

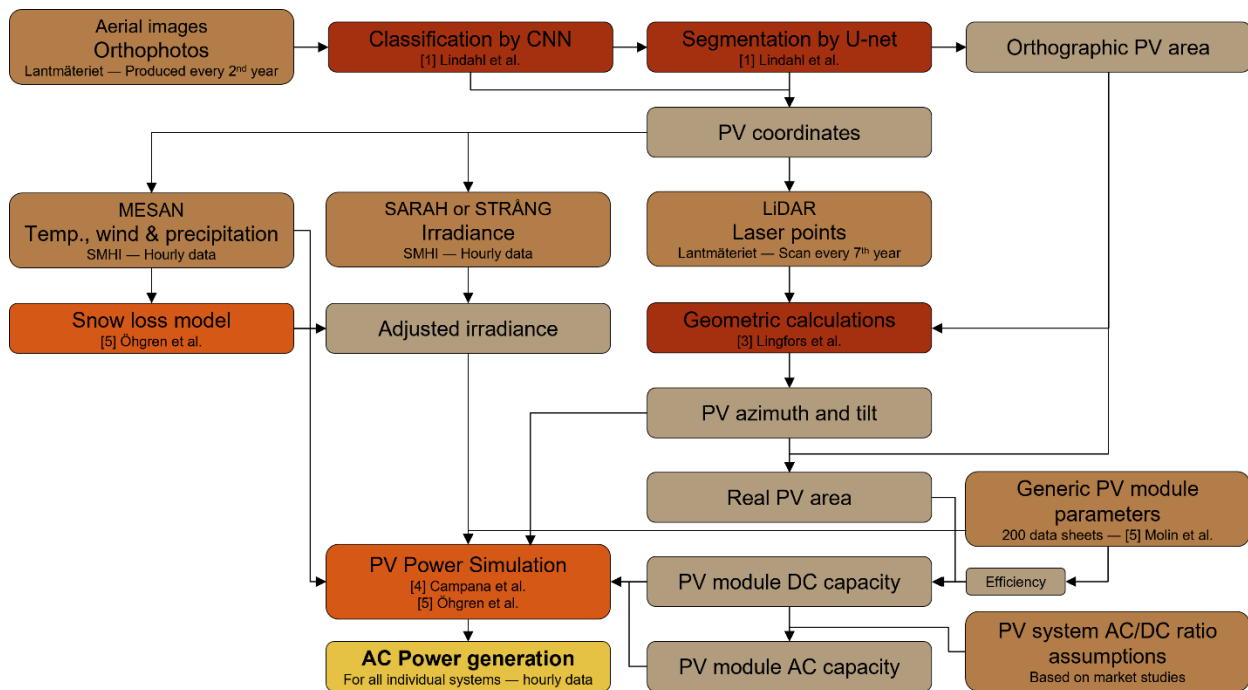
Maskininlärdd flerklassig segmentering av solenergisystem

Noggrannare kartläggning av Sveriges decentraliserade solenergisystem för en djupare insikt i den svenska solenergimarknaden

Bakgrund

Becquerel Sweden har i ett tidigare [forskningsprojekt](#), i samarbete med Uppsala universitet, utvecklat en systemarkitektur, *Alfrödull*, med slutmål att kunna utföra **storskaliga** produktionssimuleringar av decentraliserade solcellsanläggningar i Sverige. Ingångssteget till *Alfrödull* involverar en binär klassifikationsmodell i form av ett faltningsnätverk som kan detektera solenergisystem genom att skanna högupplösta flygfoton över stora områden i Sverige. Klassifikationsmodellen har än så länge använts för att skanna flygfoton över Knivsta, Uppvidinge och Falu kommun mellan år 2019–2024, vilket i kombination med manuell annotering har skapat databaser över alla solenergisystemen inom dessa kommuner, sammanställda i s.k. grundsanningar.

I nästa steg matas de positiva bilderna, dvs bilder på solenergisystem, in i en U-Net semantisk segmenteringsmodell som utför en pixelvis segmentering av bilderna. Detta genererar filtermasker som urskiljer solenergimodulerna från bakgrunden i syfte att uppskatta arean på varje individuell solcellsanläggning. Anläggningarnas areor är essentiell information, både för uppskattningen av solcellskapacitet (peak power) och för att, med hjälp av LiDAR data, kunna beräkna anläggningars orientering (tilt och azimut). När alla hittade anläggningars area och orientering är funna går det att, i samband med en statistisk bas på modulparametrar, simulera elproduktionen på timbasis för varje individuell anläggning i hela kommuner, och i framtiden över hela länder. Med dessa simuleringar kan *Alfrödull* systemet bidra till smartare elnätslösningar och elhandel med allting från lokala elnätsdimensioneringar till timvisa spotprisuppskattningar från decentraliserade solcellsanläggningar.



Figur 1. Schematisk av Alfrödulls modellarkitektur.

I nuläget är segmenteringsmodellen, som utgör det andra delen av ingångssteget till *Alfrödull*, en tvåklassig semantisk segmenteringsmodell. Detta innebär att modellen skattar varje pixel i en given bild med ett sannolikhetsvärde för att bedöma om pixeln tillhör ett solenergisystem eller om den tillhör bakgrunden.

Segmenteringsmodellens prestanda kan mätas på många olika sätt, beroende på modellens ändamål. Tidigare har modellens prestandamått varit *MAPE* (*mean absolute percentage error*), *mIoU* (*mean Intersection over Union*) samt *pixelvis precision och återkallelse*. För en flerklassig segmenteringsmodell skulle precision och återkallelse i stället kunna mätas som objektvis precision och återkallelse.

Eftersom solvärmesystem, från fågelperspektiv i flygfoton, har ett dylikt utseende till solceller har en manuell annotering av solenergisystemets typ varit nödvändig för att kunna urskilja solcellssystem från solvärme.



Figur 2. Exempel på en solvärmepanel.

Vidare finns det betydelsefulla skillnader bland solcellsanläggningar där majoriteten av anläggningar installerade innan 2018 var polykristallina (ikonisk klarblå färg) medan nya installationer efter 2018 har dominerats av monokristallina moduler (svart färg). Att urskilja solcellsmoduler från solvärmesystem, samt inom solcellsmodulerna urskilja dessa två modultyper, skulle både minska det manuella annoteringsarbetet som krävs för utvecklingen av grundsanningar såväl som tillåta noggrannare val av modulparametrar vid produktionssimuleringar. Däri motiveras övergången från en tvåklassig till en flerklassig segmenteringsmodell.



Figur 3. Polykristallin solcell (vänster) och monokristallin solcell (höger).

Syfte och frågeställning

Syftet med detta examensarbete är att generalisera *Alfrödulls* segmenteringsmodell till en modell som, antingen direkt eller indirekt genom klassisk bildbearbetning, kan kategorisera olika solenergisystem som tillhörande följande fyra klasser: solvärme, monokristallin solcell, polykristallin solcell samt bakgrund (under tröskelvärde på alla andra klasser).

Den huvudsakliga frågeställningen i detta projekt är följande:

- Hur påverkas modellens förmåga att segmentera solenergisystem, mätt i mått som *MAPE* och *mIoU*, när segmenteringsmodellen utökas från en tvåklassig till en flerklassig modell?
- Vad får en flerklassig modell för prestanda med avseende på kategorisering av solenergisystem, mätt i *precision* och *återkallelse*?
- Kan för- och efterbearbetningssteg, mätt i *precision* och *återkallelse* per energityp, effektivt stödja segmenteringsmodellen för att korrekt kategorisera solenergisystem i olika typer?
- ...
- Hur väl fungerar alternativa metoder, som bildanalys baserad på RGB-spridning, i att korrekt identifiera och klassificera solenergisystem, jämfört med U-net modellen i en kvantifierad analys av t.ex. *precision* och *återkallelse* per klass? (om tiden finns)

Avgränsningar

Studien begränsas till *Alfrödulls* segmenteringssteg och kommer därför enbart orientera sig kring utvecklingen av segmenteringsmodellen samt omkringliggande för- och efterbearbetningssteg till segmenteringsmodellen. Klassifikationsmodellen och vidare konvertering från segmenteringsmasker till polygon kommer således inte undersökas i denna studie.

Metod och material

Preliminära metoder planeras av studenten med inspel från Becquerel Sweden innan projektstart.

...

System och metoder som ska användas

Kodbasen som använts för att träna och testa den tvåklassiga U-net modellen kommer göras tillgänglig vid projektstart. I projektet kan studenten förväntas använda sig av följande mjukvara:

- Python i Pycharm eller liknande IDE
- PyTorch samt TorchVision maskininlärningsbibliotek
- ...
- ...

Data

Becquerel Sweden tillhandahåller ett dataset på 9000+ bilder över svenska decentraliserade solenergisystem varav 953 bilder är kategoriserade som solvärme. Vidare har ingen uppdelning mellan mono- och polykristallina solcellssystem utförts vid nuläget. Högst troligt är majoriteten av bilderna på monokristallina solcellsanläggningar då dessa system är bättre lämpade för kallare klimat.

Vidare finns det ett tyskt segmenteringsdataset samt ett liknande franskt dataset bestående av bilder på solcellsanläggningar med totalt 4028 respektive 13 303 bilder i vardera. Varken det tyska eller det franska dataseten har kategoriserat anläggningarna som mono- eller polykristallina. Dessa dataset är betydligt äldre och från varmare klimat, vilket innebär att många av dessa bilder troligtvis är på polykristallina anläggningar.

Sammanfattningsvis består tillgängliga dataset av följande:

- Svenskt segmenteringsdataset på 9000+ bilder varav 900+ bilder är på solvärmesystem (resterande solcellssystem)
- Tyskt segmenteringsdataset på 4028 solcellssystem.
- Franskt segmenteringsdataset på 13 303 solcellssystem.

Tidsplan

Tidsplanen sätts av studenten innan projektstart.