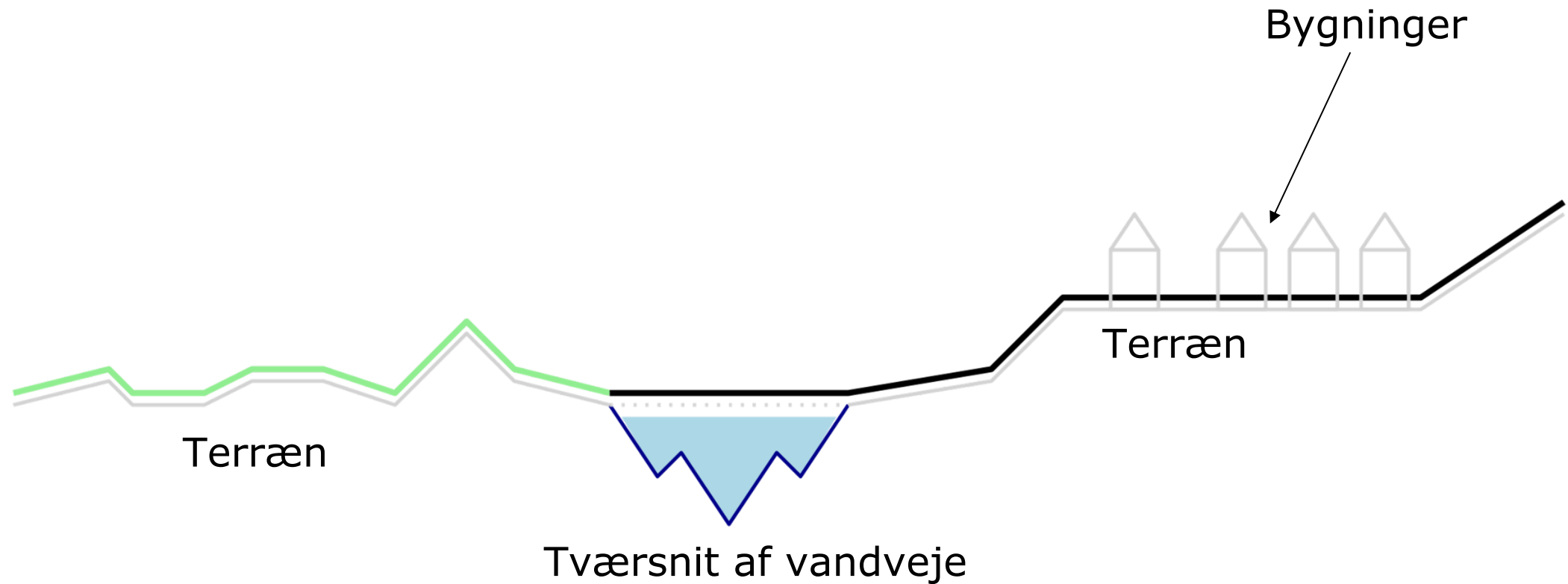


ML-baseret befæstelseskort

Teknologi, resultater og processer

Morten Revsbæk / Jonas Tranberg

Detaljeret digital model af den fysiske verden



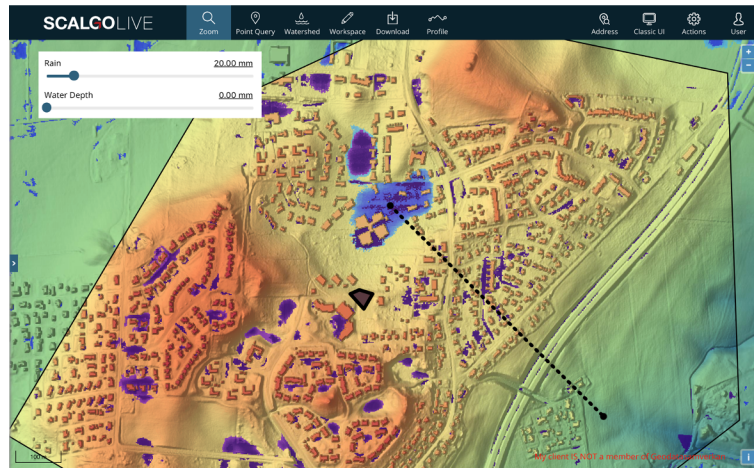
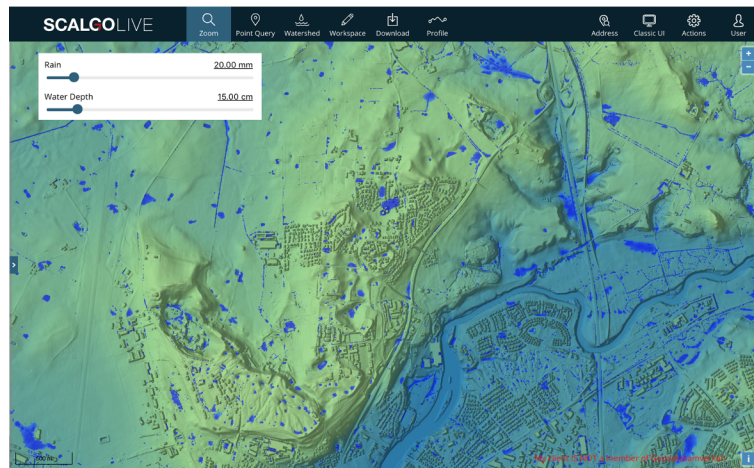
Detaljerede data er tilgængelige

- Højdemodeller i 40 cm opløsning
- Luftfoto i 10 cm opløsning
- Højere opløsning og kortere opdateringstid

Fra data til interaktiv digital model

- Indsamling af data fra offentlige databaser
- Vedligeholde opdaterede analyser og værktøjer
- Udstille i brugerflade

Overblik, indsigt og scenarier



Fra data til information og opgaveløsning

Fra detaljeret model til storskala

6 lande

> 1.000.000.000.000 elevation points

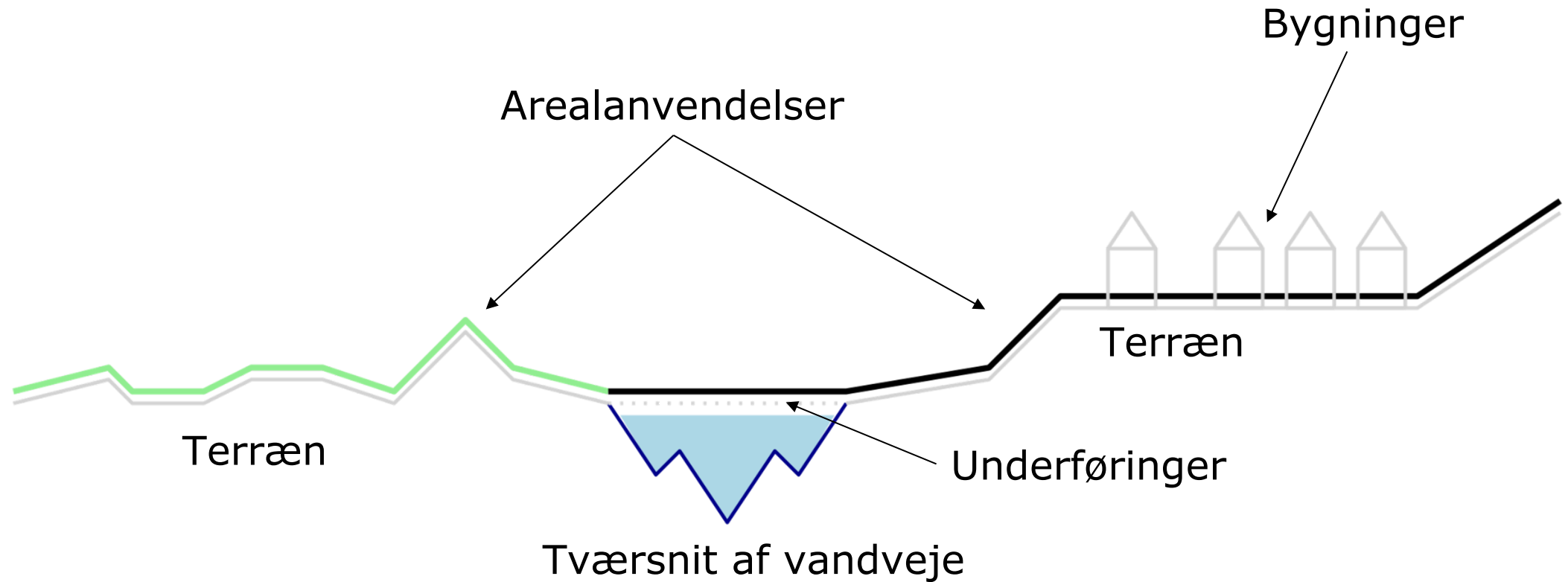
EU Open Data Directive (2023)

USGS 3DEP (2023)

(towards 10.000.000.000.000)

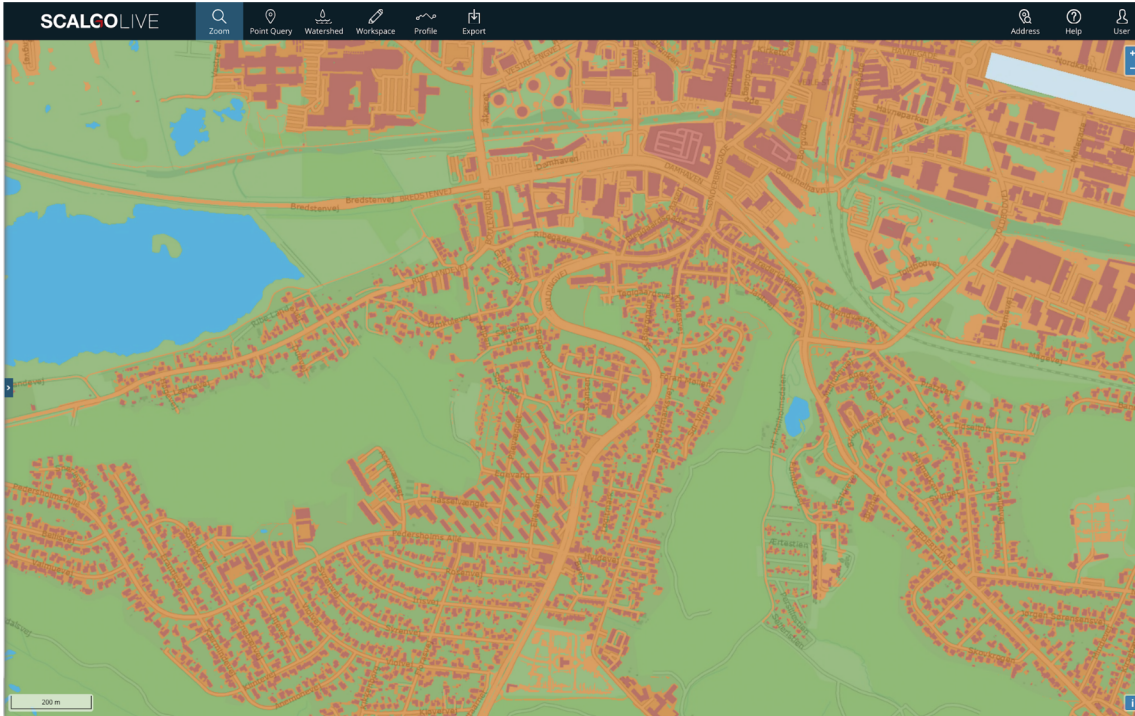


Fra storskala til mere detaljeret model

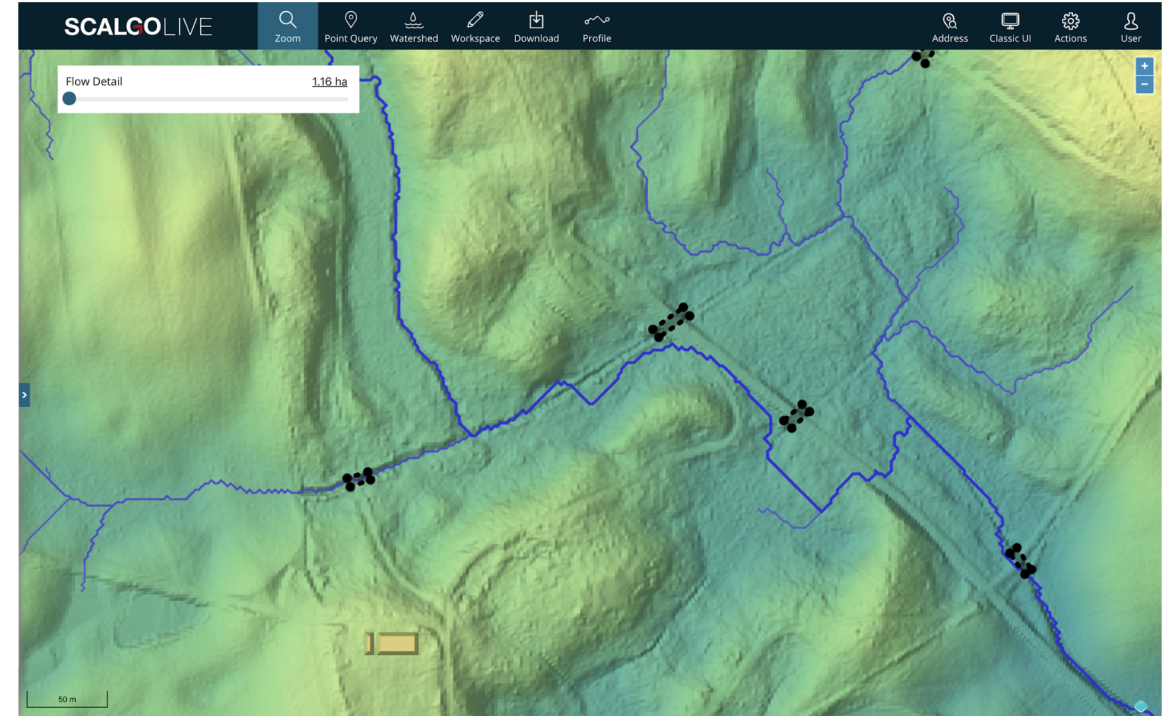


Kan vi anvende store mængder geografiske data på tværs af landegrænser til at forbedre vores model af den fysiske verden?

Konkrete eksempler på værdiskabelse med ML



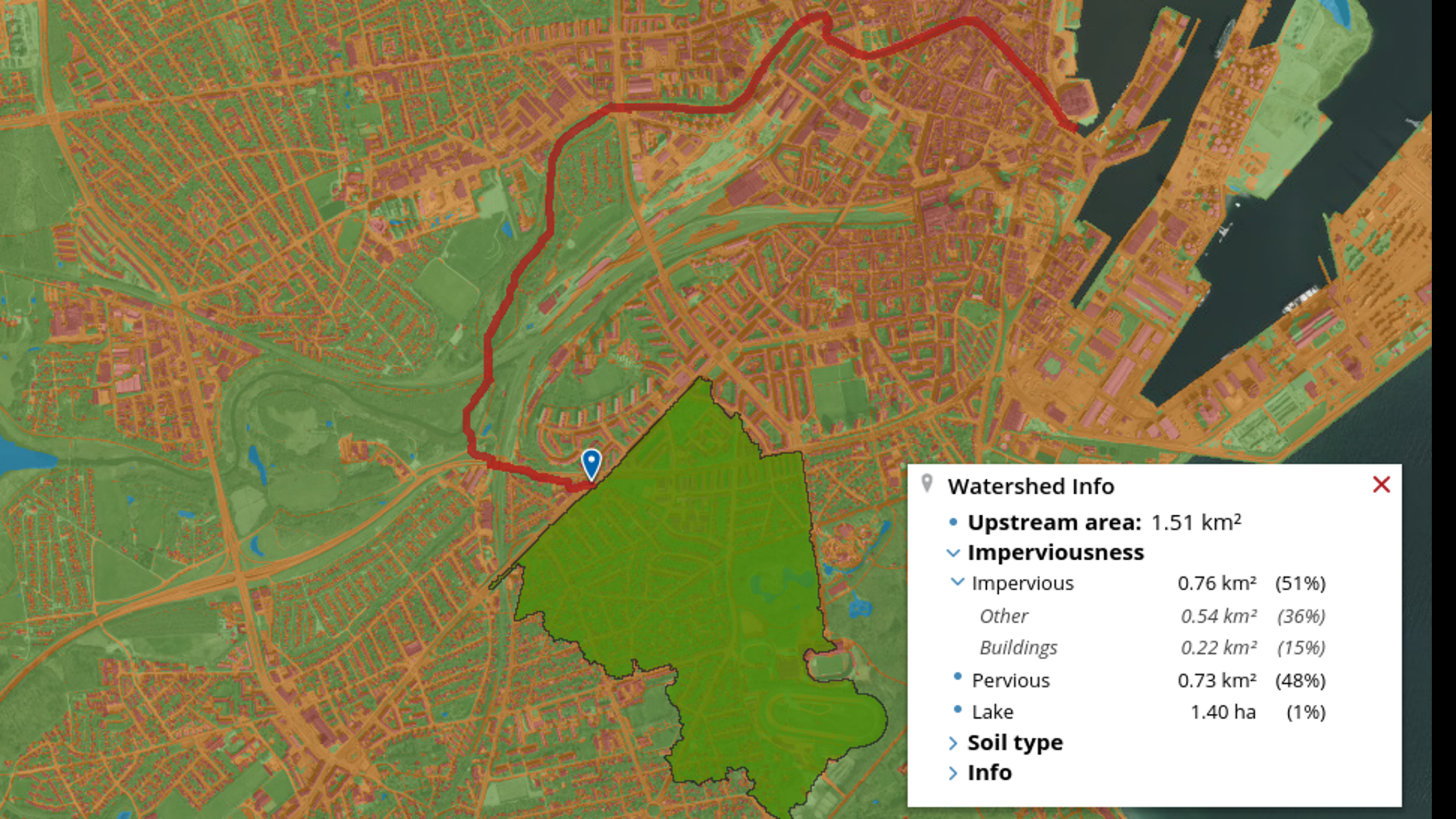
Befæstelseskort



Hydrologiske tilpasninger

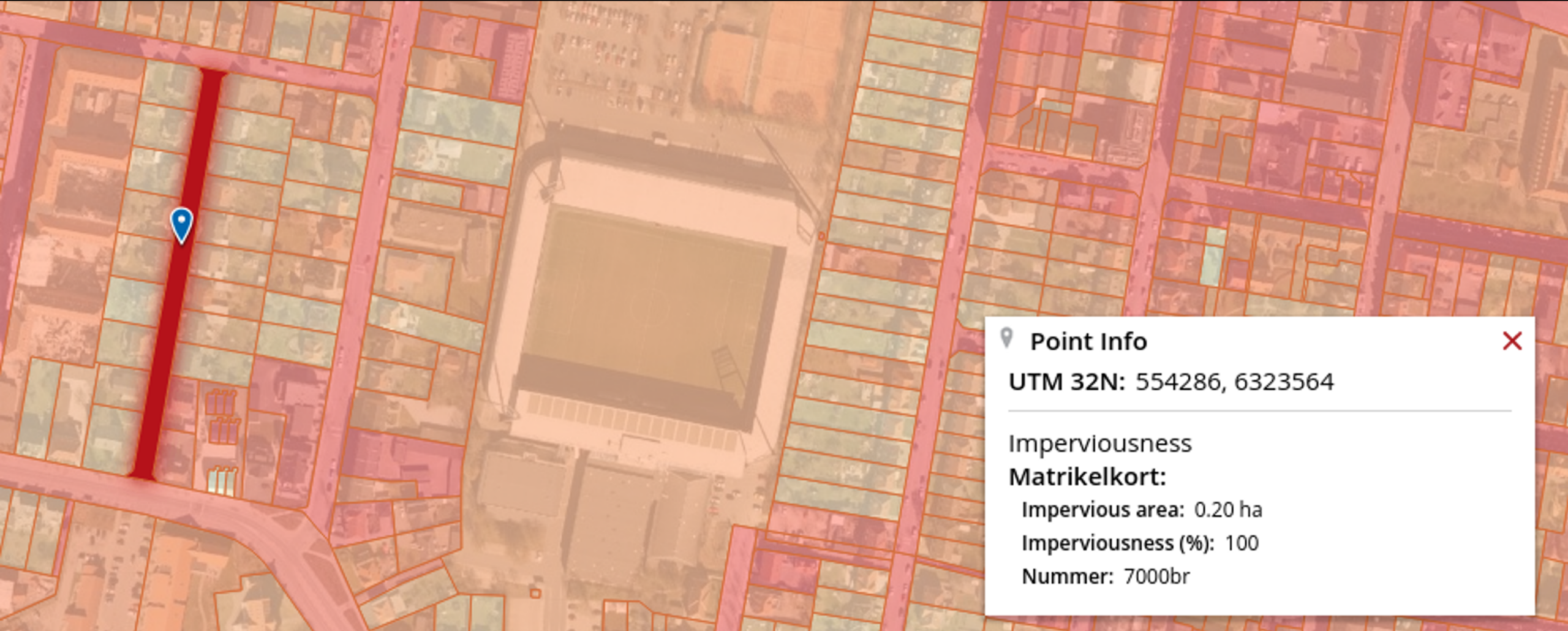
Værdiskabelse

- Forbedret automatiseret modellering i nationalskala
- Hurtigere opsætning af lokale hydrodynamisk oversvømmelsesmodellering
- Bedre information i det daglige arbejde i kommuner og forsyninger



Watershed Info ✕

- **Upstream area:** 1.51 km²
- ✓ **Imperviousness**
 - ✓ Impervious 0.76 km² (51%)
 - Other* 0.54 km² (36%)
 - Buildings* 0.22 km² (15%)
 - Pervious 0.73 km² (48%)
 - Lake 1.40 ha (1%)
- > **Soil type**
- > **Info**



Point Info ✕

UTM 32N: 554286, 6323564

Imperviousness
Matrikelkort:
Impervious area: 0.20 ha
Imperviousness (%): 100
Nummer: 7000br

To måneders forbrug

- > 12.000 minutter
- > 500 forskellige brugere
- > 120 downloads

Vigtige ingredienser i Machine Learning

- **Håndtering, indsamling og adgang til store datamængder**
 - Både input og output
- **Investering i beregningskraft**
 - Det kræver massiv beregningskraft at beregne i DK (og i Nord Europa)
- **Teknisk know how**
 - Udvikling af high-performance ML algoritmer er et specialisterområde
- **Proces frem for produkt**
 - Ikke et IT produkt som bliver specificeret og leveret
 - En proces som integreres i produkt og platform og løbende forbedres

Proces - befæstelseskortet

Evaluer på hele
danmark
(måske release)

Skaf mere
træningsdata
(annotering)

Processering af
eksisterende kilder.
Annotering af
orthophotos for nye
data (Special-minds)

Evaluer på mindre
set

Både manuelt og på
validerings sæt

Manipuler trænings sæt

Find attribuering af fejl i
træningssæt.

Træn ny model(s)

Typisk enkelt model for
indre løkke og en
ensemble for ydre.

Data kilder - Input features (full model)

- LiDAR point clouds: Normalized-DSM*, Return-Num, Classification
- Orthophoto billeder : RGB+N (fra to forskellige år for bedre at dække områder i skygge)
- Vector features: vej-midte, vej-kant

Data sources - Input features (ortho model)

- Orthophoto: RGB+N fra et år

* $\log(\text{clip}(\text{DSM} - \text{Gaussfilter}(\text{DTM})) + 1)$

Data kilder - targets

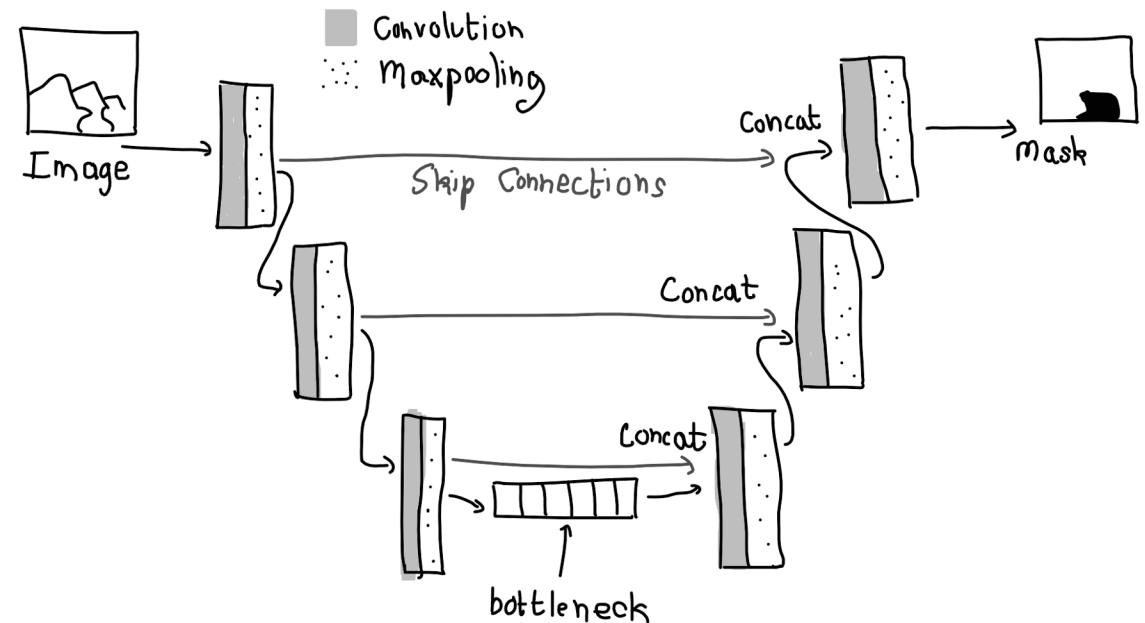
Nuværende trænings-set indeholder 1.8 milliarder pixels.

- Komplette labels fra to kommuner (Skanderborg og lejre).
- Genereret set af alle veje (fra vej-kant / vej-midte). Ikke specielt, kræver manuel inspection.
- Alle marker i Danmark
- Alle bygninger i Danmark
- Alle skove
- Over 4000 parkeringspladser genereret på baggrund af linjestykker af geodanmark af SpecialMinds (ca. 500 i brug).
- Omkring 50 anoterede baggårde (SpecialMinds).
- Udvalgte ikke-befæstede områder (strande / hede).
- På baggrund af dette laves trænings polygoner.

Model

UNET baseret “Convolutional Neural Network” med tricks.

- Bruger “depth wise separable convolutions” for at begrænse antallet af parametre uden at miste performance.
- Bruger “skip-connections (ResNet)” for bedre flow af gradienter.
- Bruger “atrous-convolutions” som øger det perceptuelle felt ved at downsample direkte i filteret.
- Er gået fra model med 1.6 mil parametre -> til omkring 6 mil parametre -> 38.5 mil parametre. (4x i ensemble)



Annotering (special minds)

- Behov for “nålestiksoperationer” i træningsdata.
- Platform for produktion af polygoner til træningsdata.
- Task baseret flow, der giver både os og brugeren flexibilitet.
- Semi-automatiseret kvalitetssikring (statistik baseret).
- Live Demo ([MI Annotation \(scalgo.com\)](https://scalgo.com))



(Målbar) Performance

Skovby

Lejre

Strand

	accuracy	recall	precision	recall false	precision false	auc	f1
Ny (2021-2)	0.9530	0.9543	0.8822	0.9526	0.9824	0.9908	0.9168
Gm (2020-11)	0.9390	0.9549	0.8415	0.9331	0.9823	0.9852	0.8947
SDFE	0.9075	0.9220	0.7781	0.9021	0.9688	0.9121	0.8439
Ny (2021-2)	0.9665	0.9577	0.9095	0.9693	0.9862	0.9944	0.9330
Gm (2020-11)	0.9452	0.9524	0.8428	0.9428	0.9840	0.9880	0.8943
SDFE	0.9381	0.8268	0.9108	0.9739	0.9458	0.9003	0.8667
Ny (2021-2)	0.9249	0.7353	0.0036	0.9250	0.9999	0.9403	0.0072
Gm (2020-11)	0.5967	0.7085	0.0006	0.5966	0.9998	0.7920	0.0013
SDFE	0.2824	0.8946	0.0005	0.2822	0.9999	0.5884	0.0009

Eksempel 1

SDFE

NY (2021-02)

GM (2020-11)

