



ORIGINAL PAPER

Open Access



Automating intersection marking data collection and condition assessment at scale with an artificial intelligence-powered system

Kun Xie^{1*} , Huiming Sun², Xiaomeng Dong¹, Hong Yang³ and Hongkai Yu²

Abstract

Intersection markings play a vital role in providing road users with guidance and information. The conditions of intersection markings will be gradually degrading due to vehicular traffic, rain, and/or snowplowing. Degraded markings can confuse drivers, leading to increased risk of traffic crashes. Timely obtaining high-quality information of intersection markings lays a foundation for making informed decisions in safety management and maintenance prioritization. However, current labor-intensive and high-cost data collection practices make it very challenging to gather intersection data on a large scale. This paper develops an automated system to intelligently detect intersection markings and to assess their degradation conditions with existing roadway Geographic information systems (GIS) data and aerial images. The system harnesses emerging artificial intelligence (AI) techniques such as deep learning and multi-task learning to enhance its robustness, accuracy, and computational efficiency. AI models were developed to detect lane-use arrows (85% mean average precision) and crosswalks (89% mean average precision) and to assess the degradation conditions of markings (91% overall accuracy for lane-use arrows and 83% for crosswalks). Data acquisition and computer vision modules developed were integrated and a graphical user interface (GUI) was built for the system. The proposed system can fully automate the processes of marking data collection and condition assessment on a large scale with almost zero cost and short processing time. The developed system has great potential to propel urban science forward by providing fundamental urban infrastructure data for analysis and decision-making across various critical areas such as data-driven safety management and prioritization of infrastructure maintenance.

Keywords Intersection markings, Artificial intelligence, Data acquisition, Degradation condition assessment, Infrastructure management

1 Introduction

Intersection markings play a vital role in providing road users with guidance and information. Maintaining an accurate inventory of intersection markings is essential

for effective transportation management. According to the Federal Highway Administration's (FHWA's) program on the Model Inventory of Roadway Elements (MIRE), roadway data, including intersection elements, are critical to data-driven highway safety management (Lefler et al., 2017). Specifically, MIRE's gap analysis has identified that existing roadway inventories have large gaps in intersection descriptors such as type and number of exclusive left turn lanes, right turn channelization, and presence of crosswalk (Mallela et al., 2012). Meanwhile, the conditions of intersection markings will be gradually degrading due to vehicular traffic, rain, and/or snowplowing. Degraded markings can confuse drivers, leading

*Correspondence:

Kun Xie
kxie@odu.edu

¹ Department of Civil and Environmental Engineering, Old Dominion University, Norfolk, VA 23529, USA

² Department of Electrical Engineering and Computer Science, Cleveland State University, Cleveland, OH 44115, USA

³ Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, VA 23529, USA



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

to increased risk of traffic crashes. Timely obtaining high-quality information of intersection markings lays a foundation for making informed decisions in safety management and maintenance prioritization.

However, many states do not process an up-to-date statewide inventory and condition information of traffic assets because the high cost of data collection offsets the benefit of having such information (Balali et al., 2015). Traffic asset data are generally collected either by field investigation or computer-based manual extraction from aerial images, street views and video logs, and both of these data collection approaches are cost prohibitive (Proulx et al., 2015). Current labor-intensive and high-cost data collection practices make it very challenging to gather intersection data on a large scale (Fiedler et al., 2013). Road markings are among the traffic assets that can easily deteriorate over time, making it even more costly to keep track of their latest conditions. To collect statewide marking data and to prioritize the replacement need have created a demand for a cost-effective and scalable tool that can efficiently and accurately track the classifications, geographic locations, and conditions of road markings.

This study aims to develop an automated and scalable system powered by artificial intelligence (AI) for urban infrastructure data collection. The system can fully automate the processes of marking data collection and condition assessment on a large scale with almost zero cost and short processing time (e.g., in a preliminary test, the processing time per intersection is less than 2 s). Urban science is a multidisciplinary domain centered around leveraging data, technology, and analytical methods to tackle complex urban challenges. In this context, the study holds significant potential for advancing urban science by introducing innovative methodologies for collecting urban infrastructure data. The system's ability to generate extensive datasets in a cost-effective manner can profoundly impact urban science in multiple critical areas:

1.1 Improves the inventory of roadway data elements

The system offers a highly cost-effective tool to enhance current roadway inventory databases while supplying fundamental data elements crucial for advancing urban science.

1.2 Advances intersection safety management

The system can provide transportation agencies demanding data for Highway Safety Improvement Program (HSIP). The availability of large-scale intersection marking data (e.g., presence of crosswalks, dedicated left-turn lanes, etc.) enables agencies to use the analytic methods provided in the American Association of State Highway

and Transportation Officials' (AASHTO's) Highway Safety Manual (HSM). It helps bridge the gaps in current modeling practices by offering critical data to support safety decision making in hotspot identification and before-after safety evaluation.

1.3 Enables infrastructure maintenance prioritization

It is estimated that state agencies spend more than \$1 billion annually in maintaining road markings in the United States and Canada (Zhang & Ge, 2012). The developed system can allow agencies to monitor the conditions of a large number of markings for better allocation of resources and timely maintenance.

1.4 Augments intelligent transportation systems (ITS)

The developed system can produce detailed intersection profiles for supporting ITS applications such as the development of high-resolution digital maps, driver-assistance systems, and safety warning systems.

1.5 Supports transportation planning modeling

The generated intersection data can help transportation planners develop more accurate planning models by incorporating detailed information on intersection configurations.

2 Literature review

Though road marking data are generally collected manually in practice, there are research efforts devoted to automating the process. Image processing techniques were widely used to identify road markings such as image segmentation (Senlet & Elgammal, 2012), geometric parameter optimization (Foucher et al., 2011) and edge detection (Ahmetovic et al., 2015). The template matching method (Liu et al., 2012; Wu & Ranganathan, 2012) was also used for road marking recognition. Despite the fast speed image processing and template matching methods can offer, their decisions rely on empirical functions, which are difficult to be generalized in a changing environment (Chen et al., 2015; Vokhidov et al., 2016). More adaptive methods are learning-based such as k-nearest neighbors (KNN) (Rebut et al., 2004), support vector machine (SVM) (Greenhalgh & Mirmehdi, 2015; Sukhwani et al., 2014), random forest (Smith et al., 2013) and artificial neural network (ANN) (Máttyus et al., 2016; Yamamoto et al., 2014).

More recent advances include the exploitation of deep learning methods that have capability to autonomously learn discriminative features from image data. For instance, Vokhidov et al. (2016) found convolutional neural network (CNN) could better recognized lane-use arrows in various environments. Wen et al. (2019) also used CNN to classify different types of road markings

with considerable differences. R-CNN (Region-based Convolutional Neural Network), proposed by Girshick et al. (2014), can not only recognizing what objects are present but also determining their precise locations by drawing bounding boxes around them. It combined selective search for region proposals and a CNN for feature extraction. R-CNN achieved impressive accuracy but was computationally expensive due to its sequential processing of regions, making it impractical for real-time applications. R-CNN was utilized by Tian et al. (2020) to detect lane-use arrows and while/yellow lane lines. Their results showed that R-CNN could robustly extract road markers under various complex traffic scene. Fast R-CNN (Girshick, 2015) addressed the computational inefficiency of R-CNN by introducing the concept of region-of-interest (ROI) pooling. It allowed feature extraction from the entire image in a single forward pass, significantly speeding up the process. Fast R-CNN demonstrated improved accuracy and efficiency over its predecessor, making it more practical for real-world applications. Qian et al. (2016) employed Fast R-CNN to detect road surface traffic signs including lane-use markings to assist automated driving.

Compared with marking recognition, much less research focused on the automatic assessment of marking conditions. Burrow et al. (2000) determined the extent of erosion by comparing present road markings with the “ideal” ones. Both Zhang and Ge (2012) and Lin et al. (2016) used image processing techniques to capture characteristics of markings such as geometric deformity, colors and edge lines and then to determine the quality level of markings.

There are several limitations of existing studies. Firstly, most learning-based methods for marking recognition

are customized for driving assistance instead of inventory management, so they use small and local datasets and are not suitable for large-scale data collection. Secondly, most existing approaches for marking recognition are still sensitive to noises on road markings such as occlusion, illumination variations and worn-out conditions. Thirdly, condition assessment of markings is still under-examined. Existing methods rely on image processing techniques and more robust and adaptive methods are needed. Fourthly, previous studies either focus on marking recognition or condition assessment, there is no integrated method available which can optimize the whole data collection process and reduce computation time. Thus, there is an immediate need to develop a more optimal and economical solution for marking data collection on a large scale.

3 Methodology

3.1 An overview of the system

This section presents an overview of the AI-powered system for intersection marking data collection. You can find a demonstration of the system at the following link: <https://youtu.be/fvHf1H7i8Wo>. Figure 1 illustrates the conceptual design of the system. The system focuses on two types of markings at intersections – lane-use arrows and crosswalks, while it has the flexibility to be extended to cover other road markings as well. The system economically utilizes roadway geographic information systems (GIS) data and aerial images as inputs, which are commonly available from transportation agencies or open sources. The use of GIS data enables fast indexing and identification of intersections and accelerate the process of aeriels image data extraction, making the proposed approach truly

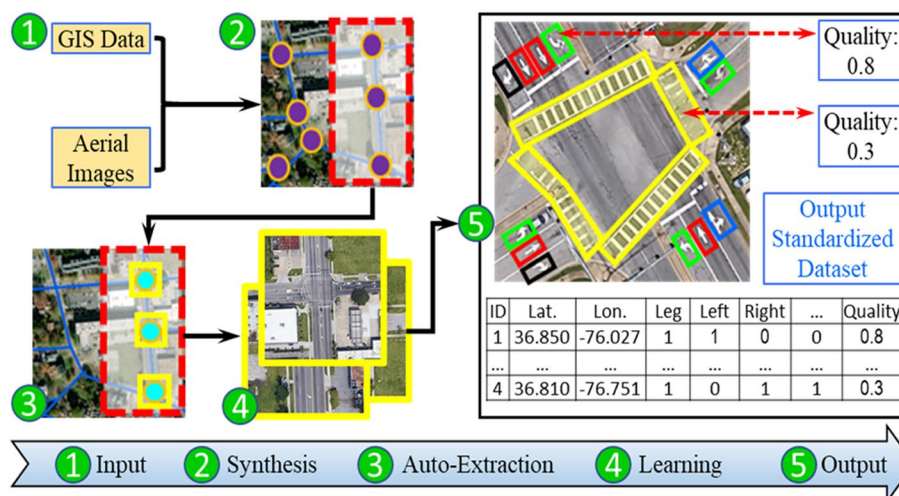


Fig. 1 Conceptual illustration of the automated system for intersection marking data collection

scalable and computationally efficient. The synthesis process entails the matching of geographic coordinates between the intersection GIS data and aerial images, allowing for auto-extraction the corresponding intersection images. The extracted intersection image data were used to train a novel computer vision model for detection, characterization, and condition assessment of intersection markings. Emerging AI techniques were harnessed to improve accuracy, robustness, and computational efficiency of the system. This system will be the foundation of future expansions to collect other roadway features such as medians and driveways to support additional data needs.

The system has innovatively addressed the limitations of existing data collection approaches from the following aspects:

1. *Seamless integration of spatial analytics with AI techniques.* With the help of existing intersections' locations, AI techniques can easily have the advantages in recognizing visual patterns. Spatial analytics helps pinpoint intersections in the target area and auto-extract their aerial images. Incorporating the spatial information can be the catalyst to greatly reduce the efforts in image segmentation and object recognition, and thus it makes the data collection process truly scalable and computationally efficient.
2. *Smart application of deep learning for condition assessment.* Humans are sensitive to visual impairments of markings, but it is very costly to apply subjective assessment on a large scale. The system leverages deep learning to generate quality scores consistent with human viewers. The multi-scale deep features of markings are fed into a regression sub-network to produce quality scores to indicate their degradation conditions.
3. *Multi-task learning for higher accuracy and computational efficiency.* The system creatively performs the joint tasks of intersection marking detection, characterization, and condition assessment in an end-to-end deep learning model. Model can better learn a new task by transferring the knowledge it has acquired by learning a related task. The simultaneous accomplishment of multiple tasks ensures its computational efficiency and inference performance for large-scale data collection practices.
4. *Enhanced system accessibility and reproducibility.* Despite the equipped advanced spatial analytics and AI components, the system has no prerequisite of knowledge and skills in imaging processing and GIS tools, and therefore enables more users to access it. In addition, it provides objective measurements for reproducible data collection.

3.2 Annotation of intersection aerial images

An annotation tool of Computer Vision Annotation Tool (CVAT) was tested and used to manually label the types of lane-use arrows (i.e., left, right, left & straight, right & straight, and straight) and crosswalks (i.e., transverse, zebra, and ladder) and their degradation conditions (i.e., low-quality and high-quality). Markings are categorized as high-quality if they are intact without any visible damage. Conversely, if a marking exhibits any form of damage or deterioration, it is classified as low-quality. Prior to data collection, all assessors underwent a thorough training session to become well-acquainted with both the annotation tool and the data collection protocol. Evaluation was conducted initially to ensure the integrity and consistency of the collected data. An example of annotation results is shown in Fig. 2.

3.3 Lane-use arrow detection

The Faster R-CNN (Ren et al., 2015) model is an object detection model that improves on Fast R-CNN by using a region proposal network (RPN) with the CNN model. The RPN shares full-image convolutional features with the detection network, enabling nearly cost-free region proposals. It's a fully convolutional network that simultaneously predicts object bounds and objectness scores at each position. The RPN is trained end-to-end to generate high-quality region proposals, which are then used by Fast R-CNN for detection. As a whole, Faster R-CNN consists of two modules: a deep fully convolutional network that proposes regions, and the Fast R-CNN detector that uses the proposed regions. The Faster RCNN model was used to detect and classify lane-use arrows (five categories: Left, Left & Straight, Straight, Right, Right & Straight) in the satellite images. The network structure is shown in Fig. 3. The backbone to extract image feature is the convolutional neural network with 16 layers (VGG16) (Simonyan & Zisserman, 2014).

3.4 Crosswalk detection

The crosswalks were classified into three types according to the Manual on Uniform Traffic Control Devices (MUTCD) standards as shown in Fig. 4.

Most crosswalk markings are arbitrary-oriented, horizontal bounding boxes used for the detection of lane-use arrows are no longer suitable. A deep learning model capable of detecting rotated objects is needed. The Box Boundary-Aware Vectors (BBAVectors) model (Yi et al., 2021) was used for oriented object detection in aerial images with Box Boundary-Aware Vectors. The BBAVectors model resulted in an outstanding performance in the Large-scale dataset for object detection in aerial images (DOTA) dataset (Xia et al., 2018),



Fig. 2 Overall annotated image with land-use arrows, crosswalks, and degradation conditions labeled

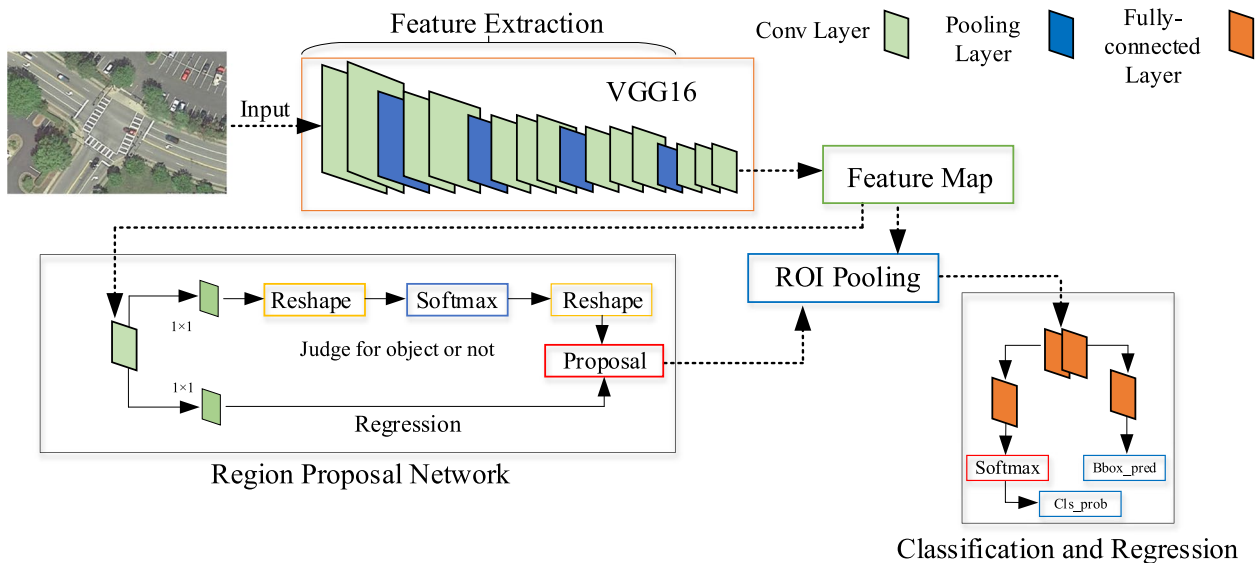


Fig. 3 Faster RCNN model for lane-use arrow detection (modified based on Ren et al. (2015))

which is a benchmark dataset for oriented object detection in computer vision. The BBAVectors model is used for detecting arbitrary-oriented objects, such as crosswalk markings in this case. This model is built upon the CenterNet (Duan et al., 2019), extending it for the oriented object detection task. The BBAVectors use a simple yet effective strategy to describe the

Oriented Bounding Box (OBB). They are measured in the same Cartesian coordinate system for all the arbitrarily oriented objects, achieving better performance than the baseline method that learns the width, height, and angle of the OBBs. The model is single-stage and anchor box free, which makes it fast and accurate. The network structure of BBA Vectors is shown in Fig. 5.

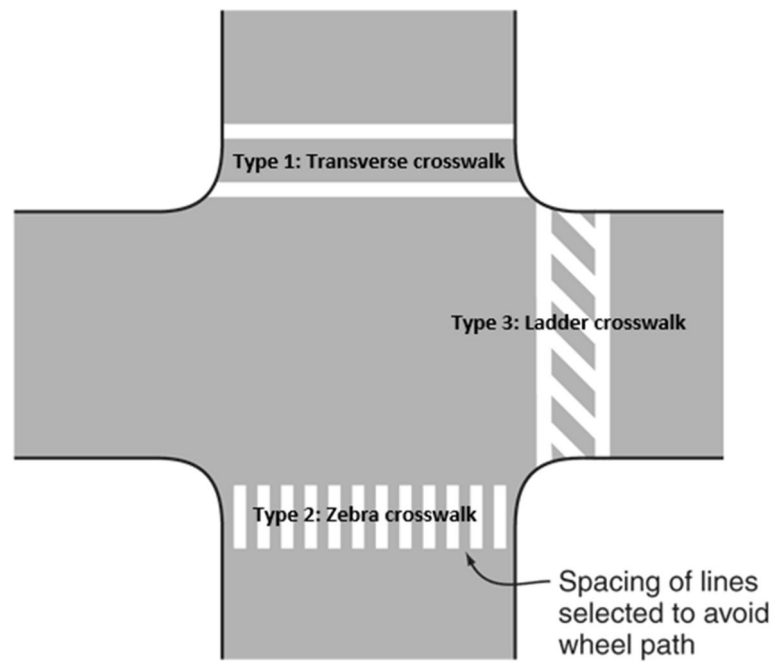


Fig. 4 Examples of crosswalk markings (T1—Transverse Crosswalk: crosswalk marker with two parallel solid white lines; T2—Zebra Crosswalk: crosswalk marker with a series of closely spaced solid white lines; T3 – Ladder Crosswalk: crosswalk marker with solid white lines between two parallel solid white lines) (FHWA, 2009)

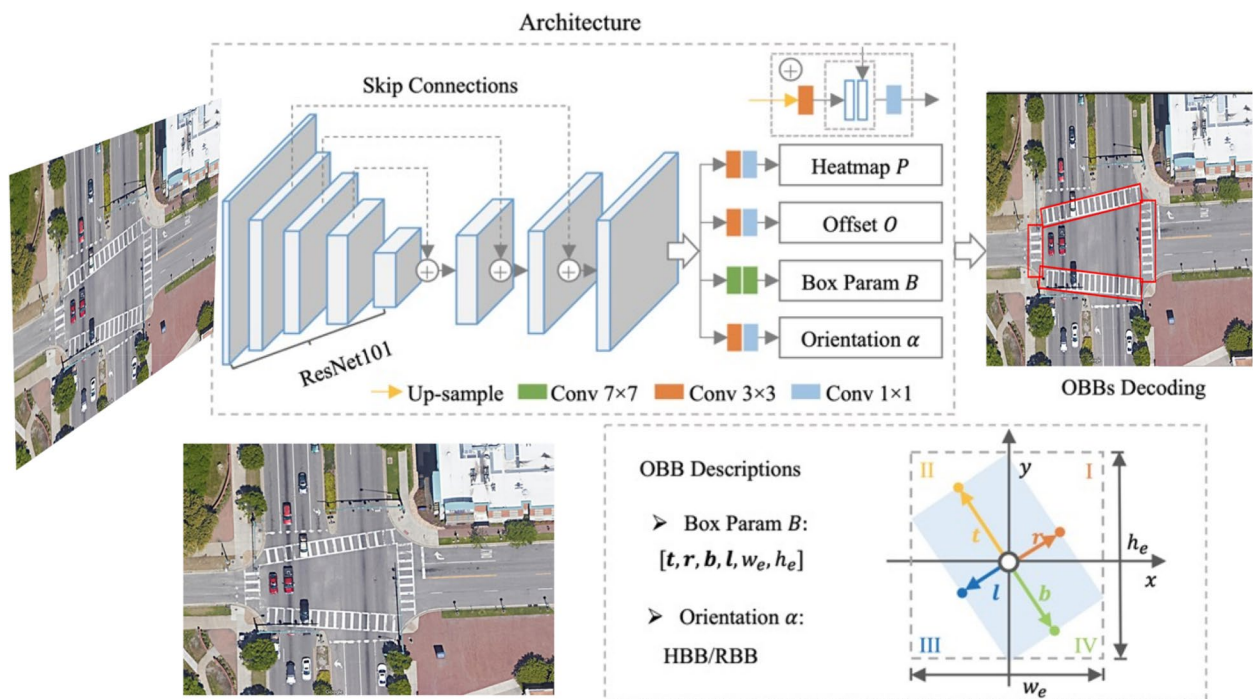


Fig. 5 BBA Vectors network structure (modified based on Yi et al. (2021))

3.5 Degradation condition assessment

Degradation conditions of markings were first manually annotated into two quality classes, i.e., low-quality and high-quality. If a marking (a lane-use arrow or crosswalk) is complete without any visible damage, it is classified as high-quality, otherwise it is classified as low-quality. Examples of low-quality and high-quality markings are presented in Fig. 6.

Kang et al. (2014) used a convolutional neural network for image quality assessment. A deep convolutional neural network model VGG16 (Simonyan & Zisserman, 2014) was developed for quality assessment. The quality score, which represents the estimated probability of a marking belonging to the high-quality category as determined by VGG16, was utilized to assess marking conditions. Quality scores range from 0 (indicating the lowest quality) to 1 (indicating the highest quality), providing a measure of marking degradation levels. VGG16 has great flexibility to learn the perception of human viewers on degradation conditions. The structure of VGG16 is presented in Fig. 7.

4 System development

4.1 System structure

Figure 8 illustrates the architecture of the system, which consists of two main components: the backend and the frontend. The backend is responsible for deploying a system that facilitates the transmission of results from the vision component to the frontend. Conversely, the

frontend is designed to display the outcomes and provide a user interface for seamless interaction.

4.2 Backend

The FastAPI (Lathkar, 2023) framework was selected as the foundation for constructing the backend system. FastAPI is a contemporary, efficient, and web-based framework designed for creating Application Programming Interfaces (APIs) using Python 3.6+ and relies on standard Python type hints. In the backend, intersection images serve as input and are processed through a computer vision module and an output module. The computer vision module performs the detection of lane-use arrows and crosswalks while assessing their degradation conditions. The resulting outputs consist of labeled intersection images and.csv files containing comprehensive marking information.

4.3 Frontend

The frontend of the web-based system was developed to provide users with a graphical user interface (GUI) for viewing and interacting with the system. JavaScript was utilized to create dynamic elements on static Hyper Text Markup Language (HTML) web pages. The Mapbox API was employed to retrieve aerial images of intersections based on the coordinates provided by users. The interface features four buttons: Input, Start, End, and Output. The Input button allows users to enter the location of the intersections, the Start button

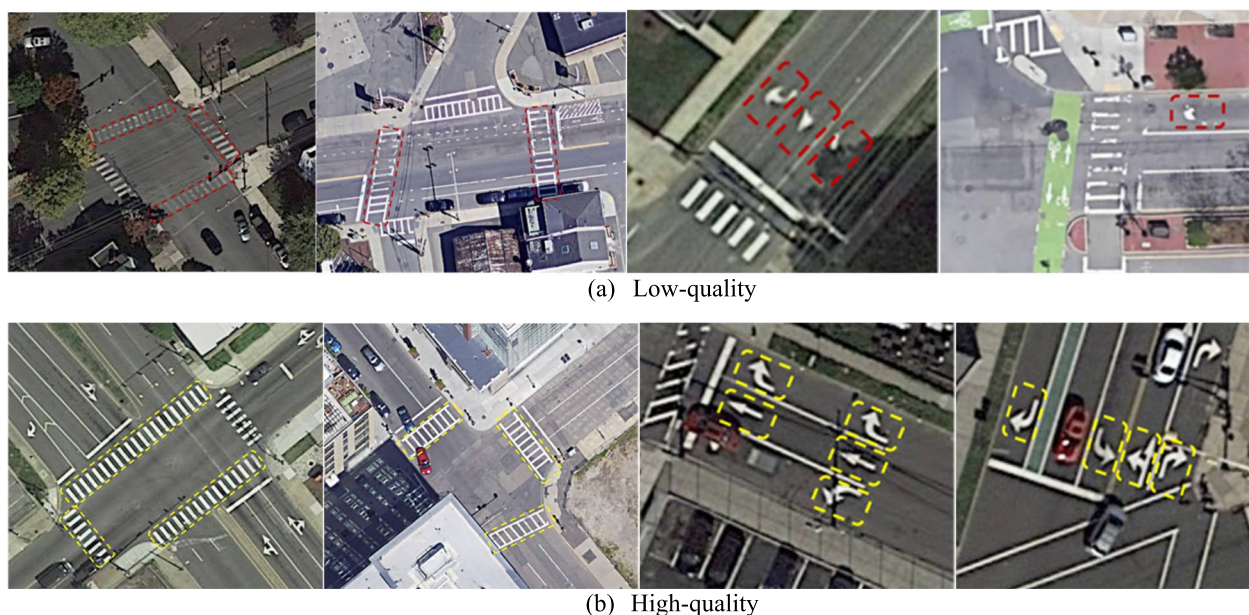


Fig. 6 Examples of low-quality (a) and high-quality (b) degradation conditions

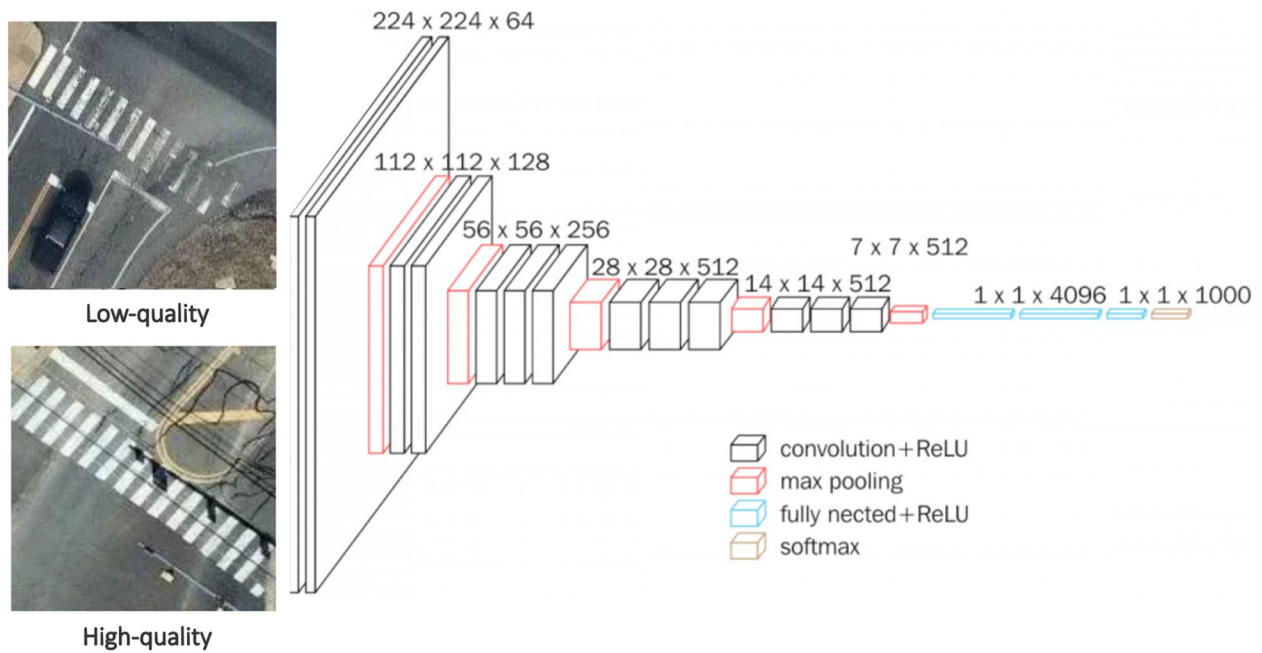


Fig. 7 The VGG16 network structure for image quality assessment (modified based on Simonyan and Zisserman (2014))

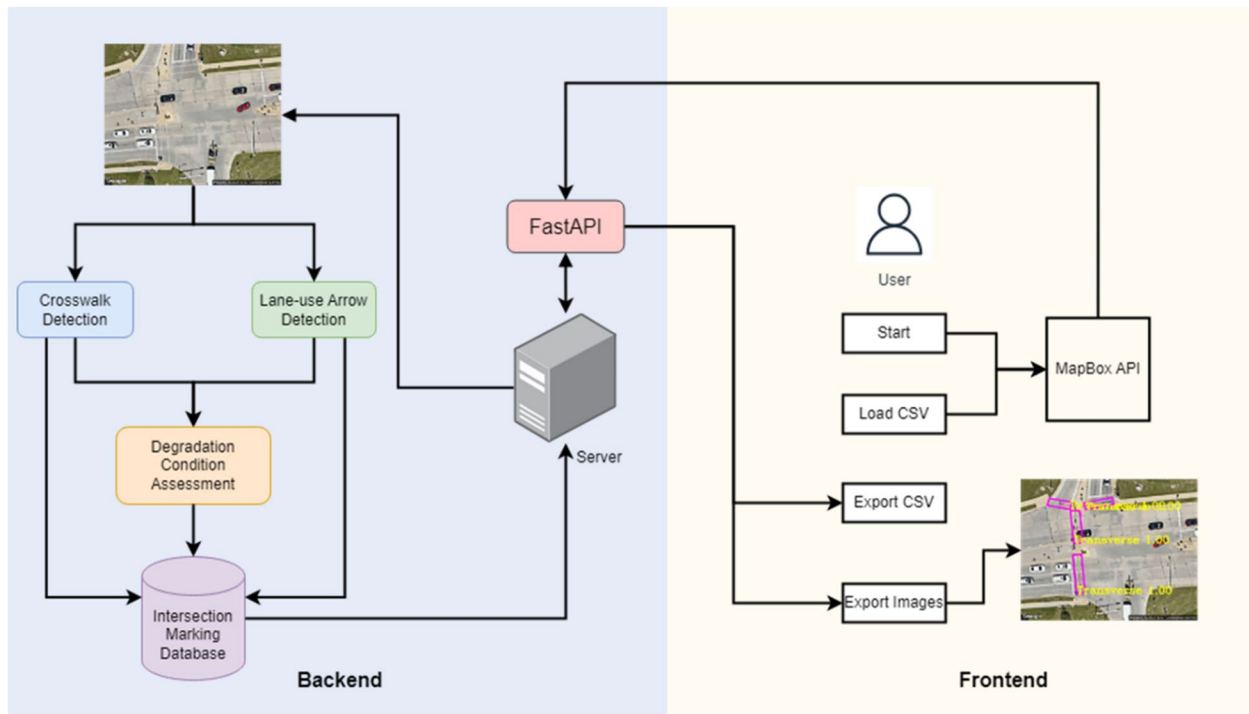


Fig. 8 System structure

Table 1 Input data format

Intersection_ID	Latitude	Longitude
1,269,888	36.8201493	-75.99975
1,085,129	36.7805423	-76.141486
1,045,006	36.7744851	-76.133553
1,037,605	36.7802918	-76.142283
541,180	36.8372126	-76.159457
541,789	36.8785388	-76.141027
...

initiates the processing, the End button halts the process, and the Output button enables the export of data.

4.4 Input, graphical user interface, and output

The input data contains intersection coordinate information and is tabulated in common.csv format. An example

of the input data derived from LRS Road Intersections (VDOT, 2017) is shown in Table 1. There are three columns including *Intersection_ID*, *Latitude*, and *Longitude*.

The graphical user interface of the system prototype is shown in Fig. 9. You can find a demonstration of the system at the following link: <https://youtu.be/fvHf1H7i8Wo>.

A sample output file in.csv format is presented in Fig. 10, with its field description listed in Table 2. The users have the option to output labeled images data for verification purposes as shown in Fig. 11.

4.5 Programming packages and analytical tools

For programming packages, the vision algorithm was made use of PyTorch (Paszke et al., 2019), a popular deep learning framework, to build and train the AI models. PyTorch provides a flexible and efficient platform for developing neural networks and conducting deep learning tasks. Additionally, other essential



Fig. 9 Graphical user interface of the system prototype

Intersection_ID	Latitude	Longitude	Marking_ID	Type	SubType	Location	Quality	Length	Width
541355	36.81321	-76.125	1	Crosswalk	Zebra	[839, 354, 785, 385, 445, 69, 499, 37]	0.77	75.4	10.4
541355	36.81321	-76.125	2	Crosswalk	Ladder	[839, 354, 785, 385, 445, 69, 499, 37]	0.75	75.4	10.4
541355	36.81321	-76.125	3	Crosswalk	Ladder	[491, 933, 436, 966, 99, 673, 153, 640]	0.99	72.6	10.6
541355	36.81321	-76.125	4	Crosswalk	Ladder	[491, 46, 201, 587, 138, 566, 427, 25]	0.72	102.7	10.8
541355	36.81321	-76.125	5	Crosswalk	Ladder	[828, 392, 606, 967, 546, 952, 769, 376]	1	103.3	10
541355	36.81321	-76.125	6	Arrow	Left	[183, 67, 183, 124, 222, 124, 222, 67]	0.98		
541355	36.81321	-76.125	9	Arrow	Left	[164, 103, 164, 151, 206, 151, 206, 103]	0.99		
541355	36.81321	-76.125	10	Arrow	Straight	[788, 1004, 788, 1050, 845, 1050, 845, 1004]	0.99		
541355	36.81321	-76.125	11	Arrow	Straight	[227, 40, 227, 84, 271, 84, 271, 40]	0.99		
541355	36.81321	-76.125	13	Arrow	Right	[744, 1056, 744, 1098, 797, 1098, 797, 1056]	0.99		
1269888	36.82015	-75.9998	14	Crosswalk	Transverse	[1127, 557, 1071, 560, 1022, 268, 1079, 264]	0.95	48.1	9.5
1269888	36.82015	-75.9998	15	Crosswalk	Transverse	[630, 612, 578, 624, 429, 263, 481, 251]	1	63.3	8.9
1269888	36.82015	-75.9998	16	Crosswalk	Transverse	[1003, 234, 993, 285, 476, 252, 487, 200]	0.97	84	8.8
1269888	36.82015	-75.9998	17	Crosswalk	Transverse	[1096, 612, 643, 676, 629, 630, 1082, 566]	0.83	76.6	7.8

Fig. 10 Sample of an exported.csv file

Table 2 Field description of the exported.csv data

Field	Description	Example
Intersection_ID	ID of intersections from the input file	541,465
Latitude	Coordinate that specifies the north–south position of an intersection	36.6830800700893
Longitude	Coordinate that specifies the east–west position of an intersection	76.0231042794389
Marking_ID	ID of markings started from 1	1
Type	Marking types such as crosswalks or lane-use arrows	Crosswalk
SubType	Crosswalks have three subtypes, i.e., transverse, zebra, and ladder; lane-ues arrows have five subtypes, i.e., left, right, left & straight, right & straight, and straight	Ladder
Location	Coordinates of four corner points of the detection bounding box. The Y-pixel coordinate increases from top to bottom and the X-pixel coordinate increases from left to right	[790, 282, 438, 313, 425, 276, 777, 245]
Quality	Measured as the probability of a marking belonging to the high-quality category. Describe marking degradation condition, ranging from 0 (lowest quality) to 1 (highest quality)	0.89
Length	The length of a crosswalk in feet	59.1
Width	The width of a crosswalk in feet	6.4

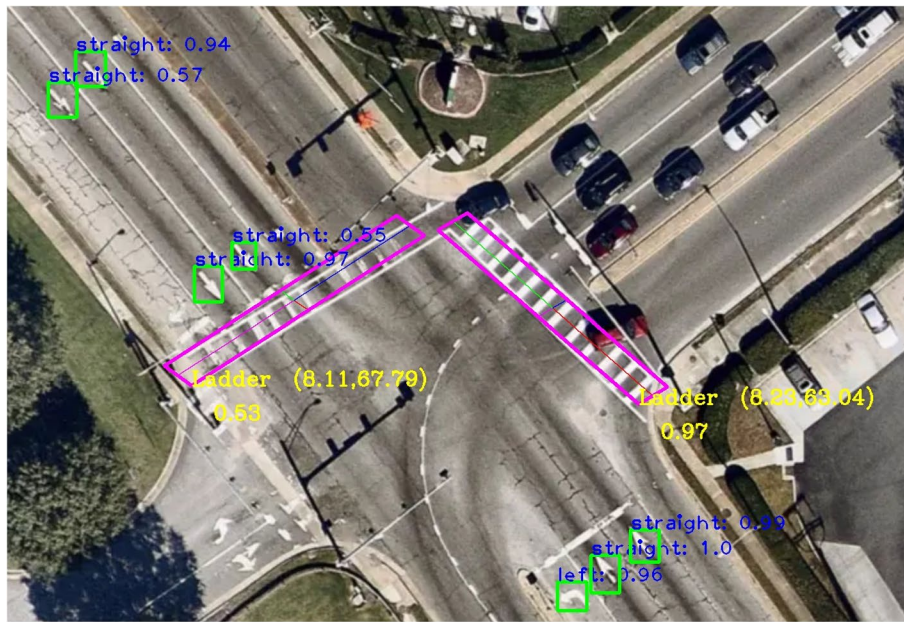


Fig. 11 Sample of an exported.jpg file (Numbers outside of parentheses are quality scores, numbers within parentheses are widths and lengths of crosswalks)

packages like NumPy, pandas, and JavaScript (JS) were employed. NumPy facilitated numerical computations, pandas enabled efficient data manipulation and analysis, while JS was used for creating dynamic elements in the frontend GUI.

4.6 Experiment setting

For the experiment settings, both the Faster RCNN (Ren et al., 2015) and BBAVectors (Yi et al., 2021) networks were trained for 100 epochs using a learning rate of 1e-4. A confidence threshold of 0.2 was set to determine the

detection. Additionally, the quality model was trained for 120 epochs for convergence. For the computational resources, the system is deployed on a 22.04 Ubuntu operating system with NVIDIA GeForce 3090 graphics card.

5 Results

5.1 Lane-use arrow detection

The downloaded aerial images from Mapbox were divided into a training set to train computer vision models and a testing set to test the trained model for

performance evaluation. Each aerial image is a 3-channel Red, green, and blue (RGB) color image with a rough resolution of 1354×967 pixels. The lane-use arrows of each image were also manually annotated.

Table 3 Lane-use arrow data distribution and detection performance

Traffic Markers	Total No	No. in Training Set	No. in Testing Set	Average Precision on Testing Set
Left	4,089	3,248	841	92%
Left & Straight	522	446	76	71%
Straight	1,136	860	276	86%
Right	2,399	1,920	479	89%
Right & Straight	676	550	126	86%
Overall	8,822	7,024	1,798	85%

Table 3 presents the distributions of the lane-use arrows in training and testing datasets.

After training the Faster RCNN model on the training set, the detection performance was evaluated on the testing set. Examples of correctly detected and incorrectly detected (e.g., misclassification, missing) lane-use arrows are presented in Fig. 12. Average precision (a.k.a., Area Under the Precision-Recall Curve) was used to evaluate the performance of each lane-use arrow class. Average precision can indicate whether the model can correctly identify all the positive examples without accidentally marking too many negative examples as positive. The mean average precision reaches 85% on the testing set as shown in Table 3.

5.2 Crosswalk detection

Over 3,000 aerial images of intersections with crosswalks were collected, which were subsequently divided into a training set to develop the deep learning model and a



(a) Correctly detected lane-use arrows



(b) Incorrectly detected lane-use arrows

Fig. 12 Example of correctly detected and incorrectly detected lane-use arrows (Numbers indicate confidence levels)

Table 4 Crosswalk data distribution and detection performance

Crosswalk Type	Total No	No. in Training Set	No. in Testing Set	Average Precision on Testing Set
T ₁ —Transverse Crosswalk	4,974	3,997	977	89%
T ₂ —Zebra Crosswalk	2,437	1,947	490	87%
T ₃ —Ladder Crosswalk	1,238	975	263	92%
Overall	8,649	6,919	1,730	89%

testing set to evaluate the model performance. All the crosswalks on these images were manually annotated. Table 4 Crosswalk data distribution and detection performance Table 4 presents the distributions of the crosswalks in training and testing datasets.

After developing the BBA Vectors model on the training set, the detection performance was evaluated on the

testing set. Examples of correctly detected and incorrectly detected (e.g., misclassification, missing) crosswalks are presented in Fig. 13. A mean average precision of 89% was achieved as shown in Table 4.

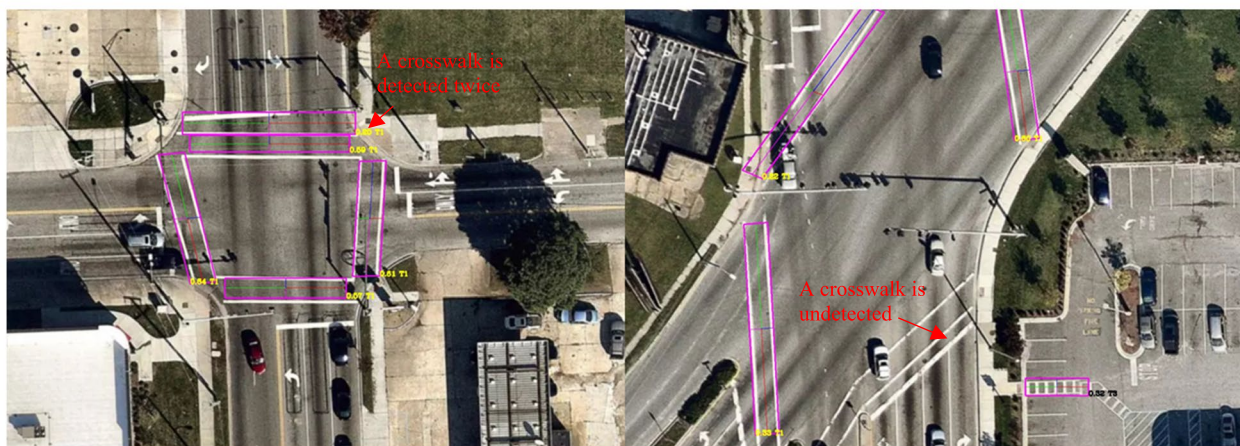
5.3 Assess the degradation conditions of markings

A total of 6,396 lane-use arrows and 5,031 crosswalks were annotated by trained reviewers. Tables 5 and 6 present the distributions of degradation conditions for lane-use arrows and crosswalks. The majority of markings (85.4% for lane-use arrows and 69.4% for crosswalks) are in the high-quality category.

After training the VGG16 model, the classification performance was evaluated on the testing sets of both lane-use arrows and crosswalks. Examples of correctly classified and incorrectly classified markings are presented in Fig. 14. Accuracy (No. of corrected classified instances/total No. of instances) was used to evaluate the performance of conditions assessment as reported in



(a) Correctly detected crosswalks



(b) Incorrectly detected crosswalks

Fig. 13 Examples of correctly detected and incorrectly detected crosswalks (Numbers indicate confidence levels)

Table 5 Degradation conditions of lane-use arrows and condition assessment performance

Quality Type	Total No	No. in Training Set	No. in Testing Set	Accuracy on Testing Set
Low-quality	1,154	923	231	80%
High-quality	6,765	5,412	1,353	94%
Overall	7,919	6,335	1,584	91%

Table 6 Degradation conditions of crosswalks

Quality Type	Total No	No. in Training Set	No. in Testing Set	Accuracy on Testing Set
Low-quality	1,327	1,061	266	81%
High-quality	3,009	2,407	602	85%
Overall	4,336	3,468	868	83%

Tables 5 and 6. The overall accuracies for lane-use arrows and crosswalks have achieved 91% and 83%, respectively.

6 Conclusions

This paper develops an automated system that utilizes advanced AI techniques to detect intersection markings and assess their condition. The system that has been developed holds immense potential for driving the progress of urban science by offering essential urban infrastructure data in a cost-effective manner, which serves as a foundation for analysis and decision-making processes. A summary of the investigation results is as follows:

1. A Faster RCNN model was developed to detect lane-use arrows. The mean average precision has achieved 85% on the testing set.
2. Developed a BBAAVectors model that can capture rotated objects to detect crosswalks and achieved a mean average precision of 89%.
3. A VGG16 model was developed to assess the degradation conditions of markings. The overall accuracies for lane-use arrows and crosswalks achieved 91% and 83%, respectively.

From the investigation, it is found that emerging AI techniques (e.g., deep learning) could deliver satisfactory data products in terms of detection, characterization, and condition assessment of intersection markings. The model performance could be further enhanced when

additional data are used for model development. The seamless integration of spatial analytics and advanced computer vision techniques makes the system truly cost-effective, scalable, and computationally efficient. The system harnesses emerging AI techniques such as multi-task deep learning to enhance its robustness, accuracy, and computational efficiency. The system is very accessible to users of different technical skills through its graphical user interface.

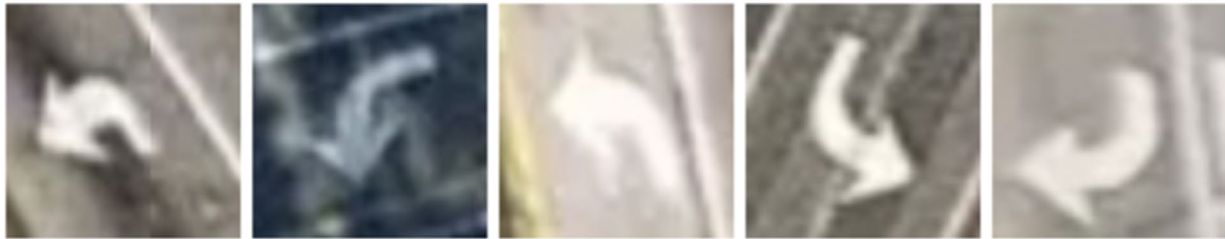
Existing intersection marking data are generally collected either by field investigation or computer-aided manual extraction from aerial images, street views, and/or video logs. These approaches cost prohibitive and only feasible for very limited data collection needs. In addition, their inherently subjective nature requires extensive training to reduce human errors. The system offers distinct advantages to innovate current practices: (a) extremely low cost, (b) extraordinary scalability, (c) timeliness and consistency, and (d) objective and high-degree reproducibility. The system can automate state-wide intersection marking data collection at almost zero cost and with machine-based objective measurements. It can enhance timeliness and consistency of roadway inventory data by rapidly processing latest aerial image data periodically. It eliminates the exposure of surveyors to hazards in field data collection. Unlike manual data collection, the system also provides objective measurements and a high-degree reproducibility of collected data.

The system can generate data elements highly expected by transportation agencies to support the Model Inventory of Roadway Elements (MIRE) program and to advance Highway Safety Improvement Programs (HSIP). Current data collection practices require transportation agencies to invest millions of dollars in contracting very time-consuming data collection services each year. By economically providing large-scale intersection marking data, this system will enable transportation agencies to empower analytic methods for data-driven safety management. The system can also assess the degradation condition of identified markings, and thus timely assist maintenance prioritization for reinforcing intersection safety.

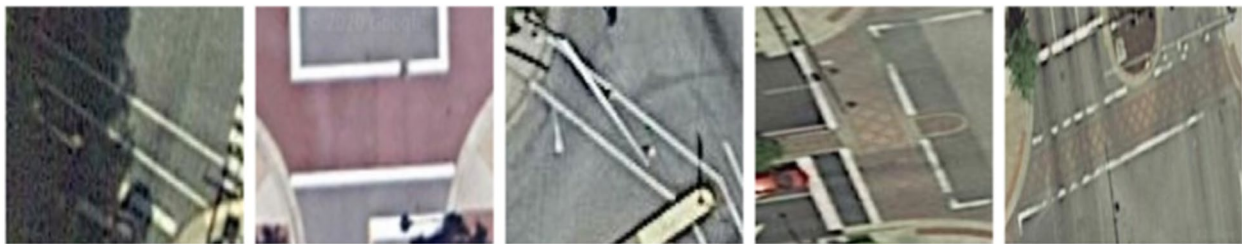
Although the system demonstrates promising performance, it is essential to acknowledge the potential limitations and challenges associated with utilizing aerial photo data in certain geographic contexts. In rural and mountainous regions, the resolution of aerial data might be insufficient, leading to potential impacts on the accuracy of detection and quality assessment outcomes. Furthermore, the less frequent updates of



(a) Correctly classified lane-use arrows (Ground truth: High, High, High, Low, Low; Predict: High, High, High, Low, Low)



(b) Incorrectly classified lane-use arrows (Ground truth: High, High, High, Low, Low; Predict: Low, Low, Low, High, High)



(c) Correctly classified crosswalks (Ground truth: High, High, High, Low, Low; Predict: High, High, High, Low, Low)



(d) Incorrectly classified crosswalks (Ground truth: High, High, High, Low, Low; Predict: Low, Low, Low, High, High)

Fig. 14 Examples of correctly classified and incorrectly classified markings based on degradation conditions

aerial data in these areas can result in outdated information, posing challenges in accurately capturing the current conditions of the markings. It is important to remain cognizant of these factors when implementing the proposed system for data collection in such areas. Additionally, the performance of computer vision models can be further improved by including more data for training.

Acknowledgements

The research team would like to express our sincere gratitude to the program manager, Inam Jawed, as well as the IDEA advisors, Paul J. Carlson and Wei Zhang, and the expert advisory panel members, In-Kyu Lim, Zhongren Wang, and Shan Di, for their invaluable guidance and assistance throughout the project.

Authors' contributions

K. Xie: Conceptualization, Methodology, Writing, Review & Editing; H. Sun: Methodology, Software, Writing; X. Dong: Data Curation, Writing, Validation; H.

Yang: Conceptualization, Methodology, Review & Editing; K. Yu: Conceptualization, Methodology, Review & Editing.

Funding

This research was supported by the NCHRP IDEA program (Project ID: NCHRP 225), with matching funds from the Virginia Department of Transportation.

Availability of data and materials

Data and material presented in this paper are not available for public access or distribution given intellectual property concerns.

Declarations

Competing interests

Authors have no competing interests to disclose.

Received: 1 May 2023 Revised: 19 June 2023 Accepted: 29 June 2023

Published online: 13 July 2023

References

- Ahmetovic, D., Manduchi, R., Coughlan, J. M., & Mascetti, S. (2015). Zebra crossing spotter: Automatic population of spatial databases for increased safety of blind travelers. *Proceedings of the 17th International ACM SIGAC-CESS Conference on Computers & Accessibility*, 251–258. <https://doi.org/10.1145/2700648.2809847>.
- Balali, V., Rad, A. A., & Golparvar-Fard, M. (2015). Detection, classification, and mapping of US traffic signs using google street view images for roadway inventory management. *Visualization in Engineering*, 3(1), 15.
- Burrow, M., Evdorides, H., & Snaith, M. (2000). Road marking assessment using digital image analysis. *Proceedings of the Institution of Civil Engineers-Transport*, 141(2), 107–112. <https://doi.org/10.1680/tran.2000.141.2.107>.
- Chen, T., Chen, Z., Shi, Q., & Huang, X. (2015). Road marking detection and classification using machine learning algorithms. *2015 IEEE Intelligent Vehicles Symposium (IV)*, 617–621. <https://doi.org/10.1109/IVS.2015.7225753>.
- Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q., & Tian, Q. (2019). Centernet: Keypoint triplets for object detection. *Proceedings of the IEEE/CVF international conference on computer vision*, 6569–6578. <https://doi.org/10.48550/arXiv.1904.08189>.
- FHWA. (2009). *Manual on Uniform Traffic Control Devices*. US Department of Transportation Federal Highway Administration. Retrieved December 16 from https://mutcd.fhwa.dot.gov/htm/2009/part3/fig3b_19_longdesc.htm
- Fiedler, R., Lefler, N., Mallela, J., Abbott, D., Smelser, D., & Becker, R. (2013). *MIRE MIS Lead Agency Data Collection*.
- Foucher, P., Sebsadjji, Y., Tarel, J., Charbonnier, P., & Nicolle, P. (2011). Detection and recognition of urban road markings using images. *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 1747–1752. <https://doi.org/10.1109/ITSC.2011.6082840>.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 580–587. <https://doi.org/10.48550/arXiv.1311.2524>.
- Girshick, R. (2015). Fast r-cnn. *Proceedings of the IEEE international conference on computer vision*, 1440–1448. <https://doi.org/10.48550/arXiv.1504.08083>.
- Greenhalgh, J., & Mirmehdi, M. (2015). Detection and Recognition of Painted Road Surface Markings. *Proceedings of the International Conference on Pattern Recognition Applications and Methods - Volume 1: ICPRAM*, 130–138. <https://doi.org/10.5220/0005273501300138>.
- Kang, L., Ye, P., Li, Y., & Doermann, D. (2014). Convolutional neural networks for no-reference image quality assessment. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1733–1740. <https://doi.org/10.1109/CVPR.2014.224>.
- Lathkar, M. (2023). Introduction to FastAPI. In *High-Performance Web Apps with FastAPI: The Asynchronous Web Framework Based on Modern Python*, 1–28. <https://doi.org/10.1007/978-1-4842-9178-8>.
- Lefler, N., Zhou, Y., Carter, D., McGee, H., Harkey, D., & Council, F. (2017). *Model Inventory of Roadway Elements MIRE 2.0*.
- Lin, K.-L., Wu, T.-C., & Wang, Y.-R. (2016). An innovative road marking quality assessment mechanism using computer vision. *Advances in Mechanical Engineering*, 8(6), 1687814016654043.
- Liu, Z., Wang, S., & Ding, X. (2012). ROI perspective transform based road marking detection and recognition. *2012 International Conference on Audio, Language and Image Processing*, 841–846. <https://doi.org/10.1109/ICALIP.2012.6376731>.
- Mallela, J., Sadasivam, S., & Lefler, N. (2012). *MIRE Element Collection Mechanisms and Gap Analysis*.
- Máttyus, G., Wang, S., Fidler, S., & Urtasun, R. (2016). Hd maps: Fine-grained road segmentation by parsing ground and aerial images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3611–3619. <https://doi.org/10.1109/CVPR.2016.393>.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., & Antiga, L. (2019). Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32.
- Proulx, F. R., Zhang, Y., & Grembek, O. (2015). Database for active transportation infrastructure and volume. *Transportation Research Record*, 2527(1), 99–106.
- Qian, R., Liu, Q., Yue, Y., Coenen, F., & Zhang, B. (2016). Road surface traffic sign detection with hybrid region proposal and fast R-CNN. *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, 555–559. <https://doi.org/10.1109/FSKD.2016.7603233>.
- Rebut, J., Bensrhair, A., & Toulminet, G. (2004). Image segmentation and pattern recognition for road marking analysis. *2004 IEEE International Symposium on Industrial Electronics*, 1, 727–732. <https://doi.org/10.1109/ISIE.2004.1571896>.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28, 91–99.
- Senlet, T., & Elgammal, A. (2012). Segmentation of occluded sidewalks in satellite images. *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, 805–808.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Smith, V., Malik, J., & Culler, D. (2013). Classification of sidewalks in street view images. *2013 International Green Computing Conference Proceedings*, 1–6. <https://doi.org/10.1109/IGCC.2013.6604476>.
- Sukhwani, M., Singh, S., Goyal, A., Behl, A., Mohapatra, P., Bharti, B. K., & Jawahar, C. (2014). Monocular vision based road marking recognition for driver assistance and safety. *2014 IEEE International Conference on Vehicular Electronics and Safety*, 11–16. <https://doi.org/10.1109/ICVES.2014.7063716>.
- Tian, J., Yuan, J., & Liu, H. (2020). Road marking detection based on mask R-CNN instance segmentation model. *2020 international conference on computer vision, image and deep learning (CVIDL)*, 246–249. <https://doi.org/10.1109/CVIDL51233.2020.00-92>.
- VDOT. (2017). *LRS Road Intersections*. Virginia Department of Transportation. Retrieved August 2 from <https://www.virginiaroads.org/datasets/VDOT:Lrs-road-intersections/explore?location=37.912663%2C-79.494919%2C8.61>
- Vokhidov, H., Hong, H. G., Kang, J. K., Hoang, T. M., & Park, K. R. (2016). Recognition of damaged arrow-road markings by visible light camera sensor based on convolutional neural network. *Sensors*, 16(12), 2160.
- Wen, C., Sun, X., Li, J., Wang, C., Guo, Y., & Habib, A. (2019). A deep learning framework for road marking extraction, classification and completion from mobile laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, 178–192.
- Wu, T., & Ranganathan, A. (2012). A practical system for road marking detection and recognition. *2012 IEEE Intelligent Vehicles Symposium*, 25–30. <https://doi.org/10.1109/IVS.2012.6232144>.
- Xia, G.-S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J., Dattu, M., Pelillo, M., & Zhang, L. (2018). DOTA: A large-scale dataset for object detection in aerial images. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3974–3983. <https://doi.org/10.48550/arXiv.1711.10398>.
- Yamamoto, J., Karungaru, S., & Terada, K. (2014). Road surface marking recognition using neural network. *2014 IEEE/SICE International Symposium on System Integration*, 484–489. <https://doi.org/10.1109/SII.2014.7028087>.
- Yi, J., Wu, P., Liu, B., Huang, Q., Qu, H., & Metaxas, D. (2021). Oriented object detection in aerial images with box boundary-aware vectors. *Proceedings*

of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2150–2159. <https://doi.org/10.48550/arXiv.2008.07043>.

Zhang, Y., & Ge, H. (2012). Assessment of presence conditions of pavement markings with image processing. *Transportation Research Record*, 2272(1), 94–102.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.