

Nationalt befæstelseskort til brug for klimatilpasning

Af Mogens Skov og Rasmus Johansson, Styrelsen for Dataforsyning og Infrastruktur

Abstract

Danmark ser ind i en fremtid med hyppigere skybrud og flere oversvømmelser. Detaljerede data og viden om overfladens belægning og dermed nedsivningsevne kan bidrage til en intelligent håndtering af regnvand. Styrelsen for Dataforsyning og Infrastruktur (SDFI) har udviklet en prototype af et nationalt befæstelseskort. Et nationalt befæstelseskort giver fx kommuner et efterspurgt datagrundlag til arbejdet med klimatilpasning og grøn omstilling.

Vi har siden 2020 arbejdet med at udvikle et nationalt befæstelseskort. Arbejdet har løbende været beskrevet og kan følges på SDFI LABS. Denne artikel fokuserer på de seneste udviklingstiltag og på de kommende aktiviteter i projektet. Vi ønsker bl.a. med AI at styrke muligheden for at generere nye og frie data ved intelligent brug af eksisterende og frie geodata.

Keywords: Befæstelse, machine learning, Artificial Neural Networks, extrabytes, træningsdata, frie data.

1. Jorden er ikke vandtæt

Danmark ser ind i en fremtid med hyppigere skybrud og flere oversvømmelser. For at undgå store skader er det vigtigt at kunne modellere, hvor vand løber og ender. Til dette benyttes allerede i dag Danmarks Højdemodel, som kan sige noget om, hvor vand løber hen. Danmarks højdemodel kan dog i dag ikke sige noget om, hvad overfladen er lavet af, hvilket er en vigtig parameter i forhold til nedsivning og hastigheden af vandstrømninger. Dette er vigtigt for at beregne, hvordan regnvand belaster afløbssystemet og påvirker grundvandet.

Udviklingen af et nationalt befæstelseskort vil reducere tidsforbrug og udgifter til planlægning og prissætning af gravearbejde ved at skabe overblik over de forskellige belægningstyper, da det har stor betydning, om der graves i asfalt, grus eller fliser. Dertil giver befæstelseskortet den enkelte kommune mulighed for at følge udviklingen i det grønne byrum som en indikator for, om kommunen bliver mere eller mindre grøn.

For at udvikle et godt datagrundlag har vi siden 2020 arbejdet med befæstelsesdata og årligt genereret opdaterede landsdækkende datasæt, som bl.a. kan ses på [SDFI LABS](#). I Digitaliseringsstrategien fra 2022-2025 har SDFI under initiativet: "Vand fra alle sider" ydermere fået en bevilling til videreudvikling af et nationalt befæstelseskort.

Et konkret resultat af denne strategi er udviklingen af en prototype for et nationalt befæstelseskort. I dette arbejder vi med AI for at styrke muligheden for at generere nye og frie befæstelsesdata ved intelligent brug af eksisterende og frie geodata såsom skråfotos, GeoDanmark ortofotos og LiDAR-data.

Endelig søger vi igennem brugerfeedback og opstart af en referencegruppe at kunne tilpasse prototypen, så den kan fungere som et ensartet beregnings- og beslutningsgrundlag på tværs af kommuner og forsyninger.

2. Ny AI-metode; nye muligheder

Det seneste år har vi arbejdet med AI-metoden, som er ændret fra maskinlæringsalgoritmen, *Random Forests*, der klassificerer et areal baseret på data i punktform, til algoritmen *Artificial Neural Networks*. Denne

klassificerer et areal ud fra arealbaserede data, hvor ikke blot punktet, men også området omkring punktet, tages i betragtning. Dette styrker nøjagtigheden, da inddatakilder som fx ortofotos kan skifte udseende fra år til år alene på grund af lysintensitet, kameratype, solvinkel og flyvehøjde. Mange gange kan to pixler have samme farve, men tilhøre helt forskellige kategorier. Fx kan en grå pixel både være ubefæstet, grus, brosten eller fliser. Ved at vurdere området omkring et punkt sikres en mere homogen klassifikation på tværs af variationer i kildematerialet.

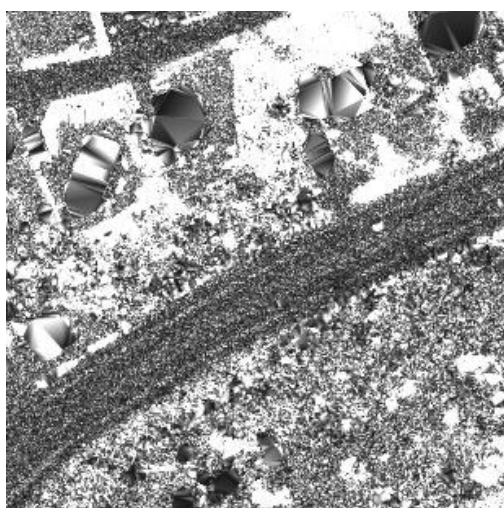
4-kanals ortofotos, der indgår som kanaler i klassifikationen, opdeles i fire enkeltbilleder: Rød, Grøn, Blå og NIR (nær infrarød), se figur 1.



Figur 1: De fire kanaler fra 4-kanalsbillede opdelt i enkeltbilleder – fra venstre: Rød, Grøn, Blå og NIR.

Ortofotos har været brugt længe, men som noget nyt forsøger vi også at styrke metoden med informationer fra højdemodellen og dennes rådata. Fx benyttes nogle af de ekstrainformationer (*extrabytes*), der opsamles ved LiDAR-registreringen, til højdemodellen. Det interessante ved disse ekstrainformationer er, at de siger noget om ruheden på de flader, som rammes, og kan supplere de øvrige kanaler og forbedre bestemmelsen af overfladetypen, se faktaboks om *extrabytes*.

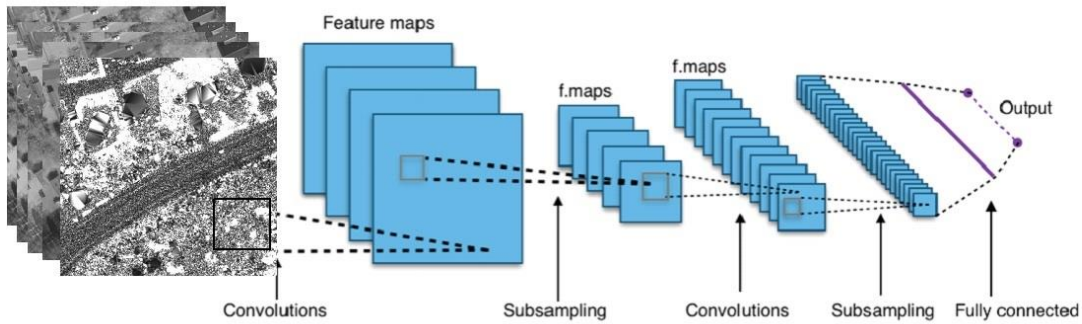
For at kunne benytte ekstrabytes som kilde, laves et TIFF-billede i samme opløsning som de øvrige datakilder, se figur 2.



Figur 2: TIFF-billede af Deviation.

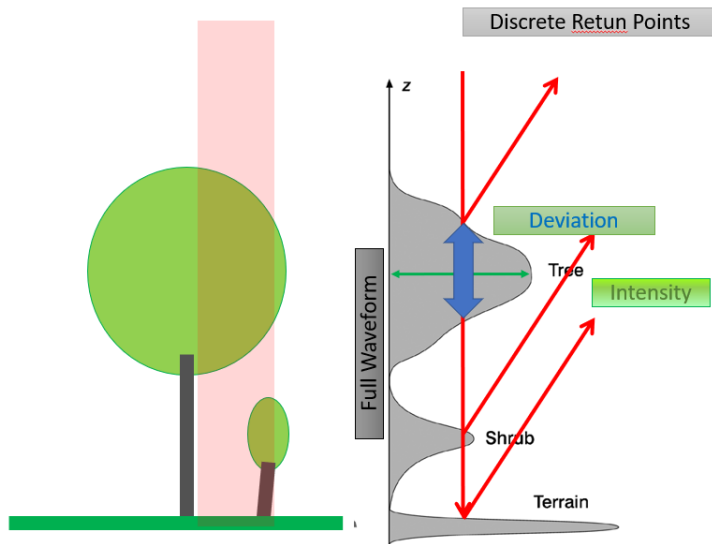
Disse i alt fem billeder sættes sammen til et billede med fem kanaler og benyttes som input til et neuralt netværk (*Unet, convolutional neural network*), se figur 3.

R-G-B-nir-Extrabytes



Figur 3: R-G-B-NIR-Extrabytes – i alt fem billeder benyttes som input til et neuralt netværk.

Faktaboks om extrabytes



Extrabytes som en del af full wave-informationen

I forbindelse med LiDAR-registrering måles det signal, der returneres (return) fra overfladen, der rammes. Den klassiske registrering af first return (= "Tree" på tegning) vil ofte afspejle punktet, hvor laserstrålen har registreret fx toppen af et træ, og last return (= "Terrain" på tegning) vil afspejle, hvor laserstrålen har registreret terrænoverfladen. Herudover registreres der med den type af laserskanner, der nu benyttes, en fuld bølge af den "pulse", som returneres til laserskanneren. Den fulde returbølge ses som det grå areal på tegningen.

Extrabytes er en fællesbetegnelse for flere forskellige informationer, herunder "Deviation" og "Intensity". Deviation (= den blå pil på tegningen) er et udtryk for den pulsbredde eller tid, som laserstrålen har brugt på dens vej gennem et objekt. Vi eksperimenterer i første omgang med at tage Deviation med som kanal for at styrke klassifikationen. Deviation viser også, hvor forskellig det retursignal, der reflekteres fra underlaget, ser ud sammenlignet med, hvordan det ville have set ud, hvis underlaget havde været helt fladt/glat. Det giver altså en indikation af, hvor ru underlaget er. Kender man ruheden, kan det sammen med de øvrige kanalers input være med til at bestemme overfladetyper.

3. Ikke kun til vandmodellering

Allerede nu ses det, at skiftet i AI-metode samt etablering og anvendelse af supplerende data har givet en bedre bestemmelse af, om et areal er befæstet eller ubefæstet. Med ændringen forventer vi at kunne bestemme, om et areal er befæstet eller ubefæstet med en sikkerhed på 92-95 %. Med den nye metode er det ikke alene trærsikkerheden, der forventes at blive bedre. Ses der på de befæstede arealer, vil det også give mulighed for yderligere klassifikation i subklasserne: asfalt, beton, fliser, brosten og grus.

En fordel ved denne subklassifikation er, at en del anvendere som fx graveaktører har stort udbytte af mere detaljeret information om belægningstypen. Med denne type information kan udgifter og planlægning af gravearbejde bedre beregnes. Herudover vil visse brugere alt efter anvendelse fx foretrække at benytte subklassen ”grus” som befæstet, mens andre anvendere vil foretrække at benytte ”grus” som ubefæstet. Begge anvendelser kan lade sig gøre ved at filtrere på attributten subklasse og efterfølgende flytte subklassens data fra klassifikationen befæstet til ubefæstet.

Som det ser ud nu, vil subklassifikationen af de befæstede arealer have en mindre nøjagtighed end klassifikationen i befæstet/ubefæstet, idet det kan være vanskeligt at bestemme, om et areal fx er befæstet med beton eller grus. Som det ser ud nu, forventes det, at subklasserne vil kunne bestemmes med en nøjagtighed på 40-90%, men også, at det vil blive bedre, som tiden går, og nye data inkorporeres.

4. Træningsdata – ”Man får, hvad man træner”

En af de vigtigste elementer ved machine learning er træningsdata. Det er de data, man ved, hvad er, og de data, som algoritmerne benytter som sandhed. Men træningsdata er også et af de dyreste elementer, da der skal være en omfattende mængde træningsdata, og da gode træningsdata kræver en sikker viden om arealernes sande befæstelse – og så er de ofte manuelt skabt.

For at skaffe flest træningsdata for færrest mulige udgifter udnytter vi flere forskellige datakilder, men et centralt datasæt er de landsdækkende skråfotos. Disse fotos har en høj opløsning, og samme areal belyses fra mange vinkler, hvilket forbedrer sandsynligheden for en troværdig manuel vurdering. En række typiske billeder udvælges og benyttes til at digitalisere/udpege befæstede og ubefæstede arealer samt subklasserne af de befæstede arealer.

Selv med gode skråfotos er træningsdata endnu ikke helt i mål, dersom skråfotos som udgangspunkt ikke er *ortorektificerede* – deres koordinater er reelt endnu ikke beregnet. Inden de digitaliserede træningsdata kan anvendes, omregner vi træningspolygonerne via såkaldte stråleligningsberegninger fra billedkoordinater til absolutte koordinater. Dette er desværre ikke en triviell proces, men kræver en vis fotogrammetrisk viden.

5. Med en trænet model

Efter machine learning-modellerne er trænet, kan de anvendes til at finde befæstelser på nye data, nye årgange eller nye arealer. Resultaterne har dog stadig nogle mangler, idet vand og bygninger ikke erkendes, og fordi uddata fra machine learning-processen skal efterbehandles ved bl.a. at fjerne isolerede pixels. Dette gøres ved at overskrive de data, som machine learning-modellen har det svært med, og i stedet at anvende data fra en anden og mere sikker kilde. Helt specifikt er det objekttyperne sø, bygning og kyst/havn, som modellerne har det svært med, og disse hentes fra GeoDanmark. Yderligere suppleres modellerne også med markblokke fra Landbrugsstyrelsen. Uanset hvilket resultat, machine learning-processen var kommet frem til i disse områder, overskrives det respektive areal med de nævnte objekttyper.

Befæstelseskortet leveres i dag som et rasterkort. Dette er til mange formål tilstrækkeligt, men ved brugerworkshops, som vi har afholdt, er det blevet klart, at der er et ønske om at kunne få befæstelsesdata leveret som polygoner i stedet for raster. Polygoner kunne fx lette en beregning af hvor stor en del af en grundejerforening eller et jordstykke, der er befæstet.

Vi har indtil videre valgt kun at distribuere data som raster. Dette valg skyldes dels, at der er ret store forskelle på med hvilken præcision, brugerne ønsker rasterne vektoriseret i. Ulempen ved brugerønskerne til

præcision er, at det ville resultere i en lang række forskellige vektordatasæt i forskellige opløsninger, hvilket vil være kompliceret at generere, vedligeholde og distribuere. Til dels skyldes dette valg også, at der i dag findes standardfunktioner til at lave raster-mod-vektor-analyser i de fleste GIS-programmer.

6. Data til bred anvendelse

Et nationalt befæstelseskort er særligt værdifuldt i områder med tæt bebyggelse, som ofte er de områder, hvor skaden ved skybrud og oversvømmelser er størst. Her vil befæstelseskortlægningen hjælpe til en bedre klimatilpasning og oversvømmelsesrisikovurdering ved at tilføje en forbedret mulighed for at kunne vurdere nedsivning i jorden.

Bydele, byer og kommuner vil kunne bruge information om befæstelsesarealet til at planlægge grønne byrum, til at overvåge udviklingen over tid og som en målbar indikator om at gøre byen mere grøn. Dertil kommer et bedre grundlag for planlægning og budgettering i forbindelse med nedgravet infrastruktur (el, vand, varme og tele) samt dimensionering af rør.

7. Adgang til data

Der er adgang til frie befæstelsesdata, der er genereret på baggrund af den forrige *Random Forests*-algoritme. Data findes som leaflet, WMS-, WCS-tjeneste eller til download via:

<https://dataforsyningen.dk/labs/2259>.

Hvis der ønskes adgang til data, der er genereret efter den nye algoritme, *Artificial Neural Networks*, så kontakt mig gerne via: mos@sdfi.dk.

For de yderligere interesserede i projektets fremdrift i 2023-2024, se faktaboksen om tidsplan.

Faktaboks om tidsplan for landsdækkende befæstelsesdata i 2023-2024

- **Foråret 2023:** Fortsat udvikling af metode og afprøvning af kildedata. Referencegruppe etableres og tilbydes iterationer af lokale data, der er genereret efter den nye metode, *Artificial Neural Networks*.
- **September 2023:** Frigivelse af landsdækkende befæstelsesdata på basis af Geodanmark ortofotos fra 2023, og baseret på den nye AI-metode samt korrigeret efter input fra referencegruppe.
- **2024:** Det afgøres, om befæstelsesdata skal være en permanent produktion af frie data med adgang fra Dataforsyningen, eller om udviklingen og LABS-udstillingen af prototypen afsluttes.