Cambridge Image Project

Purpose

This is an exploratory project to exemplify how publicly available aerial imagery can be used by public officials to target interventions aimed at reducing road traffic collisions (RTCs).

Research Question

Can unsupervised machine learning techniques be used to effectively cluster road segments from satellite images, and can these clusters then be used to generate insights into the types of locations which may prove most beneficial for interventions aimed at reducing RTCs?

Hypothesis

A better understanding of how different types of road segments in Cambridge are related to RTCs can help the city of Cambridge (council of Cambridgeshire?) target interventions aimed at reducing RTC incidents.

Objectives

1. Derive clusters of road segments from the purely visual aspect of their built environment form.
2. Descriptively explore how these clusters are differentially related to RTCs (and different types of RTCs).
3. Identify aspects of road clusters that are more highly correlated with RTCs (and different types of RTCs) to categorize the types of interventions that could prove useful to reducing the likelihood of future RTCs in these various locations/clusters.

Methods

Data:

* 25cm resolution RGB data for the extent of the city of Cambridge was gathered for three time points (2016,2017,2020)
* Ordnance Survey Road network data for the UK
  + Also have OSM data, may need to experiment with how they differ/which proves most effective
* RTC data for Cambridgeshire
  + Point Data for RTCs of the period 2016-2022
  + Potential Columns of Interest
    - Severity
    - Road Condition
    - Visibility
    - # of Casualties (dataset is only RTCs with 1+ Casualties)
    - Manoeuvre (Straight, Left Turn, Right Turn, Both Turn, etc.)
    - Time of Day
    - # of Vehicles
    - Road Type
    - Speed Limit
    - Junction Type (if applicable)
    - Crossing Type (if applicable)
    - Weather
    - Day of Week
    - Reported/Not Reported

Analysis Pipeline:

* Identify points along the road network to be centroids of the unit of analysis (RGB image)
  + May need to experiment with this, but could be every 5m, 10m, 25m, 50m, 100m, etc.
* Create two sets of square buffers around each point
  + Small buffer to capture localized road features (e.g., # of lanes, road material, width of sidewalk, existence of a crosswalk, intersection, traffic light, etc.)
    - Experiment with this, but probably 10m, 25m, 50m (maybe just same as .5\*centroid distance)
  + Larger buffer to capture localish street network features (e.g., part of a longer highway strip, rural/urban/suburban, closed residential loop, surrounded by pedestrian-only areas, etc.)
    - Experiment with this, but probably somewhere in the 100m, 200m, 300m range (maybe just same as .5\*centroid distance)
* Crop images to these buffers to create two parallel image datasets for use in ML pipeline
  + Will likely end up doing two parallel models, but will have to see if this outperforms just looking at one image per analytic data point
* ML Pipeline
  + The ML methodology will be quite similar to article by some Turing fellows in Urban Analytics <https://www.sciencedirect.com/science/article/pii/S0198971522000461>
  + Using some sort of autoencoder train a deep learning model to decompose the images into their features and then recreate them as accurately as possible
    - Experiment with a few different types of architectures, maybe U-Net generates better features?
    - Separate models for the separately sized images
  + Once model is trained, completely ignore the decoding path of the model and simply extract the latent image features that are at the bottom of the encoding pathway
  + Combine features from the two models (each image is a row with columns containing the combined features from both ML models joined together by ID)
  + Run a cluster analysis on the image features to generate clusters of road segments
    - Will experiment with a few types here, K-means may be sufficient
* Once clusters are extracted, perform descriptive analyses to see how clusters are correlated with total RTCs and the varying types of RTCs
  + Dig deeper into the patterns that seem to be popping out more
    - Some basic statistical tests should be sufficient here alongside plenty of graphics
* For patterns that seem of particular interest, look qualitatively at the types of images and types of RTCs, what interventions seem most appropriate for these clusters?
  + Does (improved signage, speed limits, traffic lights, speed bumps, cycle lanes, dividers, priority, etc.) seem to make more sense based on larger field of RTC research?
    - Could stop before this step, leave it as an exercise for the reader… might be good to do at least one case study to get the reader’s mind going about the possibilities

Expected Outputs

* Map of Clusters
* Heatmaps between clusters and different types of RTCs
* Bar graphs (stacks on stacks)
* Table of Suggestions of potential useful interventions for each cluster with higher RTC rates

Diagram

Description automatically generated

General Notes/Resources:

* <https://arxiv.org/abs/2112.03241>
  + Unsupervised Domain Adaptation for Semantic Image Segmentation: a Comprehensive Survey
* <https://www.spiedigitallibrary.org/journals/journal-of-applied-remote-sensing/volume-13/issue-3/038501/Unsupervised-remote-sensing-image-segmentation-based-on-a-dual-autoencoder/10.1117/1.JRS.13.038501.short?SSO=1>
  + Unsupervised remote sensing image segmentation based on a dual autoencoder
    - Not sure this is quite what I want, but has a few good learnings
      * Can use an autoencoder (eg UNet), then take the latent features from the lowest model point, run the cluster analysis on these latent features, and then build up the other side of the AE to see where the latent features fall spatially
      * Not sure how well this method would work with tons of clusters…. Might make sense to use some transfer learning, take a model that can do the basics of segmenting, buildings, roads, water, etc., and then run the cluster analysis on those segmented features… otherwise a CNN is basically being forced to do both steps, this could get it over the hump of figuring out that roads and farm plots should likely not be grouped together
      * This paper also isn’t thinking about discovering features, the idea is that you sort of know what you are looking for beforehand… might it make sense to get together a long list of what different features of the urban environment might be to guide the model formulation process like this article does?

Goal/Hypothesis of the Paper

* Unsupervised ML techniques can be used to discover features of the built environment that are associated with various urban outcomes (vehicular accidents, various types of crime, deprivation, satisfaction, etc.)
  + These features can provide insights for governmental agencies to better understand how to effect different urban outcomes
    - Cases Study Example
      * A better of understanding of how different types of road segments/intersections in Cambridge are related to vehicular accidents/fatalities can help the city of Cambridge target interventions (signage, speed limits, traffic lights, speed bumps, cycle lanes, dividers, priority, etc.)
  + The concept of this paper is to simply take the form of the environment at face value, some studies (Spatial Signatures - Understanding (urban) spaces through form and function) argue that there is a distinct benefit to including aspects of functionality into the analysis
    - “Some aspects of form and function, like human perception of space, are influenced by both and, while clear conceptually, they can be challenging to measure. Broadening the pool of indices that can be deployed ensures better accuracy when characterising existing patterns on the ground. In this section, we detail our proposal to understand urban form and function through what we term “spatial signatures”.”
  + Generating Spatial Units of Interest (options)
    - Administrative
      * No
    - Granular Uniform Grids
      * With enough resolution grids should be clusterable into groups that make sense for our purposes…
    - Morphometric Units
      * Street segments, plots, building footprints, greenspaces, etc.
      * Could potentially do this first through Transfer learning…. Tons of models using RGB exist already to generate these types of groupings
    - Enclosed Tesselation Cell ([Spatial Signatures paper](https://www.sciencedirect.com/science/article/pii/S0197397522001382#:~:text=spatial%20signatures%20as%3A-,A%20characterisation%20of%20space%20based%20on%20form%20and%20function%20designed,consistent%20morphological%20and%20functional%20characteristics.))
      * The portion of space that results from growing a morphological tessellation within an enclosure delineated by a series of natural or built barriers identified from the literature on urban form, function and perception.
    - Pixels
      * The fully unsupervised method using some autoencoders combined with clustering methodologies and likely some smoothing functions (otherwise it may be real spotty)
  + Look into cluster gram
    - A general challenge for k-means is assigning an appropriate number of clusters (k) that represent yet uncovered structure contained within the input data, under the assumption that there is no theoretical rational for selecting a particular value of k. To explore a suitable value for k we implemented a Clustergram (Schonlau, 2002). This presents different potential k values by plotting the weighted mean of the first component of a principal component analysis (PCA) for each individual cluster. Given that the first component of a PCA provides measures the majority of variance contained within the input data, then separation of the clusters along the Y axis gives an indicatior of their difference. Such charts also highlight how clusters are spit by moving to different values of k; with the width of the lines representing the overall size of the cluster for each solution.
  + Likely want to use a CAE Convolutional autoencoder (same as UNET?)
* This paper provides a decent general methodological approach for what I want to do
  + <https://www.sciencedirect.com/science/article/pii/S0198971522000461>
    - Estimating generalized measures of local neighbourhood context from multispectral satellite images using a convolutional neural network
* One option to narrow the scope of this project would be to just look at classifying road segments within Cambridge. In this case I could take the centre point of road segment every 10/25/50/100m and take a 10/25/50/100m buffer around each segment. This would end up giving me a lot more images and would prevent the model from clustering things I don’t care about as much
  + Under this situation I would only be clustering road segments to look at the type of areas that are related to road traffic accidents (although other road outcomes like police stops, fatalities, etc.)
  + There is a change this would give my images to narrow a scope, I might lose some of the network connections that should effect things…
    - How to include larger road network effects?
      * Possibly combine local/larger latent features? 2 models with different buffer sizes whose latent features are combined in a cluster analysis, can iterate a couple times dropping features that don’t seem to be contributing…
* Basic autoencoder vs UNet type model, which gives better latent features?
  + <https://arxiv.org/abs/2004.02788>
    - Scene de-occulsion paper might have potential code to use
  + Testing Types of autoencoders
    - <https://www.mdpi.com/1424-8220/21/13/4294/htm>
    - <https://www.mdpi.com/2072-6694/13/9/2013/htm>
    - <https://arxiv.org/pdf/2003.05991.pdf> general book chapter on autoencoders

Other traffic accident/risk prediction work

Identifying street/traffic patterns that are related to various types of vehicular accidents

* <https://ieeexplore.ieee.org/abstract/document/8621996>
  + RiskSens: A Multi-view Learning Approach to Identifying Risky Traffic Locations in Intelligent Transportation Systems Using Social and Remote Sensing
* <https://openaccess.thecvf.com/content/ICCV2021/papers/He_Inferring_High-Resolution_Traffic_Accident_Risk_Maps_Based_on_Satellite_Imagery_ICCV_2021_paper.pdf>
  + Inferring high-resolution traffic accident risk maps based on satellite imagery and GPS trajectories
* <https://ieeexplore.ieee.org/abstract/document/9158447>
  + A Multi-modal Graph Neural Network Approach to Traffic Risk Forecasting in Smart Urban Sensing