
- **Preprocess Text** ← Performs cleaning operations on text, like removal of stop-words, case normalization
- **Word2Vector** ← Converts words to values for use in NLP tasks, like recommender, named entity recognition, machine translation

Predict between several categories

Multiclass Classification

Answers complex questions with multiple possible answers

Answers questions like: Is this A or B or C or D?

- **Multiclass Logistic Regression** ← Fast training times, linear model
- **Multiclass Neural Network** ← Accuracy, long training times
- **Multiclass Decision Forest** ← Accuracy, fast training times
- **One-vs-All Multiclass** ← Depends on the two-class classifier
- **One-vs-One Multiclass** ← Depends on binary classifier, less sensitive to an imbalanced dataset with larger complexity
- **Multiclass Boosted Decision Tree** ← Non-parametric, fast training times and scalable

Predict between two categories

Two-Class Classification

Answers simple two-choice questions, like yes or no, true or false

Answers questions like: Is this A or B?

- **Two-Class Support Vector Machine** ← Under 100 features, linear model
- **Two-Class Averaged Perceptron** ← Fast training, linear model
- **Two-Class Decision Forest** ← Accurate, fast training
- **Two-Class Logistic Regression** ← Fast training, linear model
- **Two-Class Boosted Decision Tree** ← Accurate, fast training, large memory footprint
- **Two-Class Neural Network** ← Accurate, long training times

Generate recommendations

Recommenders

Predicts what someone will be interested in

Answers the question: What will they be interested in?

- **Use the Train Wide & Deep Recommender module** ← Hybrid recommender, both collaborative filtering and content-based approach
- **SVD Recommender** ← Collaborative filtering, better performance with lower cost by reducing dimensionality

Predict values

Regression

Makes forecasts by estimating the relationship between values

Answers questions like: How much or how many?

- **Fast Forest Quantile Regression** ← Predicts a distribution
- **Poisson Regression** ← Predicts event counts
- **Linear Regression** ← Fast training, linear model
- **Bayesian Linear Regression** ← Linear model, small data sets
- **Decision Forest Regression** ← Accurate, fast training times
- **Neural Network Regression** ← Accurate, long training times
- **Boosted Decision Tree Regression** ← Accurate, fast training times, large memory footprint

Discover structure

Clustering

Separates similar data points into intuitive groups

Answers questions like: How is this organized?

- **K-Means** ← Unsupervised learning

Classify images

Image Classification

Classifies images with popular networks

Answers questions like: What does this image represent?

- **ResNet** ← Modern deep learning neural network
- **DenseNet** ← Modern deep learning neural network

Find unusual occurrences

Anomaly Detection

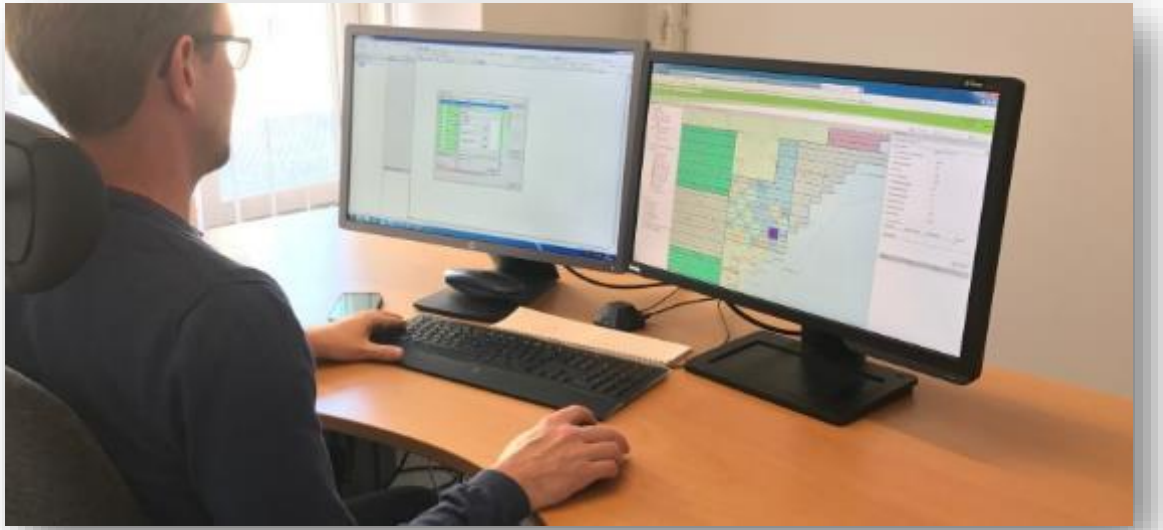
Identifies and predicts rare or unusual data points

Answers the question: Is this weird?

- **One Class SVM** ← Under 100 features, aggressive boundary
- **PCA-Based Anomaly Detection** ← Fast training times

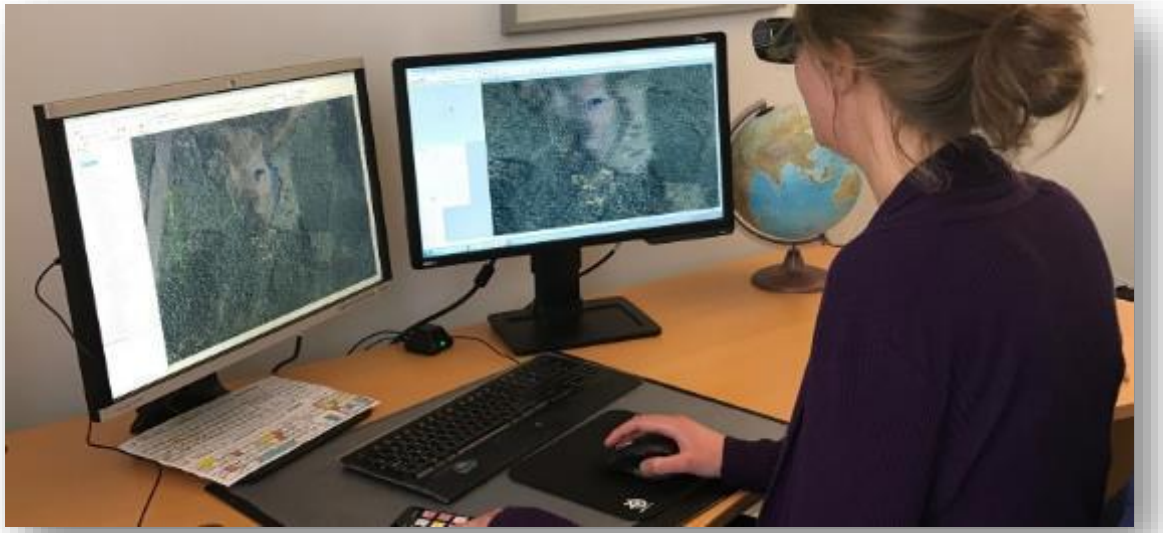
The context

- Sweden has an area of appr. 455 000 sqkm
- Appr. one third of the area is flown and photographed every year
- It is divided into three sub-areas which are updated at different intervals
- One 6-10 year region, one 4 year and one every 2 year
- The area which is updated every 2 years is captured with a 0.15m resolution
- Other areas have 0.38m resolution



Background

- Our main topographic database contains 200+ object types, within eleven themes
- The majority of these object types are captured manually, using 3D and 2D
- Object/change detection using machine learning would substantially speed up the revision time

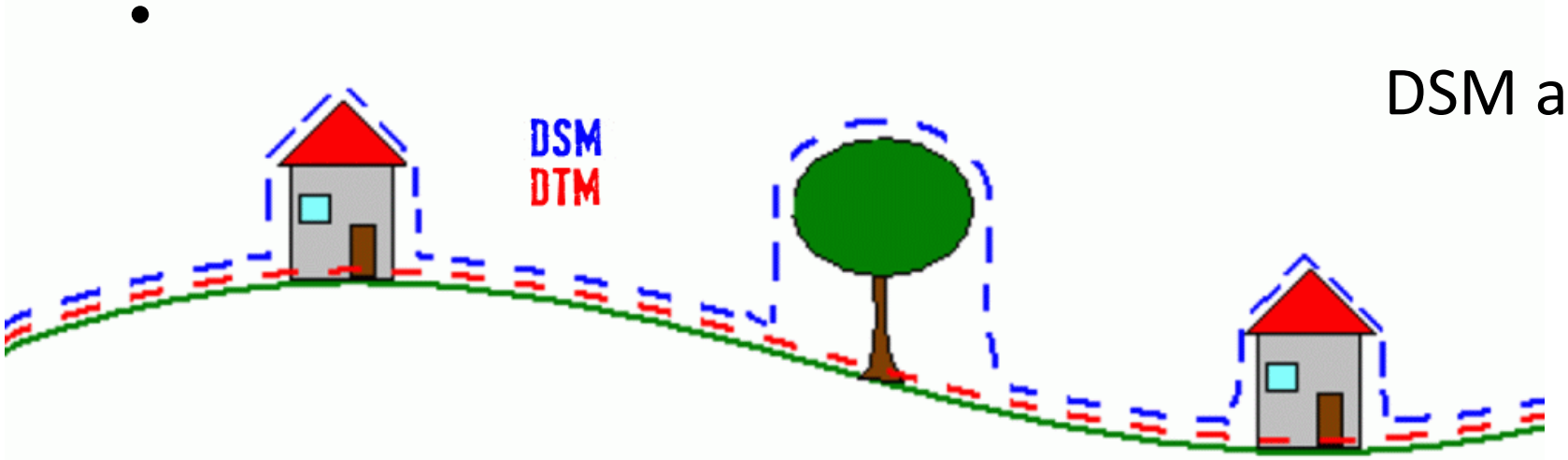


Input data

- The images is taken in R, G, B, IR bands
- To use the images in the machine learning model they are converted into true ortho images
- The true ortho images have a resolution of 0.25m
- Along with the images a normalized digital surface model is also used as input to the model

Elevation models

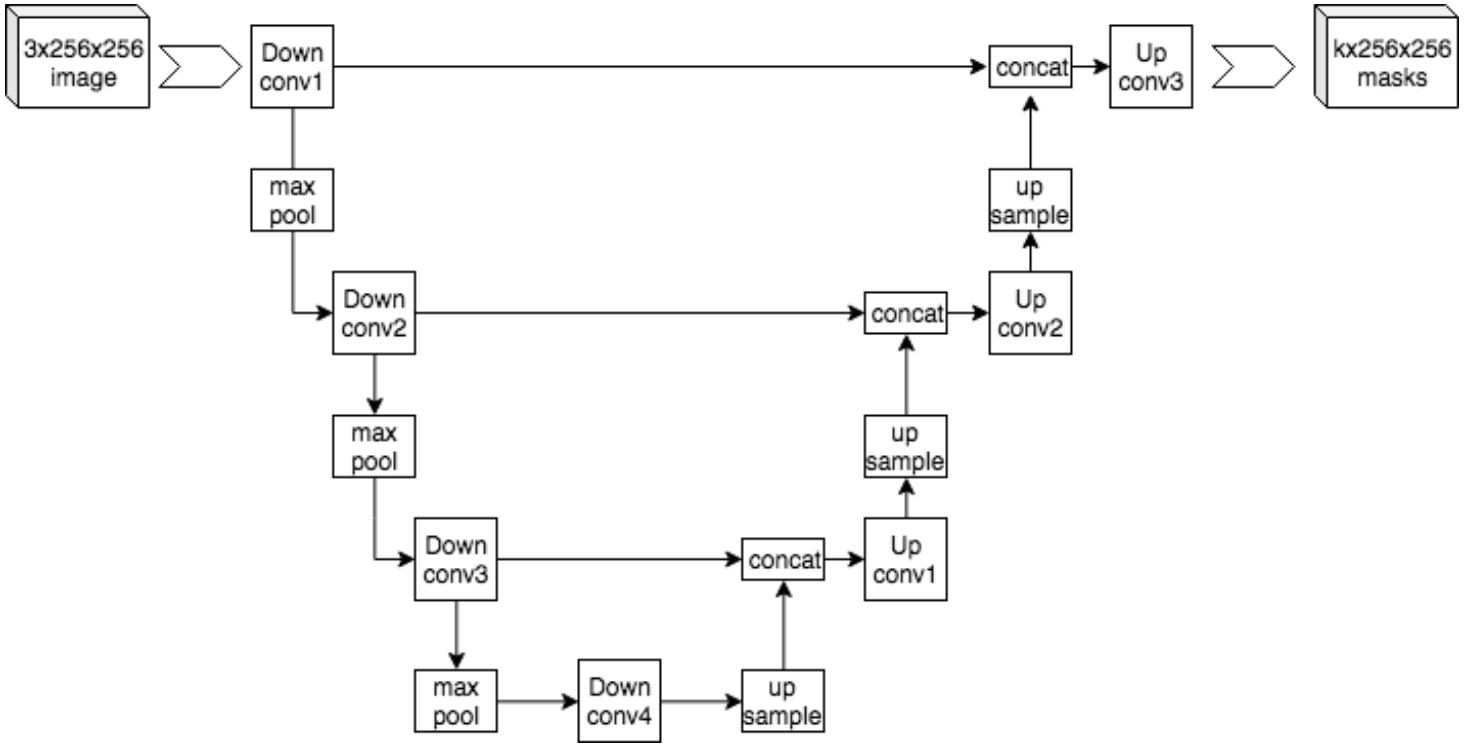
- DSM and DTM



- nDSM = DSM-DTM
- ”elevation from ground”

Model specifications

- Unet model
- Resnet101 backbone
- Relu activation function
- Tversky loss function (alpha = 0.7, beta =0.3, gamma = 0.75)
- Batch_size of 6
- Learning_rate 0.01

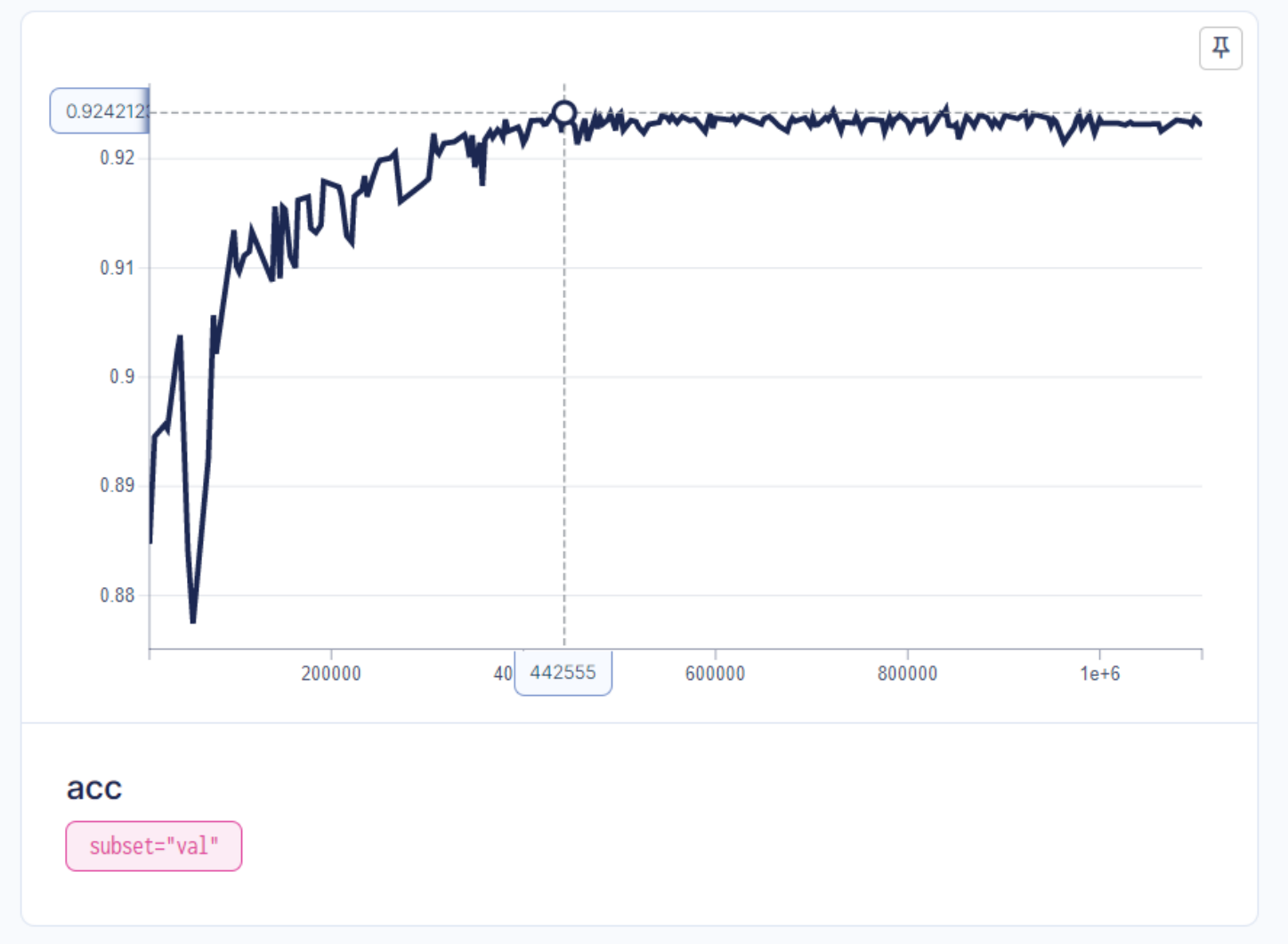


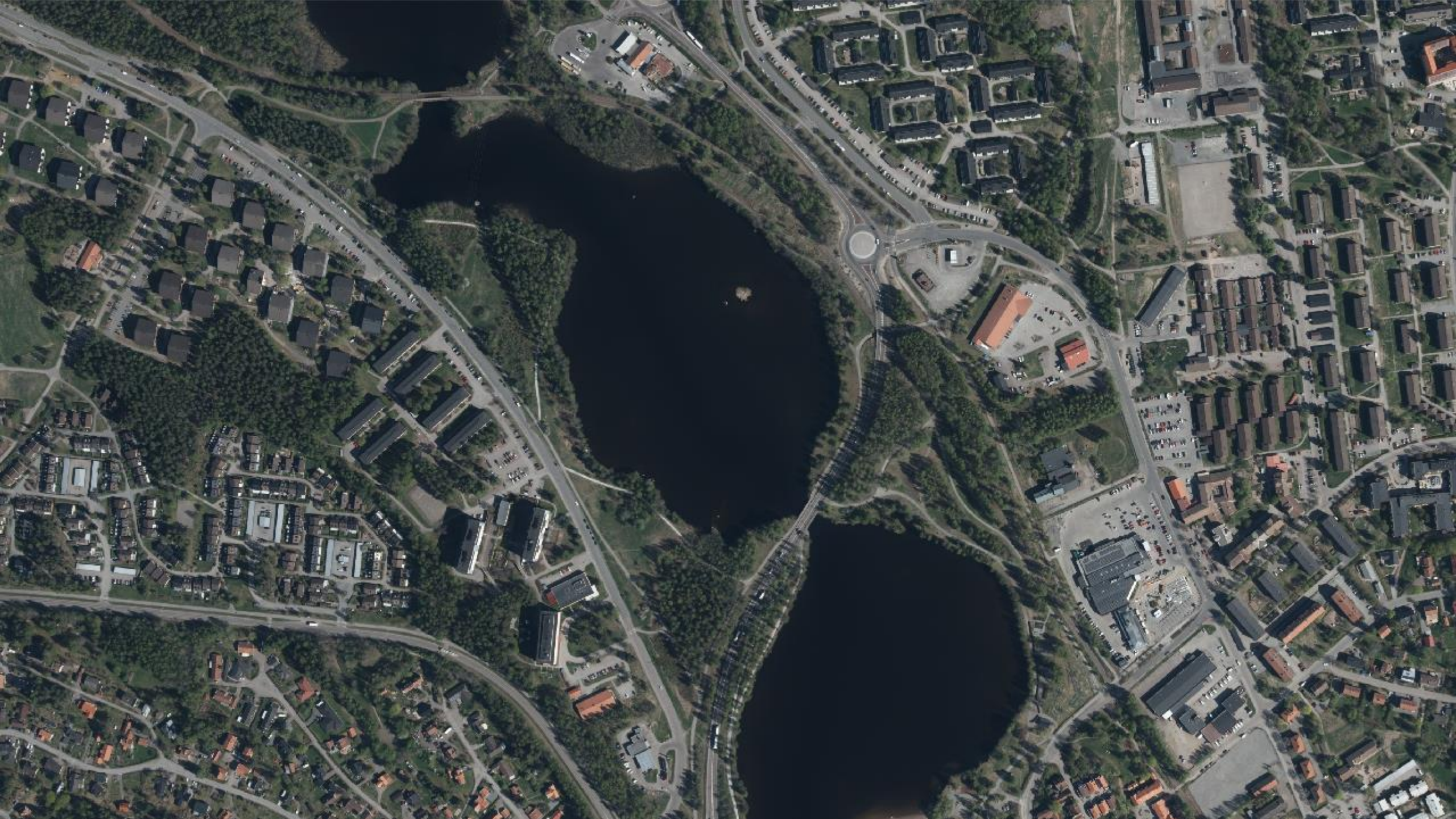
Model specifications

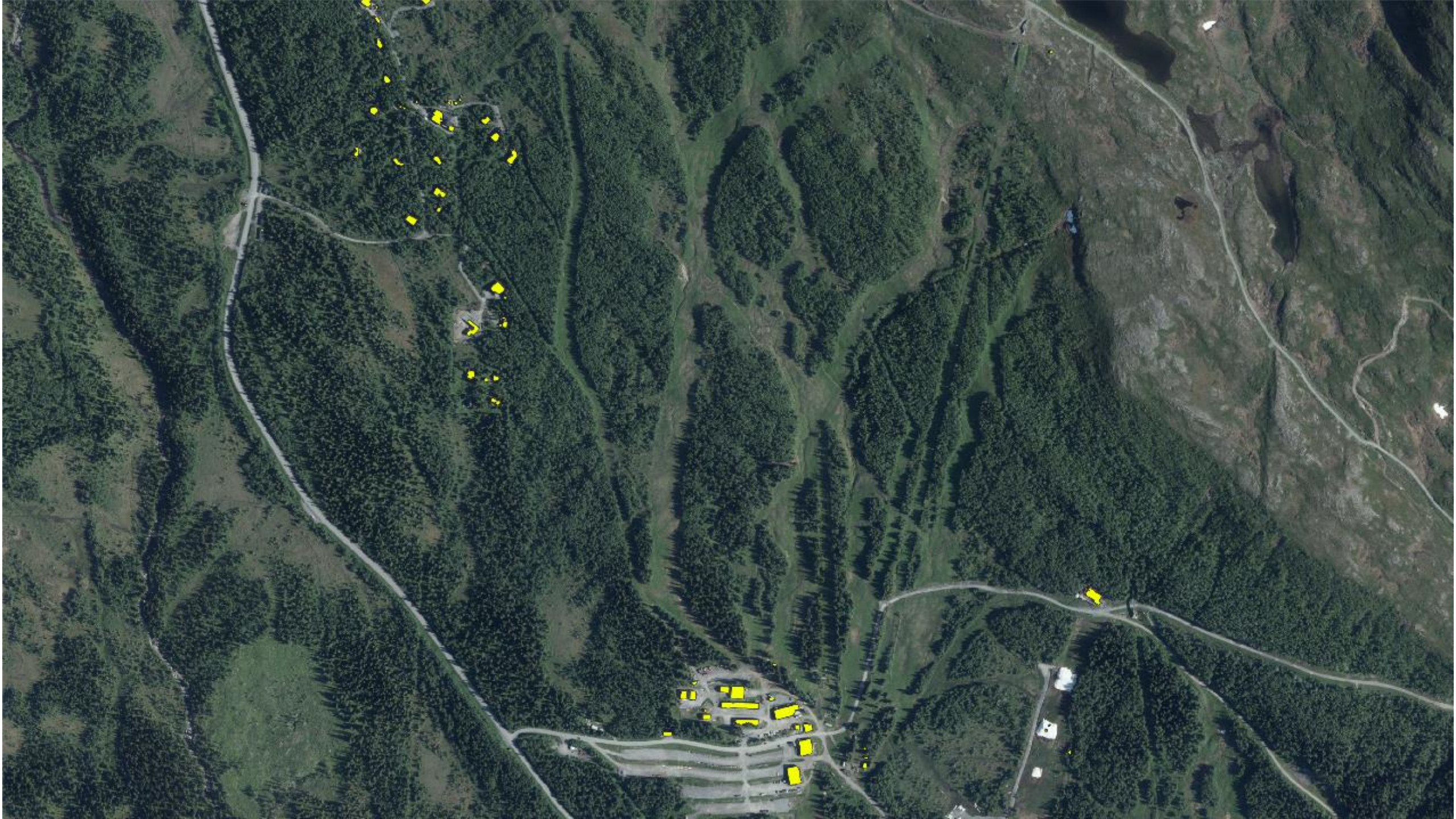
- The environment used for these results is pytorch
- Trained on 15000 True ortho photos
- Divided images into 512x512 pixels as input
- 25 cm resolution true ortho
- IR, R, G, B bands is used at input with a nDSM
- We train the model for 2-4 days
- Overall accuracy is about 92-93%

Results

- Approximately 92% accuracy
- Improvement during training

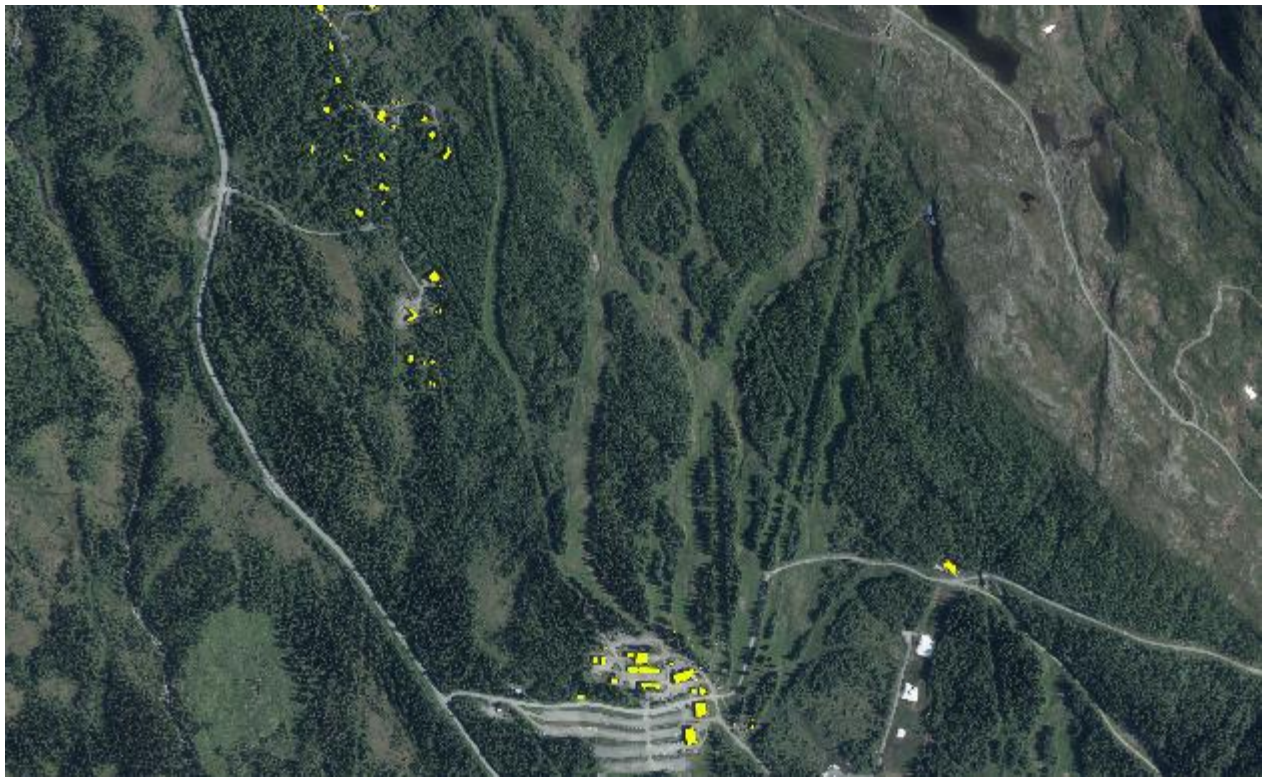






Common problems and solutions

- Ground truth data is very important, both in quality and quantity
- Big difference between urban and rural areas, important to have data of all areas ”class balancing”



Common problems and solutions

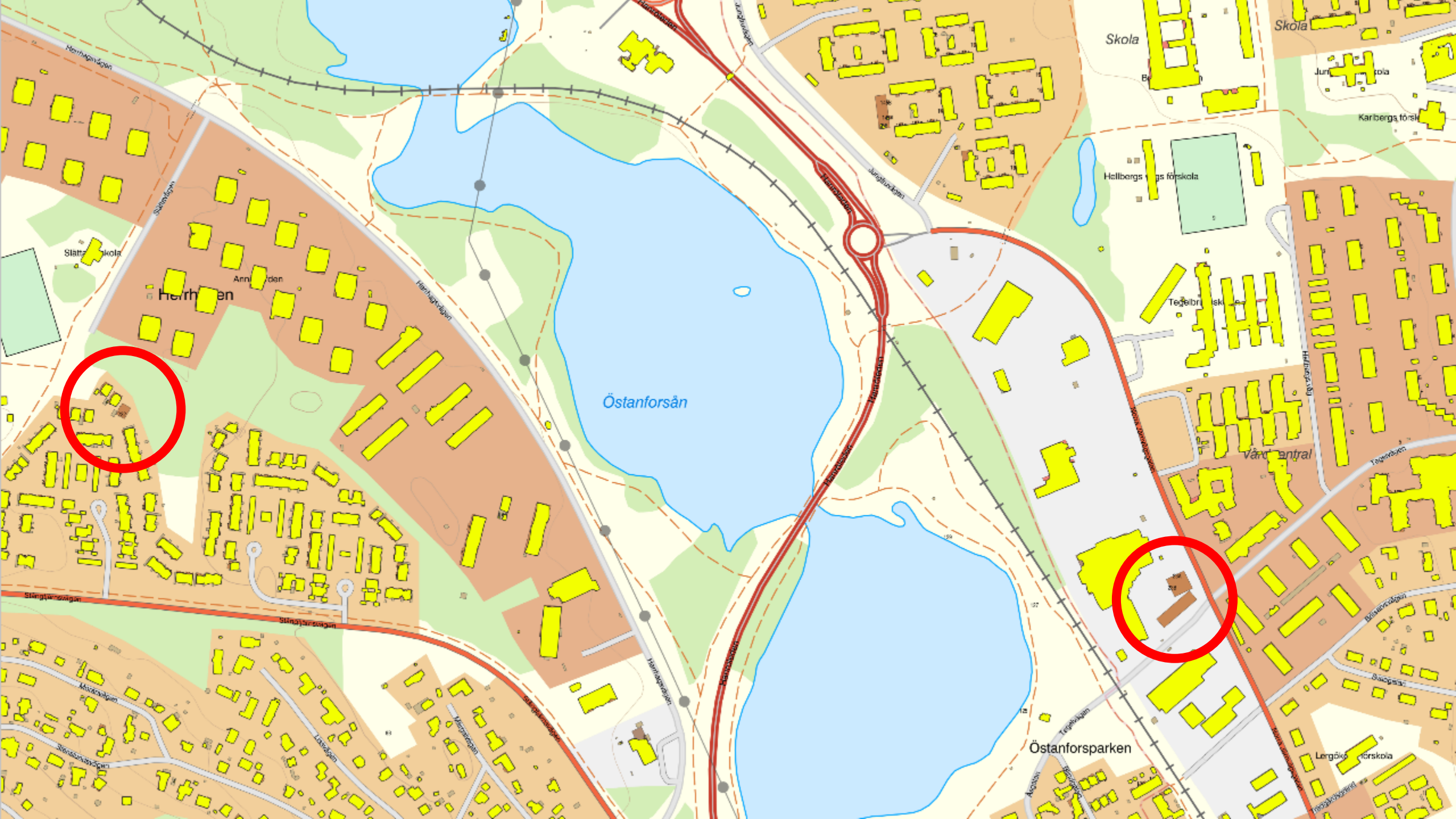
- nDSM helped a lot in Åre which is located approximately at 700 meters elevation
- Some buildings are larger than the 512x512 image (meaning the entire image is a building) – more training data

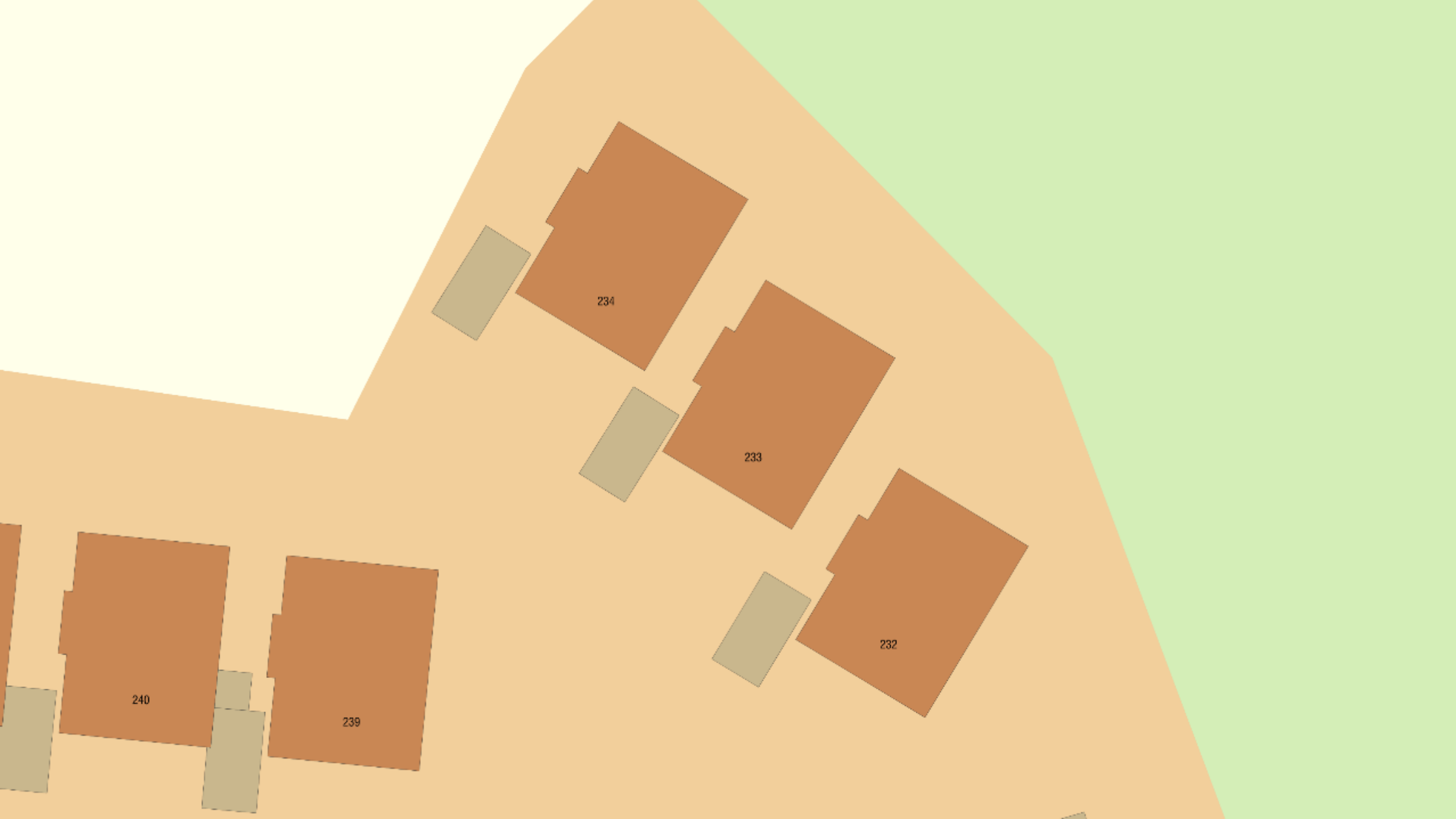
Common problems and solutions

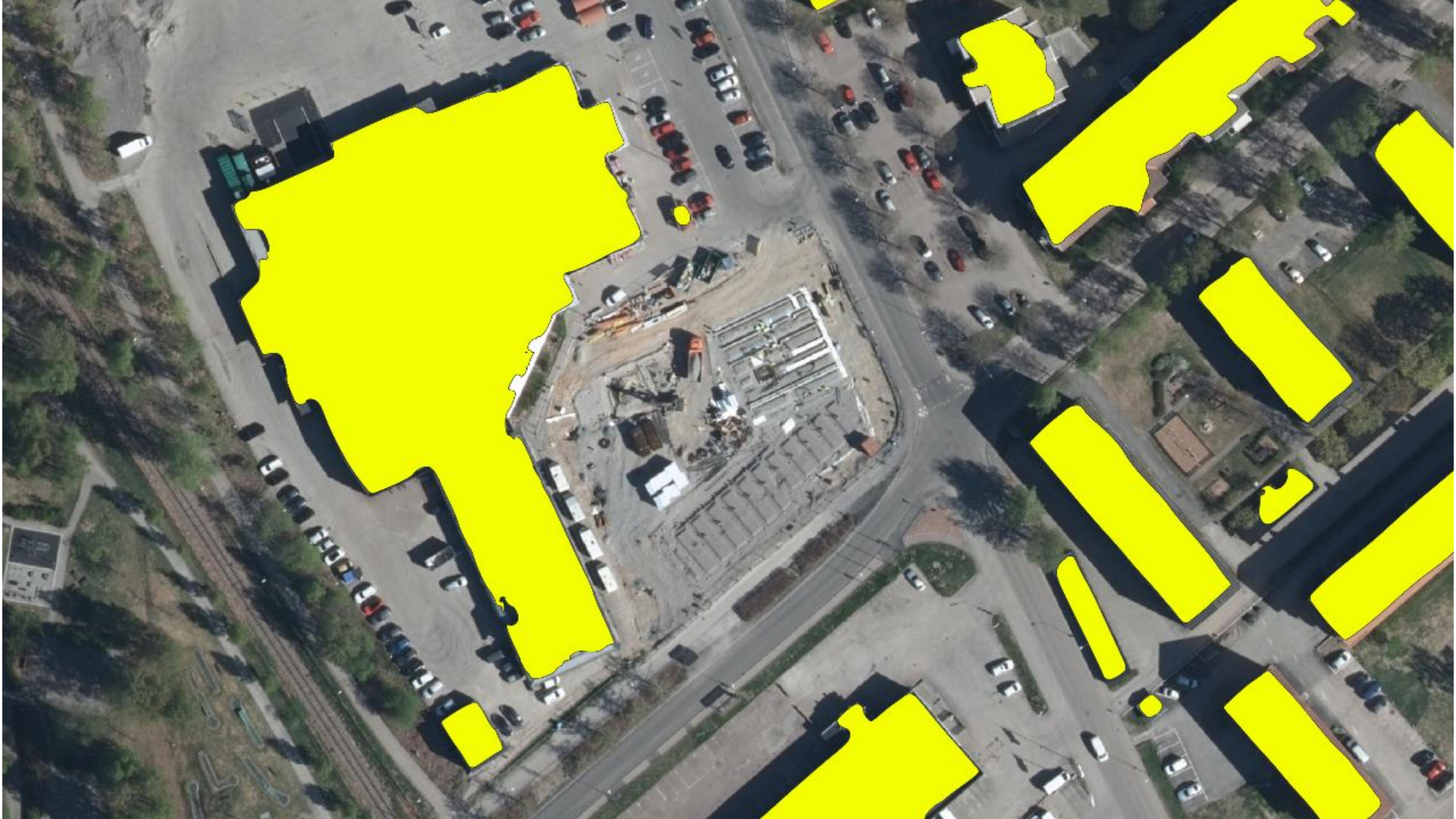
- Hard to find buildings near water – more training data

Common problems and solutions

- Example of bad data
- Example of building which is hidden


Östanforsån
Skola
Skola
Östanforsparken


234
233
232
240
239



Conclusion

- Our goal is to create a tool to improve quality and speed in aerial image analysis.
- Both to find buildings as many buildings as possible, and be as geometrically correct as possible.
- In the near future we are looking at doing the same thing on roads, powerlines and possibly landcover.

Thank you for listening

- Any questions?

- adam_hedkvist@lm.se