YALE UNIVERSITY

GREY TO GREEN: A NOVEL APPLICATION OF COMPUTER VISION TO QUANTIFY ON-STREET PARKING IN LONDON

A SENIOR PROJECT SUBMITTED TO

THE FACULTY OF THE DEPARTMENT OF THE STATISTICS & DATA SCIENCE

IN PARTIAL FULLFILMENT FOR THE DEGREE OF

BACHELOR OF ARTS

FELIX BRIDGEMAN

ADVISOR: ARIANNA SALAZAR-MIRANDA

SECOND READER: BRIAN MACDONALD

NEW HAVEN, CONNECTICUT

DECEMBER 2024

Acknowledgements

I am hugely grateful to my advisor, Arianna for her patience, support, and optimism throughout this project.

I'd like to extend a huge thank you to Audrey de Nazelle, Mandy Weng, and Alex Morgan at Imperial College London for the data about London parking that served as the basis of this work. I am also grateful to the Yale Center for Research Computing for access to their research computing infrastructure, specifically the Grace Cluster, which was invaluable the completion of this work.

Thank you to Leyli Granmayeh for being my statistical partner in crime – we made it!

Thank you to Isabel Chomnalez for inspiring me to study statistics in the first place. It has been a long road, and I never thought I would end up producing a project like this.

Lastly, I am grateful to all my friends for all their support and advice about this project and have put up with nothing but parking for the last four months!

Abstract

As cities around the world come under greater pressure from climate change and population growth in the following decades, urban space reserved for parking holds a suite of opportunities. To assess the potential for this space to be converted to more environmentally or socially valuable resources, it first needs to be quantified. Conventional approaches to quantify a city's off-street parking inventory are dependent on the availability of high-quality highway data from municipal governments. However, for many cities, especially those in developing countries, municipal data about the quantity of on- and off-street parking are unavailable for a multitude of reasons, while high-resolution satellite data are better suited to providing context about off-street parking lots.

Computer vision are a range ever-more powerful tools that facilitate the processing of enormous amounts of visual data, and have the potential to transform how we understand cites. Despite ubiquitous street view imagery like Google Street View, researchers have yet to produce a model that can accurately predict the total urban area occupied by low on-street parking spaces distinct from highway circulation space. This project presents a novel application of an object detection convolutional neural network to quantify the number of parked vehicles across street view imagery, and in turn, estimate the on-street parking stock of London through the development of a regression model that translates vehicle detections to predictions of the actual number of parking spaces.

Table of Contents

List of Figures

Introduction

The efficient allocation of land resources has become increasingly critical. As a result, parking given its poor utility and detrimental has come under greater scrutiny in recent years. Parking occupies substantial urban space, presents unique environmental burdens, and socioeconomic costs. Addressing these issues requires not only a shift in urban planning priorities but also innovative methodologies to quantify and reimagine the spatial footprint of parking.

Despite the large body of scholarly work about parking significance, efforts to actually quantify parking have received relatively little scholarly attention, and few comprehensive studies have been conducted to measure parking inventory or usage on a citywide scale. Existing methodologies, ranging from manual surveys to satellite imagery analysis, are often constrained by their geographic scope, resource intensity, or lack of precision. These limitations underscore the need for new approaches that combine accuracy, scalability, and accessibility to inform evidence-based policymaking.

This thesis introduces a novel methodology to estimate the number of on-street parking spaces in urban areas, with London serving as a case study. Leveraging advancements in computer vision and machine learning, this research employs a custom-trained convolutional neural network (CNN) based on the YOLO architecture to detect and classify vehicles as moving, parked on-street, or parked off-street. By integrating these detections with regression models, the study translates vehicle counts into estimates of parking space inventories across London boroughs.

As such, the objectives of this research are threefold:

1. To outline the environmental significance of a new, more accessible methodology to quantify parking stock.
2. To custom train a convolutional neural network to classify vehicles as moving, parked on-street, or parked off-street.
3. Using two linear regression models–a “borough-wide” regression and a “streetwise” regression–to translate predictions of vehicles parked on-street by image into a total estimation of the number of on-street parking spaces in London.

This research demonstrates the feasibility and potential of using computer vision and machine learning to estimate on-street parking inventories in urban environments. The final YOLO model achieved strong detection performance, with a mAP50 of 0.685 and within-class precision of 94% for stationary on-street vehicles, highlighting its ability to distinguish between moving and parked vehicles. The borough-wide regression effectively translated vehicle detections into estimates of controlled parking spaces, with the most robust model achieving an adjusted R-squared of 0.7855. However, model performance was influenced by borough-specific factors, including CPZ coverage rates and urban density, which limited generalizability beyond Inner London boroughs.

The streetwise regression methodology provided more granular insights into parking dynamics at the street level, revealing strong relationships between vehicle detections and reported parking spaces but also uncovering significant zero-inflation and variability across road types. While less explanatory than the borough-wide approach, the streetwise analysis highlighted areas for future refinement, including the need to heteroscedasticity and the possibility of deploying Poisson regression. These findings contribute to the broader goal of developing scalable, accurate, and globally applicable tools for urban land-use analysis.

Background and Literature Review

In this section, I first detail the pre-existing scholarly work that highlight why parking is an urban land use of concern, especially for its detrimental environmental effects. I then provide a brief background and application of convolutional neural networks including the You Only Look Once (YOLO) model before detailing the principles of training such models and the metrics used to evaluate their performance.

Why is Parking a Problem?

Since the emergence of the private automobile in the twentieth century, parking has become a ubiquitous feature of urban land use. While essential for automobile-centric transport systems, both on- and off-street parking are generally oversupplied (Mukhija and Shoup 2006), and resultantly occupy huge swathes of land, exacerbating urban sprawl (Davis et al. 2010). Shoup, in *The High Cost of Free Parking* (2005) describes how, in turn, excess parking reinforces automobile dependency, congestion, and urban sprawl.

Statistics

The supply of parking varies from region to region, influenced by factors including automobile dependency and parking requirements. In the average global city, streetspace is estimated to occupy 25–33% of urban land cover (UN-Habitat 2013). In turn, on-street parking is estimated to occupy 25–50% of this space (Furchtlehner, Lehner, and Lička 2022; Guo and Schloeter 2013). Off-street parking also consumes large amounts of space. While data providing a global overview are limited, in the United States where automobile dependency is high, parking lots are estimated to cover 20–40% of land in built-up urban areas (Akbari, Shea Rose, and Taha 2003; Marshall and Garrick 2006).

London, as the locus of this project, reflects these trends. Despite its lower automobile dependency relative to U.S. cities, parking remains abundant. According to the last large-scale study of parking in London, commissioned by Transport for London in 1999, London possesses 6.8 million parking spaces, 46% of which are on-street, 39% are residential off-street spaces, and 15% are in public or private parking lots (MVA in association with Data Collection Ltd. 2000, as cited in Leibling 2014). Given there are 2.8 million cars and trucks in London (Department for

Transport 2024), this amounts to approximately 2.4 parking spaces per vehicle, a similar level of supply to many U.S. cities (Litman 2023). Although data on the areal footprint of parking in London is limited, estimates based on UK parking bay dimensions (2.4 metres by 4.8 metres) suggest these spaces occupy approximately 79 km², or 5% of Greater London's total area (*The Traffic Signs Regulations and General Directions 2002*; British Parking Association 2016).

Why is parking so widespread?

Minimum parking requirements are one of the primary drivers of historic oversupply. Prevalent in the United States, these requirements take the form of both off-street and on-street minimums. Off-street minimums that require developers to provide a set number of spaces for residential and commercial developments are widespread (Shoup 2005; McDonnell, Madar, and Been 2011). Meanwhile, Guo and Schloeter (2013) explain that on-street minimums have effectively been embedded in highway design standards, particularly in residential areas. Importantly, many of the world's most rapidly growing cities in Asia and Latin America are frequently adopting parking minimums (Barter 2012; Ríos Flores, Vincentini, and Acevedo-Duanas 2013), exacerbating urban sprawl at the moment of acute growth.

In response to the negative effects of oversupply, European cities in Belgium, France, Germany, Italy, Switzerland, the Netherlands, as well as the UK have replaced parking minimums with parking maximums (F. Li and Guo 2014). London replaced off-street minimums with off-street maximums in 2004 (Guo and Ren 2013), and most boroughs have limited the affordability and availability of on-street parking through the creation of controlled-parking zones (Leibling 2014).

Environmental Issues

Each parking space, whether or not it is oversupplied in aggregate, imposes significant environmental costs. In addition to the valuable space it occupies, parking has manifold environmental effects that are particularly pertinent as cities struggle with the consequences of climate change. As with other paved surfaces present in the streetscape, parking replaces vegetative cover (Haaland and van den Bosch 2015), thus diminishing urban biodiversity and fragmenting remaining urban ecosystems (Russo and Cirella 2018). Parking surfaces, such as asphalt and concrete, exacerbate the urban heat island (UHI) effect as they both absorb heat and replace the shade and evaporative cooling benefits that vegetation provides (Hoehne et al. 2022;

Norton et al. 2015). Impermeable parking surfaces prevent infiltration, and increase excess stormwater runoff, elevating flood risk to surrounding local communities (Eaton 2018), Moreover, excess runoff from parking surfaces carries harmful pollutants such as motor oil, heavy metals, and surface sealants into delicate aquatic ecosystems (Müller et al. 2020; Scoggins et al. 2007; Zhao et al. 2019). Parking, as impermeable cover, disrupts natural groundwater recharge pathways that, in some climates, may limit recharge below levels necessary to replenish aquifers upon which many cities rely for municipal water supply (Lerner 1990; 2002; Vázquez-Suñé, Sánchez-Vila, and Carrera 2005).

Socioeconomic Issues

In addition to the financial costs caused by its negative environmental and public health impacts, parking is a deeply wasteful land use that places an economic burden on both city governments and residents through other means. In particular, off-street parking lots occupy often-large parcels of cities that could otherwise be used for more economically valuable uses such as housing or commercial development (Mukhija and Shoup 2006). Shoup (2005) argues that, as a result of their low relative utility and revenue, parking lots depress gross property tax receipts, and thus force cities to raise tax rates overall. This reveals a harmful truth: parking is rarely "free" but rather its costs are cross subsidized through higher taxes imposed on other city land users.

Compounded Impacts

Crucially, the harmful environmental and socioeconomic effects of parking and urban greyspace are only expected to worsen as cities come under strain from the compounding challenges of population growth and climate change. Between today and 2050, the number of people living in cities is predicted to rise from 4.22 billion in 2018 to 6.68 billion in 2050 (United Nations 2019), rendering land an ever more precious resource and diminishing the availability of green space (Haaland and van den Bosch 2015). Meanwhile, climate change will exacerbate urban heat and increase the frequency and severity of extreme weather including storms and droughts (Rosenzweig et al. 2018).

In other words, reimagining the use of parking offers an opportunity for cities to mitigate the pressing challenges they face. On-street parking can be converted into nature-based solutions (NBS) such as urban parklets, blue-green infrastructure like bioswales, or increased tree planting

(Croeser et al. 2022). Additionally, it can be converted into other resources such as traffic-calming measures that increase safety for other road users (Thigpen and Volker 2017). For-off street parking, the opportunities are even more plentiful; in addition to green space, parking lots have the potential to be converted into valuable commercial or community resources, directly addressing land scarcity, increasing walkability and supporting climate adaptation (Johansson et al. 2022).

Conventional Methodologies to Measure Parking

For a long time, the study of parking has been both professionally and academically neglected (Manville and Shoup 2005). To this day, most academic studies provide an accurate reflection of a small area, but few studies attempt to systematically estimate the count or areal coverage of parking across entire cities regions (R. Li, Helmrich, and Chester 2022). Most of the studies that have attempted to do so have taken place in the United States where few municipalities keep systematic records of parking data (Gabbe, Osman, and Manville 2021).

Studies that have attempted to calculate citywide parking inventories have tended to use one or more of the four following methodologies, listed in order of spatial scope: 1) conducting in-person surveys (Davis et al. 2010; Marshall and Garrick 2006; MVA in association with Data Collection Ltd. 2000), 2) applying parking requirements to assessor data (Chester et al. 2015; Gabbe, Osman, and Manville 2021; Hoehne et al. 2019; R. Li, Helmrich, and Chester 2022), 3) applying gross space-to-vehicle ratios (Chester, Horvath, and Madanat 2010), and 4) analysis of aerial or satellite imagery (Akbari, Shea Rose, and Taha 2003; Davis et al. 2010; Hoehne et al. 2019; McCahill and Garrick 2014; McCahill et al. 2014).

In-person survey collection and extrapolation

In-person surveys such as those conducted by Marshall and Garrick (2006) provide the most reliable estimate of parking inventory, and when municipal data is unavailable, they can be used to validate other methodologies. They provide detailed inventories of various types of parking. Typically, their application has been limited in geographic scope due to the labour intensity of in-person surveys, and while useful, their reliability decreases when generous assumptions are used to extend their findings to larger geographies.

Application of parking policy to assessor and highway data

The practice of applying parking requirements to assessor data can provide accurate estimations of parking inventory. They involve the collection of (often heterogeneous) data about a city's building stock and streetscape, and then apply policy assumptions to convert these data to estimations of parking stock. The result is rich can provide rich and spatially explicit information about parking, connected to specific parcels of land, that can inform policymaking decisions. Chester et al. (2015) and Hoehne et al. (2019) deployed this methodology to provide citywide estimates of parking inventory in Los Angeles and Phoenix respectively. More recently, Gabbe, Osman, and Manville (2021) ambitiously used the method to estimate parking inventories for seven cities in the Bay Area in California, however such an undertaking is deeply resource-intensive and complicated by the heterogeneity of the data sources used. Additionally, this study design necessitates explicit and consistent parking policy is explicit and applied consistently. Resultantly, it is inappropriate in countries with flexible parking design standards or where standards have changed considerably over time. Moreover, it is limited by the availability of detailed assessor and highway data and thus is harder to apply in developing geographies.

Gross space-to-vehicle ratio estimations

Gross space-to-vehicle methodologies like those used by Chester, Horvath, and Madanat (2010) estimate parking counts based on generalized ratios of parking space per vehicle across large geographies. However, results cannot be validated and are not spatially explicit, limiting their usefulness for precise intracity comparison and understanding the potential of greyspace conversion.

Conventional satellite imagery analysis

The analysis of aerial or satellite imagery is promising due to its global availability and inherent consistency, and it is the most accurate methodology to understand the spatial footprint of parking spaces. However, three major challenges have limited its application so far. Firstly, lower-resolution imagery makes it difficult to distinguish parking from other impervious greyspace, thereby constraining its use to the estimation of the general area of parking surfaces using measurements such as Normalized Difference Vegetation Index NDVI (Akbari, Shea Rose, and Taha 2003). Secondly, the occlusion caused by tree canopies and urban structures hinders the analysis of on-street parking, especially in cities with high levels of urban greenery. Thirdly, while higher-resolution imagery can enable detailed counts of parking spaces, its analysis is resource-intensive, and thus manual analysis has been limited to studies of smaller geographic scope.

London Parking Data and Studies

London's 28 boroughs, unlike most American municipalities, keep detailed records of their parking inventory to inform and enforce parking policy. These highly accurate data come closest to providing a "ground truth" of parking statistics in a city and make the city an attractive case study to estimate parking, given the potential to validate results. However, their scope is limited to parking that is municipally controlled.

In addition to widespread municipal data, London has been the locus of several professional and academic studies. London's 1999 parking study by MVA and commissioned by TfL (Leibling 2014) provides the most comprehensive study of London's parking inventory, estimating 6.8 million spaces. Using a survey-based methodology, the study sampled 5% of London's surface area to estimate parking counts citywide. While it is one of the few studies to account for all types of parking, both on- and off-street, its accuracy is limited by its reliance on a limited survey area. Moreover, it does not go so far as to estimate the areal footprint of the inventory, something researchers have yet to do. The study was revisited in 2005, although at a smaller scale, soon after parking maximums were imposed, to understand changing parking dynamics. Additionally, boroughs have from time to time undertaken their own research such as Westminster City Council with its recent *Parking Occupancy Survey* (2022).

New Methodologies to Measure Parking

Now, the proliferation of satellite and street view imagery (SVI), developments in machine learning (ML) and computer vision (CV), and the growing computer power to process a large number of images have opened avenues to new ways of efficiently measuring the inventory, spatial extent, and even usage of both on- and off-street parking. These new methods have the potential to combine the high accuracy of smaller scale, survey-reliant studies with the higher areal coverage of past studies that have used satellite data alongside conventional satellite processing methodologies such as NDVI analysis.

While both satellite and SVI data have been available for some time, their usefulness has historically been limited by the challenges of analysing large amounts of imagery. However, the advancement of CV techniques, namely the development of object detection and image segmentation models, have made it feasible to automate these tasks.

Satellite and SVI imagery are natural complements with individual strengths that render them more appropriate for measuring different types of parking. Satellite imagery, as already demonstrated, is well suited to identifying off-street parking lots and estimating their spatial footprint. With CV, avenues are now open to identify and enumerate individual spaces to provide off-street parking inventory counts entirely based off primary data. SVI—deployed in this project—mitigates the occlusion challenges inherent to optical satellite imagery and, as a result, has potential to provide effective measurements of on-street parking stock (Together, these data sources enable a more comprehensive approach to parking estimation.)

Crucially, both satellite and SVI data have the advantage of global coverage and inherent consistency across geographies. This opens the possibility of generating parking estimates for cities worldwide in a matter of hours. Moreover, their homogeneity enables direct cross-city comparisons, allowing parking inventories in one city to be evaluated alongside those in another city on a different continent—a critical step toward understanding interregional parking patterns and their implications.

Technical Background

Machine learning (ML) is a subset of artificial intelligence (AI) focused on creating systems that learn from data without explicit programming. Within ML, computer vision (CV) is a field that aims to create systems that can perform visual tasks such as object detection and image classification, and segmentation with levels of performance greater or equivalent to a human performing the same task. Deep learning, a subfield of ML, has driven advancements in CV using neural networks to automatically identify patterns and features from visual data. Unlike traditional algorithms, deep learning models require minimal manual engineering but rather learn directly from raw images (Sarker 2021). In particular, convolutional neural netwroks (CNNs) are specialized types of deep neural networks optimized for visual data (Indolia et al. 2018)

You Only Look Once (YOLO)

YOLO was a game-changing advancement in convolutional neural networks when first released. Instead of processing images via multiple convolutional layers that identify, classify, and refine object detections, the YOLO model "reframes object detection as a single regression problem," that uses a single convolutional layer to simultaneously predict multiple bounding boxes and their respective class probabilities in one go (Redmon et al. 2016). As a result of its cutting-edge structure, YOLO is remarkably fast and computationally efficient. In combination with advancements in hardware computing power, models like YOLO have opened the possibility to processing incredibly large image datasets and even facilitating real-time object detection. Moreover, compared to other CNN models, the YOLO model is particularly suited to the exercise of distinguishing between moving and parked vehicles since it uniquely considers background pixels outside of the bounding box. In the case of vehicle movement classification, these background pixels could provide valuable information as to the location of the car relative to the rest of the road among other things.

CNN Training Principles

In general, the training involves feeding labelled data into a CNN which iteratively adjusts class weights to minimise a loss function. Training CNNs requires large, representative datasets to avoid overfitting and address class imbalance. Moreover, selecting an appropriate confidence threshold involves balancing precision and recall based on the application's needs.

For example, a higher than necessary recall is desirable when the cost of false negatives is high, for example when detecting fraudulent payments (Johnson and Khoshgoftaar 2019) whereas a high level of precision (and thus a higher confidence threshold) is desirable when the cost of false positives is high, for example in facial recognition applications (Perkowitz 2021).

Metrics

Confidence: Box & Class

During the model training phase, as part of the loss minimization function, a CNN generates a variety of confidence scores for each detection. It is important to note that in the context of CNNs, confidence scores do not reflect a models certainty in a probabilistic sense (Wenkel et al. 2021). Some models generate both box confidence and class confidence score, while the YOLO model, given its single layer structure, generates a single combined confidence score (Redmon et al. 2016). Box confidence, in general, is the probability that an object is present. Therefore, in general it can be understood as

$$C_{\text{box}} = \mathrm{P}(\text{object}) \cdot \text{IoU}. \tag{1}$$

$\mathrm{P}(\text{object})$ is the probability of an object (of any class) being present within the predicted bounding box. IOU or Intersection over Union is a measurement of the overlap between the predicted bounding box and the ground truth bounding box given in equation 2. If the bounding boxes completely overlap, the IoU value is 1. Conversely, if there is no overlap, the IOU is 0.

$$\text{IOU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{2}$$

Class confidence on the other hand, as detailed in equation 3, is a measure of the probability of the object belonging to a certain class given it is present.

$$C_{\text{class}} = \mathrm{P}(\text{class}_i|\text{object present}) \tag{3}$$

Relevant to this project, since the YOLO model identifies (predicts bounding boxes) and classifies objects simultaneously, the model generates an overall confidence score that can be understood as

$$C = \mathrm{P}(\text{class}_i) \cdot \text{IOU}. \tag{4}$$

Thus, this overall score reflects both the likelihood of the presence of a class with the model's spatial accuracy. Given each detection has an associated confidence level, the number of object detections and thus the number of true positives, true negatives, false positives, and false negatives will vary depending on which confidence level is set at a threshold.

Precision and Recall

Precision and Recall are two of the most fundamental metrics to understand CNN performance. Precision (defined in equation 3) is a measure of the accuracy of each prediction. Recall (defined in equation 4), on the other hand, is a measurement of how well the model recalls or detects all relevant object. As the confidence threshold increases, fewer false positives are detected but more false negatives are missed; as such, precision increases and recall decreases. Given this, the trade-off between precision and recall guides the selection of an appropriate confidence threshold: precision increases and recall decreases as confidence increase.

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \tag{5}$$

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}} \tag{6}$$

F1-Score

The F1-score (equation 7) is the harmonic mean of precision and recall. As such, the score provides a single metric that equally balances the trade-off between the two values. As such, in cases where concerns over precision and recall are equal, choosing a confidence value at which F1-score is maximised is sensible.

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

Mean Average Precision (mAP)

Mean average precision (equation 6) is arguably the most frequently used metric to assess overall model performance. at different confidence thresholds. It provides a comprehensive assessment by combining precision and recall across multiple classes and IOU levels, offering a single value that reflects overall model accuracy. mAP is conventionally calculated in two forms: mAP50 which is calculated at a fixed IOU value of 0.5, and mAP50-95 which averages mAP between IOU values of 0.5 to 0.95 in 0.05 increments.

$$\text{mAP} = \frac{1}{C} \sum_{c=1}^{C} \int_{\text{IOU}_{\text{lower}}}^{\text{IOU}_{\text{upper}}} \text{P}_c \tag{8}$$

Methodology

In the first part of this section, I detail the steps I take to sample a training set of street view images using London parking data to construct a representative balance of moving and stationary vehicles. In the second part, I detail how this training set is used to custom train a YOLO object detection model to detect vehicles and classify them as moving, stationary on-street, or stationary off-street. I then detail the deployment of the model to estimate the number of cars parked on-street in London. In the last part of this section, I detail the steps taken to construct two forms of linear regression model. The first model–the "borough-wide" model–considers entire boroughs in aggregate and provides an estimate of the number of parking spaces in a borough with the predicted number of cars parked on-street generated by the object detection CNN. The second model–the "streetwise" model attempts to understand the relationship between reported parking data and CNN vehicle detections on a more spatially granular level using 500 randomly sampled street segments.

Parking in London

In London, parking regulations are administered by each of the 33 boroughs including the City of London. To control parking, boroughs have elected to impose controlled parking zones or CPZs in some to most areas of the borough. In a CPZ, residents must obtain a permit to park their vehicle on-street while visitors may only park in specifically marked pay-and-display bays. Inner London boroughs, due to population pressures, tend have higher CPZ coverage levels up to 100%, whereas outer London boroughs have CPZ coverage levels as low as 7% (see table X). Given boroughs only monitor and control parking bays within CPZs, the data obtained from boroughs used throughout this project refer exclusively to controlled parking spaces. As such, especially in boroughs where CPZ coverage is lower, it was likely that our final model would detect more parked cars than reported spaces given that controlled spaces constitute just a fraction of all parking spaces within the borough. In this study, a selection of boroughs was taken at each stage of the process dependent on data limitations (see table 1). Initially, raw parking data was available from 11 boroughs, 8 in Inner London and 3 in Outer London. Additionally, SVI images were available for all London boroughs except the City of London. Due to data limitations across various boroughs, the object detection models were trained from a set of SVI training images sampled from 9 boroughs, 6 in Inner London and 3 in Outer London, representing 19.8% of London's total area. During the deployment phase, due to limited availability of highly granular parking counts in Outer London boroughs, the streetwise regression model domain was limited to four boroughs in Inner London, while the borough-wide regression model was limited to all boroughs were aggregate counts of on-street parking were available. In the final part of the deployment phase, parking count estimations were generated for all boroughs in London excluding the City of London.

Borough	*CPZ coverage*	*Controlled on-street parking*	*Training set sampling*	*Borough-wide regression*	*Streetwise regression*
City of London[1]	100%	-			
Hackney	100%	33,451	✓	✓	✓
Islington	100%	40,857	✓	✓	✓
Kensington & Chelsea[2]	100%	36,640		✓	
Tower Hamlets	100%	28,373	✓	✓	✓
Newham	99%	72,000*		✓	
Westminster	99%	44,003	✓	✓	✓
Camden	98%	35,523	✓	✓	
Hammersmith & Fulham	92%	43,954		✓	
Lambeth	80%	-			
Wandsworth	67%	61,946	✓	✓	
Haringey	60%	-			
Southwark[3]	60%	28,717		✓	
Ealing	47%	-			
Merton	46%	-			
Waltham Forest	46%	-			
Hounslow	36%	-	✓	✓	
Greenwich	30%	-			
Barking and Dagenham	29%	-			
Harrow	27%	-	✓		

[1] SVIs unavailable for City of London.
[2] Fine-grained street-by-street data only included visitors' bays and did not detail numerous residents' bays.
[3] The statistic of CPZ coverage was an estimate provided by the borough. An actual number was not available.

Barnet	22%	-		
Lewisham	21%	-		
Richmond upon Thames	20%	-		
Kingston upon Thames	19%	-		
Sutton	16%	-	✓	✓
Enfield	15%	-		
Brent	14%	-		
Havering	14%	-		
Hillingdon	14%	-		
Redbridge	14%	-		
Bromley	10%	-		
Croydon	10%	-		
Bexley	7%	-		

Table 1. Basic statistics about CPZs and parking on a borough-by-borough basis. CPZ data found on the Parking Action Map (Possible, n.d.)

Part 1: Data Preparation

Description of SVI Data

The Street View Imagery (SVI) dataset used in this study was provided by Arianna Salazar-Miranda. It comprises imagery from all London boroughs, excluding the City of London. Each SVI image is a 400x400 pixel JPEG, offering a consistent resolution across the dataset. The dataset contains 294,074 unique points, or "panoIDs," where images were captured using a 360° camera. At each panoID, four images were taken, corresponding to the four cardinal directions (0°, 90°, 180°, and 270°). This setup results in a total of 1,176,296 individual images.

Each image in the dataset is associated with specific metadata, including:

- **Location**: Identified by the "panoID".
- **Aspect**: The camera's directional perspective (0°, 90°, 180°, or 270°).
- **Date**: The timestamp indicating when the image was captured.

Preparation of London Parking Data

Data files containing information about parking across 11 boroughs in London (see fFigure 1) were provided by Audrey de Nazelle and her associates at Imperial College London. These data were comprised of spatial data either in GeoPackage or GeoJSON formats. Given their heterogeneity, initially, data for each borough were explored in QGIS (QGIS Development Team 2024) to gain familiarity and identify any initial issues.

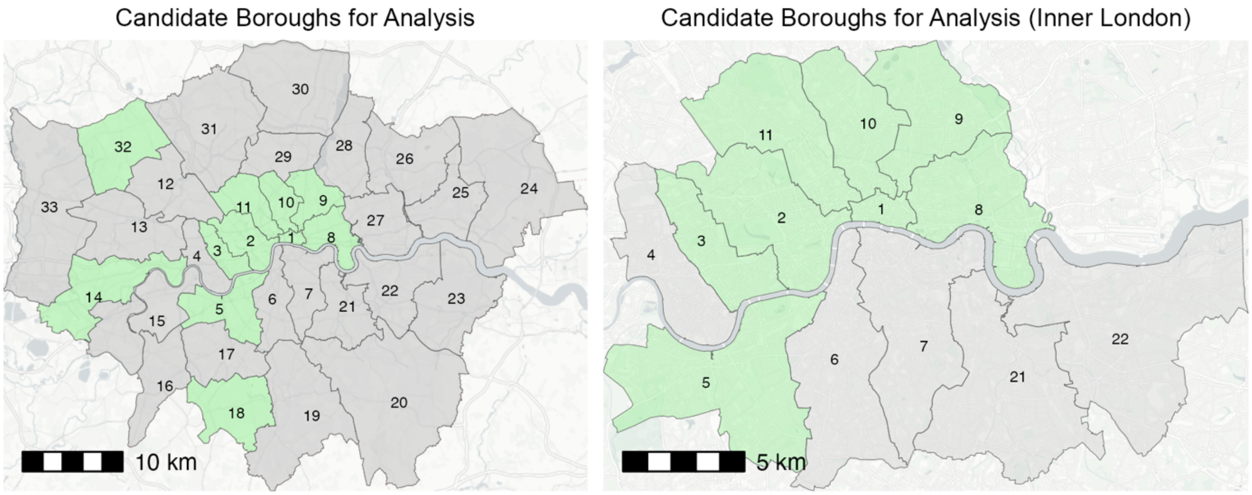

Figure 1. A map of the boroughs for which parking data was originally available.
1) **City of London**, 2) **Westminster**, 3) **Kensington & Chelsea**, 4) Hammersmith, 5) **Wandsworth**, 6) Lambeth, 7) Southwark, 8) **Tower Hamlets**, 9) **Hackney**, 10) **Islington**, 11) **Camden**, 12) Brent, 13) Ealing, 14) **Hounslow**, 15) Richmond upon Thames, 16) Kingston upon Thames, 17) Merton, 18) **Sutton**, 19) Croydon, 20) Bromley, 21) Lewisham, 22) Greenwich, 23) Bexley, 24) Havering, 25) Barking and Dagenham, 26) Redbridge, 27) Newham, 28) Waltham Forest, 29) Haringey, 30) Enfield, 31) Barnet, 32) **Harrow**, 33) Hillingdon (boroughs for which data was available marked in bold).

After exploration, the borough parking data were imported into R (version 4.3.2) for further exploration and cleaning. Using the sf package (version 1.0-19) (Pebesma 2018; Pebesma and Bivand 2023), all data objects were realigned to the same CRS projection, each parking space was given a variable for the borough it fell under and a unique ID, and then all parking datasets of the same shape type (e.g. point, polygon, or polyline) were merged. As such, the final joint datasets had the following variables: ID, number of parking spaces, borough, coordinates, and geometry. These three datasets were used as the basis of SVI image sampling described in the next part, and as the basis of regression model training described in part four.

Part 2: YOLO Model Training and Evaluation

Model Training Set Preparation

To finetune the model, I established an initial workflow that includes 3 steps. First, SVI data were divided into two sets: (1) panoIDs near parking spaces and (2) panoIDs far from parking spaces. Second, an equal number of SVIs were evenly sampled evenly from each set. Third, vehicle information was extracted from the training set using manual image labelling.

Selection of SVI panoIDs

In order to generate a dataset that contained a balance of different classes, it was necessary to split all possible SVIs between areas in high proximity to parking and in low proximity to parking. Given the high prevalence of on-street parking on most urban streets, this step ensured that there would be sufficient images with moving vehicles, or no vehicles present to help the model better distinguish between a moving and stationary vehicle. To do so, I leveraged the cleaned parking dataset, and using the sf package, calculated buffers around parking objects (points, polygons, or lines depending on borough). At this point, images from Kensington & Chelsea were excluded from the training set sampling domain since their format did not enable the generation of meaningful buffers.[4] Resultantly, as SVIs in the City of London were unavailable, the training set geographic domain consisted of 9 boroughs, 6 in Inner London and 3 in outer London. To calculate the buffer, a 5-metre distance was chosen for points and lines, while a 2-metre distance was chosen for polygons. This distance reflects the fact that the boundary of the polygon is already closer to the centre of the street where the SVI's coordinate is located.

Sampling of SVIs

Once the buffers had been generated, the SVIs were split into two sets, those within a buffer or highly likely to be close to a parking space, and those outside a buffer, and thus more likely to show a variety of other streetscapes that feature fewer on-street parked cars such as commercial streets, highways, and service roads. During the first round of SVI sampling, 50 panoIDs were sampled from each set, totalling 100 panoIDs and thus 400 images.

Ground Truth Labelling and Preprocessing

During the initial labelling attempt, a wider range of classes were included; for example, buses were differentiated from trucks, and motorcycles and bicycles were included (see Table 2). This approach was informed by the rationale that, since it was unlikely the initial model would yield satisfactory results, a more specific range of classes would provide valuable information to guide the refinement process. To improve the robustness of the model, vehicles obstructed by other objects (e.g., fences, walls, buildings, or vegetation) were also included. Roboflow (Dwyer et al. 2024), an open source labelling tool, was chosen as the labelling platform. Initially, 402 SVIs were labelled resulting in a total of 1,429 objects labelled. The labelled SVIs were then divided into training, validating and testing datasets at a ratio of 70:15:15. Figure 2 shows a selection of labelled images with different coloured bounding boxes representing different object classes.

---

4. Data from Kensington and Chelsea this borough were comprised of evenly spaced points along the centreline of each street that provided details of nearby on-street parking spaces. While this data provided a general sense of the number of parking spaces in the vicinity of any point, it was insufficiently spatially explicit to be able to generate meaningful buffer zones in which images would likely to be near to a parking space or conversely, outside which images would be likely to be far from a parking space.

	Class		$n_{initial}$	$n_{enlarged}$
Primary classes of interest	**Car**	moving	164	890
		stationary on-street	817	3,440
		stationary off-street	194	797
	Light truck	moving	47	242
		stationary on-street	98	385
		stationary off-street	23	102
Additional classes included during first round of model training.	**Heavy truck**	moving	8	77
		stationary on-street	5	45
		stationary off-street	2	10
	Bus	moving	5	64
		stationary on-street	0	6
		stationary off-street	0	0
	Bicycle	moving	4	33
		stationary on-street	28	75
		stationary off-street	2	0
	Motorcycle	moving	6	22
		stationary on-street	16	76
		stationary off-street	1	7
	Other	-	9	34

Table 2. Classes and total (pre-split) feature counts for objects in the initial and enlarged dataset.

Figure 2. A selection of images with ground truth labels and boxes made using the Supervision package (Roboflow 2024) in Python.

Model Training: First Attempt

Given the purpose of the model was to estimate reliable parking counts, high model performance was of strong importance over other model selection factors like computational efficiency. For this reason, the YOLO11x model, the largest in the YOLO series (i.e. trained on the highest number of parameters) was selected. The pre-trained YOLO11x weights were downloaded and pre-trained on Python (version 3.10.12) using the Ultralytics API. The model was custom trained using 4 GPUs on the Grace HPC cluster made available by the Yale Center for Research Computing. Using the initial dataset of 402 SVIs, the model was trained. The resultant model (referred to herein as the "first model") had a mAP50 of 0.371 and a mAP50-95 of 0.263. Given the relatively low performance of the model across all classes, it was clear that the initial dataset was of insufficient size. This was demonstrated by the "spiky" precision-confidence curves where precision actually decreases for increasing confidence levels, suggesting poor model stability for undersampled classes (e.g. for stationary-light-truck-onstreet in the bottom right of fFigure 3).

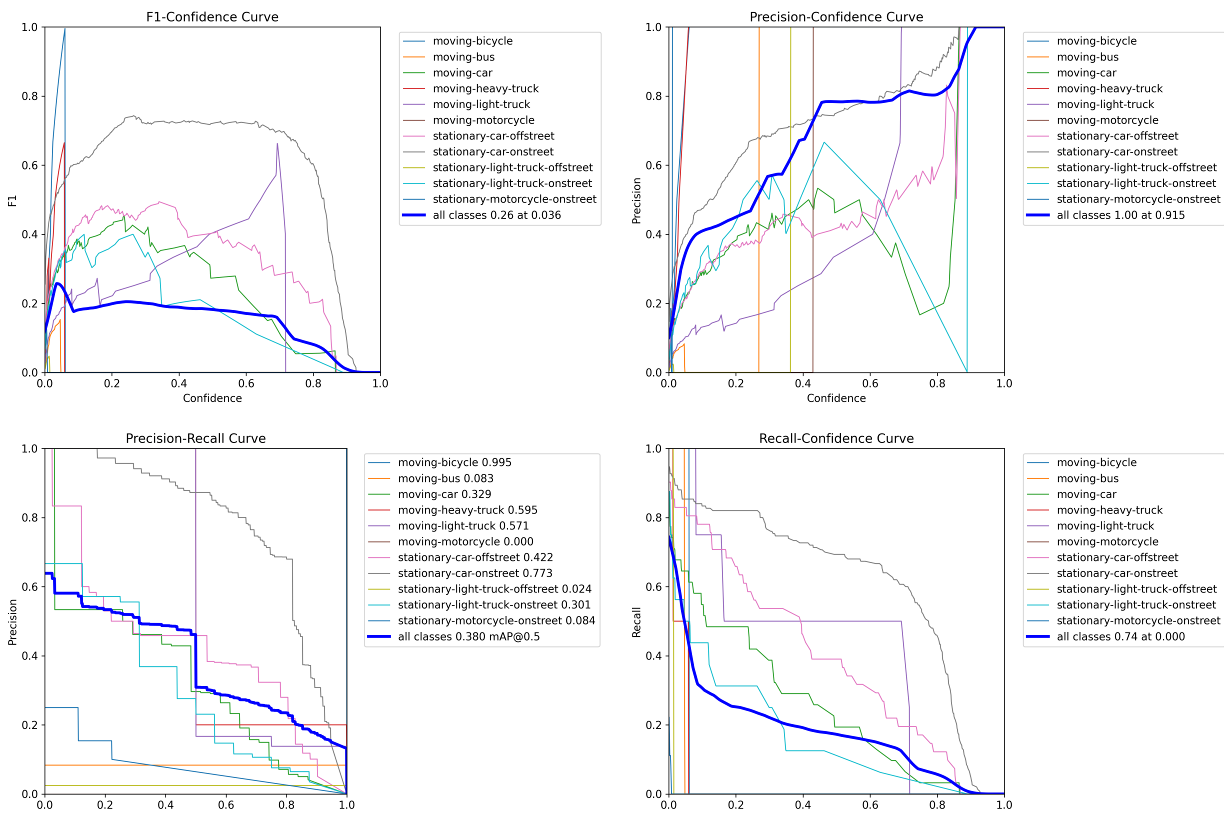

Figure 3: F1-confidence curve (top-left), precision-confidence curve (top-right), precision-recall curve (bottom-left), and recall-confidence curve (bottom-right) for initial object detection model.

Dataset Enlargement

Given the generally depressed performance of the initial YOLO11x model, 300 additional panoIDs were increasing the sample set to 396 panoIDs, amounting to 1,584 images with a total of 6,308 labelled objects. The additional sample was collected from the set of images excluding the existing sample and were labelled through the same process as before and merged with the initial dataset. As with the initial dataset, vehicle types other than cars and light trucks continued to be undersampled, and there was a strong imbalance between vehicle states within each vehicle type, however action was not taken to correct this imbalance given the importance of the model to reflect real world likelihoods.

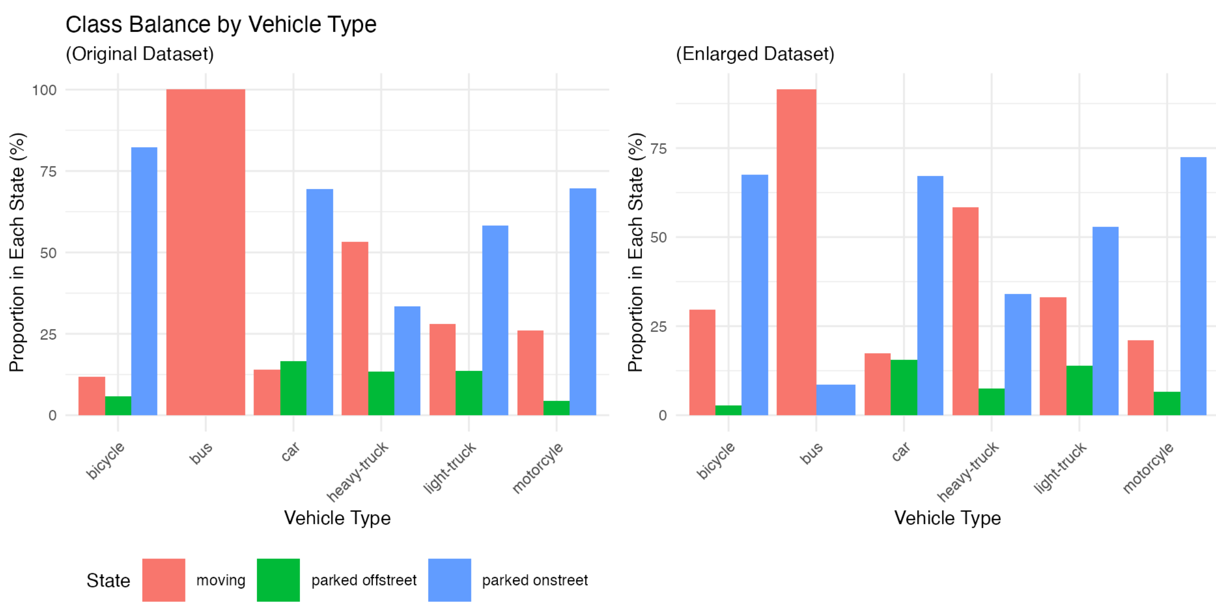

Figure 4. A bar chart showing the within-vehicle proportional class distributions of the initial and enlarged dataset of SVIs ($n_{initial}$ = 1,429, $n_{enlarged}$ = 6,308).

Retraining YOLO11x with the enlarged dataset

Once the dataset was enlarged, the same pre-trained YOLO11x weights were used to custom train the model. Despite quadrupling sample size, performance statistics metrics decreased between the first model and the enlarged model: mAP50 decreased from 0.371 to 0.353, while mAP50-95 decreased from 0.263 to 0.22. At first this decrease seemed concerning, however, the model curves (see Figure 5), in particular for light trucks in the Precision-Confidence curve, indicated that the enlarged amount of training data was mitigating overfitting, and thus the decrease in mAP represented a mitigation of performance overinflation. However, the erratic curves were not entirely alleviated, especially for smaller classes like bicycles and motorcycles indicating that training data remained insufficient for these classes. Given these classes were outside the central aim of the study, and that a distinction between motor vehicle types was not crucial to our goal, it seemed appropriate to remove these classes and condense the remaining classes.

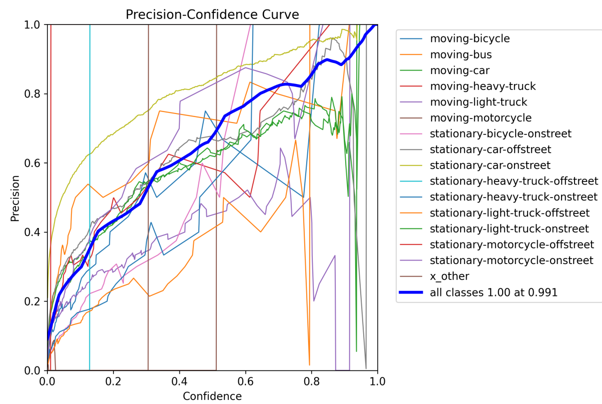

Figure 5. Precision-confidence curve of the enlarged model.

Class Simplification

Given that poor model performance was concentrated in undersampled classes, and that the distinction between different types of motor vehicle was not relevant to identify the number of parking spaces, the enlarged dataset was reclassified: cars, light trucks, heavy trucks, and buses would become vehicles, while bicycles and motorcycles would be eliminated, leaving only three classes (Table 3).[5]

Final classes		**n**
Automobile	moving	1,273
	stationary on-street	3,876
	stationary off-street	909

Table 3. Classes and respective ground truth label counts (**n**) on which the final object detection model was trained.

Retraining YOLO11x with the simplified dataset

In the third training attempt trained the same YOLO11x weights on the new simplified dataset. The resultant model (the “final model”) benefitted from dramatically improved performance: between the enlarged and final models, mAP50 increased 94% from 0.353 to 0.685, while mAP50-95 increased 207% from 0.22 to 0.457. However, “spiky” confidence lines

5. Although buses and heavy trucks take up a greater physical footprint than cars or light trucks, municipal parking data included loading bays and other spaces for such vehicles and thus their inclusion was appropriate.

that created valleys in the precision confidence curve (see Figure 6) indicated instability due to potential overfitting towards stationary-on-street vehicles since this class remained oversampled relative to stationary off-street and moving vehicles.

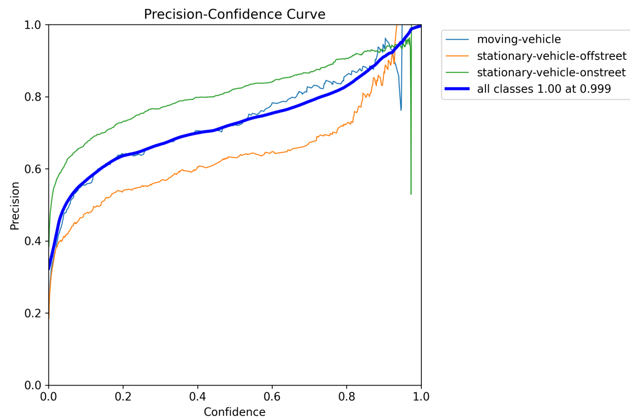

Figure 6. Precision-confidence curve of the final model. Not the downward spike around 0.9 confidence for "stationary-vehicle-onstreet."

Part 3: Model deployment and construction of Linear Regression Model

Once a successful object detection model was constructed, the last remaining challenge was to find an effective means to translate the count of predicted on-street parked vehicles to an estimation of parking spaces. To do so, two methods were deployed. First, I used ordinary least squares (OLS) regression to understand the relationship between aggregate parking counts at a borough level. Second, to understand the model at a smaller spatial scale and to compare the predictive power of the model to other more widely available metrics such as road type and road length, I constructed a street-by-street parking and vehicle detection dataset to build a regression model. First however, the model needed to be deployed on every SVI in London to produce a final dataset.

Choosing an IOU and confidence threshold for detections

To deploy the model, two parameters needed to be specified: the non-maximum suppression (NMS) IOU threshold, and the confidence level. The NMS IoU threshold specifies the maximum degree of overlap allowed between detected bounding boxes for the same object. NMS works by ranking overlapping bounding boxes and removing those with lower confidence scores to ensure that each object is represented by the single most confidence bounding box. Given that there was a high degree of overlap between adjacent parked cars in SVIs that looked along the street (images taken at 0 and 180 degree angles at each panoID), a higher (i.e., less restrictive) threshold warranted consideration. Given this, a range of NMS IOU thresholds were tested between 0.35 and 0.85. Since there was minimal difference in resultant detections between these thresholds, a higher or less restrictive IOU threshold of 0.7 was elected to ensure adjacent parked cars were detected. The model was with a confidence threshold of 0.85, reflecting a preference towards classifying of vehicles as either moving or stationary in the necessary precision-recall trade-off; the differing results are discussed later.

Borough-wide regression model

To elucidate the scaling ratio between the number of detected *stationary vehicles* and the number of *parking spaces*, and to understand how this ratio changed on a borough-to-borough basis, OLS regression was undertaken. First, however, data of borough-wide controlled parking counts were collected. The following twelve boroughs were included in our analysis (in order of decreasing CPZ coverage): Hackney, Islington, Kensington & Chelsea, Tower Hamlets, Newham, Westminster, Camden, Hammersmith & Fulham, Wandsworth, Southwark, Hounslow, and Sutton.

Alongside the CPZ coverage rate, the number of detected on-street stationary vehicles would be proportional to the total number of *controlled* parking spaces under the hypothesis that:

$$\text{total parking spaces} = K_1 * \text{onstreet stationary vehicle detections}$$

$$\text{total parking spaces in CPZ} = K_2 * \text{total parking space} * \text{CPZ coverage}$$

therefore,

$$\text{controlled parking spaces} = K_1 * K_2 * \text{total stationary vehicle detections} * \text{CPZ coverage}$$

where K_1 is some scaling ratio that reflects the fact that not a certain proportion of parking spaces are always unoccupied and that not all vehicles detected as stationary will necessarily be parked in a legal (and thus reported) on-street space, while K_2 represents the scaling factor between controlled parking spaces (i.e., those in the CPZs within the borough) and all on-street spaces throughout the borough.

Given this, two models were constructed in R for each set of detection results (set one at confidence level 0.75, and set two at confidence level 0.85):

$$\text{Reported Controlled Spaces} = \beta_0 + \beta_1(\text{Vehicle Detections}) + \epsilon \qquad \text{(M1.1)}$$

$$\text{Reported Controlled Spaces} = \beta_0 + \beta_1(\text{Vehicle Detections} * \text{CPZ Coverage Rate}) + \epsilon \tag{M1.2}$$

where $\beta_1 = K_1 * K_2$.

Before proceeding with the hypothesised regression equation, the relationships between the predictors (on-street stationary vehicle detections) and response (controlled parking spaces) were investigated to validate linearity assumptions of OLS regression. As demonstrated on the left in Figure 7, there is a clear linear relationship between reported controlled parking spaces and on-street vehicle detections. However, to mitigate for evident heteroscedasticity (although not explicitly problematic) data is normalized over borough area. The resultant linear relationship is enhanced and no longer left-skewed as shown in the figures on the right of Figure 7.

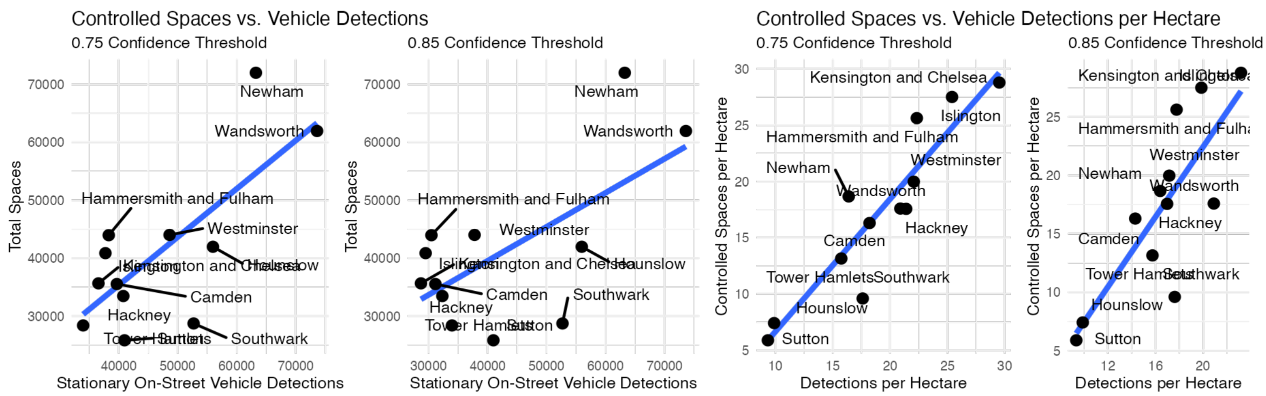

Figure 7. Plots showing vehicle detections against total controlled parking space in each borough, both actual (left) and normalised over borough area (right).

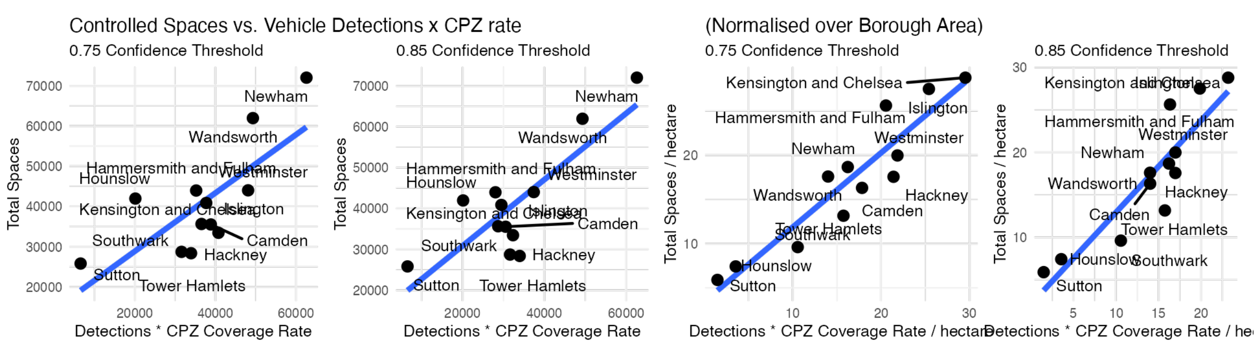

Figure 8. Plots showing vehicle detections *scaled by CPZ coverage* against reported controlled parking spaces in each borough, both actual (left) and normalised over borough area (right).

Given the strengthened linear relationship after normalising spaces and detections over area, the following new equations are additionally modelled in R:

$$\frac{\text{Reported Controlled Spaces}}{\text{Area}} = \beta_0 + \beta_1\left(\frac{\text{Vehicle Detections}}{\text{Area}}\right) + \epsilon$$

(M1.3)

$$\frac{\text{Reported Controlled Spaces}}{\text{Area}} = \beta_0 + \beta_1 \left(\frac{\text{Vehicle Detections } \bullet \text{ CPZ Coverage Rate}}{\text{Area}}\right) + \epsilon$$

(M1.4)

"Vehicle Detections" refer specifically to stationary on-street vehicle detections.

Using the resultant models, predictions were generated for the total amount of parking in London by predicting the reported controlled spaces for each borough and dividing this figure by the CPZ coverage rate. This methodology was the closest one could possibly come to estimating the total amount of parking in London, through this current methodology, however it does assume that CPZ coverage rate is the direct ratio between reported controlled spaces and total spaces.

Streetwise regression model

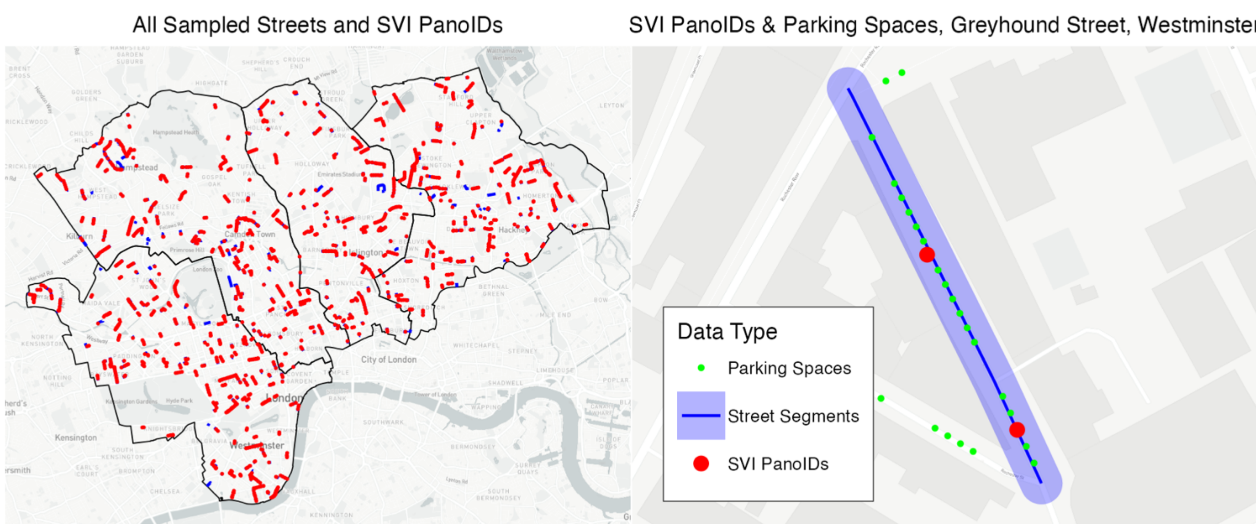

Figure 9. Map of all sampled street segments with respective SVI panoID points (left). Example street and 10-metre buffer with respective parking spaces and SVI panoID points (right).

During this stage of the project, it was necessary to limit the analysis to Camden, Islington, Hackney, Westminster, all inner London boroughs. There are two reasons for this. Firstly, street-to-street parking data was available for these boroughs. Secondly, all four boroughs have comprehensive CPZ coverages (between 98% and 100%) and thus reported controlled parking spaces reflect practically all on-street spaces in the borough. As such, for any randomly sampled street, the risk of missing ground truth data is minimised. To build a dataset for analysis, line segments for all streets in the four boroughs were collected from OpenStreetMaps (OSM) using the osmdata package (version 0.2.5) in R. From this domain, 500 streets were sampled for analysis (see Figure 10 for the distribution of each sampled streets between boroughs). Additional variables for analysis were included in the original OSM data including road length and road type: highway, primary, secondary, tertiary, or residential (see Figure 10).

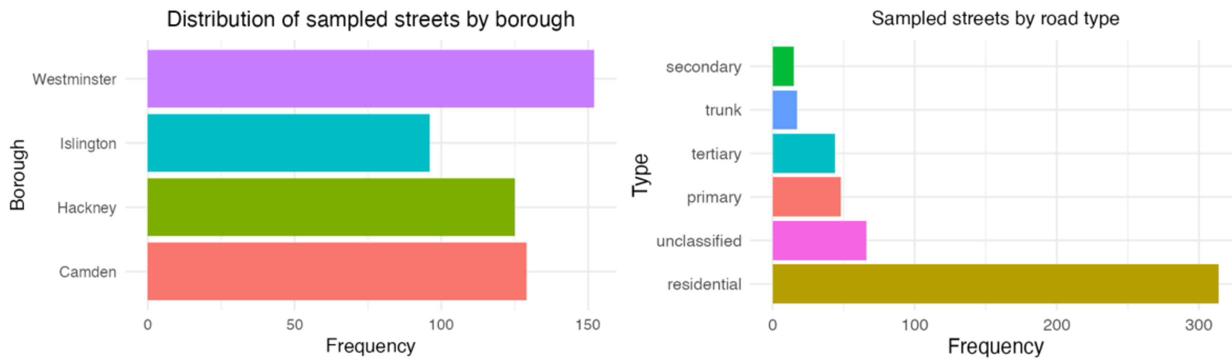

Figure 10. Distribution of sampled streets by borough and road type.

PanoIDs of SVIs were collected from a 10-metre buffer of each street and the final model was deployed to predict counts of moving, stationary on-street, and stationary off-street vehicles in each SVI. The detection results for each class were then summed to produce a total number for each street. Meanwhile, the ground truth total number of parking spaces on each street was collected by summing the total number of spaces within 10-metres of each line segment. Figure 9 shows the resultant points collected from this buffer process on an example street. From these results, OLS regression is conducted with the following equations:

$$\text{Parking Spaces} = \beta_0 + \beta_1(\text{Vehicle Detections}) + \epsilon \tag{M2.1}$$

$$\text{Parking Spaces} = \beta_0 + \beta_1(\text{Vehicle Detections}) + \sum_{i=2}^{4}(\text{Highway Category Effect})_i + \epsilon \tag{M2.2}$$

"Vehicle Detections" refer specifically to stationary on-street vehicle detections. Residential streets are chosen as the baseline highway category.

Before proceeding with regression, the relationship between predictors and response (parking spaces) is investigated. The overall relationship between detected stationary vehicles and reported parking spaces is broadly linear (see figure X), however there is apparent zero-inflation that may potentially negatively influence model performance. Moreover, observations are heavily right skewed and slightly homoscedastic.

Results and Discussion

In this section, I first explore the results of the three CNN models before investigating both the borough-wide and streetwise regressions, comparing the merits of both approaches, and identifying areas of potential further development.

Parking Detection Model Performance

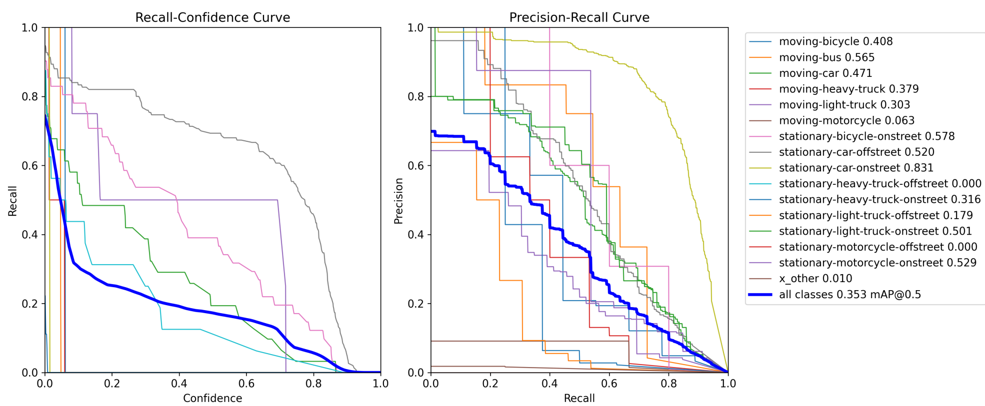

Figure 11. Precision-recall curves for the initial model (left) and enlarged model (right).

Model	Precision	Recall	mAP50	mAP50-95
Initial Model	0.388	0.334	0.371	0.263
Enlarged Model	0.635	0.32	0.353	0.22
Final Model	0.692	0.643	0.685	0.457

Table 4. Primary performance metrics for each of the three YOLO models.

The progression of mAP50 and mAP50-95 values between the initial, enlarged, and final models demonstrates the improvements that were made at each training set. It should be noted that the initial model was validated on a smaller set of data, and thus direct comparison between the overall metrics of the initial model and the enlarged or final model are limited.

	stationary car on-street		*stationary light truck on-street*		*stationary vehicle on-street*	
Model	**AP50**	**AP$_c$50-95**	**AP50**	**AP50-95**	**AP50**	**AP50-95**
Initial Model	0.77	0.549	0.301	0.214		
Enlarged Model	0.832	0.598	0.49	0.366		
Final Model					0.848	0.625

Table 5. Class-specific AP50 and AP50-95 values for selected classes in the three YOLO models.

To better understand the performance of the final model, it is useful to monitor the discrepancy between within-class precision (that excludes background false positives from its calculation) and overall precision, in addition to recall.[6] For the final model, at a confidence level of 0.85 with an NMS threshold of 0.7, stationary vehicles on-street were detected with 94% within-class precision and 92.8% overall precision, indicating few background false positives and thus providing no indication that there is any concern with the quality of the training or validation dataset.

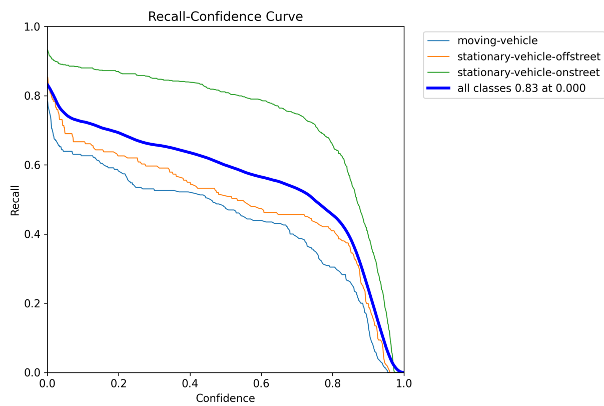

Figure 12. Precision-recall curve for the final model.

At this high confidence level, recall is severely curtailed, with a value of approximately 40% for all classes (see Figure 12) and just 55.5% for stationary on-street vehicles. The low recall statistic provides a large source of doubt for the deployment results and parking count estimates, however, as long as the final model's recall ability is consistent throughout different geographies, streetscapes, and boroughs, the low recall should be compensated for by inflated regression coefficients in the borough-wide and streetwise regressions below.

Class	Within class precision	Overall precision	Recall
Stationary vehicle on-street	0.940	0.928	0.555
Stationary vehicle off-street	0.811	0.811	0.351
Moving vehicle	0.880	0.880	0.257

Table 6. Within-class precision, overall precision, and recall for the final model at 0.85 confidence level.

[6] The within-detections precision only considers detections from other classes misallocated to the class in question as false negatives, i.e., it excludes background FPs. While the conventional methodology for measuring precision that includes background FPs, this within-detections method of calculating precision is useful to isolate the model's ability to distinguish between moving and stationary vehicles from its ability to detect vehicles in the first place.

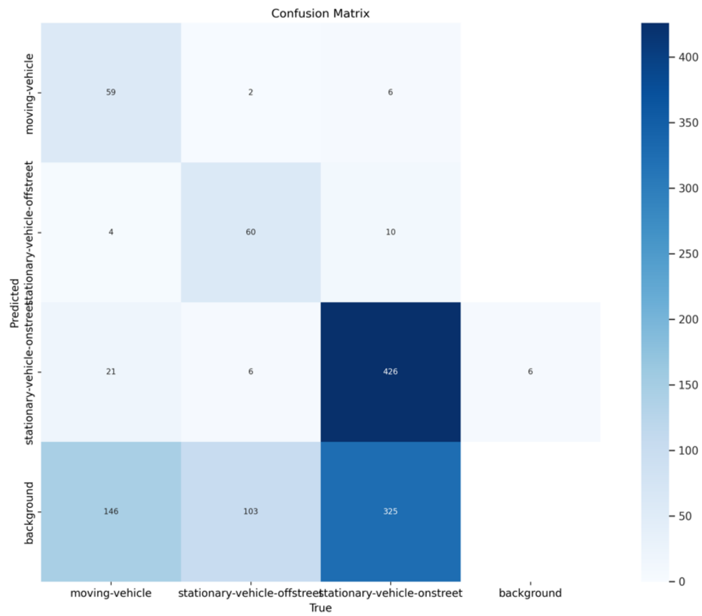

Figure 13. Confusion matrix for final model validation results at 0.85 confidence level. The model exhibits strong inter-class classification abilities but is weakened by poor recall (though this is to be expected at such a high confidence level).

Aggregate Borough Results

Model	Outcome	Predictor Term	R^2	Adj. R^2
M1.4	Spaces / hec.	(Detections * CPZ) / hec.	0.8050	0.7855
M1.2	Spaces	Detections * CPZ	0.6623	0.6285
M1.3	Spaces / hec.	Detections / hec.	0.6484	0.6133
M1.1	Spaces	Detections	0.4184	0.3603

Table 7. R-squared and adjusted R-squared values for the four borough-wide regression models.

Overall, the borough-wide regression yielded strong results. The strategy to adjust parking detections using the CPZ coverage rate yielded stronger results than exclusively relying upon detections. Moreover, normalising detections and reported spaces on an areal basis provided further performance improvements. Overall, the model with the equation M1.4 had the highest performance with an adjusted R-squared value of 0.7855 (Table 7), meaning that only 21.45% of variance was attributable to non-model variables. However, the model should be used cautiously, given its small sample size and relative bias towards Inner London boroughs (75% of the boroughs included in the regression model were in Inner London). With this said, though, there does not seem to be an obvious degree of homoscedasticity or non-normality affecting the residuals, implying the assumptions of OLS hold (Figure 14).

$$\frac{\text{Reported Controlled Spaces}}{\text{Area}} = \beta_0 + \beta_1 \left(\frac{\text{Vehicle Detections } \bullet \text{ CPZ Coverage Rate}}{\text{Area}} \right) + \epsilon$$

(M1.4)

Assessing the coefficients in tablesTable 8 and Table 9, we see that in most cases, the intercept was insignificant, indicating that if no cars are detected, then there are no parking spaces. In M1.2, however, the intercept was significant ($P<0.05$), suggesting that if there were no detected cars, one would still expect parking spaces. This is also a plausible scenario that possibly alludes the role of unoccupied parking spaces.

M1.1: controlled spaces = $\beta_0 + \beta_1$(detections) + ϵ

Predictor	**Estimate**	**Error**	**t value**	**Pr(>|T|)**	
Intercept	16,030	9,841	1.629	0.134	
(detections)	0.5875	0.219	2.682	0.023	**

M1.2: controlled spaces = $\beta_0 + \beta_1$(detections * CPZ coverage) + ϵ

Predictor	**Estimate**	**Error**	**t value**	**Pr(>|T|)**	
Intercept	14,610	6,434	2.271	0.04652	*
(detections * CPZ coverage)	0.8111	0.1831	4.429	0.00128	**

Table 8. Coefficients and their respective confidence terms for the first two borough-wide regression models.

M1.3: controlled spaces per hectare = $\beta_0 + \beta_1$(detections per hectare) + ϵ

Predictor	**Estimate**	**Error**	**t value**	**Pr(>|T|)**	
Intercept	-7.530	5.945	-1.267	0.23399	
(detections per hectare)	1.499	0.349	4.295	0.00157	**

M1.3: controlled spaces per hectare = $\beta_0 + \beta_1$(detections * CPZ coverage per hectare) + ϵ

Predictor	**Estimate**	**Error**	**t value**	**Pr(>|T|)**	
Intercept	2.1160	2.5727	0.822	0.43	
(detections * CPZ coverage per hectare)	1.0811	0.1683	6.425	7.58×10^{-5}	***

Table 9. Coefficients and their respective confidence terms for the second two, area-normalised, borough-wide regression models.

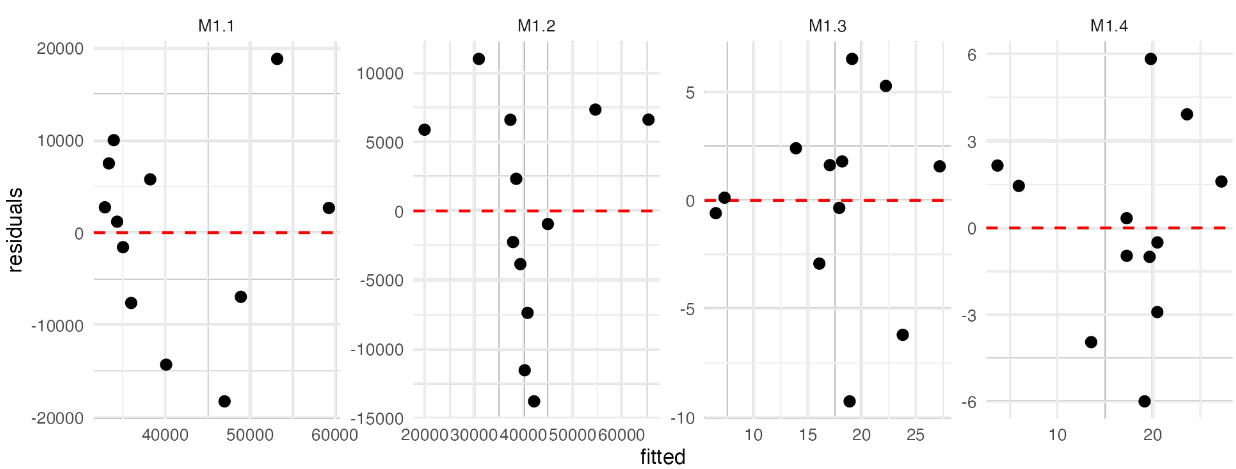

Figure 14. Residuals against fitted values for all 4 borough-wide models.

Using each of the four models, we produce four different estimates of the number of parking spaces in the whole of London excluding the City of London (due to lack of SVIs). The results are below in table X. In order to translate the predictions of controlled spaces back to predictions of actual spaces, the estimate had to be rescaled against CPZ coverage rates, and the results are displayed in Figure 16.

The last estimate of London's parking supply, conducted in 1999, estimated there were 3,152,400 (+/- 241,150) parking spaces in London. Alarmingly, M1.4 completely overestimates London's total parking. The result is so inflated since the model predicts some of the largest parking counts in boroughs with the lowest CPZ coverage rates. Resultantly, these controlled total parking estimates become even more inflated when divided back by CPZ coverage rate to estimate the total amount of parking in London. On the other hand, models M1.1, M1.2, and M1.3 provide promising results that are close to the 1999 estimate. It should be noted that our estimates for the number of controlled spaces cannot be reasonably compared with the 1999 estimate given the large expansion of CPZs in the past 25 years.

Model	**London Controlled Estimate**	**London Total Estimate**
M1.1	1,228,397	3,610,195
M1.2	1,094,028	3,620,585
M1.3	1,035,625	3,209,218
M1.4	1,464,331	6,103,949
MVA Parking Survey 1999	161,000	3,152,400

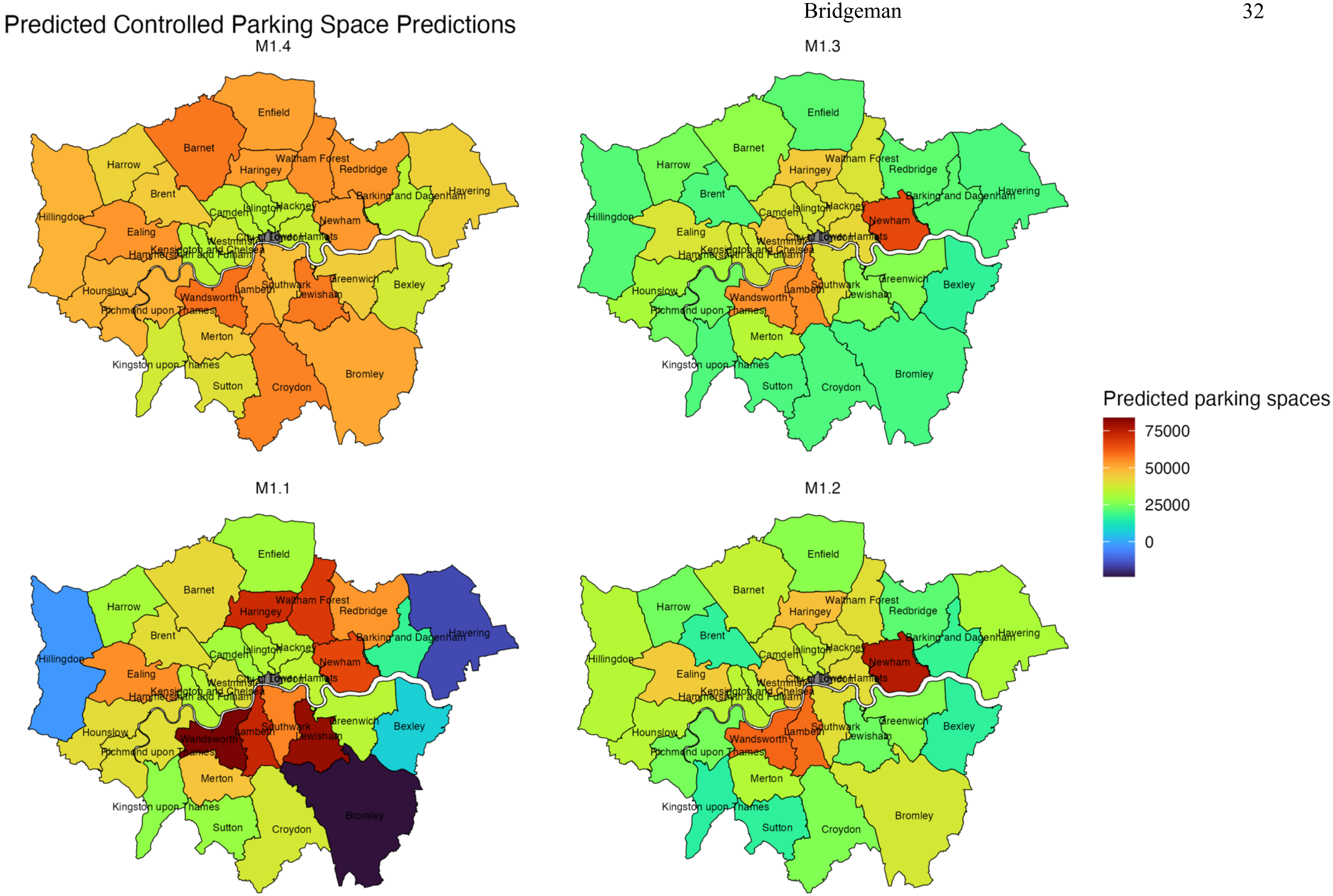

Figure 15. A chloropleth map indicating the estimated number of *controlled* parking spaces across London boroughs according to the four borough-wide models.

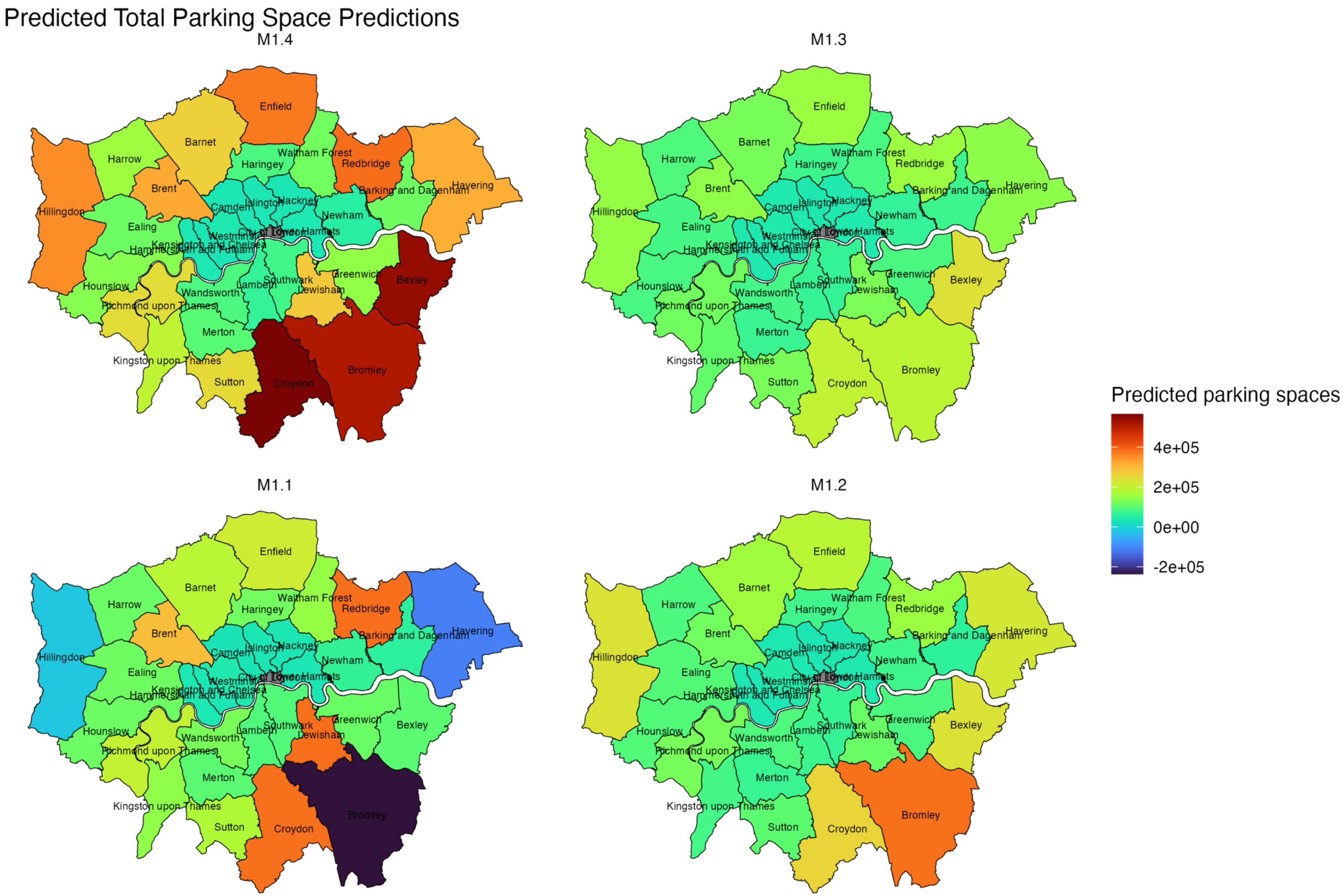

Figure 16. A chloropleth map indicating the estimated number of total parking spaces across London boroughs according to the four borough-wide models.

Streetwise Regression Results

$$\text{Parking Spaces} = \beta_0 + \beta_1(\text{Vehicle Detections}) + \epsilon \tag{M2.1}$$

$$\text{Parking Spaces} = \beta_0 + \beta_1(\text{Vehicle Detections}) + \sum_{i=2}^{4}(\text{Highway Category Effect})_i + \epsilon \tag{M2.2}$$

Model	Outcome	Predictor Term	R^2	Adj. R^2
M1.4	Spaces per street	(Detections * CPZ) / hec.	0.6459	0.6452
M1.2	Spaces per street	Detections * CPZ	0.6732	0.6693

Table 10. R-squared and adjusted R-squared values for the four borough-wide regression models.

M2.1: parking spaces = $\beta_0 + \beta_1$ (detections) + ϵ

Predictor	Estimate	Error	t value	Pr(>|T|)	
Intercept	3.4441	1.0291	3.347	0.000879	***
(detections)	1.0392	0.0342	30.383	$< 2\text{x}10^{-16}$	***

M1.2: controlled spaces = $\beta_0 + \beta_1$ (detections) + (highway category effects) + ϵ

Predictor	Estimate	Error	t value	Pr(>|T|)	
Intercept	7.35885	1.22281	6.018	$3.41\text{x}10^{-9}$	***
(detections)	1.00394	0.03396	29.563	$< 2\text{x}10^{-16}$	***
Highways(residential)	0	-	-	-	
Highway(tertiary)	-8.32413	2.69593	-3.088	0.002129	**
Highway(secondary)	-4.03063	4.42624	-0.911	0.362933	
Highway(primary)	-13.5611	2.58170	-5.253	$2.22\text{x}10^{-7}$	***
Highway(trunk)	-14.62778	4.19602	-3.486	0.000533	***
Highways(unclassified)	-4.07212	2.26328	-1.799	0.072587	.

Table 11. Coefficients and their respective confidence terms for both streetwise regression. Note that residential streets are considered the baseline categorical variable.

For the streets regression, both models yield statistically significant results. While the separation of residuals between highway categories would suggest a better fitted model as demonstrated (see

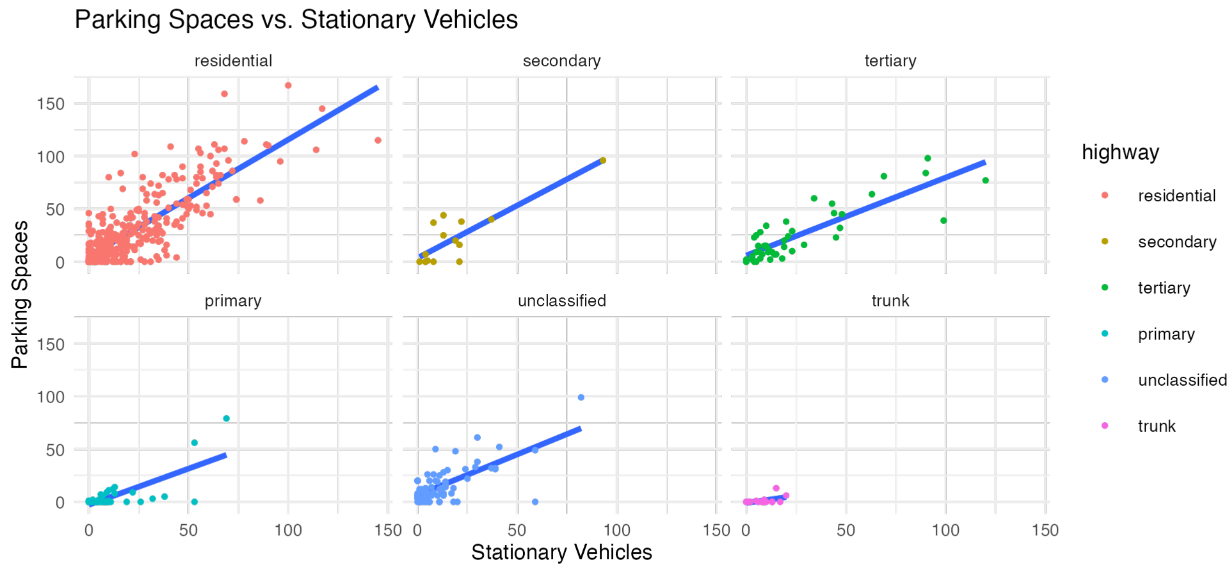

Figure 19), the adjusted R-squared value only grows marginally from 0.6452 to 0.6693. Moreover, the response and predictors possess a homoscedastic relationship that is also seen in the relationship between the fitted values and residuals violating the homoscedasticity assumption of OLS regression. Given this, the streetwise regression methodology warrants further development before generating meaningful results.

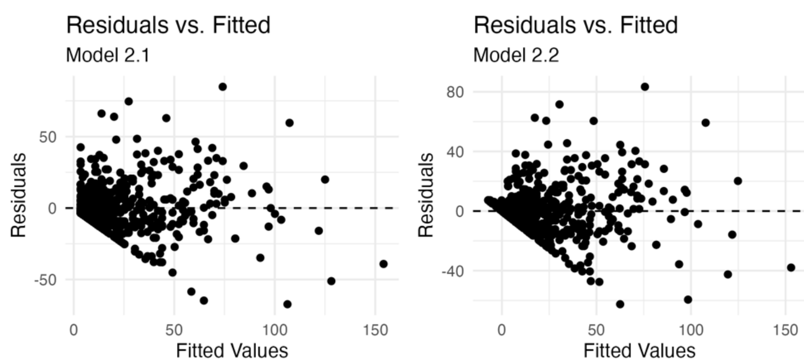

Figure 17. Plots of fitted values against residuals for M2.1 and M2.2 streetwise regression models.

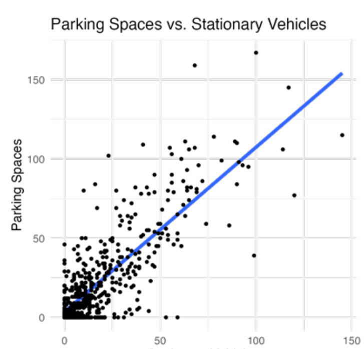

Figure 18. Detected on-street stationary vehicles at 0.85 confidence threshold against reported parking spaces in the sampled street set.Z

While the street-wise methodology is limited, it provides valuable insight into the origin of error. For example, the majority of zero-inflation occurs on primary or trunk streets, other streets do not demonstrate the same degree of zero inflation (see

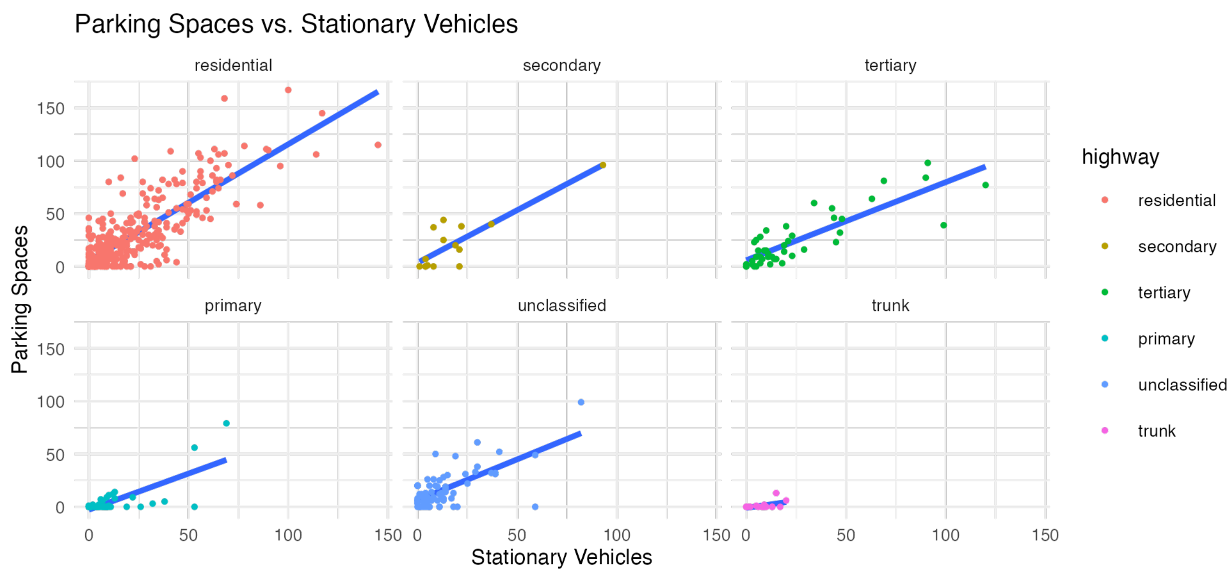

Figure 19), these are likely streets where there is zero permitted parking on the street itself. While the cause of the zero inflation cannot be directly ascertained, it is either inherited from the CNN model or stems from the nature of the streetscape on such roads. Causes for CNN-originating error could include poor classification performance on such roads, high false positive rate, and low precision on such streets. While the model showed generally strong classification abilities at the confidence level of 0.85, model performance by street type has yet to be explored. In any case, poor model performance on primary or trunk roads would not be surprising given the majority of training images that featured on-street parked cars were on quieter residential.

Other sources of error could be factors not considered during the streetwise methodology. For example, many vehicles may stop illegally to load or pick something up on primary roads, and while these cars may correctly be classified by the CNN model as "stationary," they are technically not parked – a distinction this model cannot make a distinction between. Another obvious source of discrepancy may be the fact that vehicles parked on one street may appear in the field of view of SVIs taken on other street, especially at uncrowded intersections, resulting in the overestimation of on-street results. This hypothesis is not supported however by the "detections" coefficient in both M2.1 and M2.2 which is greater than 1, however a combination of factors may be at play.

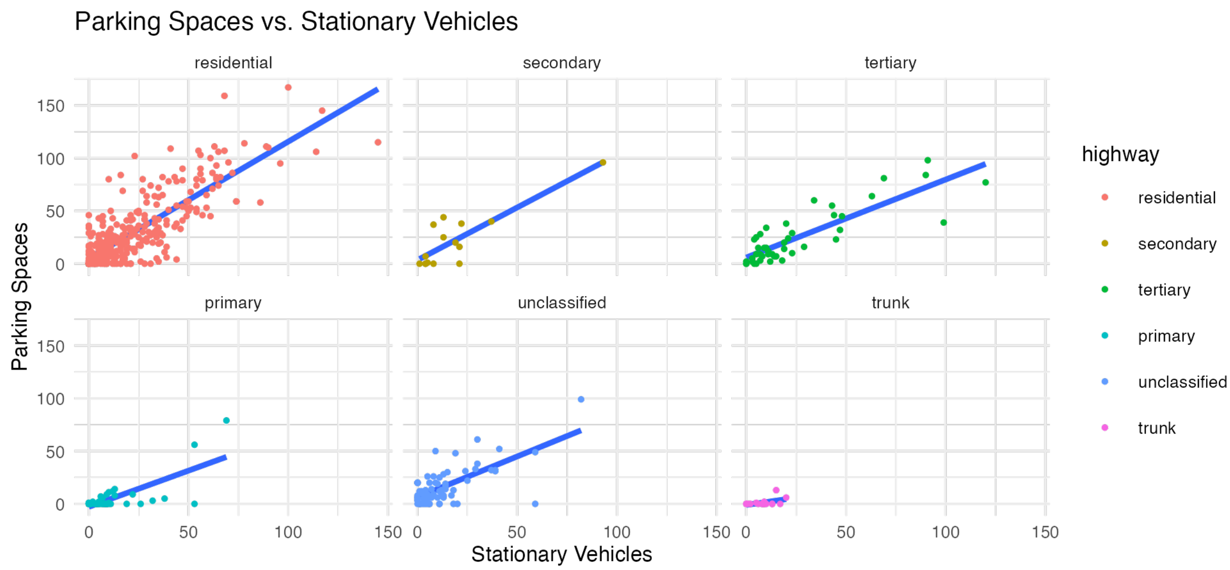

Figure 19. Detected on-street stationary vehicles at 0.85 confidence threshold against reported parking spaces *by highway type* in the sampled street set.

Regression Discussion

The regression analysis at both scales, borough-wide and streetwise has its unique limitations. The borough-wide regression models demonstrated higher adjusted R-squared values, suggesting that aggregating data at the borough level reduced unexplained variance compared to streetwise regression. However, this aggregation likely masked finer-grained variability and introduced overfitting to the Inner London boroughs that dominated the dataset. This bias became evident when applying M2.4 to Outer London boroughs, where lower CPZ coverage rates led to inflated estimates and reduced model generalizability.

Streetwise regression, while providing more detailed insights, exhibited greater variability and lacked the explanatory power of the borough-wide models. This variability highlights the need to account for localized factors such as differences in streetscape characteristics, parking regulations, or vehicle behaviour. For example, the residual patterns across different highway categories indicate that the model performed less effectively on trunk and primary roads, which typically have limited legal parking. Addressing these patterns in future iterations could enhance model accuracy.

Despite its challenges, the streetwise approach offers a promising avenue for more granular analysis of parking dynamics. Further refinements, such as incorporating additional predictors or adjusting for zero-inflation on specific road types, could significantly improve performance. Meanwhile, the borough-wide regression models, though more robust in their current form, require expanded training data to better generalize to suburban and less densely populated areas.

Conclusion

As cities face the twin challenges of urban growth and climate change, the need for precise and efficient land-use analysis has never been more critical. This thesis addresses a key but often overlooked aspect of urban land use—parking inventories—by introducing a novel methodology that harnesses computer vision and machine learning to provide deeper insights into the spatial footprint of parking. By demonstrating the potential of these technologies, this work paves the way for more informed urban planning and policymaking."

This thesis sets the foundation for a larger body of future research. The challenges encountered during the data collection process highlight the need for greater standardization and open-source collation of municipal parking data in London and beyond. The final YOLO11 model proved the feasibility of using computer vision to differentiate between stationary and moving vehicles from still image frames. However, the model's performance could be significantly improved with access to larger and more diverse training datasets. Addressing class imbalance through novel techniques would enable the deployment of the original range of vehicle classes at scale, providing richer insights into inter-vehicle parking dynamics.

The borough-wide scaling methodology presented here offers a promising approach for translating stationary on-street vehicle counts into estimates of reported parking spaces. While effective, future work should focus on developing methodologies that are universally applicable across cities, allowing for meaningful cross-city comparisons. Similarly, the streetwise sampling methodology shows potential as a generalizable technique for CNN-driven parking estimation. In contexts lacking municipal parking data, pairing CNN model outputs with in-person surveys or sampled street analyses could provide calibrated results. Nonetheless, addressing issues such as zero-inflation and heteroscedasticity will require the adoption of advanced statistical methods, such as Zero-Inflated Poisson or Negative Binomial models.

In conclusion, this thesis demonstrates the transformative potential of machine learning and computer vision for urban parking analysis. By advancing methodologies to quantify parking inventories, this work contributes to the broader goal of optimizing urban land use and enabling cities to respond more effectively to the challenges of the twenty-first century.

References

Akbari, Hashem, L. Shea Rose, and Haider Taha. 2003. “Analyzing the Land Cover of an Urban Environment Using High-Resolution Orthophotos.” *Landscape and Urban Planning* 63 (1): 1–14. https://doi.org/10.1016/S0169-2046(02)00165-2.

Barter, Paul A. 2012. “Off-Street Parking Policy Surprises in Asian Cities.” *Cities* 29 (1): 23–31. https://doi.org/10.1016/j.cities.2011.06.007.

British Parking Association. 2016. “Parking Know How: Bay Sizes.” https://www.britishparking.co.uk/write/Documents/Library%202016/Bay_Sizes_-_Jul_2016.pdf.

Chester, Mikhail, Andrew Fraser, Juan Matute, Carolyn Flower, and Ram Pendyala. 2015. “Parking Infrastructure: A Constraint on or Opportunity for Urban Redevelopment? A Study of Los Angeles County Parking Supply and Growth.” *Journal of the American Planning Association* 81 (4): 268–86. https://doi.org/10.1080/01944363.2015.1092879.

Chester, Mikhail, Arpad Horvath, and Samer Madanat. 2010. “Parking Infrastructure: Energy, Emissions, and Automobile Life-Cycle Environmental Accounting.” *Environmental Research Letters* 5 (3): e034001. https://doi.org/10.1088/1748-9326/5/3/034001.

Croeser, Thami, Georgia E. Garrard, Casey Visintin, Holly Kirk, Alessandro Ossola, Casey Furlong, Rebecca Clements, Andrew Butt, Elizabeth Taylor, and Sarah A. Bekessy. 2022. “Finding Space for Nature in Cities: The Considerable Potential of Redundant Car Parking.” *Npj Urban Sustainability* 2 (1): 27. https://doi.org/10.1038/s42949-022-00073-x.

Davis, Amélie Y., Bryan C. Pijanowski, Kimberly Robinson, and Bernard Engel. 2010. “The Environmental and Economic Costs of Sprawling Parking Lots in the United States.” *Land Use Policy* 27 (2): 255–61. https://doi.org/10.1016/j.landusepol.2009.03.002.

Department for Transport. 2024. “Licensed Vehicles by Type and Borough.” Excel Spreadsheet. Survey. London Datastore. https://data.london.gov.uk/dataset/licensed-vehicles-type-0.

Dwyer, B, J Nelson, J Solawetz, and others. 2024. “Roboflow.” https://roboflow.com.

Eaton, Timothy T. 2018. “Approach and Case-Study of Green Infrastructure Screening Analysis for Urban Stormwater Control.” *Journal of Environmental Management* 209 (March):495–504. https://doi.org/10.1016/j.jenvman.2017.12.068.

Furchtlehner, Jürgen, Daniela Lehner, and Lilli Lička. 2022. “Sustainable Streetscapes: Design Approaches and Examples of Viennese Practice.” *Sustainability* 14 (2): 961. https://doi.org/10.3390/su14020961.

Gabbe, C.J., Taner Osman, and Michael Manville. 2021. "The Opportunity Cost of Parking Requirements: Would Silicon Valley Be Richer If Its Parking Requirements Were Lower?" *Journal of Transport and Land Use* 14 (1): 277–301. https://doi.org/10.5198/jtlu.2021.1758.

Guo, Zhan, and Shuai Ren. 2013. "From Minimum to Maximum: Impact of the London Parking Reform on Residential Parking Supply from 2004 to 2010?" *Urban Studies* 50 (6): 1183–1200. https://doi.org/10.1177/0042098012460735.

Guo, Zhan, and Luis Schloeter. 2013. "Street Standards as Parking Policy: Rethinking the Provision of Residential Street Parking in American Suburbs." *Journal of Planning Education and Research* 33 (4): 456–70. https://doi.org/10.1177/0739456X13504299.

Haaland, Christine, and Cecil Konijnendijk van den Bosch. 2015. "Challenges and Strategies for Urban Green-Space Planning in Cities Undergoing Densification: A Review." *Urban Forestry & Urban Greening* 14 (4): 760–71. https://doi.org/10.1016/j.ufug.2015.07.009.

Hoehne, Christopher G., Mikhail V. Chester, Andrew M. Fraser, and David A. King. 2019. "Valley of the Sun-Drenched Parking Space: The Growth, Extent, and Implications of Parking Infrastructure in Phoenix." *Cities* 89 (June):186–98. https://doi.org/10.1016/j.cities.2019.02.007.

Hoehne, Christopher G., Mikhail V. Chester, David J. Sailor, and David A. King. 2022. "Urban Heat Implications from Parking, Roads, and Cars: A Case Study of Metro Phoenix." *Sustainable and Resilient Infrastructure* 7 (4): 272–90. https://doi.org/10.1080/23789689.2020.1773013.

Indolia, Sakshi, Anil Kumar Goswami, S. P. Mishra, and Pooja Asopa. 2018. "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach." *Procedia Computer Science*, International Conference on Computational Intelligence and Data Science, 132 (January):679–88. https://doi.org/10.1016/j.procs.2018.05.069.

Johansson, Fredrik, Jonas Åkerman, Greger Henriksson, and Pelle Envall. 2022. "A Pathway for Parking in Line with the Paris Agreement." *Case Studies on Transport Policy* 10 (2): 1223–33. https://doi.org/10.1016/j.cstp.2022.04.008.

Johnson, Justin M., and Taghi M. Khoshgoftaar. 2019. "Medicare Fraud Detection Using Neural Networks." *Journal of Big Data* 6 (1): 63. https://doi.org/10.1186/s40537-019-0225-0.

Leibling, David. 2014. "Chapter 12: Parking Supply and Urban Impacts." In *Parking: Issues and Policies*, edited by Stephen Ison and Corinne Mulley, 1st ed., 5:259–89. Transport and Sustainability. Bingley, UK: Emerald. https://doi.org/10.1108/S2044-994120140000005013.

Lerner, David N. 1990. "Groundwater Recharge in Urban Areas." *Atmospheric Environment. Part B. Urban Atmosphere* 24 (1): 29–33. https://doi.org/10.1016/0957-1272(90)90006-G.

———. 2002. "Identifying and Quantifying Urban Recharge: A Review." *Hydrogeology Journal* 10 (1): 143–52. https://doi.org/10.1007/s10040-001-0177-1.

Li, Fei, and Zhan Guo. 2014. "Do Parking Standards Matter? Evaluating the London Parking Reform with a Matched-Pair Approach." *Transportation Research Part A: Policy and Practice* 67 (September):352–65. https://doi.org/10.1016/j.tra.2014.08.001.

Li, Rui, Alysha Helmrich, and Mikhail Chester. 2022. "Parking by the Bay: The Supply and Implications of Parking Infrastructure in the San Francisco Bay Area." *Journal of Transport and Land Use* 15 (1): 527–45. https://doi.org/10.5198/jtlu.2022.2123.

Litman, Todd. 2023. "Comprehensive Parking Supply, Cost and Price Analysis." In . Montreal. https://www.vtpi.org/WCTR2023_parking.pdf.

Manville, Michael, and Donald Shoup. 2005. "Parking, People, and Cities." *Journal of Urban Planning and Development* 131 (4): 233–45. https://doi.org/10.1061/(ASCE)0733-9488(2005)131:4(233).

Marshall, Wesley E., and Norman W. Garrick. 2006. "Parking at Mixed-Use Centers in Small Cities." *Transportation Research Record* 1977 (1): 164–71. https://doi.org/10.1177/0361198106197700119.

McCahill, Christopher, and Norman Garrick. 2014. "Chapter 3: Parking Supply and Urban Impacts." In *Parking: Issues and Policies*, edited by Stephen Ison and Corinne Mulley, 1st ed., 5:33–55. Transport and Sustainability. Bingley, UK: Emerald. https://doi.org/10.1108/S2044-994120140000005017.

McCahill, Christopher, Jessica Haerter-Ratchford, Norman Garrick, and Carol Atkinson-Palombo. 2014. "Parking in Urban Centers: Policies, Supplies, and Implications in Six Cities." *Transportation Research Record: Journal of the Transportation Research Board* 2469 (1): 49–56. https://doi.org/10.3141/2469-06.

McDonnell, Simon, Josiah Madar, and Vicki Been. 2011. "Minimum Parking Requirements and Housing Affordability in New York City." *Housing Policy Debate* 21 (1): 45–68. https://doi.org/10.1080/10511482.2011.534386.

Mukhija, Vinit, and Donald Shoup. 2006. "Quantity versus Quality in Off-Street Parking Requirements." *Journal of the American Planning Association* 72 (3): 296–308. https://doi.org/10.1080/01944360608976752.

Müller, Alexandra, Heléne Österlund, Jiri Marsalek, and Maria Viklander. 2020. "The Pollution Conveyed by Urban Runoff: A Review of Sources." *Science of The Total Environment* 709 (March):136125. https://doi.org/10.1016/j.scitotenv.2019.136125.

MVA in association with Data Collection Ltd. 2000. "London Parking Supply Study: Main Report." London: Prepared for Government Office for London. TfL.

Norton, Briony A., Andrew M. Coutts, Stephen J. Livesley, Richard J. Harris, Annie M. Hunter, and Nicholas S. G. Williams. 2015. "Planning for Cooler Cities: A Framework to Prioritise Green Infrastructure to Mitigate High Temperatures in Urban Landscapes." *Landscape and Urban Planning* 134 (February):127–38. https://doi.org/10.1016/j.landurbplan.2014.10.018.

Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439–46. https://doi.org/10.32614/RJ-2018-009.

Pebesma, Edzer, and Roger Bivand. 2023. *Spatial Data Science: With Applications in R*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429459016.

Perkowitz, Sidney. 2021. "The Bias in the Machine: Facial Recognition Technology and Racial Disparities." *MIT Case Studies in Social and Ethical Responsibilities of Computing*, no. Winter 2021 (February). https://doi.org/10.21428/2c646de5.62272586.

Possible. n.d. "Parking Action Map." Accessed December 20, 2024. https://interactive.wearepossible.org/parkingaction/.

QGIS Development Team. 2024. "QGIS Geographic Information System." MacOS 14. QGIS Association. https://www.qgis.org.

Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. "You Only Look Once: Unified, Real-Time Object Detection." In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–88. https://doi.org/10.1109/CVPR.2016.91.

Ríos Flores, Ramiro Alberto, Vera Lucia Vincentini, and Rafael M. Acevedo-Duanas. 2013. "Practical Guidebook: Parking and Travel Demand Management Policies in Latin America." IDB-MG-151. Washington D.C.: Inter-American Development Bank (IDB). https://publications.iadb.org/en/practical-guidebook-parking-and-travel-demand-management-policies-latin-america.

Roboflow. 2024. "Supervision." R. https://github.com/roboflow/supervision.

Rosenzweig, C., W. Solecki, C. Braneon, P. Racco, G. Mollod, D. Bader, M. Phillips, et al. 2018. "The Future We Don't Want: How Climate Change Could Impact the World's Greatest Cities." UCCRN Technical Report. C40 Cities; Global Covenant of Mayors; Acclimatise; Urban Climate Change Research Network. https://www.c40.org/wp-content/uploads/2023/04/1789_Future_We_Dont_Want_Report_1.4_hi-res_120618.original-compressed.pdf.

Russo, Alessio, and Giuseppe T. Cirella. 2018. "Modern Compact Cities: How Much Greenery Do We Need?" *International Journal of Environmental Research and Public Health* 15 (10): 2180. https://doi.org/10.3390/ijerph15102180.

Sarker, Iqbal H. 2021. "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions." *SN Computer Science* 2 (6): 420. https://doi.org/10.1007/s42979-021-00815-1.

Scoggins, M., N. L. McClintock, L. Gosselink, and P. Bryer. 2007. “Occurrence of Polycyclic Aromatic Hydrocarbons below Coal-Tar-Sealed Parking Lots and Effects on Stream Benthic Macroinvertebrate Communities.” *Journal of the North American Benthological Society* 26 (4): 694–707. https://doi.org/10.1899/06-109.1.

Shoup, Donald C. 2005. *The High Cost of Free Parking*. 1st ed. Routledge. https://doi.org/10.4324/9781351179539.

The Traffic Signs Regulations and General Directions 2002. 2002. https://www.legislation.gov.uk/uksi/2016/362/pdfs/uksi_20160362_en.pdf.

Thigpen, Calvin G., and Jamey M. B. Volker. 2017. “Repurposing the Paving: The Case of Surplus Residential Parking in Davis, CA.” *Cities* 70 (October):111–21. https://doi.org/10.1016/j.cities.2017.06.020.

UN-Habitat. 2013. “Streets as Public Spaces and Drivers of Urban Prosperity.” U.N. Doc. HS/060/13E. Nairobi: UN-Habitat. https://unhabitat.org/sites/default/files/2020/08/streets_as_public_spaces_and_drivers_of_urban_prosperity.pdf.

United Nations, Department of Economic and Social Affairs, Population Division. 2019. “World Urbanization Prospects: The 2018 Revision.” U.N. Doc. ST/ESA/SER.A/420. Population Studies. New York: United Nations. https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf.

Vázquez-Suñé, E., X. Sánchez-Vila, and J. Carrera. 2005. “Introductory Review of Specific Factors Influencing Urban Groundwater, an Emerging Branch of Hydrogeology, with Reference to Barcelona, Spain.” *Hydrogeology Journal* 13 (3): 522–33. https://doi.org/10.1007/s10040-004-0360-2.

Wenkel, Simon, Khaled Alhazmi, Tanel Liiv, Saud Alrshoud, and Martin Simon. 2021. “Confidence Score: The Forgotten Dimension of Object Detection Performance Evaluation.” *Sensors* 21 (13): 4350. https://doi.org/10.3390/s21134350.

Westminster City Council in association with Nationwide Data Collection. 2022. “Parking Occupancy Survey 2022.” London: Westminster City Council.

Zhao, Chen, Shi-Jie Gao, Lei Zhou, Xiang Li, Xi Chen, and Chong-Chen Wang. 2019. “Dissolved Organic Matter in Urban Forestland Soil and Its Interactions with Typical Heavy Metals: A Case of Daxing District, Beijing.” *Environmental Science and Pollution Research* 26 (3): 2960–73. https://doi.org/10.1007/s11356-018-3860-7.