Exploring the Use of Machine Learning and Remote Sensing for Traffic Map Generation at Large Scale

Machine learning for feature detection and user applications
ECMWF ML workshop, virtual
1st April 2022
Speaker: Taha Alfaqheri
Introduction

- **Carbon dioxide** is the largest constituent of road traffic greenhouse gas emissions [1]:
  - Local Government Authorities (LGAs) are typically responsible for facilitating mitigation of these emissions.
  - Critical to this task is the ability to assess the impact of transport interventions on road traffic emissions for a whole network.

- **Sustainability problem** that could be generated from transformation of traditional cities to smart cities becomes increasingly important. Urban growth trends expected to increase by 70% in 2050 of the world population will live in cities [2].

- Organisations like World Health Organization (WHO) or the European Environment Agency have reported that being exposed to air pollutants (generated from road vehicles) increases the risk of early death [3].

- Air pollution is the biggest environmental threat to health in the UK, with **between 28,000 and 36,000 deaths a year attributed to long-term exposure** (Public Health England (PHE), 2019).
Main motivations

1- Frequent revisits of earth imaging satellite constellations with providing high spatial resolution support the available efforts for updating traffic maps when combined with advanced machine learning techniques.

2- Support the Local Government Authorities (LGAs) to mitigate road network emissions, and critical to this task is the ability to assess the impact of transport interventions on road traffic emissions for a whole traffic network.

3- Support Air quality prediction tasks by reducing the uncertainty level in answering the main question: Where these pollutants originate from.
4 Earth Intelligence-Satellite Air Quality Model (SAQM) project

Satellite images@2019 Maxar Tech.
World View 2/3 Barcelona city

https://business.esa.int/projects/saqm-project
4EI Intelligence-Satellite Air Quality Model (SAQM) project

Image credit: 4 Earth Intelligence Ltd, Imperial College London. Contains OS data © Crown copyright and database right (2021).
Agenda

Multi model ML:
Road segmentation and Vehicle localisation

Data analysis platform:
Average speed estimation and verification

Potential use cases applications
Speed calculation concept

- Panchromatic (PAN)
- Multi Spectral (MS)
- Identified distance of each detected vehicle
- Identified relevant shadow
- PCA analysis
- Coefficients analysis with PAN image
- Detected vehicles
- The time lag between both DBS
- Time lag in sub second
The idea of speed calculation using spaceborne data

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Title</th>
</tr>
</thead>
</table>

1. Road environmental conditions: road lines, vegetation, signboards etc.
2. Low accuracy of vehicle extraction and localisation.
3. Low spatial satellite resolution in 2011 compared to 2022.
4. Limited optimised ML approaches for road segmentation and vehicles detection tasks.
Machine Learning implementation approach

- Satellite data.
- Local government road data.
- Openly available road data.

Multi model machine learning

Speed estimation processing unit

1. Extracted road segments.
2. Estimated average speed (km/h).
3. Time in Y/M/D/H format.

Large scale traffic notification system

Data analysis platform

Air quality prediction system
Multi model Machine Learning implementation

Automatic roads extraction and vehicles localisation tasks
Some of Remote sensing datasets

<table>
<thead>
<tr>
<th>Dataset source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massachusetts Roads Dataset</td>
<td>Openly available spaceborne data, 3 channel resolution with 1 m resolution.</td>
</tr>
<tr>
<td>Cars Overhead With Context</td>
<td>Openly available aerial data, 3 channel resolution with 15 cm resolution.</td>
</tr>
<tr>
<td>(COWC) dataset</td>
<td></td>
</tr>
<tr>
<td>xView dataset</td>
<td>Openly available and one of the largest overhead data, 3 channel resolution with 30 cm resolution, 60 classes and 1 million object instances.</td>
</tr>
<tr>
<td>Worldview-4</td>
<td>Panchromatic, and 4 Multispectral bands (B, G, R, and NIR).</td>
</tr>
<tr>
<td></td>
<td>Panchromatic nadir: 0.31m Multispectral nadir: 1.24m</td>
</tr>
<tr>
<td>Worldview-3</td>
<td>Panchromatic, 8 Multispectral bands, and 8 SWIR</td>
</tr>
<tr>
<td></td>
<td>Panchromatic nadir: 0.31m Multispectral nadir: 1.24m</td>
</tr>
<tr>
<td>SkySat</td>
<td><strong>[SkySat-1, SkySat-2] - A/B Generation</strong></td>
</tr>
<tr>
<td></td>
<td>Panchromatic: 0.86 m Multispectral: 1.0 m</td>
</tr>
<tr>
<td></td>
<td><strong>[SkySat-3 - SkySat-15] - C Generation</strong></td>
</tr>
<tr>
<td></td>
<td>Panchromatic: 0.65m (0.72 m before 30/06/2020) Multispectral: 0.81m (1.0 m before 30/06/2020)</td>
</tr>
<tr>
<td></td>
<td><strong>[SkySat-16 - SkySat-21] - C Generation</strong></td>
</tr>
<tr>
<td></td>
<td>Panchromatic: 0.57m Multispectral: 0.75m</td>
</tr>
</tbody>
</table>
Vehicles detection and localisation
Data Analysis platform: Speed calculations and verification
Large scale roads traffic notification system: Multi model Machine Learning Integration Stage

Extracted road segments → Identified moving vehicles

Calculate the average speed over predefined period

Lower/Higher the threshold speed limits?

Lower → Calculate the traffic trends over the AOI → Save the event in a database

Higher → Customised large scale traffic map for air quality modeling work

Send notification to LGA to improve traffic reduction policies

Send notification to LGA to revise traffic speed management restriction
• On site measurements stations/ **Lack of spatial coverage:**
  • On the Earth’s surface, networks of measurement stations record the concentration of various chemicals predefined locations. Such networks are commonly run by environmental agencies and provide frequent measurements while often lacking in spatial coverage.

• Remote sensing spectrometers:
  The higher spatial resolutions could reach is in **kilometres range** and with little information about the pollutant’s vertical distribution.
The comparison of basic statistics information of Taiwan EPA Monitoring Station and AirBox PM$_{2.5}$ concentration data from October 14 to 27, 2016.\[15\]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Range</th>
<th>Mean</th>
<th>St. Dev</th>
<th>Sampling Frequency</th>
<th>Number of Stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM$_{2.5}$(EPA)</td>
<td>µg/m$^3$</td>
<td>[2.0, 142.0]</td>
<td>20.985</td>
<td>17.064</td>
<td>1 h</td>
<td>74</td>
</tr>
<tr>
<td>PM$_{2.5}$(AirBox)</td>
<td>µg/m$^3$</td>
<td>[1.0, 211.0]</td>
<td>36.430</td>
<td>23.585</td>
<td>5 min</td>
<td>1176</td>
</tr>
</tbody>
</table>
Use case for air pollution estimation research

Binary classification map

**Good condition class**
- Good: Range (0-50)
- Moderate: Range (51-100)

**Bad condition class**
- Unhealthy for sensitive group: Range (101-150)
- Unhealthy: Range (151-200)
- Very unhealthy: Range (201-300)
- Hazardous: Range (AQI>301)

PM2.5 measurement values
Air Quality Index (AQI)

Data correlation analysis
Prepared data samples for ML implementation

Generated traffic map
Weather data

Hybrid models
CNN
LSTM
Conclusions

• Manage traffic and estimate automobile emissions: Using satellite data with machine learning could provide more reliable decision support ability to address air quality prediction issues.

• With the advancements of satellite remote sensing sensors and ML algorithms in remote sensing industry, the automation of features extraction from spaceborne data could be a valuable input source in the following domains:
  • Large scale traffic monitoring systems.
  • Air pollution forecasting systems.
Thank you

Taha Alfaqheri, ta@4earthintelligence.com
4 Earth Intelligence (4EI)
Team names: Taha Alfaqheri, Jonathan Hendry, Callum Clarke, Alex Yeo, Sahir Khan
References


2. Obando Bobadilla, Laura Maria, María Paula Ruiz Nieto, and José Ignacio Rodríguez Molano. 2018. “From Traditional Cities to Smart Cities.” Communications in Computer and Information Science 877: 661–73. https://doi.org/10.1007/978-3-319-95204-8_55.


References

12. xView dataset: http://xviewdataset.org/
16. edge-enhanced super-resolution GAN (EESRGAN), Code and paper: https://github.com/Jakaria08/EESRGAN.