

FÖRBEREDELSEPROJEKT

Utvärdering av platstrygghet i gaturummet utifrån gatuplansbilder och maskininlärning

Nicklas Korsell



Syftet med projektet

- Att testa en automatiserad metod att utifrån gatuplansbilder och AI/Machine Learning utvärdera gators och offentliga platser potential till platstrygghet
- Bakomliggande syftet: att synliggöra för kommunsektorn potentialen i gatuplansbilder som ett dataunderlag för mätning och monitorering
- Platstrygghet (i denna rapport): en "platskvalité" $\approx$ platsens förmåga att utifrån dess utformning och skick få människor att känna sig trygga på platsen

Gatuplansbilder från ett IoT?

- Många streetviewbilder kommer från det kanske största IoT: alla uppkopplade smartphonse
- Några få kommuner har testat att rigga vanliga mobilkameror på sopbilar/servicefordon som tillsammans bildar ett nätverk av rörliga "sensorer"

OBS!

- Projektet handlar **inte** om att identifiera otrygga platser pga. faktiska händelser, brottsutsatthet och/eller rörelsemönster av människor, utan bara utifrån platsens "rumsliga utformning och skick".
- Hur människor **faktiskt** upplever tryggheten på en plats färgas annars starkt av förkunskapen kring den: rykten, nyheter, tidigare erfarenheter.
- Detta är ansatsens styrka: den bygger på ett träningsmaterial av gatuplansbilder från platser som individerna som rankat bilderna inte känner till.

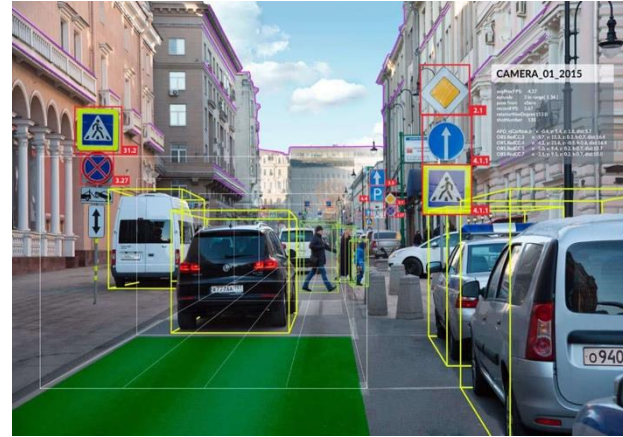
Grundläggande idén:

- Om en stadsutvecklare kan gå runt i en stad och utvärdera olika platsers kvalitéer visuellt: skulle inte en tränad maskin-/djupinlärningsmodell kunna "gå runt" bland gatuplansbilderna över staden och utvärdera/mäta?
- Gatuplansbilder ("Street View Imagery") = Foton tagna i markplan: ofta sammansatta till 360-graders panoramabilder
- Maskin-/Djupinlärningsmodeller = Storskaliga statistiska modeller med väldigt många parametrar som utifrån (en stor mängd) data kan "organisera sig själv" till en bra prediktionsmodell

Computer vision, Djupinlärning och gatuplansbilder ("Street View Imagery")

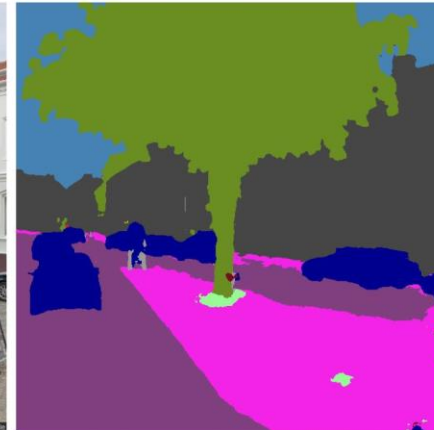
Djupinlärningsmodeller kan lära sig ("tränas") att:

- känna igen specifika objekt i bilder (**object detection**)



Target 13, Tomelilla. Latitude: 55.544 (North), Longitud: 13.952 (East)

- Klassificera delar av bilder till olika kategorier (**semantic segmentation**)



(Potentiella) nyttor + andra tillämpningar

- mätningar över stora områden med rent maskinella resurser – ej resurskrävande manuell inventering/inspektion
- gatuplanperspektivet ger ett bättre perspektiv än fågelperspektiv ur trafikantens perspektiv
- kartera platser/områdets [sett ur specifika grupperns behov](#): t.ex. personer med fysisk funktionsnedsättning.
- Inventering urbana grönområden: t.ex. spåra trädplatser, identifiera arter, bedöma trädens hälsa samt behov av skötsel.
- ? (Allt som behöver inventeras/karteras över större områden i urbana miljöer)...

Tillämpningar i Sverige

- Helsingborg använde liknande metod i ett prisbelönt projekt för att monitorera gators skick utifrån egengenererade gatuplansbilder
- Stockholm/Unvrse
- Sandviken fotograferade trottoarer/trottoarkanter från fordon



Genomförande & metoder

Två tänkbara ansatser:

1) "Checklista/protokoll"-approachen

- På varje plats: identifiera och checka av med specifika objekt/attribut som forskningen pekar på borde vara viktiga
- Aggregera till sammanvägd "trygghetspoäng" utifrån förekomsten av dessa olika platsattribut på något sätt

2) Träna en modell på en populations preferenser avs. platstrygghet

- Träna en maskin/-djupinlärningsmodell på hur människor rankat/poängsatt olika bilder av gator m.a.p. trygghet (y-variabeln)
- Använd t.ex. Place Pulse 2.0-datasetet (MIT)
- Modellen får träna att prediktera trygghetspoäng/ranking (y-variabeln) utifrån en stor mängd attribut extraherade ur bilderna (x-variable)

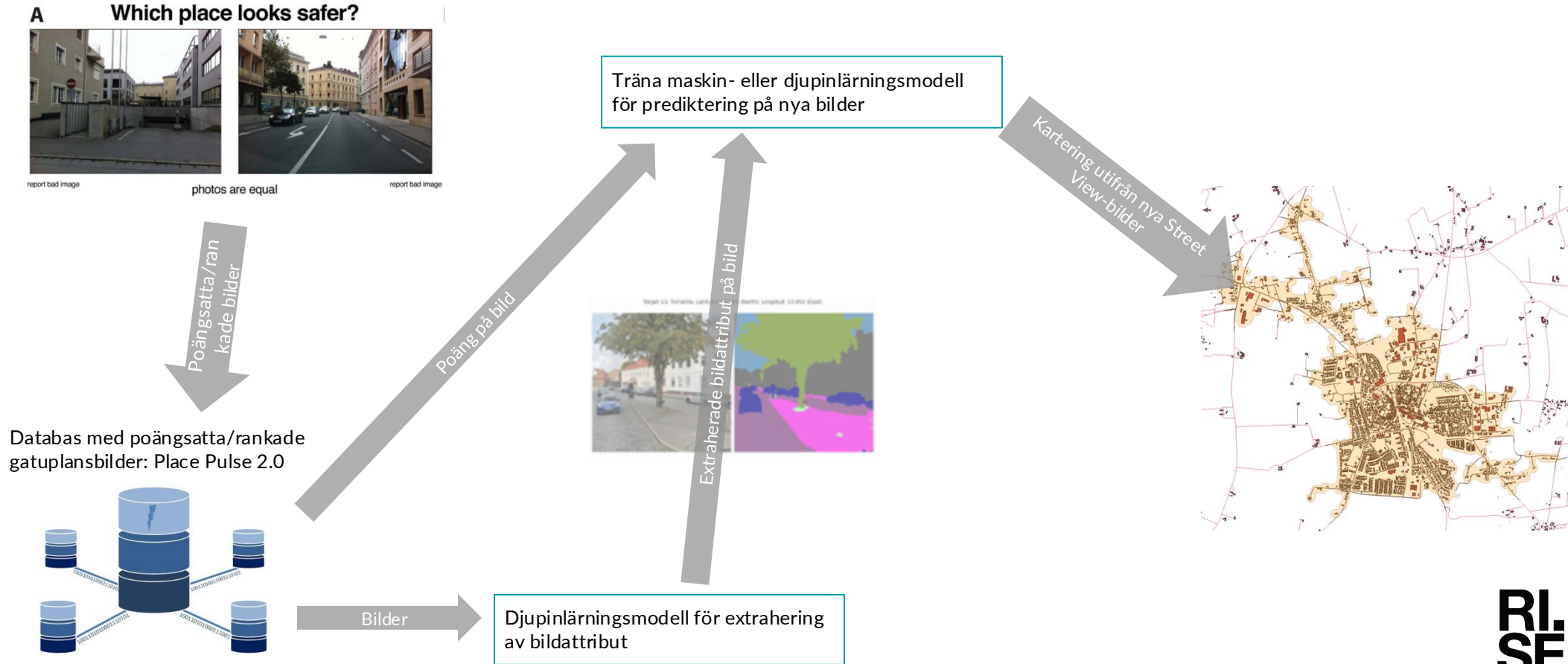
- Ansatsen innebär i princip att man tränar en ML/DL-modell på en populations "preferenser" för platstrygghet
- Vi använder Place Pulse 2.0 datasetet (Dubey et al., 2016) från MIT Lab, ett forskningsdataset med över 110 000 gatuplansbilder från 56 städer, med 1 170 000 parvisa jämförelser av gatuplansbilder gjorda av knappt 82 000 respondenter online.
- Jämförelserna gjordes med avseende på 6 olika platskvaliteter: "safe", "lively", "boring", "wealthy", "depressing" och "beautiful".
- De parvisa jämförelserna översätts till en ranking av de olika bilderna ("TrueSkill") (Herbrich et. al 2006)



Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. (2016). Deep Learning the City: Quantifying Urban Perception At A Global Scale (Vol. 9905). https://doi.org/10.1007/978-3-319-46448-0_12

Herbrich, R., Minka, T., & Graepel, T. (2006). TrueSkill™: A Bayesian Skill Rating System. In B. Schölkopf, J. Platt, & T. Hoffman (Eds.), *Advances in Neural Information Processing Systems* (Vol. 19). MIT Press. https://proceedings.neurips.cc/paper_files/paper/2006/file/f44ee263952e65b3610b8ba51229d1f9-Paper.pdf

Upplägg på genomförandet

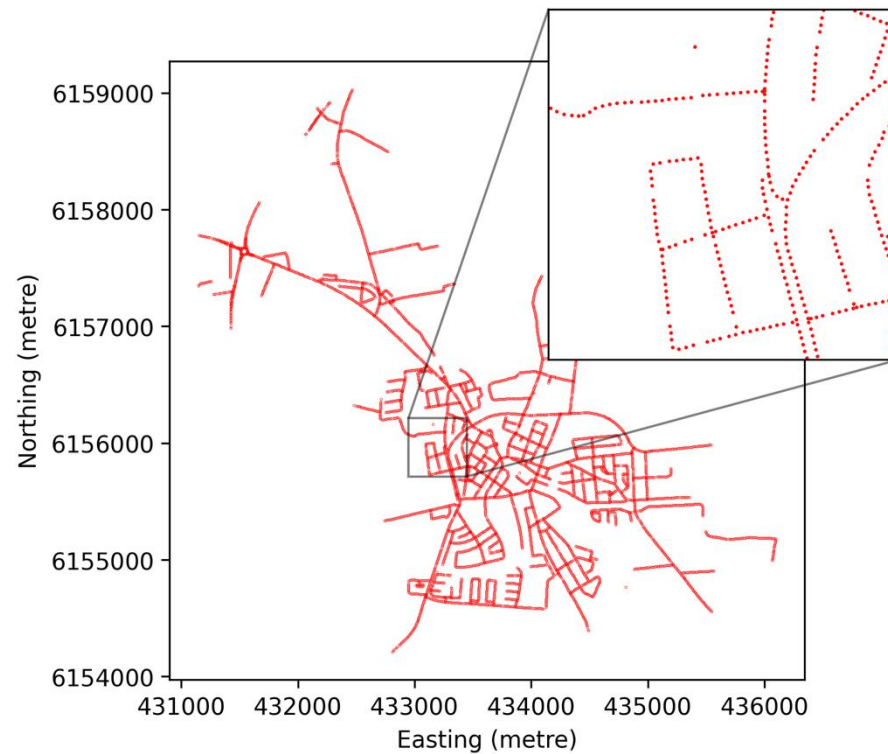


- Vi använder alltså två modeller:
 - En **förtränad djupinlärningsmodell** av typen Convolutional Neural Network för extrahera bildattribut ur gatuplansbilder med “Semantisk segmentering”
 - En **maskininlärningsmodell** att prediktera “trygghetspoäng” för en panoramaplats utifrån extraherade bildattribut
 - Vi prövade ett antal olika maskininlärningsmodeller. Den som predikterade bäst var av typen “Gradient Boosting Regression”

Mätområdet: Tomelilla "centraltätort"



Platser med gatuplansbilder i Google Street View Static API



Panoramabilder över Tomelilla från Google Streetview

- För varje plats där det fanns gatuplansbilder tillgängliga i Google Street View API hämtades panoramabilden ned som 8 bilder med 45 graders siktvinkel vardera.
- Sammanlagt laddades 6 197 panoramabilder ned från Google Street View Static API.

Exempel på 360-graders panoramabilder från två platser i Tomelilla. Bilden består av 8 bilder med 45 graders siktvinkel nedladdade från Google Streetview Tiles API och sedan ihopsatta till en panoramabild



Vilka attribut extraheras ur bilderna?

Objekt som extraherades ur bilderna med semantisk segmentering.
Objektklasserna är ett urval av klasserna i dataseten ADE20K och Cityscape

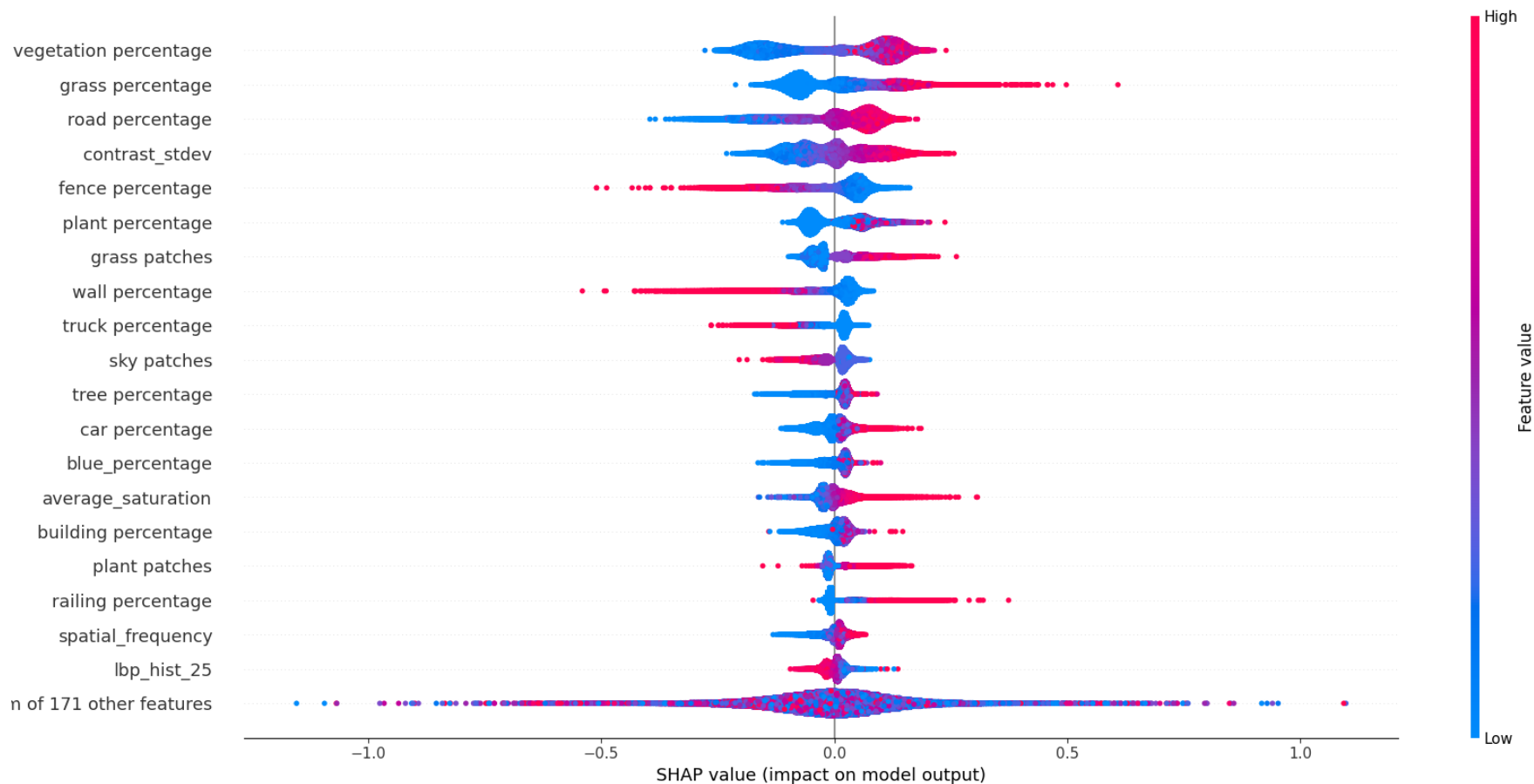
Kategori	Klass
Byggnadsdetaljer	Vägg, Golv, Dörr, Staket, Räcke, Trappa, Skärmdörr, Trapphall, Bänk, Markis, Stång, Fontän, Tak, Skräpkorg, Tält, Skulptur, Flagga
Byggnader	Byggnad, Hus, Torn, Pir
Gatuobjekt	Väg, Gata, Trottoar, Bro, Grusväg, Gatlykta, Trafikljus, Trafikmärke
Växtlighet, terräng etc.	Träd, Gräs, Växt, Fält, Flod/Å, Himmel, Kulle, Sjö, Vegetation, Terräng
Varelser	Person, Djur,
Fordon	Bil, Båt, Buss, Skåpbil, Skepp, Cykel, Lastbil, Ryttare, Tåg, Motorcykel

- 3 typer av attribut:
 1. Klasser i de två bilddatseten ADE20K och Cityscape (se tabell)
 2. Aggregerade grafiska egenskaper hos varje enskild bild
 3. ("low level") bildattribut av typen enkla mönster el geometriska former
- Sammanlagt över 1400 olika attribut/variabler

Grafiska egenskaper ("features") som extraherades ur gatubilder

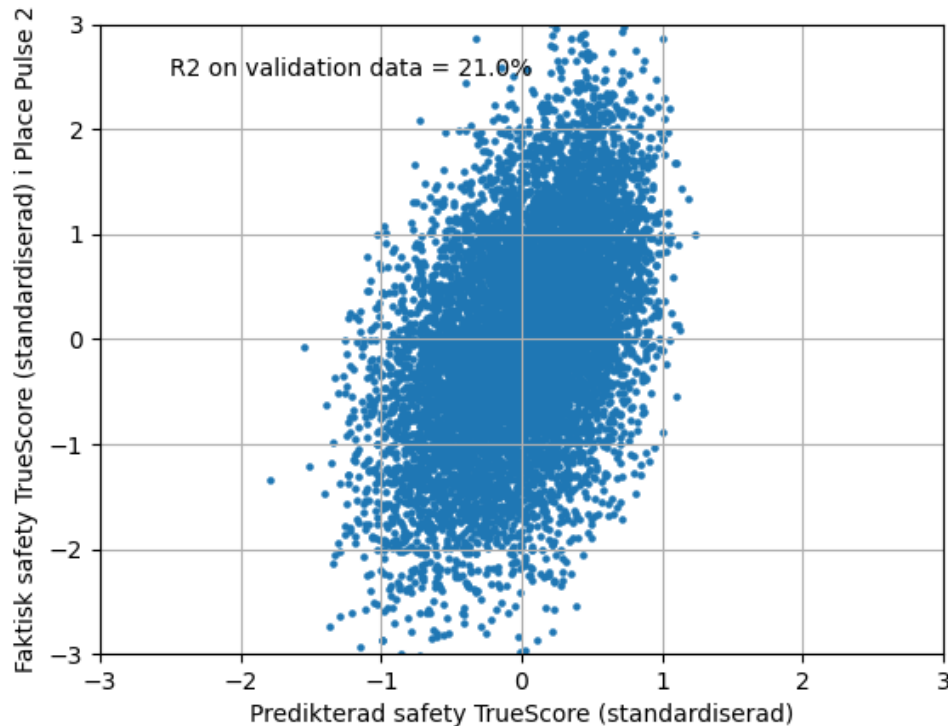
Namn	Beskrivning
Ljusstyrka, medelvärde och standardavvikelse	Genomsnittlig ljusstyrka över alla pixlar i bilden
Kontrast, medelvärde och standardavvikelse	Kontrast = skillnad i ljusstyrka mellan olika delar av bild
Nyans, medelvärde och standardavvikelse	
Mättnad, medelvärde och standardavvikelse	
Skärpa, medelvärde och standardavvikelse	
Entropi	Ett mått på ojämnheten i bildens gråskalefördelning
Kantdensitet (edge density),	Andel pixlar i en bild som ingår i en "kant"
Färgvarians	Varians av färgvärdena i RGB-kodningen

Vilka attribut/variabler är viktiga i prediktionen av potentiell platstrygghet?



Vad är prediktionsmodellens förklaringsförmåga?

- Determinationskoefficienten = "Andel av total variation i trygghetspoängen som bildattributen kan förklara"
- Maskininlärningsmodellen har en relativt låg förklaringsförmåga av TrueSkill poängen för trygghet (21-23% beroende på vilken modell som används)
- Svårsmätt storhet: "potential till upplevd trygghet"
- Alla de tränade maskininlärningsmodellerna tenderar att prediktera mot medelvärdet när de inte hittar starka samband

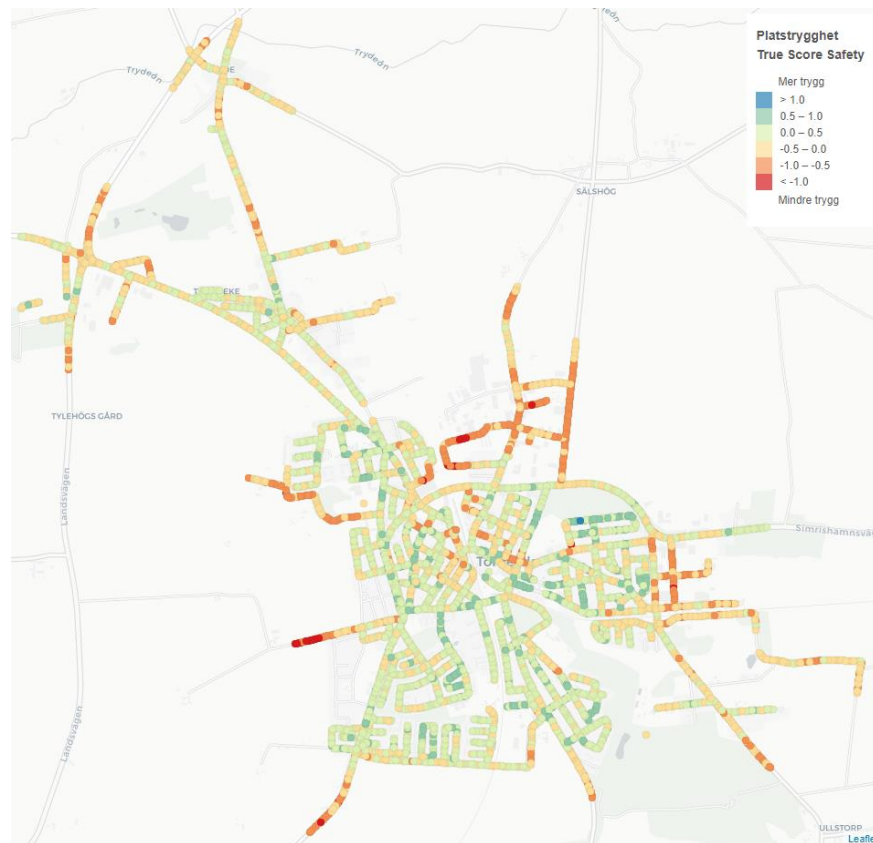




Resultat och analys

Karteringen av Tomelilla centraltätort

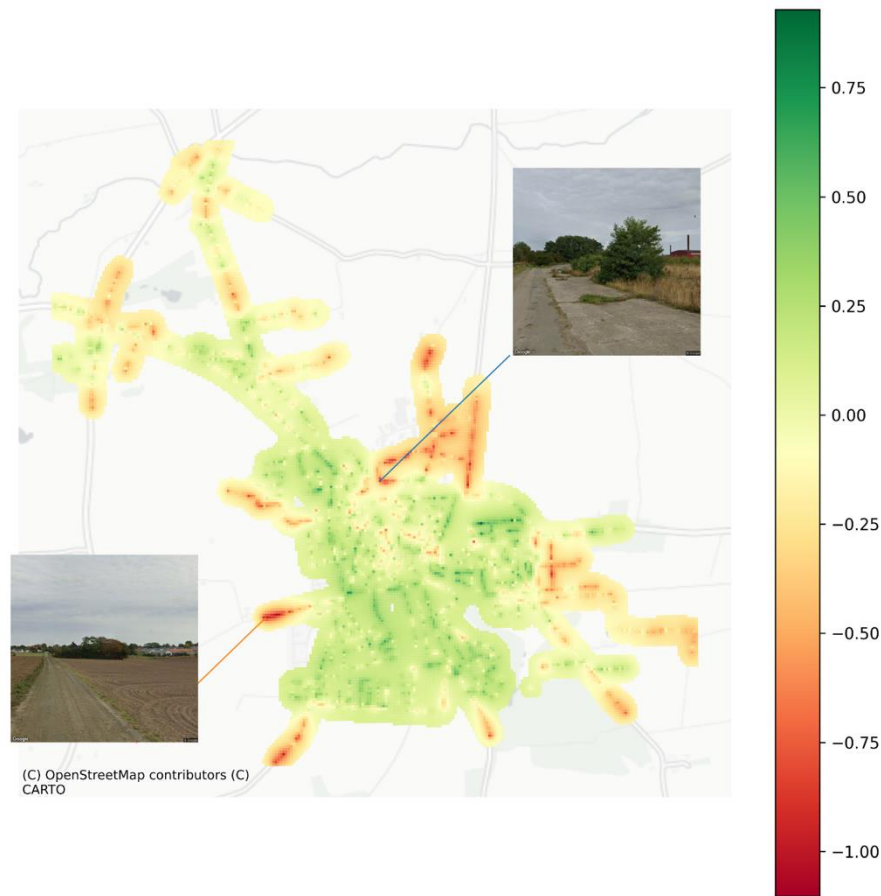
- Bild från den interaktiva visualiseringen av i Platser i Tomelilla där maskininlärningsmodellen predikterat platstrygghet utifrån panoramabild från Google Street View Static API.
- Observera att färgskalan medvetet är överdriven: Den standardiserade TrueScore-poängen löper mellan -3 till +3 ungefär. Mätskalan över Tomelillas platser ligger däremot mellan -1.0 och +1.0. De allra flesta platser uppfattas av modellen som relativt trygga.



<https://nppd.se/tomelilla-platstrygghet/>

Utgjämnade värden

- Utjämning av predikterade värden: en sorts "heat map"
- Notera generellt lägre predikterad trygghet i "ändområdena"
- Observera att färgskalan medvetet är överdriven: Den standardiserade TrueScore-poängen löper mellan -3 till +3 ungefär. De predikterade värdena över Tomelillas platser ligger däremot mellan -1.0 och +1.0. De allra flesta platser uppfattas av modellen som relativt trygga.



<https://nppd.se/tomelilla-platstrygghet/>

Exempel på platser i Tomelilla med relativt **hög** predikterad platstrygghet

- Välordnade, bebodda platser
- Tecken på skötsel av platsen: gräsmattor
- Många synliga bostadsbyggnader med fönster – god insyn på platsen
- Ofta gott om trottoarer/gångvägar, vägskyltar, gatulampor

Platstrygghet: 1.0735



Platstrygghet: 1.0403



Platstrygghet: 0.9919



Exempel på platser i Tomelilla med relativt **låg** predikterad platstrygghet

- Obebodda platser: industriområden. öppen terräng, fält, avskärmade gatudelar
- ingen insyn från bostadshus
- Ofta stängsel, murar eller tomma väggar

Platstrygghet: -1.0452



Platstrygghet: -1.2223



Platstrygghet: -1.2027



Ett exempel på en central gata med (relativt) låg predikterad platstrygghet

- Relativt avskärmd gatustump
- Långt mellan gatulampor
- Stängsel och höga staket
- Låg insyn från boende?





Slutsatser och lärdomar

Lärdomar

- Potential till upplevd platstrygghet är en svårämät storhet!
- Precisionen och förklaringsförmåga på enskilda platser är relativt låg: (iaf med den begränsade uppsättning platsattribut som används i våra modelleringar)
- För trygghet skulle det troligen behöva extraheras mer specifika platsattribut/egenskaper: t.ex. Fönster, urskilja boningshus, klotter/nedskräpning, skick på väggar och gator m.m. Kräver flera olika djupinlärningsmodeller.
- Men: modellen kan göra meningsfulla karteringar över stora områden väldigt snabbt

Förslag på utveckling

- För platstrygghet: Kombinera flera modeller för object detection/semantisk segmentation och fånga flera specifika attribut (t.ex. "fönster", klotter, upptäcka byggnader/vägar i dåligt skick skulle troligen förbättra precision
- (I en större empirisk studie): Validera ML-modellens prediktioner av platstrygghet mot hur verkliga personer bedömer och rankar gatubilderna från Tomelilla.
- En intressant utveckling av metoden vore att [även ta med attribut hos respondenterna](#) både i beräkningen av TrueSkill-värdet och sedan i prediktionsmodellen för densamma -> [olika kartor för olika grupper](#).
- En övergripande slutsats: maskinella/AI-baserade "karteringar" av platskvalitéer i urbana områden utifrån gatuplansbilder är en utvärderingsmetod med potential

Nästa steg

- Använda approachen för **lite lättare mätstorheter**: inventering/monitorering av
 - klotter/skadegörelse/nedskräpning
 - skick på soptunnor, återvinningsstationer, busskurer etc
 - skick på gator, trafikmarkeringar, skyltar etc
 - trafikflödesmätning, trafiksäkerhet (trafiksituation vid skolor olika klockslag t.ex.)

Frågor, reflektioner, diskussion, brainstorming

- ?

Tack!

Nicklas Korsell

Forskare

Avdelning Mätteknik

Enhet Mäta i samhället

D: +46 10 516 55 58 | M: +46 73 070 64 19

nicklas.korsell@ri.se

RISE Research Institutes of Sweden | ri.se