Datum
2023-04-21

Dnr
2020-001614

Projektnr
50265-1

Energimyndighetens titel på projektet – svenska Automatisk kartläggning av solpaneler och generering av solelsprognoser från flygbilder och maskininlärning	
Energimyndighetens titel på projektet – engelska Automatic mapping of solar panels and generation of solar forecasts through aerial imagery and machine learning	
Universitet/högskola/företag Becquerel Sweden AB	Avdelning/institution -
Adress Staffansvägen 14, 741 42 Knivsta	
Namn på projektledare Johan Lindahl	
Namn på eventuella övriga projektdeltagare Robert Johansson, Amelia Oller Westerberg, Sofia Ekbring, Sara Ericson, Lisa Molin, David Lingfors, Azeddine Frimane, Joakim Munkhammar	
Nyckelord: 5–7 stycken Decentraliserad solenergi, fjärranalys, maskininlärning, orientering, solelproduktionssimulering, socioekonomisk analys	

Förord

Detta forskningsprojekt har utförts Becquerel Sweden AB tillsammans med Uppsala universitet med hjälp av underkonsulten Competic AB mellan augusti 2020 och april 2023. Dr. Johan Lindahl vid Becquerel Sweden har varit initiativtagare och projektledare för projekt och har tillsammans med Amelia Oller Westerberg, underkonsulten Robert Johansson vid Competic AB och examensarbetarna Sofia Ekbring, Sara Ericson och Lisa Molin varit de som bidragit i arbetet från Becquerel Sweden. Från Uppsala universitet har Dr. David Lingfors, Dr. Âzeddine Frimane och Dr. Joakim Munkhammar alla varit involverade och bidragit till projektet.

Deltagarna i projektet tackar tacksamt för finansieringen från Energimyndigheten (bidragsnummer P50265-1). Utöver finansiering från Energimyndigheten har projektet delvis finansierats med hjälp av interna resurser och medel hos Becquerel Sweden AB.

Dessutom vill projektdeltagarna tacka Nicholas Etherden på Vattenfall R&D och Oscar Willén på Falu Energi & Vatten AB för hjälpen med att korskontrollera våra resultat mot deras databaser.

Innehållsförteckning

Sammanfattning 5
Summary 6
Inledning/Bakgrund 7
Solenergimarknaden 7
Systemidentifiering genom fjärranalys och djup maskininlärning 8
Kartläggning av solpaneler från flygbilder och maskininlärning i Sverige 9
Vidare metodutveckling inom projektet 10
U-net modell 10
Orienterings- och skuggningsmodell 10
Elproduktionssimulering 10
Projektmål 11
Genomförande 11
Skanning av ortofoton 11
Klassificering 13
CNN-klassificeringsalgoritmen 14
Framtagandet av en grundsanning 15
Utvärderingsmetoder 16
Träningsdata 19
Socioekonomisk analys 19
Kommunernas egenskaper 19
Socioekonomiska data 20
Variabler och korrelationsanalys 22
Automatisk segmentering av solenergianläggningar 23
U-net-arkitekturen 24
Modellutvärdering 24
Mjuk- och hårdvara 25
Beräkning av solenergianläggningarnas orientering 25
Justering av solenergisystempolygonen 26
Orienteringsberäkning 27
Undantagsfall 28
Simulering av producerad energi 28
Resultat 30
Kartläggning av solpaneler utifrån flygbilder och djup maskininlärning 30
Icke-detekterbara solenergisystem 30
CNN-klassificeringsalgoritmen 33
Andra skanningen av Knivsta 37
Statistikgenerering 38
Resultat socioekonomisk 42
Demografiska statistikområden 42
Korrelationsanalys 43
Litteraturanalys 45
Egenskaper hos hushåll med solenergisystem 47
Resultat U-net 48
Resultat för orientering och skuggning 54
Preliminära resultat för simulering av producerad energi 61

Diskussion....63
Faltningsnätverk för statistikgenerering....63
Socioekonomisk studie....65
U-net....66
Orientering och skuggning....66
Publikationslista....67
Referenser....68
Bilaga A....73
Ortofoton....73
Bilaga B....74
Skatteverkets typkoder....74
Bilaga C....76
R-script av U-netmodellen....76

Sammanfattning

Det här forskningsprojektet har utvärderat om det går att fastställa var solcellsanläggningarna i Sverige finns installerade, samt deras effekt och utformning, genom fjärranalys som använder sig av flygbilder och en bildanalysmetod som begagnar maskininlärning.

I ett första steg bevisades att ett faltningsnätverk för identifikation av decentraliserade solcellssystem kunde identifiera 96% av solcellsanläggningarna och 81% av solvärmeanläggningarna i en hel kommun efter att succesivt ha tränats med genererade träningsbilder och bilder från två tidigare skannade kommuner. För att i ett andra steg kunna uppskatta arean hos solenergianläggningarna utvecklades en metod för automatisk segmentering med en prestanda som visade sig vara bättre än vad som tidigare publicerats inom det internationella vetenskapliga samfundet. I ett tredje steg vidareutvecklades en automatisk metod för att uppskatta orienteringen hos de identifierade solenergianläggningarna. I sin nuvarande utformning kunde metoden, som är baserad på nationell tillgänglig LiDAR-data, uppskatta en representativ azimut för minst 90% av solenergianläggningarna och en representativ vinkel för minst 75% av anläggningarna.

I projektet har således de metoder som behövs för att automatiskt skapa databaser över majoriteten av alla solenergianläggningar, deras orientering och uppskattade kapacitet inom stora geografiska områden utvecklats. Den grunddata som genereras genom dessa metoder möjliggör simuleringar av alla enskilda existerande solcellsanläggningars solkraftsproduktion inom ett geografiskt område där hänsyn tas till den faktiska orienteringen hos anläggningar, något som tidigare inte har kunnat göras, varken i Sverige eller internationellt. Sådana aggregerade simuleringar fanns det dock inte utrymme att utföra inom detta forskningsprojekt, utan är nästa naturliga utvecklingssteg. Vid publikationen av denna rapport genomförs dock, genom ett examensarbete, en potentiell simuleringsmodell för sådana aggregerade simuleringar utvärderas, och de preliminära resultaten visar att det går att simulera solelproduktion från villasystem som relativt väl stämmer överens med den faktisk uppmätta solelproduktionen då determinationskoefficienter på mellan 0.7–0.8 har uppmätts för flera referensanläggningar.

Vidare har de metoder som utvecklats inom projektet inneburit att tidigare ej tillgänglig statistik över solenergianläggningar kunnat genereras för de tre kommuner som skannats av. Denna statistik har, utöver en mycket hög geografisk upplösning, även mycket hög detaljnivå gällande byggnadsändamål, vilket gör att marknadssegmentstatistik som överträffar tidigare källor kunnat sammanställas. Vidare kopplades den genererade databasen med identifierade solenergianläggningar till socioekonomisk information och en kvalitativ analys genomfördes kring vilka socioekonomiska faktorer hos hushåll som styr utbyggnaden av solenergianläggningar på småhus.

Summary

This research project has evaluated whether it is possible to determine where solar energy systems in Sweden are installed, their capacity and orientation through remote sensing that uses aerial images and an image analysis tool based on machine learning.

In a first step, it was proven that a convolutional neural network for the identification of decentralized solar energy systems was able to identify 96% of all the PV systems and 81% of all the solar thermal systems in a municipality, after being incrementally trained with generated training images, and images from two previously scanned municipalities. To be able to estimate the area of the solar energy systems, a method was developed for automatic segmentation, with a performance that proved to be better than what has previously been published in the scientific community. In a third step, an automatic method for estimating the orientation of the identified solar energy systems was further developed. In its current shape, the method, which is based on nationally available LiDAR data, was able to estimate a representative azimuth for at least 90% of solar energy systems and a representative tilt for at least 75% of solar energy systems.

Thus, the project has developed the methods needed to automatically create databases covering the majority of all solar energy systems, along with their orientation and estimated capacity, within large geographic areas. The data generated with the help of these methods enables simulations of the solar power production of all individual existing PV installations, with the actual orientation of systems considered, something that has been impossible, both in Sweden and internationally, up until now. There was no room to carry out such aggregated simulations within this research project, but they represent the next natural step in its development. However, at the publication of this report a diploma thesis project validates a potential simulation model for such aggregated simulations, and the preliminary results are that it is possible to simulate PV electricity production from distributed residential systems that agree relatively well with the actual measured solar electricity production as determination coefficients of between 0.7– 0.8 has been calculated for several reference PV systems.

Furthermore, the methods developed within the project made it possible to generate previously unavailable statistics. In addition to a very high geographical resolution, the statistics over solar energy systems also had a very high level of detail regarding building purposes, which means that market segment statistics that surpass previous sources could be compiled. Furthermore, the generated database with identified solar energy systems was complemented with socio-economic information. This enabled for a qualitative analysis to be conducted, regarding the influences of different socio-economic factors of households on the deployment of solar energy systems on detached houses.

Inledning/Bakgrund

Solenergimarknaden

Solenergi är samlingsnamnet för kraft och värme som utvinns ur solinstrålningen. I den här kategorin finns solfångare, som koncentrerar solinstrålningens energi till nyttig värme genom uppvärmning av ett ofta vattenbaserat medium. Dessa system byggs ofta i mindre skala, där systemen vanligtvis monteras på tak eller i anslutning till en byggnad och används för uppvärmning av själva huset eller speciella anläggningar kopplade till byggnaden, till exempel simbassänger. I vissa fall används solfångare också i större skala, för fjärrvärme och industriella tillämpningar. De sistnämnda har vanligtvis en mer komplex design, kräver långvarig värmelagring och behöver ofta arbeta vid högre temperaturer. Globalt uppskattas 407 TWh av världens 260 300 TWh uppvärmning av byggnader ha utgjorts av solvärme år 2020 [1].

Den andra vanliga solenergitekniken är solceller, där energin från solinstrålningen omvandlas direkt till likströmselektricitet genom den s.k. fotovoltaiska effekten i solcellerna. Solcellerna är normalt monterade i moduler, som tillsammans med växelriktare, monteringssystem, annan kraftelektronik och eventuella batterier utgör det som vanligtvis kallas ett solcellssystem. Den globala solcellsmarknaden har växt i en takt som överträffar både solvärmemarknadens tillväxt [2] och de allmänna förväntningarna under det senaste decenniet [3]. Solelproduktionen uppgick till cirka 997 TWh år 2020 [2], vilket motsvarar ungefär 4% av det globala elbehovet. Utbyggnaden av solceller i världen kan vara icke nätansluten, både i pico-system på 0–10 W_p och i system i storleksordningen 10–100 W_p som förser de grundläggande elbehoven för cirka 250 miljoner människor (år 2019) i områden utanför det allmänna elnätet. Solceller kan även vara integrerade i olika fordon eller komponenter för att minska behovet av extern energiförsörjning, de kan vara del i decentraliserade takmonterade system för självkonsumtion och försäljning av överskottsel, och i stora markmonterade solcellsparker byggda för traditionell centraliserad kraftproduktion.

Gemensamt för både solfångare och solceller är deras modulära egenskaper, som i sin tur gör teknologin mycket skalbar. Det har gjort det möjligt, och sänkt risken, för ett stort antal intressenter att köpa, installera och använda solenergiteknologierna över hela världen [4]. Småskaliga solvärmesystem för enfamiljshus, flerfamiljshus, hotell och offentliga byggnader utgör cirka 60% av solvärmemarknaden [1] och cirka 40% av den globala solcellsmarknaden utgörs likaledes av så kallade distribuerade solcellssystem [2].

Samtidigt som skalbarheten är en av fördelarna med solenergi, utgör den en stor utmaning för myndigheter och andra aktörer med intresse att följa och dokumentera solcellsutbyggnaden, både historiskt och i realtid [5]. Det stora antalet decentraliserade, ibland icke nätanslutna, system som installeras innebär i praktiken att det är omöjligt att följa utvecklingen på global skala med någon större noggrannhet. Den administrativa kapaciteten att följa installationen av solcellsanläggningar saknas i de flesta länder, och bedöms orealistisk att uppnå

inom en överskådlig framtid. Därtill förväntas utbyggnaden av solenergi accelerera under de kommande åren, i synnerhet i växande ekonomier där solcellsmarknaden börjat ta fart först nu [3], [6], [7]. Samtidigt är solkraften ett av de kraftslag som bedöms central i den globala energiomställningen. Även solvärmemarknaden förväntas växa under de kommande åren, om än inte lika fort som solcellsmarknaden [8]. Det är således viktigt att utforska och utveckla alternativa metoder för att uppskatta solcellsinstallationsvolymer.

Systemidentifiering genom fjärranalys och djup maskininlärning

Ett potentiellt alternativ till att manuellt registrera solenergisystem är att identifiera dem genom fjärranalys av ortofoton (geometriskt korrigerade flygbilder som sätts ihop för att skapa en skalriktig avbildning av marken), med hjälp av djup maskininlärning [9]. Det har visat sig möjligt att generera databaser över solcellssystem, deras exakta plats [10]–[19], och på vilken typ av byggnad [10]–[20] eller mark de anlagts på [13], [21], med denna metod. Därtill kan den genererade informationen kombineras med socioekonomiska data för vidare analys av utbyggnaden [11], [19], [20]. Till skillnad från konventionell statistikinsamling av nätanslutna anläggningar genom rapportering från elnätsägaren, har metoden även potential att hitta icke-nätanslutna system, samt system som antingen är felrapporterade till nätägaren eller som av andra anledningar skiljer sig från dess inrapporterade dimensioner [10]. Det kan till exempel vara fallet att ett solcellssystems installerade kapacitet kraftigt överstiger den abonnerade nätanslutningen, vilket sannolikt kan förklaras av en elektrisk last innanför elmätaren eller ett batteri.

För att solenergisystem ska kunna identifieras med den här metoden krävs däremot att systemen är synliga från himlen och att de har de karaktäristiska dragen hos ett solcells- eller solvärmesystem. Detta innebär alltså att exempelvis vertikalt installerade moduler eller vissa byggnadsintegrerade solceller inte kan detekteras med hjälp av den här metoden. Tidigare forskning har framförallt fokuserat på generering av träningsbilder [22], [23], metodutveckling och validering [12], [14], [15], [21], [24]–[28], snarare än statistikgenerering. Några projekt har dock använt detta tillvägagångssätt för att generera solcellssysteminventeringar; *DeepSolar*-projektet som upptäckte 1,47 miljoner solcellssystem i hela USA [11], *DeepSolar++*-modellen som genererade data för 383 727 system inklusive dess installationsår [19], och en inventering [13] som kartlade 68 661 solcellsinstallationer i kommersiella, industriella respektive offentliga system på en global skala genom flygfotoanalys mellan juni 2016 och september 2018, samt ett projekt utfört av Cadastre (Nederländernas fastighetsregister och kartmyndighet) som identifierade 156 637 byggnader med solcellssystem i tre Nederländska regioner [10].

I två studier, [17], [18], har dessutom neural nätverksklassificering i flygbilder och 3D-byggnadsdata kombinerats för att generera plats, produktionskapacitet, lutning och orientering för befintliga solcellssystem för att sedan jämföras med det officiella tyska solcellsregistret innehållande fler än en miljon byggnader. Dessutom utvärderades i [16] precisionen för identifiering av solcellssystem ur

flygfoton, samt uppskattades dess installerade kapacitet jämfört med ett ofullständigt datamaterial på 33 114 solcellssystem i Connecticut, USA.

Kartläggning av solpaneler från flygbilder och maskininlärning i Sverige

I det här projektet har tre svenska kommuners totala yta skannats för att undersöka och utvärdera metoden i en svensk kontext samt potentiella fördelar för identifiering av solenergisystem med hjälp av maskininlärning. En betydande styrka med tillvägagångssättet i det här projektet i jämförelse med tidigare internationella studier är upprättandet av en tillförlitlig s.k. *grundsanning* (eng: ground truth*)*, dvs. en närapå komplett lista över samtliga solenergisystem i dessa kommuner. En tillförlitlig *grundsanning* möjliggör en grundlig utvärdering av metodikens precision. Ingen av de tidigare nämnda studierna hade någon tillförlitlig grundsanning att jämföra sina identifierade system med undantag för [14]som uppvisade en precision på 81–86% och så kallad återkallelser (eng. recall) på 84–87% för neurala nätverksklassificeringar av solcellssystem i San Francisco och Boston. Ett stort bidrag till den vetenskapliga litteraturen från detta forskningsprojekt är således en utvärdering av hur ett *faltningsnätverk* (convolutional neural network, *CNN*), i detta fall en CNN-flygbildsklassificeringsalgoritm kan användas i praktiken som ett inventeringsverktyg.

Möjliga tillämpningsområden för den här typen av verktyg är för elsystemövervakning och som en statistisk genereringsmetod för att kunna utföra fördjupade analyser av solenergimarknadsutvecklingen. Dessutom har användningen av flygbilder och djup maskininlärning för att identifiera solvärmesystem, vilket gjordes inom projektet, enligt författarnas vetskap ännu inte utforskats i den vetenskapliga litteraturen.

Metodiken möjliggör även detaljerade och mer omfattande socioekonomiska och demografiska analyser. I det här projektet har ett antal oberoende variabler som privata solenergisystemägares inkomst, kön, födelseursprung och utbildningsnivå analyserats på hushållsnivå. Även parametrar som hur många solenergianläggningar det finns per invånare, per hushåll samt per område som klassas som landsbygd, urban miljö eller stadscentrum studerades.

I det här projektet undersöks alltså metoden att identifiera och klassificera decentraliserade nätanslutna och icke-nätanslutna solenergisystem genom ortofoton och djup maskininlärning. Den allmänt tillgängliga databasen och algoritmerna DeepSolar, som använder ett konvolutionerande neuralt nätverk (engelska: Convolutional Neural Network (*CNN)),* närmare bestämt bildigenkänningsarkitekturen *Inception v3,* har tillämpats på Sverige. Flygfoton från tre svenska kommuner har skannats med hjälp av den här metoden och resultaten från skanningen har validerats genom att upprätta en grundsanning innehållande närapå samtliga solcellssystem i respektive kommun. Den svenska marknaden utgör ett gott studieobjekt eftersom det är en solenergimarknad med förhållandevis höga andelar stationära icke-nätanslutna solcellssystem, små decentraliserade solcellssystem och småskaliga solvärmesystem.

Vidare metodutveckling inom projektet

U-net modell

Som tidigare nämnts användes i det här projektet en redan befintlig klassificieringsmodell från ett annat forskningsprojekt (DeepSolar [11]) för identifieringen av solenergianläggningar. Utöver det, utvecklades i projektet en ny segmenteringsmodell, baserad på en s.k. *U-nätverksarkitektur*, som segmenterar pixlarna i ortofotona [29]. Unät, eller *U-net*, är en arkitektur som kan läggas till ett CNN för snabb och exakt bildsegmentering. I det här projektet utvecklades denna modell för att utföra arealuppskattning av de identifierade solenergisystemen. Areaberäkningen är en viktig faktor eftersom det är det första steget till en potentiell metod för att uppskatta kapacitet hos identifierade solenergisystem.

Orienterings- och skuggningsmodell

Polygonerna som genereras av U-net modellen, kan i ett nästa steg användas för att bedöma orientering och skuggning av modulerna. Från modulernas lutning kan solenergisystemets faktiska aktiva yta uppskattas genom grundläggande geometri. Därför vidareutvecklades i projektet en metodik [30] baserad på *Light detection and ranging*-data, LiDAR, för att uppskatta systemens lutning och azimut. Eftersom byggnader i ortofotona inte är ortogonalt projicerade, används i den vidareutvecklade metoden, LiDAR-data för att justera positionen på solenergisystemet, och för att finna de byggnadsapplicerade solenergisystemens lutning och azimut.

På många platser i världen är egenanvändningen vid soliga dagar redan mycket tydligt synlig som en minskning av efterfrågan på kraft från elnäten. För en effektiv drift av elsystemet kommer betydelsen av att förstå och förutsäga egenanvänd solkraft på en aggregerad nivå öka i takt med att penetrationen av solcellsanläggningar ökar [31], [32]. En av de största utmaningarna med att simulera aggregerad produktion från distribuerade solcellsanläggningar är att orienteringen (lutning och azimut) är okänd i Sverige [32]. Det är obligatoriskt att meddela nätägaren när en solcellsanläggning ansluts till nätet, men det är vanligtvis bara växelriktarens kapacitet som anges tillsammans med driftsättningsdatum och inom vilken fastighet systemet är installerat. Följaktligen saknar nätägarna information om den exakta placeringen samt anläggningarnas orientering, vilket denna del av projektet är ett första steg att försöka avhjälpa.

Elproduktionssimulering

När själva arean, lutningen och azimuten för solcellssystemet är känt kan den installerade effekten uppskattas genom dess geografiska position och genom att anta en typisk effektivitet för solcellsmodulerna och växelriktaren. För detta ändamål har en GIS-OptiCE solkraftproduktionsmodell [33], som har vidareutvecklats av Becquerel och Mälardalens Högskola i ett annat forskningsprojekt (Energiforskprojektet EVIB40501), modifierats för att simulera solkraftsproduktionen. Denna modell använder väderdata från de av SMHI framtagna vädermodellerna STRÅNG [34] och MESAN [35]. STRÅNG- och MESAN-datan täcker hela Norden med en timupplösning. Från mars 2017 har

dessa modeller en rumslig upplösning på 2,5×2,5 km. Därmed är den rumsliga upplösningen av väderdata hög, även om den inte matchar den mycket höga rumsliga upplösningen för platserna för de distribuerade solcellssystemen.

Projektmål

Projektet delades inledningsvis in fyra delmål:

1) Implementera och utvärdera befintlig metod för detektering av solenergi-anläggningar.
2) Insamling av träningsdata från Sverige för att förbättra systemets precision med målet är att systemet detektera minst 95% av solcellsanläggningarna som presenteras för systemet under tester.
3) Utveckla och testa systemet för att kunna skilja på solceller och solfångare samt för att undersöka möjligheten att skilja mellan polykristallina, monokristallina och tunnfilmssolceller.
4) Undersöka om även elproduktionen kan beräknas genom lokala väderdata för utvalda anläggningar.

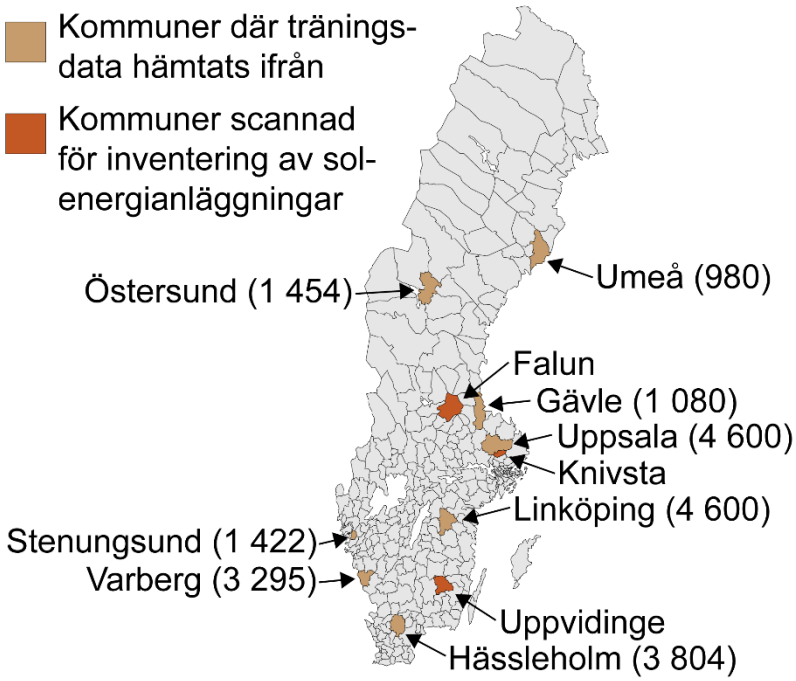
Genomförande

Projektet har genomförts i fyra arbetspaket (AP1-AP4), vilka motsvarar de fyra delmålen. Projektets tekniska delar har utvecklats med ett agilt arbetssätt, där utvecklingen av systemet framskridit med kontinuerlig kontroll av resultaten och där justeringar gjorts i metodiken allt eftersom nya lärdomar dragits.

Skanning av ortofoton

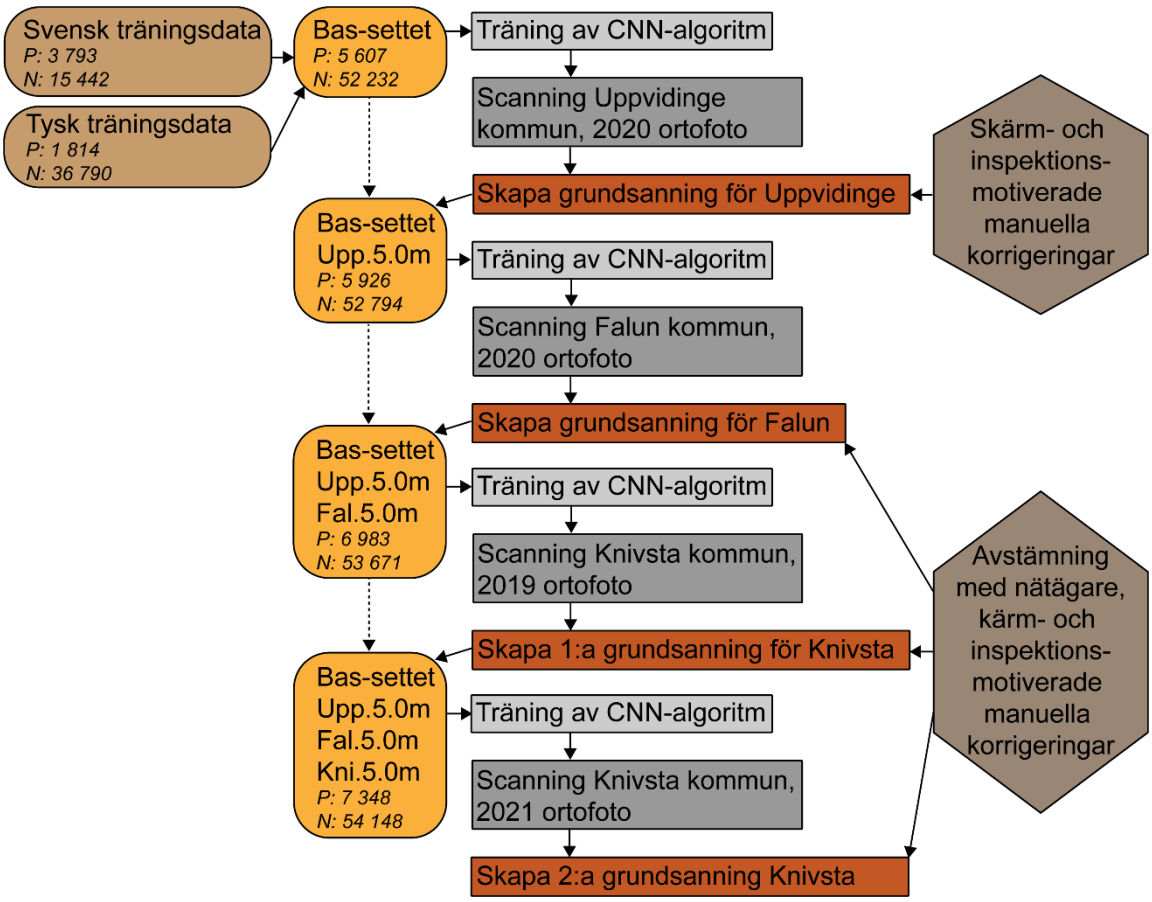
Som nämnt i inledningen är studien baserad på den öppet tillgängliga och välbeprövade klassificeringsgrenen av CNN DeepSolar-ramverket [11]. En något modifierad version av denna tillämpades på tre svenska kommuner i det här projektet, vars ortofoton således skannades av för att identifiera solenergisystem. De tre kommunerna som valts ut för studien är Uppvidinge, Falun och Knivsta (se Figur 1). Urvalet för dessa kommuner beskrivs närmare i *Socioekonomisk analys.*

Python har använts som programmeringsspråk för projektet och programmet QGIS har använts för att visualisera, redigera och analysera insamlad och skapad geospatial data.



Figur 1: Kommunerna (brunmarkerade) som användes för att inhämta träningsdata från, med antalet bildrutor från varje kommun inom parentes, och kommunerna (rödmarkerade) som skannades för att inventera utbredningen av solceller och solfångare.

I Figur 2 illustreras det övergripande arbetsflödet för att skapa databasen, dvs. grundsanningen, innehållande samtliga solenergisystem i respektive kommun. Genomförandet av de illustrerade stegen kommer följaktligen beskrivas i det här kapitlet.



Figur 2: Schematisk illustration av arbetsflödet för att skapa databasen över solenergisystem i de tre utvalda kommunerna. Bokstaven P indikerar antalet positiva bilder och N för antalet negativa bilder.

Varje ortofoto från Lantmäteriet täcker en yta på 2,5×2,5 km, vilket motsvarar 15 625×15 625 pixlar. Detta ger en upplösning på 0,16 m/pixel. En detaljerad beskrivning av ortofotona finns i Bilaga A. Ortofotona delas i sin tur in i rutor (*eng: tiles*) som omformades till 299×299 pixlar från det initiala 115×115 pixlar för att matcha formatet i DeepSolar. Detta ger en slutgiltig upplösning på 0.0615 m/pixel. Flygbilderna, som ligger till grund för ortofotona, togs över Uppvidinge år 2020 antingen 2020-05-31 eller 2020-06-01, de av Falun för 2020 togs antingen 2020-05-21, 2020-06-11 eller 2020-06-14, de av Knivsta för 2019 togs antingen 2019-07-19 eller 2019-07-20 och flygbilderna över Knivsta 2021 togs antingen 2021-06-18, 2021-06-21 eller 2021-08-28.

Klassificering

Klassificeringen av de identifierade byggnaderna och fastigheterna utfördes genom att nyttja officiella klassificeringsregister med standarddefinitioner över både fastigheter och byggnader. För det första klassificerar Skatteverket alla fastigheter i Sverige vid fastighetstaxeringen med en typkod, baserad på karaktäristiken hos och aktiviteten som utförs i byggnaden (se Bilaga A för en detaljerad beskrivning av de åtta huvudkategorierna av typkoder).

För det andra erbjuder Lantmäteriet geodataprodukter inom kartor och geografisk information samt fastighetsinformation. I detta projekt användes; (1) AY — Ytskikt med fastigheter och samfälligheter och (2) BY — Ytskikt med byggnader inklusive dess koder, syften och beskrivningar.

Två olika datamängder kombinerades för att träna CNN-klassificeringsalgoritmen innan den första fullständiga skanningen av en kommun genomfördes. Den första källan är en datauppsättning från delstaten Nordrhein-Westfalen i Tyskland som innehåller 1 814 positiva (dvs. som innehåller ett solcellsystem) och 36 790 negativa bildrutor, med en *uppsamplad* upplösning på 0,05 m/pixel. Den här datauppsättningen är fri att använda och tillgång fås genom OpenNRW_train_16 [25].

Därtill skapades en svensk uppsättning bilder för att kalibrera DeepSolar CNN-klassificeringsalgoritmen [11] till svenska förhållanden. Denna uppsättning skapades genom att använda databasen för det svenska investeringsstödet för solceller, där solcellssystem som fått ta del av det så kallade *solcellsstödet* finns registrerade med sin fastighetsbeteckning. Detta gjordes för åtta kommuner med en god geografisk spridning, se Figur 1. Därefter extraherades ortofotorutor som överlappade dessa system och märktes manuellt som "positiva" om de innehöll ett solcellssystem eller "negativa" om de inte gjorde det. Detta gjordes enligt ett liknande tillvägagångssätt som föreslagits av [22], med två granskare (ingenjörstudenter) som gick igenom samtliga bildrutor. Om båda granskarna var överens om en bildrutas klassificering så märktes den därefter. I de fall det rådde oenighet mellan de två granskarna gjorde en tredje granskare (forskningsprojektets projektledare) den slutliga bedömningen om bildrutan skulle klassas som positiv eller negativ. Avgörandet av den tredje granskaren ägde rum för cirka 0,5% av bilderna, resten var de två ursprungliga granskarna överens om. Antalet bildrutor från varje kommun anges inom parentes i Figur 1. Totalt genererades 3 793 positiva och 15 442 negativa bildrutor med den här metodiken. Det bör noteras att detta endast avser solcellssystem. Någon uppsättning referensbilder för identifiering av solvärmesystem kunde inte skapas på motsvarande sätt eftersom det inte finns något register över installerade solfångare i Sverige.

Bilderna i den tyska träningsuppsättningen, OpenNRW_train_16-setet, skalades ner från 320×320 pixlar till 299×299 pixlar så att storleken på bildrutorna fick samma storlek som de svenska. Som jämförelse är således upplösningen på de tyska bildrutorna 0,0545 m/pixel och den svenska 0,0615 m/pixel. Dessa två datauppsättningar kombinerades för att träna CNN-klassificeringsalgoritmen, och vi hänvisar till denna kombinerade träningsuppsättning som *basuppsättningen* i resten av rapporten. Basuppsättningen innehåller därmed totalt 5 607 positiva och 52 232 negativa bildrutor.

CNN-klassificeringsalgoritmen

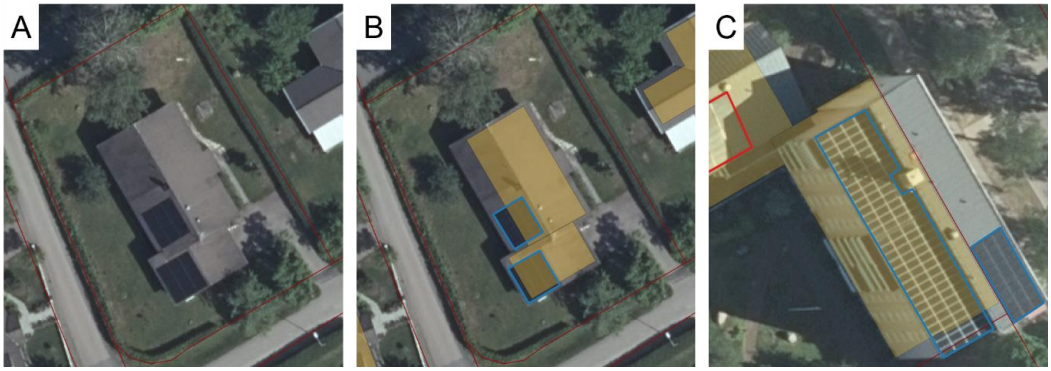
Som nämnts tidigare, användes CNN-klassificeringsalgoritmen som den beskrivs i [11] men modifierad för att köras med PyTorch[1]. Av de tre PyTorch-modellerna

[1] Koden kan laddas ned här: https://github.com/wangzhecheng/deepsolar_pytorch

som finns tillgängliga i [25] och [17] användes *DeepSolar_GoogleMaps _classification.tar* som utgångspunkt, då den presterade bäst på en testkörning med 2000 bilder från Sverige baserad på träningen med basuppsättningen.

Framtagandet av en grundsanning

Det första steget i att skapa en "*grundsanning*" för en kommun är att skanna alla kommunens ortofoton med CNN-klassificeringsalgoritmen. Alla ortofotorutor som klassificerats som innehållande ett solenergisystem lades till i en shape-fil. Denna shape-fil korrigerades manuellt, utifrån visuell kontroll som styrktes med uppgifter om solcellsinstallationer inom specifika fastigheter från databasen för det svenska investeringsstödet för solceller. Platsen och systemgränserna för de identifierade solcells- och solvärmesystemen markerades genom att skapa polygonobjekt för själva panelerna i en separat shape-fil. Dessutom tilldelades polygonobjektet manuellt sin respektive teknologi, d.v.s. om det var ett solcells- eller solvärmesystem. Genom polygonobjektet får de identifierade systemen koordinater och en 2-dimensionell area sett uppifrån. Exempel på identifierade solcells- och solvärmesystempolygoner illustreras i Figur 3.



Figur 3: Exempel på identifierade solcellssystem (blå polygongränser) och solvärmesystem (röda polygongränser) ur ortofoton från Lantmäteriet. I (a) är solcellsmoduler installerade på ett bostadshus huvudbyggnad samt på ett intilliggande garage med AY- Polygonskikt från Lantmäteriet som markerar gränserna för fastigheten, som visas med en tunn mörkröd linje. (b) visar samma solcellsanläggning och byggnader, men med BY-Polygon-skiktet från ©Lantmäteriet markerat med halvtransparenta gula former ovanpå flygbilden. (c) visar en 656systempolygon (röd) och två solcellssystemspolygoner på två olika flerfamiljshus inom samma fastighet.

Genom AY-polygonskiktet för fastighetsbeteckningar, illustrerat med de tunna mörkröda linjerna i Figur 3, kan fastighetsbeteckningen tilldelas solenergipolygonobjekten. På liknande sätt används BY-polygonlagret för att para ihop solenergipolygonobjekten till typkoden för byggnaden som de är installerade på. BY-polygonskiktet visas som ett gult-halvtransparent område i Figur 3b och c.

Vissa solenergisystem är markmonterade. För att hantera dessa system har dessa tilldelats typkoden för den närmaste byggnaden om det finns inom 5 meter (verkligt avstånd) och tilldelas annars typen "mark". Anledningen till detta tillvägagångssätt är en förskjutningsdefekt som ibland uppstår mellan BY-polygonskiktet och ortofotot, specifikt i fall där flygbilden inte togs direkt ovanifrån (tidigare

diskuterats av [10]). En sådan förskjutning illustreras i Figur 3c, där en sida av ett högt flerfamiljshus med balkonger syns på bilden. I vissa fall där klassificeringsmetodiken givit fel resultat, har solenergisystemets polygonobjekt fått korrigeras och tilldelas rätt fastighetsbeteckning manuellt.

Fastighetsbeteckningen är den identifieringsnyckel som används för att koppla de identifierade solcellsanläggningarna till den information som finns för enskilda solcellssystem i solcellsstödets databas samt i den lokala elnätsägarens databaser. De identifierade solcellssystemen korskontrollerades med de två lokala elnätsägarnas interna databaser i Falun och Knivsta kommun. Tyvärr var detta inte möjligt för Uppvidinge kommun, eftersom den lokala elnätsägaren ansåg att detta skulle strida mot deras interna juridiska regler. Genom korskontroll kunde vissa systems teknologi korrigeras (solcell- respektive solvärmesystem) och vissa solcellssystem som inte upptäcktes kunde identifieras och läggas till som polygonobjekt genom att titta i flygbilderna på fastighetsbeteckningen där elnätsägaren hade registrerat ett solcellssystem som inte identifierats av CNN-klassificeringsalgoritmen.

Avsaknaden av ett register över solvärmesystem gjorde det omöjligt att dubbelkontrollera de identifierade systemen på samma sätt som för solcellssystem. Men de solvärmesystem som upptäckts manuellt i flygbilderna, men som inte identifierats av CNN-klassificeringsalgoritmen, lades också till som polygonobjekt.

I fall där förekomst eller typ av teknologi inte kunnat verifieras med säkerhet utfördes fysisk inspektion på plats som ett sista steg och solenergipolygonobjekten kunde korrigerades i vissa fall. De slutliga shape-filerna som innehåller tillagda och korrigerade polygonobjekt över solenergisystem kallas härefter för "*grundsanningen*".

Utvärderingsmetoder

Precisions/noggrannhetsanalyser är centralt i utvärderingen av modeller för djup maskininlärning. Eftersom vi i det här projektet använt oss av en något modifierad version av CNN *DeepSolar*-ramverket, används samma metod som i [11], [25] för att utvärdera klassificeringen, d.v.s. att använda mätvärden baserade på sanna/falska positiva/negativa ortofotorutor, där en sann positiv (eng: *true positive*, TP) motsvarar att CNN-klassificeringsalgoritmen hittar ett faktiskt solenergisystem i en bildruta, en falskt positiv (eng: *false positive*, FP) motsvarar att CNN-klassificeringsalgoritmen identifierat något som inte är ett solenergisystem som ett solenergisystem, en sann negativ (eng: *true negative*, TN) motsvarar att CNN-klassificeringsalgoritmen inte hittat ett solenergisystem i en bildruta som inte innehåller något solenergisystem och falsk negativ (eng: *false negative*, FN) motsvarar att CNN-klassificeringsalgoritmen inte hittat ett solenergisystem i en bildruta som faktiskt innehåller ett solenergisystem.

Från det totala antalet TP, FP, TN och FN utvärderas körningarna för varje kommun i termer av *precision* som definieras som:

$$precision = \frac{TP}{TP + FP} \tag{1}$$

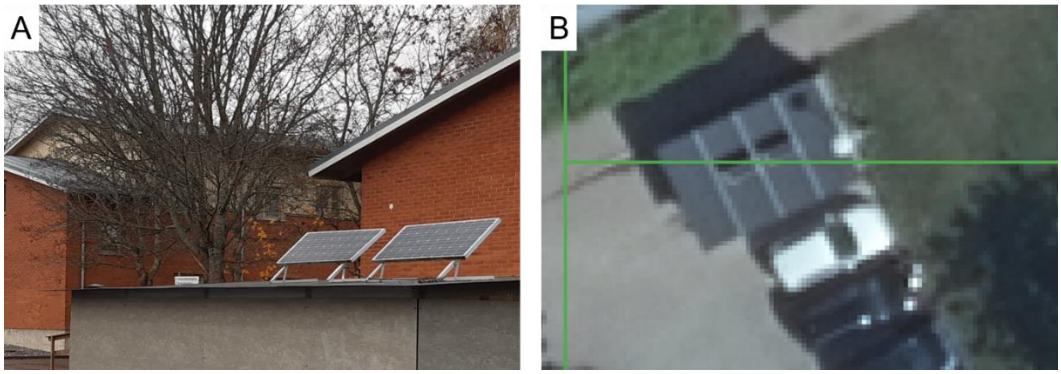
och återkallelsen (eng: *recall*) som definieras som:

$$recall = \frac{TP}{TP + FN} \tag{2}$$

Precision mäter förhållandet mellan sanna positiva bland samtliga positiva identifikationer, medan *recall* mäter förhållandet mellan alla sanna positiva och samtliga positiva prover i dataunderlaget (det vill säga summan av sanna positiva och falska negativa) [11]. Dessutom beräknas även F1-värdet (eng: *F-score),* [10], [15], [24], vilket är ett övergripande mått på en modells noggrannhet definierad som det harmoniska medelvärdet av *precision* och *recall*;

$$F1 = 2 \times \frac{Precision \ \times Recall}{Precision + Recall} \tag{3}$$

Eftersom bildrutorna som utvärderas vid en kommunskanning är placerade sida vid sida är det oundvikligt att vissa rutor endast kommer att visa en mycket liten del av ett solenergisystem. Logiskt sett är det svårare för CNN-klassificeringsalgoritmen att korrekt identifiera ett solenergisystem i en bildruta när endast ett litet fragment av systemet ligger inom bildrutan. I vissa fall visade sig emellertid CNN-algoritmen vara kapabel att identifiera små system som bara innehåller en eller två paneler, till exempel i Figur 4b där den nedre bildrutan endast 0,7 m^2.

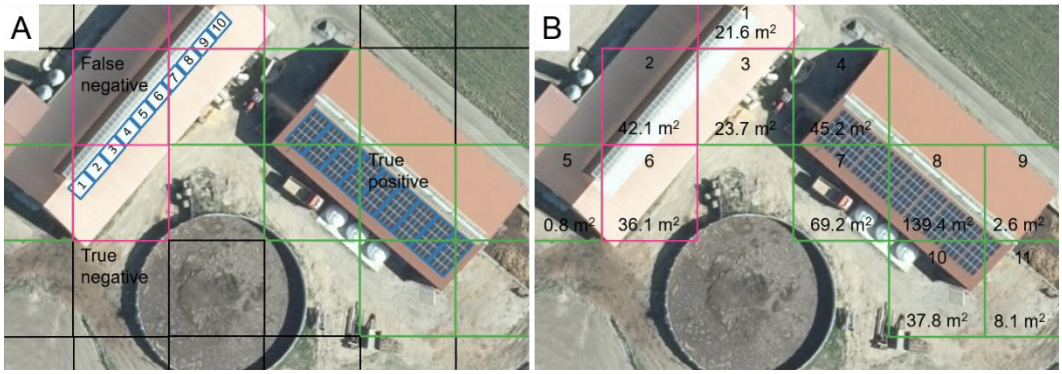


Figur 4: Ett off-grid solcellssystem på en liten sopsortieringsbyggnad bestående av två solcellsmoduler. (A) Visar en bild på systemet taget vid fysisk inspektion och (B) visar dess ortofoto ©, från Lantmäteriet, och de två bildrutorna som överlappar systemet.

För att hantera osäkerheten kring bilder som bara innehåller mycket små delar av ett solenergisystem definierar vi en kategori av bildrutor som ”gränsrutor”. Dessa gränsrutor representerar alla bildrutor som endast täcker 0–0,5 m^2 av ett solenergisystempolygonobjekt från grundsanningen. Gränsrutorna anses i projektet varken vara positiva, eftersom det i många fall skulle vara orealistiskt att förvänta sig att CNN-klassificeringsalgoritmen korrekt ska identifiera en så liten del av en solenergimodul, eller negativa, eftersom de faktiskt innehåller en del av ett

solenergisystem. Gränsrutorna tas därför bort från den totala uppsättningen av bildrutor i noggrannhetsutvärderingen, men listas ändå i resultatet för referens.

Precis som [10] påpekar, är klassificeringsnoggrannhet förhållandet mellan korrekta förutsägelser och det totala antalet bildrutor, och eftersom de flesta bildrutor inte innehåller något solenergisystem skulle hög noggrannhet likväl uppnås även om alla rutor klassificerades som negativa. Så, förutom analyserna av individuella bildrutor (precision, recall och F1-värde), introduceras i projektet en objektbaserad metod, vilket är att föredra vid utvärdering av klassificeringsnoggrannheten för solenergisystemsidentifikation [16]. Eftersom målet med denna studie är att utvärdera möjligheten att använda faltningsnätverk för att generera solenergistatistik och att använda metoden som ett inventeringsverktyg, är det av större vikt att åtminstone en del av ett solenergisystem identifieras än att alla delar av det identifieras. Figur 5 illustrerar ett sådant exempel. I Figur 5a klassificerades alla sex bildrutor som täcker solcellspolygonobjekten på den högra ladugården korrekt som positiva, d.v.s. TP, medan tre av bildrutorna över den vänstra ladugården klassificerades som negativa av CNN-klassificeringsalgoritmen, dvs FN. Solcellspolygonobjekten i Figur 5a markerade 1, 7, 8 och 9 täcks delvis av en bildruta som CNN-klassificeringsalgoritmen har märkt som positiv. Ett polygonobjekt som helt eller delvis överlappas av en eller flera sanna positiva, TP, bildrutor, som i det nämnda exemplet, definieras som ett "detekterad polygon" i vår objektbaserade metod.



Figur 5. Solcellspolygon (blåmarkerade i bild A) på två ladugårdar bestående av flera uppsättningar solcellsmoduler med mellanrum emellan i ett ortofoto © från Lantmäteriet. (A) Exempel på sanna positiva bildrutor (gröna), falska negativa bildrutor (magenta) och sanna negativa (svarta) och tio numrerade solcellspolygoner på den vänstra ladugården och nio polygonobjekt på den högra ladugården. (B) De positiva bildrutorna är numrerade i övre delen av varje bildruta och den totala arean som täcks av ett solcellspolygonobjekt visas i dess nedre del.

Vidare definierar vi ett ’system’ som en grupp av moduler placerade på samma byggnad. Om någon av polygonen på byggnaden överlappas av minst en sann positivbildruta, definieras den som ett ”detekterat system”.

Dessutom kan en fastighet bestå av flera byggnader med ’system’ på flera av byggnaderna. I Sverige ser nätägaren allt bakom en enskild nätanslutningspunkt som ett enda elsystem och generellt sett har en fastighet endast en anslutningspunkt. Således skapas en tredje polygongrupp för att möjliggöra jämförelse mellan de identifierade systemen och nätägarens register över solcellsanläggningar. Den gruppen innehåller samtliga solenergipolygonobjekt inom en och samma fastighet.

På samma sätt som tidigare definieras det som en ”detekterad fastighet” om en sådan fastighetspolygongrupp överlappas av en eller flera sanna positiva bildrutor.

Slutligen kan det objektbaserade polygondetektionsmåttet jämföras med det totala antalet polygonobjekt i en kommuns grundsanning, medan polygongrupperna *detekterade system* och *detekterade fastigheter* jämförs med det totala antalet byggnader respektive fastigheter i en kommun som enligt grundsanningen har solenergisystem.

Träningsdata

Efter upprättandet av grundsanningen för en kommun kan ytterligare träningsdata genereras, eftersom alla bildrutor som täcker kommunen är klassificerade som antingen TP, TN, FP eller FN. För att hitta rätt balans mellan hur stor yta av ett solenergisystem som en positiv bildruta ska innehålla för att det ska vara en bra bildruta att träna CNN-klassificeringsalgoritmen på gjordes en utvärdering efter den första kommunskanningen av Uppvidinge, som använde basuppsättningen som träningsdata.

Utvärderingen gjordes på den andra kommunen (Falun) där CNN-klassificeringsalgoritmen tränades med basuppsättningen plus kompletterande bildrutor från Uppvidinge. Utifrån Uppvidinge kommuns grundsannnings-polygonlager definierades positiva träningsprov som bildrutor som innehöll minst en area X = [0,5, 2,5, 5, 7,5] m^2 av en del, men inte hela, solenergipolygonen. För bildrutor innehållande hela solenergisystemets polygonobjekt, sattes en gräns på minst 0,5 m^2 i alla körningar. Samtliga falska positiva, FP, rutor lades till träningsdatan som negativa bildrutor.

Eftersom X varierades i de fyra körningarna, gjorde så även antalet positiva bildrutor i vardera av de fyra träningsdatauppsättningarna som användes för Falun. För att konkretisera detta i ett exempel innebär det att alla elva positiva (gröna) bildrutor i Figur 5b klassas som positiva träningsprov vid X = 0,5 m^2. Vid X = 2,5 m^2 utelämnas ruta fem från träningen och vid X = 5 m^2 elimineras även ruta nio. De fyra extra träningsdatauppsättningarna genererade från Uppvidinge på detta vis betecknas *Upp.0.5m*, *Upp.2.5m*, *Upp.5.0m* respektive *Upp.7.5m*.

Socioekonomisk analys

Framtagandet av grundsanningen kunde i sin tur användas som underlag för en socioekonomisk delstudie som utförts inom projektet. Eftersom lokala faktorer tidigare har visat sig påverka spridningen av solcellssystem i Sverige [36], bidrog den här demografiska och socioekonomiska studien till valet av kommuner som scannades, vilket som sedan tidigare nämnt slutligen tillföll Falun, Knivsta och Uppvidinge, eftersom dessa tre påvisar en god spridning, såväl till storlek som till egenskaper och geografisk plats i landet.

Kommunernas egenskaper

Falun är en till ytan stor kommun, med en area på 2 040 km^2, belägen i mellersta Sverige. 59 837 invånare bodde i kommunen år 2020 [37]. Kommunen har en centralort, Falun, som utgör ett starkt regionalt centrum, kring vilken tio mindre

omgivande tätorter, eller ”byar”, finns belägna. Tillsammans är dessa 11 tätorter hem för 87,3% av kommunens invånare.

Knivsta är å andra sidan en relativt liten kommun, både med sin yta på 295 km^2 och med sina 19 818 invånare år 2020. Kommunen ligger nära huvudstaden Stockholm, samt Uppsala, Sveriges fjärde största stad. 72,8% av kommunens befolkning bor i de två tätorterna i kommunen, Knivsta och Alsike, som båda kan klassas som typiska pendlarorter. Knivsta innehar även en av de lägsta medelåldrarna bland alla Sveriges kommuner [38].

Slutligen valdes också Uppvidinge som studieobjekt, som i sin tur är en till ytan förhållandevis stor kommun med sina 1 178 km^2 och är belägen i södra Sverige. Den låga befolkningen på 9 449 invånare år 2020 vittnar om en kommun som domineras av jordbruksmark och skog. Kommunen saknar en större stad, men 74,3% bor i tätortsområden, inklusive 2 700 invånare i centralorten Åseda [39].

I projektet representerar Falun således en stor stads- och landsbygdskommun centrerad kring en stor centralort, Knivsta en förortskommun och Uppvidinge en landsbygdskommun. Några generella nyckeltal över de tre kommunerna är sammanfattade i Tabell 1.

Tabell 1: Befolkningstäthet [antal invånare/km^2], medelålder [år], medelinkomst för alla kommunens invånare> 20 år [kr/månad] och sysselsättningsgrad [%] för kommunerna Falun, Knivsta and Uppvidinge år 2020.

	Befolkningstäthet	Medelålder	Medelinkomst	Sysselsättning
Falun	29,2	42,4	27 358	82,6
Knivsta	32,2	36,9	33 408	87,2
Uppvidinge	16,2	43,5	23 617	81,6

Socioekonomiska data

Med hjälp av grundsanningarna, innehållande samtliga solenergianläggningar i dessa tre kommuner, kan uppgifterna om de fastigheter solenergianläggningarna finns installerade på användas för att extrahera socioekonomiska data, och baserat på analyser av denna data undersöka drivkrafterna bakom solenergiutbyggnaden. För att hämta in nödvändiga data används återigen fastighetsbeteckningen som nyckel. Samtliga fastigheter som ägs av företag exkluderas från den socioekonomiska studien, och detsamma gäller hushåll som förvärvat fastigheten efter att flygbilderna togs, eftersom ingen socioekonomisk information om tidigare ägaren då finns tillgängliga. Därtill exkluderades ett par fastigheter kring vilka det rådde osäkerhet till följd av specifika ägarsituationer, som t.ex. att egendomen ägs av någon som bor utanför kommunen, av flera personer som är bosatta på olika adresser eller ägs av flera ägare som samtliga bor inom kommunen men på olika adresser.

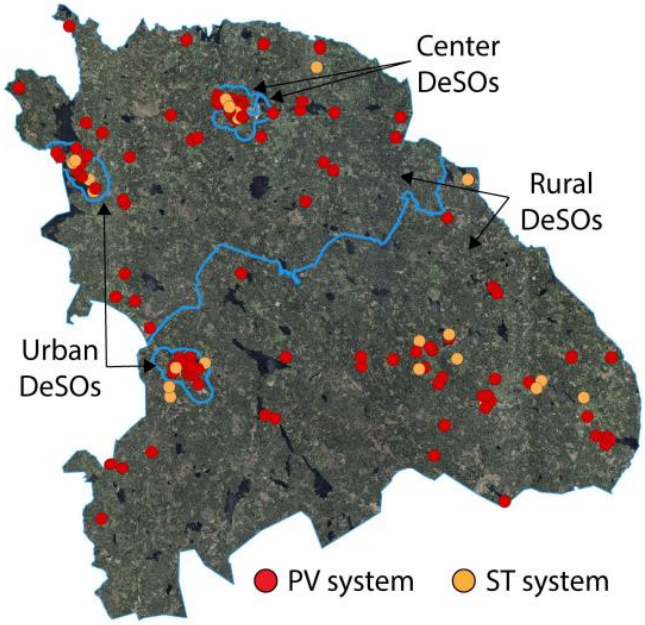
Den socioekonomiska data inhämtades från tre källor; Statistiska centralbyrån (SCB), Lantmäteriet och företaget Ratsit AB som tar fram en årlig taxeringskalender med inkomst och kapitalägande för privatpersoner.

Utöver på hushållsnivå fins vissa demografiska uppgifter tillgängliga på två aggregerade nivåer, där den högsta nivån är kommunal nivå. Den näst högsta aggregerade nivån är, precis som i [40], SCB:s regionala indelning DeSO – Demografiska statistikområden. Storleken på ett DeSO bestäms utifrån befolknings- och byggnadskoncentrationen och kan därför variera kraftigt till ytan, de innehar vanligtvis omkring 1 500 invånare, men kan variera mellan 600 och 3 500 [41]. Det finns tre olika kategorier av DeSO:

A. Landsbygd – DeSO som ligger till största delen utanför större befolkningskoncentrationer eller tätorter.

B. Urban – DeSO som ligger till största delen i en befolkningskoncentration eller tätort, men inte i kommunens centralort

C. Centrum – DeSO som till största delen ligger i kommunens centralort.



Figur 6: Ortofoto över Uppvidinge kommun med DeSo markerade med blåa linjer samt alla solcells- respektive solvärmesystem installerade i kommunen i juli 2020 markerade som röda eller orangea punkter.

Som ett illustrativt exempel på hur DeSO används för att utröna eventuella faktorer kopplade till solenergiutbyggnad visar Figur 6 ortofoto av Uppvidinge kommun, där gränserna för de olika DeSO är markerade med blå linjer. Därtill är samtliga solenergisystem markerade som röda och orangea punkter.

Det totala antalet DeSO och fördelningen mellan de olika kategorierna A-C i de valda kommunerna presenteras i Tabell 2 jämte antal invånare, antal hushåll och medelantalet invånare per hushåll i de olika DeSO-kategorierna.

Tabell 2: Antal DeSO - demografiska statistikområden samt antal invånare, antal hushåll och medelantal invånare per hushåll i de olika DeSO-kategorierna i Falun, Knivsta och Uppvidinge.

	Landsbygd	Urban	Centrum	Totalt
Antal DeSO				
Falun	6	5	26	37
Knivsta	3	0	7	10
Uppvidinge	2	2	2	6
Antal invånare				
Falun	9 641	7 925	41 962	59 528
Knivsta	5 249	-	13 857	19 106
Uppvidinge	3 673	3 052	2 773	9 498
Antal hushåll				
Falun	4 119	3 648	20 222	27 989
Knivsta	2 082	-	5 463	7 545
Uppvidinge	1 724	1 448	1 219	4 391
Medelantal invånare per hushåll				
Falun	2,3	2,2	2,1	2,1
Knivsta	2,5	-	2,5	2,5
Uppvidinge	2,1	2,1	2,3	2,2

I tidigare socioekonomisk solcellsforskning har systemen identifierats genom olika tillvägagångssätt och data analyserat på olika nivåer och med varierande detaljnivå. I den här studien studerades de socioekonomiska faktorerna med en granularitet på hushållsnivå för vissa faktorer och på DeSO- och kommunnivå för korrelationsanalysen.

Variabler och korrelationsanalys

För att utvärdera olika socioekonomiska och demografiska faktorers relation till, och påverkan på, solenergiutbyggnaden i det privata marknadssegmentet behöver koncentrationstermer definieras. Dessa uttryckes i termer av antal system per hushåll och utgör den beroende variabeln i studien. De oberoende variablerna som analyseras i relation till den beroende variabeln valdes baserat på tidigare forskning och tillgång av data och är; (1) Ålder, (2) Kön, (3) Födelseursprung, (4) Utbildningsnivå (5) Sysselsättningsgrad och (6) Andra ekonomiska förhållanden, sammanfattade i Tabell 3.

Ålder, (eng: *Age*), representeras genom fem oberoende variabler motsvarande andel av befolkningen i åldersgrupperna 0–15, 16–24, 25–44, 45–64 och >65 år. Kön, (eng: *Sex*), uttrycks som den procentuella andelen män. Födelseursprung som en oberoende variabel förkortas *BOri* (från eng: *Birth origin*) och representerar procentandelen individer som är födda i Sverige. Utbildningsnivå annoteras *Educ* (från eng: *Education*) och representerar andelen av befolkningen med en eftergymnasial utbildning. Sysselsättningsgraden betecknas *Emp* (från eng:

Employment rate) och uttrycks som andelen av befolkningen som är sysselsatt. Hushållens ekonomiska situation representeras i sin tur av flera oberoende variabler som samlats in från de olika datakällor som anges i Tabell 3. Den första är medelinkomsten, *AInc* (från eng: *Average Income*), för DeSO, definierad som den aggregerade beskattningsbara förvärvsinkomsten, vilket avser inkomst från anställning, företagande, pension, sjuklön och andra skattepliktiga inbetalningar. Den andra variabeln är den genomsnittliga ekonomiska standarden, *EcSt* (från eng: *Economic standard*), som beräknas som disponibel inkomst per konsumtionsenhet.

Tabell 3: Valda socioekonomiska variabler och dess tillgängliga detaljnivå, granularitetsnivå och datakälla. SLS anser i det här fallet Lantmäteriet.

		Granularitetsnivå			
	Detaljnivå	Kommun	DeSO	Hushåll	Källa
Befolkning	Totalt antal hushåll	X	X	-	SCB
	För all med en ålder >20 år	-	-	X	Ratsit
Ålder	7 åldersgrupper	X	X	-	SCB
	För all med en ålder >20 år	-	-	X	Ratsit
Kön	Total distribution	X	X	-	SCB
	För all med en ålder >20 år	-	-	X	Ratsit
Födelseursprung	Sverige, Europa, RoW	X	X	-	SCB
Utbildning	För all med en ålder på 25–64 år	X	X	-	SCB
Sysselsättning	För all med en ålder på 20–64 år	X	X	-	SCB
Medelinkomst	För all med en ålder ≥20 år	X	X	-	SCB
	För all med en ålder ≥20 år	-	-	X	Ratsit
ekonomisk standard	4 grupper För all med en ålder ≥20 år	X	X	-	SCB
Fastighetsändamål	Generell och detaljerad	-	-	X	Lantmäteriet
Fastighetsägare	Beskattad ägare	-	-	X	Lantmäteriet
Taxeringsvärde	Byggnader, mark och totalt	-	-	X	Lantmäteriet

För att analysera oberoende variablernas inbördes korrelation beräknades Pearson-koefficienten för samtliga variabler. En stark korrelation definieras som ett absolutbelopp för korrelationskoefficienten på >0,7, mellankorrelation för värden 0,5–0,7, måttlig korrelation för värden på 0,3–0,5 och svag korrelation för absolutbelopp på 0,1–0,3. Att presentera de beräknade Pearson-korrelationskoefficienter i en matris möjliggör också kollinearitetsanalys mellan de oberoende variablerna. Kollinearitet hänvisar till variablers icke-oberoende och uppstår när två variabler korrelerar på grund av ömsesidiga underliggande faktorer, eller när data som används är helt eller delvis gemensam för variablerna. Om absolutbeloppet av korrelationskoefficienterna mellan två oberoende variabler är högre än 0,7 indikerar det en stark kollinearitet [42].

Automatisk segmentering av solenergianläggningar

För att separera förgrunden (solenergisystem) från bakgrunden (allt annat) på ortofotona utför U-net-modellen en så kallad finkornig klassificering istället för en grovkornig dito, vilket typiskt används av faltningsnätverk för klassificeringsuppgifter. Genom detta förfarande tilldelas en klassetikett till varje pixel med den i projektet utvecklade U-Net-modellen.

U-net-arkitekturen

Det utmärkande för U-net-modellen är dess symmetriska U-form på *encoder-decoder* kaskadarkitekturen, stabiliserad av horisontella "broar" över alla block (eller nivåer). Kodaren (eng: *encoder*) anropar den kontrakterande banan (eng: *contracting path*) där ett antal faltningsprocesser utförs för att fånga objektens kontextuella egenskaper eller särdrag. Avkodaren (eng: *decoder*) anropar å andra sidan den expansiva banan (eng: *expansive path*) där bilden är förstorad till sin ursprungliga storlek så att objekten kan lokaliseras. Ett R-skript för den i projektet utvecklade U-net-modellen återfinns i Bilaga C.

Modellutvärdering

Fyra olika jämförelsemått används för att utvärdera modellprestandan för den här modellen och jämföra den med andra tidigare publicerade metoder och resultat. Det första är det genomsnittliga absoluta procentuella felet (MAPE), som definieras som;

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{S_i - \hat{S}_i}{S_i}\right| \tag{4}$$

där S_i är den manuellt införda arean för solenergianläggningen i ortofotot grundsanningen för bild i, $\hat{S}_i$ är i sin tur den uppskattade arean i samma ortofoto av modellen för samma bild i och N är antal bilder. Det andra jämförelsemåttet är det genomsnittliga absoluta felet (MAE), definierat som:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|S_i - \hat{S}_i\right| \tag{5}$$

med samma parametrar som i ekvation 4. Det tredje måttet är bias (B), som mäter över- respektive underskattning i segmenteringsprocessen, vilket kan beräknas som:

$$B = \frac{\sum_{i=1}^{N} S_i - \hat{S}_i}{\sum_{i=1}^{N} S_i} \tag{6}$$

Slutligen beräknas den genomsnittliga skärningspunkten över unionen (mIoU), vilket beräknas som kvoten av den totala överlappningens area mellan modellens bildsegmentering och grundsanningens manuella mask och den totala arean av båda dessa. För två klasser kan den skrivas som:

$$mIoU = \frac{1}{2}\sum_{k=1}^{2}\frac{TP_k}{FN_k + FP_k + TP_k} \tag{7}$$

där TP (eng: *True positives*) i det här fallet representerar den korrekta areasegmenteringen jämfört med grundsanningens areauppskattning, FP (eng:

False positives) är arean av den inkorrekta segmentering utanför grundsanningens areamarkering och FN (eng: *False negative*) representerar området inom grundsanningsmasken som modellen inte segmenterade korrekt.

Mjuk- och hårdvara

Programmeringsspråket R har använts för det här delprojektet där i synnerhet R-Torch-paketet nyttjats, eftersom det har stöd för snabb tensorberäkning med GPU-acceleration. All nödvändig programvara har paketerats i en Dockerhub[2]. Det ger öppen tillgång till skriptet, som kan köras konsekvent i vilken infrastruktur som helst, utan krav på förkunskaper eller behov av att installera en R-miljö. Se Bilaga C för instruktioner om hur du använder modellen.

Träningen av modellen genomfördes på Alvis-klustret, del av *Chalmers centre for computational science and engineering* (C3SE), specifikt avsett för forskning i artificiell intelligens och maskininlärning. C3SE är ett av sex centrum som bildar *the National Infrastructure for Computing* (SNIC), en nationell datorinfrastruktur som tillhandahåller datorresurser för storskaliga beräkningar och lagring inom alla ämnesområden.

Beräkning av solenergianläggningarnas orientering

Det fjärde delprojektet avser utvecklingen av en modell för vertikala avståndsberäkningar i flygfoton, som möjliggör uppskattning av solenergisystemens orientering (lutning och azimut). Detta kan göras genom att använda LiDAR-data (eng: *Light Detection And Ranging*); genom att laserpulser reflekteras mot en yta kan avståndet till ytan beräknas som en funktion av tiden det tar för laserpulsen att återvända till källan. Tekniken används bland annat inom bilindustrin, men för projektets ändamål används LiDAR-data som samlats in av Lantmäteriet med hjälp av flygplan på 3000 meters höjd över hela Sverige [43]. Det punktmoln av LiDAR-data som samlas in ger en bild av topografin, samt höjd på exempelvis träd och byggnader. I det här projektet används LiDAR-datat för att bestämma orienteringen av solenergianläggningarna.

Eftersom byggnader inte är ortogonalt projicerade i Lantmäteriets ortofoton, används även LiDAR-datat för en ortogonal justering i ett första steg.

Lantmäteriet tillhandahåller flygbilderna både i råformat och som ortofoton, dvs. terrängen i ett ortofoto är justerad (*ortogonalt projicerad*), så att den förefaller vara avbildad rakt uppifrån, trots att så inte är fallet. Någon justering för byggnaders höjd görs dock inte. Detta innebär att i de fall där byggnaderna inte är i nadir relativt kameran i fotoögonblicket, kommer takets placering, även i ortofotot, vara förskjutet. Ju högre byggnad och desto längre från nadir, ju större är förskjutningen (se Figur 3c) och solenergisystem på (framförallt brant) lutande tak kommer vara skjuvade (ej rätvinkliga). För att bättre kunna uppskatta solenergianläggningarnas orientering och form måste först de identifierade solenergisystempolygonerna

[2] Se: https://hub.docker.com/u/frimane

justeras så att de får rätt geografisk placering. Detta görs också med hjälp av LiDAR-data och beskrivs i nästa avsnitt.

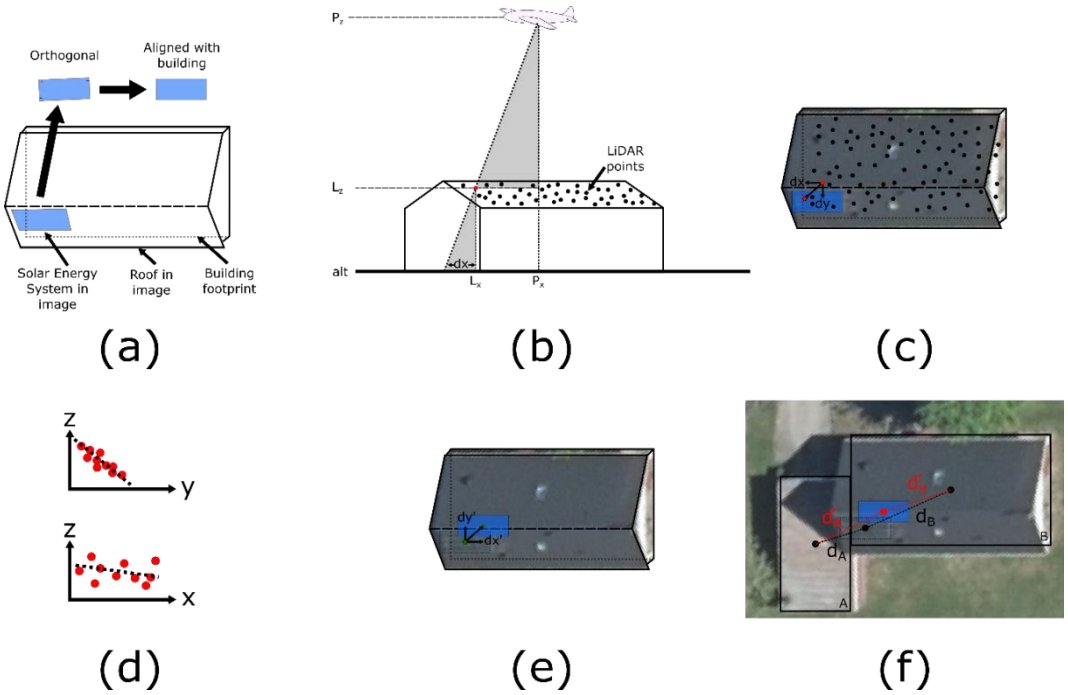
Justering av solenergisystempolygonen

Huvudstegen för att anpassa polygonen till LiDAR-datat illustreras i de fem stegen (a-e) i Figur 7. Eftersom solenergisystemen i de flesta fall lutar och dessutom inte befinner sig i nadir relativt flygplanets kamera när bilden tas, kommer solenergisystempolygonen vara sneda (skjuvade) i bilden, dvs. att hörnen inte är 90˚. Därför är det första steget i justeringen att ge polygonen räta hörn (Figur 7a). För detta antas systemen vara parallella med byggnaden de är installerade på, och således orienterade i samma riktigt som dess huvudsakliga form [30], och polygonen roteras därefter. I det andra steget (Figur 7b) uppskattas avståndet mellan kamerans position (P i Figur 7b) och objektet baserat på LiDAR-punkternas (L i Figur 7b) höjd relativt marken (alt i Figur 7b).

Markhöjden bestäms därefter som medelvärdet av z-värdena för LiDAR-punkterna på marken inom 10 meter från byggnadens ytterkant [30]. Den röda LiDAR-punkten i Figur 7b och c illustrerar processen; i det tredje steget (c), flyttas den med ett avstånd dx och dy i x- respektive y-riktningen. Eftersom de grå trianglarna i Figur 7b är likformiga, kan dx beräknas som:

$$dx = L_z \frac{L_x - P_x}{P_z - L_z} \tag{7}$$

där L_z och P_z är höjden relativt marken för LiDAR-punkten/punkterna respektive flygplanet. De LiDAR-punkter som nu befinner sig inuti polygonen, antas representera dess orientering.



Figur 7: Illustration över anpassningsstegen för form och position för en identifierad solenergisystempolygon.

Orienteringsberäkning

Solenergisystempolygonens lutning beräknas genom linjär-regression, med hjälp av *robustfit*-funktionen i Matlab. I fall då färre än fem LiDAR-punkter hamnar inom polygonen ingår även punkter inom 1 meter från polygonen i regressionsanalysen. Detta eftersom det har visat sig vara en bra balans mellan att ha ett tillräckligt antal LiDAR-punkter för den linjära regressionen, samtidigt som risken är låg för att inkludera punkter som representerar takdelar med en annan orientering än solenergisystemspolygonen, och därmed resultera i en mindre exakt uppskattning av orienteringen.

Solenergipolygonen och LiDAR-punkterna roteras så att polygonen är i linje med xy-planet (för att förenkla illustrationen är exempelbyggnaden i Figur 7 redan i linje med xy-planet). På detta sätt kan den linjära regressionen (Figur 7d) göras i xz- respektive yz-planen, där det lägsta regressionsfelet i respektive plan ger lutning och azimut för solenergipolygonen.

Solenergipolygonen har nu xyz-koordinater. Nästa steg (Figur 7e) är att flytta in polygonen inom byggnadens gränser baserat på dess mittpunkt (*dx'*, respektive *dy'*), på samma sätt som i (Figur 7c), men i motsatt riktning (Figur 7e).

Eftersom höjden relativt marken för LiDAR-punkterna är okänd innan analysen utförts, upprepas steg (a)-(e) för de fem byggnaderna som ligger närmast. I ovanliga fall, är förskjutningen av solenergipolygonen nämligen så stor att det inte är närmaste byggnaden som ska användas för analysen.

Undantagsfall

I vissa fall är LiDAR-datan otillräckliga och då tilldelar modellen solenergisystemet samma orientering som den sydligaste långsidan av närmsta byggnad. Höjden över marken sätts då till 4,26 m och lutningen till 26˚, baserat på medelvärden bland en testuppsättning på 100 solcellssystem från Knivsta.

Om det till exempel inte skulle finnas tillräckliga LiDAR-data i Figur 7f, skulle solenergisystemet tilldelas samma azimut som hus A, dvs väst- eller östsidan beroende på vad som är sydligast, eftersom hus A är närmast ($d_A < d_B$ i Figur 7f). För att undvika sådan felklassificering görs 1) antagandet av byggnadens höjd och 2) en förflyttning enligt Figur 7e alltid som ett försteg till antagandet av orienteringen. I exemplet i Figur 7f skulle mycket riktigt hus B falla ut som närmast ($d'_A > d'_B$) och felklassificeringen undviks. Däremot är detta ett specialfall, i de flesta fall finns tillräckligt med LiDAR-data inom solenergipolygonen. Då kommer lutningen och azimut för solenergipolygonen anpassas efter LiDAR-data enligt beskrivningen i föregående stycken.

Om avståndet mellan solenergipolygonen och närmaste byggnaden är större än 2 m, antas polygonet vara markförlagt. I dessa fall antas azimuten följa polygonens mest sydliga sida och lutningen sätts till 30 grader, vilket är en vanlig lutning. Ett alternativ hade varit att använda LiDAR-punkterna inom solenergipolygonen, men då är risken stor att punkter som representerar omkringliggande (låg) vegetation inkluderas och att regressions-analysen därför ger en missvisande lutning.

Simulering av producerad energi

För att simulera solelproduktionen från de identifierade solcellsanläggningarna utvärderas en modell ursprungligen utvecklande av Mälardalens Högskola [33]och som vidareutvecklats av Becquerel Sweden och Mälardalens Högskola i ett nyligen avslutat forskningsprojekt [44]. Denna modell använder rumsliga och temporala klimatologiska data från Sveriges Meteorologiska och Hydrologiska Institutet (SMHI), närmare bestämt SMHI-modellerna STRÅNG och MESAN. STRÅNG-modellen genererar data om globalstrålning, direkt strålning, fotosyntetisk aktiv strålning och CIE-vägd UV-strålning som täcker större delen av Skandinavien [34]. Modellerna STRÅNG och MESAN genererar kontinuerligt data för hela den nordiska och baltiska regionen med en spatial upplösning på 2,5×2,5 km och en tidsupplösning per timme, medan kommersiella programvaror, så som PVGIS baseras på antingen på interpolerade data (version 4) eller genom att analysera om redan existerande data och analyser (version 5).

I det nyligen avslutade forskningsprojektet [44] byggdes en databas upp som möjliggör för användaren av modellen att ladda ner historiska spatiala väderdata för åren 2018–2022 för alla 2,5×2,5 km stora koordinatpunkter i hela Norden och Baltikum, och baserat på denna väderdata simulera den historiska solelproduktionen för dessa år.

För att utvärdera produktionssimuleringsmodellen för de identifierade decentraliserade solcellsanläggningarna jämförs (vid skrivandet av denna rapport) den historiska simulerade solelproduktionen med en timupplösning med den

faktiska uppmätta solelproduktionen från referenssystem belägna i de tre olika kommunerna i ett pågående examensarbete. Den uppmätta solelproduktionen kommer från några av CheckWatt AB´s kunder, som är villiga att dela datat för detta forskningssyfte. Valideringen av simuleringsmodellen utvärderas med hjälp av några standardmått för statistiska jämförelser; determinationskoefficienten (R^2), medelvärdet av standardavvikelserna (eng: *Root Mean Square Error*)(RMSE), genomsnittligt absolut fel (eng: *Mean Absolute Error*)(MAE) och medelfel (eng: *Mean Bias Error*)(MBE).

Resultat

I följande kapitel beskrivs de resultat och slutsatser som genererats inom ramen för projektet.

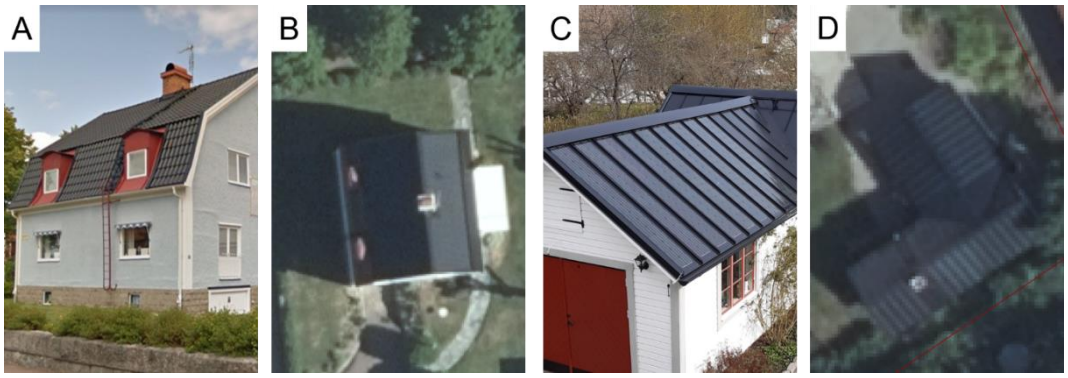
Kartläggning av solpaneler utifrån flygbilder och djup maskininlärning

Som beskrivits tidigare består delprojektet för utvärdering av faltningsnätverk som verktyg för statistikgenerering av ett antal metodiksteg, resultaten av dessa utvärderas och presenteras i följande avsnitt.

Icke-detekterbara solenergisystem

Vid korskontrollen av de solenergisystem som CNN-klassificeringsalgoritmen identifierat och elnätsägarens register visade sig vissa typer av system vara svåra eller omöjliga att identifiera.

Två konkreta systemkategorier kunde identifieras bland dessa icke identifierade system, och anses oidentifierbara med denna metod, nämligen byggnadsintegrerade solceller (BIPV) och vertikalt installerade solenergimoduler. Som namnet antyder så är byggnadsintegrerade celler/moduler designade för att smälta in i den bebyggda miljön och byggnaderna de är installerade på. Det ligger således i denna teknologis natur att de, speciellt från långa avstånd, är svåra att detektera med blotta ögat. Från ett avstånd så långt som på ett flygfoto kan dessa betraktas som omöjliga att hitta, vilket illustreras i Figur 8. Likaså kan vertikalt installerade solenergimoduler inte detekteras i en bild som är tagen uppifrån.



Figur 8: Exempel på byggnadsintegrerade solceller (BIPV) som anses som omöjliga att identifiera med fjärranalys av flygbilder. (A) visar solcellstakpannor från företaget Rustabo AB och (B) ortofotot © över villan i fråga. (C) visar tunnfilmssolceller med samma falsbredd som ett klassiskt falsat plåttak från Midsummer AB och (D) visar ortofoto © över garaget som visas i (C).

Dessa är alltså de två systemtyper som konsekvent inte gick att identifiera, utan fick detekteras med hjälp av elnätsägarens register och med fysisk inspektion. Detta utgör en uppenbar nackdel med metodiken, om den används i syfte för statistikgenerering.

Däremot är förekomsten av de här typerna av system förhållandevis låg. Genom korskontrollen med lokala elnätsägares register över nätanslutna solcellssystem och

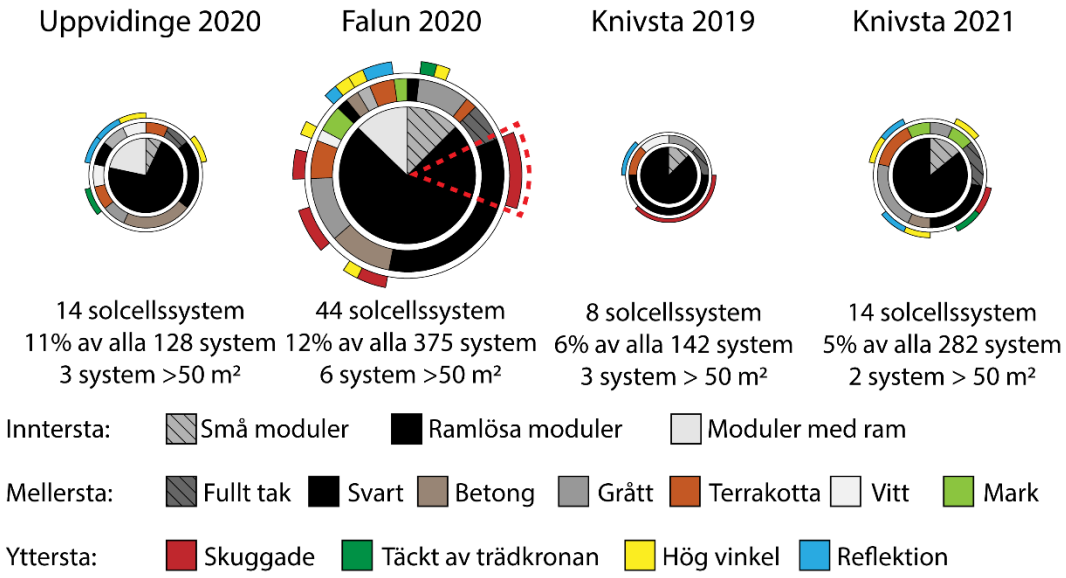
fysiska kontroller i Falun och Knivsta kommun identifierades endast 10 BIPV-system, 5 vertikala nätanslutna solcellssystem och 2 vertikala solvärmesystem. Det innebär att antalet BIPV och vertikala solenergisystem i dessa kommuner är mycket få till antalet jämfört med vanliga *byggnadsapplicerade* takmonterade solenergisystem. De identifierade BIPV och vertikala systemen lades till solenergisystemspolygonlagren i de olika kommunerna och den genererade databasen för solenergisystem, men utelämnades för generering av ytterligare träningsdata och klassificeras således som negativa bildrutor i utvärderingar av CNN-klassificeringsalgoritmen i nästa avsnitt.

Vidare undersöktes samtliga system som inte upptäcktes av CNN-klassificeringsalgoritmen närmare för att identifiera anledningen till felklassificeringen och eventuella gemensamma karaktärsdrag. Först delades de in baserat på modulegenskaper; (1) Små moduler — moduler som i ortofotot tycktes betydligt mindre än normalstora moduler (i nästan alla fall användes dessa i stationära icke nätuppkopplade system), (2) Ramlösa moduler — moduler som inte har den karaktäristiska ramen hos en solenergimodul, och (3) moduler med ram — moduler med en distinkt ram, traditionellt i aluminium.

I ett andra steg klassificerades takets färg i sex vanliga takfärger och en sjunde kategori om det var markmonterat. Dessa var; (1) Fullt tak — solenergimodulerna täcker helt en del av taket så att materialet under inte syns, (2) Svart — svartfärgade tak, (3) Betong — en brungrå färg som är typiskt för betongpannor, (4) Grått — olika nyanser av grått, vilket inkluderar tak som ser ut att vara gjorda av stål- eller aluminiumtakplåt, (5) Terrakotta — färgen på traditionella tegelpannor, (6) Vitt — vita eller mycket ljusgrå tak, och (7) Mark — för markmonterade solcellssystem.

Dessutom identifierades fyra andra gemensamma försvårande karaktärsdrag för de oupptäckta solcellssystemen i analysen, vilka var; (1) Skugga — hela, eller en del, av solcellsanläggningen skuggades av omgivande träd, byggnader eller en del av byggnaden den installerats på, (2) täckt av trädkronor — hela eller en del av solcellssystemet täcks av en trädkrona i ortofotot, (3) Hög lutning — moduler som har installerats med en hög lutning, vanligtvis >45°, vilket gör att de uppträder med olika (ofta skeva) dimensioner i ortofoton, jämfört med system installerade med lägre lutning, och (4) Reflektion — moduler som ser ljust vita ut på grund av ett sammanträffande där vinkeln mellan solen, systemens orientering och kameran gör att det bildats en stark reflektion när flygbilden tas.

Alla oupptäckta solcellssystem klassificerades med en av de tre modulegenskaperna, en av de sju tak-eller markfärgerna, och de försvårande egenskaperna om de var aktuella i det specifika fallet. Kvantitativt presenteras resultatet av denna klassificering i Figur 9.



Figur 9: Illustrativ sammanfattning över försvårande karaktärsdrag för identifiering av solenergisystem med fjärranalys av flygbilder. Storleken på cirkeldiagrammen representerar antal oidentifierade system i respektive kommun, Tårtbitarna i diagrammets ytterkant representerar antalet system som delar den försvårande karaktärsdrag i fråga och är även kopplade till den innersta och mittersta delen i diagrammet som representerar modulegenskaper och tak/markfärg. Till exempel visar den streckade röda ytan i cirkeldiagrammet för Falun att 5 oidentifierade system i kommunen utgjordes av ramlösa moduler, installerade på ett svart tak som var skuggat vid tillfället när flygfotot togs, som sedan användes för att skapa ortofotot.

Den allmänna slutsatsen från Figur 9 är att små moduler, oavsett takets färg, ger upphov till ett antal oidentifierade system, men att de flesta system som inte identifierats består av ramlösa moduler när de installeras på antingen svarta tak eller som *heltak*. Vidare är kombinationen ramlösa moduler installerade på en kontrasterande takfärg (som betong, grå, terrakotta eller vit), ofta i kombination med någon av de försvårande karaktärsdragen, också en anledning till att system inte kan identifieras, där skuggning är den vanligast förekommande orsaken. Slutligen kan konstateras att moduler med ram som inte detekterats av CNN-klassificeringsmetoden nästan uteslutande beror på att en eller flera av de fyra försvårande karaktärsdragen är närvarande, vilket är logistiskt eftersom ramen tydligt avskärmar modulerna från omgivningen och således borde vara de moduler som är enklast att identifiera.

Förutom utseendets inverkan på identifieringen, noterades också att mindre solcellssystem är något överrepresenterade som oupptäckta. Av system med en yta motsvarande >50 m^2 i ortofoton var endast 4% oupptäckta i Uppvidinge, 4% i Falun, 5% i Knivsta 2019 skanningen och 2% i Knivstas 2021 skanningen, vilket kan jämföras med den totala *oupptäcktsfrekvensen* på 11% i Uppvidinge, 12% i Falun, 6% vid Knivstas 2019 skanningen och 5% vid Knivstas 2021 skanningen.

Ingen analys gjordes för de oupptäckta solvärmesystemen eftersom en tydlig trend av förbättrad detekteringsgrad kan utläsas från Tabell 4 och Tabell 6 i ett senare stycke, vilket är en konsekvens av att adderad träningsdata över solvärmesystem efter varje kommunskanning, och vi anser därför att den största orsaken till att solvärmesystem fortfarande inte identifieras är bristen på träningsbilder över sådana.

CNN-klassificeringsalgoritmen

Varje kommunskanning analyserades utifrån noggrannhetsmåtten som beskrivs i *Utvärderingsmetoder*, och resultaten presenteras i Tabell 4.

Solenergipolygonlagren innehåller information om huruvida ett polygonobjekt är ett solcells- eller solvärmesystem. Detta gör det möjligt att dela upp antalet TP-bildrutor och FN-bildrutor i de två solenergiteknologierna. Det bör noteras att summan av antalet sanna positiva solcellsbildrutor och sanna positiva solvärmebildrutor i Tabell 4 inte överensstämmer men det angivna antalet totala TP-bildrutor för båda teknologierna för Falun och Knivsta. Anledningen är att det för dessa kommuner fanns 14, 3 respektive 4 bildrutor som överlappade med både ett solcells- och solvärmepolygon och därmed ingick som TP för båda teknologierna i Tabell 4. Detsamma gäller falska negativa i Falun, där två bildrutor som överlappar både solcells- och solvärmepolygon förblev oupptäckta i alla fyra körningarna.

Tabell 4. Noggrannhetsanalys av samtliga körningar baserat på träningsdatauppsättningarna och dess prestanda enligt arbetsflödet presenterat i Figur 2, som i tabellen motsvarar från höger till vänster kolumn.

		Uppvidinge 2020	**Falun 2020**					**Knivsta 2019**		**Knivsta 2021**	
Träningsdatauppsättning	**Teknologi**	Bas	Bas	Bas Upp.0.5m	Bas Upp.2.5m	Bas Upp.5.0m	Bas Upp.7.5m	Bas	Bas Upp.5.0m Fal.5.0m	Bas	Bas Upp.5.0m Fal.5.0m Kni.5.0m
Totalt # bildrutor	-	3 645 323	6 767 141	6 767 141	6 767 141	6 767 141	6 767 141	877 142	877 142	877 142	877 142
# Gränsrutor	-	31	108	108	108	108	108	32	32	56	56
# Sanna Positiva	Båda	267	944	931	936	934	940	353	359	601	625
	Solceller	248	677	675	667	674	672	305	301	558	569
	Solvärme	19	281	270	283	274	282	51	61	46	60
# Falska Positiva	-	562	1 582	1 268	1 501	877	862	647	474	601	353
# Sanna Negativa	-	3 644 353	6 764 126	6 764 440	6 764 207	6 764 831	6 764 846	876 003	876 176	875 706	875 954
# Falska Negativa	Båda	110	381	394	389	391	385	107	101	178	154
	Solceller	93	222	224	232	225	227	67	71	138	127
	Solvärme	17	161	172	159	168	160	40	30	41	27
Precision	Båda	32,2%	37,4%	42,3%	38,4%	51,6%	52,2%	35,3%	43,1%	50,0%	63,9%
Recall	Båda	70,8%	71,2%	70,3%	70,6%	70,5%	70,9%	76,7%	78,0%	77,2%	80,2%
	Solceller	72,7%	75,3%	75,1%	74,2%	75,0%	74,7%	82,0%	80,9%	80,2%	81,8%
	Solvärme	52,8%	63,6%	61,1%	64,0%	62,0%	63,8%	56,0%	67,0%	52,9%	69,0%
F1	Båda	44,3%	49,0%	52,8%	49,8%	59,6%	60,1%	48,4%	55,5%	60,7%	71,1%

De fyra skanningarna på Falun kommun, baserade på basuppsättningen plus var och en av träningsdatauppsättningarna från Uppvidinge, utvärderades mot grundsanningen i Falun genom att beräkna precisionen, återkallelsen (*recall*) och F1-värdet, samtliga presenterade i Tabell 4. Ur resultaten kan det utläsas att antalet FP-bildrutor minskade mellan körningarna där CNN-klassificeringsalgoritmen hade tränats med basuppsättningen plus Upp.2,5m och med basuppsättningen plus Upp.5,0m, trots endast marginella förändringar i antalet TP-bildrutor. Därigenom förbättrades precisionen från 38,4% till 51,6%, med endast en försumbar inverkan på återkallelsen, när minimigränsen sattes till 5,0 m^2 för generering av ytterligare träningsdata jämfört med gränsen på 2,5 m^2. Ingen betydande minskning i antalet FP-bildrutor åstadkoms när minimigränsen sattes till 7,5 m^2. Resultatet är logiskt, eftersom antalet ytterligare träningsbildrutor minskar när minimitröskeln ökas (t.ex. från 377 till 341 positiva bilder mellan Upp.2,5m och Upp.5,0m), medan en för låg tröskel leder till att olämpliga bildrutor ingår träningsdatauppsättningen. Med bakgrund i detta användes följaktligen en träningsdatauppsättning för Knivsta bestående av basuppsättningen samt positiva träningsprov genererade från tidigare kommunskanningar med bildrutor med ett helt solenergipolygonobjekt på minst 0,5 m^2 och bildrutor som överlappar en del av ett solenergisystem med minst 5,0 m^2. De ytterligare uppsättningarna träningsdata som läggs till efter varje kommunskanning betecknas följaktligen Upp.5,0m, Fal.5,0m respektive Kni.5,0m.

Eftersom både antalet sanna positiva och falska negativa bildrutor kan tilldelas någon av de två teknologierna, kan återkallelsen för varje teknologi beräknas. Som presenterat i Tabell 4 är återkallelsen för solceller högre än för solvärme i alla skanningar. Detta är ett väntat resultat, eftersom det inte identifierades något effektivt sätt att generera träningsdata över solvärmesystem i förväg, som ledde till att basuppsättningen inte innehåller några bildrutor av solvärmesystem. Efter varje kommunskanning ökar däremot återkallelsen för solvärme avsevärt som ett direkt resultat av att träningsprov för solvärmesystem läggs till i träningsuppsättningarna efter varje kommunskanning.

Precisionsutvärderingen kan å andra sidan inte beräknas för de två individuella teknologierna separat, eftersom FP-bildrutor inte kan allokeras till vare sig solvärme- eller solcellsteknologierna, och följaktligen kan endast ett gemensamt F1-värde beräknas.

I allmänhet dras slutsatsen att tillvägagångssättet att generera ytterligare träningsdata efter varje kommunskanning och att träna om CNN-algoritmen med bilderna (som tidigare blev fel) successivt förbättrar noggrannheten, jämfört med skanningarna där endast basuppsättningen använts i träningen. Som Tabell 4 visar var den absoluta förändringen i precision en ökning med 13,9% (absolut), medan återkallelsen för båda teknologierna och F1-värdet förbättrades med 3,1% (absolut) respektive 10,5% (absolut) för den sista skanningen av Knivsta.

Även om delprojektets fokus är statistiskgenerering är det intressant att jämföra våra noggrannhetsmått med tidigare internationella studiers. Några forskningsstudier har haft fokuset på att upptäcka stora solcellsparker, som [12], [13], [21], [26], [27] vilket skiljer sig från detektion av decentraliserade

solcellssystem. Studierna baserade på DeepSolar-projektet [11], [17], [18], [25], tillsammans med [10], [14], [15] fokuserar dock alla på decentraliserade solcellssystem i den byggda miljön. De noggrannheter som uppnåtts i dessa studier presenteras i Tabell 5, tillsammans med vårt resultat för den sista skanningen i denna studie (Knivsta 2021) i fetstil. Dessutom kördes CNN-algoritmen, tränade på samma sätt som i den senaste kommunskanningen (Basuppsättning + Upp.5.0m + Fal.5.0m + Kni.5.0m), på OpenNRW_Test_16-testsetet tillhandahållet av [25] för jämförelse, och resultaten presenteras på den näst sista raden i kursiv stil i Tabell 5. Även om domängap orsakade av skillnader i geografi och typ av bilder [9], [16] bör beaktas är det värt att notera att tillägget av träningsdatat från Sverige verkar ge en förbättrad precision jämfört med resultaten från [17], [18].

Tabell 5. Noggrannhetsjämförelse mellan kända studier av att använda faltningsnätverk för identifikation av decentraliserade solcellssystem I den bebyggda miljön, samt information om flygfotoupplösningen, antalet skannade bildrutor och andel positiva prov bland bildrutorna.

Studie	Modell	Land	Precision [%]	Recall [%]	F1 [%]	Flygfoto-upplösning [m/pixel]	Antal skannade bildrutor	Andel positiva prov bland bild-rutorna [%]
Yuan et al., 2016 [14]	ConvNet	USA (1)	81.2/ 85.5	84.0 / 87.3	82.6 / 86.4	0.30	-	-
Malof et al., 2019 [15]	SolarMapper	USA	0.76	0.77	0.76	0.30	-	-
Yu et al., 2018 [11]	DeepSolar	USA (3)	93.1 / 93.7	88.5 / 90.5	90.7 / 92.1	0.15	93 500	1.31
Mayer et al., 2020 [25]	DeepSolar	USA	91.0	98.1	94.4	0.05	3 798	4.08
Kausika et al., 2021 [10]	TernausNet	Nederländerna	93.1	90.7	91.9	0.10	2 791 904 (3)	5.61
Mayer et al., 2022; Rausch et al., 2020 [17], [18]	DeepSolar	Tyskland	87.3	87.5	87.4	0.10	45 060	0.85
Detta projekt	*DeepSolar — CNN*	Tyskland	*93.4*	*81.3*	*86.9*	*0.10*	*45 060*	*0.85*
Detta projekt	**DeepSolar — CNN**	**Sverige**	**63.9**	**81.8**	**71.1**	**0.16**	**877 142**	**0.09**

(1) I Boston respektive San Francisco
(3) I bostads-respektive icke-bostadsområden
(3) Antal byggnader, inte bilder

En möjlig anmärkning är att den andra skanningen av Knivsta skulle vara subjektiv, eftersom det har genererats positiva träningsbildrutor över 209 (de system som fanns i juli 2019) av de totalt 348 solenergisystem som CNN-klassificeringsalgoritmen har till uppgift att identifiera. Men eftersom flygbilderna från 2019 respektive 2021 är tagna vid olika tider på året, tid på dygnet, vid olika väderförhållanden och olika solintensitet, skiljer sig flera aspekter av bilderna på samma objekt åt mellan de två skanningarna. Detta inkluderar till exempel hur

skuggorna faller, upplevda färger på både solenergisystem och byggnader, storlek på trädkronor och flyttning av icke-stationära föremål eller förändringar av byggnader. På grund av dessa förändringar i bilderna, som är traditionella aspekter för domängap som ger upphov till sämre prestanda när en klassificeringsalgoritm tillämpas på en ny plats [9], [15], [16], bedöms eventuell påverkan som försumbar.

Resultaten av de tre objektsbaserade måtten är listade i Tabell 6. Precis som för precision och återkallelse är andelen *detekterade polygon*, *detekterade system* och *detekterade fastigheter* högre för solcellsteknologin för varje skanning än för solvärme, av samma skäl som tidigare diskuterats. Andelen identifierade solvärmesystem ökar däremot stadigt när CNN-klassificeringsalgoritmen succesivt tränas med fler positiva solvärmebilder. I slutändan hittades 95% respektive 80% av alla detekterbara solcells-respektive solvärmesystem i 2021-skanningen av Knivsta. För solcellssystemen är det åtminstone ett tillräckligt högt antal för att metoden ska kunna användas för att generera databaser för statistiska ändamål. Det bör noteras att dessa siffror inte inkluderar de två "*odetekterbara*" systemkategorierna BIPV och vertikala system, som enligt projektutförarnas vetskap var två av varje till antalet i Knivsta.

Tabell 6: Noggrannhetsanalys av de objektbaserade måtten *detekterade polygon*, *detekterade system* och *detekterade fastigheter* för de olika kommunskanningarna jämfört med grundsanningen för körningarna enligt arbetsflödet presenteras i Figur 2, vilket representeras som vänster till höger kolumn i tabellen.

		Uppvidinge 2020		**Falun 2020**		**Knivsta 2019**		**Knivsta 2021**	
Teknologi		Solceller	Solvärme	Solceller	Solvärme	Solceller	Solvärme	Solceller	Solvärme
Polygon	Grundsanning	238	30	849	373	282	77	562	77
	Identifierade	216	18	756	271	271	60	533	62
	Andel	91%	60%	89%	73%	96%	78%	95%	81%
System	Grundsanning	128	27	375	313	142	67	282	66
	Identifierade	114	16	328	231	134	52	268	53
	Andel	89%	59%	87%	74%	94%	78%	95%	80%
Fastighet	Grundsanning	116	26	329	312	124	65	240	64
	Identifierade	105	16	285	231	120	51	230	52
	Andel	91%	62%	87%	74%	97%	78%	96%	81%

Andra skanningen av Knivsta

Som tidigare nämnt skannades både bilder från Knivsta tagna år 2019 och år 2021, vilket gör det möjligt att studera förändringarna i solenergisystemsflottan mellan dessa år. Som väntat ökade antalet system avsevärt (se Tabell 6) och bekräftar att identifiering av solenergisystem i flygbilder genom djup maskininlärning är ett verktyg som kan användas för att analysera drivkrafter bakom solcellsutrullningen [19] och följa marknadsutvecklingen med en hög rumslig upplösning.

Ett specialfall av tillbyggnaden mellan de två skanningarna är fastigheten som visas i Figur 5 som lade till ytterligare en uppsättning moduler, vilket kan ses i Figur 10a. Den här typen av kapacitetsutökning inom en fastighet känner nätägarna i Sverige inte nödvändigtvis till, eftersom rapportering endast är obligatorisk när en fastighet tillför initial elproduktionskapacitet eller när säkringen till fastigheten ändrats. Två

sådana utbyggnader noterades i Knivsta mellan de två flygbildsskanningarna. Vid korskontroll med elnätsägarens register hittas dessutom två tidigare förmodade kapacitetsutbyggnader av solcellsanläggningar i Knivsta och en i Falun, som verkar ha skett innan flygbilderna togs.

Figur 10: Ortofoton © från Lantmäteriet över solcellssystem I Knivsta. (A) Visar samma fastighet som i Figur 5, men där ytterligare solcellsmoduler nu installerats på takets norra sida. (B) Visar en byggnad på vilken solcellssystemet inte kunde identifieras vid 2019s skanning och (C) visar samma byggnad och system som i (B), men vid ortofotot 2021, där CNN-klassificeringsalgoritmen korrekt identifierade solcellssystemet.

Dessutom möjliggör efterföljande skanningar av en kommun att av solenergisystem som tas ur drift mellan uppsättningarna flygbilder identifieras. Genom att jämföra resultaten från den första och andra skanningen av Knivsta visade det sig att två solvärmesystem hade demonterats under perioden.

CNN-klassificeringsalgoritmen kunde i den andra skanningen av Knivsta korrekt klassificera 7 av de 8 solcellssystemen och 8 av de 15 solvärmesystemen som inte upptäcktes av CNN-klassificeringsalgoritmen i den första skanningen. Sannolikt kan detta dels förklaras av den ytterligare träningsdata som tillkommit från den första skanningen, som inkluderar bildrutor av dessa saknade system, och dels av ändrade förhållanden i ortofoton på grund av de varierande tidpunkterna när flygfotona togs. Figur 10 illustrerar ett exempel på det senare, där ett solcellssystem befann sig i en fullständig skugga i ortofotot från 2019 (full grönska på trädet vid flygfototillfället) vilket resulterade i en felaktig klassificering (Figur 10b). I Figur 10c är solcellssystemet väl synligt i ortofotot från 2021, och kunde därför korrekt identifieras av CNN-klassificeringsalgoritmen.

Sammanfattningsvis är slutsatsen att kontinuerliga skanningar av en kommun möjliggöra analyser av utrullning och demontering av solenergisystem, samt identifikation av tidigare oidentifierade system.

Statistikgenerering

Slutligen sammanfattar detta avsnitt vilken typ av statistik som är möjlig att extrahera med denna metod. För det första är den uppenbara slutsatsen att den rumsliga upplösningen för metoden är mycket hög. Solenergisystemen kan tilldelas med en exakt koordinat och en areautbredning sett uppifrån. Detta gör det möjligt att knyta dem till specifika byggnader (eller platser på marken), vilket är en högre upplösning än vad de flesta av de svenska elnätsägarna har, då de bara får

information om att det installeras ett solcellssystem med en viss kapacitet bakom en viss nätanslutningspunkt, vilket vanligtvis ger en upplösning per fastighet. För statistisk kring solvärmeutbredningen finns endast aggregerad försäljningsstatistik på nationell nivå i Sverige, så den rumsliga upplösningen av denna metodik överstiger vida den nuvarande tillgängliga statistiken.

Vidare gör möjligheten att allokera de identifierade solenergisystem till vissa byggnader och fastigheter det möjligt att tilldela dem sina respektive byggnadsändamål, detaljerade byggnadsändamål och typkoder. En sammanställning av resultatet baserat på dessa parametrar för system i respektive kommun presenteras för solceller i Tabell 7 och för solvärme i Tabell 8. Till exempel finns det största antalet solenergisystem på byggnadstyperna *Bostad; Småhus friliggande* och *Komplementbyggnad; Ospecificerad*. Systemen i den första kategorin är enkla att härleda till marknadssegmentet ”privatmarknaden” eller ”villamarknaden”. Den andra kategorin kan däremot åsyfta flera olika typer av system, till exempel ett garage (Figur 3) som hör till privatmarknaden, eller en ladugård (Figur 5) som snarare hör till jordbrukssegmentet av marknaden. En tydlig indikation på vilken typ av komplementbyggnad det rör sig om kan dock erhållas med tilläggsinformationen om typkoden för fastigheter från Skatteverket.

Tabell 7. Statistisk sammanfattning över antal byggnader inom respektive byggnadsändamål, detaljerade ändamål och typkod som allokerats till solcellssystems i projektets grundsanning i de tre respektive kommunerna.

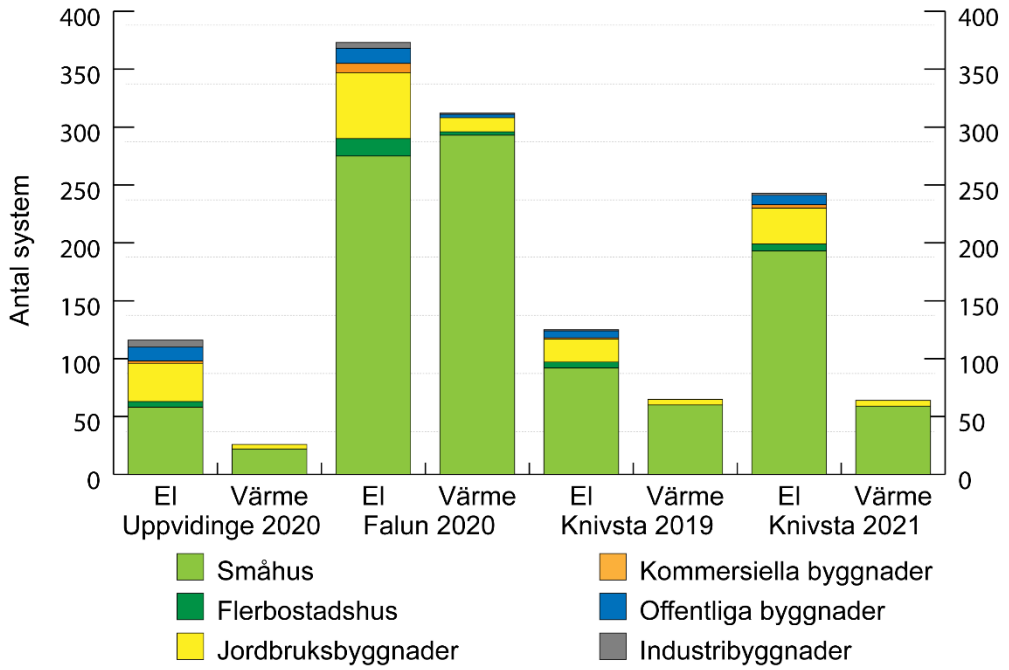
Byggnads-ändamål	**Detaljerat byggnadsändamål**	**Typkoder för fastigheter**	**Antal system**			
			Upp. 2020	Falun 2020	Knivsta 2019	Knivsta 2021
Mark	-	Småhusenheter				2
		Lantbruksenhet		2	1	1
Bostad	Småhus, friliggande	Småhusenheter	27	199	74	159
		Hyreshusenheter	1			
		Lantbruksenhet	6	11		
		Industrienheter		1		
		Specialenheter		1		
	Småhus, kedjehus	Småhusenheter		18	4	7
	Småhus, radhus	Småhusenheter		2	1	3
		Hyreshusenheter			1	1
	Småhus med flera lägenheter	Hyreshusenheter		1		
	Flerfamiljshus	Hyreshusenheter	3	12	3	3
		Småhusenheter	1			
		Specialenheter	1			1
	Ospecificerad	Småhusenheter	3			
		Lantbruksenhet	1	1		
Komplement-byggnad	Ospecificerad	Småhusenheter	20	41	13	22
		Hyreshusenheter	1	2		
		Lantbruksenhet	31	47	18	25
		Industrienheter	2	2		
		Elproduktionsenheter	1			
		Specialenheter	3			1
Ekonomi byggnad	Ospecificerad	Småhusenheter		2		1
		Lantbruksenhet		3		3
Industri	Trävaruindustri	Industrienheter	1			
	Annan tillverkningsindustri	Hyreshusenheter	1	1		1
	Övrig industribyggnad	Industrienheter	1			
	Ospecificerad	Industrienheter	1	1	1	1
		Småhusenheter		1		
Samhällsfunktion	Skola	Specialenheter		5	3	4
	Högskola	Industrienheter		1		
	Sjukhus	Specialenheter		2		
	Vårdcentral	Hyreshusenheter	1			
	Brandstation	Specialenheter		1		
	Kommunhus	Specialenheter		1		
	Kulturhus	Specialenheter	1			
	Samfund	Hyreshusenheter		1		
	Sporthall	Specialenheter			1	1
	Ishall	Industrienheter	1			
		Specialenheter	1			
	Badhus	Specialenheter	1			
	Ridhus	Lantbruksenhet			1	1
		Specialenheter	1			
	Ospecificerad	Specialenheter	2		1	1
		Hyreshusenheter	1	1		
Verksamhet	Ospecificerad	Småhusenheter		1		
		Hyreshusenheter		2		1
		Industrienheter	2	5	1	2
		Specialenheter		1		
Övrig byggnad	Ospecificerad	Småhusenheter		1		
		Hyreshusenheter			1	1
		Lantbruksenhet			1	1
Totalt			116	370	125	243

Tabell 8: Statistisk sammanfattning över antal byggnader inom respektive byggnadsändamål, detaljerade ändamål och typkod som allokerats till solvärmesystems i projektets grundsanning i de tre respektive kommunerna.

Byggnads-ändamål	**Detaljerat byggnads-ändamål**	**Typkoder för fastigheter**	**Antal system**			
			Upp. 2020	Falun 2020	Knivsta 2019	Knivsta 2021
Mark	-	Småhusenheter			2	2
	-	Lantbruksenhet		1		
Bostad	Småhus, friliggande	Småhusenheter	14	196	38	36
		Hyreshusenheter		1		
		Lantbruksenhet	3	36	7	8
		Elproduktionsenheter		1		
	Småhus, kedjehus	Småhusenheter		9		
	Flerfamiljshus	Hyreshusenheter		3		
	Ospecificerad	Småhusenheter			2	2
		Lantbruksenhet	2			
Komplement-byggnad	Ospecificerad	Småhusenheter	5	50	11	11
		Lantbruksenhet	2	11	5	5
	Övrig industribyggnad	Industrienheter		1		
Samhällsfunktion	Skola	Specialenheter		1		
	Sporthall	Specialenheter		1		
Verksamhet	Ospecificerad	Hyreshusenheter		1		
Totalt			26	312	65	64

Som Tabell / och Tabell 8 visar så kan en mycket hög upplösning på marknadssegmenten uppnås i Sverige, eftersom uppgifter om fastigheter och byggnader finns på en sådan detaljerad nivå. För att tydliggöra resultaten har en aggregerad nivå introducerats som illustreras i färgkoder i Tabell 7 och Tabell 8, där sex mer klassiskt definierade marknadssegment visas, nämligen (1) Småhus — ljusgrön, (2) Flerbostadshus — mörkgrön, (3) Jordbruksbyggnader — gul, (4) Kommersiella byggnader — orange, (5) Offentliga byggnader — blå och (6) Industribyggnader — grå. Resultaten av denna indelning presenteras i Figur 11.

Det är mycket tydligt att Småhus är det enskilt största marknadssegmentet för solenergi i dessa tre kommuner sett till antalet system. Det bör noteras att flera icke nätuppkopplade system ingår i uppställningen. Därefter kommer marknadssegmenten jordbruksbyggnader och kommersiella byggnader. Detta är en intressant eftersom det inte har varit möjligt att separera dessa två marknadssegment från varandra i de befintliga statistikdatabaserna i Sverige förrän nu. Offentliga anläggningar och industrianläggningar står endast för ett fåtal solenergisystem i varje kommun, men varje sådant system är generellt sett mycket större än systemen för enfamiljshus, så kapacitetsmässigt skulle de sannolikt stå för större marknadsandelar.



Figur 11: Sammanställning av antalet solenergisystem som kan kopplas till respektive byggnadstyp i de tre kommunernas grundsanningar.

Resultat socioekonomisk

Med kunskapen som förvärvats genom kartläggningsaktiviteterna i föregående avsnitt kan socioekonomiska parametrar kopplas till hushållen i enfamiljshusen, och en analys av korrelationsfaktorer till solenergiutbredning utföras.

Demografiska statistikområden

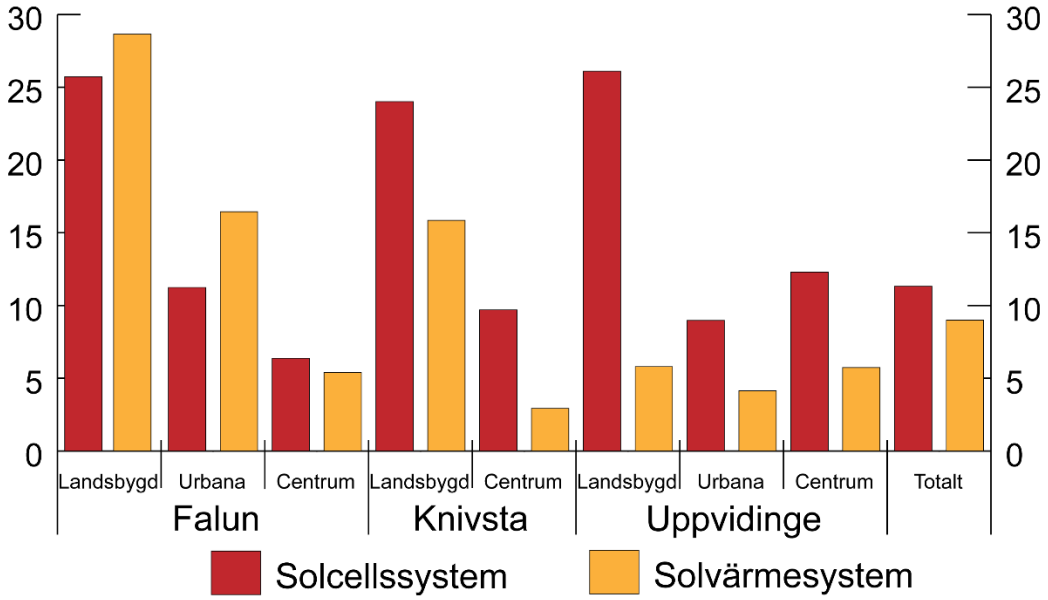
Vid kategorisering av de privatägda solcellssystemen i de olika DeSo-områdena visar resultatet att 201 av de 452 solcellssystemen (44,5%) återfanns i kommunernas olika landsbygds-DeSOn, 54 (11,9%) i urbana DeSOn och 197 (43,6%) i centrum-DeSOn. Likaså återfanns 161 av de 359 solvärmesystemen (44,8%) kommunernas olika landsbygds-DeSOn, 66 (18,4%), i urbana DeSOn och 132 (36,8%), i centrum-DeSOn. Se sammanställning i Tabell 9.

Tabell 9: Totalt antal privata solcells- och solvärmesystem i de tre studerade kommunerna, uppdelat på respektive DeSO-typ. DeSO_L, DeSO_U och DeSO_C avser Landsbygds-DeSo, Urbant DeSo respektive Centrum-DeSo.

	Antal solcellssystem			**Antal solvärmesystem**		
	DeSO_L	**DeSO_U**	**DeSO_C**	**DeSO_L**	**DeSO_U**	**DeSO_C**
Falun	201	54	197	161	66	132
Totalt	452			359		

Som redan nämnt behövde ett antal system uteslutas ur analysen på grund av avvikande ägarförhållanden eller andra osäkra bakgrundsdata. 37 system utelämnades i Falun, 8 i Knivsta och 19 i Uppvidinge på grund av att fastighetsägaren tycktes bo utanför kommunen i fråga eller förvärvade egendomen efter 2020-12-31. Dessutom utelämnades ytterligare 12 system i Falun på grund av brist på data eller andra svårigheter att erhålla en komplett socioekonomisk bild. Borttagandet av dessa system betraktas ha låg inverkan på slutresultatet.

Som framgår av antalet privata solenergisystem per 1 000 hushåll, illustrerat i Figur 12, är koncentrationen av båda solenergiteknologierna betydligt högre på landsbygden än i urbana och centrala områden. Anledningen är sannolikt att andelen hushåll som bor i flerbostadshus är betydligt högre i de sistnämnda kategorierna. En annan observation är att koncentrationen av solvärmesystem visade sig vara högre än för solcellssystem i Faluns landsbygds- och urbana områden, en motsatt trend jämfört med i övriga kommuner. Detta indikerar att lokala faktorer troligen har haft en viktig roll i att driva spridningen av även solvärmesystem, något som Palm identifierat för solcellssystem [36].



Figur 12: Medelantalet privata solenergisystem per 1 000 hushåll i de tre studerade kommunerna, uppdelat på de olika typerna av DeSOn.

Korrelationsanalys

De beroende variablerna *antal solcellssystem per hushåll, #PV,* och *antal solvärmesystem per hushåll, #ST,* undersöktes i relation till flertalet oberoende korrelationsparametrar enligt beskrivningen i avsnitt Variabler och korrelationsanalys.

#PV	1	-												
#ST	0.36	0.61	1											
0–15	-0.09	0.06	-0.24	1										
16–24	-0.44	-0.31	-0.43	0.16	1									
25–44	-0.41	-0.42	-0.51	0.50	0.21	1								
45–64	0.52	0.59	0.58	-0.39	-0.35	-0.74	1							
>65	0.18	0.00	0.31	-0.90	-0.34	-0.67	0.34	1						
Sex	0.56	0.50	0.36	0.01	-0.23	-0.20	0.51	-0.07	1					
BOri	0.32	0.36	0.56	-0.07	-0.51	-0.35	0.42	0.19	0.05	1				
Edu	-0.22	-0.06	-0.26	0.38	0.28	0.23	-0.22	-0.39	-0.49	0.28	1			
Emp	0.15	0.35	0.27	0.34	-0.35	-0.07	0.19	-0.21	-0.02	0.75	0.50	1		
AInc	0.00	0.21	-0.07	0.66	0.14	0.21	-0.01	-0.65	-0.07	0.33	0.76	0.71	1	
EcSt	0.18	0.43	0.06	0.38	-0.04	-0.05	0.27	-0.37	0.01	0.39	0.66	0.71	0.85	1
	#PV	#PV*	#ST	0–15	16–24	25–44	45–64	>65	Sex	BOri	Edu	Emp	AInc	EcSt

Figur 13: Korrelationsmatris som visar Pearson-koefficienten mellan antalet solcells- respektive solvärmesystem per hushåll i kommunernas DeSOn och de oberoende parametrarna presenterade i Tabell 3.

Korrelationsanalysen som presenteras i Figur 13 baseras på Pearson-koefficient mellan de insamlade datamängderna från samtliga 47 DeSOn där minst ett privat solenergisystem identifierades, och på socioekonomiska och demografiska data för varje DeSO. Två av dessa 47 DeSOn uppvisade tydligt högre penetrationsnivåer (ett landsbygds-DeSO i Falun med 0,046 PV system per hushåll och ett landsbygds-DeSO i Uppvidinge) i jämförelse med 0,001–0,033 solcellssystem per hushåll i övriga. Utelämnandet av dessa två avvikande DeSOn resulterade i ytterligare en beroende variabel, #PV*, som visar installerade solcellssystem i de återstående 45 DeSOrna.

Som Figur 13 visar så identifierades en negativ kollinearitet för de oberoende variablerna (1) åldersgrupperna 0–15 och >65, och (2) åldersgrupperna 25–44 och 45–64. Positiv kollinearitet återfanns å andra sidan för (1) anställning, Emp, och födelseursprung, BOri, (2) utbildningsnivå, Edu, och medelinkomst, AInc, (3) sysselsättning, Emp, och medelinkomst, AInc, (4) sysselsättning, Emp, och genomsnittlig ekonomisk standard, EcST, och (5) medelinkomst, AInc och genomsnittlig ekonomisk standard, EcST.

Det huvudsakliga bidraget till det globala kunskapsläget är den måttligt positiva korrelation (0,36) som identifierades mellan #PV och #ST, vilket ökade till en medelstark positiv korrelation när de två extrem-DeSOrna uteslöts (0,61). Det är första gången relationen mellan de två solenergiteknologierna har undersökts.

Utöver detta förekommer det också en medelstark positiv korrelation mellan både #PV och #ST och åldersgruppen 45–64, medan åldersgrupperna 16–24 och 25–44 tycks vara måttligt negativt korrelerade till solenergisystemspenetrationen på DeSO-nivå. Andelen män i ett demografiskt område verkar också ha en positiv effekt på solenergiutrullningen, med en medelstark positiv korrelation till #PV och

måttlig positiv korrelation till #ST. För de andra oberoende demografiska variablerna visade sig korrelationen vara svag eller obefintlig, undantaget om de två DeSO:erna med exceptionellt hög solcellspenetration utelämnades, vilket resulterade i att korrelationen mellan #PV* och sysselsättning och genomsnitt ekonomisk standard steg till en måttlig positiv korrelation. Den presenterade korrelationsmatrisen i Figur 13 gör inte skillnad mellan allmänna demografiska faktorer i de tre kommunerna, t.ex. generell inkomst- och utbildningsnivå, vilket kan ha resulterat i en svagare korrelation mellan de olika variablerna. Pearsons korrelationskoefficient mellan antal solcellssystem per hushåll och varje separat oberoende variabel presenteras emellertid för de tre kommunerna separat i Figur 14.

#PV	#ST	0–15	16–24	25–44	45–64	>65	Sex	BOri	Edu	Emp	AInc	EcSt
Falun	0.49	-0.08	-0.38	-0.41	0.49	0.19	0.66	0.43	-0.24	0.16	0.03	0.09
Knivsta	0.45	-0.24	-0.50	-0.68	0.69	0.37	0.74	0.59	-0.20	0.27	0.00	0.62
Uppvidinge	-0.57	-0.15	-0.81	-0.51	0.63	0.38	-0.22	0.87	0.62	0.56	0.42	0.79
#ST	#PV	0–15	16–24	25–44	45–64	>65	Sex	BOri	Edu	Emp	AInc	EcSt
Falun	0.49	-0.18	-0.43	-0.58	0.65	0.31	0.56	0.55	-0.41	0.31	-0.02	0.14
Knivsta	0.45	-0.33	-0.65	-0.45	0.73	0.33	0.81	0.55	-0.76	0.14	-0.16	0.22
Uppvidinge	-0.57	0.12	0.26	0.24	-0.25	-0.18	0.32	-0.64	0.12	-0.65	-0.70	-0.62

Figur 14: Pearsonkorrelationskoefficienten mellan de oberoende variablerna antal solcellssystem per hushåll (PV, övre delen) och solvärmesystem per hushåll (ST, nedre delen) och för de tre studerade kommunerna.

Angående solcellssystem är resultatet som presenteras i Figur 14 är förhållandevis konsekvent mellan de tre kommunerna vad gäller korrelationen med de olika åldersgrupperna och födelseursprunget. Däremot finns skillnader avseende kön och utbildningsnivå, där Uppvidinge sticker ut med en motsatt korrelation än de andra kommunerna. Dessutom förekommer stark och medelstark korrelation till genomsnittlig ekonomisk standard i Knivsta och Uppvidinge, medan den korrelationen endast är svag i Falun.

För solvärmeteknologin är resultaten snarlika de för solceller i Knivsta och Falun. Den medelstarka och starka negativa korrelationen mellan antalet solvärmesystem och utbildningsnivå sticker däremot ut. Solvärmesystemenen i Uppvidinge tycks däremot korrelera på ett annat sätt än både de två andra kommunerna och med korrelationerna för solceller. En anledning till detta kan vara att Uppvidinge är kommunen med svagast dataunderlag med endast 23 solvärmeanläggningar och 6 DeSOn (att jämföra med 287 system i Falun och 49 i Knivsta, med 37 respektive 10 DeSOn).

Litteraturanalys

Faktum är att resultaten från den socioekonomiska studien kring den obefintliga eller svaga korrelationen mellan solenergipenetration och medelinkomst på aggregerad nivå ligger i linje med flera kvantitativa studier. Svaga eller obetydliga effekter av medelinkomst för solcellsinstallationer inom stora rumsliga områden

presenteras i [45]–[51], och en till och med en negativ effekt påvisas i [52]–[54]. Dessutom, precis som i denna studie, har det tidigare visats att genomsnittlig ekonomisk standard, eller ackumulerat kapital, kan vara en demografisk faktor med högre betydelse för solcellsutbredning än medelinkomsten på aggregerad arealnivå [55].

En medelstark positiv korrelation mellan båda solenergiteknologierna och åldersgruppen 45–64 år kunde påvisas, precis som tidigare rapporterats i Uppsala kommun [40]. Detta tyder på att områden med en högre andel av sin befolkning i denna ålder tenderar att ha en högre solenergipenetration. Motsatsen verkar vara fallet för unga, åldersgruppen 16–24, och medelålders vuxna, åldersgruppen 25–44, vilka båda presenterar en måttlig till medelstark negativ korrelation till antal solenergisystem per hushåll på DeSO-nivå. Det är värt att notera att medelåldern i Knivsta är lägre (36,9 år) än i Falun (42,4 år) och i Uppvidinge (43,5 år), vilket kan påverka resultatet i Figur 14. Dessa fynd överensstämmer i stor utsträckning med tidigare forskning, där mani USA [56], [57] och i Nederländerna [58], påvisat negativ korrelation för åldersgrupperna 20–45 och >65 på postortsnivå. Även i Australien [50], Tyskland [47], Grekland [59] och Storbritannien [49] har studier visat att en högre andel invånare i medelåldern i ett område ökar förekomsten av solcellssystem. Angående åldersgruppen >65 år, drog [60] slutsatsen att pensionerade personer i Storbritannien är mindre motiverade att komplettera sitt el- eller värmesystem med ett solenergisystem på grund av den höga investeringskostnaden och långa avbetalningstiden. Dessutom fann [61] att benägenheten att investera i solvärme minskade med hushållets medelålder under den tidiga marknadsutvecklingen i Tyskland, även om det också finns litteratur [45], [62] som visar att ålder har en mindre betydelse än tidigare trott.

Kön visade sig ha en medelstark positiv korrelation till både utbyggnaden av solcell- och solvärmesystem (0,56 respektive 0,50) och till åldersgruppen 45–64 år (0,51) i denna studie. Det första indikerar en högre benägenhet hos män att installera solenergisystem, vilket studier på postortsnivå i Kalifornien också antyder [56].

Det måttliga respektive medelstarka sambandet mellan födelseursprung och utbyggnaden av solceller och solfångare kan i analyseras i relation till den starka positiva korrelationen (0.75) mellan födelseursprung och sysselsättningsgrad. Detta eftersom sysselsättningsgrad i sin tur har en högre korrelation till solenergiutbyggnad. Det kan tolkas som att det snarare än födelseutsprunget är sysselsättningsgraden som kan kopplas till investeringen i ett solenergisystem. Denna hypotes stärks delvis av resultaten i [55]–[57], [62], [63], i USA, som presenterar endast svaga bevis på en högre solcellspenetration i områden med en högre andel vita invånare.

Föga förvånande är variablerna Sysselsättningsgrad, *Emp*, kopplade till de två ekonomiska variablerna Medelinkomst, *AInc*, och Ekonomisk standard, *EcSt*. Korrelationen till antalet installerade solenergisystem konstaterades svag för kommunerna i helhet, men starkare för solceller när kommunerna studerades var för sig. Den svaga positiva korrelationen anses ligga i linje med tidigare internationella studier [47], [55], [64]. Variabeln Utbildningsnivå visade sig ha en svag negativ korrelation till solenergiutbredningen och en stark korrelation till

Medelinkomst. När kommunerna undersöktes separat visade sig utbildningsnivån ha en negativ relation till båda solenergiteknikerna i Falun och Knivsta och en positiv korrelation i Uppvidinge. På DeSo-nivå visade sig däremot variabeln ha en positiv korrelation till ”urbana” och ”centrum”-DeSon. Med undantag för en studie [45], är en positiv relation till utbildningsnivå konstaterad i flertalet internationella studier [40], [47], [48], [54], [57], [62], [64], [65].

Egenskaper hos hushåll med solenergisystem

Uppgifter om medelålder och antal år som ett hushåll i genomsnitt är bosatt i en bostad gick erhålla genom Ratsit AB och resultaten presenteras i Tabell 10. I den kan även utläsas att solcellsägare i snitt är något äldre än ägare av solvärmesystem. Detta överensstämmer med slutsatsen att hushåll med en medelålder på >40 är överrepresenterade bland solcellsägare i Danmark [65], men resultaten är motsägelsefulla i relation till en svensk studie där ålder visade sig ha en, om än låg, negativ effekt på solcellspenetrationen [66].

Tabell 10. Medelbosättningstiden för hushåll med solcell- respektive solvärmesystem [år] samt medelåldern för de bosatta (>20 år) i hushåll med solcell- respektive solvärmesystem och i genomsnitt i kommunerna Falun, Knivsta och Uppvidinge.

	Bosättningstid, Solcell	**Bosättningstid, Solvärme**	**Medelålder, Solcell**	**Medelålder, Solvärme**	**Medelålder, total**
Falun	22,1	22,6	57,9	57,9	51,8
Knivsta	17,1	17,8	54,4	52,3	48,3
Uppvidinge	23,7	28,6	58,9	55,4	53,3

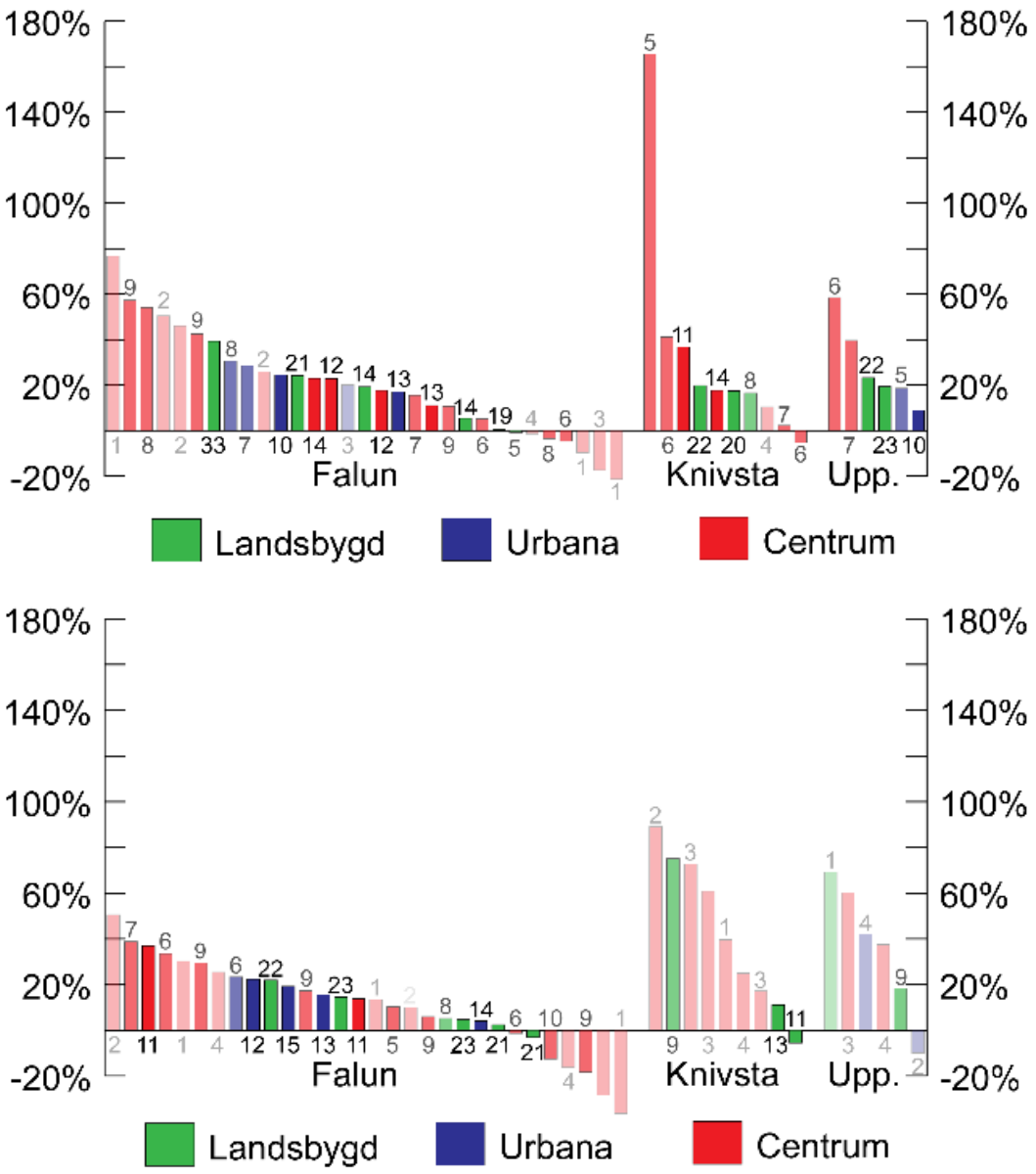
Dessutom undersöktes könsfördelningen i hushållen med solenergisystem, se Tabell 11. Det bör däremot noteras att medelålder och bosättningstid baseras på invånare >20 år, medan information om kön finns att tillgå för samtliga medlemmar i hushållen. Detta resulterar i en viss diskrepans mellan dataunderlagen. Slutsatsen är att andelen män i hushåll med solenergisystem är något högre än snittet bland alla hushåll i de tre kommunerna. Detta ligger i linje med en tidigare svensk enkätstudie om solcellsutbyggnaden [66], tidigare forskning på aggregerad nivå, och med en studie från Danmark, som påvisade att registrerade ägare av solcellssystem oftare var män (85,8%) och att män i mycket högre grad tar själva investeringsbeslutet för byggnationen av solcellssystem [65]. Det har också slagits fast att män i Tyskland är mer benägna att investera i solvärmesystem än kvinnor [61].

Tabell 11. Medelkönet, vilket avser andelen män, för hushåll med solcell- respektive solvärmesystem samt totalt i kommunerna Falun, Knivsta och Uppvidinge.

	Medelkön, Solcell	**Medelkön, Solvärme**	**Medelkön, Totalt**
Falun	49,9%	54,8%	49,6%
Knivsta	51,1%	51,4%	51,0%
Uppvidinge	54,8%	55,1%	52,3%

Slutligen visar resultaten att medelinkomsten för hushåll med solenergisystem ofta är högre än för samtliga hushåll i respektive DeSO, vilket illustreras i Figur 15. Detta bekräftar i sin tur att hushållens medelinkomst är en viktig faktor för privata solcellsinstallationer, vilket dokumenterats i många länder [40], [45], [55]–[59],

[61], [65], [67]–[69] och för privata solvärmesystem, vilket påvisats i tyska studier [69], [70].



Figur 15: Den procentuella skillnaden mellan medelinkomsten för hushåll med solenergisystem och den totala medelinkomsten i samma DeSo. Siffran ovanför staplarna indikerar antalet hushåll med privata solenergisystem i det DeSot. DeSon med <10 respektive <5 hushåll med solcellssystem visas med olika stark färgintensitet för att indikera resultatets statistiska säkerhet. Figur a) t.v. visar hushåll med solcellssystem och figur b) t.h. visar hushåll med solvärmesystem.

Resultat U-net

Den i det här projektet utvecklade U-net-modellen är tränad och testad på svenska och tyska datauppsättningar av varierande storlek, inklusive blandningen av båda datamängderna, för utvärdering av en högre generaliserbarhet av modellen. Dessutom testas alla inställningar på två olika bildupplösningar: 128×128 och 256×256 pixlar. Målet här är att utvärdera effekten av upplösningen på modellens

prestanda, och därefter utforma rekommendationer om vilken upplösning som ska väljas. Så länge resultaten är tillräckligt goda bör alltid en lägre bildupplösning väljas på grund av den höga hårdvarupåfrestningen som krävs för att driva en sådana körningar.

Modellvalidering

Modellens prestanda presenteras i radardiagrammen i Figur 16. Resultaten visar tydligt att U-net-modellen presterar bättre när den tillämpas på den tyska datamängden än den svenska. Detta är särskilt tydligt om man ser till *Dice loss*, DL-värdet, som är ett mått på den relativa överlappningen mellan den genererade segmenteringen och den manuellt skapade *masken*. En förklaring till detta är att DL-värdet inte gör skillnad på små och stora system och kännetecknas därför av en stor variation. Exempelvis kommer samma DL-värde ges till en segmentering som missat en enda pixel av ett litet system i en bild som en betydande del av ett stort solenergisystem. Eftersom den svenska datauppsättningen innehåller många bilder där endast en liten del av ett helt solenergisystem syns, som tidigare avhandlat, är det till nackdel för DL-värdet.

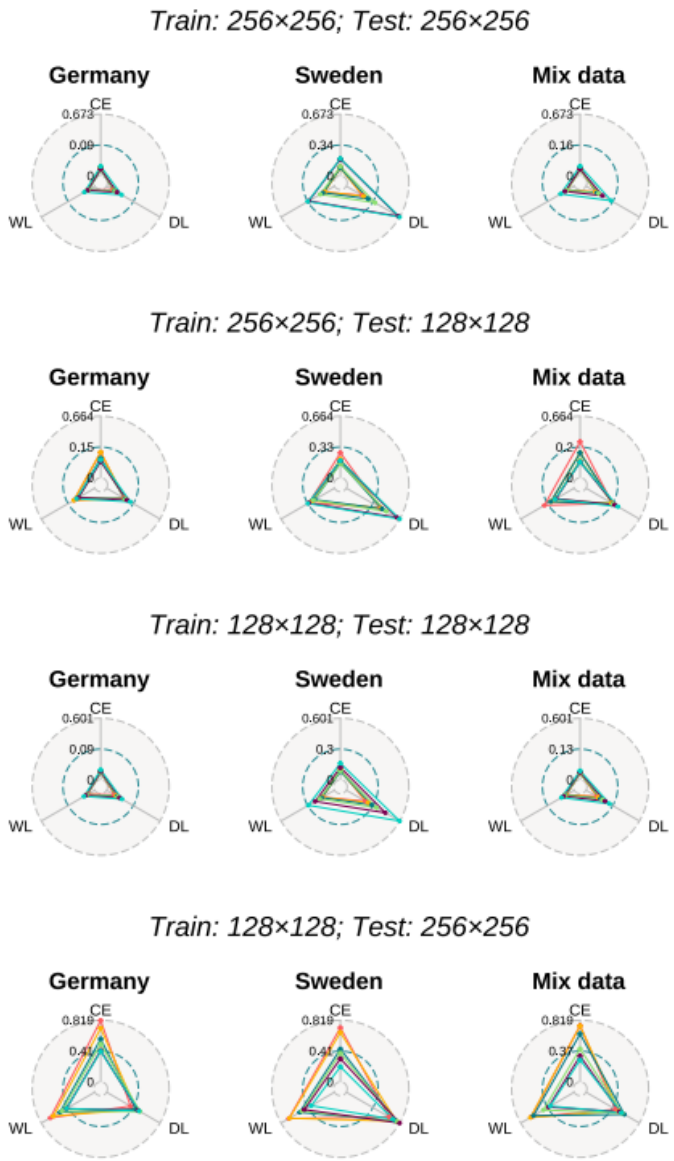
Kordentropin, CE-värdet, utvärderar däremot varje enskild pixel och ger ett mer konsekvent resultat mellan de två datauppsättningarna. Den viktade förlusten, eller viktade diskrepansen, mellan segmenteringen och den manuella masken betecknas *Weighted loss*, WL, och inkluderas i Figur 16 för att presentera en så rättvis bild så möjligt. I samma figur kan det även utläsas att den blandade datauppsättningen möjliggör en högre generaliserbarhet och således en god balans mellan segmentering av solenergisystem från båda platser med en bättre genomsnittlig noggrannhet.

Slutligen kan utläsas att trots att det förekommer en viss skillnad i noggrannhet mellan modeller som tränats på bildupplösningen 256×256 och 128×128, är det inte signifikant. Därför kan man ur resultatet dra slutsatsen att det kan vara fördelaktigt att använda 128×128 från hårdvaruanvändningssynpunkt, eftersom fler pixlar innebär fler parametrar att lära in sig på.

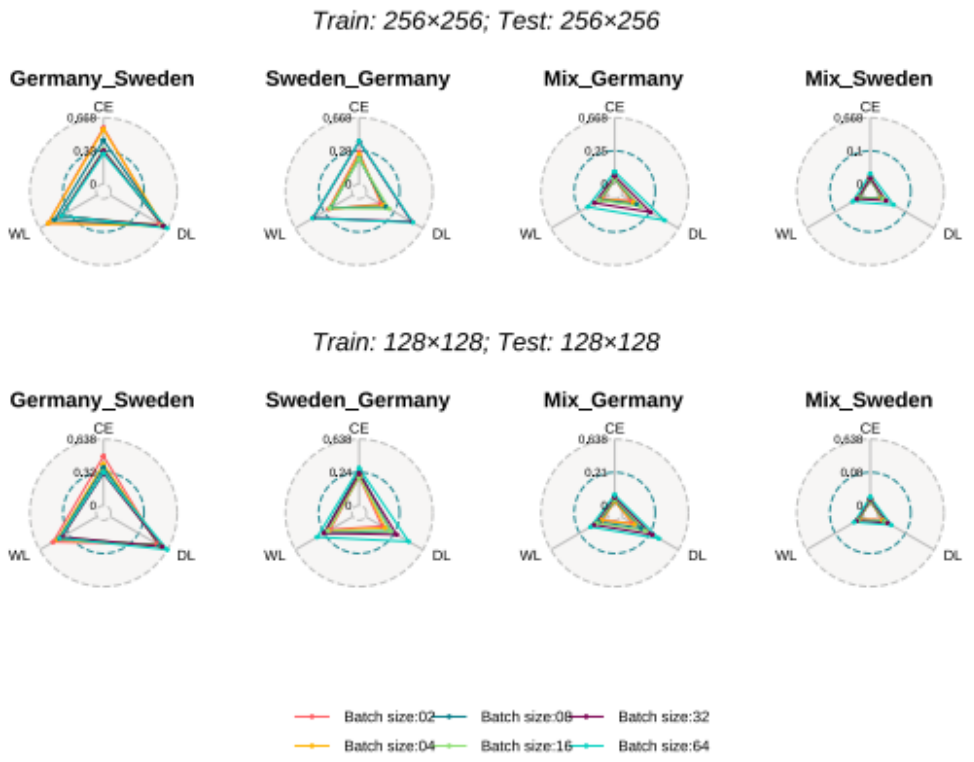
Korsvalidering

Till skillnad från modellvalidering, innebär korsvalidering att modellen tränas på en uppsättning data, för att sedan testas på en annan. Figur 17 presenterar DL, CE och WL-värdena för dessa korskörningar mellan de totalt fyra uppsättningarna. Resultaten är något överraskande, eftersom U-net-modellen som tränats på svenska data och körts på tyska data visar sig vara den med bäst sammanvägd noggrannhet. Förklaringen till detta är det faktum att de tyska bilderna är valda så att de innehåller tydliga och stora system, vilket innebär färre bilder som bara innehåller delar av systemen. Man kan också utröna att träning av modellen på den blandade datauppsättningen gör modellen mer generaliserbar och mer allmänt tillämplig.

Precis som i modellvalideringen kan det konstateras att en högre upplösning inte nämnvärt förbättrar noggrannheten av modellen, utan snarare förbrukar mer minne och beräkningskapacitet i förhållande till nyttan det genererar.



Figur 16: Radardiagram över korsentropin (*Cross entropy*, CE), dice-loss värde (DL) och viktad förlust (*Weighted loss*, WL) för alla datauppsättningar, testade i samtliga tillgängliga bildupplösningar. Notera att diagramaxlarna varierar något mellan raderna.



Figur 17. Radardiagram över korsentropin (*Cross entropy*, CE), dice-loss värde (DL) och viktad förlust (*Weighted loss*, WL) för den genererade segmenteringen när modellen tränats på en datauppsättning (t.ex. Tyskland) och sedan testats på en annan (t.ex. Sverige) för två olika bildupplösningar. Notera att diagramaxlarna varierar något mellan raderna.

Jämförelse med andra studier

För att kunna jämföra resultaten med tidigare resultat i litteraturen beräknas ett antal statistiska jämförelsemått för samtliga modellkörningar som presenteras i Figur 16. Resultaten varierar kraftigt mellan olika datauppsättningar och bildupplösningar, se Tabell 12, men visar konsekvent bättre resultat för den tyska datan än den svenska. Dessutom uppnås lägre värden för statistisk säkerhet när modellen tränas och körs med bilder av olika upplösning, speciellt om den tränas på data med lägre upplösning än den sedan testas på.

Tabell 12: De statistiska jämförelsemåtten *det genomsnittliga absoluta procentuella felet* (MAPE), *det genomsnittliga absoluta felet* (MAE), *bias* (B) och den *genomsnittliga*

skärningspunkten över unionen (mIoU) för samtliga datauppsättningar, testade och körda i samtliga bildupplösningar.

Train resolution	Test resolution	Batch size	German				Swedish				Mix data			
			MAPE	MAE	mIoU	B	MAPE	MAE	mIoU	B	MAPE	MAE	mIoU	B
256×256	256×256	02	0.08	638.81	0.92	0.01	0.21	645.49	0.86	0.06	0.14	624.68	0.90	0.01
		04	0.09	778.45	0.91	0.01	0.27	796.39	0.85	0.04	0.14	609.35	0.90	0.01
		08	0.09	787.62	0.91	0.01	0.35	998.83	0.81	0.04	0.16	705.84	0.89	0.01
		16	0.12	883.77	0.90	0.01	0.43	1487.14	0.74	0.06	0.19	785.79	0.87	-0.00
		32	0.15	1096.09	0.88	0.01	1.18	3194.51	0.57	-0.09	0.24	1025.04	0.83	-0.00
		64	0.24	1725.68	0.83	0.01	1.25	3315.18	0.56	-0.11	0.40	1731.00	0.76	0.00
256×256	128×128	02	0.23	551.36	0.83	0.03	0.53	483.10	0.73	0.12	0.35	551.54	0.78	0.09
		04	0.23	637.67	0.81	0.06	0.55	472.51	0.72	0.10	0.35	560.18	0.77	0.09
		08	0.24	620.86	0.81	0.06	0.64	513.70	0.70	0.02	0.35	618.96	0.76	0.12
		16	0.27	633.56	0.79	0.06	0.75	571.11	0.66	-0.03	0.36	619.19	0.75	0.10
		32	0.30	670.52	0.78	0.05	1.54	920.18	0.57	-0.32	0.37	610.31	0.74	0.07
		64	0.40	823.68	0.74	0.02	1.39	867.58	0.56	-0.19	0.49	700.53	0.70	0.03
128×128	128×128	02	0.08	152.49	0.91	0.00	0.26	198.82	0.84	0.07	0.14	147.61	0.89	0.01
		04	0.10	180.90	0.90	0.01	0.31	215.25	0.82	0.04	0.14	159.81	0.88	0.01
		08	0.11	208.16	0.89	0.01	0.35	248.31	0.79	0.04	0.19	193.63	0.86	0.00
		16	0.14	233.27	0.88	0.00	0.46	322.48	0.76	0.01	0.19	207.31	0.85	0.00
		32	0.16	285.99	0.86	0.01	0.58	502.65	0.69	0.03	0.24	265.42	0.83	-0.01
		64	0.20	363.49	0.84	0.00	0.98	712.52	0.60	-0.02	0.33	343.82	0.79	-0.02
128×128	256×256	02	0.44	5695.65	0.71	0.42	0.68	3980.89	0.60	0.72	0.51	5324.13	0.67	0.51
		04	0.53	6783.89	0.66	0.51	0.79	4431.84	0.54	0.80	0.57	5875.94	0.64	0.56
		08	0.51	6732.88	0.66	0.51	0.75	4293.79	0.55	0.77	0.63	6570.71	0.60	0.63
		16	0.58	7794.21	0.62	0.59	0.79	4550.40	0.54	0.82	0.59	6143.63	0.61	0.59
		32	0.54	7225.34	0.63	0.54	0.75	4510.06	0.51	0.81	0.56	5566.12	0.63	0.53
		64	0.56	7446.43	0.61	0.55	0.75	3171.14	0.55	0.34	0.57	5231.05	0.61	0.49

Utvärderingen bekräftar även korsvalideringens slutsats att modellen presterar bättre om den tränas på svenska data och körs på tyska, jämfört med vice versa, vilket kan utläsas i Tabell 13.

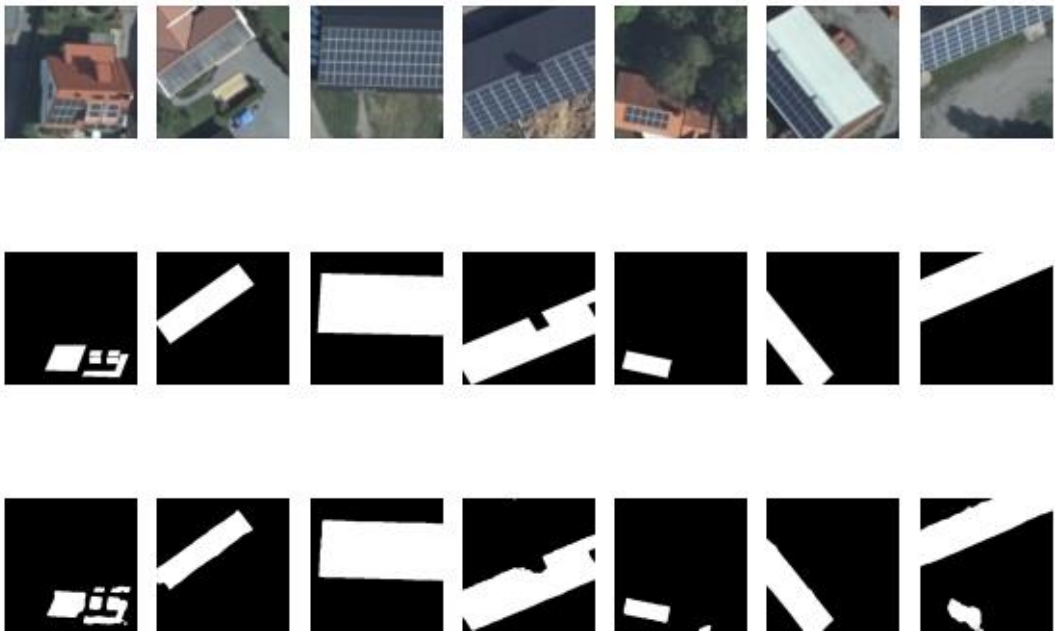
Tabell 13: De statistiska jämförelsemåtten *det genomsnittliga absoluta procentuella felet* (MAPE), *det genomsnittliga absoluta felet* (MAE), *bias* (B) och den *genomsnittliga skärningspunkten över unionen* (mIoU) för samtliga datauppsättningar, tränad på en datauppsättning och körd med en annan.

Train resolution	Test resolution	Batch size	German-Swedish				Swedish-German				Mix data-German				Mix data-Swedish			
			MAPE (%)	MAE	mIoU	B	MAPE (%)	MAE	mIoU	B	MAPE (%)	MAE	mIoU	B				
256×256	256×256	02	0.61	2504.53	0.66	0.44	0.31	1953.95	0.85	-0.01	0.14	408.62	0.91	0.03	0.05	384.22	0.94	-0.00
		04	0.64	2725.84	0.65	0.48	0.37	2544.75	0.82	0.01	0.18	468.90	0.89	0.04	0.05	392.63	0.94	-0.00
		08	0.65	3127.68	0.63	0.56	0.44	2920.97	0.80	-0.02	0.22	560.97	0.86	0.03	0.05	427.03	0.93	-0.00
		16	0.63	2966.54	0.64	0.53	0.67	3875.22	0.75	-0.15	0.29	739.74	0.82	0.04	0.07	492.12	0.92	-0.00
		32	0.61	3177.75	0.63	0.56	1.65	7030.18	0.55	-0.24	0.39	1276.17	0.76	0.08	0.13	799.45	0.89	-0.01
		64	0.65	3260.47	0.60	0.57	1.54	6823.66	0.54	-0.18	0.53	2037.68	0.66	0.20	0.27	1741.14	0.82	-0.02
128×128	128×128	021	0.56	559.36	0.69	0.37	0.26	454.36	0.83	-0.00	0.13	91.94	0.90	0.03	0.05	95.73	0.93	-0.01
		04	0.59	590.68	0.67	0.41	0.41	607.99	0.81	-0.01	0.18	111.25	0.88	0.03	0.06	98.21	0.92	-0.01
		08	0.55	635.79	0.67	0.45	0.54	886.73	0.76	0.04	0.24	160.43	0.84	0.04	0.07	114.90	0.91	-0.01
		16	0.60	702.88	0.65	0.49	0.69	935.87	0.76	-0.15	0.27	201.68	0.81	0.06	0.09	147.46	0.90	-0.02
		32	0.57	694.81	0.64	0.48	1.06	1288.80	0.69	-0.27	0.38	287.86	0.76	0.08	0.12	190.71	0.88	-0.02
		64	0.60	789.22	0.62	0.54	1.35	1507.14	0.60	-0.21	0.45	387.75	0.71	0.11	0.22	314.82	0.84	-0.04

Slutligen visar jämförelsen med andra studier att modellen presterar bättre än erhållna resultat i tidigare studier. Tabell 14 presenterar den statistiska jämförelsen och Figur 18 ett exempel på hur den semantiska segmenteringen kan se ut.

Tabell 14: Jämförelse mellan modellen (med en batchstorlek på 8) och ett antal relevanta forskningsresultat.

Studie/körning	MAPE [%]	MAE	mIoU [%]	B [%]
Yu et al. (2018) [11]	24,6	-	-	3,0
Castello et al. (2019) [24]	-	-	64	-
Zhuang et al. (2020) [71]	-	-	75	-
Parhar et al. (2020) [72]	-	-	82	-
Kasmi et al. (2022) [73]	-	-	86	-
Mayer et al. (2022) [18]	18,5	-	74	3,9
Our German 256×256	9	787,62	91	1
Our German 128×128	10	180,90	90	1
Our Swedish 256×256	35	998,83	81	4
Our Swedish 128×128	35	248,31	79	4
Our Mix 256×256	16	705,84	89	1
Our Mix 128×128	19	193,63	86	0,1



Figur 18: Exempel på bildsegmentering (ur den svenska datauppsättningen). Översta raden visar flygbilder innehållande solenergisystem, den mellersta raden visar den manuella segmenteringen av samma flygbilder och den tredje raden visar segmenteringen från den valda körningen av U-netmodellen.

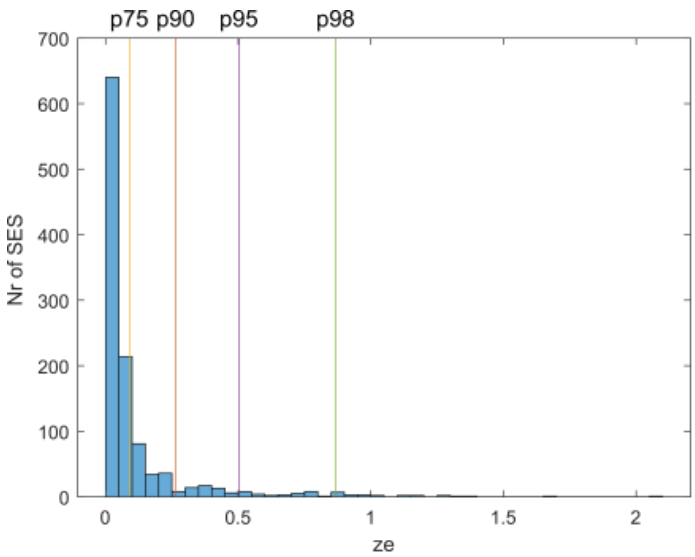
Resultat för orientering och skuggning

De riktningar (azimut) som modellen beräknar utvärderades genom att undersöka ett urval av 100 (av 1234) icke-vertikala solenergisystem belägna i Falu kommun. Tre system tilldelades fel azimut; de första två på grund av otillräcklig LiDAR-data och att de inte orienterades enligt den längsta sidan av närmaste byggnad och den tredje på grund av låg LiDAR-punkttäthet (vilket resulterade i hög ze-värde).
Dessutom, för platta tak, där azimuten inte är viktig, om man antar att ingen ställning för att öka lutningen används, returnerade modellen azimut även till norr. En utvidgning av modellen skulle vara att anta en azimut, till exempel i riktning mot den ”sydligaste” långsidan av solenergisystemet (se Figur 19a) och även anta att ett monteringssystem används, vilket resulterar i en lutning på cirka 10 grader, eftersom detta är vanlig praxis för att minska dammansamling på panelerna, men samtidigt begränsa intern skuggning och vindlast. När den sydligaste långsidan är nära väst eller öst är det troligt att solenergisystemet är vänd mot två riktningar (öst-väst) som visas i Figur 19b, vilket ibland är att föredra för en flackare produktionsprofil och tätare placerade paneler.

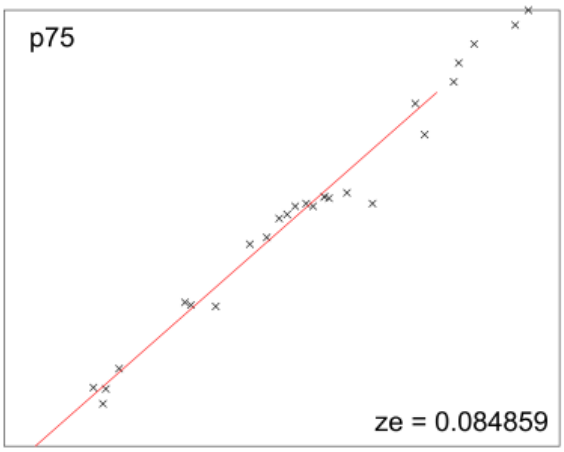
(a) (b)

Figur 19. Typiska arrangemang av solcellsmoduler på platta tak; (a) följer byggnadens sydligaste orientering eller (b) både öster och väster, och i båda fallen lutade ungefär 10 grader med hjälp av ställning. Ortofoto, 0.16 m färg © Lantmäteriet (2022).

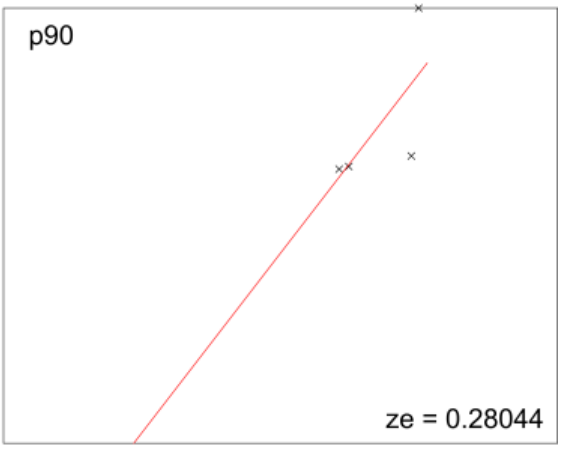
Hur väl modellen uppskattar lutningen är svårare att undersöka från flygbilder. Ett lågt ze-värde betyder dock en bra regression av LiDAR-punkterna och indikerar en trovärdig lutning. Figur 20a visar fördelningen av ze-värden för 1119 av 1276 solenergisystem i Falun. För de återstående 157 systemen var antalet LiDAR-punkter otillräckligt för en linjär regression och lutningen sattes därför till 26˚. Figur 20b) och c) är exempel på solenergisystem nära 75:e respektive 90:e percentilen. Av detta kan man dra slutsatsen att den modellerade lutningen förmodligen är representativ för minst 75 % av solenergisystemen.



(a)

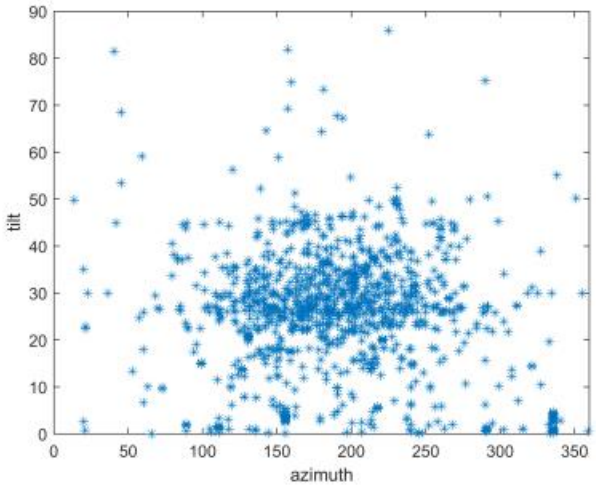


(b)

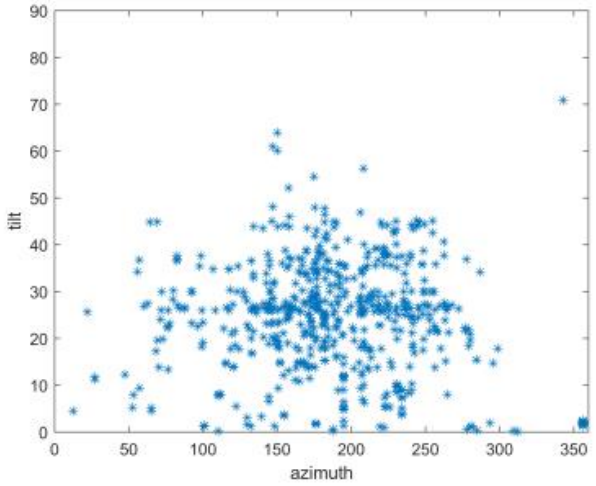


(c)

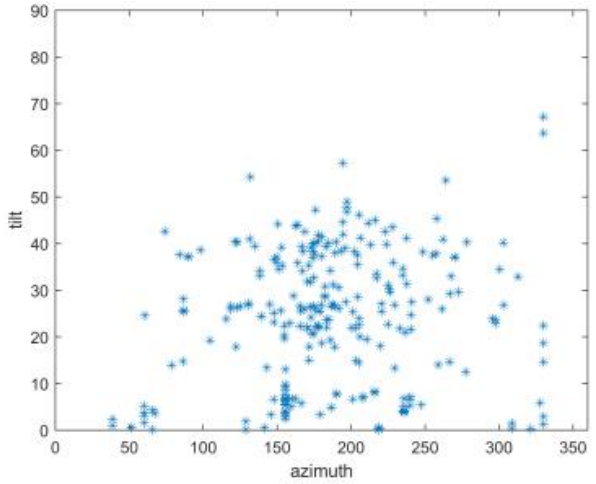
Figur 20. I (a) ett histogram av ze-värden för alla 1119 (av 1276) solenergisystem. I (b) och (c) exempel på linjärregression för SES nära 75:e respektive 90:e percentilen.



(a)



(b)



(c)

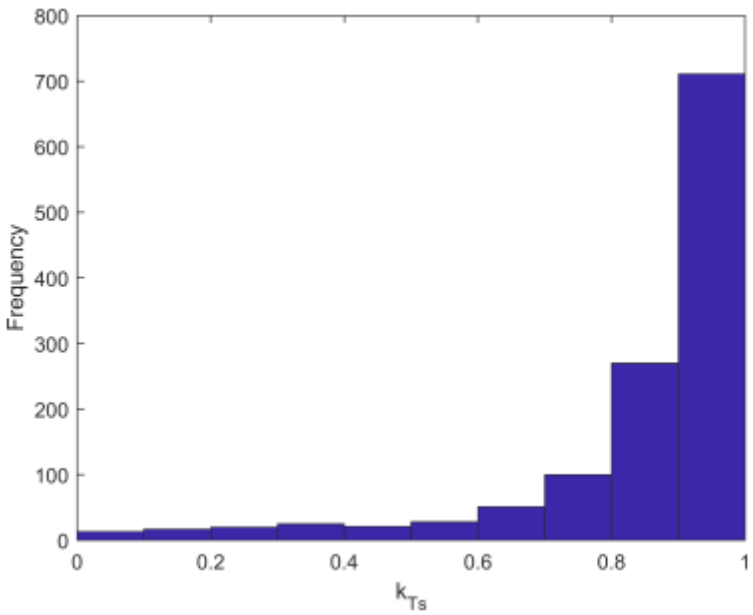
Figur 21. Tilt och azimut för varje solenergisystem i (a) Falun, (b) Knivsta respektive (c) Uppvidinge (där 180˚ är mot söder).

Med tanke på att lutningen och azimuten inte alltid, men förmodligen i de flesta fall, är väl modellerade, är det intressant att ta en titt på spridningen av lutningen och azimuten för varje kommun, vilket Figur 21 illustrerar. Ett solenergisystem med en lutning på mindre än 10˚ ligger förmodligen på ett platt tak. Det finns också system med otillräckliga LiDAR-data, mest utmärkande i Figur 21a som en koncentration av markörer för lutningar lika med 26˚.

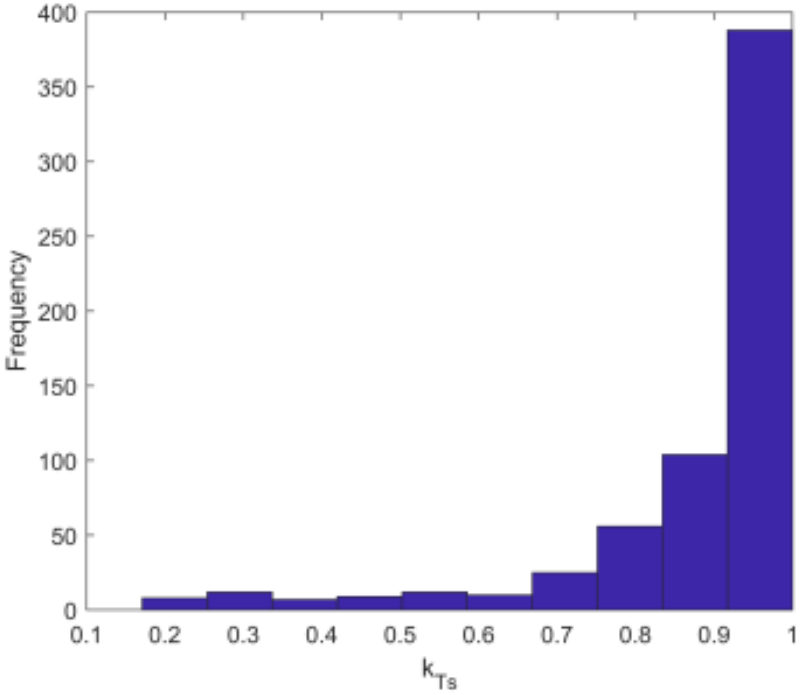
Utbytet från ett solenergisystem påverkas inte bara av orienteringen, utan också av skuggning från omgivningen, det vill säga strukturer, träd och gradienter. Påverkan från skuggning kan studeras genom att använda ett *skuggindex* [30], [74]:

$$k_{Ts} = \frac{I_{gs}}{I_g}, \qquad (1)$$

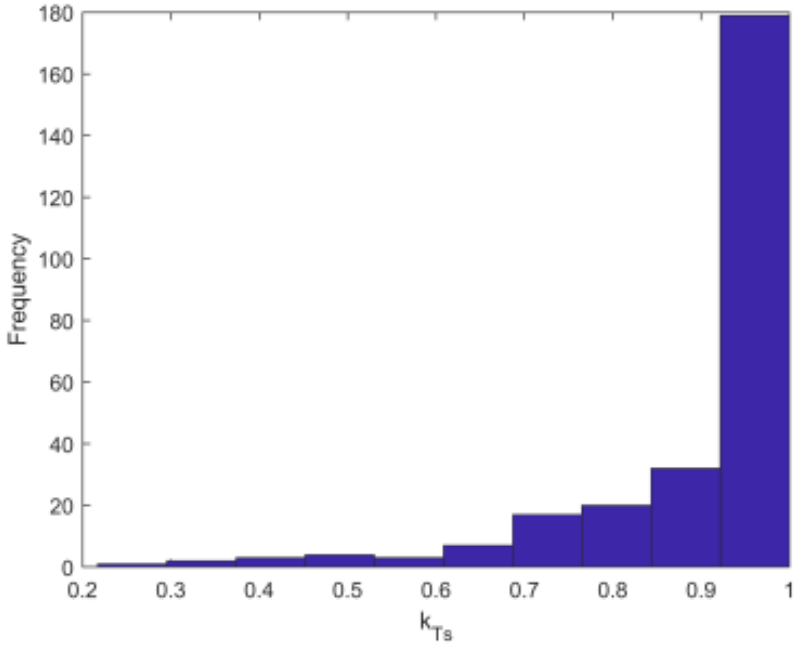
där I_{gs} and I_g är den skuggade respektive oskuggade globala årliga solinstrålningen på det lutade planet, här modellerad med hjälp av data från STRÅNG för 2021 [34]. Figur 22 illustrerar fördelningen av k_{Ts} för de tre kommunerna. k_{Ts}<0,4 beror i de flesta fall på modellfel; (i) för få LiDAR-punkter, vilket resulterar i en hög lutning, eller (ii) byggnader byggda efter LiDAR-undersökningen genomfördes, det vill säga solenergisystem som är placerade nära marken eller nära ett träd som fanns där när LiDAR-undersökningen ägde rum. För Falun gavs (42) vertikala solenergisystem en mycket låg k_{Ts}, till och med 0 för 13 av dem, vilket är naturligt eftersom modellen inte kan fånga vertikala system på ett bra sätt. En betydande del av solenergisystemen har dock en k_{Ts} mellan 0,4 och 0,8, vilket i allvarliga problem med skuggning. Det är troligt att träd, i vissa fall, har fällts efter LiDAR-undersökningen, men manuell inspektion av de (senare) flygbilderna tyder på att träden i de flesta fall finns kvar.



(a)



(b)



(c)

Figur 22. Histogram över skuggindex, k_{Ts}, för (a) Falun, (b) Knivsta respektive (c) Uppvidinge.

Anledningarna till att placera solenergisystem där de utsätts för skuggning eller orienteringen är långt ifrån optimal kan vara många, till exempel bristande

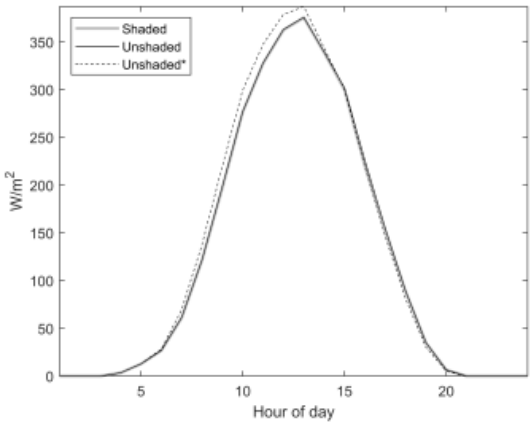
förståelse eller inget bättre alternativ för fastighetsägaren. Att studera alternativa platser i närheten med bättre orientering och mindre skuggning kan vara ett sätt att utforska hur energigemenskaper kan leda till en mer resurseffektiv användning av solenergisystem.

Slutligen jämfördes resultaten från modellen med uppmätta data från ett fåtal solcellssystem (se Tabell 15). Det finns inga mätningar av de direkta och diffusa solinstrålningskomponenterna nära de tre kommunerna, men den mesoskaliga solinstrålningsmodellen STRÅNG har dessa delkomponenter [34] som behövs för att modellera den totala instrålningen och därmed energiutbytet på det lutade planet. Men eftersom STRÅNG-modellen har låg noggrannhet på timbasis, men stämmer relativt väl på årsbasis undersöktes de uppmätta och modellerade genomsnittliga dygnsinstrålningsprofilerna. I Figur 23a och b presenteras de modellerade skuggade och oskuggade genomsnittliga dygnsinstrålningsprofilerna för två av systemen (1A respektive 4C) från Tabell 15. Den streckade linjen representerar azimut och lutning, angivna av ägaren. För system 1A (Figur 23a) avviker den modellerade och rapporterade azimuten och därför förskjuts den prickade linjen åt vänster (dvs. instrålningen är högre på morgonen) och middagstoppen är högre (azimut mer mot söder). Skuggningen i detta fall är försumbar (k_{Ts} = 99,4), medan den för system 4C (Figur 23b) är signifikant (k_{TS} = 56,3) på grund av ett närliggande stort träd, se Figur 23c.

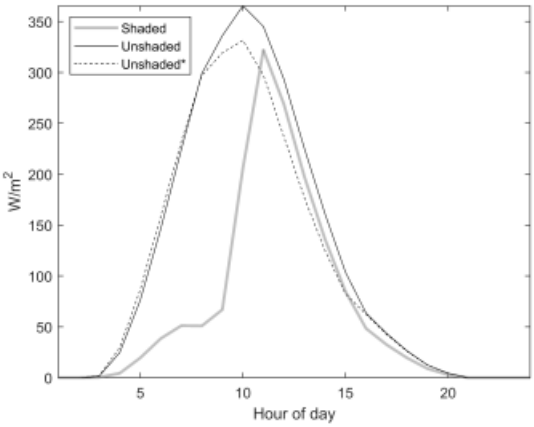
Tabell 15. Rapporterad och modellerad tilt, azimut och installerad (DC-)effekt för valideringsdatasetet. Stora bokstäver i ID-kolumnen representerar subsystem bakom mätaren.

		Rapporterad			Modellerad					Skillnader			
ID	**Kommun**	**DC (kW_p)**	**Tilt**	**Azimut**	**Area (m^2)**	**Tilt**	**Azimut**	**ze**	**DC* (kW_p)**	**Δ_{tilt}**	**Δ_{azimut}**	**Δ_{DC}**	**k_{Ts}**
1A	Upp.	11,4	22	40	72,29	21,8	52,91	0,023	14,46	-0,2	12,91	3,06	99,4
2A	Upp.	10,08	27	0	56,39	26,59	6,46	0,045	11,28	-0,41	6,46	1,20	98,6
3A	Knivsta	2,55	22	-10	19,1	20,49	-11,96	0,038	3,82	-1,51	-1,96	1,27	97,5
3B	Knivsta	5,95	22	80	41,4	21,54	78,04	0,079	8,28	-0,46	-1,96	2,33	99,1
4A	Knivsta	2,52	25	-100	17,56	23,26	-87,64	0,039	3,51	-1,74	12,36	0,99	67,4
4B	Knivsta	2,52	25	80	14,36	26,32	92,36	0,079	2,87	1,32	12,36	0,35	85,8
4C	Knivsta	8,93	42	-95	60,83	35,39	-82,28	0,058	12,17	-6,61	12,72	3,24	56,6
4D	Knivsta	8,93	42	85	41,14	36,88	97,72	0,059	8,29	-5,12	12,72	-0,70	93,3
5A	Falun	14,26	38	40	60,03	35,80	39,38	0,24	12,01	-2,20	-0,62	-2,25	93,2
5B	Falun	5,27	38	-50	28,60	36,08	-50,34	0,45	5,72	-1,92	-0,34	0,45	96,3

*Antar 20 % PV modulverkningsgrad vid Standard Test Conditions (STC).



(a)



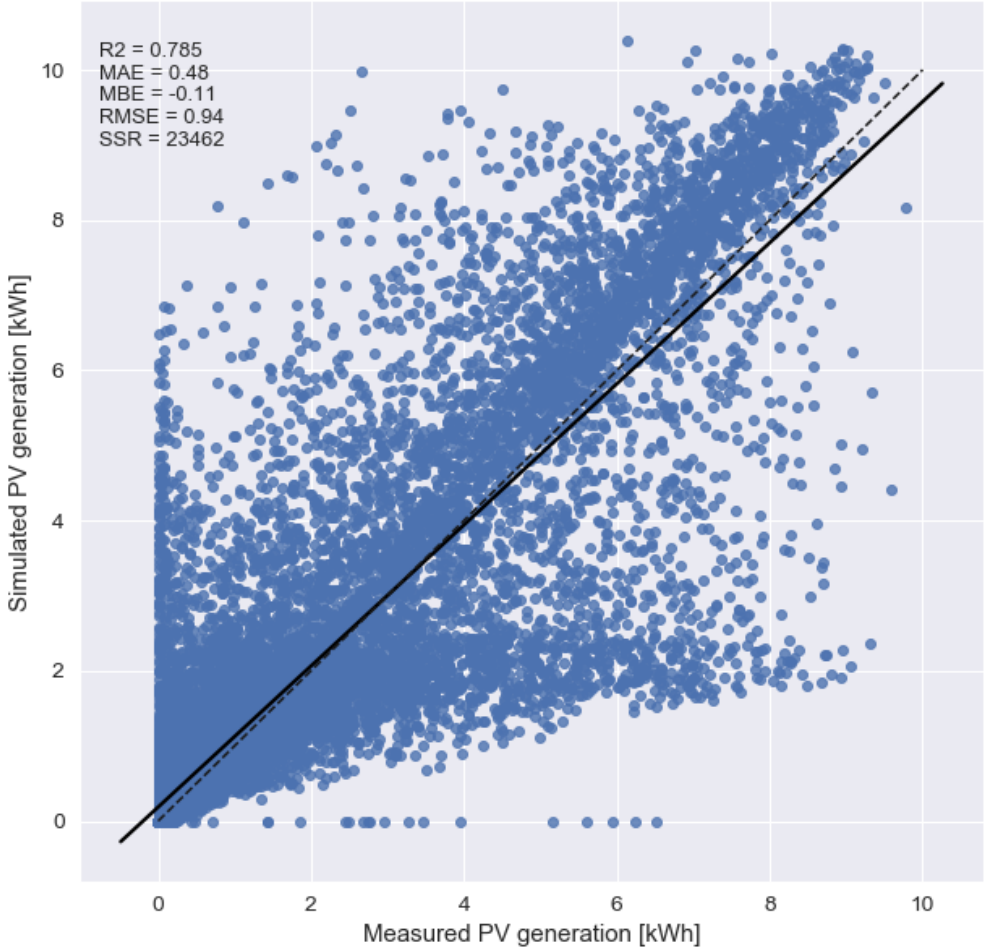
(b)

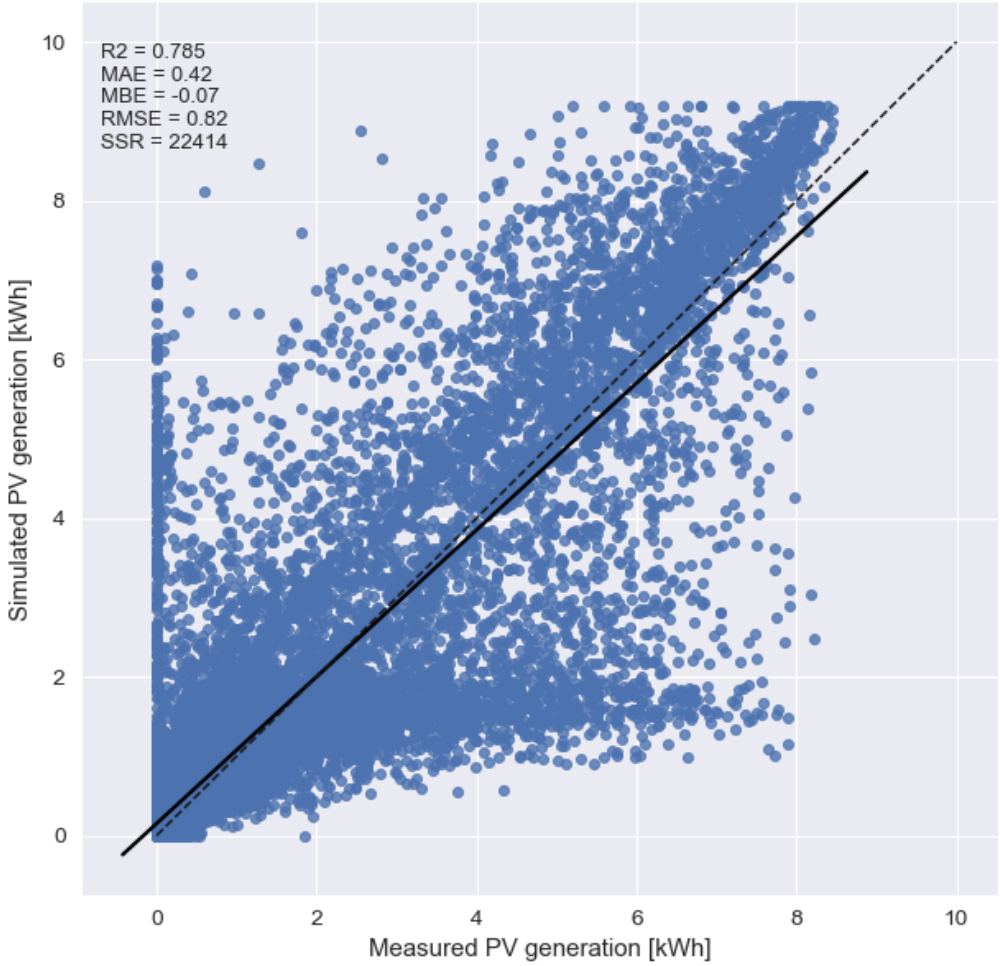
(c)

Figur 23. Genomsnittlig dygnsinstrålningsprofil för solcellsanläggningar i (a) Uppvidinge och (b) Knivsta, om man inkluderar skuggning (grå) eller inte (svart). Den streckade kurvan representerar den oskuggade instrålningen baserat på den inrapporterade lutningen och azimuten från solcellsanläggningens ägare. I (c) är det östvända systemet detsamma som i (b) och där ett stort träd begränsar effektuttaget på morgontimmarna.

Preliminära resultat för simulering av producerad energi

Den grunddata som genereras genom de tidigare beskrivna stegen/metoderna möjliggör simuleringar av alla enskilda existerande solcellsanläggningars solkraftsproduktion inom ett geografiskt område där hänsyn tas till den faktiska orienteringen hos anläggningar, något som tidigare inte har kunnat göras i Sverige [32] och med begränsad framgång internationellt [74]. Att utföra aggregerade simuleringar på kommunbasis fanns det dock inte utrymme att utföra inom detta forskningsprojekt. Vid publikationen av denna rapport hade däremot det examensarbete som beskrivs i under avsnittet Genomförande dock kommit så långt att en preliminär valideringen av simuleringsmodellen utvecklad i [44] kan presenteras. Simuleringen av solelproduktionen per timme för två distribuerade småskaliga villasystem belägna i Uppvidinge kommun valideras mot den faktiskt uppmätta produktionen på timbasis i Figur 24 för åren 2018–2021. Som synes visar de preliminära resultaten är att den simulerade solelproduktion från villasystemen stämmer relativt väl med den faktisk uppmätta solelproduktionen då determinationskoefficienter (R^2) på mellan 0,785 uppnåddes för båda referenssystemen, vilket i sammanhanget är tämligen höga värden. Detta är lovande resultat och som motiverar en fortsatt utveckling mot att kunna simulera den aggregerade solkraftsproduktionen på kommun- eller lågspänningsnivå.





Figur 24. Modellvalidering för solelproduktionssimuleringsmodellen från [44] per timme för åren 2018–2021 jämfört med den uppmätta kraftproduktionen för två villareferenssystem i Uppvidinge, med STRÅNGs globala horisontella instrålning och diffus horisontell instrålning som indata för transponeringsmodellen.

Diskussion

Faltningsnätverk för statistikgenerering

Till skillnad från majoriteten av litteraturen på området så har huvudfokus i det här projektet inte varit att förbättra CNN-klassificeringsalgoritmen för att nå maximal noggrannhet. Snarare har fokus legat på att utvärdera faltningsnätverk och flygbilder som verktyg för att skanna kommuner i syfte att upprätta kompletta databaser för statistiska ändamål.

Vid studiens genomförande innehöll processen initialt två steg som utfördes manuellt, men som teoretiskt går att automatisera. Det första var skapandet av ett polygon för areautbredningen av solenergisystemen i ortofotona. Genom utvecklingen av U-net modellen finns nu möjligheten att detta kan göras automatiskt vid eventuella framtida skanningar av kommuner.

Det andra manuella steget var klassificeringen av solenergiteknologierna, det vill säga som reguljära solcellsanläggningar, byggnadsintegrerade solcellsanläggningar eller solvärmeanläggningar. Den främsta anledningen till att det behövdes göras manuellt var bristen på träningsdata för solvärmeteknologin. Från de fyra kommunskanningarna i denna studie har däremot totalt 656 bildrutor över solvärmesystem skapats. Om fler kommuner skannas på samma sätt kommer antalet solvärmebilder att öka, vilket så småningom skulle uppgå till tillräckligt många bilder för att träna modellen för att automatiskt skilja de två teknologierna åt.

Utöver det fanns initialt en ambition att undersöka möjligheten att skilja mellan poly kristallina, monokristallina och tunnfilmssolceller. Dock identifierades ingen praktisk möjlighet att klassificera solcellsanläggningarna i träningsbilderna mellan de olika solcellsteknologierna polykristallina, monokristallina och tunnfilmssolceller. Utifrån den uppnådda praktiskt erfarenheten av metodens kapacitet under projektets gång bedöms förmågan att framöver kunna skilja mellan de väldigt lika solcellsteknologierna polykristallina, monokristallina och tunnfilmssolceller som mycket utmanande till gränsen på omöjligt.

Den lägre noggrannheten i detta projekt jämfört med vissa tidigare studier, se Tabell 5, kan förklaras av flera faktorer. En uppenbar faktor är flygfotoupplösningen, eftersom hög upplösning behövs för att tillförlitligt klassificera finkorniga objekt som solenergimoduler.

En annan förklaring är att den här studien omfattar kompletta kommunskanningar i glesbebyggda Sverige, vilket innebär en mycket högre andel negativa bildprover än i de nämnda studierna. Som Tabell 5 visar är andelen positiva bildrutor en faktor på ~10 till ~500 lägre för de svenska kommunskanningarna, i jämförelse med de tidigare internationella studierna. En större andel negativa bildrutor kommer naturligt att generera relativt fler FP-bilder jämfört med TP-bilder, vilket resulterar i en sämre precision. Denna förklaring bekräftas till stor del av att en betydligt högre precision, 93,4%, jämfört med tidigare 63,9%, uppnåddes vid körning av den identiskt tränade CNN-algoritmen på OpenNRW_Test_16-testuppsättningen från [25] som har en nästan 10 gånger högre andel positiva bilder.

När det gäller återkallelsen är studiens 81,8% lägre jämfört med tidigare studier, men inte så låg som precisionsmåttet. Eftersom huvudmålet är att utvärdera hur användbar kartläggning av decentraliserade solenergisystem med hjälp av flygbilder och en CNN-klassificeringsalgoritm är för att skapa en så omfattande databas som möjligt, är en hög återkallande viktigare än en hög precision, då återkallelsen mäter förhållandet mellan alla korrekt identifierade solenergisystem och det totala antalet faktiska existerande solenergisystem.

En möjlig anledning till den något lägre återkallelsen kan vara att en mer korrekt grundsanningen uppnåts i det här projekt i jämförelse med tidigare internationella studier. Enligt vår vetskap så har alla studierna i Tabell 5 skapat sina grundsanningar genom att manuellt registrera solcellsanläggningar i flygfoton. Endast två studier [17], [18] hade sedan möjligheten att jämföra sina skapade databaser över anläggningar med register över platser för kända solcellssystems, vilket dessutom var register med kända betydande fel enligt [16], [17].

Följaktligen är det troligt att vissa mycket svåra fall (som ramlösa moduler på svart tak med till exempel hög lutning, skuggor eller reflektioner) inte har klassificerats korrekt som positiva i de tidigare internationella studierna, vars resultat jämförts med den här studiens i Tabell 5. Det innebär att det typiska antagandet att grundsanningen är ofelbar är inkorrekt i dessa studier, eftersom felfrekvensen för annotatorerna är omöjlig att veta utan några kända platser att jämföra med [16]. Till exempel krävs det en mycket skicklig eller tursam person för att upptäcka solcellssystemet i Figur 3. I [16] uppskattar de att annotatorerna missade cirka 3,8% av de synliga panelerna och 5,0% om även icke-synliga system ingick. Eftersom två olika register användes för att korskontrollera resultaten i det här projektet, och därtill även inspektioner på plats, bedöms det här projektets grundsanning som mer tillförlitlig än i någon av de tidigare internationella studierna på området, då det innehåller positiva bildrutor som det är mycket svårt för en person (eller en CNN-klassificeringsalgoritm) att upptäcka.

Ett sätt att förbättra processen, och möjligen noggrannheten, kan vara att implementera en mosaikmetod (eng.: *mosaicking with sliding windows approach*), vilket gjordes i [26] med förbättrat resultat vid ett minskat stegvärde. En mosaikmetod resulterar i överlappande bildrutor, vilket skulle göra det möjligt att utesluta gränsrutor från analysen utan att förlora information. Ett sådant tillvägagångssätt skulle dock öka den nödvändiga datakapaciteten för att utföra beräkningarna [26], vilket, beroende på det använda stegvärdet, skulle kunna begränsa den praktiska möjligheten att skanna så stora områden som kommuner.

Ett annat sätt att förbättra noggrannheten skulle kunna vara att implementera förbearbetningssteg, som att endast skanna bilder som överlappar byggnader, eller efterbearbetningssteg, som [10] som raderade alla bildrutor över växthus. Att bara utvärdera bilder som täcker byggnader skulle ha haft en positiv inverkan på noggrannheten, eftersom många FP-bildrutor i de första skanningarna täckte markdetaljer som symmetriskt plöjda åkrar, vägar, sjöar med vågor och kraftledningar. En sådan åtgärd skulle dock inneburit att 8 markmonterade solenergisystemen som upptäcktes i de tre kommunerna skulle ha missats, och skulle därför ha resulterat i lägre noggrannhet i de objektbaserade måtten. I motsats

till Nederländerna, där det finns många växthus [10], identifierades ingen specifik byggnadstyp som ofta ledde till FP-klassificering. Ett potentiellt efterbearbetningssteg som skulle ha förbättrat noggrannheten i Sverige hade i stället varit att ta bort FP på tydligt norrvända tak.

Socioekonomisk studie

I den socioekonomiska studien gjordes ett antal antaganden som försvagar säkerheten i resultatet. Dels antogs ägaren till bostaden vara den som äger solenergisystemen, trots att information om hur systemet finansierats inte fanns tillgängliga. Dels fanns förvisso både anskaffningsdatum av bostäderna och driftsättningsdatum för solcellssystem att tillgå, men eftersom inget register finns för solvärme kunde information solvärmesystemens driftsättningsdatum inte utrönas. Det innebär att det heller inte gick att fastställa huruvida det nuvarande hushållet skaffat solvärmesystemet eller om det gjordes av en tidigare ägare. Vi förväntar oss således större osäkerheter för resultaten gällande solvärmesystemen. Det bör också noteras att datauppsättningen för denna studie är betydligt mindre jämfört med flera av studierna som resultatet jämförs med, vilket kan vara flera hundratusentals system [45]–[47], [50], [53], [58], [64], [75], [76]. Dock kan man argumentera för att vår datauppsättning väger upp sin storlek med att de socioekonomiska variablerna inkomst, ålder och kön kunde presenteras på hushållsnivå.

Slutsatsen att koncentrationen av solenergisystem är högre på landsbygden än i urbana och centrala områden är inte unik, utan har också observerats i flera andra länder, såsom Australien [75], Belgien [45], Danmark [65], Tyskland [52], Nederländerna [58], Storbritannien [48], [49], [53], [67], [77] och USA [55], [57], [64]. Enligt vår vetskap är det däremot den första studien som observerat samma mönster för solvärme.

Detta globala mönster kan vid första anblick tyckas oförenligt med de rapporterade observationerna om positiv korrelation mellan solenergiutrullning och hög inkomst och befolkningstäthet. Dessa faktorer förknippas nämligen med städer och storstadsområden. Denna studie visar mycket riktigt att den genomsnittliga månadsinkomsten för alla Landsbygds-DeSOn var 27 920 SEK, jämfört med 29 459 SEK in Centrum-DeSOn. Förklaringen förväntas vara den samma som lanserats i [45], [48], [50], [52], [55], [75], nämligen att även om det generellt är hushåll med hög inkomst och områden med hög befolkningstäthet i städer, så bor många fler människor i flerbostadshus, med begränsade möjligheter att installera solenergisystem i anslutning till sin bostad.

Det mest betydande observationen i denna studie är sannolikt den måttligt positiva (0,36), som vid borttagande av extremvärden blev medelstark positiv (0,61), korrelationen mellan solcells- och solvärmeutbyggnaden. Det stärker tidigare dokumentation som beskriver hur kunskap om en teknologi (i det här fallet solfångare), ökar sannolikheten för upptagandet av en ny, likartad, teknologi (i det här fallet solceller) [78]. Vidare bör noteras att relativt många hushåll har installerat både solcell- och solvärmesystem. Antagandet att solvärme installerades först baseras på det faktum att solvärmemarknaden generellt startade ett decennium före

solcellsutrullningen i Sverige. Hela 6,7% av hushållen med solvärme har även installerat ett solcellssystem (5,6 % i Falun, 14,3 % i Knivsta och 4,3 % i Uppvidinge). Det här skulle kunna ses som någon slags hushållsintern *spill over peer-effekt* mellan de två solenergiteknologierna.

U-net

Givet att CNN-baserade segmenteringsmetoder numera är praxis för automatisk bildigenkänning och identifiering av solenergisystem, och att U-net-modeller har blivit mycket populära för global lokalisering och kontextkartläggning, är det logiskt att para ihop dessa. Precis som i de andra projektstegen har dataunderlaget sina utmaningar, i det här fallet kombinationen mellan brist på data och nätverksparametriseringen som metoden kräver. Trots detta genererar modellen utvecklade i det här projektet betydligt bättre resultat än tidigare publicerade forskningsstudier. Den här studien har förbättrat förståelsen för solenergisystemsidentifikation genom flygfoton och vilken arkitektur som krävs för att erhålla goda resultat. Med den erfarenheten som projektet genererat rekommenderas att använda den modifierade versionen av U-net-modellen, se Bilaga C, för att öka noggrannheten i eventuella vidare studier.

För att öka generaliserbarheten vore det önskvärt att utveckla och addera en så kallad multi-klass U-net-modell som kan separera solfångare från solceller. Det är däremot, precis som i de andra delprojekten, inte helt trivialt eftersom det saknas stora träningsdatauppsättningar med en avskiljning mellan dessa teknologier, främst på grund av dess snarlika utseenden. Den enda datauppsättningen som särskiljer dessa två teknologier inom den internationella forskningen är enligt författarnas vetskap den som genereras inom det här projektet, närmare bestämt under arbetspaketet med *Framtagandet av en grundsanning*.

Eftersom den här studien bevisat att man med hög noggrannhet kan uppskatta arean för solenergianläggningar genom flygfoton är en tydlig vidareutveckling att para ihop den med 3D-data eller LiDAR-data för att erhålla orientering och vinkel på systemen i fråga. På så vis kan kapaciteten uppskattas med hjälp av antaganden om moduleffektivitet.

Orientering och skuggning

Det finns ett antal förbättringsåtgärder för modellen, samtidigt som man måste vara försiktig med att göra den för specialiserad mot den begränsade mängden data från de tre kommunerna som varit tillgängliga i projektet.

Icke-ortogonala solenergisystem bör inte forceras till ortogonalitet, medan mycket skeva solenergisystempolygon på grund av stort avstånd (och brant tak) från nadir borde göra det. Detta är en känslig balans, men kanske löses detta genom att efter den linjära regressionen flytta hörnen av den ursprungliga polygonen istället för den ortogonaliserade och ortogonalisera polygonen efter att den har flyttats. På detta sätt är storleken på solenergisystemen sannolikt mer korrekt (och i sin tur den installerade kapaciteten), och tröskeln (nu 35 grader, vilket är väldigt högt) för ortogonaliseringen kan göras mer restriktiv.

(Nära) platta solenergisystem ges nu en azimut i en godtycklig riktning, men en låg lutning (på en stor byggnad) skulle istället kunna resultera i en fast lutning (på 10 grader) på långsidan av solenergisystemen, och om det är nära till öst/väst, kan man anta att den också är orienterad i både väst och öst, annars endast ”sydligast” (vilket åskådliggörs i Figur 19).

Spridningen vad gäller både orientering och skuggning utgör väldigt intressanta resultat då det indikerar att suboptimala tak används ur ett systemperspektiv. Om man inför starkare styrmedel, exempelvis som stimulerar energigemenskaper där energi delas mellan byggnader och elanvändare, tycks det finnas en potential att utnyttja solenergisystem mer effektivt genom att de bästa taken (eller markerna) används, dvs. där orienteringen är mer gynnsam och skuggningen försumbar.

Publikationslista

Inom forskningsprojektet har följande publikationer tagits fram.

Fullskaliga vetenskapliga tidskriftsartiklar:

1. Frimane Â, Johansson R, Munkhammar J, Lingfors D, Lindahl J. Semantic segmentation of solar energy systems in aerial images using the U-net model [Unpublished manuscript] 2023. *Inskickad till tidskriften Solar Energy och accepterad med ändringar.*

2. Lindahl J, Johansson R, Lingfors D. Mapping of decentralised photovoltaic and solar thermal systems by remote sensing aerial imagery and deep machine learning for statistic generation [Unpublished manuscript] 2023. *Inskickad till tidskriften Applied Energy.*

Vetenskaplig konferensartikel:

1. Lindahl J, Ekbring S, Johansson R, Lingfors D, Munkhammar J. Socioeconomic and demographic factors behind the deployment of domestic photovoltaic and solar thermal systems in three Swedish municipalities. 8th World Conference on Photovoltaic Energy Conversion, 2022, p. 1530–40. https://doi.org/10.4229/WCPEC-82022-5DO.10.5

Examensarbete:

1. Ekbring S. Key Socioeconomic factors for domestic solar energy — An interdisciplinary analysis of the characteristics of photovoltaic and solar thermal installations in three Swedish municipalities, 2022, Uppsala Universitet. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1697980&dswid=4290

Referenser

[1] W. Weiss and M. Spörk-Dür, “Solar Heat Worldwide – 2021 Edition,” 2021.
[2] IEA PVPS task 1 *et al.*, “Trends in Photovoltaic Applications — 2022,” 2022. [Online]. Available: https://iea-pvps.org/wp-content/uploads/2020/02/5319-iea-pvps-report-2019-08-lr.pdf
[3] M. Jaxa-Rozen and E. Trutnevyte, “Sources of uncertainty in long-term global scenarios of solar photovoltaic technology,” *Nat Clim Chang*, vol. 11, no. 3, pp. 266–273, 2021, doi: 10.1038/s41558-021-00998-8.
[4] C. Wilson, A. Grubler, N. Bento, S. Healey, S. De Stercke, and C. Zimm, “Granular technologies to accelerate decarbonization,” *Science (1979)*, vol. 368, no. 6486, pp. 36–39, 2020.
[5] IEA PVPS *et al.*, *Data model for PV systems — Data model and data acquisition for PV registration schemes and grid connection evaluations — Best practice and recommendations*. 2020.
[6] SolarPower Europe, “Global Market Outlook for Solar Power — 2021–2025,” 2021. [Online]. Available: http://www.solarpowereurope.org/insights/global-market-outlook/
[7] International Energy Agency (IEA), “World Energy Outlook 2020,” 2020. [Online]. Available: https://www.oecd-ilibrary.org/energy/world-energy-outlook-2020_557a761b-en
[8] Mordor Intelligence, “Global solar thermal market (2022 - 2027),” 2022.
[9] S. Ren *et al.*, “Automated Extraction of Energy Systems Information from Remotely Sensed Data: A Review and Analysis,” *Appl Energy*, vol. 326, no. September, p. 119876, 2022, doi: 10.1016/j.apenergy.2022.119876.
[10] B. B. Kausika, D. Nijmeijer, I. Reimerink, P. Brouwer, and V. Liem, “GeoAI for detection of solar photovoltaic installations in the Netherlands,” *Energy and AI*, vol. 6, no. August, p. 100111, 2021, doi: 10.1016/j.egyai.2021.100111.
[11] J. Yu, Z. Wang, A. Majumdar, and R. Rajagopal, “DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States,” *Joule*, vol. 2, no. 12, pp. 2605–2617, 2018, doi: 10.1016/j.joule.2018.11.021.
[12] X. Hou, B. Wang, W. Hu, L. Yin, and H. Wu, “SolarNet: A Deep Learning Framework to Map Solar Power Plants In China From Satellite Imagery,” in *ICLR 2020 Workshop on Tackling Climate Change with Machine Learning*, 2020. [Online]. Available: http://arxiv.org/abs/1912.03685
[13] L. Kruitwagen, K. T. Story, J. Friedrich, L. Byers, S. Skillman, and C. Hepburn, “A global inventory of photovoltaic solar energy generating units,” *Nature*, vol. 598, pp. 604–610, 2021, doi: https://doi.org/10.1038/s41586-021-03957-7.
[14] J. Yuan, H. H. L. Yang, O. A. Omitaomu, and B. L. Bhaduri, “Large-scale solar panel mapping from aerial images using deep convolutional networks,” in *2016 IEEE International Conference on Big Data, Big Data 2016*, IEEE, 2016, pp. 2703–2708. doi: 10.1109/BigData.2016.7840915.
[15] J. M. Malof, B. Li, B. Huang, K. Bradbury, and A. Stretslov, “Mapping solar array location, size, and capacity using deep learning and overhead imagery,” *ArXiv*, vol. abs/1902.1, 2019, [Online]. Available: http://arxiv.org/abs/1902.10895
[16] W. Hu *et al.*, “What you get is not always what you see—pitfalls in solar array assessment using overhead imagery,” *Appl Energy*, vol. 327, no. July, p. 120143, 2022, doi: 10.1016/j.apenergy.2022.120143.
[17] B. Rausch *et al.*, “An enriched automated PV registry: Combining image recognition and 3D building data,” *ArXiv*, no. December, 2020.
[18] K. Mayer *et al.*, “3D-PV-Locator: Large-scale detection of rooftop-mounted photovoltaic systems in 3D,” *Appl Energy*, vol. 310, no. December 2021, p. 118469, 2022, doi: 10.1016/j.apenergy.2021.118469.
[19] Z. Wang, M.-L. Arlt, C. Zanocco, A. Majumdar, and R. Rajagopal, “DeepSolar++: Understanding residential solar adoption trajectories with computer vision and technology diffusion models,” *Joule*, pp. 1–15, 2022, doi: 10.1016/j.joule.2022.09.011.
[20] J. Lindahl, S. Ekbring, R. Johansson, D. Lingfors, and J. Munkhammar, “Socioeconomic and demographic factors behind the deployment of domestic photovoltaic and solar thermal systems in three Swedish municipalities,” in *8th World Conference on*

Photovoltaic Energy Conversion, 2022, pp. 1530–1540. doi: 10.4229/WCPEC-82022-5DO.10.5.

[21] Z. Xia *et al.*, "Mapping the rapid development of photovoltaic power stations in northwestern China using remote sensing," *Energy Reports*, vol. 8, pp. 4117–4127, 2022, doi: 10.1016/j.egyr.2022.03.039.

[22] K. Bradbury *et al.*, "Distributed solar photovoltaic array location and extent dataset for remote sensing object identification," *Sci Data*, vol. 3, no. December, pp. 1–9, 2016, doi: 10.1038/sdata.2016.106.

[23] G. Kasmi *et al.*, "A crowdsourced dataset of aerial images with annotated solar photovoltaic arrays and installation metadata," *Sci Data*, vol. 10, no. 59, pp. 2052–4463, 2023, doi: https://doi.org/10.1038/s41597-023-01951-4.

[24] R. Castello, S. Roquette, M. Esguerra, A. Guerra, and J. L. Scartezzini, "Deep learning in the built environment: Automatic detection of rooftop solar panels using Convolutional Neural Networks," *J Phys Conf Ser*, vol. 1343, no. 1, pp. 1–6, 2019, doi: 10.1088/1742-6596/1343/1/012034.

[25] K. Mayer, Z. Wang, M.-L. Arlt, D. Neumann, and R. Rajagopal, "DeepSolar for Germany: A deep learning framework for PV system mapping from aerial imagery," in *Proceedings of the 3rd International Conference on Smart Energy Systems and Technologies (SEST)*, 2020, pp. 11–16. doi: 10.1109/SEST48500.2020.9203258.

[26] M. V. C. V. da Costa *et al.*, "Remote sensing for monitoring photovoltaic solar plants in brazil using deep semantic segmentation," *Energies (Basel)*, vol. 14, no. 10, pp. 1–15, 2021, doi: 10.3390/en14102960.

[27] X. Zhang, M. Zeraatpisheh, M. M. Rahman, S. Wang, and M. Xu, "Texture is important in improving the accuracy of mapping photovoltaic power plants: A case study of Ningxia autonomous region, China," *Remote Sens (Basel)*, vol. 13, no. 19, 2021, doi: 10.3390/rs13193909.

[28] J. M. Malof, L. M. Collins, and K. Bradbury, "A deep convolutional neural network, with pre-training, for solar photovoltaic array detection in aerial imagery," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2017-July, pp. 874–877, 2017, doi: 10.1109/IGARSS.2017.8127092.

[29] Â. Frimane, R. Johansson, J. Munkhammar, D. Lingfors, and J. Lindahl, "Identifying small decentralized solar systems in aerial images using deep learning," *Solar Energy*, 2023.

[30] D. Lingfors, J. Bright, N. Engerer, J. Ahlberg, S. Killinger, and J. Widén, "Comparing the capability of low- and high-resolution LiDAR data with application to solar resource assessment, roof type classification and shading analysis," *Appl Energy*, vol. 205, no. June, pp. 1216–1230, 2017, doi: 10.1016/j.apenergy.2017.08.045.

[31] N. Etherden, J. Ahlberg, D. Lingfors, and K. Kvamme, "Calculating the hosting capacity of electrical network with high penetration of solar PV," in *Proceedings ot the 8th International Workshop on the Integration of Solar Power into Power Systems*, 2018.

[32] U. H. Ramadhani, D. Lingfors, J. Munkhammar, and J. Widén, "On the properties of residential rooftop azimuth and tilt uncertainties for photovoltaic power generation modeling and hosting capacity analysis," *Solar Energy Advances*, vol. 3, no. February, p. 100036, 2023, doi: 10.1016/j.seja.2023.100036.

[33] P. E. Campana *et al.*, "A gridded optimization model for photovoltaic applications," *Solar Energy*, vol. 202, no. March, pp. 465–484, 2020, doi: 10.1016/j.solener.2020.03.076.

[34] SMHI, "STRÅNG - a mesoscale model for solar radiation," 1999. https://strang.smhi.se/ (accessed Jan. 30, 2023).

[35] SMHI, "Analysmodell (MESAN)," 2013. https://www.smhi.se/data/oppna-data/meteorologiska-data/analysmodell-mesan-1.30445 (accessed Jan. 30, 2023).

[36] A. Palm, "Local factors driving the diffusion of solar photovoltaics in Sweden: A case study of five municipalities in an early market," *Energy Res Soc Sci*, vol. 14, pp. 1–12, 2016, doi: 10.1016/j.erss.2015.12.027.

[37] Falu municipality, "Falun — Kommunfakta," 2022. https://www.falun.se/kommun--demokrati/kommunfakta.html (accessed Aug. 22, 2022).

[38] Statistics Sweden, "Average age of the population by region and sex. Year 1998 - 2021," 2022.

https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__BE__BE0101__BE0101B/BefolkningMedelAlder/table/tableViewLayout1/ (accessed Aug. 22, 2022).

[39] Uppvidinge municipality, "Uppvidinge — Kommunfakta." https://www.uppvidinge.se/kommun-och-politik/kommunfakta.html (accessed Aug. 22, 2022).

[40] U. H. Ramadhani, B. Tsegai, C. Apelryd, S. Ekbring, J. Munkhammar, and J. Widén, "Estimating the Socio-economic drivers of residential photovoltaic systems adoption in a Swedish city," in *11th Solar & Storage Power System Integration Workshop*, 2021, pp. 1–6.

[41] Statistics Sweden, "Demografiska statistikområden, en ny regional indelning under kommuner." https://www.scb.se/hitta-statistik/artiklar/2018/demografiska-statistikomraden-en-ny-regional-indelning-under-kommuner/ (accessed Aug. 22, 2022).

[42] C. F. Dormann *et al.*, "Collinearity: A review of methods to deal with it and a simulation study evaluating their performance," *Ecography*, vol. 36, no. 1, pp. 27–46, 2013, doi: 10.1111/j.1600-0587.2012.07348.x.

[43] Lantmäteriet, "Laserdata Nedladdning, skog." https://www.lantmateriet.se/sv/geodata/vara-produkter/produktlista/laserdata-nedladdning-skog/#anchor-0 (accessed Apr. 18, 2023).

[44] J. Lindahl, P. E. Campana, and M. Kofod-Hansen, "Beräkningsverktyg för optimering av solcellsparker," 2023. [Online]. Available: https://energiforsk.se/media/32376/2023-933-bera-kningsverktyg-fo-r-optimering-av-solcellsparker.pdf

[45] O. De Groote, G. Pepermans, and F. Verboven, "Heterogeneity in the adoption of photovoltaic systems in Flanders," *Energy Econ*, vol. 59, pp. 45–57, 2016, doi: 10.1016/j.eneco.2016.07.008.

[46] A. J. Schaffer and S. Brun, "Beyond the sun—Socioeconomic drivers of the adoption of small-scale photovoltaic installations in Germany," *Energy Res Soc Sci*, vol. 10, pp. 220–227, 2015, doi: 10.1016/j.erss.2015.06.010.

[47] S. Dharshing, "Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic (PV) systems in Germany," *Energy Res Soc Sci*, vol. 23, pp. 113–124, 2017, doi: 10.1016/j.erss.2016.10.012.

[48] N. Balta-Ozkan, J. Yildirim, and P. M. Connor, "Regional distribution of photovoltaic deployment in the UK and its determinants: A spatial econometric approach," *Energy Econ*, vol. 51, no. 2015, pp. 417–429, 2015, doi: 10.1016/j.eneco.2015.08.003.

[49] L.-L. Richter, "Social Effects in the Diffusion of Solar Photovoltaic Technology in the UK," *Cambridge Working Papers in Economics*, 2013.

[50] J. Sommerfeld, L. Buys, K. Mengersen, and D. Vine, "Influence of demographic variables on uptake of domestic solar photovoltaic technology," *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 315–323, 2017, doi: 10.1016/j.rser.2016.09.009.

[51] Y. Zhang, J. Song, and S. Hamori, "Impact of subsidy policies on diffusion of photovoltaic power generation," *Energy Policy*, vol. 39, no. 4, pp. 1958–1964, 2011, doi: 10.1016/j.enpol.2011.01.021.

[52] S. Müller and J. Rode, "The adoption of photovoltaic systems in Wiesbaden, Germany," *Economics of Innovation and New Technology*, vol. 22, no. 5. pp. 519–535, 2013. doi: 10.1080/10438599.2013.804333.

[53] N. Balta-Ozkan, J. Yildirim, P. M. Connor, I. Truckell, and P. Hart, "Energy transition at local level: Analyzing the role of peer effects and socio-economic factors on UK solar photovoltaic deployment," *Energy Policy*, vol. 148, no. PB, p. 112004, 2021, doi: 10.1016/j.enpol.2020.112004.

[54] T. Islam and N. Meade, "The impact of attribute preferences on adoption timing: The case of photo-voltaic (PV) solar cells for household electricity generation," *Energy Policy*, vol. 55, pp. 521–530, 2013, doi: 10.1016/j.enpol.2012.12.041.

[55] M. Graziano and K. Gillingham, "Spatial patterns of solar photovoltaic system adoption: The influence of neighbors and the built environment," *J Econ Geogr*, vol. 15, no. 4, pp. 815–839, 2015, doi: 10.1093/jeg/lbu036.

[56] B. Bollinger and K. Gillingham, "Peer Effects in the Diffusion of Solar Photovoltaic Panels," *Marketing Science*, vol. 31, no. 6, pp. 900–912, 2012, doi: 10.1287/mksc.1120.0727.

[57] C. L. Kwan, “Influence of local environmental, social, economic and political variables on the spatial distribution of residential solar PV arrays across the United States,” *Energy Policy*, vol. 47, pp. 332–344, 2012, doi: 10.1016/j.enpol.2012.04.074.
[58] R. Bernards, J. Morren, and H. Slootweg, “Development and Implementation of Statistical Models for Estimating Diversified Adoption of Energy Transition Technologies,” *IEEE Trans Sustain Energy*, vol. 9, no. 4, pp. 1540–1554, 2018, doi: 10.1109/TSTE.2018.2794579.
[59] E. Sardianou and P. Genoudi, “Which factors affect the willingness of consumers to adopt renewable energies?,” *Renew Energy*, vol. 57, pp. 1–4, 2013, doi: 10.1016/j.renene.2013.01.031.
[60] K. Willis, R. Scarpa, R. Gilroy, and N. Hamza, “Renewable energy adoption in an ageing population: Heterogeneity in preferences for micro-generation technology adoption,” *Energy Policy*, vol. 39, no. 10, pp. 6021–6029, 2011, doi: 10.1016/j.enpol.2011.06.066.
[61] H. Welsch and J. Kühling, “Determinants of pro-environmental consumption: The role of reference groups and routine behavior,” *Ecological Economics*, vol. 69, no. 1, pp. 166–176, 2009, doi: 10.1016/j.ecolecon.2009.08.009.
[62] C. Davidson, E. Drury, A. Lopez, R. Elmore, and R. Margolis, “Modeling photovoltaic diffusion: An analysis of geospatial datasets,” *Environmental Research Letters*, vol. 9, no. 7, 2014, doi: 10.1088/1748-9326/9/7/074009.
[63] M. Graziano, M. Fiaschetti, and C. Atkinson-Palombo, “Peer effects in the adoption of solar energy technologies in the United States: An urban case study,” *Energy Res Soc Sci*, vol. 48, no. September 2018, pp. 75–84, 2019, doi: 10.1016/j.erss.2018.09.002.
[64] B. Lukanov and E. Krieger, “Distributed solar and environmental justice: Exploring the demographic and socio-economic trends of residential PV adoption in California,” *Energy Policy*, vol. 134, no. August, p. 110935, 2019, doi: 10.1016/j.enpol.2019.110935.
[65] A. R. Hansen, M. H. Jacobsen, and K. Gram-Hanssen, “Characterizing the Danish energy prosumer: Who buys solar PV systems and why do they buy them?,” *Ecological Economics*, vol. 193, no. December 2021, p. 107333, 2022, doi: 10.1016/j.ecolecon.2021.107333.
[66] L. Mundaca and M. Samahita, “What drives home solar PV uptake? Subsidies, peer effects and visibility in Sweden,” *Energy Res Soc Sci*, vol. 60, no. October 2019, p. 101319, 2020, doi: 10.1016/j.erss.2019.101319.
[67] J. R. Snape, “Spatial and temporal characteristics of PV adoption in the UK and their implications for the smart grid,” *Energies (Basel)*, vol. 9, no. 3, pp. 1–18, 2016, doi: 10.3390/en9030210.
[68] C. A. Bollino, “The willingness to pay for renewable energy sources: The case of italy with socio-demographic determinants,” *Energy Journal*, vol. 30, no. 2, pp. 81–96, 2009, doi: 10.5547/issn0195-6574-ej-vol30-no2-4.
[69] B. Mills and J. Schleich, “Profits or preferences? Assessing the adoption of residential solar thermal technologies,” *Energy Policy*, vol. 37, no. 10, pp. 4145–4154, 2009, doi: 10.1016/j.enpol.2009.05.014.
[70] J. S. Woersdorfer and W. Kaus, “Will nonowners follow pioneer consumers in the adoption of solar thermal systems? Empirical evidence for northwestern Germany,” *Ecological Economics*, vol. 70, no. 12, pp. 2282–2291, 2011, doi: 10.1016/j.ecolecon.2011.04.005.
[71] L. Zhuang, Z. Zhang, and L. Wang, “The automatic segmentation of residential solar panels based on satellite images: A cross learning driven U-Net method,” *Applied Soft Computing Journal*, vol. 92, p. 106283, 2020, doi: 10.1016/j.asoc.2020.106283.
[72] P. Parhar *et al.*, “HyperionSolarNet: Solar Panel Detection from Aerial Images,” *ArXiv*, pp. 1–16, 2022, [Online]. Available: http://arxiv.org/abs/2201.02107
[73] G. Kasmi, L. Dubus, P. Blanc, and Y.-M. Saint-Drenan, “Towards unsupervised assessment with open-source data of the accuracy of deep learning-based distributed PV mapping,” *ArXiv*, pp. 1–12, 2022, doi: 10.48550/arXiv.2207.07466.
[74] S. Killinger *et al.*, “On the search for representative characteristics of PV systems: Data collection and analysis of PV system azimuth, tilt, capacity, yield and shading,” *Solar Energy*, vol. 173, no. August, pp. 1087–1106, 2018, doi: 10.1016/j.solener.2018.08.051.

[75] H. Lan, Z. Gou, and Y. Lu, “Machine learning approach to understand regional disparity of residential solar adoption in Australia,” *Renewable and Sustainable Energy Reviews*, vol. 136, no. September 2019, 2021, doi: 10.1016/j.rser.2020.110458.

[76] J. Rode and A. Weber, “Does localized imitation drive technology adoption? A case study on rooftop photovoltaic systems in Germany,” *J Environ Econ Manage*, vol. 78, pp. 38–48, 2016, doi: 10.1016/j.jeem.2016.02.001.

[77] R. Snape and C. Rynikiewicz, “Peer effect and social learning in micro-generation adoption and urban smarter grids development?,” *Network Industries Quarterly*, vol. 14, no. 2 & 3, pp. 24–27, 2012, [Online]. Available: http://newsletter.epfl.ch/mir/index.php?module=epflfiles&func=getFile&fid=292&inline=1

[78] A. Kaplan, “From passive to active about solar electricity: Innovation decision process and photovoltaic interest generation,” *Technovation*, vol. 19, no. 8, pp. 467–481, 1999, doi: 10.1016/S0166-4972(98)00128-X.

[79] Skatteverket, “Typkoder för fastigheter,” 2023. https://www.skatteverket.se/foretag/skatterochavdrag/fastighet/fastighetstaxering/typkoder.4.3f4496fd14864cc5ac9126d.html#accordion-0-item-0 (accessed Apr. 17, 2023).

Bilaga A

Ortofoton

För en mer omfattande beskrivning, se Lantmäteriets produktbeskrivning Ortofoto[3] (Dokumentversion 3.9).

Ortofoton är *radiometriskt bearbetade flygfotodata*, geometriskt projicerade till en ortogonal kartprojektion med hjälp av en höjdmodell. I ett ortofoto påverkas inte skalan eller avståndet mellan punkterna av variationer i terrängen, vilket är fallet på ett flygfoto med central projektion.

Produkten Ortofoto innehåller ortogonala projicerade flygfoton (ortofoton) med olika kombinationer av *våglängdsband* och upplösningar. Sedan 2019 finns ortofoton med en upplösning på 0,1 m/pixel, och 4-kanals (röd, grön, blå, infraröd), för Sverige och det är dessa ortofoton som har använts i detta projekt.

Ortofoton är tagna på en flyghöjd av 3000 meter, och varje Ortofoto täcker 2,5×2,5 km (15 625×15 625 pixlar) med ungefärliga horisontella standardfel på 0,2 m.

Flygfototillfället har stor inverkan på Ortofotots kvalitéer, där tex bilder tagna före lövsprickning innehåller mycket lite IR-information. Därför kan bilderna vara mer eller mindre lämpliga för olika användningsområden, beroende på tidpunkten för fotograferingen. Andra faktorer kan också ge en viss variation i bilderna. Exempel på sådana faktorer är dimma, solvinkel och förhållanden på marken, tex torka, vid fotograferingstillfället. Speciellt solvinkeln är av betydelse för detta projekt då vissa solcellssystem endast är vita/glänsande i ortofoton när solvinkeln och modulernas lutning och orientering är i sådan position att direkt solljus reflekteras in i kameran.

Beroende på var på flygfotot en byggnad ligger, det vill säga hur långt från en byggnad flyglinjen är, kan mer eller mindre av fasaden synas på ortofotot. När en byggnad ligger mitt i ett flygfoto så tas bilden direkt ovanför byggnaden och då ser man taket på byggnaden och ingenting av fasaden. När en byggnad ligger långt bort från flyglinjen, det vill säga nära kanten på ett flygfoto, ser man byggnaden snett ovanifrån, och således även relativt mycket av fasaden. Av denna anledning kan skillnader uppstå mellan olika flygår, när samma område har flugits över med olika flyglinjer och sidotäckning. På ett Ortofoto från ett specifikt flygår kanske inte fasaden på en viss byggnad syns över huvud taget medan ett Ortofoto från ett annat flygår visar relativt mycket av fasaden på samma byggnad.

Lantmäteriets ambition är att fotografera cirka 30% av landet varje år; oftare och med högre upplösning i tätt bebyggda områden i södra Sverige och längs Norrlandskusten, men mer sällan och med lägre upplösning i Norrlands inland och i norra fjällkedjorna.

[3] https://www.lantmateriet.se/globalassets/geodata/geodataprodukter/flyg--och-satellitbilder/pb_ortofoto.pdf

Bilaga B

Skatteverkets typkoder

Vid fastighetstaxeringen fastställer Skatteverket en specifik typkod för varje fastighet. Genom att använda informationen som tillhandahålls i fastighetsdeklarationen beräknar Skatteverket taxeringsvärdet, vilket utgör grunden för fastighetsskatten eller fastighetsavgiften som fastighetsägaren är skyldig att betala [79]. Typkoderna har i denna studie möjliggjort en mer detaljerad analys av solenergimarknadssegmenten än vad som tidigare varit möjligt, eftersom de med hög detaljnivå kan ange vilken typ av verksamhet en byggnad är avsedd för.

Se nedan de olika kategorierna för typkoder.

Tabell B.1: Skatteverkets typkoder, de olika kategorierna i vilka fastigheter indelas i Sverige och en beskrivning av kategorierna [79].

Beskrivning	Typkoder
Lantbruksenhet	110 Lantbruksenhet, obebyggd 113 Lantbruksenhet, bostadsbyggnadsvärde under 50 000 kronor 120 Lantbruksenhet, bebyggd 121 Lantbruksenhet, enbart ekonomibyggnad 122 Lantbruksenhet, bebyggd enbart med ekonomibyggnad 180 Lantbruksenhet, skatte-/avgiftsfri (3 kap. 2 § 2 st. FTL) 181 Lantbruksenhet, skatte-/avgiftsfri (3 kap. 3 § FTL) 197 Lantbruksenhet, övrig mark utanför tätort, taxeringsvärde 0 kronor 199 Lantbruksenhet, taxeringsvärde under 1 000 kronor
Småhusenheter	210 Småhusenhet, tomtmark 213 Småhusenhet, byggnadsvärde under 50 000 kr 220 Småhusenhet, bebyggd 223 Småhusenhet, bostadsbyggnad med lokaler 225 Småhusenhet, småhus på ofri grund 230 Småhusenhet, grupphusområde enligt 12 kap. 3 § FTL 240 Småhusenhet, bostadsbyggnad på vattenfastighet 280 Småhusenhet, skatte-/avgiftsfri (3 kap. 2 § 2 st. FTL) 281 Småhusenhet, skatte-/avgiftsfri (3 kap. 4 § FTL) 299 Småhusenhet, taxeringsvärde under 1 000 kr
Hyreshusenheter	310 Hyreshusenhet, tomtmark 311 Hyreshusenhet, med saneringsbyggnad 313 Hyreshusenhet, byggnadsvärde under 50 000 kr 320 Hyreshusenhet, bostäder 321 Hyreshusenhet, bostäder och lokaler 322 Hyreshusenhet, hotell eller restaurangbyggnad 323 Hyreshusenhet, kiosk 324 Hyreshusenhet, parkeringshus/garage 325 Hyreshusenhet, lokaler 326 Hyreshusenhet, kontor inom industrimark 340 Hyreshusenhet, byggnad på vattenfastighet 380 Hyreshusenhet, skatte-/avgiftsfri (3 kap. 2 § 2 st. FTL) 381 Hyreshusenhet, skatte-/avgiftsfri (3 kap. 4 § FTL) 399 Hyreshusenhet, taxeringsvärde under 1 000 kr
Industrienheter	411 Industrienhet, tomtmark 412 Industrienhet, med saneringsbyggnad 413 Industrienhet, upplag eller uppställningsplats 414 Industrienhet, byggnadsvärde under 50 000 kr 415 Industrienhet, fiskefastighet 420 Industrienhet, industrihotell 421 Industrienhet, kemisk industri

	422 Industrienhet, livsmedelsindustri 423 Industrienhet, metall- och maskinindustri 424 Industrienhet, textil- och beklädnadsindustri 425 Industrienhet, trävaruindustri 426 Industrienhet, annan tillverkningsindustri 430 Industrienhet, bensinstation 431 Industrienhet, reparationsverkstad 432 Industrienhet, lager 433 Industrienhet, annan övrig byggnad 443 Industrienhet, övrig byggnad f. radiokommunikation 480 Industrienhet, skatte-/avgiftsfri (3 kap. 2 § 2 st. FTL) 481 Industrienhet, skatte-/avgiftsfri (3 kap. 4 § FTL) 497 Industrienhet, övrig mark inom tätort, taxeringsvärde 0 kr 498 Industrienhet, övrig mark (gatu- och parkmark), taxeringsvärde 0 kr 499 Industrienhet, taxeringsvärde under 1 000 kr
Ägarlägenhetsenheter	510 Ägarlägenhetsenhet, obebyggd 513 Ägarlägenhetsenhet, byggnadsvärde under 50 000 kr 520 Ägarlägenhetsenhet, bostad 580 Ägarlägenhetsenhet, skatte-/avgiftsfri (3 kap. 2 § 2 st. FTL) 581 Ägarlägenhetsenhet, skatte-/avgiftsfri (3 kap. 4 § FTL)
Täkter	610 Industrienhet, obebyggd täktmark 613 Industrienhet, täkt med byggnadsvärde < 50 000 kr 620 Industrienhet, bebyggd täktmark 699 Industrienhet, täktmark med taxeringsvärde under 1 000 kr
Elproduktionsenheter	711 Elproduktionsenhet, outbyggt strömfall 712 Elproduktionsenhet, tomt till kraftvärmeverk 713 Elproduktionsenhet, (vatten) byggn.värde < 50 000 714 Elproduktionsenhet, ersättningskraft 715 Elproduktionsenhet, tomt till kärnkraftverk 716 Elproduktionsenhet, tomt till kondenskraftverk 717 Elproduktionsenhet, tomt till gasturbinanläggning 718 Elproduktionsenhet, tomt till vindkraftverk 719 Elproduktionsenhet, (värme) byggn.värde < 50 000 720 Elproduktionsenhet, vattenkraftverk 730 Elproduktionsenhet, kärnkraftverk 731 Elproduktionsenhet, kondenskraftverk 732 Elproduktionsenhet, kraftvärmeverk 733 Elproduktionsenhet, gasturbinanläggning 734 Elproduktionsenhet, vindkraftverk 739 Elproduktionsenhet, vindkraftverk byggn.värde <50 000 740 Elproduktionsenhet, vindkraftverk i vattenområde 790 Elproduktionsenhet, vindkraftverk som inte tagits i drift 799 Elproduktionsenhet, taxeringsvärde under 1 000 kr
Specialenheter	810 Specialenhet, tomtmark till specialbyggnad 820 Specialenhet, distributionsbyggnad 821 Specialenhet, reningsanläggning 822 Specialenhet, värmecentral 823 Specialenhet, vårdbyggnad 824 Specialenhet, bad-, sport- och idrottsanläggning 825 Specialenhet, skolbyggnad 826 Specialenhet, kulturbyggnad 827 Specialenhet, ecklesiastikbyggnad 828 Specialenhet, allmän byggnad 829 Specialenhet, kommunikationsbyggnad 890 Specialenhet, försvarsbyggnad

Bilaga C

R-script av U-netmodellen

I den här bilagan förklaras hur man rent praktiskt kan använda den version av U-netmodellen som tagits fram i det här projektet, från nedladdning till tillämpning.

Det enklaste tillvägagångssättet är att använda den så kallade *Docker-containern* som innehåller samtliga nödvändiga tillägg för att köra GPU-versionen av U-netmodellen i R på datorer som är kompatibla GPU-enhet (den körs på CPU om ingen GPU hittas). För att göra detta utför man följande kommando:

docker pull frimane/azfunetr:latest

Det bör observeras att Docker-containern är oberoende av U-net-modellskriptet så att den ska kunna köras med vad som finns tillgängligt. För att köra U-netmodellen ovanpå denna Docker-container måste filer och kataloger ges en hierarki. Den hierarkidesign som använts i den projektet presenteras i Figur C.1, men den kan givetvis tilldelas en annan hierarki om önskat.

U-net-skriptet skrivet i 5 stycken R-filer: *optimization.R*, *pvSegDsClass.R*, *training.R*, *UNetContainers.R* och *UnetTorch.R*. Dessa skriptfiler är samlade i en katalog som döpts till *Rscript*. Datasetet delas upp i test-, tränings- och valideringskataloger. Var och en av dem delas i sin tur upp i bild- och maskkataloger. Hela datasetet bör hållas i katalogen som heter *pvData*. De slutliga resultaten kommer lagras i en katalog som kallas *output* när U-netkörningen är klar.

Slutligen, för att starta U-netmodellen kan följande skript användas:

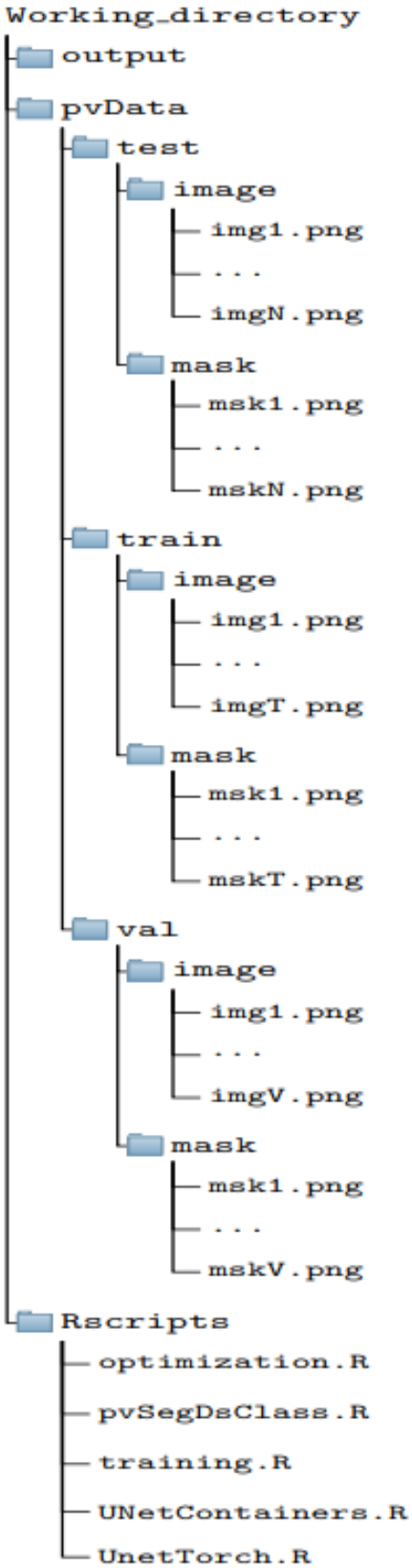
```
#!/ bin/ bash echo
" Starting Docker service "
systemctl restart docker

echo " Starting U - net model "
```

För körningen kan följande skript användas:

```
docker run - it -- gpus all -- rm -v Working_directory / Rscripts :/ Rscripts -v
Working_directory / pvData :/ pvData -v Working_directory / output :/ output
path_to / azfunetr : latest ‘


# Working_directory : the folder where all the work is performed
# path_to : the path to the directory where the container is saved
```



Figur C.1. Filträd som visar den designade hierarkin som behövs för att använda U-netmodellen över *Docker-containern*.