

Liikennemerkkien alueellinen jakautuminen Helsingissä – konenäköavusteinen tutkimus

Pyry Pohjanoksa, 2023

Sisällysluettelo

1 JOHDANTO.....	3
2 DATAN KERÄÄMINEN JA VALMISTELU	3
2.1 OMAN AINEISTON HANKKIMINEN	3
2.2 MUU KOULUTUSAINESTO	6
3 TUTKIMUSMENETELMÄT	6
3.1 AINEISTON LISÄVALMISTELUT	6
3.1.1 <i>Copy-paste-metodi</i>	7
3.1.2 <i>Mapillary-aineiston valmistelu</i>	8
3.2 KONEOPPIMISMALLIN KOULUTTAMINEN	8
3.3 MERKKIEN TUNNISTAMINEN	8
4 TULOKSET JA TULKINNAT	9
4.1 MALLIN TARKKUUSMITTARIT	9
4.2 VERTAILUAINESTO	10
4.3 LIIKENNEMERKKIEN YLEISYYDEN VERTAILU	11
4.4 YLEISIMMÄT LIIKENNEMERKIT	14
5 TULOSTEN POHDINTA.....	14
5.1 KANTAKAUPUNGIN MERKITYS LIIKENNEMERKKIEN JAKAUTUMISESSA	14
5.2 KANTAKAUPUNKI EROTTUU MUUSTA KAUPUNGISTA	15
6 TUTKIMUKSEN RAJOITUKSET	16
7 LÄHTEET JA VIITTEET	16

1 Johdanto

Autolla ajamisen kokemus vaihtelee eri puolilla Helsinkiä. Tämä tutkimus pyrkii selvittämään, mitä liikennemerkit paljastavat kaupungista. Tutkimuksen päätavoite on siis vertailla, miten eri liikennemerkkejä esiintyy eri kaupunginosissa.

Kesällä 2020 voimaan astunut uusi tieliikennelaki (729/2018) velvoittaa tienpitäjiä toimittamaan tietoja asetetuista liikenteenohjauslaitteista, kuten liikennemerkeistä, ajoratamaalauksista ja liikennevaloista Väylävirastolle (vayla.fi, päivitetty 14.4.2023). Väylävirasto hallinnoi tietoja Digiroad-tietojärjestelmässä, johon on koottu Suomen tie- ja katuverkon keskilinjageometria ja tärkeimmät ominaisuustiedot.

Tutkimuksen kirjoittamisen aikaan (09/2023) Digiroad-järjestelmän tiedot pääkaupunkiseudun liikennemerkeistä olivat melko puutteellisia. Sen sijaan pienempien kuntien, kuten Järvenpään, Keravan ja Orimattilan rekisterit liikennemerkeistä olivat hyvin kattavat.

Kattavampaa paikkatietoa Helsingin liikennemerkeistä tarjoaa esimerkiksi ruotsalaisen Mapillary AB:n *Mapillary*-palvelu. Mapillary kerää katonäkymäkuvia käyttäjiltään ja konenäköä hyödyntäen inventoi katuominaisuuksia, kuten katuvaloja, liikennevaloja ja liikennemerkkejä.

Tämän tutkimuksen tarkoituksena on luoda oma tutkimusaineisto ja tunnistaa liikennemerkkejä konenäön avulla. Mapillary-palvelun paikkatietoa Helsingin liikennemerkeistä käytetään tutkimuksessa vertailuaineistona (*ground truth*), johon omia tuloksia verrataan.

2 Datan kerääminen ja valmistelu

Liikennemerkkien esiintyvyyksien selvittämiseksi oli kerättävä kuvia ympäri Helsinkiä, joista koneoppimismalli tunnistaisi kuvissa näkyvät liikennemerkit. Jotta malli ”oppisi” osoittamaan ja tunnistamaan kuvista liikennemerkkejä, oli sille syötettävä suuri määrä kuvia, joihin on jo valmiiksi merkattu ja nimetty kuvassa näkyvät liikennemerkit. Tätä aineistoa kutsutaan *koulutusaineistoksi*.

Tutkimusta varten kerättiin siis aineistoa sekä koneoppimismallin kouluttamista että lopullista liikennemerkkien tunnistusta varten. Aineistoa kerättiin Google Street View -API:n avulla sekä hyödyntämällä netistä löytyviä valmiita liikennemerkkien tunnistamiseen tarkoitettuja koulutusaineistoja.

2.1 Oman aineiston hankkiminen

Käytössä oli Helsingin kaupungin karttapalvelun aineisto Helsingin peruspiireistä ja Digiroad-aineisto tieverkosta.

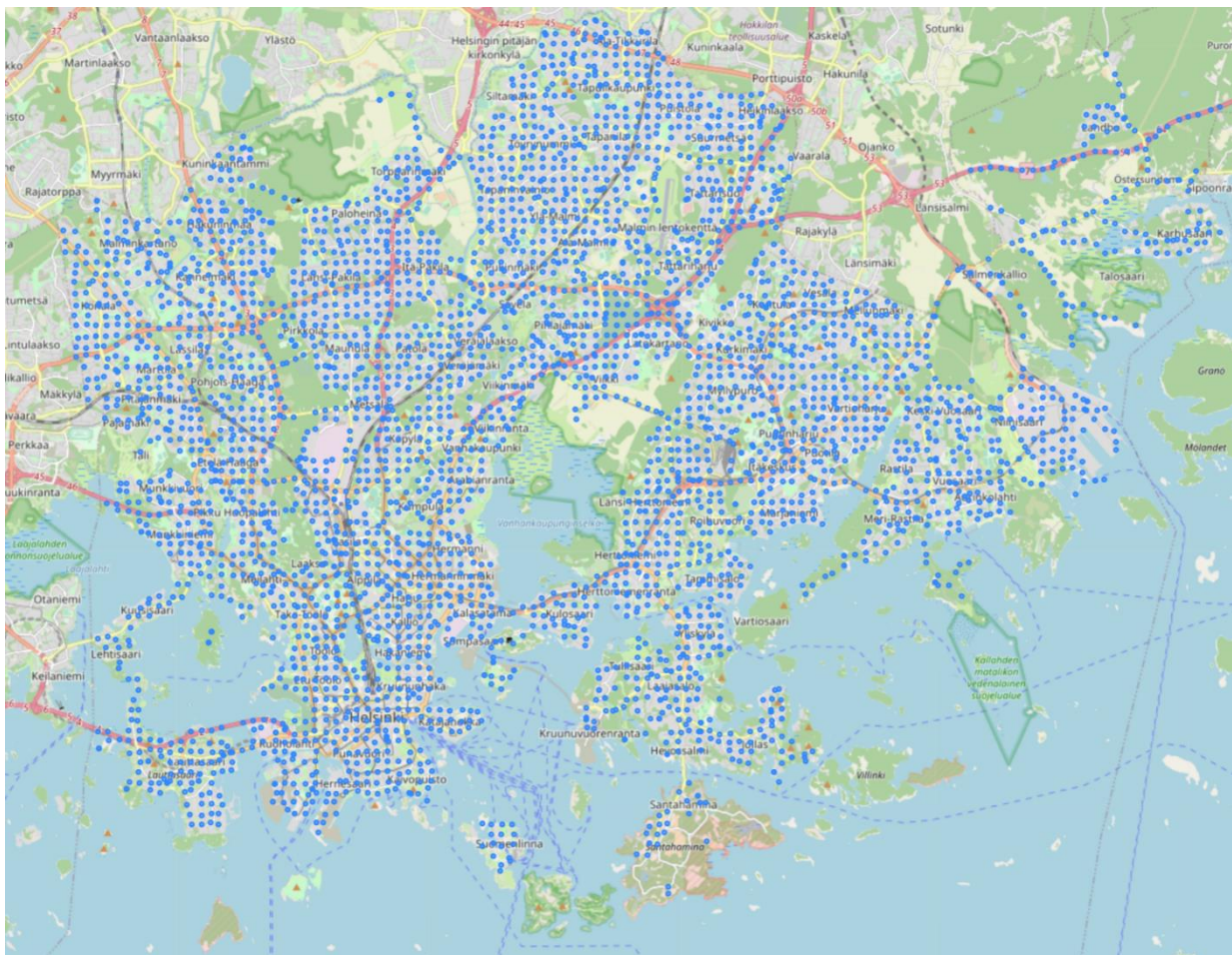
Oman aineiston kerääminen aloitettiin asettamalla Helsingin kartan päälle pisteruudukko, jossa pisteiden väliset etäisyydet olivat 200 m. Näin saatiin vertailukelpoinen satunnaisosotus jokaiselta peruspiiriltä. Pisteiden välisiksi etäisyyksiksi valittiin 200 m Google Street View -API:n käyttöön liittyvien resurssirajoitteiden vuoksi.

Seuraavaksi pisteet, jotka eivät olleet teiden lähellä, suodatettiin pois. Lopuksi pisteet siirrettiin lähimmän tien päälle ja laskettiin alla olevan tiesegmentin kompassisuunta. Näin saatiin jokaiselle pisteelle seuraavat attributit: ID-luku, koordinaatit, peruspiiri ja tien suunta. Pisteitä oli 4053. (Kuva 1)

Seuraavaksi jokaisen GPS-pisteen kohdalta otettiin kuvia Google Street View -API:n avulla tien suuntaisesti eteen- ja taaksepäin. Kuvien resoluutio oli vain 640 x 640 pikseliä, minkä vuoksi kuvia otettiin kustakin suunnasta kolme. Kuvien näkökenttä (fov) oli 20°. Tällä menetelmällä saatiin mahdollisimman paljon liikennemerkkejä kuvattua kustakin GPS-pisteestä. 403 pisteestä ei ollut saatavilla kuvia, joten kuvia saatiin otettua yhteensä $6 \cdot 3650 = 21\,900$. (Kuva 2)

Tekoälyn koulutusta varten valittiin jokaisen GPS-pisteen kohdalta kaikki toiseen suuntaan otetut kuvat eli kuvista valittiin puolet. Seuraavaksi kuvaan manuaalisesti merkattiin niissä näkyviä liikennemerkkejä (Kuva 3). Merkkauksprosessissa keskityttiin vain 18 erityyppiseen liikennemerkkiin. Kaikkiin kuvaan ei siis ole merkattu kaikkia niissä näkyviä merkkejä.

Koulutusaineistoa varten luotiin vielä toinen lista GPS-pisteitä aiemmin kuvatulla prosessilla, kuitenkin niin, että pisteruudukkoa oli siirretty 50 m etelään ja länteen. Kuvia otettiin vain yhteen suuntaan näkökentällä 18°. Näillä menetelmillä pyrittiin välttämään siitä, ettei tässä uudessa aineistossa ole identtisiä kuvia alkuperäisen koulutusaineiston kanssa. Kuvia otettiin 3703, joista 905 merkattiin liikennemerkkejä.



Kuva 1: 4053 GPS-pistettä Helsingin rajojen sisäpuolella (Kuva: Leaflet/OpenStreetView)



Kuva 2: Kolme kuvaa samasta suunnasta (Kuva: Google Street View)



Kuva 3: Liikennemerkkien manuaalinen nimeäminen

2.2 Muu koulutusaineisto

Koneoppimismallin kouluttamiseen käytettiin itse kerätyn aineiston lisäksi netistä löytyviä valmiita liikennemerkkien tunnistamiseen tarkoitettuja koulutusaineistoja. Merkkien oli näytettävä tarpeeksi samanlaisilta suomalaisten liikennemerkkien kanssa.

Näistä ensimmäinen ja suurin oli *Mapillary*-palvelun liikennemerkkiaineisto. Aineistossa on yli 40 000 kuvaa, joihin on merkattu kaikki kuvassa näkyvät liikennemerkit. Erilaisia liikennemerkkejä on yli 300 (*class, luokka*).

Toinen aineisto oli *Kaggle*-sivustolta löytyvä aineisto ruotsalaisista liikennemerkeistä. Aineistoon sisältyy 3112 kuvaa pelkistä liikennemerkeistä. Kuvissa ei siis näy mitään muuta kuin liikennemerkki, toisin kuin Google Street View -aineistossa tai Mapillaryn aineistossa, joissa liikennemerkit näkyvät oikeassa kontekstissa (ks. kuva 3). Merkkejä on 16 eri luokasta.

3 Tutkimusmenetelmät

Liikennemerkkien tunnistamiseen kuvista käytettiin Yolov7-kohteentunnistusalgoritmia. Algoritmi pyrkii paikantamaan kuvista liikennemerkkejä piirtämällä niiden ympärille kehyslaatikoita (ks. kuva 3) sekä luokittelemaan ne eli tunnistamaan mikä merkki on kyseessä.

3.1 Aineiston lisävalmistelut

Itse kerätystä Google Street View -koulutusaineistosta 2914 kuvaan oli merkattu liikennemerkkejä. Aineisto jaettiin koulutus-, validointi- ja testiaineistoon. Validointiaineistoon sijoitettiin 200 kuvaa, testiaineistoon 200 ja loput 2514 koulutusaineistoon.

Koulutusaineistoa käytetään kohteentunnistusalgoritmin kouluttamiseen, validointiaineiston avulla seurataan mallin suorituskyykyyn liittyviä mittareita koulutusvaiheen aikana ja testiaineistolla varmistutaan siitä, että malli kykenee tunnistamaan kohteita kuvista, joita se ei ole ennen nähnyt.

Koulutusaineiston 2514 kuvassa oli merkattuna yhteensä 4702 eri liikennemerkkiä. Taulukossa 1 on esitetty näiden liikennemerkkien määrät luokittain. Huomataan, että luokkajakauma on hyvin epätasainen. Tämä voi johtaa siihen, että kohteentunnistusmalli vinoutuu enemmistöluokkia kohtaan, jolloin se suoriutuu hyvin enemmistöluokkien tunnistamisessa, mutta huonosti vähemmistöluokkien tunnistamisessa.

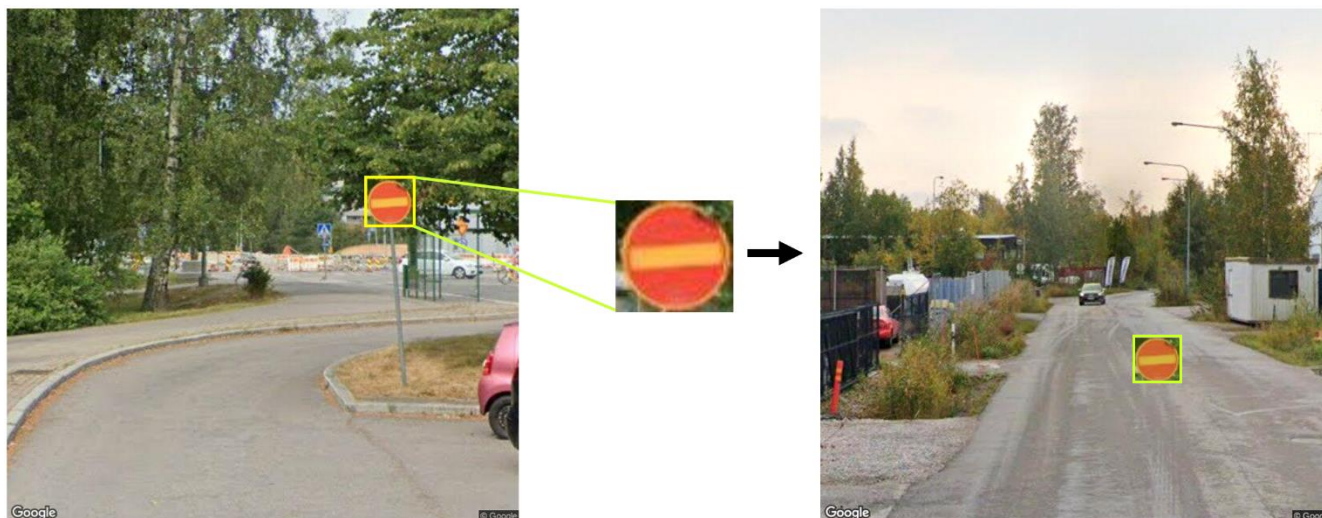
Luokkajakauman tasoittaminen korjattiin ylinäytteisyydellä (*oversampling*) täydentämällä aineistoa copy-paste-metodilla.

Taulukko 1: Koulutusaineiston luokkamäärät

Luokka	Määrä
Suojatie	1943
Pysäköinti kielletty	575
Liikenteenjakaja (oikealle)	506
Pysäköintipaikka	303
Väistämisvelvollisuus risteyksessä	296
Yhdistetty pyörätie ja jalkakäytävä	241
Pyörätie ja jalkakäytävä rinnakkain	142
Työmaa	125
Pysäköintikieltoalue	120
Moottorikäyttöisellä ajoneuvolla ajo kielletty	106
Pysäyttäminen kielletty	87
Kielletty ajosuunta	67
Varo lapsia	56
Vasemmalle kääntyminen kielletty	34
Liikenneympyrä	34
Etuaajo-oikeutettu tie	26
Oikealle kääntyminen kielletty	21
Ajoneuvolla ajaminen kielletty	20
Yhteensä	4702

3.1.1 Copy-paste-metodi

Copy-paste-metodissa kuva kohteesta liitetään uuden taustan päälle. Näin saadaan luotua jo olemassa olevan datan avulla lisää dataa. Omasta Google Street View -aineistosta leikattiin irti liikennemerkit ja liitettiin tyhjiin taustakuviin. Taustakuvat olivat samasta aineistosta kuvia, joissa ei näkynyt liikennemerkkejä. Käytössä oli myös Kaggle-sivustolta saatu aineisto ruotsalaisista liikennemerkeistä.



Kuva 4: Copy-paste-metodi

Jokaiseen taustakuvaan liimatiin päälle 1–3 liikennemerkkiä. Liikennemerkkejä vääristettiin (*skew*) ja niiden kokoa, rotaatiota ja sijaintia varioitiin. (*data augmentation*). Liikennemerkkien kirkkausaste on pyritty korjaamaan suurin piirtein samaksi kuin taustakuvan kirkkaus. (Kuva 5)

Koulutusaineiston luokkamäärät on pyritty tällä copy-paste-metodilla tasaamaan niin, että aineistossa jokaista liikennemerkkiä olisi noin 800, lukuun ottamatta suojatiemerkkejä, joiden määrä pysyi muuttamattomana.

Copy-paste-kuvia muodostettiin yhteensä 3647 kuvaa. Keskenään erilaisia taustakuvia oli 548.



Kuva 5: Copy-paste-metodi ja data-augmentaatio

3.1.2 Mapillary-aineiston valmistelu

Mapillaryn liikennemerkkiaineistossa oli yhteensä 41 909 kuvaa. Kuvia oli ympäri maailmaa, joten liikennemerkkejäkin oli hyvin erilaisia. Mukaan tutkimuksen koulutusaineistoon valittiin vain kuvia, joiden liikennemerkit näyttivät riittävän samanlaisilta suomalaisten liikennemerkkien kanssa.

Lisäksi Mapillary-aineiston kuvat olivat resoluutioltaan liian suuria ja kuvasuhteeltaan vääränlaisia. Yolov7-malliin on syötettävä 640 x 640 pikselin kokoisia kuvia. Tämä korjattiin ensin rajaamalla jokainen kuva kuvasuhteeseen 2:1. Tällöin kuva voitiin jakaa vasempaan ja oikeaan neliön muotoiseen osaan. Kuvien resoluutio muutettiin 640 x 640 pikselin kokoon. Kuvat, joissa ei ollut liikennemerkkejä tai näkyvät merkit olivat vääränlaisia, suodatettiin pois.

Näiden suodatustoimenpiteiden ja muokkauksien jälkeen jäljelle jäi 10 084 liikennemerkillistä kuvaa. Näitä kuvia ei käytetty copy-paste-datan luomisessa.

3.2 Koneoppimismallin kouluttaminen

Kohteentunnistusmallin ytimessä oleva neuroverkko pyrkii jäljittelemään ihmisaivoja. Neuroverkko koostuu neuroneista ja synapseista. Kun koulutusaineiston syöttää neuroverkolle, se pyrkii ennustamaan kuvissa näkyviä merkkejä. Mallin ennustuksien onnistumisien perusteella neuroverkko muuttaa neuronien välisten synapsien painokertoimia. Jokaista kertaa, kun neuroverkko käy koko koulutusaineiston läpi, kutsutaan *epochiksi*.

Kohteentunnistusmallin kouluttaminen aloitettiin syöttämällä pelkästään Mapillary-koulutusaineistoa. Tämän jälkeen malli koulutettiin uudestaan omalla Google Street View- ja copy-paste-datalla käyttäen Mapillary-datalla koulutetun neuroverkon painokertoimia pohjana. Tätä menetelmää kutsutaan kaksivaiheiseksi siirto-oppimiseksi (*Two-Stage Transfer Learning*).

Validointiaineistossa ja testiaineistossa oli kummassakin 200 Google Street View -kuvaa. Koulutusaineisto koostui Google Street View -kuvista, joissa näkyi liikennemerkkejä, copy-paste kuvista sekä Google Street View -kuvista, joissa ei näkynyt yhtään liikennemerkkiä. Taulukossa 2 näkyy koulutusaineiston koostumus. Tyhjät kuvat lisättiin, jotta malli oppisi paremmin tunnistamaan, mitkä objektit eivät ole liikennemerkkejä.

Taulukko 2: Koulutusaineiston koostumus

Lähde	Määrä	%
Google Street View	2576	40,5 %
Copy-paste-kuva	3647	57,4 %
Tyhjä kuva	131	2,1 %
Yhteensä	8798	100 %

Ennen koulutusaineiston kuvien syöttämistä neuroverkon käsiteltäväksi eli ennen jokaista epochia kuvia vielä vääristettiin (*skew*) ja niiden sijaintia, rotaatiota ja kokoa varioitiin. Myös kuvien kontrastia, kirkkautta ja värejä varioitiin.

Kohteentunnistusmallin kouluttaminen oli iteratiivinen prosessi, jossa mallin hyperparametreja ja koulutusaineiston kokoonpanoa muutettiin jokaisen koulutuskerran jälkeen mallin suoriutumisen perusteella.

3.3 Merkkien tunnistaminen

Valmiin kohteentunnistusmallin avulla tunnistettiin liikennemerkit Helsingistä kerätyistä Google Street View -kuvista. Tunnistamiseen käytettiin kaikkia kuvia, myös mallin kouluttamiseen käytettyjä kuvia. Malli tunnsti liikennemerkkejä yhteensä 25 605 kuvasta.

4 Tulokset ja tulkinnat

4.1 Mallin tarkkuusmittarit

Kohteentunnistusmallin kykyä tunnistaa liikennemerkkejä mitattiin usealla mittarilla. Kyseiset mittarit on laskettu ensin yhdistämällä validointiaineisto ja testiaineisto uudeksi testiaineistoksi, sitten antamalla mallin tunnistaa liikennemerkkejä tämän uuden testiaineiston kuvista. Kuvia oli yhteensä 400. Vertaamalla mallin antamia ennusteita testiaineiston todelliseen dataan, saadaan laskettua tarkkuusmittarit.

Ensimmäinen mittari oli tarkkuus (*Precision, P*). Tarkkuus mittaa mallin kykyä tunnistaa liikennemerkkejä oikein. Tarkkuusluku kertoo, kuinka suuri osuus mallin tunnistamista merkeistä on tunnistettu oikein kussakin kategoriassa. Korkea tarkkuus siis tarkoittaa, että malli tekee vähän virheitä merkkien luokittelussa.

Toinen mittari oli herkkyys (*Recall, R*). Herkkyys mittaa mallin kykyä löytää kunkin kategorian liikennemerkkejä kuvista. Korkea herkkyys tarkoittaa, että malli löytää useimmat liikennemerkit kuvista. Toisin sanoen, herkkyys kertoo, kuinka monta oikeaa asiaa malli pystyy löytämään suhteessa kaikkiin oikeisiin asioihin, jotka se voisi löytää.

Kolmas mittari oli *Mean Average Precision (mAP)*. MAP on monimutkaisempi mittari, joka yhdistää tarkkuuden ja herkkyyden. Korkea mAP arvo tarkoittaa, että malli suoriutuu liikennemerkkien löytämisestä ja luokittelusta hyvin. MAP toimii hyvänä yleismittarina. MAP@.5 ja mAP@.5:.95: eroavat toisistaan niin, että ne käyttävät eri kynnysarvoja kohteen tunnistamisessa. MAP@.5: arvioi mallia vain yhdellä kynnysarvolla, kun taas mAP@.5:.95: ottaa huomioon useita eri kynnysarvoja ja näin ollen tarjoaa tarkemman ja monipuolisemman arvion mallin suorituskyvystä.

Oheisessa taulukossa on esitetty tarkastelussa olevien liikennemerkkien tarkkuusmittarit. Osa tunnistetuista merkeistä on jätetty tarkastelun ulkopuolelle.

Taulukko 3: Kohteentunnistusmallin tarkkuusmittarit

Luokka	P	R	mAP@.5	mAP@.5:.95:
Pysäköinti kielletty	0,83	0,85	0,90	0,61
Vasemmalle kääntyminen kielletty	0,43	1,00	0,83	0,56
Pysäköintipaikka	0,92	0,72	0,81	0,56
Moottorikäyttöisellä ajoneuvolla ajo kielletty	0,81	0,63	0,74	0,56
Liikenteenjakaja (oikealle)	0,96	0,87	0,91	0,55
Suojatie	0,94	0,81	0,86	0,54
Väistämisvelvollisuus risteyksessä	1,00	0,75	0,86	0,54
Kielletty ajosuunta	0,84	0,85	0,85	0,53
Liikenneympyrä	0,73	0,75	0,81	0,51
Yhdistetty pyörätie ja jalkakäytävä	0,80	0,65	0,68	0,44
Pyörätie ja jalkakäytävä rinnakkain	0,88	0,54	0,67	0,42
Etujajo-oikeutettu tie	1,00	0,75	0,76	0,35
Yhteensä	0,81	0,73	0,76	0,47

4.2 Vertailuaineisto

Vertaamalla omia tuloksia vertailuaineistoon saadaan parempi käsitys tulosten luotettavuudesta ja käytetyn kohteentunnistusmallin suorituskyvystä. Tässä tutkimuksessa oman kohteentunnistusmallin tuloksia Helsingin liikennemerkkien jakaumasta verrattiin Mapillary-palvelun paikkatietoon. Mapillaryn aineistoon on pyritty merkitsemään jokaisen Helsingin liikennemerkkin sijainti ja tyyppi. Mapillaryn aineistoa käsitellään tässä tutkimuksessa ”todellisuutena” eli sen oletetaan kuvaavan todellisuutta täydellisesti (*ground truth*).

Omia tuloksia voidaan verrata vertailuaineistoon laskemalla niiden välinen keskineliövirhe (*mean squared error, MSE*). Kun n on havaintojen määrä, Y todellinen arvo ja \hat{Y} ennustettu arvo, saadaan keskineliövirhe

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

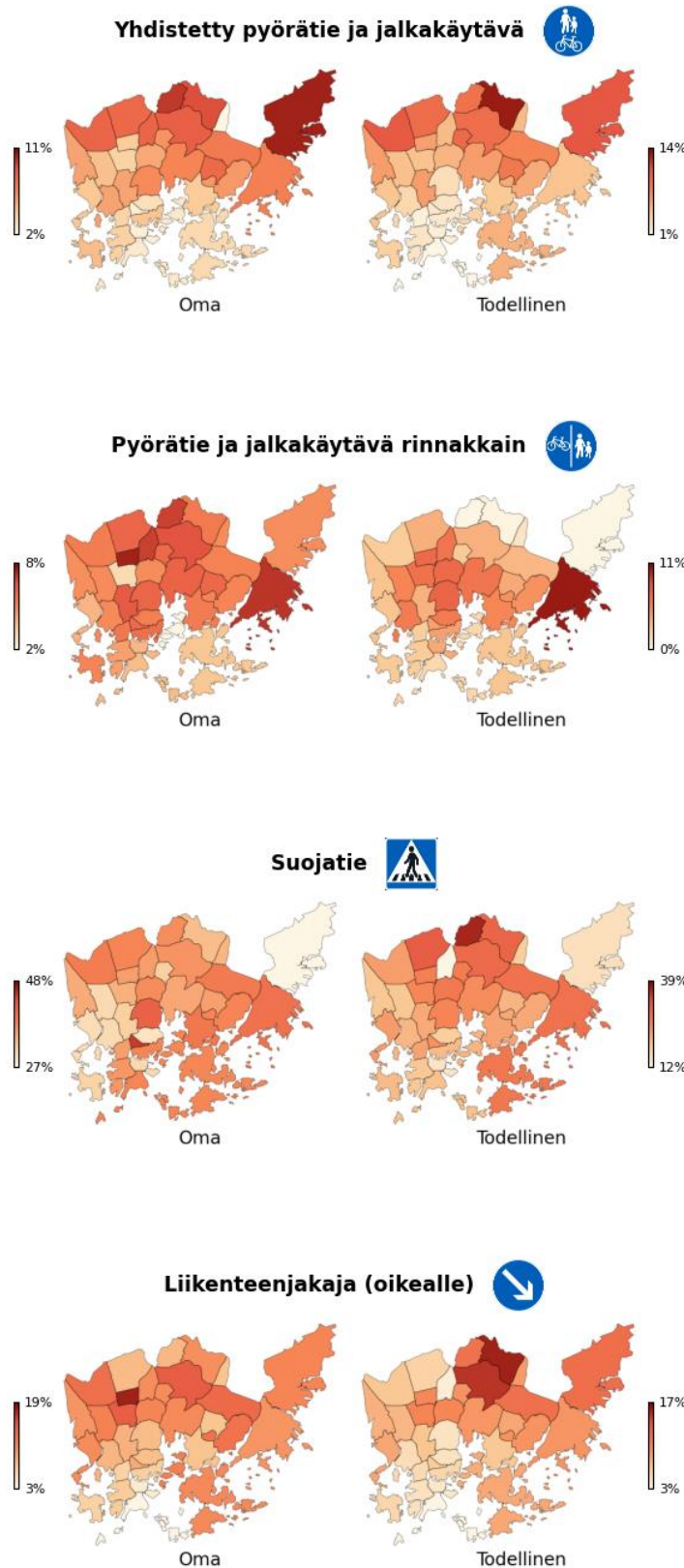
MSE on siis todellisen arvon ja ennustetun arvon erotuksen neliö. Pienemmät luvut tarkoittavat siis pienempää virhettä. MSE-luku on laskettu jokaisen luokan kohdalla erikseen. Ennen MSE-luvun laskemista, jokainen luokka on normalisoitu *Min-Max* skaalauksella välille 0–1. Arvoja on siis siirretty ja skaalattu uudelleen niin, että niiden arvoalueet ovat samalla skaalalla ja alkuperäisessä datassa olevat arvojen väliset suhteet säilyvät. Koska jokaisen luokan luvut on muutettu samalle asteikolle, luokkien MSE luvut ovat nyt keskenään vertailukelpoisia.

Käytännössä MSE luku siis kuvaa, kuinka samanlaisia omat tulokset ovat vertailuaineiston kanssa. Tässä tapauksessa siis sitä, kuinka samalta oma kartta ja vertailuaineiston kartta näyttävät. Pienempi luku tarkoittaa parempaa yhtäläisyyttä.

Oheiseen taulukkoon on merkattu jokaisen luokan MSE-luku sekä mAP arvot.

Taulukko 4: Keskineliövirhe (MSE) luokittain

Luokka	MSE	mAP@.5	mAP@.5:.95:
Pysäköinti kielletty	0,02	0,90	0,61
Yhdistetty pyörätie ja jalkakäytävä	0,03	0,68	0,44
Liikenteenjakaja (oikealle)	0,04	0,91	0,55
Suojatie	0,05	0,86	0,54
Väistämisvelvollisuus risteyksessä	0,05	0,86	0,54
Etuaajo-oikeutettu tie	0,05	0,76	0,35
Pysäköintipaikka	0,06	0,81	0,56
Kielletty ajosuunta	0,06	0,85	0,53
Vasemmalle kääntyminen kielletty	0,08	0,83	0,56
Liikenneympyrä	0,08	0,81	0,51
Moottorikäyttöisellä ajoneuvolla ajo kielletty	0,10	0,74	0,56
Pyörätie ja jalkakäytävä rinnakkain	0,11	0,67	0,42



Kuva 6: Liikennemerkkien yleisyys kartalla

4.3 Liikennemerkkien yleisyyden vertailu

Jokaisen liikennemerkin yleisyys on ilmoitettu prosenttilukuna peruspiireittäin. Prosenttiluku ilmaisee, kuinka suuri osa alueen kaikista liikennemerkeistä on käsittelyssä olevaa liikennemerkkiä. Toisin sanoen

$$f\% = \frac{f}{f_{\text{kaikki}}}$$

jossa $f\%$ on liikennemerkin esiintyvyyss prosentti, f liikennemerkin lukumäärä alueella ja f_{kaikki} kaikkien havaittujen liikennemerkkien kokonaismäärä alueella.

Nämä esiintyvyyss prosentit on esitetty kartalla. Tummemmilla alueilla liikennemerkki on yleisempi kuin vaaleammilla alueilla.

Oheisista kartoista huomataan, että **yhdistetty pyörätie ja jalkakäytävä** on yleisempi kaupungin ulkoreunalla kuin keskustassa. Myös **pyörätie ja jalkakäytävä rinnakkain** on yleisempi keskustan ulkopuolella. Huomataan kuitenkin, että yhdistetyn pyörätien ja jalkakäytävän esiintyvyys on suurin aivan kaupungin ulkoreunalla, kun taas pyörätien ja jalkakäytävän rinnakkain esiintyvyys on suurin hieman lähempänä keskustaa.

Myös **suojatien ja liikenteenjakajan (oikealle)** kohdalla huomataan, että kyseisten merkkien esiintyvyys kasvaa kohti kaupungin ulkoreunoja.

Omat tulokset vastaavat melko hyvin todellista dataa, vaikka joitakin eroja kuitenkin on esimerkiksi suojatien ($MSE\ 0,05$) ja pyörätien ja jalkakäytävän rinnakkain ($MSE\ 0,11$) kohdalla.

Luokka	MSE	mAP@.5:.95:
Yhdistetty pyörätie ja jalkakäytävä	0,03	0,68
Pyörätie ja jalkakäytävä rinnakkain	0,11	0,67
Suojatie	0,05	0,86
Liikenteenjakaja (oikealle)	0,04	0,91

Kielletty ajosuunta -merkkien yleisyyden huomataan olevan suurin kantakaupungin alueella. Yleisyys on suuri myös Östersundomissa.

Tarkastelua varten **pysäköinti kielletty** ja pysäköintikieltoalue-merkit on yhdistetty yhteiseksi pysäköinti kielletty -kartaksi. Huomataan, että näiden merkkien esiintyvyys on suurin kantakaupungin alueella.

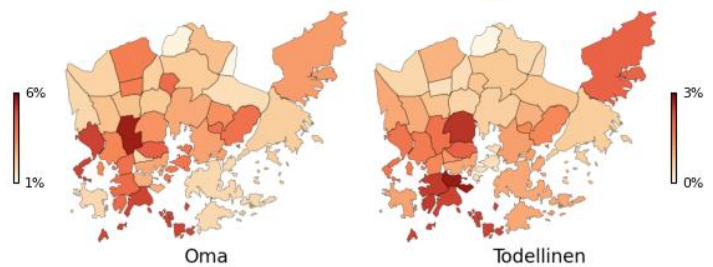
Vasemmalle kääntyminen kielletty -merkit ovat myös yleisimpiä keskustan alueella.

Moottorikäyttöisellä ajoneuvolla ajo kielletty -merkkien kohdalla ei taas ole erityisen vahvaa trendiä kaupungin keskustan suhteen. Huomataan kuitenkin, että merkin esiintyvyys on suurin keskustan ja kaupungin ulkoreunan välisellä alueella.

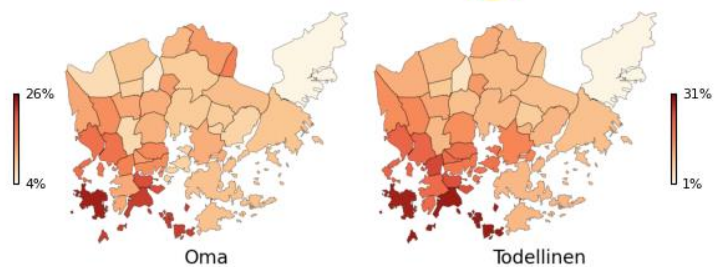
Omat tulokset vastaavat melko hyvin todellista dataa. Joitakin huomattavia eroja kuitenkin on. Oheisista kuvista parhaiten todellisuutta vastaa pysäköinti kielletty -merkin tulokset (*MSE 0,02*).

Luokka	MSE	mAP@.5:.95:
Kielletty ajosuunta	0,06	0,85
Pysäköinti kielletty	0,02	0,90
Vasemmalle kääntyminen kielletty	0,08	0,83
Moottorikäyttöisellä ajoneuvolla ajo kielletty	0,10	0,74

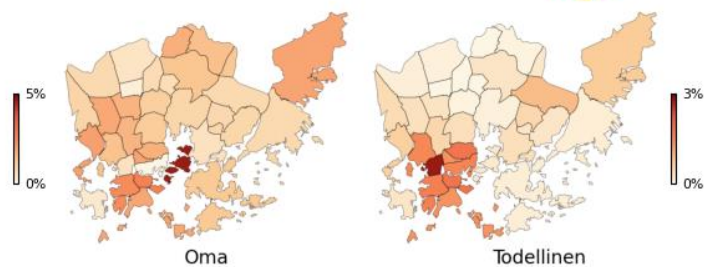
Kielletty ajosuunta



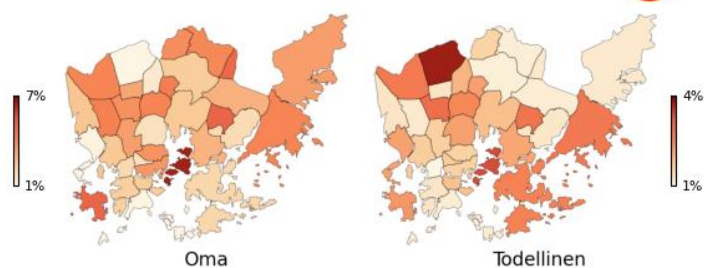
Pysäköinti kielletty



Vasemmalle kääntyminen kielletty

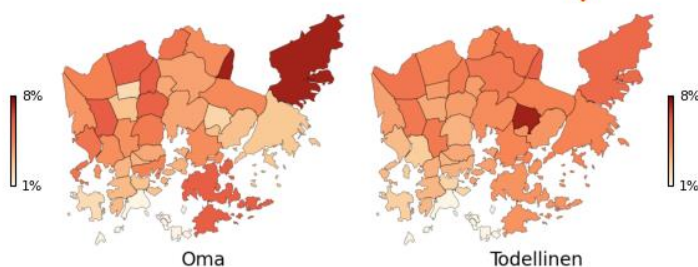


Moottorikäyttöisellä ajoneuvolla ajo kielletty



Kuva 7: Liikennemerkkien yleisyys kartalla

Väistämisvelvollisuus risteyksessä



Väistämisvelvollisuus risteyksessä -merkin esiintyvyys on suurin keskustan ulkopuolella. Esiintyvyys on erityisen suuri esimerkiksi Östersundomissa.

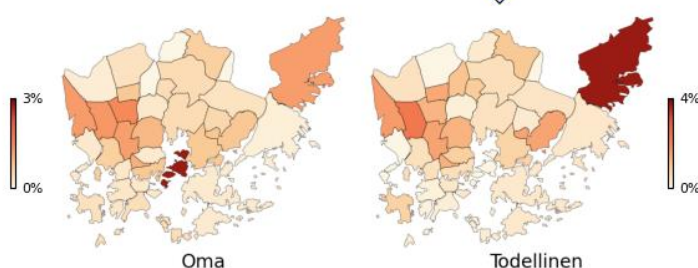
Myös **etuajo-oikeutettu tie** -merkit ovat yleisempiä keskustan ulkopuolella – erityisesti suoraan keskustan pohjoispuolella Pitäjänmäki-Vanhakaupunki-akselilla. Esiintyvyys on suurin Östersundomissa.

Pysäköintipaikka -merkin kohdalla huomataan esiintyvyyden olevan suurinta kaupungin eteläisissä osissa.

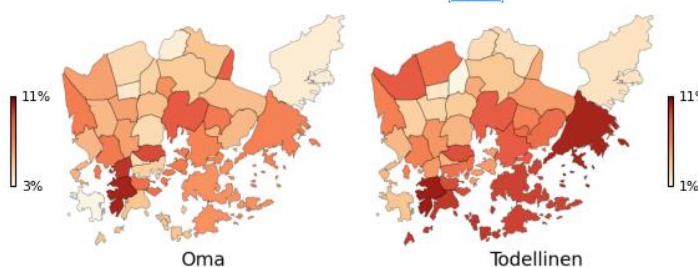
Liikenneympyröiden kohdalla ei ole erityisen vahvaa trendiä. Esiintyvyys on kuitenkin hieman suurempaa kaupungin ulkoreunoilla kuin keskustan alueella.

Omat tulokset vastaavat todellista dataa jossain määrin. Esimerkiksi liikenneympyrän kohdalla omat tulokset eivät ole luotettavia ($MSE\ 0,08$).

Etujajo-oikeutettu tie

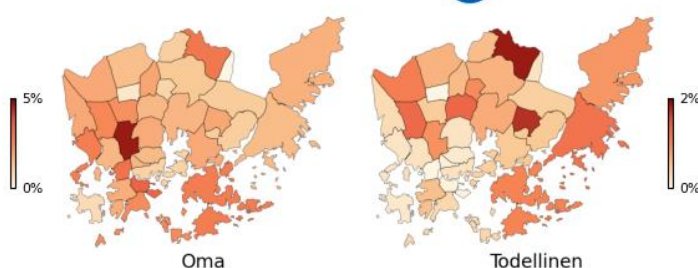


Pysäköintipaikka



Luokka	MSE	mAP@.5:.95:
Väistämisvelvollisuus risteyksessä	0,05	0,86
Etujajo-oikeutettu tie	0,05	0,76
Pysäköintipaikka	0,06	0,81
Liikenneympyrä	0,08	0,81

Liikenneympyrä



Kuva 8: Liikennemerkkien yleisyys kartalla

4.4 Yleisimmät liikennemerkit

Katsotaan vielä tunnistettujen merkkien lukumääriä. Huomataan, että koko kaupungin yleisin liikennemerkki on suojatienmerkki. Toiseksi yleisin on pysäköinti kielletty -merkki ja kolmanneksi yleisin liikenteenjakaja oikealle. Omien havaintojen jakauma noudattaa vertailuaineistoa melko hyvin.

Huomaa, että taulukkoon ei ole merkitty kaikkia tunnistettuja merkkejä. Merkkien yleisyysprosentin laskemiseen käytettyyn f_{kaikki} -lukuun sisältyy kuiteinkin kaikki merkit.

Taulukko 5: Havaittujen merkkien lukumäärä

Luokka	Vertailuaineisto	Oma
Suojatie	20377	5981
Pysäköinti kielletty	14957	2085
Liikenteenjakaja (oikealle)	6328	1732
Pysäköintipaikka	5994	1050
Yhdistetty pyörätie ja jalkakäytävä	3883	988
Pyörätie ja jalkakäytävä rinnakkain	3839	938
Väistämisvelvollisuus risteyksessä	3605	740
Moottorikäyttöisellä ajoneuvolla ajo kielletty	1443	557
Kielletty ajosuunta	1137	463
Etuaajo-oikeutettu tie	768	97
Vasemmalle kääntyminen kielletty	703	234
Liikenneympyrä	497	296

5 Tulosten pohdinta

5.1 Kantakaupungin merkitys liikennemerkkien jakautumisessa

Tuloksista huomattiin, että joidenkin liikennemerkkien esiintyvyys korreloi selvästi etäisyyden kaupungin keskustasta kanssa. Joidenkin merkkien kohdalla erityisen selvää riippuvuussuhdetta ei huomattu. Taulukossa 6 on esitetty liikennemerkkien yleisyyden ja etäisyyden kaupungin keskustasta väliset korrelaatiokertoimet ja p-luvut liikennemerkillä. Piste, johon etäisyydet on laskettu, on Mannerheimintien ja Eteläesplanadin risteys.

Taulukko 6: Liikennemerkin yleisyyden ja etäisyyden kaupungin keskustasta välinen korrelaatio r sekä p -luku

Luokka	r	p
Yhdistetty pyörätie ja jalkakäytävä	0,82	0,00
Pysäköinti kielletty	-0,82	0,00
Liikenteenjakaja (oikealle)	0,71	0,00
Väistämisvelvollisuus risteyksessä	0,7	0,00
Vasemmalle kääntyminen kielletty	-0,63	0,00
Kielletty ajosuunta	-0,53	0,00
Liikenneympyrä	0,49	0,00
Pysäköintipaikka	-0,38	0,03
Suojatie	0,33	0,06
Etuaajo-oikeutettu tie	0,3	0,08
Pyörätie ja jalkakäytävä rinnakkain	-0,18	0,32
Moottorikäyttöisellä ajoneuvolla ajo kielletty	-0,01	0,94

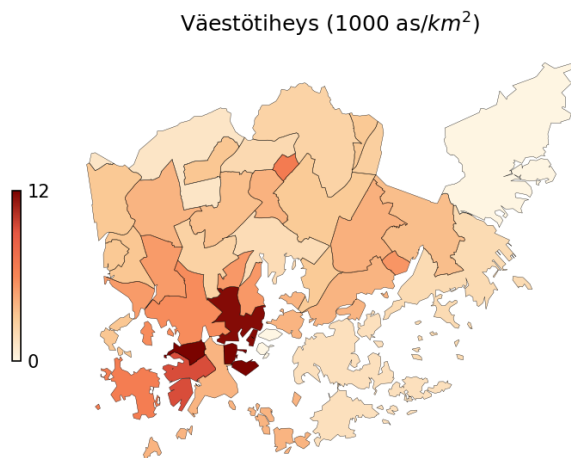
Huomataan, että useiden liikennemerkkien kohdalla korrelaatio on hyvin voimakas. Osa liikennemerkeistä, kuten yhdistetty pyörätie ja jalkakäytävä, väistämismääräisyys risteyksessä ja liikenneympyrä ovat yleisempiä kauempana keskustasta. Osa merkeistä, kuten pysäköinti kielletty, vasemmalle kääntyminen kielletty ja kielletty ajosuunta ovat taas yleisempiä kantakaupungin tuntumassa.

Huomaa kuitenkin, että korrelaatiokerroin heijastaa lineaarisen riippuvuuden voimakkuutta. Esimerkiksi moottorikäyttöisellä ajoneuvolla ajo kielletty -merkin kohdalla huomattiin, että sen esiintyvyys on suurin keskustan ja kaupungin ulkoreunan välissä. Merkin esiintyvyydellä on siis selvä yhteys etäisyyteen keskustasta. Suhde ei ole kuitenkaan lineaarinen, minkä vuoksi korrelaatiokerroin on hyvin alhainen.

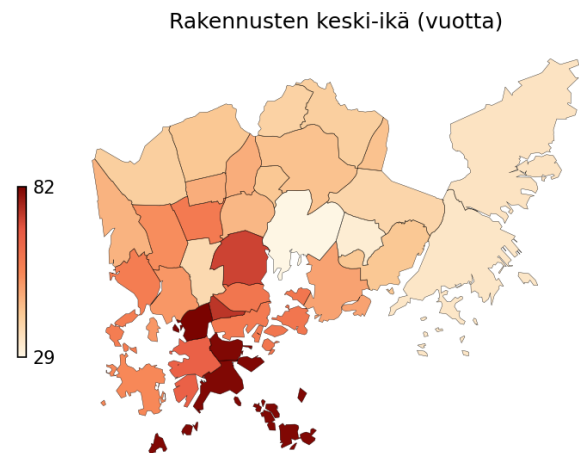
5.2 Kantakaupunki erottuu muusta kaupungista

Kaupungin keskustan merkityksellisyys johtuu todennäköisesti erilaisesta – vanhemmasta ja tiiviimmästä – kaupunkisuunnittelusta. Kuvissa 9 ja 10 on esitetty kaupungin väestötiheys ja rakennuskannan keski-ikä peruspiireittäin. Näiden muuttujien kautta tarkasteltuna kantakaupunki erottuu muusta kaupungista vanhalla rakennuskannalla ja suurella väestötiheydellä.

Satoja vuosia vanhan kaupunkitiestön sovittaminen autoilulle näkyy esimerkiksi useina yksisuuntaisina teinä ja rajoituksina pysäköinnille ja käännöksille, kuten taulukosta 5 huomattiin.



Kuva 9: Helsingin väestötiheys peruspiireittäin



Kuva 10: Helsingin rakennusten keski-ikä peruspiireittäin

Rakennusten keski-ikä laskettiin hyödyntämällä tilastoa Helsingin rakennuskannasta. Tilastossa on ilmoitettu rakennusten rakennusvuosi vuosikymmenittäin. Rakennusten ikä on siis arvioitu kymmenen vuoden tarkkuudella. Rakennusvuosikymmen on ilmoitettu rakennusvuodesta 1920 alkaen. Ennen vuotta 1920 rakennetut rakennukset on oheista tarkastelua varten oletettu 100-vuotiaiksi. Rakennusten numeerinen ikä ei ole tarkastelua varten oleellinen. Keskeistä on erot kaupungin alueiden välillä.

6 Tutkimuksen rajoitukset

Kuten tulokset-osiossa huomattiin, omat tulokset eivät täysin vastaa todellisuutta osin kohteentunnistusmallin heikon suorituskyvyn ja osin tutkimusmenetelmien vuoksi. Omien tuloksien luotettavuuteen vaikuttavat esimerkiksi rajallinen otos havaintopisteitä, rajallinen koulutusaineisto sekä osin suuri määrä copy-paste-metodilla luotua dataa. Toisaalta myöskään Mapillaryn vertailuaineiston paikkansapitävyydestä ei ole täyttä varmuutta.

7 Lähteet ja viitteet

Käytetty kohteentunnistusalgoritmi: Yolov7 (<https://github.com/WongKinYiu/yolov7>, ladattu 12.8.2023)

Käytetty kuva-aineisto:

- Oma kuva-aineisto: Google Street View API, kuvat otettu 7.7, 8.7 ja 9.9.2023
- Ruotsalaiset liikennemerkkit: <https://www.kaggle.com/competitions/sweden-traffic-signs-classification-eng/data>, ladattu 23.7.2023
- Mapillaryn koulutusaineisto: <https://www.mapillary.com/dataset/trafficsign>, ladattu 9.9.2023

Käytetty paikkatietoaineisto:

- Peruspiirit: Helsingin kaupungin karttapalvelu (<http://kartta.hel.fi/>, ladattu 4.3.2023)
- Helsingin maa-alan paikkatietoaineisto: Maanmittauslaitos (<https://asiointi.maanmittauslaitos.fi/karttapaikka/tiedostopalvelu>) ja Väyläviraston Digiroad (<https://suomenvaylat.vayla.fi/>), ladattu 5.3.2023
- Väestötilastot ja rakennuskantatilastot: Helsingin kaupungin karttapalvelu (<http://kartta.hel.fi/>, ladattu 21.9.2023)