



AI-baserad segmentering av fasader för att optimera renovering i en större skala

Downloaded from: <https://research.chalmers.se>, 2026-04-06 19:21 UTC

Citation for the original published paper (version of record):

Somanath, S., Yu, Y., Nordmark, N. et al (2021). AI-baserad segmentering av fasader för att optimera renovering i en större skala. *Bygg och teknik*, 2021(2): 26-29

N.B. When citing this work, cite the original published paper.

AI-baserad segmentering av fasader för att optimera renovering i en större skala



Figur 1: Fönster på olika byggnader som har identifierats med maskininläring.

Hur kan vi på ett automatiskt sätt skapa mer detaljerade 3D modeller av byggnader i digitala tvillingar och förbättra indata för att beräkna energibesparingspotentialer i befintliga byggnader? I en pilotstudie har vi undersökt hur maskininläring kan användas för att extrahera information om fönstersättning och storlek i befintliga byggnader. Vi har utvecklat en modell som har "tränats" att känna igen och segmentera fönster från bilder med byggnadsfasader och på så sätt skapa digitala och mer detaljerade data för befintliga byggnader. Vårt långsiktiga mål är att utveckla en helautomatisk metod för analyser av renoveringspotentialer för byggnader och fastighetsportföljer.

Digitalisering för bättre energisimuleringar och optimerade renoveringslösningar

Det befintliga byggnadsbeståndet har stor inverkan på både människors välbefinnande och miljön. Cirka 12,2 mil-

joner ton eller 19 procent av Sveriges utsläpp av växthusgaser är relaterade till bygg- och anläggningssektorn [1], en stor del kopplad till befintliga byggnader. För att uppnå Sveriges klimatmål för

2045 krävs det bland annat storskaliga renoveringar av byggnadsbeståndet. Att planera renoveringar är rätt så tidskrävande, inte minst på grund av omfattande manuella inventeringar och beräkningar som behöver genomföras för att ta fram klimat- eller energibesparingspotentialerna i störrebyggnadsbestånd. Som ett resultat renoveras byggnader och byggnadsbestånd inte på ett optimalt sätt.

Det finns ett antal olika arbetsflöden för att identifiera optimala renoveringslösningar som är oftast kombinationer av åtgärder som tilläggsisolering, nya fönster, installation av solceller och byte av värmesystem [2]. Dessa lösningar presenteras till exempel i Pareto-kurvor för att visa avvägningen mellan kostnader och utsläpp av växthusgaser, figur 2 [3]. För en optimal simulering av byggnadernas energianvändning behövs det dock en 3D-modell med data om fasader (storlek, fönsteröppningar, etc.) och tak, se figur 3. Eftersom information om befintliga byggnader ofta bara finns i 2D-pappersritningar är framtagningen av en digital 3D-modell tidskrävande.

Det vanligaste sättet att skapa 3D modeller är att 3D skanna byggnader med LiDAR (Light detection and ranging) och skapa punktmoln, men sådana system är antingen dyra att använda i större skala eller ger inte den nödvändiga punktätheten. Ett annat tillvägagångssätt är baserat på digitala tvillingar av en stad. Flera städer, inklusive Göteborg och Helsingborg, har utvecklat en digital tvilling av städerna som inkluderar byggnadsvolymer och intresset för att skapa digitala tvillingar ökar i landets kommuner. Inom det nationella Vinnova kompetenscentret Digital Twin City Centre (DTCC) [4] som leds av Chalmers arbetar man med att ta fram arbetsflöden för att skapa digitala tvillingar av städer på ett generiskt sätt. Modellerna baseras



Sanjay Somanath
Chalmers Tekniska Högskola



Mola Ayenew
Chalmers Tekniska Högskola



Yinan Yu
Chalmers Tekniska Högskola



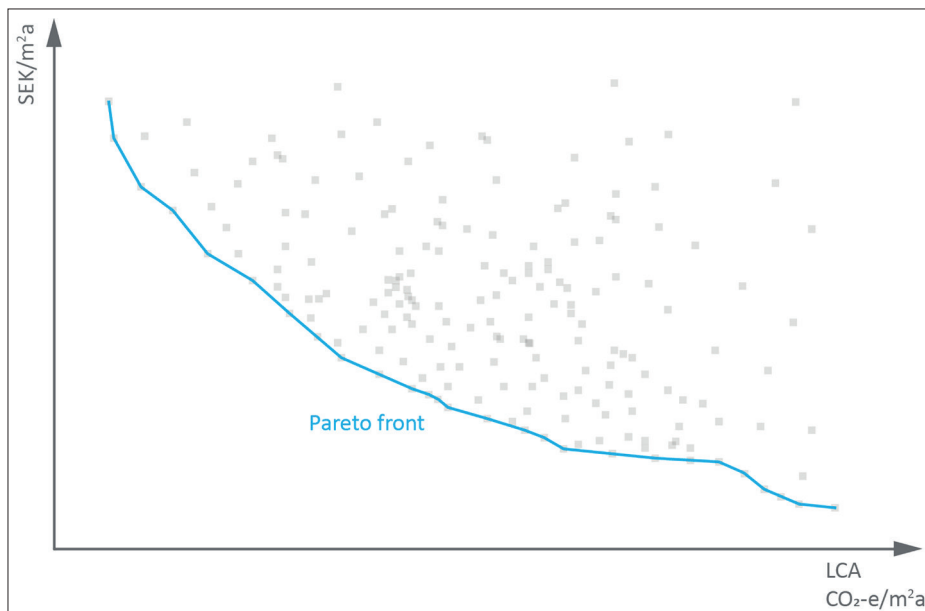
Liane Thuvander
Chalmers Tekniska Högskola



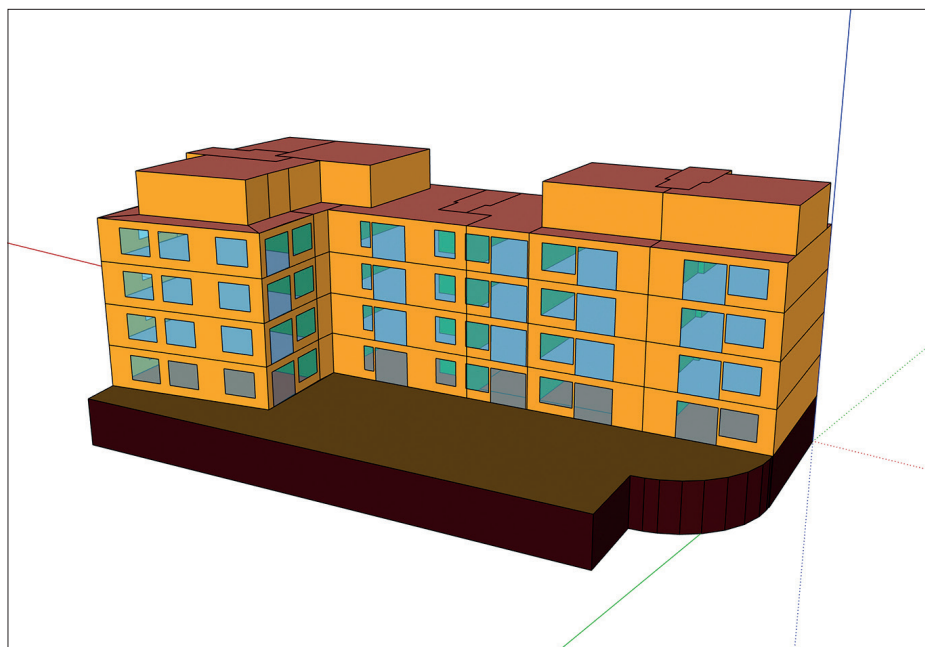
Nils Nordmark
Chalmers Tekniska Högskola



Alexander Hollberg
Chalmers Tekniska Högskola



Figur 2: Pareto-kurva som visar kostnader för renovering och relaterade utsläpp av växthusgaser i CO₂-equivalent.



Figur 3: Enkel 3D model för simulering av byggnadernas energibehov.

bland annat på byggnadsdata från öppna källor som OpenStreetMap eller från Lantmäteriet. Men dessa data innehåller ingen information om öppningar i fasaderna, fönster eller dörrar, och ofta saknas de exakta formerna på taken. Den typ av information är dock viktig för simuleringen av byggnadernas energibehov, speciellt förhållandet mellan fönster och vägg (WWR – Window Wall Ratio).

I vår studie har därför testat en metod för att snabbt och automatiskt generera data för fönstersättning i fasader i en större skala. Mer specifikt har vi undersökt potentialen att använda maskininläring för att extrahera data om fönster i byggnader baserat på bilder tagna på gatunivå. Tanken är att data

och metoden ska integreras i digitala tvillingar och ge underlag för förbättrade energisimuleringar.

Maskininläring, objekt-detektering och segmentering

Maskininläring (ML) är ett område inom *artificiell intelligens* och handlar om metoder för att med data "träna" datorer att upptäcka och "lära" sig regler för att lösa en uppgift, utan att datorerna har programmerats med regler för just den uppgiften. I studien använder vi ML för bildigenkänning genom att tillämpa två metoder inom datorseende: objekt-detektering och bildsegmentering.

Objekt-detektering är processen att upptäcka förekomster av objekt i en bild. Populära applikationer för objekt-

detektering är ansiktsgenkänningsfunktionen på mobiltelefoner eller i smile-to-capture-funktionen i moderna kameror. Vanliga algoritmer för objekt-detektering är så kallade Regioner med Convolutional Neural Networks (R-CNN) [5]. Algoritmen utför två uppgifter, lokalisering och klassificering, där den första ritat en avgränsningsruta runt föremålen i bilden och den andra identifierar klassen objektet tillhör.

Bildsegmentering är processen att dela upp en digital bild i flera segment av pixlar. Segmentering används vanligtvis för att identifiera distinkta funktioner i en bild såsom linjer, kanter och gränser. På senare tid har träningsbara segmenteringsmetoder tillämpats för att identifiera distinkta objekt i bilder genom att utnyttja utvecklingen i neurala nätverk och ökad beräkningsförmåga. Träningsbara segmenteringsmetoder "lärt sig" sambanden mellan pixlarna för den önskade funktionen utifrån träningsdata som matas in i bildsegmenteringsmodellen.

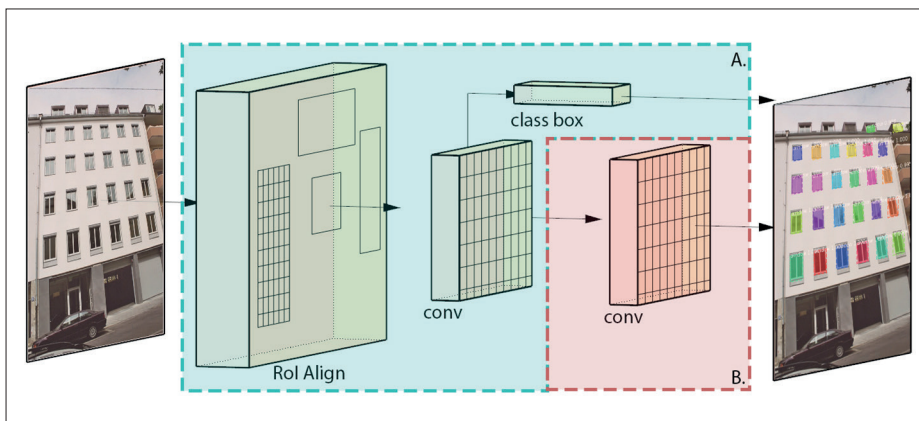
Objekt-detekteringsmodellen Mask R-CNN och algoritmer

Vi har använt en objekt-detekteringsmodell som heter Mask R-CNN, den senaste modellen inom R-CNN-familjen för objekt-detekteringsmodeller. Den tränade modellen kan inte bara identifiera objektet den tränades på utan också rita en ram runt varje förekommande objekt. Modellen identifierar först en region (RoI – Region of Interest) kring möjliga objekt i bilden. Sedan klassificerar modellen de föreslagna regionerna, i vårt fall fönster, skapar en avgränsningsruta runt regionerna och drar en ram runt det identifierade objektet, se figur 4.

Träning av algoritmer

ML-algoritmer bygger på den information som ges till dem. Kvaliteten på träningsdata är alltså avgörande för en korrekt identifiering och klassificering av objekt, i vårt fall fönster. För att träna algoritmerna använde vi två steg: förberedelse av träningsdata och överföringsinläring.

Då vår modell tillämpas för Göteborg måste den utbildas med bilder som är representativa för stadens arkitektoniska stil. De flesta allmänt tillgängliga data (till exempel [4], [5]) som har taggade fasadbilder skiljer sig från den arkitektoniska stilen i Göteborg. Oftast är dessa bilder också korrigerade och beskurna och har en vridning i kameravinkeln. Eftersom vi behöver använda bilder med gatuvy är ovan nämnda datamängder olämpliga för vår applikation. Istället har vi extraherat



Figur 4: Från originalbild med en byggnadsfasad till en fasadbild med taggade fönster. I del A genomförs objekt-detekteringen och en avgränsningsruta ritas. Först identifieras Region of Interests (RoI Align) innan så kallade faltningsnätverk (conv = convolutional neural network) används för att tilldela objektet (här fönster) en avgränsningsruta. Därefter klassas det identifierade objektet. I del B, en process som pågår parallellt, genereras med hjälp av ett annat faltningsnätverk en ram för taggning av fönstren.



Figur 5: Resultat från fasadsegmenteringen med hjälp av Mask R-CNN modellen. Den vänstra kolumnen visar originalbilderna, den andra vänstra kolumnen visar de manuellt taggade bilderna (ground truth), den tredje kolumnen visar algoritmnernas prediktion och den fjärde kolumnen visar överlappningen mellan "ground truth" och prediktion. Matchande fönster är markerade i orange.

bilder med gatuvy från Google Street View och beskurit dem så att bilderna representerar ett antal olika byggnader som är typiska för Göteborg. Bilderna taggades sedan manuellt med hjälp av

bild- och videomärkningsverktyget VGG Image Annotator [6]. Data setet utökades gradvist till en samling av 100 bilder med 1540 fönster. Samlingen delades sedan upp i en träningsuppsättning (80 bilder)

och en valideringsuppsättning (20 bilder).

I ett nästa steg kommer vi använda oss av överföringsinlärning vilket innebär att vi inte behöver träna en ny modell från början utan använder redan inlärd funktionerna från ett befintligt dataset, i vårt fall COCO [7] som innehåller cirka 120 000 bilder.

Resultat och utveckling framöver

Vi har totalt sett tränat och utvärderat mer än 60 konfigurationer för algoritmen i R-CNN. Den bästa uppnådde pixel-noggrannheter är på 94,8 procent. Exempel på resultaten visas i figur 4. Den vänstra kolumnen visar originalbilderna, den andra kolumnen visar den manuellt taggade bilden, den tredje kolumnen visar algoritmnernas prediktion och den fjärde kolumnen visar överlappningen mellan kolumn två och tre. Det vi kan se är att prediktionen är bra för stora fasadområden men i de fall där vegetation blockerar fönstren (exempelvis byggnaden på andra raden) eller dörrar ingår (exempelvis byggnaden på femte raden) uppstod fel. Felen kan hanteras genom att använda flera bilder från samma byggnad från olika vinklar och genom att lägga till fler klasser till modellen, som till exempel dörrar.

Vår studie visar att fönsteröppningar och deras placering kan extraheras från bilder med byggnadsfasader tagna på gatunivå. Vi har fokuserat på att träna och använda modellen för fönster men i ett nästa steg kommer vi att utöka modellen och klassningen till att inkludera andra fasaddelar såsom dörrar och balkonger. Längre fram kommer även olika fönstertyper integreras (spröjsar, enlufts-, tvålufts-fönster, etc.) för att förfinas indata med högre detaljnivå och därmed ge högre noggrannhet i energisimuleringar. Att lägga till ytterligare klasser i modellen (dörrar och balkonger) bidrar också till mer detaljerade 3D stadsmodeller och digitala tvillingar vilket möjliggör bättre visualiseringar som i sin tur kan stödja bättre kommunikation mellan olika intressenter.

En annan viktig egenskap för att bedöma renoveringsbehov och energiparpotentialer hos en befintlig byggnad är dessa uppgifter är ofta svåra att samla in på ett automatiserat sätt. En potentiell tillämpning av vårt tillvägagångssätt är att identifiera fasadmateriell, till exempel tegel eller trä, samt byggnadens skick, till exempel tillväxt av alger eller behovet av att måla om fasaderna, om modellen kompletteras med ytterligare bilddata. Att kunna detektera typ av

fasadmateriel kan också bidra till att upprätta en digital inventering av stadens materiallager, det vill säga inbyggt material i våra byggnadsbestånd, som kan underlätta återanvändning av material då informationen kring materialets placering och art blir tillgängligt på ett strukturerat sätt. Detektering av byggnadsmaterial kommer att undersökas som en del av ett nyligen startat projekt inom DTCC kallat "Circularity".

Sammanlagt, så kommer resultaten från studien att användas för att både skapa mer detaljerade 3D modeller av städer i en digital tvilling och bidra till snabba energisimuleringar med hög detaljnivå i större skala. Uppgifterna om fönster ger information om fönster-vägg förhållanden i fasaden, vilket gör att vi kan öka noggrannheten i våra analysmodeller. På så sätt ges också möjligheten att fatta välgrundade beslut om storskaliga renoveringsstrategier i byggnadsbestånd. ■

Läs mer:

Digital Twin Cities Centre (DTCC): <https://dtcc.chalmers.se/>

DTCC som har sin bas på Chalmers är ett Vinnova kompetenscentrum be-

stående av ett brett konsortium med 31 svenska och internationella intressenter. Avancerad forskning bedrivs inom åtta olika områden som täcker olika skalor och tillämpningar. Målet är att använda tekniken för nya tillvägagångssätt inom stadsplanering & design, arkitektur och digital konstruktion.

Referenser

- [1] Boverket, 'Emissions of greenhouse gases from the construction and real estate sector', National Board of Housing, Building and Planning, 2020. [Online]. Available: <https://www.boverket.se/sv/byggnad/hallbart-byggande-och-forvaltning/miljoindikatorer---aktuell-status/vaxthushusgaser/>. [Accessed: 08-Aug-2020].
- [2] A. Hollberg and J. Ruth, 'LCA in architectural design—a parametric approach', *Int. J. Life Cycle Assess.*, vol. 21, no. 7, pp. 943–960, Jul. 2016.
- [3] N. Klüber, A. Hollberg, and J. Ruth, 'Life cycle optimized application of renewable raw materials for retrofitting measures', in *World Sustainable Building*, 2014.
- [4] F. Korc and W. Förstner, 'TRIMS Image Database for Interpreting Images of Man-Made Scenes', 2009.

[5] R. Tylecek and R. Sára, 'Spatial Pattern Templates for Recognition of Objects with Regular Structure', in *Pattern Recognition*, 2013, pp. 364–374.

[6] A. Dutta, A. Gupta, and A. Zisserman, 'VGG Image Annotator (VIA)', 2016. [Online]. Available: <https://www.robots.ox.ac.uk/~vgg/software/via/>. [Accessed: 04-Feb-2021].

[7] T. Y. Lin et al., 'Microsoft COCO: Common objects in context', *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8693 LNCS, no. PART 5, pp. 740–755, 2014.