Multi-Modal Mobility Demand Generation for Urban Hubs: Preliminary Work

1st Yuri Frey Hudak
Faculty of Mathematics and Computer Sciences
Eindhoven University of Technology
Eindhoven, The Netherlands
y.f.hudak@tue.nl

1st Anup Vasu Padaki
Faculty of Mathematics and Computer Sciences
Eindhoven University of Technology
Eindhoven, The Netherlands
a.v.padaki@tue.nl

2nd Fabio Paparella
Faculty of Mechanical Engineering, Control Systems Technology
Eindhoven University of Technology
Eindhoven, The Netherlands
f.paparella@tue.nl

3rd Mauro Salazar
Faculty of Mechanical Engineering, Control Systems Technology
Eindhoven University of Technology
Eindhoven, The Netherlands
m.r.u.salazer@tue.nl

Abstract—The design of sustainable multi-modal mobility networks for urban hubs requires further understanding of the demand for different modes of transport and the relationship between demand and temporal, spatial, and external factors, such as weather. In this paper, we present a preliminary work on the generation of synthetic demand for taxi and shared bike services in urban hubs. We propose a conditional synthetic demand generation approach that uses a generative adversarial network to output synthetic taxi and shared bike demand for the New York City Metripolitan area conditioned on temporal and weather features. We also compare performance with two traditional variations of a multi-output regressor model for the purpose of providing a benchmark for the generative approach. Our preliminary results show that the proposed conditional synthetic demand generation approach is able to generate synthetic taxi and shared bike demand data conditioned on the basis of weather and temporal features, however further work is required to improve the quality of the generated data.

Index Terms—mobility, demand, synthetic data, taxi, bike, shared mobility, urban hubs, NYC

I. INTRODUCTION

A. Background:

As populations migrate away from rural areas and toward urban hubs, mobility networks are undergoing large-scale transformations in order to meet this increase in demand [1]. These transformations must be designed with sustainable objectives in mind, and a number of initiatives have investigated sustainable urban mobility systems from a sustainable design perspective [2] [3] [4]. Large amounts of funding have been dedicated to advancing sustainable urban mobility systems, for example the European Institute of Technology (EIT) recently dedicated 400 million Euros to investigate the topic [5].

Numerous initiatives are aimed at methods of incentivizing commuters to adopt multi-modal forms of transportation, for instance a combination of a subway ride followed by a short e-bike ride. The integration of micro-mobility systems within larger urban mobility networks can help cities meet increasing demand and reduce traffic congestion from automobiles in a cost-effective manner. The introduction of multi-modal micro-mobility systems such as e-scooters and e-bikes has also shifted the perception of what form future urban mobility systems might take. It has been estimated that micro-mobility providers will comprise a 300- 500 billion US dollar market by 2030 [6].

However, the complex nature of multi-modal mobility in densely-populated urban hubs makes it difficult to analyze different trends and preferences of the user toward a mode of transportation. The mode of transport that the user chooses is also dependent on various external factors, such as weather conditions and time of day, as well as factors internal to the mobility system, such as average wait time and total trip time [7] [8]. More accurate forecasting models of mobility systems have shown promise as a means of decreasing the negative impacts of these internal factors [9]. These models can help the development of optimal rebalancing and fleet management strategies for fleet vehicles, as well as aid further research in the dynamics of urban multi-modal mobility systems. Synthetic datasets which combine multi-modal mobility systems with external factors such as weather information can also help future research in the development of more effective algorithms to address these problems in mobility system design.

B. Literature Review:

In order to provide an overview of the current state of the art in multi-modal mobility demand, we review literature in two parts: the development of forecasting models for mobility demand, and the generation of synthetic data for mobility demand.

Forecasting:: There have been several investigations in the areas of short-term taxi demand forecasting using Auto-Regressive Integrated Moving Average (ARIMA), a time-series analysis method to predict short-term taxi demand [10]. With advances in high-performance computational power, Machine Learning and Deep Learning methods [11] have been more popular, and have yielded promising results [12] [13] [14] [15] [16]. Hybrid approaches where hierarchical clustering algorithms have also been demonstrated, for example which showcases a clustering algorithm based on density and distance used to construct virtual service points and a hybrid network based on convolutional and time-series models used for demand prediction. The usage of deep learning attention modules, and specifically transformers [17], have also paved the way to use architectures like BERT [18], where demand has been predicted by designing the network to learn varying taxi demand patterns between different functional regions as well as the trends in taxi demand on dynamic daily and weekly time scales. A multi-source information-based spatiotemporal neural network (MSI-STNN) deep learning architecture to predict short-term taxi demand has also been proposed [19]. This model fuses pick-up and drop-off time-series data, weather information, and location popularity data using three deep-learning models, including a stacked convolutional long short-term memory (ConvLSTM) model, a self-attention model, and a convolutional neural network (CNN) model. An end-to-end framework has been proposed in [13], which models the spatiotemporal correlations with semantic correlations in traffic networks. This approach models five types of spatial and semantic correlations into multi-location graphs, which are then fused by a multi-graph convolutional network. Temporal correlations between future and past taxi demand values are then captured by an LSTM in an encoder-decoder network. Multi-stage learning with LSTMs has also been used to capture pick-up and drop-off correlations together with spatiotemporal features to predict pick-up and drop-off demand [20]. Deep Reinforcement Learning has also been used in taxi demand prediction to reduce mean waiting times by integrating the derived context through analysis of zone-level taxi demands [21]. Travel context has also been modeled from different data sources to make demand predictions [22]. Some investigations have modeled point of interest (POI) regions to study taxi demand around urban regions with POIs, such as parks or popular destinations [18] [23] Generative Adversarial Networks (GANs) have also been used recently to solve time-series forecasting problems [24] [25]. Specifically, in [26] a stochastic spatio-temporal generative model (D-GAN) has been proposed which adopts a GAN-based structure for more accurate spatiotemporal prediction in multiple future time steps. This proposed network also supports the fusion of external factors through explicit objectives to improve the model learning process.

Synthetic Data Generation:: Synthetic data generation is a subdomain of data science which has multiple applications including data augmentation, data security, style transfer and manipulation, and signal super resolution [27]. As opposed to forecasting network architectures which are designed to predict data for future time steps, data generation architectures are intended to produce data which resembles the original distribution of the training data. Traditional large dataset generation methods for deep learning applications include the variational autoencoder architecture [28] [29] which has been superseded by the generative adversarial network (GAN) due to its improved performance, however recently the two architectures have been combined with promising results [30]. A traditional GAN network is comprised of two main components: the generator (G) and the discriminator (D) [24]. The former is an individual network which learns to generate synthetic data based as closely as possible on the statistical distribution of the real dataset. The latter is a separate individual network whose sole objective is to correctly identify output from the generator as either real or 'fake'. Together, the Generator and Discriminator essentially adhere to the 'zero-sum game' model from game theory in which it is theorized that both networks strive for a Nash equilibrium. The two are trained iteratively according to the min-max loss function

$$\min_G \max_D \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[1 - \log D(G(z))] \tag{1}$$

where $p_{\text{data}}(x)$ is the distribution of the real data, $p_z(z)$ is the distribution of the latent space, and $G(z)$ is the output of the generator network. The generator network is trained to minimize the loss function. The discriminator network is, conversely, trained to maximize the loss function. A more formal definition of the binary-crossentropy loss function and its use in the context of GANs is out of the scope of this paper, however the reader is directed to [24] for an in-depth discussion.

In general, there are no constraints on the underlying form that the generator or discriminator networks take, however best practices have emerged based on particular domains. The two networks are typically constructed with different architectures due to their unique objectives. Training is typically performed in mini-batched samples of the training dataset x and the prior noise distribution z. The network weights of G and D are updated using backpropagation, with only one network being trained at once. Architectures vary according to the type of data being generated, however deep convolutional networks are common and are typically used for image generation.

The development of GANs for spatiotemporal data is a relatively new area of research, but some promising models have emerged. Gao et. al. provide an extensive review on the subject of variations of GAN architecture for the spatiotemporal domain [31]. Saxena and Cao proposed a deep generative adversarial network-based model, called D-GAN,

which includes a deep spatiotemporal feature learning network and a fusion module for the addition of external factors such as weather, points of interest, and weekend vs. weekday [32].

The development of the conditional GAN (cGAN) builds upon the traditional GAN architecture with the addition of a layer for both the generator and discriminator which provides extra information, for example in the form of class labels or data from other modalities, which is embedded and combined with the original data [33]. In the generator network, the prior input noise $p_z(z)$ is combined with the conditional data in a joint hidden representation. Importantly, the conditional GAN allows the targeted generation of synthetic data based on desired class characteristics intrinsic to the latent space. In the context of urban mobility planning, this architecture provides the potential for targeted mobility demand generation based on time, location, weather, and other features. To the best of the knowledge of the authors, the development of a conditional GAN for the generation of mobility demand based on time, location, and weather classes has not yet been published.

C. Statement of Contribution:

Predicting multi-modal mobility demand (e.g. taxi, bike, etc.) is a difficult problem and it is dependent on numerous factors. Recent developments in deep learning demonstrate that taxi and bike demand can be predicted more accurately through more complex deep learning architectures, together with modules for fusing mobility features, weather information, and spatiotemporal features, similar to [34] [35]. Synthetic data generation for spatiotemporal data can be constrained through the addition of conditional class labels, and in this way high-dimensional data distributions can be learned and generated [33]. Therefore, the main contributions of this paper are as follows:

- A preliminary conditional GAN-based deep learning network to fuse spatiotemporal mobility features and weather features which can generate synthetic urban taxi and shared bike demand data for variable time periods conditioned on weather class, date, and time of day.
- A comparison of the proposed conditional GAN architecture with traditional feedforward deep learning regression architectures.

The aforementioned contributions are demonstrated through the application toward two specific mobility networks: the New York City (NYC) Yellow Cab system and the New York City CitiBike docked shared bicycle system. All work is available on our public GitLab repository [1].

D. Organization:

The remainder of this paper is structured as follows: Section II presents the Methodology, including the architecture of the preliminary conditional GAN and feedforward networks and a discussion. Section III presents the Results, including performance metrics for the proposed conditional GAN system against the benchmark architectures and findings. Section IV

---

[1]https://gitlab.tue.nl/20200365/mobilityforecast

presents the Conclusion, which discusses the most important outcomes and their interpretations as well as an outlook on applications of the results and proposals for future work.

II. PRELIMINARY MODELING INVESTIGATION

The following section presents a description of the design of the proposed conditional data generation algorithm, benchmarking algorithm designs, and a discussion of the methods.

A. System Design:

The system design encompasses the development of the entire conditional data generation algorithm, beginning with preprocessing of the datasets, design of the deep learning architecture, and finally methods for training and evaluating the model.

1) Data Preparation and Preprocessing: Data for the NYC Yellow Cab system were obtained from the NYC OpenData website [36], which is a free public data repository made available by New York City municipal agencies and partners. The selected dataset contained all trip records for the Yellow Cab system in PARQUET format for the entire 2021 year. The data includes features for pick-up and drop-off dates and times, locations, trip distances, fare information, and passenger counts. Data for the Citi Bike shared docked bike system were obtained from the Citi Bike NYC website [37], which is a data repository hosted by the Citi Bike initiative. The data includes features for unique ride identification, start and end time and location, and docking station IDs. The selected dataset contained all trip records for the Citi Bike system in CSV format for the entire 2021 year. Weather data was obtained from the Weather Underground website [38], which provides publicly available historical weather data on an hourly basis for weather stations within the Weather Underground network. The selected dataset contained hourly weather conditions for the KWO35 National Weather Service station over the entire 2021 year.

Custom data preparation methods were developed in order to read the raw datasets and translate them into consistent formats for use in the deep learning architecture. Please see Figure 1 for a graphical overview of the data processing pipeline. Erroneous data entries in the form of NaN values or empty values were eliminated from the datasets. For each feature, values that fell outside of 3 standard deviations from the mean were considered outliers and were removed. Exploratory data analysis was conducted to determine which features held more significance relative to taxi and bike demand than others. For the taxi and bike datasets, features for pick-up and drop-off times and locations were selected. For the weather dataset, features for time and three categorical weather conditions (clear, rainy, or snowy) were found to be the most significant.

In order to simplify the data space, taxi trip locational data was converted from unique zone ID (a format used by the NYC Taxi and Limousine Commission) to corresponding latitude and longitude, since the Citi Bike dataset was already presented with locations in the form of latitude and longitude data. This made correlation and structuring simpler. Both sets
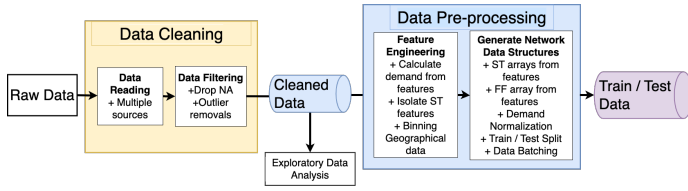
Fig. 1. Data processing pipeline for cGAN and regression networks.

of locational data were binned into discrete sets of latitude and longitude groups of equal dimension in order to make demand calculation simpler.

Demand for taxi and bike trips was calculated from the feature sets by aggregating trips on a monthly, daily, and hourly basis and counting the number of trips for each combination of pick-up and drop-off location for both taxi and bike data. For the cGAN network, the data was constructed into a 3D series of spatiotemporal arrays, with X and Y dimensions representing pickup and dropoff locations, values representing demand, and the Z dimension representing an array for each individual hour in the set. A unique 3D array was constructed for taxi and bike data, and both sets were engineered to match spatial bins and temporal bins to allow for synchronous generation of taxi and bike demand. A separate 3D 'class' array is constructed which contains the month and day as well as the weather conditions for each corresponding hour. A single class array is used for both taxi and bike data, since the two datasets were correlated along temporal and spatial dimensions. For the feedforward network, the data is constructed into a 2D array with a column for month, day, hour, unique pickup-dropoff combination, weather conditions, and finally taxi and bike demand. Each row represents an individual hour of data. Both sets of data for the cGAN and feedforward networks contain data for every hour over the course of one full year, in this case 2021.

2) Network Design: The conditional generative adversarial network (cGAN) consists of two main blocks: the generator and the discriminator, who together act as adversaries during the training process. An overview network architecture and training diagram can be seen in Figure 2. The generator has to goal of generating synthetic data which matches the training dataset distribution. The discriminator has the goal of discriminating whether the input data is 'real' (from the original distribution) or 'fake' (produced by the generator). In this preliminary work, the generator network design follows a traditional convolutional neural network (CNN) architecture, which was chosen because the data is essentially comprised of low resolution images. Throughout this text, the terms image and spatiotemporal array will be used interchangeably. The generator also utilizes an embedding layer for the class labels. The generator takes as inputs a latent space dimension and the number of classes. The generator outputs a single 'image' array (a 2D spatiotemporal array of demand for corresponding pickup and dropoff locations) based on the input class labels. The discriminator network design also follows a traditional
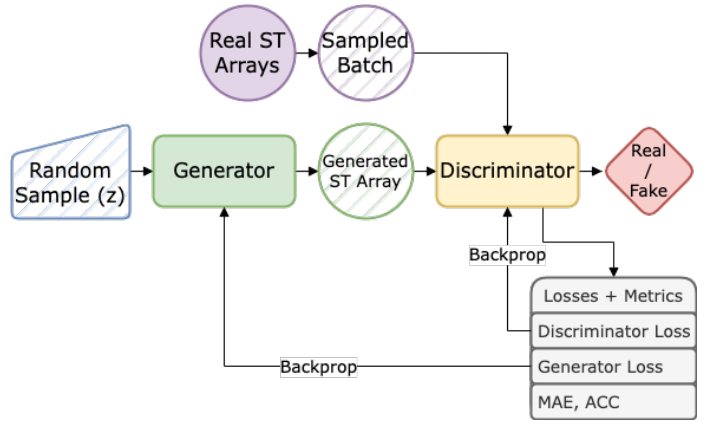

Fig. 2. GAN network architecture and training diagram.

CNN architecture with an embedding layer for the class labels. The discriminator takes as input an 'image' (a 2D spatiotemporal array) and the number of classes. The discriminator outputs a single sigmoid prediction in the range of 0 to 1, which predicts whether the input image was 'fake' or 'real' (synthetically generated or from the original distribution). Both of these networks are then combined in a simple cGAN network which takes as inputs the generator and discriminator models and tracks training metrics. The conditional inputs for both the generator and discriminator networks are the class labels for each of the temporal, spatial, and weather classes (8 in total). This input is constructed as a 1-dimensional array with a single value for each class label. An embedding layer in each network maps the 1D array to a 2D array, and this embedding layer is then fed to a fully connected layer, which is then reshaped and concatenated with the spatiotemporal array.

The feedforward regression network, shown in Figure 3, for this preliminary work is formulated as a nonlinear regression network with multiple outputs. It follows a traditional multi-layer perceptron (MLP) architecture, with sequential fully-connected layers which reduce the shape of the network space to two sigmoidal outputs which individually predict a value between zero and one for the normalized taxi and bike demand. The network takes as an input a 1D array containing an individual row from the 2D array of training features. A simple multi-output regression model was also constructed using the Scikit-learn library [39] for comparison.

3) Model Training: The cGAN network was trained using binary crossentropy as the loss function, which essentially gives the network the objective to maximize the probability that the spatiotemporal arrays output by the generator are assigned a value of 1, or 'real' by the discriminator. The performance metric computed for the cGAN network was mean absolute error (MAE), which applies to the discriminator's class predictions as well as the spatiotemporal array distributions output by the generator compared with the training data. The cGAN was trained for 50 epochs, and a separate model was trained for taxi and bike demand data. The custom feedforward regression network was trained using the mean
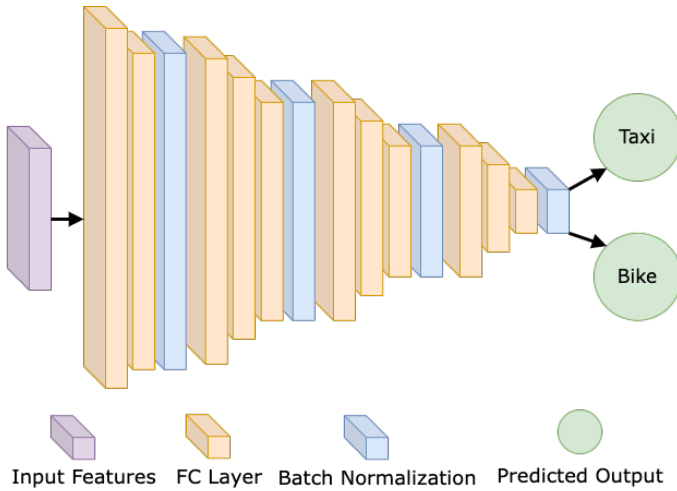
Fig. 3. Fully-connected feedforward regression network architecture.

Fig. 4. Sample of predicted hour of taxi demand with corresponding sample from the testing data.

Metric	GAN		Regression	
	Generator	*Discriminator*	*FF*	*Scikit-Learn*
BCE Loss	26.4	0.001	N/A	N/A
MAE	0.99	0.001	0.0027	0.0042

absolute error for predicted taxi and bike demand values as the loss function, which gave the network the objective to minimize the difference between predicted and real values for taxi and bike demand. Mean absolute error was also used as the performance metric. The custom feedforward regression network was trained for 50 epochs at a learning rate of 0.1 for both taxi and bike demand data simultaneously. The multi-output regression model from Scikit-learn was also fit to the data prepared for the custom regression network and MAE was used as the performance metric.

B. Results and Discussion:

A summary of performance metric results for the cGAN model and the custom regression model is presented in Table I. As an additional point of comparison, a simple multi-output Regressor model from the Scikit-Learn Python library [39] was fit to the data used for the custom regression network, and its performance metrics are also included in Table I. Both the custom regression network model and the multi-output Regressor model performed reasonably well at predicting values for taxi and bike demand at single points in time. The cGAN taxi and bike models both showed poor performance at the task of generating spatiotemporal arrays of pickup-dropoff demand. A sample of a generated array for an hour of both taxi and bike demand as well as a corresponding sample from the data used for testing the models can be seen in Figure 4. This poor performance could be due to the large degree of sparsity seen in the training data, for which convolutional-based neural network architectures typically underperform. Future work will investigate methods for addressing this inherent sparsity.

Further context is warranted to clarify some design choices for the reader. First, the structures given to the training data for both the cGAN network and the feedforward regression network were designed according to best practices for both formulation approaches, however there is room for exploration in both cases. Binning of the spatial data into discrete containers of latitude and longitude values is a hyperparameter that will
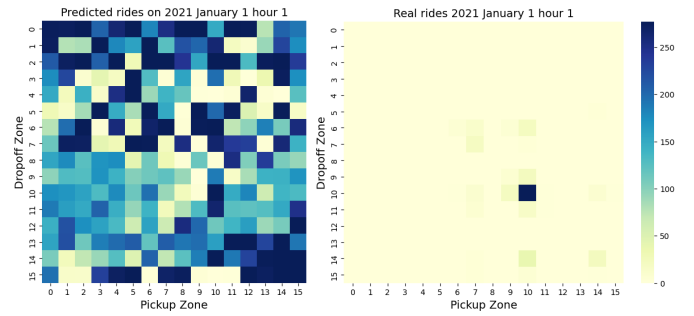
be explored further in future work. In the present preliminary investigation, the spatial data was discretized into 16 bins. Future work will investigate higher resolution bins as well as multi-scale bins of variable resolution in urban areas with higher expected demand. In addition, performance comparison of the two different network architectures was limited to the mean absolute error metric for the output demand values for the cGAN and feedforward networks. However, due to the high contrast between the two architectures and their respective learning processes, the authors will investigate more advances methods of performance comparison in future work. An additional choice made for the training process of the cGAN model was the decision to train the generator network 10 times for every iteration of training the discriminator network. This was done in order to compensate for behavior seen in previous training attempts, where the discriminator clearly dominated the adversarial relationship between the two networks and the generator was unable to produce spatiotemporal arrays capable of fooling the discriminator. This parameter will be investigated further in future work. Lastly, it is important to clarify an assumption made during the data processing and feature engineering stage of this preliminary investigation: the weather features were assumed to be consistent across all spatial dimensions. Due to the small geographical area being considered in the data, the authors chose to simplify the space of the conditional classes for weather by assigning conditions from a single weather station located in the Manhattan neighborhood to all locations at each time step.

III. CONCLUSION

This paper explored preliminary deep learning architectures for the generation of synthetic data for taxi and bike demand in the NYC Metropolitan area across spatial and temporal dimensions together with weather conditions. A preliminary conditional generative adversarial network was

designed, trained, and tested. For the purpose of contextual comparison, a traditional multi-layer perceptron feedforward regression network was also designed, trained, and tested on the task of prediction of taxi and bike demand values, formulated as a nonlinear regression problem. The performance of the resulting models demonstrated the need for further investigation of the architecture for the conditional GAN network, especially with regard to handling the inherent sparsity of the data. However, the cGAN architecture showed promise in its ability to handle complex multidimensional spatiotemporal data and fuse it with weather conditions and date-time labels for the conditional generation of synthetic demand data. Upon further development, this approach has the potential to provide significant aid in the research of multi-modal mobility networks. This preliminary investigation will be extended as follows: first, we would like to explore more complex architectures for the generator and discriminator blocks of the cGAN network. Architectures that allow for better handling of sparse data will be investigated, similar to the investigations presented in [40] [41] [11]. Second, we will investigate the potential for the cGAN network to be used as a non-deterministic forecasting tool for predicting future taxi and bike demand, rather than generating demand for the time period of the training data. Third, it is of interest to further explore the effect of design choices and hyperparameters including the use of location-dependent weather conditions as well as spatial bin resolution, multi-scale binning, and the ratio of generator to discriminator training steps per iteration.

REFERENCES

[1] Secretary General. Sustainable cities, human mobility, and international migration. Technical report, United Nations Economic and Social Council, April 2018.

[2] Omid Ghorbanzadeh, Sarbast Moslem, Thomas Blaschke, and Szabolcs Duleba. Sustainable Urban Transport Planning Considering Different Stakeholder Groups by an Interval-AHP Decision Support Model. *Sustainability*, 11(1):9, January 2019. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[3] Fatih Canitez, Pelin Alpkokin, and Sabahat Topuz Kiremitci. Sustainable urban mobility in Istanbul: Challenges and prospects. *Case Studies on Transport Policy*, 8(4):1148–1157, December 2020.

[4] Juan de Dios Ortúzar. Sustainable Urban Mobility: What Can Be Done to Achieve It? *Journal of the Indian Institute of Science*, 99(4):683–693, December 2019.

[5] EIT Urban Mobility. https://eit.europa.eu/our-communities/eit-urban-mobility.

[6] The future of micromobility: Ridership and revenue after a crisis | McKinsey. https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/the-future-of-micromobility-ridership-and-revenue-after-a-crisis.

[7] Lawrence Frank, Mark Bradley, Sarah Kavage, James Chapman, and T. Keith Lawton. Urban form, travel time, and cost relationships with tour complexity and mode choice. *Transportation*, 35(1):37–54, January 2008.

[8] Edward A. Beimborn, Michael J. Greenwald, and Xia Jin. Accessibility, Connectivity, and Captivity: Impacts on Transit Choice. *Transportation Research Record*, 1835(1):1–9, January 2003. Publisher: SAGE Publications Inc.

[9] Ramon Iglesias, Federico Rossi, Kevin Wang, David Hallac, Jure Leskovec, and Marco Pavone. Data-Driven Model Predictive Control of Autonomous Mobility-on-Demand Systems. Technical Report arXiv:1709.07032, arXiv, September 2017. arXiv:1709.07032 [cs, stat] type: article.

[10] Luis Moreira-Matias, João Gama, Michel Ferreira, João Mendes-Moreira, and Luis Damas. On Predicting the Taxi-Passenger Demand: A Real-Time Approach. In Luís Correia, Luís Paulo Reis, and José Cascalho, editors, *Progress in Artificial Intelligence*, Lecture Notes in Computer Science, pages 54–65, Berlin, Heidelberg, 2013. Springer.

[11] Jonas Uhrig, Nick Schneider, Lukas Schneider, Uwe Franke, Thomas Brox, and Andreas Geiger. Sparsity Invariant CNNs, August 2017. arXiv:1708.06500 [cs].

[12] Chizhan Zhang, Fenghua Zhu, Yisheng Lv, Peijun Ye, and Fei-Yue Wang. MLRNN: Taxi Demand Prediction Based on Multi-Level Deep Learning and Regional Heterogeneity Analysis. *IEEE Transactions on Intelligent Transportation Systems*, 23(7):8412–8422, July 2022.

[13] Genxuan Hong, Zhanquan Wang, Taoli Han, and Hengming Ji. Spatiotemporal Multi-Graph Convolutional Network for Taxi Demand Prediction. In *2021 11th International Conference on Information Science and Technology (ICIST)*, pages 242–250, May 2021. ISSN: 2573-3311.

[14] Dongchang Liu, Jia Mou, Yu Liu, and Yiping Yang. Improved Prediction of High Taxi Demand: A Deep Spatiotemporal Network for Hyper-imbalanced Data. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–8, September 2020.

[15] Filipe Rodrigues and Francisco C. Pereira. Beyond Expectation: Deep Joint Mean and Quantile Regression for Spatiotemporal Problems. *IEEE Transactions on Neural Networks and Learning Systems*, 31(12):5377–5389, December 2020.

[16] Yongxuan Lai, Kaixin Zhang, Junqiang Lin, Fan Yang, and Yi Fan. Taxi Demand Prediction with LSTM-Based Combination Model. In *2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, pages 944–950, December 2019.

[17] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need, December 2017. arXiv:1706.03762 [cs].

[18] Dun Cao, Kai Zeng, Jin Wang, Pradip Kumar Sharma, Xiaomin Ma, Yonghe Liu, and Siyuan Zhou. BERT-Based Deep Spatial-Temporal Network for Taxi Demand Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 23(7):9442–9454, July 2022.

[19] Ziyan Chen. Multi-Source Information Based Short-term Taxi Demand Prediction Using Deep-Learning Approaches. *Journal of Physics*, page 15, 2021.

[20] Chizhan Zhang, Fenghua Zhu, Xiao Wang, Leilei Sun, Haina Tang, and Yisheng Lv. Taxi Demand Prediction Using Parallel Multi-Task Learning Model. *IEEE Transactions on Intelligent Transportation Systems*, 23(2):794–803, February 2022.

[21] Z. Liu, J. Li, and K. Wu. Context-Aware Taxi Dispatching at City-Scale Using Deep Reinforcement Learning. *IEEE Transactions on Intelligent Transportation Systems*, 23(3):1996–2009, 2022. Publisher: Institute of Electrical and Electronics Engineers Inc.

[22] Lingbo Liu, Zhilin Qiu, Guanbin Li, Qing Wang, Wanli Ouyang, and Liang Lin. Contextualized Spatial–Temporal Network for Taxi Origin-Destination Demand Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 20(10):3875–3887, October 2019.

[23] Bahman Askari, Tai Le Quy, and Eirini Ntoutsi. Taxi Demand Prediction using an LSTM-Based Deep Sequence Model and Points of Interest. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 1719–1724, July 2020. ISSN: 0730-3157.

[24] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks, June 2014. arXiv:1406.2661 [cs, stat].

[25] Congmei Jiang, Yize Chen, Yongfang Mao, Yi Chai, and Mingbiao Yu. Forecasting Spatio-Temporal Renewable Scenarios: a Deep Generative Approach, March 2019. arXiv:1903.05274 [cs, math].

[26] D. Saxena and J. Cao. 12D Spatio-Temporal Prediction with Stochastic Adversarial Networks. *ACM Transactions on Intelligent Systems and Technology*, 13(2), 2022. Publisher: Association for Computing Machinery.

[27] H Surendra. A Review Of Synthetic Data Generation Methods For Privacy Preserving Data Publishing, 2017.

[28] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic Backpropagation and Approximate Inference in Deep Generative Models, May 2014. arXiv:1401.4082 [cs, stat].

[29] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes, May 2014. arXiv:1312.6114 [cs, stat].

[30] Anders Boesen Lindbo Larsen, Søren Kaae Sønderby, Hugo Larochelle, and Ole Winther. Autoencoding beyond pixels using a learned similarity metric, February 2016. arXiv:1512.09300 [cs, stat].

[31] Nan Gao, Hao Xue, Wei Shao, Sichen Zhao, Kyle Kai Qin, Arian Prabowo, Mohammad Rahaman, and Flora Salim. Generative Adversarial Networks for Spatio-temporal Data: A Survey. *ACM Transactions on Intelligent Systems and Technology*, 13:1–25, April 2022.

[32] Divya Saxena and Jiannong Cao. D-GAN: deep generative adversarial nets for spatio-temporal prediction. *CoRR*, abs/1907.08556, 2019.

[33] Mehdi Mirza and Simon Osindero. Conditional Generative Adversarial Nets, November 2014. arXiv:1411.1784 [cs, stat].

[34] Xian Zhou, Yanyan Shen, Yanmin Zhu, and Linpeng Huang. Predicting Multi-step Citywide Passenger Demands Using Attention-based Neural Networks. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, WSDM '18, pages 736–744, New York, NY, USA, February 2018. Association for Computing Machinery.

[35] Ignacio-Iker Prado-Rujas, Emilio Serrano, Antonio García-Dopico, M. Luisa Córdoba, and María S. Pérez. Combining heterogeneous data sources for spatio-temporal mobility demand forecasting. *Information Fusion*, 91:1–12, March 2023.

[36] NYC Open Data -. https://opendata.cityofnewyork.us/.

[37] Citi Bike System Data | Citi Bike NYC. https://ride.citibikenyc.com/system-data.

[38] Weather History & Data Archive | Weather Underground. https://www.wunderground.com/history.

[39] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[40] Charlie Nash, Jacob Menick, Sander Dieleman, and Peter W. Battaglia. Generating Images with Sparse Representations, March 2021. arXiv:2103.03841 [cs, stat].

[41] Yan Yan, Yuxing Mao, and Bo Li. SECOND: Sparsely Embedded Convolutional Detection. *Sensors*, 18(10):3337, October 2018.