Proceedings of the Seventh International Conference on
Artificial Intelligence, Soft Computing, Machine Learning and Optimization,
in Civil, Structural and Environmental Engineering
Edited by: P. Iványi, J. Kruis and B.H.V. Topping
Civil-Comp Conferences, Volume 11, Paper 3.1
Civil-Comp Press, Edinburgh, United Kingdom, 2025
ISSN: 2753-3239, doi: 10.4203/ccc.11.3.1

Develop a Street Speed Bump Extraction and Mapping Framework From Street Level Imagery Using Deep Learning

M. Abdel Karim[1], A. Alazmi[2] and T. Alhadidi[3]

[1]School of Business and Economics, Vrije Universiteit Amsterdam, Netherland
[2]Department of Construction Project, Ministry of Public Works of Kuwait, South Surra, Kuwait
[3]Civil Engineering Department, Al-Ahliyya Amman University, Jordan

Abstract

Developing smart infrastructure needs innovative road detection solutions to monitor road conditions. The location of speed bumps must be carried out by institutions so that they do not cause accidents and have negative consequences for road users. The aim of this research is to define a framework to detect the location of bumps using deep learning. This paper is proposing an automated way to extract, map, and geo-enable street speed bumps from street view imagery. Using machine learning computer vision-based models demonstrated that street view imagery can provide efficient, high-quality, and cost-effective solutions for large-scale mapping of street speed bumps which can be extended to include other street furniture types. The proposed framework demonstrates superior performance in accuracy and coverage metrics. It achieves an average precision score of 0.93.

Keywords: machine learning, street view imagery, GIS, street speed bump, GPS, ground truth formalisation, sustainable road.

1 Introduction

Many research papers have addressed the increased use of street view imagery for different applications [1], [2], [3], [4]. Some of these identified key areas are related to the environment, the landscape, for example, greenery, and street features. The researchers highlighted the role of machine learning, which tries to make computers learn using structures inferred from data. The tools opened the door for more research in this field [1], [5], [6], [7]. The researchers also expressed that the applications of street-level imagery have evolved from image classification, scene segmentation, and location identification to the physical understanding of the surrounding environment [1]. The researchers confirmed that cost-effectiveness, open access, and global coverage can be considered benefits of using street-level imagery to retrieve street information [8], [9], [10]. Few research papers focusing on feature extraction from street-level imagery were identified. For example, Laumer et al. (2020) proposed a cost-effective solution using street-level imagery to update tree inventories by reverse geocoding trees and matching them with their respective inventory records [8]. Other researchers use street-level imagery to extract traffic lights and telephone poles [11], [12]. Sundaraperumal et al. (2024) used live photos captured from moving automobiles and image files to create a real-time detection system to recognize speed bumps. They create a YOLO algorithm that can locate and identify multiple objects (ex., Puddles, speed bumps, and lighting changes) within an image [11]. On the other hand, the driver's behavior can be used to detect objects on the road. Mosleh et al. (2024) create a system that can identify driver behaviors by monitoring road conditions and define driving patterns. They developed an ML model that can anticipate driving over speed bumps and circular yellow speed bumps [12]

Focusing on shortlisted papers which address the same or similar type of application regarding street-level features, it was noticed that applying deep learning models to extract features from images (trees, traffic lights, light poles) was common, however the researcher apply different machine learning models to achieve the same purpose [3], [8], [11], [12], [13]. On the other hand, the researcher shows different ways for the geo-enablement of features, for example Krylov et al. (2018) proposed a novel MRF-based process for geo-tagged features which relies on the depth estimate for the extracted features[11]. Another applied a mathematical (photogrammetry) based calculation to identify the traffic sign location based on camera position, similarly [12]. Laumer et al. (2020) applied a mathematical solution to calculate the distance between the camera position and detected trees, and they have considered the tree inventory records as their ground truth data [8].

The random distribution of road bumps may have negative consequences for road users. Smart cities and digital twins are getting more attention in many countries and government organizations. The spatial aspect of these initiatives plays a crucial role in their success and failure. This recognized importance of spatial data opened the door for the need for new data layers which were not necessarily part of GIS data repositories. At the same time, creating these new datasets using the traditional tools and reference data can be either expensive for example conducting field surveys, or cannot be achieved using remote sensing data like satellite and aerial imagery due to cost, image resolution, coverage, and other factors. Due to budget and resources limitation to get a higher satellite image resolution, or conduct field surveys to collect street speed bumps, and the outdated available paper maps, the proposed framework was created.

2 Methods

The theoretical framework that used to qualify the feasibility of using street-level imagery to extract and document street speed bumps and use it to enrich the GIS data repository divided into the following parts to address street-level imagery usability from different angles.

2.1 Image Data Allocation and Gathering

Getting controlled street-level imagery that can be used to build the processing workflow will help in identifying the key factors that might contribute to the results. This means having all parameters related to the capture device and image properties, such as resolution and focal length, consistency, and others.

The state of Kuwait held a comprehensive street-level image capture project that covers nearly the whole country. This data was collected for different years and using different cameras, and Google Street View did not have the same coverage for the country. However, this research used the captured data from the State of Kuwait street-level imagery as a sample dataset to assess the proposed processing workflow for the following reasons:

- Different capturing devices were used
- Seasonality of capturing, as these images were captured multiple times
- Known capturing devices parameters, and configuration
- Different image resolutions and quality
- Availability of panoramas along with their respective image planners

Figure 1: Kuwait street-level imagery.

Figure 1 shows an example of different images of the pumps from various distances.

2.2 Ground Truth Formalization

Establishing ground truth data that will be used to qualify the accuracy of the extracted data is needed to assess and quantify the quality of the extracted data. The data was surveyed and measured using RTK-based high-accuracy GPS devices. At the same time, distance measuring benchmarks were established physically during the camera capture to assess the quality of distance calculations during data extraction from imagery.

2.3. Selecting and Fine-Tuning Deep Learning Model

Deep learning models were used to extract image features representing the desired data from the imagery. The selection of the model depends on the expected results. Two candidate models were used: instance segmentation and object detection models. Each one has its advantages and disadvantages, but selecting the model will define how the data for the model is prepared.
A monocular depth estimation model was used as a second stage for image processing; this model was used to infer the detected object's distance from the image.
Applying the deep learning workflow consisted of two stages: object detection for image features and depth inference for detected objects.

2.4. Data Labeling

The type of features required to be extracted from the street-level imagery was identified and used to create the labeled ground truth data using the image sample data. The labeled data should be representative enough to do model training and gain the expected results. The data labeling method relies on the selected deep learning model, which was used for feature extraction.

2.5. Data Postprocessing

The extracted data from the deep learning model includes attributes about the feature bounding box and inferred distance. This data was processed to convert from the image coordinate to the GIS coordinate space.
The Python programming language was used for data processing. The postprocessing workflow involved:

- Converting features coordinates, and distance attributes.
- Geo-locating features based on inferred distance and camera location.
- Create a weighted distance matrix for the different feature classes for data clustering.
- Applying the clustering method, DBSCAN, in extracted features to eliminate feature duplication based on the weighted clustering matrix created (Figure 3).
- Quantify and apply benchmark analysis between the extracted features and the established ground truth survey data for quality assessment.

- Convert and create GIS feature layers for the corresponding extracted features.

Figure 2: DBSCAN clustering (a: shows the bump detection before clustering; b: shows the bump detections after clustering).

3 Results

This section presents the outcomes of applying deep learning techniques to detect street pump locations from the collected dataset. This research used the captured data from the State of Kuwait street-level imagery as a sample dataset. Our best-performing model achieved an average precision score of **0.93**, indicating strong performance in accurately identifying pump locations.

The methodology was developed to determine the most effective approach for accurate and reliable detection. Visualizations of prediction outputs were provided to illustrate the effectiveness of the proposed approach further. Figure 2 shows the training loss curve. The training loss curve measures the dissimilarity between the model's projected output and the actual output, providing us with information about how the model's performance evolves over time.

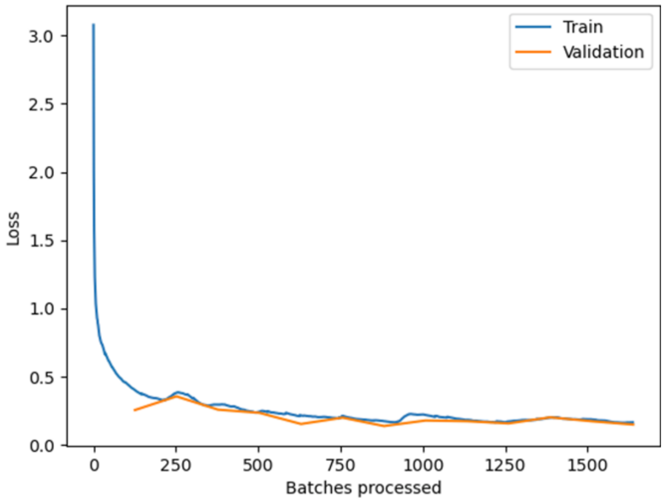

Figure 3: Training loss curve.

Figure 3 illustrates the training and validation loss curves over the course of model training using a learning rate of **0.0003**. As shown, both losses decreased steadily during the initial epoch, indicating effective learning. The training loss continues to decline consistently, while the validation loss begins to plateau and stabilize after approximately batch 700. This suggests that the model is learning generalized features without overfitting. The smooth convergence of both curves demonstrates that the selected learning rate of 0.0003 was appropriate for ensuring stable and efficient optimization.

The accuracy of the speed bumps' location was analyzed. The distribution of the distance difference between the prediction and the ground truth bumps shows that the mean absolute error of the speed bump location was 4.62 m.

Figure 4: The bump locations labeling.

Figure 4 shows the model label location of the bumps. It is clear that some of the modelss can detect the location of the bumps within 4 meters. The model

4 Conclusions and Contributions

It is necessary to have a digital version of road data which was not updated for a long time. The random distribution of the road bumps may effect negative consequences for road users.

The data should be available as a GIS layer so that the decision maker can take an actions based on the field findings. Due to budget, and resources limitation to get a higher satellite image resolution, or conduct field surveys to collect street speed bumps, and the outdated available paper maps, the proposed framework was created.

The limitation of resource availability and the increase in the demand for detailed data and data conversion time are factors that each decision-maker faces in their day-to-day tasks. Successful integration of street-level imagery with the current different GIS data sources will open the door for large-scale mapping projects with reduced cost, effort, resources, details, and time.

Using street-level imagery as a new data source for GIS data proved it promising, especially regarding the data details extracted from the profile view that street-level imagery can provide. These details can help to create new data layers which cannot be extracted from the current remote sensing for example satellite imagery due to different reasons.

Applying new emerging technologies like deep learning can help speed up and automate the data extraction process. The big challenge is making the extracted data geo-enabled to be utilized in GIS with acceptable positional accuracy.

Acknowledgement

We wish to acknowledge the support of Kuwait Foundation for the Advancement of Sciences (KFAS).

References

[1] N. He and G. Li, "Urban neighbourhood environment assessment based on street view image processing: A review of research trends," *Environmental Challenges*, vol. 4, 2021, doi: 10.1016/j.envc.2021.100090.

[2] X. Li, C. Zhang, W. Li, R. Ricard, Q. Meng, and W. Zhang, "Assessing street-level urban greenery using Google Street View and a modified green view index," *Urban For Urban Green*, vol. 14, no. 3, 2015, doi: 10.1016/j.ufug.2015.06.006.

[3] J. H. Kim, S. Lee, J. R. Hipp, and D. Ki, "Decoding urban landscapes: Google street view and measurement sensitivity," *Comput Environ Urban Syst*, vol. 88, 2021, doi: 10.1016/j.compenvurbsys.2021.101626.

[4] A. Abu-Khadrah, A. Al-Qerem, M. R. Hassan, A. M. Ali, and M. Jarrah, "Drone-assisted adaptive object detection and privacy-preserving surveillance in smart cities using whale-optimized deep reinforcement learning techniques," *Sci Rep*, vol. 15, no. 1, p. 9931, 2025.

[5] A. Alazmi and H. Rakha, “Assessing and Validating the Ability of Machine Learning to Handle Unrefined Particle Air Pollution Mobile Monitoring Data Randomly, Spatially, and Spatiotemporally,” *Int J Environ Res Public Health*, vol. 19, no. 16, 2022, doi: 10.3390/ijerph191610098.

[6] A. Alazmi and B. S. Al-Anzi, “Assessment of Machine Learning Algorithms for Predicting Air Entrainment Rates in a Confined Plunging Liquid Jet Reactor,” *Sustainability (Switzerland)*, vol. 15, no. 18, 2023, doi: 10.3390/su151813802.

[7] F. Alazemi, A. Alazmi, M. Alrumaidhi, and N. Molden, “Predicting Fuel Consumption and Emissions Using GPS-Based Machine Learning Models for Gasoline and Diesel Vehicles,” *Sustainability*, vol. 17, no. 6, p. 2395, Mar. 2025, doi: 10.3390/su17062395.

[8] D. Laumer, N. Lang, N. van Doorn, O. Mac Aodha, P. Perona, and J. D. Wegner, “Geocoding of trees from street addresses and street-level images,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 162, 2020, doi: 10.1016/j.isprsjprs.2020.02.001.

[9] S. Zou and L. Wang, “Detecting individual abandoned houses from google street view: A hierarchical deep learning approach,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 175, 2021, doi: 10.1016/j.isprsjprs.2021.03.020.

[10] J. Al-Nabulsi, A. Mesleh, and A. Yunis, “Traffic light detection for colorblind individuals,” in *2017 IEEE Jordan conference on applied electrical engineering and computing technologies (AEECT)*, IEEE, 2017, pp. 1–6.

[11] V. A. Krylov, E. Kenny, and R. Dahyot, “Automatic discovery and geotagging of objects from street view imagery,” *Remote Sens (Basel)*, vol. 10, no. 5, 2018, doi: 10.3390/rs10050661.

[12] A. Campbell, A. Both, and Q. (Chayn) Sun, “Detecting and mapping traffic signs from Google Street View images using deep learning and GIS,” *Comput Environ Urban Syst*, vol. 77, 2019, doi: 10.1016/j.compenvurbsys.2019.101350.

[13] T. Alhadidi, A. Jaber, S. Jaradat, H. I. Ashqar, and M. Elhenawy, “Object Detection using Oriented Window Learning Vi-sion Transformer: Roadway Assets Recognition,” Jun. 2024.