# Deep Learning Methods in Transportation and Urban Planning

Advancing data collection and inference methods

Tanay Rastogi

***Dedicated to .......***

# Abstract

Rapid urbanisation and the overwhelming amount of data from connected sensors, mobile devices, and open datasets challenge conventional transport-analysis tools, which rely mostly on rigid mathematical structures, static assumptions, and sparse measurements. This thesis investigates how deep learning (DL) can systematically enhance both data acquisition and analytical inference within three core domains of the Land Use–Transport Interaction (LUTI) cycle—traffic management, population synthesis, and workplace-location choice. Through five research papers, it demonstrates that DL techniques can complement or outperform established methods, yielding richer data, more accurate models, and actionable insights for planners. The research systematically explores novel deep learning applications across three key domains — traffic management, population synthesis, and workplace location choice—demonstrating how each can advance data collection and inference in urban systems. The contribution of this thesis is organized into two main thematic areas: data acquisition and analytical inference.

The **data-acquisition** theme introduces approaches that turn overlooked or incomplete information into high-value inputs for transport models. Paper 1 repurposes street-view video from vehicle-mounted cameras, integrating YOLOv5 object detection, StrongSORT tracking, photogrammetry, and geodesy to automate the construction of link-wide time–space diagrams. Validation on the KITTI benchmark shows mean positional errors below 3 meters, proving that ordinary vehicles can form a low-cost virtual sensor network with far denser spatial coverage than fixed detectors or GPS

floating-car data. Paper 3 tackles missing-attribute surveys by embedding a binary mask into a Wasserstein GAN (WGAN) training routine. The masked WGAN learns directly from incomplete microsamples and generates synthetic populations that match marginal and joint distributions almost as closely as models trained on complete data, thereby reclaiming survey entries that would otherwise be discarded.

The **analytical-inference** theme pairs these novel data sources with DL-enabled or hybrid models to solve long-standing estimation problems. Paper 2 fuses the Cell Transmission Model (CTM) with a two-stage Genetic Algorithm (GA) that jointly calibrates fundamental-diagram parameters and boundary conditions from partial camera trajectories. Across three simulated arterial links, free-flow speed and critical density are recovered within 5 km/h and 10 veh/km, while unobserved densities are estimated with RMSEs near 5 veh/km, demonstrating that sparse vision data can support robust, link-level traffic state estimation. Paper 4 employs a Conditional Tabular GAN (CT-GAN) to generate target-year synthetic populations directly from zonal marginals; a hybrid CT-GAN + Fitness-Based Combinatorial Optimisation pipeline further refines marginal fit. During experimentation, CT-GAN alone achieves perfect Total-Variation and Category-Adherence scores for conditioned attributes, surpassing a baseline IPF-style model in both convergence and fidelity. Paper 5 proposes a custom Deep Neural Network with "zone blocks" that ingest occupation type, socio-economic variables, and multimodal accessibility to predict workplace choice among >1 000 alternatives. Compared with a two-level nested logit model, the DNN attains higher log-likelihoods and reproduces observed distance distributions more faithfully, while requiring no hand-crafted utility specification.

Collectively, the thesis delivers four principal contributions. (1) It establishes vehicle-mounted cameras and incomplete surveys as viable, high-resolution data sources when coupled with DL pipelines. (2) It extends traffic state estimation to settings lacking boundary conditions or full trajectories, using GA-calibrated CTMs. (3) It advances synthetic population generation, offering the first evidence that GANs can satisfy future marginal constraints and learn from sparse, masked data. (4) It scales deep choice modelling to thousand-alternative sets, demonstrating that DNNs can equal or exceed traditional discrete-choice models in both accuracy and behavioural realism.

The thesis also analyses limitations - sensor calibration errors, modelling assumptions, issues with generalisability and scalability of DL models; and proposes remedies for these limitations. Thesis also discusses policy implications of the research and provide a recommendations of the use of DL, with with lessons that are transferable to other sectors like energy, public health, and other spatial systems.

In sum, this thesis advances the state of the art by demonstrating the feasibility and value of integrating modern deep learning techniques into the core of urban transport modelling, pointing the way toward more data-rich, adaptive, and socially responsive transport-planning tools in an increasingly complex world. It demonstrates that the fusion of domain knowledge with scalable DL architectures can unlock the full potential of modern urban data and pave the way towards more efficient, sustainable, and equitable mobility systems.

# Sammanfattning

Snabb urbanisering och den överväldigande mängden data från uppkopplade sensorer, mobila enheter och öppna datamängder utmanar konventionella verktyg för transportanalys, vilka till största delen bygger på rigida matematiska strukturer, statiska antaganden och glesa mätningar. Denna avhandling undersöker hur djupinlärning (DL) systematiskt kan förbättra både datainsamling och analytisk slutledning inom tre centrala domäner i cykeln för Land Use–Transport Interaction (LUTI): affic management, population synthesis, and workplace-location choice. Genom fem forskningsartiklar visas att DL-tekniker kan komplettera eller överträffa etablerade metoder, vilket ger rikare data, mer exakta modeller och användbara insikter för planerare. Forskningen utforskar systematiskt nya tillämpningar av djupinlärning inom dessa tre nyckelområden – trafikledning, populationssyntes och arbetsplatsval – och visar hur var och en kan förbättra datainsamling och analytisk inferens i urbana system. Avhandlingens bidrag är organiserat i två huvudteman: datainsamling och analytisk inferens.

Temat **data-acquisition** introducerar metoder som omvandlar förbisedd eller ofullständig information till högvärdiga insatsdata för transportmodeller. Paper 1 återanvänder videomaterial från fordonsmonterade kameror och integrerar YOLOv5-objektdetektion, StrongSORT-spårning, fotogrammetri och geodesi för att automatisera konstruktionen av länkvisa tids–rumsdiagram. Validering på KITTI-benchmarket visar medelpositioneringsfel under 3 meter, vilket bevisar att vanliga fordon kan fungera som ett kostnadseffektivt virtuellt sensorsystem med betydligt tätare geografisk täckning än fasta detektorer eller GPS–baserade flytande databilar. Paper 3 hanterar undersökningar med

saknade attribut genom att införa en binär mask i en Wasserstein GAN (WGAN)-träningsrutin. Den maskerade WGAN-modellen lär sig direkt från ofullständiga mikrourval och genererar syntetiska populationer som nästan lika väl uppfyller marginal- och samfördelningar som modeller tränade på fullständig data, vilket återvinner enkätposter som annars skulle ha kasserats.

Temat **analytical-inference** kopplar samman dessa nya datakällor med DL-baserade eller hybrida modeller för att lösa långvariga uppskattningsproblem. Paper 2 integrerar Cell Transmission Model (CTM) med en tvåstegs genetic algorithm (GA) som tillsammans kalibrerar parametrar för den fundamentala diagrammet och randvillkor utifrån partiella kamerabaserade trajektorier. På tre simulerade trafiksträckor återfinns frihastighet och kritisk densitet inom 5 km/h respektive 10 fordon/km, medan oupptäckta densiteter uppskattas med RMSE nära 5 fordon/km, vilket visar att glesa visuelladata kan stödja robust uppskattning av trafikläget på länk-nivå. Paper 4 använder en Conditional Tabular GAN (CT-GAN) för att generera syntetiska populationer för ett målår direkt utifrån zonmarginaler; en hybridpipeline med CT-GAN + fitnessbaserad kombinatorisk optimering förbättrar ytterligare passningen mot marginaler. Vid experiment uppvisar CT-GAN ensamt perfekta totalvariations- och kategoritillhörighetspoäng för de villkorade attributen, och överträffar en IPF-liknande basmodell både avseende konvergens och trovärdighet. Paper 5 föreslår ett anpassat djupneuronätverk med "zone blocks" som tar in yrkestyp, socioekonomiska variabler och multimodal tillgänglighet för att förutsäga arbetsplatsval bland fler än 1 000 alternativ. Jämfört med en tvåstegs nestad logitmodell uppnår DNN:et högre log-likelihood och återger observerade avståndsfördelningar mer tillförlitligt, utan behov av handgjord nyttofunktion.

Sammanfattningsvis levererar avhandlingen fyra huvudsakliga bidrag. (1) Den fastslår att fordonsmonterade kameror och ofullständiga enkäter är användbara, högupplösta datakällor när de kombineras med DL-pipelines. (2) Den utökar trafiklägesuppskattning till kontexter där randvillkor eller kompletta trajektorier saknas, genom GA-kalibrerade CTM:er. (3) Den utvecklar generering av syntetiska populationer och presenterar det första beviset på att GAN:er kan tillfredsställa framtida marginalvillkor och lära av gles, maskerad data. (4) Den skalar upp valmodellering med djupinlärning till tusentals alternativ och visar att DNN:er kan mäta sig med eller överträffa traditionella diskreta valmodeller vad gäller både precision och beteenderealisim.

Avhandlingen analyserar även begränsningar – fel vid sensorkalibrering, modellantaganden, problem med generaliserbarhet och skalbarhet för DL-modeller – och föreslår lösningar på dessa begränsningar. Avhandlingen diskuterar också politiska implikationer av forskningen och ger rekommendationer för användning av DL, med lärdomar som är överförbara till andra sektorer såsom energi, folkhälsa och andra rumsliga system.

Sammanfattningsvis för avhandlingen forskningsfronten framåt genom att visa på möjligheten och värdet av att integrera moderna djupinlärningstekniker i kärnan av urban transportmodellering, vilket visar vägen mot mer datarika, adaptiva och socialt responsiva verktyg för transportplanering i en alltmer komplex värld. Den visar att samverkan mellan domänkunskap och skalbara DL-arkitekturer kan frigöra den fulla potentialen i modern urban data och bana väg för mer effektiva, hållbara och rättvisa mobilitetssystem.

# Acknowledgments

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

# List of papers

- **Paper 1:** *Automated Construction of Time-Space Diagrams for Traffic Analysis Using Street-View Video Sequences.*
  Tanay Rastogi and Mårten Björkman.
  Published in ITSC 2023 proceedings. DOI: 10.1109/ITSC57777.2023.10421867

- **Paper 2:** *Model-based traffic state estimation for link traffic using moving cameras.*
  Tanay Rastogi, Michele D. Simone and Anders Karlström.
  Presented in VEHITS 2023 and submitted to journal of ETRR

- **Paper 3:** *Population Synthesis Using Incomplete Microsample.*
  Tanay Rastogi and Daniel Jonsson. Published in Transport Research Procedia. DOI:10.1016/j.trpro.2025.04.011.

- **Paper 4:** *Target-Year synthetic population using Deep Generative Models.*
  Tanay Rastogi and Daniel Jonsson. Submitted to journal of TRR.

- **Paper 5:** *Workplace location choice modelling using DNN.*
  Tanay Rastogi and Anders Karlström.

## Declaration of contribution

Tanay Rastogi contributed to all papers in the following roles: conceptualization, data curation, investigation, methodology, formal analysis, software, visualization, validation, writing - original draft, review and editing.

# Table of content

# 1 Introduction

Rapid urbanisation and the increasing complexity of modern cities have presented significant challenges for transportation and urban planning. According to the United Nations, more people now live in urban areas than in rural ones, with 55% of the world's population residing in cities in 2018 and projections rising to 68% by 2050 (United Nation, 2018). Traditional transportation and urban planning approaches—such as traffic management systems, population synthesis for agent-based transportation models, and discrete choice models—often struggle to keep pace with the dynamic, interconnected nature of urban systems. These methods typically rely on static assumptions, limited data sources, or rigid mathematical structures, which restrict their ability to capture the full variability and intricate dependencies of real-world environments. For instance, conventional traffic management systems frequently face difficulties in accurately predicting congestion and optimising signal timings, largely due to insufficient or outdated data Papageorgiou et al. (2003). Similarly, agent-based models (ABMs) depend heavily on the quality and granularity of synthetic populations, which are often limited by data availability and representativeness. Bastarianto et al. (2023). Discrete choice models further illustrate these constraints, as they may oversimplify the complex decision-making processes of individuals and households, neglecting spatial correlations and joint behaviours that shape urban mobility patterns Amalan et al. (2023).

These problems are compounded by a surge of diverse data—from sensors and mobile devices to social media and live mobility feeds—that traditional methods struggle to integrate. While this explosion of data has the potential to revolutionise urban analysis, it has quickly outpaced the capabilities of conventional analytical tools. As a result, much of this information remains underused, and opportunities for proactive intervention are frequently missed. Persistent challenges related to data quality, privacy, and interoperability further hinder the integration and scalability of models across diverse and rapidly evolving urban contexts Abirami et al. (2024).

To address these interconnected challenges, there is a clear need for innovative, data-driven methodologies that can fully harness the richness of modern urban data, adapt to emerging trends, and provide actionable insights for more efficient, equitable, and sustainable urban environments. Deep learning (DL) methods have emerged as particularly promising in this regard. Recent surveys Yuan et al. (2022); Zhang et al. (2024) highlight a significant rise in DL applications for intelligent transportation system (ITS) tasks, reflecting the growing recognition of their transformative potential. Unlike traditional physical models, which rely on predefined mathematical structures rooted in causality, DL methods excel at learning complex correlations directly from large, diverse datasets. This adaptive learning capability enables DL models to make more accurate predictions and support dynamic, real-time decision-making—capabilities that are often unattainable with hand-engineered, rule-based approaches. Rather than depending on rigid mathematical formulations, DL models can flexibly incorporate domain knowledge through data-driven constraints or by leveraging auxiliary information sources, thereby enhancing their flexibility and scalability.

This thesis critically examines the advancements and practical applications of deep learning in transportation and urban planning. In particular, we focus on how these approaches can advance data collection and inference, offering solutions that go beyond the limitations of traditional methods and paving the way for the next generation of data-driven urban management.

## 1.1 Aim

The aim of this research is to:

> *investigate how DL methods can be systematically applied to enhance data collection and inference in transportation and urban planning, thereby addressing the limitations of traditional analytical models.*

The thesis advance understanding of the applicability of DL methods through systematic exploration of their potential to enhance, complement, or outperform established analytical approaches. Rather than developing entirely new models, it investigates the intersection of existing DL techniques—some previously untested in transportation contexts— to address critical social challenges in transport and urban planning, described above. As shown in Figure 1, this approach emphasizes the strategic intersection of deep learning and domain-specific applications, with the goal of identifying which DL models are best suited for different types of transportation challenges.

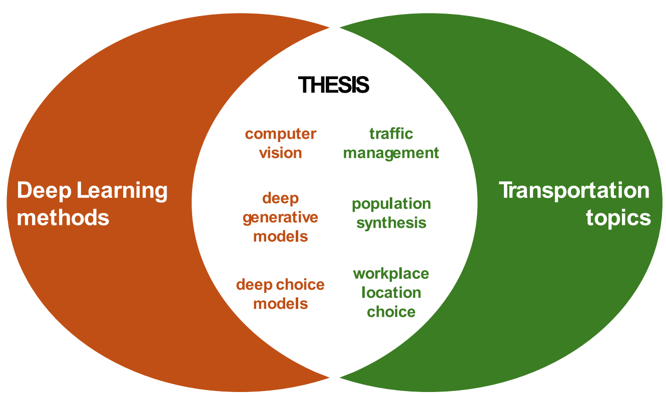



**Figure 1: Illustration highlighting the intersection of the DL models and transportation topics that are part of this thesis.**

As illustrated in the Figure 2, this aim encompasses two complementary aspects: 1). improving the quality and utility of data used in transport modelling by applying DL techniques to extract, generate, or repurpose data from both novel sources and re-purposing data that is often overlooked or discarded; and 2). developing adaptive inference methods that can learn from such enhanced data to support more accurate, dynamic, and context-aware decision-making. By integrating these objectives, the research not only proposes methodological innovations but also ensures their practical relevance to contemporary issues in traffic management, agent-based modelling, and spatial choice behaviour. This dual focus on improving data sources and creating innovative analytical methods is essential for ensuring that the integration of new data leads to meaningful solutions within the transportation field.

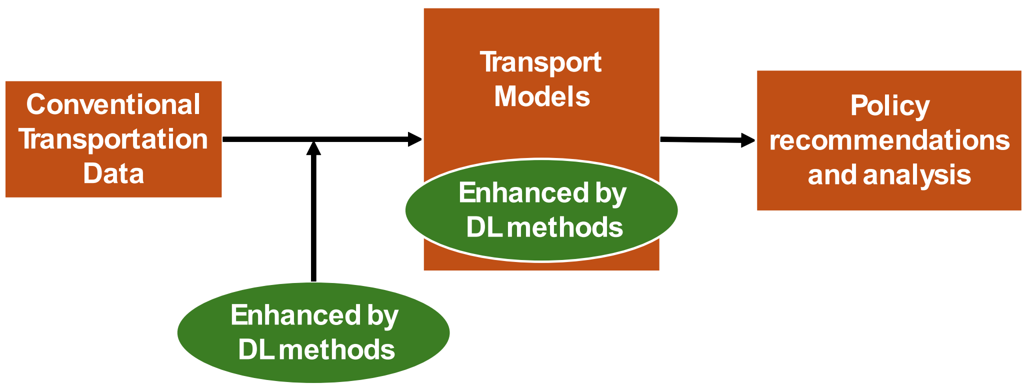



**Figure 2: Illustration highlighting the thesis's approach in using the DL-methods to enhance the transportation modelling process.**

## 1.2 Specific Research Objectives

This thesis explores the development and application of DL methods in transportation and urban planning in five interconnected papers. Throughout these papers, the research focuses on either complementing or surpassing traditional methods in the field by employing innovative applications of existing deep learning techniques. The contributions span three distinct transportation domains and are organised into two main thematic areas: **data acquisition** and **analytical inference**. Table 1 presents an overview of how each paper in the thesis is positioned within the respective domains and thematic areas.

The first theme, data acquisition, focuses on DL methods that improve how we collect data by creating new sources or making better use of existing ones, often by using information that is usually ignored or thrown away. These methods include innovations in sensor technologies as well as advanced data imputation techniques. The second theme, analytical inference, is dedicated to developing and applying advanced analytical techniques to extract insights from transportation data. This approach utilises state-of-the-art methods in DL and statistical modelling to process and interpret large-scale transportation data effectively for planning purposes.

**Table 1: A cross-sectional overview of the thesis highlighting its research coverage across three distinct transportation and urban planning domains, with contributions organized under two main thematic areas.**

| | **Themes** | |
|---|---|---|
| **Topics** | **Data Acquisition** | **Analytical Inference** |
| Traffic management | Paper 1 | Paper 2 |
| Population synthesis | Paper 3 | Paper 4 |
| Workplace location choice | | Paper 5 |

## Data Acquisition

This thematic area covers two topics in transportation and urban planning: traffic management and population synthesis. Under the topic of traffic management, the research explores the use of novel data sources (**Paper 1**), while under the topic of population synthesis, it seeks to enhance the quality of data used in transportation modelling (**Paper 3**).

In modern traffic management applications, recent advances in computer vision have positioned vision-based systems as a prominent data source (Dilek and Dener, 2023). These systems provide high-resolution spatial and temporal data at a fraction of the cost of traditional sensing technologies, thus enabling a broader range of deployment scenarios. The growing accessibility of lightweight, portable sensors—such as smartphones, dashboard cameras, and embedded vehicle devices—further facilitates continuous traffic monitoring. Aligned with this trend, **Paper 1** investigates deep learning techniques for extracting meaningful traffic information from street-view video sequences captured by vehicle-mounted cameras. This approach transforms ordinary vehicles into mobile traffic sensors, expanding data collection capabilities without dependence on fixed infrastructure.

In transportation modelling, agent-based models (ABMs) are widely used to simulate travel behaviour, forecast demand, and evaluate policy scenarios (explained further in Section 2.3). These models depend on richly detailed population data describing individuals' socio-demographic and household attributes across various geographic scales Bastarianto et al. (2023). Typically, synthetic populations are generated using micro-samples and aggregated data. However, conventional population synthesis methods assume access to complete micro-samples from a single data source—an assumption often violated in practice. Data collection errors, privacy-related omissions, and difficulties in merging multiple datasets frequently result in missing attribute information, undermining the effectiveness of traditional synthesis techniques. **Paper 3** addresses this challenge by developing a robust DL-based methodology for population synthesis capable of handling incomplete micro-sample data. The objective is to construct a framework that can effectively handle incomplete data while maintaining accuracy comparable to models trained on complete samples.

### Analytical Inference

This thematic area covers three topics in transportation and urban planning: traffic management (**Paper 2**), population synthesis (**Paper 4**) and workplace location choice (**Paper 5**). Under the topic of traffic management (explained further in Section 2.2), the thesis focuses on creating novel methodologies that can effectively utilise unique street-view data. Under the topics of population synthesis and workplace location choice, the thesis focuses on developing analytical methods using DL that improve on established approaches.

To implement intelligent traffic management strategies and operations, it is important to measure traffic conditions on the road network accurately. However, no traffic sensor can continuously collect data over an entire road network for extended periods of time. To address this limitation, traffic states in unobserved regions could be inferred using partially observed data through Traffic State Estimation (TSE) methods. However, traditional TSE methods cannot use the partial space-time trajectory data collected from on-board vehicle cameras. Hence, this thesis presents innovative approaches for using street-view traffic data collected from on-board vehicle cameras to estimate traffic states for road links in urban networks, improving our understanding of traffic conditions without relying on fixed sensor infrastructure. Specifically, **Paper 2** develops an innovative approach for estimating traffic states for unobserved regions of space-time diagrams for a road link, using the data collected from the on-board vehicle cameras in **Paper 1**. The main objective of **Paper 2** is to develop methods that effectively utilise these novel data sources and demonstrate their practical applications in real-world transportation systems, moving beyond theoretical advances to provide actionable insights.

**Paper 4** contributes to improving ABMs, which are employed for scenario planning, where urban planners develop plausible future scenarios based on current trends and uncertainties. They help decision-makers anticipate challenges and design adaptable strategies. To meet ABMs' detailed data needs for regional populations, population synthesis generates synthetic populations for both current and future scenarios. In this domain, deep generative models (DGMs) like variational autoencoders (Borysov et al., 2019) and generative adversarial networks (Garrido et al., 2020; Kim and Bansal, 2023) have gained traction. Building upon this previous research, **Paper 4** aims to provide advancement in target-population synthesis with deep generative modelling, demonstrating how DL-based methods can be adapted for this problem. Research also analyses the advantages and disadvantages of combining DL-based methods with conventional methods such that the hybrid model can overcome the problems that conventional models face.

**Paper 5** addresses a critical component of the Integrated Land-Use and Transportation Models (ILUTM) is the workplace location choice model, which establishes a connection between job location decisions in the land-use model and travel behaviours in the transportation model (explained further in Section 2.1). Traditionally, researchers have used discrete choice models, which are often based on logit models, to study where people choose to work, giving important information about how these decisions are made in different locations. However, DCMs encounter difficulties in accurately capturing individuals' decision-making behaviours, especially when faced with complex patterns involving numerous alternatives (Torres et al., 2011; Ben-Akiva et al., 2002). The research aims to address these limitations by developing and evaluating a deep neural network (DNN) approach to modelling workplace location choices. Specifically, **Paper 5** aims to explore whether DNNs can serve as a robust alternative to DCMs by leveraging their ability to learn flexible representations directly from data, without relying on pre-established theoretical assumptions. This research investigates how custom deep learning models can enhance workplace location choice predictions and improve our understanding of commuting patterns and their implications for transportation demand.

## 1.3 Thesis Structure

The remainder of this thesis is structured as follows. Chapter 2 provides essential background on transportation modelling, offering a comprehensive overview of conventional methodologies within each transport topic covered in this thesis while identifying associated research gaps. Subsequently, Chapter 3 presents the various DL-based methods employed throughout this research, demonstrating how each method addresses specific transport topics within the thesis framework. Chapter 4 then synthesises the research contributions, positioning each paper within the context of the thesis objectives while critically examining the research limitations and proposing future directions for each study. Finally, Chapter 5 presents a summary of the thesis and delivers broader conclusions regarding this research endeavour.

# 2 Background and Research gaps

In this chapter, we begin by introducing urban transportation models, specifically the Land Use Transport Interaction (LUTI) models. These models encompass various aspects, including infrastructure, vehicles, and traffic flow, and aim to understand and predict travel behaviour and demand. Additionally, the chapter outlines how each of the topics, namely - traffic management, population synthesis and workplace location choice, in this thesis are connected to steps within this LUTI framework, providing a brief background on conventional approaches currently in use for each of the topics, discussing their methodologies and highlighting the research gaps associated with these approaches.

## 2.1 Land Use Transport Interaction

Land Use and Transport Interaction (LUTI) models are a fundamental approach for understanding the complex, bidirectional relationship between land use patterns and transportation systems in urban environments. They operate on the theoretical premise that land use determines travel demand, while transport supply influences the location of activities. Consequently, these models recognise that the distribution of housing, jobs, and services influences travel behaviour, and conversely, the accessibility provided by the transport network shapes land development.

The conceptual foundations of LUTI modelling trace back to the 1950s and early 1960s. A seminal contribution was made by Lowry (1964) with "A

Model of Metropolis," widely considered the first operational LUTI model. Developed for the Pittsburgh region, Lowry's model established a foundational framework linking a basic employment sector, a population-serving retail sector, and a residential sector, thereby providing a comprehensive method for understanding metropolitan-scale interactions. Building upon this pioneering work, the 1970s and 1980s saw significant refinements. The Garin-Lowry model (Garin, 1966) introduced more sophisticated spatial interaction mechanisms, while the "entropy maximisation" framework by Wilson (1970) offered a more robust theoretical basis for these interactions. The advent of greater computing power and data availability from the early 1990s onwards boosted further advancements, leading to the development of disaggregate, activity-based models introduced by figures such as Wegener (2004). This shift allowed for a more granular and behaviourally realistic representation of individual and household choices.

In the early 21st century, the evolution of LUTI models continued, marked by two major methodological shifts: the move from aggregated to disaggregated data and the transition from static to dynamic modelling (Acheampong and Silva, 2015). The first shift was largely enabled by Random Utility Theory, introduced by McFadden (1973), which facilitated disaggregate choice models that account for individual preferences and socioeconomic characteristics. The second shift towards dynamic modelling has been driven by computational innovations and the availability of extensive datasets. Modern LUTI models now frequently incorporate microsimulation and agent-based techniques to capture detailed household behaviours and the sequential decision processes linking short-term travel choices with long-term land-use changes (Ardeshiri and Vij, 2019). Furthermore, they often include time-dependent accessibility metrics that evolve with infrastructure investments (Fransen et al., 2015) and account for behavioural adaptations, for example by using prospect theory to model commuters' departure-time choices (Geng et al., 2023).

Despite these sophisticated advancements, contemporary LUTI models still face significant limitations (Amalan et al., 2023; Parishwad et al., 2023). A primary challenge is their dependence on hand-engineered features, such as predefined accessibility metrics and utility functions, which can introduce modeller bias and overlook non-linear interactions. Data-related issues are also prevalent, including gaps and quality problems that can propagate errors and distort policy insights. Furthermore, these models struggle with capturing the full complexity of dynamic behaviours, face substantial computational burdens, especially for microsimulations, and often exhibit poor transferability between different urban contexts. The difficulty in integrating real-time data streams for dynamic updates further compounds these problems. These challenges stem from structural rigidities in model design, data limitations, and scalability constraints, which collectively hinder effective and adaptive urban planning.

Deep learning (DL) offers promising solutions to many of these persistent challenges. By automating the extraction of relevant features from raw, high-dimensional data, DL models can reduce the reliance on hand-engineered inputs and minimise modeller bias. For instance, techniques like autoencoders (Kempinska and Murcio, 2019) and graph neural networks (GNNs) (Liu and Meidani, 2024) can learn latent spatial and structural relationships directly from data, replacing predefined accessibility metrics with richer, data-driven representations of urban form. This capability also enhances model transferability; a DL model trained in one city can often be fine-tuned for another with limited local data, significantly reducing calibration efforts (Mo and Gong, 2023). To address data sparsity and quality issues, generative models such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) can synthesise plausible population microdata, correcting for gaps and improving representativeness (Borysov et al., 2019; Garrido et al., 2020). Finally, to overcome the limitations of static models, online learning architectures can continually ingest real-time data streams, allowing models to adapt dynamically to emergent travel patterns and disruptions (Barry et al., 2023).

In summary, integrating these deep learning capabilities into LUTI frameworks can significantly enhance model robustness, scalability, and responsiveness, paving the way for a new generation of more efficient and realistic urban system models. This thesis contributes to this advancement through five research papers, each addressing a distinct step within the LUTI process. The papers are organised across three domains—traffic management, population synthesis, and workplace location choice—which correspond to key stages within the LUTI feedback cycle (see Figure 3). By targeting critical methodological gaps in these areas, this research contributes to both data and methodological advancements within the cycle. Collectively, these studies advance the overarching goal of developing more robust and data-driven approaches for transportation and urban planning.

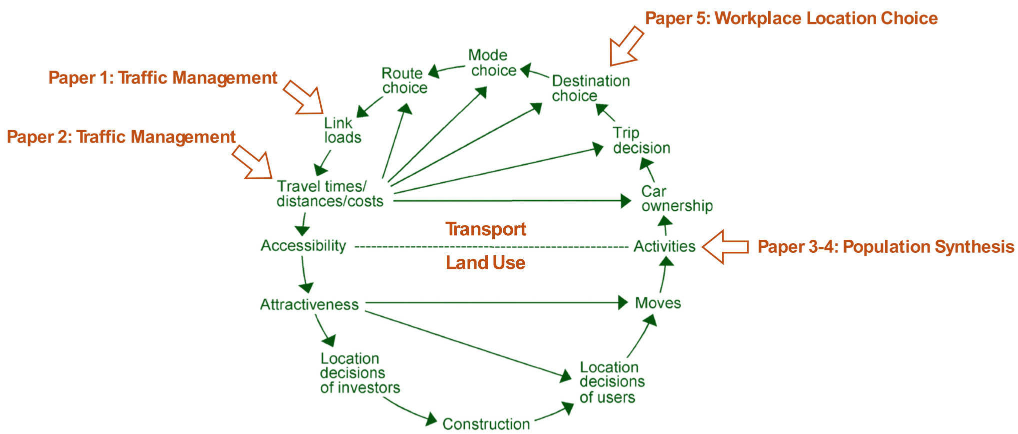



**Figure 3: Conceptual framework linking the research topics to the LUTI model by Wegener and Fuerst (2004), with the topics summarized in Table 1**

## 2.2 Traffic Management

Traffic management represents a critical component within the LUTI feedback cycle, directly influencing link loads and travel times as illustrated in Figure 3. Precise, real-time traffic measurements and reliable traffic state prediction across network links is essential for understanding accessibility impacts on land use and for calibrating LUTI models' transport modules.

The thesis mostly focused on gathering traffic data in the form of "time-space diagrams", which are a graphical representations of the trajectories of all vehicles that move through a specific road link over time. They provide insights into microscopic traffic characteristics such as time headways and space headways as well as macroscopic traffic characteristics such as density, flow, and mean speed Treiber and Kesting (2013), as illustrated as black lines in Figure 4.

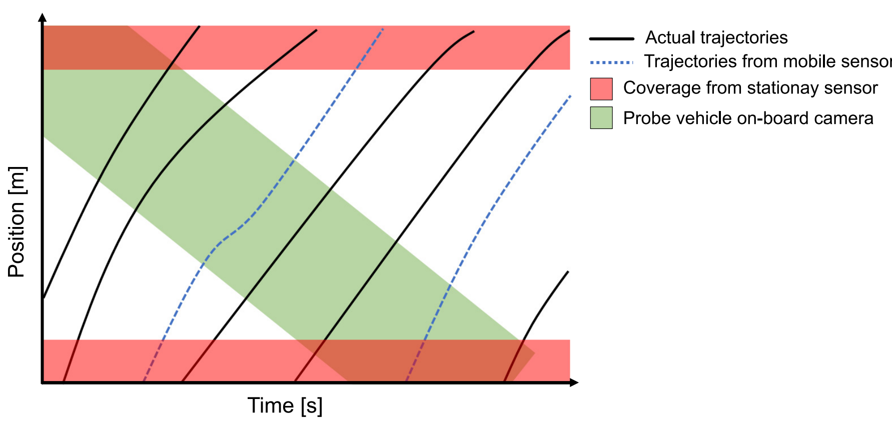



**Figure 4: Example of time-space diagram highlighting different sections of the diagram measured by various types of sensors.**

Traffic data collection methods have evolved significantly to address spatial coverage limitations in traffic management systems. Conventional approaches rely on stationary roadside infrastructure such as inductive loops, infrared detectors, and fixed cameras to measure traffic volume, occupancy, and speed at specific points (Sun and Ritchie, 1999; Atluri et al., 2009; Ua-areemitr et al., 2019). However, these fixed sensors, represented as "RED" bands in Figure 4, provide only localized measurements that prove inadequate for comprehensive network analysis, particularly on arterial roads with signal-controlled traffic flow. To overcome these limitations, mobile sensing through Floating Car Data (FCD) from GPS-equipped probe vehicles has emerged as an alternative approach, enabling traffic data collection across expanded spatial and temporal ranges. These systems, illustrated as "BLUE" trajectories in Figure 4, collect traffic state information along entire links and have been extensively used to estimate critical parameters including flow, speed, and density (Herrera and Bayen, 2010; Sunderrajan et al., 2016; Yuan et al., 2021). However, these sensors generate spatially sparse data distributions and require prior knowledge of probe vehicle penetration rates, which Rahmani et al. (2010) identified as difficult to obtain yet essential for meaningful traffic analysis. More recently, lightweight portable cameras such as smartphones, dash-cams, and on-board cameras have gained prominence for their ability to provide even broader spatial resolution through continuous interaction with the surrounding environment. Researchers including Guerrieri et al. (2019); Kumar et al. (2021a); Cao et al. (2011) have utilized this technology for traffic data collection and flow estimation. The "GREEN" band in Figure 4 demonstrates that camera-based data collection methods offer significantly wider sensor coverage than both fixed sensors and traditional mobile sensing approaches.

Any effective traffic management strategy relies on the accurate and real-time measurement of traffic conditions on the entire road network. However, despite these advances, no single sensor suite provides continuous, link-wide observations. As illustrated in Figure 4, a considerable portion of the time-space diagram remains unobserved (blank "WHITE" space) from any type of sensors. Traffic State Estimation (TSE) methods bridge the gaps by fusing partial data with traffic-flow models to reconstruct flow, density, and speed over the entire time–space domain (Seo et al., 2017). However, conventional TSE methods that uses input data from stationary or FCD data leverages Kalman-filter variants with Cell Transmission Models (CTM), assuming complete vehicle trajectories, calibrated fundamental diagrams and known boundary conditions on the link for the estimation process (Munoz et al., 2004; Seo et al., 2015; Tampere and Immers, 2007). In contrast, when vehicle trajectory data is collected using on-board cameras, the trajectories are often incomplete, and researchers typically lack access to either the boundary conditions or initial traffic states required for TSE.

### Research Gaps

As highlighted above, mobile camera-based data collection methods are significantly superior to stationary and GPS-enabled FCD sensors due to their wider sensor coverage. These cameras not only offer much higher spatial resolution compared to stationary sensors but also eliminate the need for prior knowledge of penetration rates for traffic analysis. However, research using onboard cameras as primary data sources remains limited.

Moreover, there is a notable scarcity of TSE techniques capable of leveraging the data collected from these mobile cameras. Typically, such data lack boundary conditions, initial traffic states, or ground truth at regular time intervals—elements that are usually essential for conventional estimation procedures. These limitations underscore the need for novel approaches that maximise the utility of incomplete visual trajectory data while accommodating the practical constraints of real-world traffic monitoring systems.

In this thesis, **Paper 1** presents advancements in traffic data collection by generating space-time diagrams to capture traffic flow characteristics using an in-vehicle camera mounted on a moving vehicle. The proposed method employs state-of-the-art computer vision algorithms based on deep neural networks (DNNs), photogrammetry, and geodesy to construct time-space diagrams from video sequences. Subsequently, **Paper 2** introduces a methodological advancement in TSE using vehicle trajectories obtained from cameras mounted on vehicles travelling in the opposite lane. This research proposes an innovative estimation method that combines the Cell Transmission Model (CTM) with a Genetic Algorithm (GA) to calibrate fundamental diagram parameters and boundary conditions.

## 2.3 Population Synthesis

Traditional LUTI models have historically operated at an aggregate level, connecting land use changes to transport flows through factors like accessibility and travel expenses. These models, often employing a sequential four-stage approach, struggle to capture the nuances of individual decision-making. In recent decades, agent-based models (ABMs) have emerged as computational frameworks in transportation and urban planning, simulating the actions and interactions of autonomous agents—such as individuals, households, firms, or vehicles—within defined environments to capture urban mobility complexity Bastarianto et al. (2023). This capability enables detailed analysis of daily activity patterns and their impacts on urban systems. Prominent examples include UrbanSim (Batty, 2012), TRANSIMS (Smith et al., 1995), MATSim (Horni et al., 2016), and Scaper (Västberg et al., 2020), all designed to replicate social systems and individual travel behaviours.

A core component of any of the ABM transport models is ABM is the need for detailed information about agents. These ABS usually need detailed information about populations' social and household characteristics from an area covering cities, towns, or even countries as input data. Additionally, for performing scenario or policy analysis using these ABM, requires the detailed population information both base-scenario as well as all the possible target scenarios. However, acquiring such granular, individual-level data is challenging. Issues such as privacy concerns, as well as the technical and financial constraints of data gathering, often impede accessibility to detailed data. Hence, "population synthesis" plays crucial role in ABM framework to provide these models with a reliable alternative to actual populations. Population synthesis techniques generate a comprehensive list of a simulated population, each accompanied by corresponding attribute data. The goal of population synthesis is to find the best way to use different data sources to create agents in social and geographical spaces that are very close to the underlying population structure and meet certain scenario criteria set by the user, such as the correlation structure and control totals (Axhausen et al., 2010; Ma and Srinivasan, 2016).

According to Sun et al. (2018); Fabrice Yaméogo et al. (2020) the conventional methods for population synthesis fall into two main categories: synthetic reconstruction and combinatorial optimization. Synthetic reconstruction approaches, including Iterative Proportional Fitting (IPF) and Iterative Proportional Update, integrate sample data and aggregate statistics to calculate weights representing each sample agent's significance in specific zones. Several researchers like, Rich (2018); Prédhumeau and Manley (2023); Beckman et al. (1996); Zhu and Ferreira (2014) have employed these methods for generating both baseline and target populations. The Combination Optimization (CO) algorithm, like IPF, iteratively replaces households with a new set of individuals and households until a goodness-of-fit indicator converges toward specified stopping criteria. One key advantage of CO methods is that their data requirements remain less restrictive than those for IPF methods. Additionally, CO directly generates a list of households that match multiple multilevel controls without needing to create a joint multi-way distribution. Williamson et al. (1998); Ryan et al. (2009); Ma and Srinivasan (2015) have used variation of CO-based FBS approach for generating synthetic population that can efficiently handle multilevel controls, demonstrating its feasibility and improved performance compared to traditional IPF methods.

## Research Gaps

There are several significant limitations in conventional population synthesis approaches. First, fitting large numbers of individual attributes quickly becomes computationally and memory-intensive for both IPF and CO algorithms. As highlighted by Choupani and Mamdoohi (2016), IPF methods are particularly challenged by high-dimensional data, often encountering zero-cell problems that lead to convergence issues and division-by-zero errors. Second, because these methods are deterministic, they cannot generate synthetic agents that do not exist in the original data, which limits population heterogeneity and diversity.

These challenges have motivated researchers to explore simulation-based probabilistic models using statistical learning, especially for synthesising base populations. Several studies Farooq et al. (2013); Sun and Erath (2015); Borysov et al. (2019); Garrido et al. (2020) have employed various statistical learning techniques, including Markov Chain Monte Carlo (MCMC) based on Gibbs sampling, Bayesian networks, and deep generative models such as Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN).

However, important gaps remain that limit the wider adoption of these newer methods. Most simulation-based probabilistic models have only been tested on complete micro-samples from single data sources, whereas real-world micro-samples are often sparse and incomplete, failing to capture the full range of attribute variations present in actual populations. Another major limitation is that these methods often do not satisfy the conditional distributions required for target populations, as most experiments focus solely on generating base populations without considering alignment with future targets. These issues demonstrate the importance of more robust approaches that can handle incomplete data while maintaining population diversity and meeting target conditional distributions.

In this thesis, **Paper 3** presents an advancement in data quality improvement by generating synthetic populations from incomplete micro-samples. The research proposes a novel technique for training a Wasserstein GAN (WGAN) model that enables effective learning from incomplete training data. Building on these findings, **Paper 4** introduces a methodological advancement using Conditional Tabular GAN (CT-GAN) to synthesise target populations, either directly from marginal constraints or through a hybrid approach that combines CT-GAN with Fitness-based Synthesis Combinatorial Optimisation (FBS-CO). These methods demonstrate that deep generative approaches can produce synthetic populations closely resembling actual data even when trained on incomplete samples, offering promising solutions for realistic and scalable population synthesis in transport modelling for both baseline and future scenarios.

## 2.4 Workplace Location Choice

Workplace location choice models are integral to LUTI frameworks, serving as connectors between transportation systems and urban spatial dynamics. These models analyse how individuals select employment destinations based on accessibility, demographic factors, and transportation infrastructure, thereby linking residential patterns, commuting behaviour, and land use changes (Levine, 1998). Within LUTI models, destination choice models (DCMs) simulate how workplace decisions influence broader urban systems—for example, a new subway line may increase job concentration in specific zones, altering commuting flows and residential demand (Sarri et al., 2023). By incorporating variables such as travel time, zone attractiveness, and labor market characteristics, these models enable policymakers to assess the ripple effects of transportation investments on employment distribution and urban form (Ibeas et al., 2013).

Conventional methods for workplace location choice rely on discrete choice models (DCMs), particularly multinomial logit (MNL) and nested logit (NL) frameworks grounded in random utility maximization (RUM) theory. RUM posits that individuals select destinations maximizing their utility, expressed as

$$U = v_{n,j} + \epsilon_{n,j} \tag{1}$$

where $v_{n,j}$, captures observable factors like travel cost and job density, while $\epsilon_{n,j}$ represents unobserved preferences (Train, 2009). The nested logit (NL) model, widely adopted for workplace choice, groups alternatives into hierarchical nests (e.g., by occupation or geographic proximity) to relax the independence of irrelevant alternatives (IIA) assumption inherent in MNL (Mcfadden, 1977). For instance, Naqavi et al. (2023) formulated a utility function integrating spare-time accessibility and occupational-specific constants:

$$v_{n,ij} = (\beta_A + \beta_{Acr} 1_n(cr)) A_{n,ij} + \lambda \log \sum_{k=1}^{K} e^{\alpha_k/\lambda + \log(N_{jk})} \tag{2}$$

where, $\beta_A$ is the parameter for the spare time accessibility measure $A_{n,ij}$, individual characteristics is represented by $1_n(cr)$ as a dummy variable indicating if the individual has access to personal car with a parameter $\beta_{Acr}$, and $\alpha_k$ represents occupation-specific constants for different types of occupations. The log-sum term represents the expected utility of choosing any of the workplaces available in zone $j$, where these workplaces belong to an occupational type ($k \in 1, ...K$) and there are $N_{jk}$ such workplaces. This approach accounts for both individual constraints (e.g., car access) and zonal characteristics (e.g., job availability). Empirical applications, such as Ho and Hensher (2016), demonstrate how NL models capture competition effects between employment hubs, while Inoa et al. (2015) use multi-level nests to model interdependencies between residential and workplace choices.

## Research Gaps

Despite their widespread use, DCMs face several critical limitations in accurately modelling workplace location decisions. These models often impose restrictive assumptions about utility functions, decision rules, and choice sets, leading to potential misspecifications and reduced predictive accuracy, particularly in complex or unfamiliar contexts (van Cranenburgh et al., 2022). Even advanced variants such as Mixed Logit and Latent Class Choice Models struggle with capturing taste heterogeneity without strong a priori assumptions about systematic utility and error structures (Ben-Akiva et al., 2002). These models are also computationally intensive and require extensive high-quality data, making scalability difficult as the number of alternatives increases. Given the increasing complexity of urban systems, there is a growing need for more flexible and behaviourally realistic approaches—potentially integrating data-driven methods or simulation-based agent-based models (ABMs)—to enhance workplace location modelling within LUTI frameworks.

DNNs are capable of approximating virtually any function directly from data, thereby learning flexible representations from large datasets—capabilities that often exceed those attainable with hand-engineered features crafted by domain experts. van Cranenburgh et al. (2022) provides an comprehensive review of various DNN-based models applied to choice modeling tasks—ranging from travel mode and vehicle type to train type—and compares them to traditional DCMs such as the Multinomial Logit (MNL), Nested Logit (NL), and Latent Class MNL models. However, these research primarily focus on the application of DNNs in discrete choice scenarios with relatively small choice sets (typically 5–10 alternatives), such as travel mode selection. In contrast, workplace location choice involves a substantially larger number of alternatives. Hence, the **Paper 5**, address this gap by presenting a custom DNN specifically for workplace location choice that have comparable or better performance to the hand-engineered DCM models.

# 3 Deep Learning Models in Transport

Deep learning (DL) has emerged as a transformative force in transportation science, offering sophisticated data-driven solutions to complex challenges in traffic management and urban planning. The proliferation of large-scale data, coupled with advancements in computational power, has enabled DL models to uncover intricate patterns within transportation systems, moving beyond the limitations of traditional analytical or statistical methods (Haghighat et al., 2020). The application of deep learning in transportation represents a paradigm shift from technology-driven independent systems to data-driven integrated systems of systems. As highlighted in the review by (Wang et al., 2019; Zhang et al., 2024). A diverse array of DL architectures now provides a powerful toolkit for automating model development through the systematic discovery of significant patterns in data.

This chapter surveys the DL methods that form the methodological backbone of this thesis. Each section is tailored to explain the advanced methods used to address specific research objectives. Section 3.1 details the Deep Choice Networks applied to behaviour modelling in **Paper 5**; Section 3.2 examines the computer vision algorithms used for automated data collection in **Paper 1** and **Paper 2**; and Section 3.3 explores the Deep Generative Models developed for synthetic data generation in **Paper 3** and **Paper 4**.

## 3.1 Deep Choice Networks

A Deep Neural Network (DNN) is a class of machine learning algorithms that uses multiple layers to progressively extract higher-level features from the raw input. Unlike traditional neural networks with only a few layers, DNNs, characterized by their depth (i.e., a significant number of layers), can model complex non-linear relationships in data (Lecun et al., 2015). This capability has led to their increasing adoption in various fields, including transportation modelling.

In the context of transportation models, specifically *Discrete Choice Models (DCMs)*, DNNs offer a powerful alternative or complement to traditional econometric models like the Multinomial Logit (MNL) model. For example, for a standard MNL, the utility $U_{ni}$ for individual $n$ choosing alternative $i$ in an MNL model is typically formulated as:

$$U_{ni} = V_{ni} + \epsilon_{ni} = \beta' X_{ni} + \epsilon_{ni} \tag{3}$$

where $X_{ni}$ is a vector of observed attributes, $\beta$ is a vector of parameters to be estimated, and $\epsilon_{ni}$ is an unobserved error term assumed to be independently and identically distributed (IID) Gumbel. The choice probability is then:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \tag{4}$$

DNNs enhance this framework by replacing the linear utility specification with a non-linear function approximator. The systematic utility is reformulated as $V_{ni} = f(\mathbf{X}_{ni}; \mathbf{W}, \mathbf{b})$, where $f$ represents the DNN with parameters $\mathbf{W}, \mathbf{b}$, and the network output feeds into the logit formula (Aouad and Désir, 2022). The fundamental architecture of a common DNN is the feed-forward neural network. In this structure, information flows in one direction, from the input layer, through one or more hidden layers, to the output layer. Each neuron (or node) in a layer receives inputs from neurons in the previous layer, computes a weighted sum of these inputs, adds a bias term, and then applies an activation function to produce its output.

For a network with layers $l = 1, 2..k - 1$, where each layer $l$ contains $h_l$ neurons, the layer computations are defined as:

$$z_l = W_l \cdot a_{l-1} + b_l \tag{5}$$

$$a_l = \phi(z_l) \tag{6}$$

The final output representing the systematic utility for zone $j$, is obtained at layer $l = k$:

$$V_j = \phi(W_k \cdot a_k) \tag{7}$$

In these equations, the matrix $W_l$ represents the weights of layer $l$, with dimensions $[h_l, h_{l-1}]$, and the vector $b_l$ represents the biases for layer $l$, with dimensions $h_l$. For the input layer $l = 0$, we have $a = \mathbf{X}$, and the weight matrix dimensions are $[h_l, m]$. The non-linear function $\phi$ serves as the activation function (e.g., sigmoid, ReLU, tanh) (Goodfellow et al., 2016), which modifies values between layers.

Training a DNN involves adjusting its weights and biases to minimize a loss function, which quantifies the difference between the network's predictions and the actual target values. The most common training algorithm for DNNs is back-propagation (Rumelhart et al., 1986). This method iteratively computes the gradient of the loss function with respect to each weight and bias in the network. The process starts at the output layer, calculating the error and propagating it backward through the network, layer by layer. For a given training sample, if $L$ is the loss function, the gradient for a weight $W_l$ is calculated using the chain rule:

$$\frac{\partial L}{\partial W_l} = \frac{\partial L}{\partial a_l}\frac{\partial a_l}{\partial z_l}\frac{\partial z_l}{\partial W_{l-1}} \tag{8}$$

These gradients are then used by an optimization algorithm, such as stochastic gradient descent (SGD) or its variants (e.g., Adam), to update the weights and biases in the direction that minimizes the loss:

$$W_l^{(new)} = W_l^{(old)} - \eta \frac{\partial L}{\partial W_l} \tag{9}$$

where $\eta$ is the learning rate.

In DNNs can be used to directly estimate the choice probabilities or the systematic utility $V_{ni}$ without imposing strong a priori assumptions about the functional form of the utility or the distribution of the error terms (Sifringer et al., 2020). The input layer of the DNN would receive attributes of the decision-maker and the alternatives, while the output layer would provide the probability of choosing each alternative, often using a softmax activation function to ensure probabilities sum to one, analogous to the MNL choice probability formula. By learning complex, non-linear relationships between input features and choice outcomes directly from data, DNNs can potentially capture more nuanced aspects of travel behaviour and achieve higher predictive accuracy compared to traditional Logit models, especially when dealing with large and complex datasets (Alwosheel et al., 2018).

## 3.2 Computer Vision

Traffic management systems have evolved significantly through the integration of computer vision techniques. This evolution has fundamentally changed how transportation data is collected and analysed. Modern ITS use sophisticated object detection and tracking algorithms to automate monitoring processes. These systems enable real-time traffic analysis, speed measurement, accident detection, and comprehensive traffic flow assessment (Buch et al., 2011). This technological advancement has replaced traditional manual monitoring approaches with automated systems. These automated systems can process vast amounts of visual data with remarkable accuracy and efficiency.

Monocular visual traffic surveillance systems used in ITS consist of multiple structured steps. These steps include object detection, object tracking, camera calibration, and various downstream applications. The overall system typically

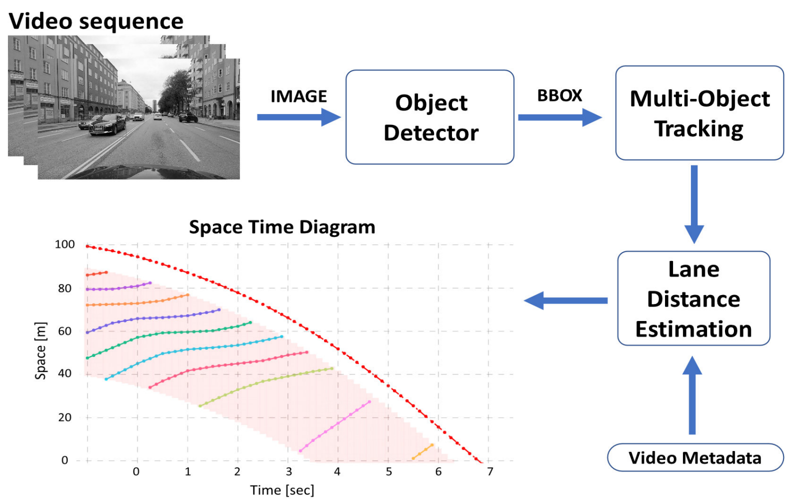



**Figure 5: Flow chart that illustrates the methodology in Paper 1 for extracting the vehicle trajectory from a street-view video sequence using different computer vision algorithms.**

consists of several interconnected components. Object detection identifies vehicles and other road users in each frame, along with their type, position, bounding box, or mask. Object tracking then links detection results in consecutive frames to derive vehicle trajectories (Ghahremannezhad et al., 2023). Figure 5 shows the steps used in **Paper 1** for extracting vehicle trajectories from an onboard moving camera. The procedure highlights the importance of object detection and tracking in this research.

Object detection and tracking in ITS enable a comprehensive range of applications that are critical for modern traffic management and safety. Detection algorithms serve as a foundation by localising and classifying objects within surveillance camera feeds. These algorithms support several key functions: Automatic Number Plate Recognition (ANPR) for vehicle identification, traffic sign detection for autonomous vehicle guidance, vehicle counting for traffic volume analysis, and vehicle type classification (cars, buses, trucks, motorcycles) for infrastructure planning (Zhang et al., 2022). Once objects are detected, tracking algorithms maintain consistent identities across frames to generate trajectories. These trajectories enable speed estimation through displacement analysis (Kumar et al., 2021b), traffic flow analysis, including vehicle interactions and queue formation for congestion monitoring (Datondji et al., 2016), behaviour learning and anomaly detection to identify illegal manoeuvres or accidents, incident management through stopped vehicle detection, and pedestrian monitoring for safety applications at intersections (Dilek and Dener, 2023). These integrated detection and tracking capabilities collectively support diverse ITS functions. These functions range from real-time traffic surveillance and congestion management to autonomous vehicle operation and emergency response systems.

### 3.2.1 YOLO

Within the methodology presented in **Paper 1**, we employed the YOLOv5 object detector network, developed by Glenn Jocher (2022), which belongs to the You Only Look Once (YOLO) family of single-shot detectors. The YOLO family of object detectors has gained prominence in computer vision due to its effective balance between speed and accuracy, making it particularly suitable for real-time applications in Intelligent Transportation Systems (ITS). As highlighted by Nepal and Eslamiat (2022); Ge et al. (2021), the YOLO architecture proves especially well-suited for processing video streams from traffic surveillance cameras, enabling the collection of crucial data for traffic management, safety analysis, and autonomous systems.

YOLOv5, which builds upon YOLOv4 from Bochkovskiy et al. (2020), comprises three essential components that work together to detect objects efficiently, as illustrated in Figure 6. The input image is first processed by the *backbone* of the network, typically a Convolutional Neural Network (CNN) such as CSPDarknet53 (Cross Stage Partial Darknet). The backbone's primary function involves extracting hierarchical feature maps at different spatial resolutions. These feature maps capture increasingly complex visual information, progressing from low-level edges and textures to higher-level object components.

Subsequently, the *neck* of the YOLOv5 architecture employs structures such as a Path Aggregation Network (PANet) or a Bidirectional Feature Pyramid Network (BiFPN) to aggregate features from various levels of the backbone. This multi-scale feature fusion proves critical in ITS applications because traffic scenes inherently contain objects of vastly different sizes. Effective feature representation across these scales ensures robust detection performance across diverse object categories.

The final stage involves the *head*, which performs the actual detection task. YOLO divides the input image into an $S \times S$ grid, where each cell predicts $B$ bounding boxes along with their associated confidence scores (Redmon et al., 2015). The confidence score quantifies both the probability of an object existing within the cell $\Pr(\text{Object})$ and the accuracy of the predicted bounding box, measured by the Intersection over Union (IoU) with the ground truth:

$$Confidence = \Pr(\text{Object}) * IoU_{\text{pred}}^{\text{truth}} \tag{10}$$

Each grid cell simultaneously predicts conditional class probabilities $Pr(\text{Class}_i \mid \text{Object})$, which enables the generation of class-specific confidence scores for each bounding box. The final output tensor comprises the four coordinates of the bounding box and the confidence score, indicating the likelihood that the box contains an object.

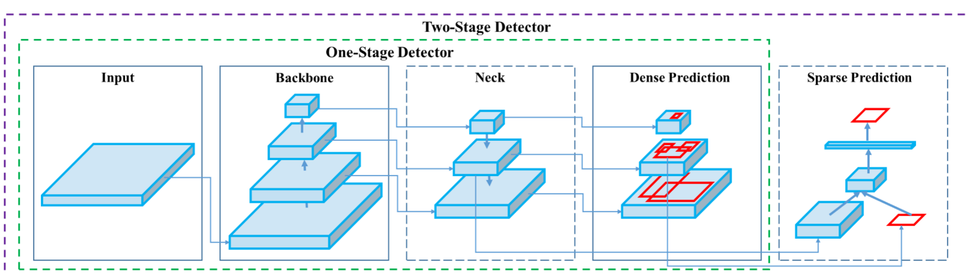



**Figure 6: Object Detection structure in YOLO (Bochkovskiy et al., 2020)**

### 3.2.2 DeepSORT

Following object detection in each frame by YOLOv5, an object tracking algorithm becomes essential for associating detections across successive frames and maintaining consistent identities over time. DeepSORT represents an enhanced tracking-by-detection algorithm that extends the Simple Online and Realtime Tracking (SORT) framework through the integration of deep learning-based appearance features. This advancement significantly improves tracking robustness in complex traffic scenarios, particularly during occlusions (Du et al., 2023). The SORT algorithm typically comprises four key stages: detection, estimation, target association, and track identity creation/deletion. Figure 7 illustrates the key stages involved in this object tracking methodology.

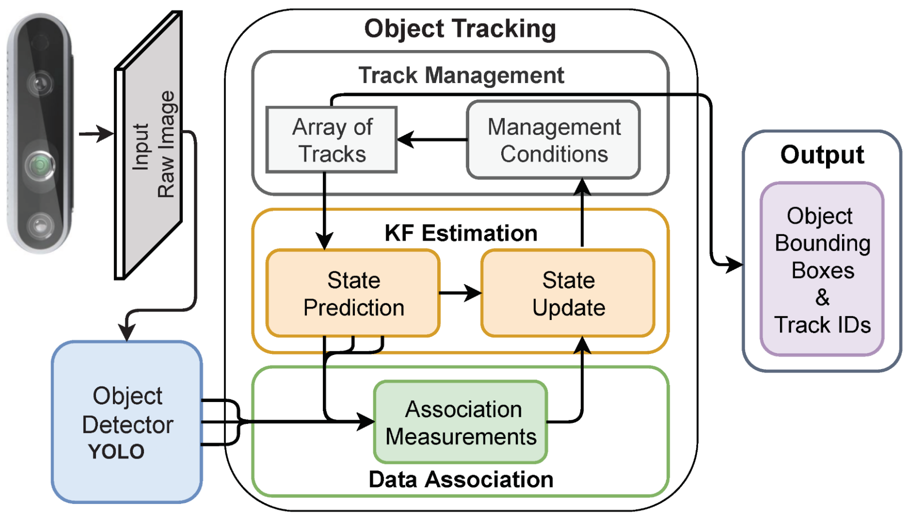



**Figure 7: Overview of object detection using DeepSORT (Pereira et al., 2022)**

The tracking process initiates with *object detection* in each video frame, performed by the YOLO detector, which generates a set of bounding boxes $B_t = (x_i, y_i, w_i, h_i)$ for every detected object $i$ at frame $t$. Here, $(x, y)$ denote the box's centre coordinates, while $(w, h)$ represent its width and height. DeepSORT's core functionality revolves around a recursive cycle of estimation and association.

For *motion estimation*, the algorithm uses a Kalman filter to predict future object states. The state vector for each tracked object is $\mathbf{x}_t = [x, y, \gamma, h, \dot{x}, \dot{y}, \dot{\gamma}, \dot{h}]^T$ incorporates positional parameters $(x, y)$, aspect ratio $\gamma$, height $h$, and their respective velocity components. The prediction step employs:

$$\mathbf{x}_{t|t-1} = \mathbf{F}\mathbf{x}_{t-1}$$

where $\mathbf{F}$ denotes the state transition matrix responsible for propagating positional states based on velocity estimates.

The *data association* phase combines both motion characteristics and appearance features. Motion compatibility between predictions and detections is quantified through the Mahalanobis distance $d^{(m)}$:

$$d^{(m)} = (\mathbf{z}_t - \mathbf{H}\mathbf{x}_{t|t-1})^T \mathbf{S}^{-1} (\mathbf{z}_t - \mathbf{H}\mathbf{x}_{t|t-1})$$

where $\mathbf{z}_t$ corresponds to the detection measurement and $\mathbf{H}$ represents the observation matrix. Concurrently, a deep association metric calculates the cosine similarity $d^{(a)}$ between the appearance features $\mathbf{f}_t$ of detected objects and stored appearance descriptors $\mathbf{f}_k$, extracted using CNN-based networks:

$$d^{(a)} = 1 - \frac{\mathbf{f}_t \cdot \mathbf{f}_k}{||\mathbf{f}_t|| \cdot ||\mathbf{f}_k||}$$

These metrics combine through a weighted cost function $C = \lambda d^{(m)} + (1 - \lambda) d^{(a)}$, where $\lambda$ balances the relative contributions of motion and appearance features. The Hungarian algorithm then resolves this bipartite matching problem to associate detections with existing tracks.

*Track management* employs confirmation thresholds (typically three consecutive matches) to upgrade tentative tracks to confirmed status. Tracks that remain unmatched for $T_{\max}$ frames (usually 5-10) are terminated to prevent spurious tracking artefacts.

The StrongSORT variant employed in **Paper 1** (Du et al., 2023) introduces several enhancements to the original DeepSORT framework. The appearance branch incorporates an OSNet architecture pre-trained on ImageNet, replacing the standard CNN to enable more discriminative feature extraction. This modification is complemented by an exponential moving average (EMA) technique in the matching strategy. Within the motion branch, StrongSORT implements camera motion compensation via the Enhanced Correlation Coefficient (ECC) algorithm and substitutes the conventional Kalman filter with a Noise-Scaled Adaptive (NSA) Kalman filter for improved motion prediction accuracy.

## 3.3 Deep Generative Models

For effective Agent-Based Models (ABMs) in transportation and urban planning, generating realistic synthetic populations is crucial. Studies such as those by Kim and Bansal (2023); Borysov et al. (2019); Garrido et al. (2020) have demonstrated that Deep Generative Models (DGMs) are powerful tools for this purpose. These models learn underlying patterns from existing data sources like census records or surveys to generate new, unobserved yet plausible individuals with realistic attributes.

As illustrated in Figure 8, a standard Generative Adversarial Network (GAN) comprises two neural networks: a generator $G$ that synthesises synthetic populations $\hat{P}(X)$ from random noise vectors $Z$, and a discriminator $D$ that distinguishes these synthetic samples $\hat{P}(X)$ from real population data $P(X)$. While GANs represent a significant innovation, they often suffer from training instability and "mode collapse," where the generator produces limited sample diversity—a critical drawback for applications requiring heterogeneous synthetic populations[1][5].

The Wasserstein GAN (WGAN) addresses these limitations by redefining the loss function. As proposed by Gulrajani et al. (2017), WGAN replaces the discriminator with a "critic" (still denoted $D$) that outputs real-valued scores rather than probabilities. For a training batch of size $M$, the critic aims to maximise the following loss function:

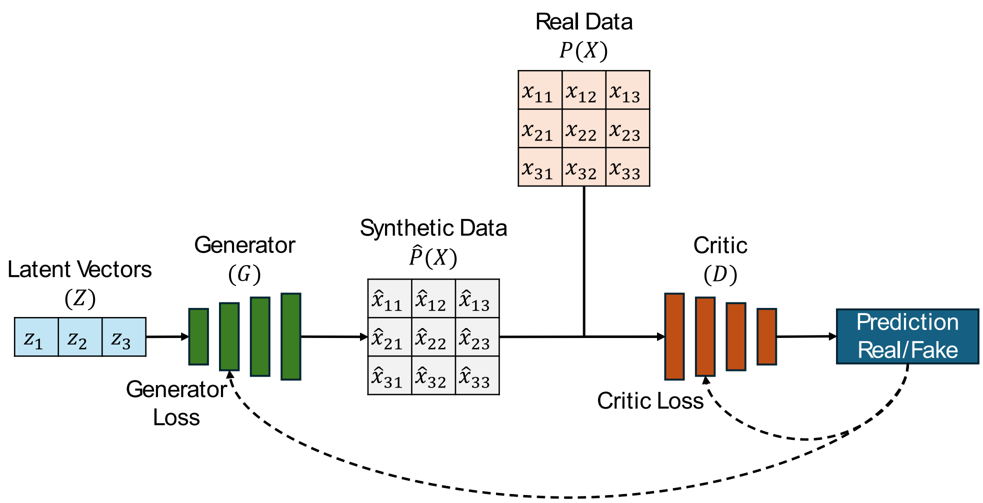



**Figure 8: Architecture and loss function of the Wasserstein GAN model.**

$$\begin{aligned}\mathscr{L}_D = &\frac{1}{M}\sum_{i=1}^{M} \underbrace{-D(X_i)}_{\text{Real score}} + \underbrace{D(G(Z_i))}_{\text{Synthetic score}} \\ &+ \lambda \cdot \left( \frac{1}{M}\sum_{i=1}^{M} \underbrace{(\left\| \nabla_{\tilde{X}_i} D(\tilde{X}_i) \right\|_2 - 1)^2}_{\text{Gradient penalty}} \right)\end{aligned} \tag{11}$$

The gradient penalty term approximates the Wasserstein-1 distance (Earth Mover's distance), where:

$$\begin{aligned}\tilde{X}_i = \alpha X_i + (1-\alpha)\hat{X}_i \\ \alpha \sim U[1],\end{aligned} \tag{12}$$

and $\lambda$ controls the penalty's intensity. Here, $\|\cdot\|_2$ denotes the Euclidean norm, while $X_i$ and $\hat{X}_i$ represent real and generated data points, respectively.

A higher $\lambda$ increases regularization, penalizing abrupt gradient changes that could destabilise training.

The generator $G$ minimises the critic's scores for synthetic samples, aligning $D(G(Z))$ with $D(X)$ via the Wasserstein metric:

$$\mathscr{L}_G = \frac{1}{M}\sum_{i=1}^{M} -D(G(Z_i)) \tag{13}$$

Training alternates between multiple critic updates and single generator updates. Unlike the Jensen-Shannon divergence in traditional GANs, the Wasserstein distance provides smoother loss gradients, enhancing training stability and reducing mode collapse. Empirical studies show this approach improves the correlation between loss metrics and synthetic data quality, particularly in high-dimensional spaces.

A key advantage of WGANs over methods like Iterative Proportional Fitting (IPF) is their ability to generate "sampling zeros" that are plausible data points absent from training samples but present in real populations. However, WGANs may also produce "structural zeros" that are implausible combinations non-existent in reality. To mitigate this, Kim and Bansal (2023) introduced boundary distance ($R_{BD}$) and average distance ($R_{AD}$) regularization terms:

$$R_{BD} = \frac{1}{M}\sum_{i=1}^{M} \min_{j\in\{1:N\}} \left( DIST(\widehat{X_i}, X_j) \right) \tag{14}$$

$$R_{AD} = -\frac{1}{NM}\sum_{i=1}^{M}\sum_{j=1}^{N} DIST(\widehat{X_i}, X_j) \tag{15}$$

where $DIST(X_i, X_j) = \sqrt{(X_i - X_j)^2}$, $\widehat{X_i}$ denotes $M$ generated samples, and $X_j$ represents $N$ training points. $R_{BD}$ encourages proximity to observed data clusters, while $R_{AD}$ discourages over-concentration in sparsely populated regions. This dual approach balances novelty and plausibility, addressing a fundamental challenge in synthetic population generation.

### 3.3.1 WGAN for Missing Data

In **Paper 3** of the thesis, we employed the masking approach presented by Neves et al. (2022) to handle missing data during the training of the WGAN model. During WGAN-GP training, we introduce a new matrix called the mask matrix as an additional input. The mask, denoted as $Y$, serves as the mask for the training data $X$, where missing values in $X$ are represented by zeros and non-missing values are represented by ones. Thus, the mask is a binary matrix of the same size as $X$.

The primary difference from the original WGAN training algorithm lies in how the matrix generated by the generator, $G(Z)$, is processed. Specifically, this generated matrix is multiplied by the mask $Y$ before being provided as input to the discriminator $D$ to obtain scores for fake samples. Figure 9 provides an illustration of the proposed training method, showcasing the training routine for both the generator and the discriminator.

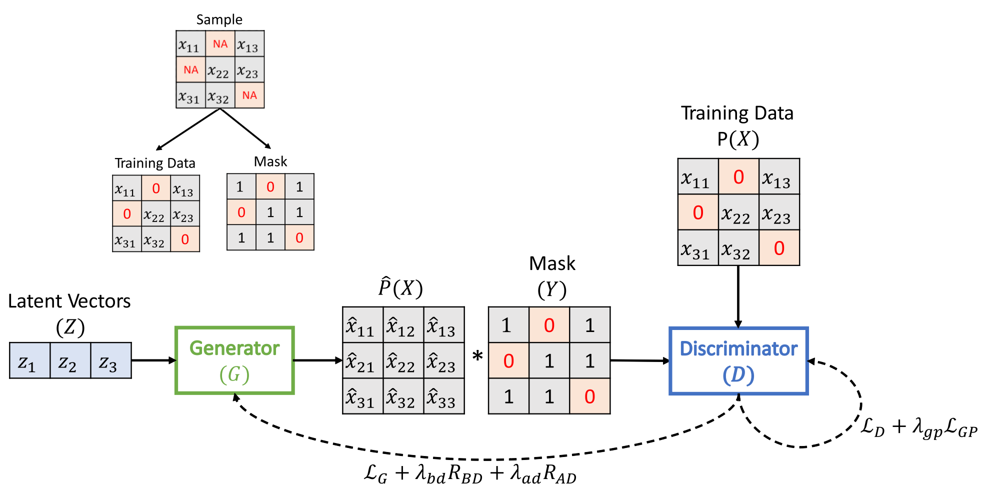



**Figure 9: Illustration for the proposed training process of WGAN with missing data.**

### 3.3.2 Conditional Tabular GAN (CT-GAN)

In **Paper 4** of the thesis, we utilised the Conditional Tabular GAN (CTGAN), proposed by Xu et al. (2019) to synthesise populations under certain future constraints. CTGAN enhances WGAN frameworks by introducing conditional generation and mode-specific normalisation to address tabular data challenges in population synthesis. This approach is particularly valuable for future scenario modelling, where capturing rare demographic combinations and complex attribute correlations is critical.

Whilst WGANs stabilise training through Wasserstein distance minimisation, CTGAN employs a conditional generator that uses training-by-sampling to balance imbalanced categorical variables, such as rare travel behaviours. Additionally, it incorporates a variational Gaussian mixture model to handle multi-modal continuous features, such as income distributions. This design enables CTGAN to synthesise future populations with realistic inter-variable dependencies, including emerging mobility patterns and demographic shifts, while maintaining statistical fidelity across both sparse and dense attribute categories. Consequently, CTGAN outperforms standard WGANs in preserving global column correlations.

# 4 Contributions and Discussions

This thesis advances the rapidly evolving field of DL applications in transportation and urban planning through five interconnected studies. Collectively, these papers address methodological gaps in current approaches in transportation and urban planning contributing to the broader objective of developing robust, data-driven DL solutions for transportation challenges. Table 2 summarizes each paper's contributions across three thematic areas within the transportation and urban planning domain, illustrating their collective impact on advancing analytical frameworks.

**Table 2: Summarizing the contribution of each paper in this thesis.**

| Paper | Topic | Model | Contribution |
|---|---|---|---|
| P1 | Traffic Management | YOLO + StrongSORT | Introducing a novel data source derived from on-board vehicle cameras using existing DL techniques. |
| P2 | Traffic Management | CTM + GA | Integrating DL-enabled data sources with analytical models for estimating traffic states. |
| P3 | Population Synthesis | WGAN | Improving the utility of incomplete data-samples through DL applications. |
| P4 | Population Synthesis | CT-GAN | Repurposing DGM to improve upon the limitations of traditional models for target synthesis |
| P5 | Workplace location choice | DNN | Using DNN for choice modelling to handle large choice sets and data-driven utility specification. |

In subsequent sections, this work elaborates on its alignment with the research objectives detailed in Section 1.2. The contributions are structured around two central themes— *data acquisition* and *analytical inference*, analysed in depth within Sections 4.1 and 4.2, respectively. Section 4.3 concludes by critically examining the research's limitations and proposing avenues for future works in the field.

## 4.1 Data Acquisition

The work presented in this thesis significantly broadens the scope of data sources available for transportation research by applying DL methods to extract and integrate information from datasets that are frequently overlooked or incomplete. The research in **Paper 1** and **Paper 3** demonstrate the versatility of these methods in enhancing data collection for transport analysis, illustrating how moving-camera video can function as a virtual sensor network and how incomplete surveys can be transformed into detailed synthetic populations through DL frameworks.

### Paper 1

The approach proposed in **Paper 1** exemplifies advancements in data acquisition through DL, by automating the construction of time–space diagrams from video sequences captured by vehicle-mounted cameras. Rather than developing new specialised sensors, this method repurposes existing street-view videos from on-board cameras, which are typically used for insurance claims or automated driving systems before being discarded.

To extract vehicle trajectories from these pre-existing video resources, the method integrates advanced computer vision techniques—specifically, the YOLOv5 multi-object detector and StrongSORT object tracker described in Section 3.2—with supplementary data such as GPS, photogrammetry, and geodesy. Empirical evaluation on the standardised KITTI dataset confirms that this pipeline generates time–space diagrams closely aligned with ground-truth measurements. As illustrated in Figure 10, the trajectories produced by our methodology for a KITTI video closely mirror the actual data. This precision is further quantified by photogrammetric analysis, which finds the trajectory reconstructions have a mean positional error of only 2.97 metres (standard deviation: 1.91 m), a level of accuracy sufficient for a wide range of traffic-state estimation applications.

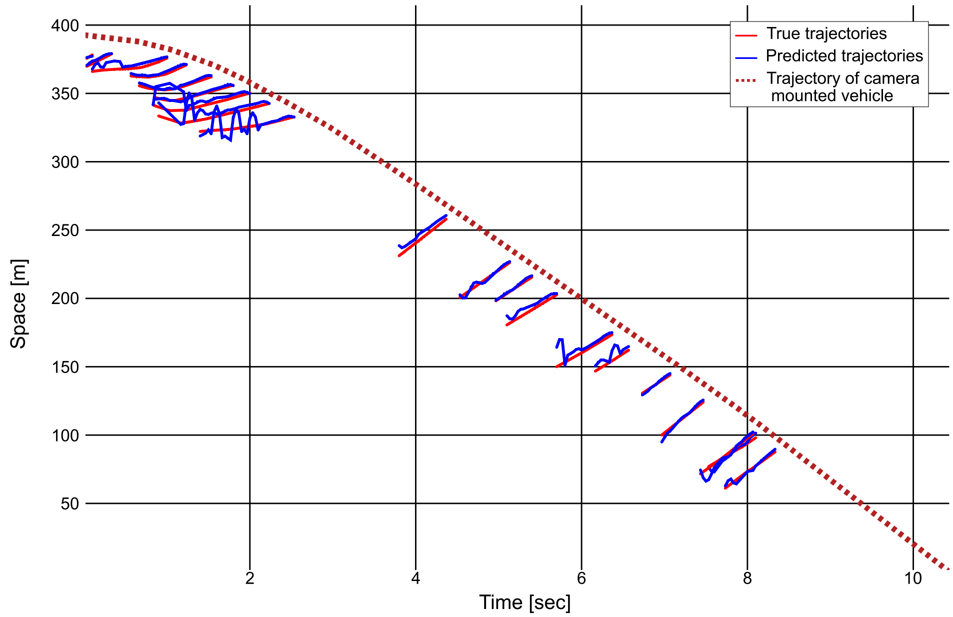


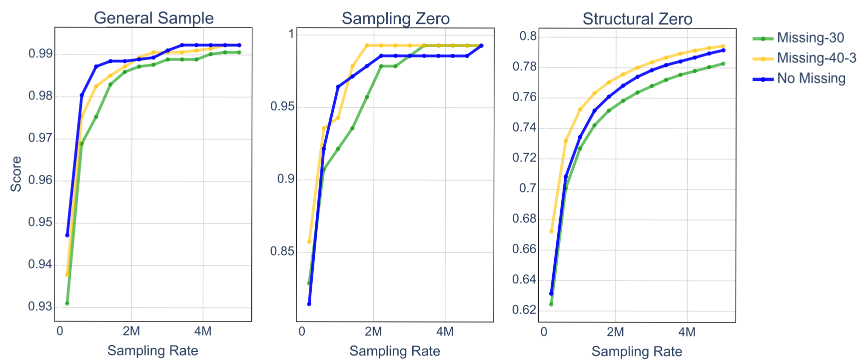

**Figure 10: Time-space diagram generated for video from KITTI dataset with true trajectories (in RED) and predicted trajectories (in BLUE).**

These outcomes fulfil the objectives of this thesis, as stated in Section 1.2, by establishing a novel data source for traffic analysis and validating a robust method for extracting valuable information using existing DL techniques.

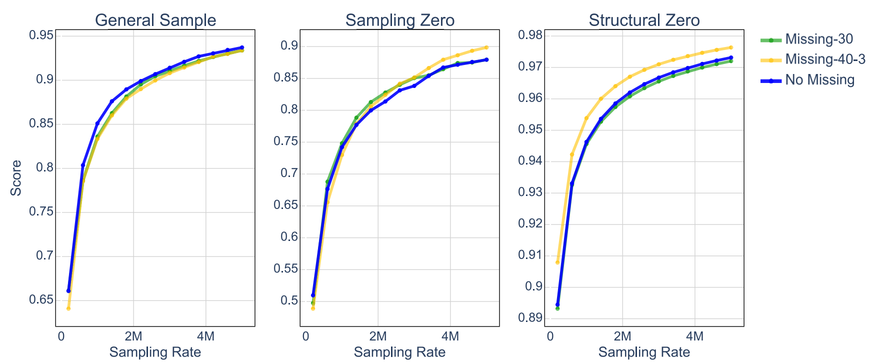
## Paper 3

Complementing this, the population synthesis work in Paper 3 focuses on enhancing incomplete datasets through deep learning. Addressing the challenge of incomplete survey microsamples, the study incorporates a binary mask into a WGAN model, enabling it to learn directly from fragmented data and reconstruct complete distributions. This approach repurposes incomplete travel-survey entries, which were frequently excluded, by embedding missing-data indicators directly into the generative process.

When evaluated on the Swedish national travel survey data, the methods accurately replicate both marginal and joint distributions of key demographic




(a) **16k-dim joint**




(b) **7M-dim joint**

**Figure 11: Plots with the ratio of general sample, sampling zero, structural zero for 16k and 7M dimensional joint data at different sampling levels.**

and travel variables, achieving performance metrics only marginally inferior to models trained on complete datasets. The WGAN model trained on partial data exhibits high values for metrics such as category coverage and total variation complement score. Figure 11 displays the numbers of sampling zeros and structural zeros across different sampling levels for two high-dimensional categorical joints—16k-dimensional and 7M-dimensional. The synthetic populations Missing-30 and Missing-40-3, generated using WGAN models trained on 30% and 40% missing data respectively, demonstrate performance very similar to the benchmark synthetic population generated by a model trained on complete data. Even when evaluated using standardised RMSE and $R^2$, the values for higher-order joint distributions for Missing-30 and Missing-40-3 are only slightly below those of the benchmark dataset. A more in-depth analysis of variations in the synthetic population is provided in **Paper 3**.

In alignment with the objectives outlined in Section 1.2, this research demonstrates a method to enhance the utility of existing microsamples. By leveraging data entries that would otherwise be discarded, this work reduces the reliance on costly or sensitive data collection efforts through strategic deep learning applications.

## 4.2 Analytical Inference

The research conducted under the theme of analytical inference demonstrates how DL approaches can systematically enhance traditional analytical methods, thereby improving traffic and urban system modelling. Specifically, **Paper 2** develops novel methodologies to effectively utilise unique data from moving vehicle cameras, while **Paper 4** and **Paper 5** focus on DL networks that address critical gaps in data utilisation and model flexibility, advancing established analytical frameworks.

## Paper 2

In alignment with the theme of analytical inference, **Paper 2** introduces a robust framework that combines the Cell Transmission Model (CTM) with a Genetic Algorithm (GA) for link-level traffic state estimation. This method leverages space-time diagrams from moving vehicle cameras to estimate traffic density in unobserved regions through a two-step optimisation process:

- **FD Calibration**: Initially, a GA calibrates the FD parameters by optimising them against the observed cell densities.
- **Density Estimation**: With the optimal FD parameters established, a second GA-driven process estimates the boundary conditions for the entire space-time diagram.

The final output is a complete space-time diagram, where traffic states in the previously unobserved cells are fully estimated using the optimised FD parameters and boundary conditions.

Validation using SUMO simulations across three distinct arterial link types demonstrates the method's ability to use partially observed mobile camera data to both calibrate the FD and estimate density. As detailed in Table 3, the FD calibration recovers free-flow speeds and optimal densities, with mean absolute errors below 5 km/h and 10 veh/km, respectively. Furthermore, the density estimation stage achieves a good level of accuracy with RMSEs between 4.8 veh/km and 5.6 veh/km across all probe vehicle runs. The consistent performance of both the calibration and estimation phases across various space-time matrices confirms the proposed method's scalability and its applicability to a wide range of traffic scenarios. These findings affirm the framework's capability to leverage sparse, camera-derived data for comprehensive link-level traffic estimation.

By successfully transforming partial camera trajectories into actionable information, this work fulfils a key objective of the thesis outlined in Section 1.2. It demonstrates the effective integration of novel, DL-enabled data sources with analytical models, advancing intelligent traffic management by bridging the gap between theoretical frameworks and practical applications.

**Table 3: Results for the FD calibration and Density estimation across 3 SUMO simulated links using the proposed methodology.**

| | **Link 1** | **Link 2** | **Link 3** |
|---|---|---|---|
| N probe runs | 1477 | 362 | 292 |
| N quartets | 4080 | 1896 | 1745 |
| **FD Calibration Results** | | | |
| Estim $v_f$ (km/hr) | 46.55 | 52.06 | 49.61 |
| Estim $k_c$ (veh/km) | 55.71 | 40.57 | 64.92 |
| $\delta v_f$ | 4.55 | 8.06 | 7.61 |
| $\delta k_c$ | 11.71 | 0.43 | 4.92 |
| **Density Estimation Results** | | | |
| Mean-RMSE (veh/km) | 5.28+3.84 | 4.82+1.93 | 5.62+3.31 |

## Paper 4

In **Paper 4**, the analytical innovation lies in the application of a Conditional Tabular Generative Adversarial Network (CT-GAN) to enhance population synthesis methods. The study introduces two key adaptations to address the limitations of traditional approaches. The first is the deployment of CT-GAN as a standalone model to synthesise populations conditioned on zone-level marginals, thereby overcoming high-dimensional constraints. The second is a novel hybrid pipeline, where CT-GAN generates a base population that is subsequently refined by Fitness-Based Synthesis Combinatorial Optimisation (FBS-CO). This hybrid strategy is designed to combine CT-GAN's strength in modelling complex attribute combinations with the precision of FBS-CO in enforcing exact marginal fits, presenting a scalable and robust alternative to conventional techniques.

The framework's performance was evaluated using two distinct datasets: travel surveys from different years and zone-level data containing aggregated marginals. The results demonstrate CT-GAN's reliability in reproducing critical demographic attributes. As shown in Table 4, the model generated synthetic populations that precisely matched conditional marginals, achieving perfect Total Variation Complement (TVC) and Category Adherence (CA) scores. For unconditioned attributes, the synthetic data also closely approximated the actual distributions, as evidenced by high TVC and CA scores.

**Table 4: Attribute level analysis for all CT-GAN synthetic population against corresponding target population.**

| | Conditioned Variables | | Un-conditioned Variables | |
|---|---|---|---|---|
| | **TVC** | **CA** | **TVC** | **CA** |
| 2011 | 1 | 1 | 0.898 | 1 |
| 2012 | 1 | 1 | 0.903 | 1 |
| 2013 | 1 | 1 | 0.903 | 1 |
| 2014 | 1 | 1 | 0.885 | 1 |
| 2015 | 1 | 1 | 0.720 | 0.783 |
| 2016 | 1 | 1 | 0.731 | 0.789 |

A zonal analysis for 89 zones in the Umeå region, summarised in Table 5, further confirms these findings. The standalone CT-GAN rigorously adhered to marginal constraints, achieving successful convergence in 84 of the 89 zones with perfect TVC scores of 1.0 for key variables. By comparison, the baseline FBS-CO model converged in only 62 zones with inferior performance metrics, while the hybrid model showed only a modest improvement over the baseline, converging in 65 zones.

By repurposing deep generative models to overcome the limitations of traditional population synthesis, this work exemplifies the thesis's aim descibed in Section 1.2. It delivers a scalable and effective method for producing high-quality synthetic populations, which are essential for robust agent-based transport simulations.

**Table 5: Result of the population generation for 89 SAMS zone in Umeå using three different models.**

| Status | #Zones | Avg. RSSZ | Avg. TVC | | |
|---|---|---|---|---|---|
| | | | AGE | SEX | WORK |
| | | **FBS-CO [Baseline]** | | | |
| Successful | 62 | 0.767 ± 0.207 | 0.930 ± 0.110 | 0.985 ± 0.073 | 0.904 ± 0.126 |
| Un-successful | 22 | 10.746 ± 28.021 | 0.896 ± 0.050 | 0.965 ± 0.002 | 0.891 ± 0.082 |
| No Population | 5 | - | - | - | - |
| | | **CT-GAN [Stand-alone]** | | | |
| Successful | 84 | - | 1.0 ± 0 | 1.0 ± 0 | 0.918 ± 0.138 |
| Un-successful | 0 | - | 0 | 0 | 0 |
| No Population | 5 | - | - | - | - |
| | | **CT-GAN + FBS-CO [Hybrid]** | | | |
| Successful | 65 | 0.799 ± 0.201 | 0.957 ± 0.119 | 0.986 ± 0.057 | 0.918 ± 0.127 |
| Un-successful | 19 | 12.315 ± 30.671 | 0.901 ± 0.053 | 0.975 ± 0.001 | 0.907 ± 0.088 |
| No Population | 5 | - | - | - | - |

## Paper 5

**Paper 5** introduces a novel deep neural network (DNN) architecture for workplace location choice prediction, extending the application of such models to choice sets with over a thousand alternatives—a scale significantly larger than in previous studies. A key methodological innovation is the design of a "zone block," which processes diverse inputs, including occupation type, socioeconomic attributes, and accessibility measures. This feature allows the network to replace the rigid, systematic-utility specifications of traditional models with a more flexible, data-driven approach.

A rigorous comparison between the proposed DNN and a two-level nested logit Discrete Choice Model (DCM) was conducted to evaluate its performance. The results show that two DNN variants—one using a binary car ownership input (DNN-Car) and another using the complete attribute set (DNN-All)—outperformed the DCM on log-likelihood metrics. Pearson correlation analysis further confirmed that the DNNs more accurately replicate the observed relationships between workplace opportunities and choices. As illustrated in Figure 12, which shows the probability distribution of travel distances, the DCM excels at short-distance predictions, whereas the DNNs match or exceed its performance at longer distances by capturing non-linear effects. Subgroup analyses by gender and car ownership also revealed that the DNN-All model achieves a comparable level of fidelity to the DCM, despite the latter's reliance on structured assumptions.

By demonstrating that a data-driven DNN can effectively handle large-scale alternative sets and complex inputs without pre-defined utility functions, this paper aligns with the thesis objective defined in Section 1.2 of developing DL methods that can complement or surpass established analytical approaches in transport modelling.

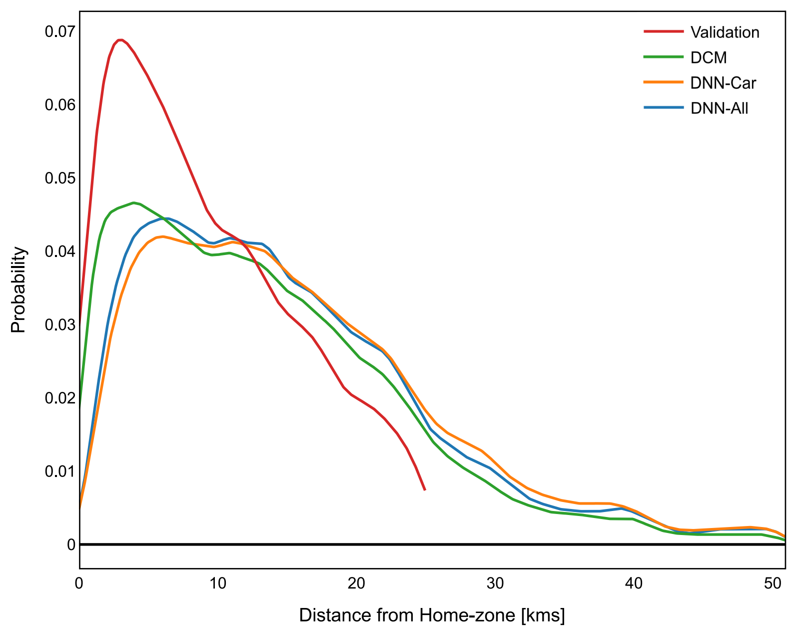



**Figure 12: Probability distribution of distance of chosen work from given home.**

## 4.3 Limitations and Future Work

Despite the advancements presented in this thesis, the research is subject to four key limitations, reflecting both paper-specific constraints and broader challenges in applying deep learning to transportation. These limitations, which also inform avenues for future research, are categorised as: (1) sensor assumptions and estimation biases, (2) modelling assumptions, (3) generalisability, and (4) scalability.

### Sensor Assumptions and Estimation Biases

Critical limitations in **Paper 1** and **Paper 2** stem from their reliance on camera-based data. **Paper 1**'s use of monocular cameras introduces photogrammetric vulnerabilities, as its similarity-triangle method assumes ideal camera calibration and uniform vehicle dimensions. Minor inaccuracies in calibration or variations in vehicle size propagate into significant depth-estimation errors at greater distances. These errors are compounded by imperfect bounding-box localisation and tracker instability, which distort trajectory reconstructions. **Paper 2** is constrained by the KITTI dataset's 146-metre detection limit, which restricts spatial discretisation and prevents CTM calibration for distant cells. Furthermore, adherence to the Courant–Friedrichs–Lewy (CFL) condition imposes cell sizes that can conflict with free-flow speed requirements, leading to biased density estimates. Collectively, these sensor-related issues undermine the accuracy of the derived traffic metrics.

Future work should focus on mitigating these issues by integrating monocular depth networks or hybrid LiDAR-camera systems to reduce dependency on photogrammetric assumptions. Tracker performance could be enhanced with trajectory-smoothing algorithms and domain-specific training. For **Paper 2**, expanding spatial coverage through multi-camera fusion or vehicle-to-everything (V2X) communication is essential, alongside rigorous real-world testing to validate system robustness beyond controlled benchmarks.

## Modelling Assumptions

The primary limitations of **Paper 2** are rooted in the assumptions of the CTM and GA optimisation. The CTM's reliance on a steady-state, triangular fundamental diagram and uniform conditions within each cell prevents it from capturing transient traffic dynamics. This simplification leads to aggregation errors, particularly in congested or rapidly changing conditions. The GA-based calibration is also sensitive to initial conditions; significant deviations between initial estimates and probe-derived data can cause error propagation during the simulation. These assumptions collectively reduce the model's accuracy in representing non-equilibrium traffic states.

To address these constraints, future research should explore more advanced traffic flow models, such as stochastic CTMs or higher-order Lighthill-Whitham-Richards (LWR) models, which can better represent transient behaviour. The GA calibration could be improved using adaptive fitness functions or surrogate-based optimisation. Furthermore, adopting dynamic discretisation schemes or models that do not require state discretisation could yield more accurate density reconstructions across all traffic regimes.

## Generalisability

A key challenge across all studies is generalisability, as each relies on specific datasets or simulated environments. Validation for **Paper 1** is confined to the KITTI dataset, which may not represent the full diversity of real-world driving scenarios. Similarly, **Paper 2**'s exclusive use of SUMO simulations fails to capture the complexities of live traffic. The population synthesis methods in **Papers 3** and **Paper 4** are tested only on Swedish travel survey data, limiting their proven applicability to different demographic contexts. Finally, the workplace choice model in **Paper 5** is developed for Stockholm, and its performance in other urban settings remains unknown. This reliance on narrow data sources limits confidence in the methods' robustness.

To improve generalisability, future research must validate these methods on broader, more diverse data. This includes testing **Paper 1** on various video datasets, conducting field trials for **Paper 2** in real-world traffic, applying the models from **Papers 3** and **Paper 4** to diverse datasets, and evaluating **Paper 5**'s DNN in different metropolitan areas. Integrating these methods into operational transport simulations will be crucial for demonstrating their practical applicability.

### Scalability

**Paper 2** and **Paper 4** face significant scalability challenges. In **Paper 2**, the computational load of the GA-based CTM calibration grows exponentially with network size, hindering real-time application. The model's reliance on short-term prediction to infer current traffic states further limits its utility for live traffic management. In **Paper 4**, the conditional rejection sampling in the CT-GAN becomes computationally intensive as the number of conditioning attributes or the granularity of marginals increases, which can lead to performance bottlenecks and sampling bias.

To address these limitations, future work on **Paper 2** could explore higher-order traffic flow models that permit coarser, more efficient discretisation. For **Paper 4**, investigating more efficient conditional sampling algorithms or alternative generative frameworks, such as Conditional Tabular Variational Autoencoders (CT-VAEs), could reduce computational overhead. For both models, leveraging parallel processing on GPU clusters is essential for enabling practical deployment across large-scale networks and populations.

## 4.4 Broader Implications

The deep-learning methods developed in this thesis carry important policy implications for urban and transport authorities. By enabling real-time traffic state estimation from low-cost cameras, cities can adopt more dynamic congestion-pricing schemes or adaptive signal controls that directly tie network management to demand patterns and emissions targets. Likewise, more accurate population synthesis and workplace-choice models support equity-focused land-use policies, ensuring that new housing or transit investments reflect the needs of diverse demographic groups rather than reinforcing existing disparities. However, these capabilities demand robust governance: policies must mandate data-privacy safeguards (e.g. anonymization protocols, secure storage), algorithmic-fairness audits, and transparent reporting so that no community is unintentionally disadvantaged by biased training data or opaque decision rules.

Deep learning is most appropriate when planners face large, heterogeneous datasets and require flexible, pattern-recognition tools—such as vision-based traffic monitoring or demand forecasting from mobile-device traces. In contrast, simpler statistical or rule-based models may be preferable in low-data settings or safety-critical applications where interpretability and certifiable guarantees are paramount. Key criteria for DL adoption include data volume and variety, computational resources, the criticality of real-time updates, and the degree of acceptable "black-box" opacity. Beyond the topics studied here, DL shows promise in optimizing public-transit schedules, enhancing freight-logistics routing, predicting infrastructure maintenance needs, and refining shared-mobility services. Policy frameworks should therefore be adaptive, establishing clear thresholds for when to deploy DL tools, what performance validations are required, and how to integrate them responsibly into existing planning processes.

Finally, the transport domain offers valuable lessons for the broader deep-learning community. Spatio-temporal flow dynamics, network constraints, and demand elasticity all encourage the development of graph-based and physics-informed neural architectures—advances that can translate to other fields such as energy systems or epidemiology. Moreover, the ethical and equity challenges encountered in transportation planning drive innovations in fairness-aware algorithms, transparency measures, and participatory model governance. By sharing these methodologies and insights, transport practitioners can help shape more robust, interpretable, and socially responsible deep-learning paradigms across diverse application areas.

# 5 Conclusion

This thesis has fulfilled its aim of investigating the systematic application of deep learning (DL) methods to enhance data collection and inference in transportation and urban planning. Through five interconnected research papers, the work has addressed fundamental limitations of traditional analytical models by exploring how deep learning techniques can complement or outperform established approaches across three distinct transportation domains: traffic management, population synthesis, and workplace location choice.

The research has been structured around two complementary themes—data acquisition and analytical inference—each contributing to the overarching objective of developing more robust, data-driven solutions for contemporary transportation challenges. This concluding chapter presents the key findings, evaluates the fulfilment of research objectives, and reflects on the broader implications of this work for the future of transportation and urban planning research.

## 5.1 Research Objectives and Achievements

The thesis has successfully addressed its research objectives through systematic exploration of DL applications across multiple transportation domains.

### Theme 1: Data Acquisition

**Paper 1** established a novel traffic data source through the innovative use of street-view video processing from vehicle-mounted cameras. By integrating YOLOv5 object detection with StrongSORT tracking algorithms, the research demonstrated how ordinary vehicles can be transformed into mobile traffic sensors, generating time-space diagrams with remarkable accuracy. This achievement represents a fundamental shift from traditional fixed-infrastructure approaches to dynamic, mobile sensing networks.

**Paper 3** advanced population synthesis methodologies by developing a robust framework for handling incomplete micro-sample data using WGAN with masking techniques. The research successfully demonstrated that synthetic populations generated from incomplete data achieved performance metrics comparable to models trained on complete datasets. This breakthrough addresses a critical limitation in traditional population synthesis, where incomplete data entries were typically discarded, thereby reducing the utility of expensive survey data.

### Theme 1: Analytical Inference

**Paper 2** developed innovative traffic state estimation methodologies that effectively utilise street-view trajectory data through the combination of Cell Transmission Models with Genetic Algorithm optimisation. The research achieved consistent performance across multiple traffic scenarios. This represents a significant advancement in traffic state estimation using novel data.

**Paper 4** enhanced population synthesis through the application of Conditional Tabular GANs for target-year population generation. The research demonstrated superior performance in generating synthetic populations that precisely matched conditional marginals. This work addresses the critical need for future scenario planning in agent-based transportation models.

**Paper 5** advanced workplace location choice modelling by developing custom deep neural network architectures capable of handling large-scale choice sets with more that 1000 alternatives. The research demonstrated that data-driven DNN approaches could match or exceed the performance of traditional discrete choice models whilst eliminating the need for pre-specified utility functions.

## 5.2 Key Contributions and Findings

The collective contributions of this research span three critical transportation domains, each demonstrating how deep learning can systematically enhance existing frameworks.

### Traffic Management Domain

The integration of computer vision with traditional traffic flow models represents a paradigm shift in traffic data collection and analysis. The research established that vehicle-mounted cameras can serve as comprehensive traffic sensors, providing spatial coverage far exceeding traditional fixed sensors or GPS-based floating car data. The successful integration of these novel data sources with established traffic flow models through genetic algorithm optimisation demonstrates their practical applications in real-world transportation systems, moving beyond theoretical advances to provide actionable insights.

### Population Synthesis Domain

The research has fundamentally addressed two persistent challenges in population synthesis: handling incomplete training data and generating populations for future scenarios. The development of masking techniques for WGAN training enables the utilisation of previously discarded incomplete survey data, effectively increasing the utility of expensive data collection efforts. The application of conditional generation through CT-GAN provides a scalable solution for target-year population synthesis, essential for long-term transportation planning and policy evaluation.

### Workplace Location Choice Domain

The successful application of deep neural networks to large-scale discrete choice problems represents a significant methodological advancement. The research demonstrated that custom DNN architectures can effectively handle choice sets with over 1,000 alternatives whilst maintaining comparable or superior performance to traditional discrete choice models. This breakthrough enables more realistic representation of spatial choice processes in urban systems.

## 5.3 Concluding Remarks

This thesis has demonstrated that deep learning methods can systematically enhance data collection and inference in transportation and urban planning, effectively addressing many limitations of traditional analytical models. Through the development of novel data sources, advanced analytical techniques, and hybrid methodological approaches, the research has contributed to the evolution of transportation science towards more data-driven, adaptive, and scalable solutions.

The work establishes a foundation for the next generation of transportation analytics, where the integration of deep learning with domain expertise creates opportunities for more accurate, efficient, and responsive urban management systems. Whilst challenges remain in terms of generalisability and scalability, the demonstrated potential of these approaches provides a clear pathway for continued advancement in the field.

The contributions of this research extend beyond the specific applications studied, offering methodological frameworks that can be adapted to address emerging challenges in transportation and urban planning. As cities continue to grow in complexity and data availability expands, the principles and techniques developed in this thesis will remain relevant for creating more intelligent, sustainable, and equitable urban transportation systems.

# References


Abirami, S., Pethuraj, M., Uthayakumar, M., and Chitra, P. (2024). A systematic survey on big data and artificial intelligence algorithms for intelligent transportation system. *Case Studies on Transport Policy*, 17:101247.

Acheampong, R. A. and Silva, E. (2015). Land use–transport interaction modeling: A review of the literature and future research directions. *Journal of Transport and Land Use*, 8(3):11–38.

Alwosheel, A., van Cranenburgh, S., and Chorus, C. G. (2018). Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling*, 28:167–182.

Amalan, T. P., Du, B., Kumarage, A. S., Wickramasuriya, R., and Perez, P. (2023). High frequency data in Land use and transport integrated model: A review of sources and application. *Asian Transport Studies*, 9:100092.

Aouad, A. and Désir, A. (2022). Representing Random Utility Choice Models with Neural Networks. *arXiv preprint arXiv:2207.12877*.

Ardeshiri, A. and Vij, A. (2019). Lifestyles, residential location, and transport mode use: A hierarchical latent class choice model. *Transportation Research Part A: Policy and Practice*, 126:342–359.

Atluri, M., Chowdhury, M., Kanhere, N., Fries, R., Sarasua, W., and Ogle, J. (2009). Development of a sensor system for traffic data collection. *Journal of Advanced Transportation*, 43(1):1–20.

Axhausen, K. W. ., Müller, and Kirill (2010). Population synthesis for microsimulation State of the art. *Arbeitsberichte Verkehrs- und Raumplanung*, 638.

Barry, M., Montiel, J., Bifet, A., Wadkar, S., Manchev, N., Halford, M., Chiky, R., Jaouhari, S. E., Shakman, K. B., Fehaily, J. A., Le Deit, F., Tran, V. T., and

Guerizec, E. (2023). StreamMLOps: Operationalizing Online Learning for Big Data Streaming & Real-Time Applications. *Proceedings - International Conference on Data Engineering*, 2023-April:3508–3521.

Bastarianto, F. F., Hancock, T. O., Choudhury, C. F., and Manley, E. (2023). Agent-based models in urban transportation: review, challenges, and opportunities. *European Transport Research Review*, 15(1):19.

Batty, M. (2012). A generic framework for computational spatial modelling. *Agent-Based Models of Geographical Systems*, pages 19–50.

Beckman, R. J., Baggerly, K. A., and McKay, M. D. (1996). Creating synthetic baseline populations. *Transportation Research Part A: Policy and Practice*, 30(6):415–429.

Ben-Akiva, M., Mcfadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D. S., Daly, A., De Palma, A., Gopinath, D., Karlstrom, A., and Munizaga, M. A. (2002). Hybrid Choice Models: Progress and Challenges. *Marketing Letters*, 13(3):163–175.

Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *ArXiv*.

Borysov, S. S., Rich, J., and Pereira, F. C. (2019). How to generate micro-agents? A deep generative modeling approach to population synthesis. *Transportation Research Part C: Emerging Technologies*, 106:73–97.

Buch, N., Velastin, S. A., and Orwell, J. (2011). A review of computer vision techniques for the analysis of urban traffic.

Cao, M., Zhu, W., and Barth, M. (2011). Mobile traffic surveillance system for dynamic roadway and vehicle traffic data integration. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 771–776. IEEE.

Choupani, A. A. and Mamdoohi, A. R. (2016). Population Synthesis Using Iterative Proportional Fitting (IPF): A Review and Future Research. In *Transportation Research Procedia*, volume 17, pages 223–233. Elsevier B.V.

Datondji, S. R. E., Dupuis, Y., Subirats, P., and Vasseur, P. (2016). A Survey of Vision-Based Traffic Monitoring of Road Intersections. *IEEE Transactions on Intelligent Transportation Systems*, 17(10):2681–2698.

Dilek, E. and Dener, M. (2023). Computer Vision Applications in Intelligent Transportation Systems: A Survey. *Sensors*, 23(6):2938.

Du, Y., Zhao, Z., Song, Y., Zhao, Y., Su, F., Gong, T., and Meng, H. (2023). StrongSORT: Make DeepSORT Great Again. *IEEE Transactions on Multimedia*, 25:8725–8737.

Fabrice Yaméogo, B., Gastineau, P., Hankach, P., and Vandanjon, P. O. (2020). Comparing Methods for Generating a Two-Layered Synthetic Population. *Transportation Research Record*, 2675(1):136–147.

Farooq, B., Bierlaire, M., Hurtubia, R., and Flötteröd, G. (2013). Simulation based population synthesis. *Transportation Research Part B: Methodological*, 58:243–263.

Fransen, K., Neutens, T., Farber, S., De Maeyer, P., Deruyter, G., and Witlox, F. (2015). Identifying public transport gaps using time-dependent accessibility levels. *Journal of Transport Geography*, 48:176–187.

Garin, R. A. (1966). Research Note: A Matrix Formulation of the Lowry Model for Intrametropolitan Activity Allocation. *Journal of the American Institute of Planners*, 32(6):361–364.

Garrido, S., Borysov, S. S., Pereira, F. C., and Rich, J. (2020). Prediction of rare feature combinations in population synthesis: Application of deep generative modelling. *Transportation Research Part C: Emerging Technologies*, 120.

Ge, Z., Liu, S., Wang, F., Li, Z., and Sun, J. (2021). YOLOX: Exceeding YOLO Series in 2021. *arXiv preprint arXiv:2107.08430v2*, 5:12.

Geng, K., Wang, Y., Cherchi, E., and Guarda, P. (2023). Commuter departure time choice behavior under congestion charge: Analysis based on cumulative prospect theory. *Transportation Research Part A: Policy and Practice*, 168:103564.

Ghahremannezhad, H., Shi, H., and Liu, C. (2023). Object Detection in Traffic Videos: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, 24(7):6780–6799.

Glenn Jocher (2022). YOLOv5.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*, volume 1. MIT Press, 1 edition.

Guerrieri, M., Parla, G., and Mauro, R. (2019). Traffic Flow Variables Estimation: An Automated Procedure Based on Moving Observer Method. Potential Application for Autonomous Vehicles. *Transport and Telecommunication*, 20(3):205–214.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. (2017). Improved Training of Wasserstein GANs. In I. Guyon and U. Von Luxburg and S. Bengio and H. Wallach and R. Fergus and S. Vishwanathan and R. Garnett, editor, *Advances in Neural Information Processing Systems*. Curran Associates, Inc.

Haghighat, A. K., Ravichandra-Mouli, V., Chakraborty, P., Esfandiari, Y., Arabi, S., and Sharma, A. (2020). Applications of Deep Learning in Intelligent Transportation Systems. *Journal of Big Data Analytics in Transportation*, 2(2):115–145.

Herrera, J. C. and Bayen, A. M. (2010). Incorporation of Lagrangian measurements in freeway traffic state estimation. *Transportation Research Part B: Methodological*, 44(4):460–481.

Ho, C. Q. and Hensher, D. A. (2016). A workplace choice model accounting for spatial competition and agglomeration effects. *Journal of Transport Geography*, 51:193–203.

Horni, A., Nagel, K., and Axhausen, K. W. (2016). The Multi-Agent Transport Simulation MATSim. *The Multi-Agent Transport Simulation MATSim*, page 618.

Ibeas, A., Cordera, R., Dell'Olio, L., and Coppola, P. (2013). Modelling the spatial interactions between workplace and residential location. *Transportation Research Part A: Policy and Practice*, 49:110–122.

Inoa, I. A., Picard, N., and de Palma, A. (2015). Effect of an Accessibility Measure in a Model for Choice of Residential Location, Workplace, and Type of Employment. *Mathematical Population Studies*, 22(1):4–36.

Kempinska, K. and Murcio, R. (2019). Modelling urban networks using Variational Autoencoders. *Applied Network Science*, 4(1):1–11.

Kim, E. J. and Bansal, P. (2023). A deep generative model for feasible and diverse population synthesis. *Transportation Research Part C: Emerging Technologies*, 148.

Kumar, A., Kashiyama, T., Maeda, H., and Sekimoto, Y. (2021a). Citywide reconstruction of cross-sectional traffic flow from moving camera videos. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 1670–1678. IEEE.

Kumar, M., Ray, S., and Yadav, D. K. (2021b). A Survey on Moving Object Detection in Video Using a Moving Camera for Smart Surveillance System. In

*Proceedings of International Conference on Machine Intelligence and Data Science Applications*, pages 241–253. Springer Nature Link.

Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature 2015 521:7553*, 521(7553):436–444.

Levine, J. (1998). Rethinking Accessibility and Jobs-Housing Balance. *Journal of the American Planning Association*, 64(2):133–149.

Liu, T. and Meidani, H. (2024). End-to-end heterogeneous graph neural networks for traffic assignment. *Transportation Research Part C: Emerging Technologies*, 165:104695.

Lowry, I. S. (1964). A MODEL OF METROPOLIS,. Technical report, RAND CORP SANTA MONICA CALIF.

Ma, L. and Srinivasan, S. (2015). Synthetic Population Generation with Multilevel Controls: A Fitness-Based Synthesis Approach and Validations. *Computer-Aided Civil and Infrastructure Engineering*, 30(2):135–150.

Ma, L. and Srinivasan, S. (2016). An empirical assessment of factors affecting the accuracy of target-year synthetic populations. *Transportation Research Part A: Policy and Practice*, 85:247–264.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Economics. New York: Academic Press.*

Mcfadden, D. (1977). Modelling the Choice of Residential Location. *Cowles Foundation Discussion Papers.*

Mo, J. and Gong, Z. (2023). Cross-City Multi-Granular Adaptive Transfer Learning for Traffic Flow Prediction. *IEEE Transactions on Knowledge and Data Engineering*, 35(11):11246–11258.

Munoz, L., Xiaotian Sun, Dengfeng Sun, Gomes, G., and Horowitz, R. (2004). Methodological calibration of the cell transmission model. In *Proceedings of the 2004 American Control Conference*, volume 1, pages 798–803. IEEE.

Naqavi, F., Sundberg, M., Västberg, O. B., Karlström, A., and Hugosson, M. B. (2023). Mobility constraints and accessibility to work: Application to Stockholm. *Transportation Research Part A: Policy and Practice*, 175.

Nepal, U. and Eslamiat, H. (2022). Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs. *Sensors (Basel, Switzerland)*, 22(2).

Neves, D. T., Alves, J., Naik, M. G., Proença, A. J., and Prasser, F. (2022). From Missing Data Imputation to Data Generation. *Journal of Computational Science*, 61.

Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., and Wang, Y. (2003). Review of road traffic control strategies. *Proceedings of the IEEE*, 91(12):2043–2065.

Parishwad, O., Jiang, S., and Gao, K. (2023). Investigating machine learning for simulating urban transport patterns: A comparison with traditional macro-models. *Multimodal Transportation*, 2(3):100085.

Pereira, R., Carvalho, G., Garrote, L., and Nunes, U. J. (2022). Sort and Deep-SORT Based Multi-Object Tracking for Mobile Robotics: Evaluation with New Data Association Metrics. *Applied Sciences*, 12(3):1319.

Prédhumeau, M. and Manley, E. (2023). A synthetic population for agent-based modelling in Canada. *Scientific Data*, 10(1).

Rahmani, M., Koutsopoulos, H. N., and Ranganathan, A. (2010). Requirements and potential of GPS-based floating car data for traffic management: Stockholm case study. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 730–735. IEEE.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2015). You Only Look Once: Unified, Real-Time Object Detection. *ArXiv*.

Rich, J. (2018). Large-scale spatial population synthesis for Denmark. *European Transport Research Review*, 10(2).

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature 1986 323:6088*, 323(6088):533–536.

Ryan, J., Maoh, H., and Kanaroglou, P. (2009). Population Synthesis: Comparing the Major Techniques Using a Small, Complete Population of Firms. *Geographical Analysis*, 41(2):181–203.

Sarri, P., Kaparias, I., Preston, J., and Simmonds, D. (2023). Using Land Use and Transportation Interaction (LUTI) models to determine land use effects from new vehicle transportation technologies; a regional scale of analysis. *Transport Policy*, 135:91–111.

Seo, T., Bayen, A. M., Kusakabe, T., and Asakura, Y. (2017). Traffic state estimation on highway: A comprehensive survey.

Seo, T., Kusakabe, T., and Asakura, Y. (2015). Traffic State Estimation with the Advanced Probe Vehicles Using Data Assimilation. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, volume 2015-October, pages 824–830. Institute of Electrical and Electronics Engineers Inc.

Sifringer, B., Lurkin, V., and Alahi, A. (2020). Enhancing discrete choice models with representation learning. *Transportation Research Part B: Methodological*, 140:236–261.

Smith, L., Beckman, R., and Baggerly, K. (1995). TRANSIMS: Transportation analysis and simulation system. Technical report, Los Alamos National Laboratory (LANL), Los Alamos, NM.

Sun, C. and Ritchie, S. G. (1999). Individual Vehicle Speed Estimation Using Single Loop Inductive Waveforms. *Journal of Transportation Engineering*, 125(6):531–538.

Sun, L. and Erath, A. (2015). A Bayesian network approach for population synthesis. *Transportation Research Part C: Emerging Technologies*, 61:49–62.

Sun, L., Erath, A., and Cai, M. (2018). A hierarchical mixture modeling framework for population synthesis. *Transportation Research Part B: Methodological*, 114:199–212.

Sunderrajan, A., Viswanathan, V., Cai, W., and Knoll, A. (2016). Traffic State Estimation Using Floating Car Data. *Procedia Computer Science*, 80:2008–2018.

Tampere, C. M. and Immers, L. H. (2007). An Extended Kalman Filter Application for Traffic State Estimation Using CTM with Implicit Mode Switching and Dynamic Parameters. In *2007 IEEE Intelligent Transportation Systems Conference*, pages 209–216. IEEE.

Torres, C., Hanley, N., and Riera, A. (2011). How wrong can you be? Implications of incorrect utility function specification for welfare measurement in choice experiments. *Journal of Environmental Economics and Management*, 62(1):111–121.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*, volume 2. Cambridge University Press.

Treiber, M. and Kesting, A. (2013). *Traffic Flow Dynamics*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Ua-areemitr, E., Sumalee, A., and Lam, W. H. (2019). Low-Cost Road Traffic State Estimation System Using Time-Spatial Image Processing. *IEEE Intelligent Transportation Systems Magazine*, 11(3):69–79.

United Nation (2018). World Urbanization Prospects 2018.

van Cranenburgh, S., Wang, S., Vij, A., Pereira, F., and Walker, J. (2022). Choice modelling in the age of machine learning - Discussion paper. *Journal of Choice Modelling*, 42.

Västberg, O. B., Karlström, A., Jonsson, D., and Sundberg, M. (2020). A dynamic discrete choice activity-based travel demand model. *Transportation Science*, 54(1):21–41.

Wang, Y., Zhang, D., Liu, Y., Dai, B., and Lee, L. H. (2019). Enhancing transportation systems via deep learning: A survey.

Wegener, M. (2004). Overview of Land Use Transport Models. *Handbook of Transport Geography and Spatial Systems*, pages 127–146.

Wegener, M. and Fuerst, F. (2004). Land-Use Transport Interaction: State of the Art. *SSRN Electronic Journal*.

Williamson, P., Birkin, M., and Rees, P. H. (1998). The Estimation of Population Microdata by Using Data from Small Area Statistics and Samples of Anonymised Records. *Environment and Planning A: Economy and Space*, 30(5):785–816.

Wilson, A. G. (1970). Entropy in urban and regional modelling. *Entropy in Urban and Regional Modelling*, pages 1–166.

Xu, L., Skoularidou, M., Cuesta-Infante, A., and Veeramachaneni, K. (2019). Modeling Tabular data using Conditional GAN. *Advances in Neural Information Processing Systems (NeurIPS)*, 32:7335–7345.

Yuan, T., Da Rocha Neto, W., Rothenberg, C. E., Obraczka, K., Barakat, C., and Turletti, T. (2022). Machine learning for next-generation intelligent transportation systems: A survey. *Transactions on Emerging Telecommunications Technologies*, 33(4):e4427.

Yuan, Y., Zhang, W., Yang, X., Liu, Y., Liu, Z., and Wang, W. (2021). Traffic state classification and prediction based on trajectory data. *Journal of Intelligent Transportation Systems*, pages 1–15.

Zhang, J., Wang, J., Zang, H., Ma, N., Skitmore, M., Qu, Z., Skulmoski, G., and Chen, J. (2024). The Application of Machine Learning and Deep Learning in Intelligent Transportation: A Scientometric Analysis and Qualitative Review of Research Trends. *Sustainability 2024, Vol. 16, Page 5879*, 16(14):5879.

Zhang, X., Feng, Y., Angeloudis, P., and Demiris, Y. (2022). Monocular Visual Traffic Surveillance: A Review. *IEEE Transactions on Intelligent Transportation Systems*, 23(9):14148–14165.

Zhu, Y. and Ferreira, J. (2014). Synthetic Population Generation at Disaggregated Spatial Scales for Land Use and Transportation Microsimulation. *Transportation Research Record: Journal of the Transportation Research Board*, 2429(1):168–177.