1. Title:

**Building a Large-Scale Micro-Simulation Transport Scenario Using Big Data**

Citation:

Schweizer, J., Poliziani, C., Rupi, F., Morgano, D. and Magi, M., (2021) Building a large-scale micro-simulation transport scenario using big data. ISPRS International Journal of Geo-Information, 103.

Link to the paper:

<https://doi.org/10.3390/ijgi10030165>

Description:

Large-scale microsimulations can be seen as an interdisciplinary project where transport planners and technology developers can work together on the same scenario; big data from OpenStreetMap, traffic surveys, GPS traces, traffic counts, and transit details are merged into a unique transport scenario. The employed activity-based demand model can simulate and evaluate door-to-door trip times while testing different mobility strategies. Indeed, a utility-based mode choice model is calibrated to match the official modal split. The scenario is implemented and analyzed with the software SUMOPy/SUMO which is an open-source software, available on GitHub. The simulated traffic flows are compared with flows from traffic counters using different indicators. The determination coefficient has been 0.7 for larger roads (width greater than seven meters). The present work shows that it is possible to build realistic microsimulation scenarios for larger urban areas. A higher precision of the results could be achieved by using more coherent data and by merging different data sources. The road network data contains the directed road network graph made of links and nodes; each link consists of one or several lanes. The most important lane attributes are maximum speed, width, and access rights; all the values are determined by analyzing the OSM attributes of the respective way. Moreover, SUMO assigns a priority level to each link which depends on the link attributes and ranges from 1 (footpath) up to 13 (national motorway). The connectivity of lanes at intersections is also derived from OSM or guessed from heuristics; all connections have been manually checked, together with road attributes and geometry. The concept of mobility strategies allows adding any kind of new technology or service. SUMO with SUMOPy enables easy access to microsimulations, edit scenarios, and track all simulation events, step by step, through a user-friendly interface, and a rich spectrum of analysis tools.

Even though the present scenario-building is a special case and leaves ample room for improvements, it starts narrowing the gap between different research areas and allows planners, data scientists, and technology developers to work together more effectively on the same transport scenario with the common goal to realistically evaluate and improve future sustainable transport systems.

1. Title:

**Analysis of Classifier Training on Synthetic Data for Cross-Domain Datasets**

Link to the paper:

<https://ieeexplore.ieee.org/abstract/document/9151398>

Citation:

A. Cortés, C. Rodríguez, G. Vélez, J. Barandiarán and M. Nieto, "Analysis of Classifier Training on Synthetic Data for Cross-Domain Datasets," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 1, pp. 190-199, Jan. 2022, doi: 10.1109/TITS.2020.3009186.

keywords: {Training;Data models;Machine learning;Pipelines;Detectors;Vehicles;Synthetic datasets;deep learning;traffic sign recognition},

Description:

A major challenge of deep learning (DL) is the necessity to collect huge amounts of training data. Often, the lack of a sufficiently large dataset discourages the use of DL in certain applications. Typically, acquiring the required amounts of data costs considerable time, material and effort. To mitigate this problem, the use of synthetic images combined with real data is a popular approach, widely adopted in the scientific community to effectively train various detectors. In this study, we examined the potential of synthetic data-based training in the field of intelligent transportation systems. The proposed augmentation pipeline of synthetic datasets includes novel augmentation processes such as structured shadows and gaussian specular highlights. A well-known DL model was trained with different datasets to compare the performance of synthetic and real image-based trained models. Synthetic images are generated using a semi-supervised errors-guide method which is also described. Our experiments showed that a synthetic image-based approach outperforms in most cases real image-based training when applied to cross-domain test datasets (+10% precision for GTSRB dataset) and consequently, the generalization of the model is improved decreasing the cost of acquiring images.

1. Title:

**Generating synthetic mobility data for a realistic population with RNNs to improve utility and privacy.**

Link to the paper: <https://dl.acm.org/doi/abs/10.1145/3477314.3507230>

Description:

The data often represent a limited sample of the population and use of the data jeopardizes privacy.

To address these issues, we present and evaluate a system for generating synthetic mobility data using a deep recurrent neural network (RNN) which is trained on real location data. The system takes a population distribution as input and generates mobility traces for a corresponding synthetic population.

1. Title:

**Visualization of Traffic Density on Graph using Trajectory Data**

Link to the paper: <https://ieeexplore.ieee.org/document/10074196>

Citation:

S. More, A. Matange and J. Abraham, "Visualization of Traffic Density on Graph using Trajectory Data," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2022, pp. 2357-2361, doi: 10.1109/ICAC3N56670.2022.10074196. keywords: {Couplings;Data privacy;Publishing;Roads;Data visualization;Feature extraction;Trajectory;Trajectory data;trajectory data mining;utility;road graph;junction-based trajectory data},

Description:

Several applications collect users’ trajectory data to provide better location-based services. The current trajectory data mining procedures take all the raw information from the trajectory data without considering the users’ privacy. This type of data mining can be prone to linkage attacks. This research proposes a novel mechanism that uses OpenStreetMap and its geo-entities recognition feature to extract and cluster the traffic signals to form vertices of a graph. From the trajectory data, vehicle type and the edge in the graph is determined. The vehicle count for each vehicle type on the edge of a graph is iteratively updated and it is visually represented on the respective road edge between two junction vertices. The visualization of the weighted road network graph enables analysts to infer useful information as the road graph itself preserves the statistical information of the road density but hides the individual users’ trajectory.