ASCE

Development of a Car-Free Street Mapping Model Using an Integrated System with Unmanned Aerial Vehicles, Aerial Mapping Cameras, and a Deep Learning Algorithm

Seungho Lee[1]; Seokhyeon Kim[2]; and Sungkon Moon[3]

Abstract: Road condition and quality are critical road maintenance and risk reduction factors. Most existing road monitoring systems include regular on-site surveys and maintenance. However, major roads in urban areas are generally complicated and have heavy traffic during the daytime, so such field investigation can be significantly limited. Moreover, any road work at nighttime can be risky and dangerous and incur excessive expenses. Based on a review of existing systems for monitoring road conditions, this study focuses on overcoming two unsolved challenges: the capacity of the monitoring range and the avoidance techniques to ensure traffic is not hindered. To solve these challenges, these paper proposes an integrated road monitoring system called Car-free Street Mapping (CfSM) using unmanned aerial vehicles (UAV), aerial mapping cameras, and deep learning (DL) algorithms. The use of the aerial mapping camera mounted on the UAV is to widen the monitoring viewing range, and general-purpose drones are used in this study rather than expensive special equipment. Since the drone-taken images include many passing vehicles that conceal the road surface from the camera vision, the DL model was applied to detect the vehicles and their shadows and then remove them from the images. To train the DL model, two image datasets were used: publicly available cars overhead with context (COWC) images and orthoimages additionally taken for the project to further improve the accuracy. The two datasets consist of 298,623 labeled objects on 9,331 images in total. The tests resulted in a mean average precision (mAP) of 89.57% for trucks, 95.77% for passenger vehicles, and 76.51% for buses. Finally, the object-removed images were composited into one whole car-free image. The CfSM was applied to two areas in Yeouido and Sangam-dong, Seoul, Korea. The car-free images in both regions show a spatial resolution of 10 mm and can be used for various purposes such as road maintenance and management and autonomous vehicle roadmaps.
DOI: 10.1061/(ASCE)CP.1943-5487.0001013.

Author keywords: Road surface model; Deep learning (DL); Image recognition; Mean average precision (mAP); Unmanned aerial vehicle (UAV); Drone; Infrastructure; Car-free street mapping (CfSM).

Background and Introduction

Real-time road conditions play a significant role in most modern transportation management systems (Gong et al. 2012). One of the major issues is the safety measures involved when driving vehicles (Malin et al. 2019). More recently, cutting-edge technologies, such as self-driving cars, rely on the corresponding road conditions in real-time (Chowdhury et al. 2020). The safety considerations of both human-driving and self-driving vehicles considerably depend on road condition and quality (Singh and Suman 2012). It has been reported that road conditions and quality are major factors that can decrease vehicle-related accidents (Bhatt et al. 2017).

Systems for vehicle driving control and dynamics require a wider scale of mapping and identification (Tang et al. 2014). At the same time, the systems must be able to operate in real-time and with a fast-tracking mode since road and related infrastructure can change from moment to moment because of emergency and situational conditions (Katsuma and Yoshida 2018). In addition, traffic interference has been one of the major issues when road mapping and maintenance have been conducted in urban areas (Zekavat et al. 2014; Zhang et al. 2019). Roads in urban areas are crowded with heavy traffic and frequent construction, which hinders road inspection and maintenance (Saleh et al. 2017; Zekavat et al. 2015). In short, a system must be developed to enable coverage of a wider range of areas with shorter processing times. It needs to meet the complex requirements of cities to minimize the frequency of traffic restrictions.

The existing systems for road inspection and maintenance have limitations when applied to this sort of heavy traffic situation. For example, Park et al. (2019) developed a system named PotholeEye, which uses vehicle-mounted cameras with GPS receivers. The critical limitation of this system is its viewing range capabilities as it depends on a ground-moving vehicle. Yulianandha Mabrur (2019) developed another system using drones to assess road surface conditions. The system, called DroneDeploy, is for road pavement inspection and can cover a wider range of roadways for data collection and model development. Zhang and Elaksher (2012) also presented an unmanned aerial vehicle (UAV)-based system for generating road images by which a resolution of less than 5 mm was achieved. However, the system was applied to unpaved road conditions rather than to an urban area. Most of existing developments did not consider the economical aspect of the implementation,

[1]Chief Executive Officer, 4S Mapper, 815 Daewangpangyo, Seongnam, 13449, Republic of Korea; DroMii Co., Ltd., 78 Mapo-daero, Mapo-gu, Seoul, 04168, Republic of Korea. Email: rsgis@paran.com

[2]Assistant Professor, Dept. of Civil Engineering, Kyung Hee Univ., Deogyeong-daero, Yongin, 17104, Republic of Korea. ORCID: https://orcid.org/0000-0002-1819-8022. Email: seokhyn.kim@gmail.com

[3]Associate Professor, Dept. of Civil Systems Engineering, Ajou Univ., Suwon 16499, Republic of Korea (corresponding author). Email: sungkon.moon@gmail.com

Note. This manuscript was submitted on August 3, 2021; approved on November 19, 2021; published online on February 7, 2022. Discussion period open until July 7, 2022; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Computing in Civil Engineering***, © ASCE, ISSN 0887-3801.

although this is one of the critical components of new system development (Moon et al. 2020).

Accepting the usable aspects of the existing road surface monitoring systems and overcoming their limitations, this research project presents a novel system, named Car-free Street Mapping (CfSM). CfSM aims to comprehensively monitor road surface conditions in urban areas without road traffic interference and is composed of three key technologies: unmanned aerial vehicles (UAV), aerial mapping cameras, and a deep learning (DL) algorithm. For CfSM, general-purpose drones are used to secure a wide range of vision from a distance. The aerial mapping cameras mounted on the drones capture road surface images at less than 10-mm resolution accuracy. The images captured by the aerial mapping cameras are processed to identify and remove objects (cars and their shadows) and then composited as a car-free image. The car-free images can be widely used for road maintenance and management, as well as basic roadmaps for autonomous vehicles or actual roadmaps for existing mapping systems. In addition, a comparative analysis with the existing systems will present the advantages and cost effectiveness of CfSM.

The remaining parts of this paper are organized as follows. In the "Related Works" section, a review of related works is presented, including road monitoring systems and technologies. Sections on "System Development and Research Design" and "System Implementation and Field Tests" provide details for system development and implementation, respectively. In the "CfSM System Verification" section, verification of CfSM is presented and includes the DL model accuracy and resolution of the car-free image. Lastly, in the "Implication, Discussion and Future Works" and "Conclusion and Research Limitation" sections, we discuss the results and future works with concluding remarks.

Related Works

Road surface monitoring techniques can be categorized by data and sensor type. The former refers to data types (e.g., images and signals) to assess road surface conditions, and the latter relates to types of sensors (e.g., optical image sensor, GPS receiver and accelerometer) and special or general-purpose equipment for mounting (e.g., satellite, vehicle, drone and smartphone). Such techniques have been developed together with advances in computational capability (speed and algorithm), data network (speed, station density/coverage and crowdsourcing), and equipment (accuracy/precision and generality). In this study, we present the development of road surface monitoring technology in chronological order according to the type of data: image and signal. The following shows the evolution of various sensors, mounted equipment, and data processing algorithms by category of image and signal.

Image-Based Methods

There have been many studies on road surface monitoring that are limited to specific events (e.g., earthquake) or use specific equipment designed for a specific purpose. Li et al. (2010) presented a framework to assess road damage based on the damage dimensions (i.e., length, area, and width) detected on a remotely-sensed image taken after an earthquake. Zhang (2008) presented a UAV-based remote sensing system to obtain road images that provide road conditions to identify issues, such as corrugation, potholes, and gravel loss. Li et al. (2011) proposed a method to detect earthquake damage to urban road networks based on both spectral and texture information from very high resolution (VHR) satellite images. In this method, once road regions are extracted from a pre-event road map, the damaged regions are identified from post-event VHR images by a one-class classifier, called a one-class support vector machine (OCSVM).

Mertz (2011) introduced a road monitoring system to detect damage, such as potholes and cracks, consisting of a structured light sensor and a camera mounted on a vehicle that regularly traveled on the road (e.g., a bus). The system continuously collects and analyzes data from numerous vehicles for road maintenance operations. Gong et al. (2012) presented a simple method to detect road damage extraction based on an object-oriented change detection algorithm using pre-established vector data of roads (i.e., lines and polygons) superimposed on post-earthquake remote sensing images. Wang et al. (2015) proposed a knowledge-based method for detecting road damage solely using post-disaster high-resolution remote sensing images. In this framework, road centerlines are extracted from pre-surveyed seed points, and road features are selected from remote sensing image-based knowledge of the roads.

More recently, increased smartphone use and the development of machine learning techniques have led to advances in various technologies for road surface monitoring. Strazdins et al. (2011) presented a framework to detect potholes using information, such as location and acceleration, received from multiple smartphones. Chang et al. (2012) presented a smartphone-based system for monitoring and reporting road conditions, called Road Monitoring and Reporting System (RMRS). Road users can easily report images of road surface conditions via RMRS on their mobile phones, and the data are transferred to on-site road engineers for maintenance and rehabilitation works. Alfarrarjeh et al. (2018) and Maeda et al. (2018) proposed DL-based models to monitor road damage. The DL model is first trained by various image examples showing damage types and then used to detect and classify various types of road damage in images taken on smartphones.

Signal-Based Methods

In general, the signals for detecting abnormal road surface conditions are obtained from Global Positioning System (GPS) receivers and accelerometers. Eriksson et al. (2008) proposed a mobile-sensing system to identify and inform the surface conditions of roads. This system collects vibration data through GPS from several sensor-equipped participating vehicles in a region, from which the locations of potholes can be detected. Suda et al. (2010) proposed an algorithm to detect road conditions, such as irregularity and friction from tire vibration. El-Wakeel et al. (2017) demonstrated the results of monitoring road surface types and anomalies. The proposed model is based on wavelet analysis with a statistical approach using measurements of land vehicle-mounted inertial sensors.

With advances in smartphones, which generally have inbuilt GPS and an accelerometer, many studies have used smartphone-collected data to detect road surface anomalies. Mohan et al. (2008) presented a system to monitor road conditions using information collected from mobile phone-mounted sensors. Chen et al. (2013) proposed a crowdsourcing-based road surface monitoring system (CRSM) that can effectively detect road potholes and evaluate road roughness levels using vibration pattern, location, and vehicle velocity received from vehicles operating over a wide area (e.g., taxis).

Alessandroni et al. (2014) proposed a crowdsourcing collaborative system to monitor road surface quality. In this system, surface roughness is estimated by processing signals received from smartphones. Vittorio et al. (2014) presented a system to monitor road surface anomalies, such as potholes and speed bumps, by using GPS-based vehicle location and three-axis accelerometer-based acceleration records obtained from the passenger's mobile phone.

Alqudah and Sababha (2017) proposed a method to detect abnormal road conditions, such as potholes and cracks, using gyro rotation data obtained from a smartphone-equipped sensor. The gyro rotation data are collected from multiple round trips over a path a few km in length, and the abnormality is represented as the moving variance of the data.

Harikrishnan and Gopi (2017) proposed a method to monitor abnormal road surfaces, such as potholes or humps, and predicted the severity based on vertical vibration signals of a moving vehicle measured by a smartphone accelerometer. Kumar et al. (2017) presented a method for sensing road surface conditions using acceleration data received from smartphones of users traveling on local roads. Once the data are collected at a central server, the road roughness is classified by a fuzzy-logic algorithm and visualized on Google Maps.

Chuang et al. (2019) proposed a road surface monitoring system using vertical and lateral acceleration data collected from smartphones. The data are acquired through a smartphone web application and converted to the perception of riding comfort highly correlated with road surface quality. Varona et al. (2019) proposed a deep learning approach for the identification of types of road surface and potholes. Here, the potholes are distinguished from destabilizations produced by speed bumps or driver actions in the crowdsensing-based application context. Abbondati et al. (2021) proposed a system called SmartRoadSense to automatically detect road surface conditions using data from GPS receivers and triaxial accelerometers equipped in mobile devices.

Overview of the Proposed Car-free Street Mapping

Despite these advances in road monitoring and relevant technology, providing a comprehensive survey of major roads in urban areas is still challenging. The challenge mainly results from high traffic volumes or shortcomings of existing approaches such as limited cost and time. To overcome such limitations, this study proposes an integrated road monitoring system called Car-free Street Mapping (CfSM), which uses an aerial mapping camera mounted on general-purpose drones and deep learning techniques.

The objectives of CfSM are to: (1) provide an alternative to on-site road surveys, (2) provide comprehensive records of road conditions for maintenance and management, (3) generate materials for building basic roadmaps for autonomous vehicles, and (4) function as real road maps with mobile mapping system (MMS) data.

As shown in Fig. 1, the proposed CfSM consists of four stages: (1) aerial photography by drones, (2) object detection by deep learning algorithms, (3) the masking of detected objects, and (4) image production by compositing biaxially overlapped multiple images. This paper includes two field works in Seoul, Korea, in which the system was implemented to generate a model of over 40 km of road surface. The field works aim to provide a practical method to address existing challenges by achieving a superior performance regarding the system's efficacy and cost effectiveness. Further details for each stage are presented in the following sections.

System Development and Research Design

For the first stage, aerial photography by drones has been used, which are two general-purpose drones, DJI Phantom 4 Pro and DJI Mavic 2 Pro (SZ DJI Technology, Guangdong, China). The former drone was used for road images, and the latter was for panorama images. Detailed specifications of the two drones are available at SZ DJI Technology Co., Ltd (2022a, b), respectively. Then, image acquisition conditions to meet the spatial resolution requirement, i.e., 10 mm, were determined by literature review and field testing over two project sites: Yeouido and Sangam in Seoul, Korea (see section "System Implementation and Field Tests" for details). It was found that the flight altitude of the drone should remain stable at 30 m to meet the resolution requirements. Images were taken with an overlap of 80% or more in both longitudinal and transverse directions of the roads to prevent voids in the final product.

The next stage is the detection of objects (i.e., cars and their shadows) in the drone-taken images (see section "Accuracy of Object Detection" for details). For this, two widely-used deep learning (DL) algorithms, RetinaNet (Lin et al. 2017) and ResNet (He et al. 2016), have been tested. We chose RetinaNet, which provided better model performance for object identification on road surfaces, such as vehicles and their shadows. The DL model was trained by object samples labeled on oblique images from the Cars Overhead With Context (COWC) datasets (Mundhenk et al. 2016), which is publicly available at https://gdo152.llnl.gov/cowc/. For model training, 80% of the dataset was used and the remaining 20% was used to assess the model accuracy, which was limited to 20% to 40%. This is because, while the car-free image we aim to generate is based on drone-taken orthoimages, the provided samples used for initially training the DL model are based on non-orthoimages. To overcome this limitation, we additionally trained the model by using object samples labeled on drone-taken images and improved the accuracy to 87%.

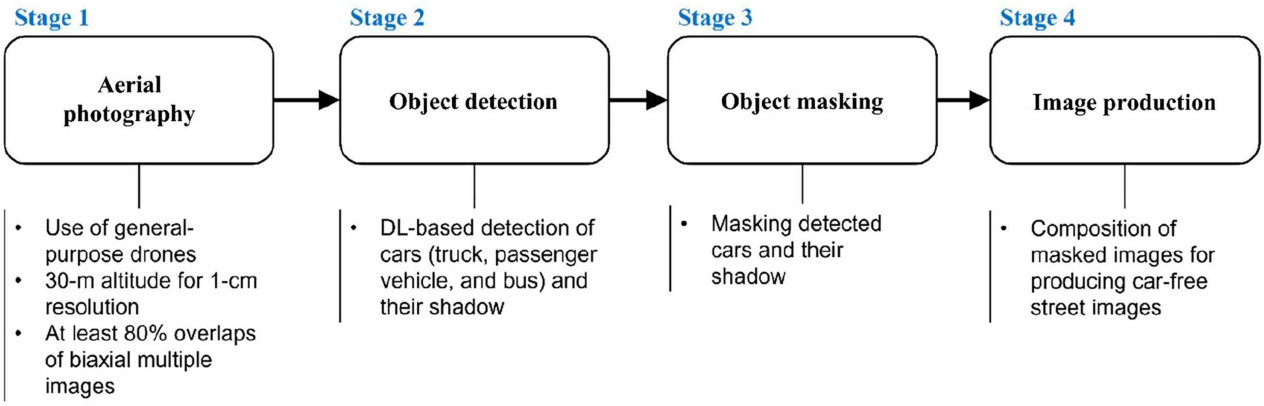


Fig. 1. Four stages of Car-free Street Mapping (CfSM).

J. Comput. Civ. Eng., 2022, 36(3): 04022003

The following stages are for masking the detected objects from the images and compositing the processed images into one whole-road surface image (see section "CfSM System Resolution and Readability" for details). Once the objects have been identified through the trained DL algorithm, they are masked from the original images. Then, the images with voids are composited with one where cars are now shown, and the composite image of the road surfaces is inspected with the original image of the road to verify visual quality. Also, it is required to check unremoved vehicles by which another flight may be needed to re-take images of the parts.

System Implementation and Field Tests

System Implementation to Two Regional Areas in Seoul

The CfSM system was applied to two regional areas in Seoul city: Yeouido and Sangam (Fig. 2). The test for Yeouido was designed to cover a road length of 23.3 km, and the test for Sangam was designed for a length of 20.7 km. Since the drones generally cover 0.1 km to 0.3 km in one flight, 50 and 48 flight/panorama points were planned for Yeouido and Sangam, respectively, as shown in Figs. 2(c and d).

The required number of flights at each point was determined by the number of lanes of roads, considering each flight can cover two lanes to take images. For example, a four-lane road required two times of flights, by generating two lanes of the image by each flight. When the speed limit in the areas is under 60 km/h, the difference between the drone's flying speed and that of passing vehicle(s) is insufficient to generate a gap between images. Namely, the images of vehicles and shadows from the picture of different angles are largely overlapped and result in a failure of object detection. Thus, the area(s) where the speed limit is set below 60 km/h required more than two flights. Additional flights were required in certain areas due to the following circumstances: (1) vehicles stopped at an intersection by traffic signals; (2) vehicles traveling slower than 10 km/h; (3) illegally stopped or parked vehicles; (4) bad weather conditions, such as strong winds, rain, or fog.

Multi-Angle Orthoimage for Photogrammetry

The flight plan for the Yeouido area included 128 flights to cover the sections making up 23.3 km in total. Each flight covered the typical distance between two intersections, 0.1–0.3 km, by considering the regulation and communication range. In the inner city of Seoul, drone operations are only allowed within the visibility range, and the pre-test struggled with frequent communication disruptions because of high-rise buildings. On the other hand, this short division helped to manage the data designed for each section of the target range.

The plan for Sangam similarly required 139 flights to cover 20.7 km. Figs. 3(a and b) shows a sample of the images taken by the drone and the pre-processing of the images. Fig. 3(a) shows overlaid multi-angle orthoimages in two-dimensional photogrammetry, and Fig. 3(b) presents the overlay with transparent detection, which is for the passing vehicles and their shadows, to attain the image to reveal the ground surface state. The image data shows a maximum of four overlaid object(s), which required at least five different angles of orthoimages.

CfSM System Verification

Accuracy of Object Detection

The mean average precision (mAP) was adopted to assess the accuracy of the object detection based on testing two DL algorithms, RetinaNet and ResNet, and RetinaNet, which provided better model performance, was finally selected for the model development. The times of two of the model training and assessments were

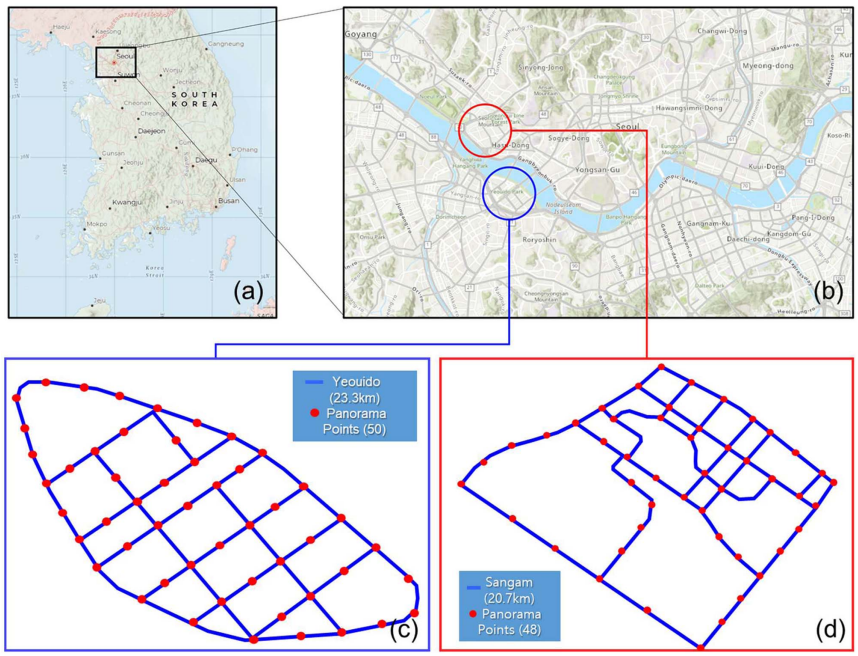
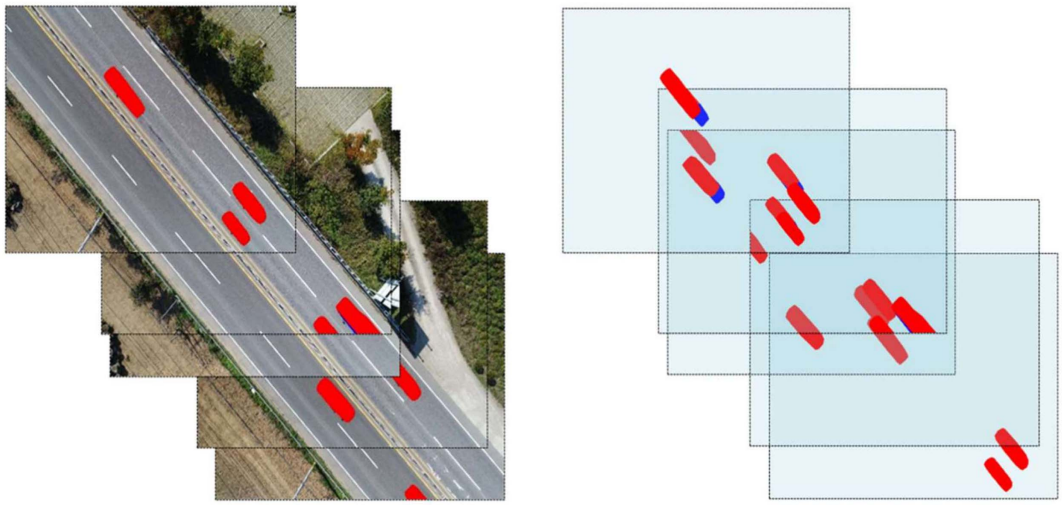


Fig. 2. Two CfSM system implementation areas, Seoul, Korea and Panorama Points: (a) map of South Korea; (b) two locations of field implementation; (c) section in Yeouido; and (d) section in Sangam. [Sources (a and b): Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community.]



(a) (b)

Fig. 3. Image processing sample: (a) overlay multi-angle orthoimage for photogrammetry (images by authors); and (b) orthoimage processing.

conducted by using 298,623 labeled objects (280,716 + 17,907) in 9,331 images (7,443 + 1,888), and 80% and 20% of the datasets have been used for model training and validation, respectively. The first dataset (i.e., COWC) has a large number of objects but consists of oblique images while the second one has a small amount but is composed of the orthogonal images that were recently taken by the drones to improve the accuracy. Note that the drone-taken orthoimages are more suitable for training the DL model because the CfSM system is based on the orthoimages. The accuracy test is categorized into three different vehicles: car (passenger vehicle), bus, and truck. The absolute numbers of each vehicle are 3,972, 74, and 784, respectively. A lower accuracy performance was initially expected by the limited number of buses.

Fig. 4 shows assessments for the three types of vehicles (truck, passenger vehicle, and bus) by plotting the precision against the recall value. The x-axis represents the index of how accurately the positives are founded by the recall, which is calculated by the ratio between the founded positives and all positives:

$$Recall = True\ positive/(True\ positive + False\ negative) \tag{1}$$

On the other hand, the y-axis represents how accurately the model finds the positives, which shows the ratio between the accurate positives by all founded positives:

$$Recall = True\ positive/(True\ positive + False\ positive) \tag{2}$$

The area under the precision-recall curve is generally defined as the value of the average precision (AP):

$$AP = \int_0^1 f(x)dx \tag{3}$$

Where x is the recall value, f(x) is a function for determining the precision-recall curve, and AP is the average precision value. The recall is determined by the proportional value of true positives out of total cases; the precision is obtained by the proportional value of true positives out of total positive cases, which is a positive predictive value. (Sammut and Webb 2011; Aggarwal 2018; Etten 2019).

The mean average precision, *mAP*, is obtained by averaging the values for all classes, where N is the number of classes, and n is the number of each class:

$$mAP = \frac{1}{N}\sum_{n=1}^{N} AP_n \tag{4}$$

Figs. 4(a and b) depict the accuracy test of the truck detection. The result shows an improved accuracy from 20.97% to 89.57%, after further training. The COWC data set and further training resulted in increased detection accuracy by 68.6%. The precision value has a zigzag pattern according to the true positive value, and the recall increases with a decrease in the precision value. During the first attempt, when the recall value exceeded 0.2, the precision value decreased to 0.86. This pattern shows that there were sharp drops in the number of true positives in the object detection during the first attempt. On the other hand, during the second attempt, the gap of this drop was much less than the first. When the recall value exceeded 0.8, the precision value still sustained 0.97, which then resulted in a mAP of 89.57%.

Figs. 4(c and d) shows the accuracy test of car (passenger vehicle) detection. The figure shows high accuracy on the first attempt of 90.88%. It is assumed that a greater number of samples (3,972) resulted in a comparably high mAP. The training and further labeling still showed an increase from 90.88% to 95.77%. When the recall value reached 0.8, the precision value did not fall from 0.995. This pattern eventually results in a mAP of 95.77%.

Figs. 4(e and f) presents the accuracy test of the bus detection. The smaller number of samples accounts for the lower mAP. Only 74 buses were detected during the entire field test. However, there was still a significant improvement after the training with the COWC data set and further labeling, which obtained a mAP of 76.51%, a marked improvement from 3.16% on the first attempt.

CfSM System Resolution and Readability

Fig. 5 presents four steps of the CfSM system. First, the original orthoimages were collected and prepared for pre-processing. The group of images was divided into different categories in accordance with each section of the drone-flight stage. Different angle orthoimages were then overlapped to create a 2D-photogrammetry, as

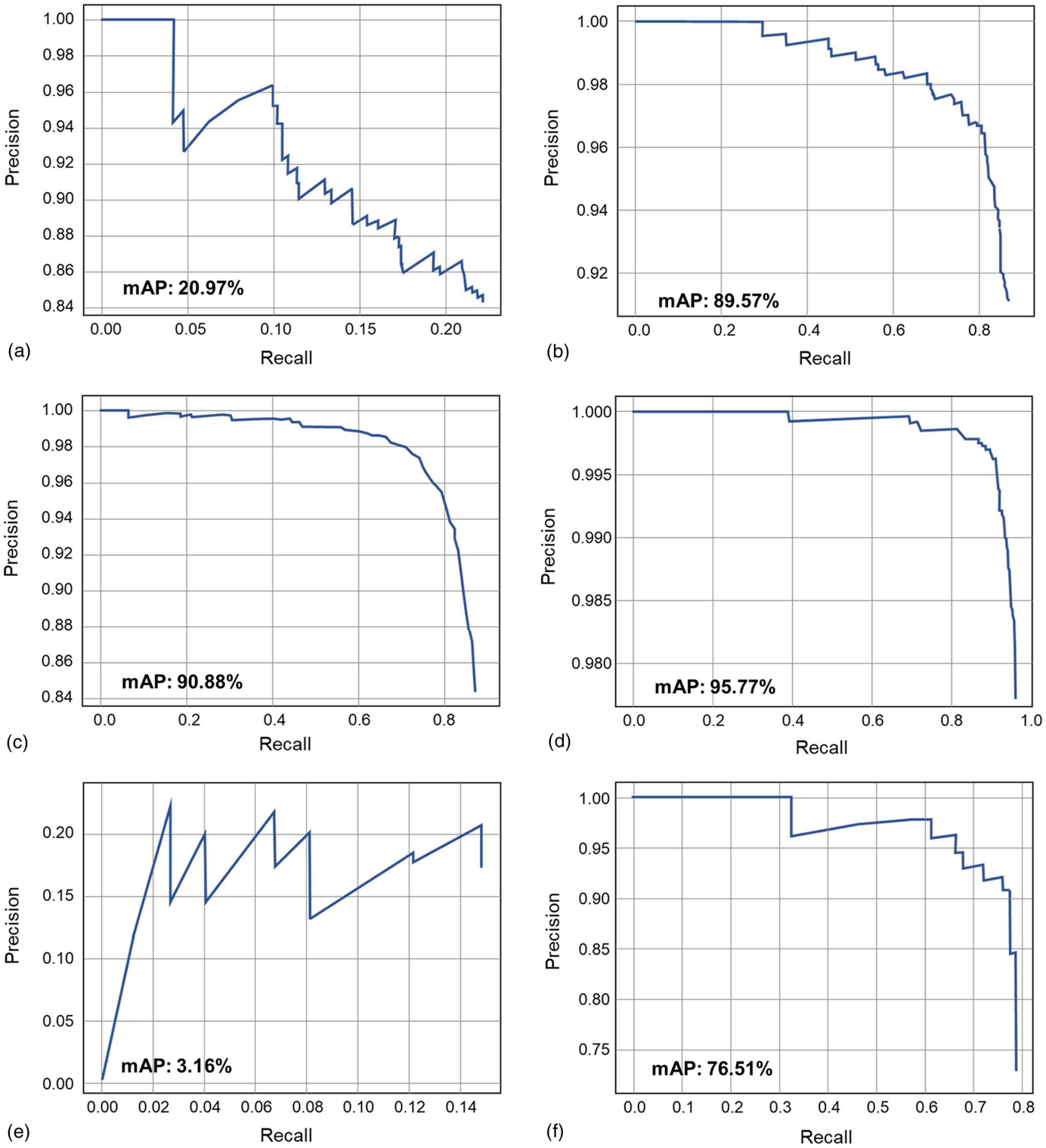


Fig. 4. mAP-based accuracy test results of object detection by deep learning: (a) first attempt for trucks; (b) second attempt for trucks; (c) first attempt for passenger vehicles; (d) second attempt for passenger vehicles; (e) first attempt for buses; and (f) second attempt for buses.

shown in the second step of the figure. A number of smudges came up during the overlapping, since the multiple angles of different orthoimages created some superpositions. These superpositions include the vehicles and their shadow images. From the second to third steps of the CfSM system, a deep learning application detected these passing vehicles and shadows to remove them from the original 2D-photogrammetry. The passing vehicles, including cars, buses, and trucks, were eliminated during the third step, and the fourth step finally identified the shadows to complete the transparent surface model of the road.

As discussed in the previous section, the detection accuracies were 89.57% for trucks, 95.77% for passenger vehicles, and 76.51% for buses, so another post-process was required to pinpoint the undetected parts from the original model. Fig. 6 presents a

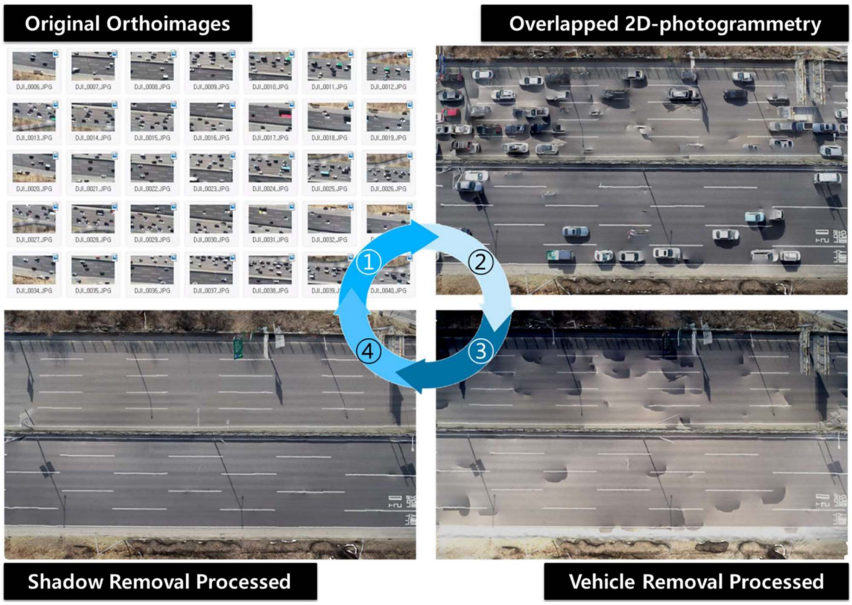


Fig. 5. Execution of CfSM system. (Images by authors.)

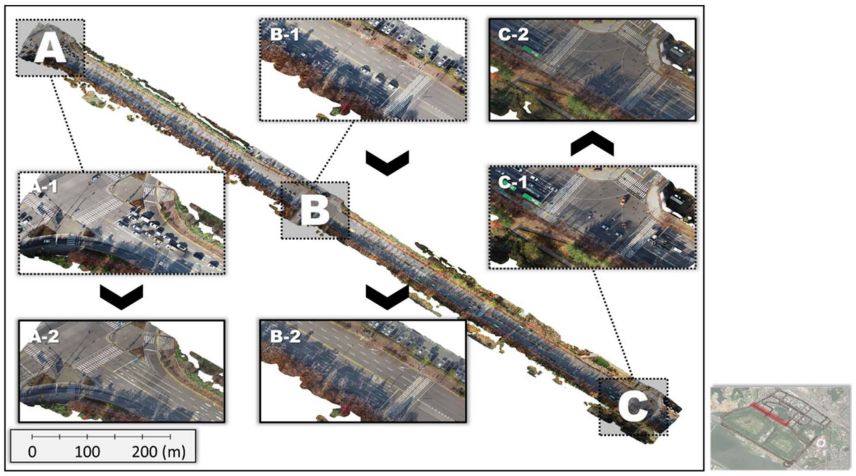


Fig. 6. Three sections of a part of the Sangam's CfSM model. (Images by authors.)

segment of Sangam's model, as seen in the mini-map on the bottom right-hand side. Sections A-C are three intersections of the selected segment. Section A first presents the intersection of the upper side of the segment. As seen as A-1 in the figure, approximately 20 vehicles were waiting for a signal and several more were passing by on the center of the intersection. A-2 represents 100% of detection and removal and a transparent surface model of the A section.

Section B, on the other hand, had a dozen unremoved vehicles at the upper side of the car park. Three cars waiting for a signal were successfully identified when the drone came back to the starting point. However, the cars at the car park remained stationary during the return-way of the flying drone, so it was not possible to remove them. Section C resulted in another quality model at the final stage, but one bus was not eliminated. This is an example that required a post-process to correct the error of the system. The center of the intersection, and even the cars stopped at the signal appeared clearly and were identified accurately, which in turn presented a transparent surface of the section. The following inspection determined that the object features were also identified, including the dimension and number of manhole lids.

Implication, Discussion, and Future Works

A core relevance of the CfSM system lies in its feasibility in creating a high-definition map (HD map). The HD map is a highly accurate road map, one of the essential components of many innovative technologies, such as autonomous driving and safety. The existing systems, such as the mobile mapping system, depend on ground-level geospatial data, which create a considerable hurdle to executing continuous updates. Traffic and the high cost of these systems still cannot be overlooked. On the other hand, existing aerial photograph technologies can be considered as a resource to generating road models. However, it is reported that sufficient resolution cannot be retained with an aerial photograph. Therefore, the CfSM system and its implementation in this paper verified its

Table 1. Comparative analysis between existing- and CfSM systems

Comparisons	Existing systems	CfSM system
Sensor cost	Approximately 500,000 USD PMS[a] and 1,000,000 USD MMS[b]	50,000 USD or less
Data acquisition	Single lane for each attempt	Double lanes for each flight, and maximum of ten lanes from multidirectional flight
Traffic interference	Yes	No
System immediacy	Marginal	Faster
System interoperability	Intranet system	Big data-based cloud platform

[a]Pavement management system.
[b]Mobile mapping system.

feasibility as an ideal technology to generate the local accuracy for the HD map.

The computational contribution of the CfSM system stems from improved detection accuracy. The two steps of the DL model training confirmed the significance of the compatibility between the detection target and the data attribute. The first dataset (i.e., COWC) consisted of oblique images, while the second was composed of orthogonal images. Despite the fact that the first dataset (280,716 objects) was larger than the second (17,907 objects), the results demonstrated that orthogonal image sets are better suited for training the DL model in the CfSM system. The detection accuracy of the bus, for example, improved significantly after further labeling with orthogonal image sets, with a mAP of 76.51% compared to 3.16% on the first attempt.

Another core novelty of the CfSM system is its capacity in preventative mode. Most of the existing systems are in a reactive mode, which means they operate when a problem is detected or reported. The higher cost of these systems accounts for this reactiveness. On the other hand, the CfSM proposes an inexpensive alternative with higher performance than the existing systems. Table 1 summarizes a comparative analysis between existing systems and the CfSM system. The CfSM system costs less than a tenth of two existing systems, but still performed well by showing a resolution accuracy of under 10 mm. Compared to the existing aerial photography systems, the CfSM operates at a lower altitude, which helped obtain higher precision.

The range of data acquisition is another critical difference. The proposed CfSM system adopts a UAV, which allows for a wider vision of data acquisition. It was able to cover multiple lanes using the UAV's multidirectional flight, while two existing systems, PMS and MMS, can only do a single lane in each attempt. More importantly, the UAV-based acquisition did not interfere with the passing traffic at all, but we could not avoid the traffic with the two existing systems. Thus, the immediacy of the CfSM system shows higher performance than the two other models.

The data and surface model by the CfSM system can be further developed by a big data-based cloud platform. This ongoing research project will next focus on deep diagnostics for detecting potholes and cracks on car-free images using a convolutional neural network (CNN), which is a deep learning application used to assess the road condition and quality by a self-learning scheme. The CNN is already a verified method, so we can expect a promising outcome with the CfSM system and acquired orthoimage/data. As a starting point, the present serviceability index (PSI) must be set up, and a labeling process will be required in accordance with a pre-set PSI.

Conclusions and Research Limitation

A highly accurate is more than required than ever before. The map, such as a high-definition map (HD map), is an essential part of today's innovative development in the infrastructure and transportation industry. This project presents a review of relevant works and suggests a newly developed system to generate the road surface model, called the CfSM system. Field implementation was also presented as system verification.

The principle of the CfSM system stems from multi-angled aerial orthoimages by general-purpose drones and post-processing through DL techniques. The proposed system has been implemented and tested over two urban areas, Yeouido and Sangam in Seoul, South Korea, and provided comprehensive car-free images. The detection accuracies achieved by the developed DL model in this project were 89.57% for trucks, 95.77% for passenger vehicles, and 76.51% for buses, which came from 298,623 labeled objects on 9,331 images. In addition, the system was able to perform 10-mm accuracy of resolution from the ground level. In particular, the total cost was only 5% to 10% of the existing systems, such as pavement management system (PMS) and mobile mapping system (MMS). The developed CfSM system was successfully verified by the generated models of two areas in Seoul and the detection accuracy test, in the "CfSM System Verification" and "Implication, Discussion and Future Works" sections.

These vehicle-free images can serve as a comprehensive record of road conditions for the maintenance and management of roads as a viable alternative to on-site road surveys. In addition, they can be used as important pieces of information for building a basic roadmap for autonomous vehicles, which is one of the pivotal industries in the era of artificial intelligence. The limitation of the project is that the generated photogrammetric model is based on 2D images. This can cause difficulties in identifying the precise degree of faulty conditions on roads. Potential cracks, for instance, can be detected by this CfSM system regarding their length and width, but not their depth.

Data Availability Statement

All data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

This study was conducted with the Seoul Innovation Challenge 2017 (IC170011) and the Public Testbed for Technology Innovation 2018 (IU180005), which are projects of Seoul Metropolitan City and Seoul Business Agency (SBA), respectively. This paper is also a part of the project "Innovation in Construction Automation & Technologies" by the Australian Government through the Department of Foreign Affairs and Trade.

References

Abbondati, F., S. A. Biancardo, R. Veropalumbo, and G. Dell'Acqua. 2021. "Surface monitoring of road pavements using mobile crowdsensing technology." *Measurement* 171 (Feb): 108763. https://doi.org/10.1016/j.measurement.2020.108763.

Aggarwal, C. C. 2018. *Outlier analysis*. New York: Springer.

Alessandroni, G., L. Klopfenstein, S. Delpriori, M. Dromedari, G. Luchetti, B. Paolini, A. Seraghiti, E. Lattanzi, V. Freschi, and A. Carini. 2014. "Smartroadsense: Collaborative road surface condition monitoring." In *Proc., UBICOMM*: 210–215. Wilmington, DE: International Academy, Research, and Industry Association.

Alfarrarjeh, A., D. Trivedi, S. H. Kim, and C. Shahabi. 2018. "A deep learning approach for road damage detection from smartphone images." In *Proc., 2018 IEEE Int. Conf. on Big Data (Big Data)*. New York: IEEE.

Alqudah, Y. A., and B. H. Sababha. 2017. On the analysis of road surface conditions using embedded smartphone sensors." In *Proc., 2017 8th Int. Conf. on Information and Communication Systems (ICICS)*. New York: IEEE.

Bhatt, P., S. Gupta, P. Singh, and P. Dhiman. 2017. "Accident and road quality assessment using android Google maps API." In *Proc., 2017 Int. Conf. on Computing, Communication and Automation (ICCCA)*. New York: IEEE.

Chang, J.-R., H.-M. Hsu, and S.-J. Chao. 2012. "Development of a road monitoring and reporting system based on location-based services and augmented-reality technologies." *J. Perform. Constr. Facil.* 26 (6): 812–823. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000272.

Chen, K., M. Lu, G. Tan, and J. Wu. 2013. "CRSM: Crowdsourcing based road surface monitoring." In *Proc., 2013 IEEE 10th Int. Conf. on High Performance Computing and Communications & 2013 IEEE Int. Conf. on Embedded and Ubiquitous Computing*. New York: IEEE.

Chowdhury, A., G. Karmakar, J. Kamruzzaman, and S. Islam. 2020. "Trustworthiness of self-driving vehicles for intelligent transportation systems in industry applications." *IEEE Trans. Ind. Inf.* 17 (2): 961–970. https://doi.org/10.1109/TII.2020.2987431.

Chuang, T.-Y., N.-H. Perng, and J.-Y. Han. 2019. "Pavement performance monitoring and anomaly recognition based on crowdsourcing spatio-temporal data." *Autom. Constr.* 106 (Oct): 102882. https://doi.org/10.1016/j.autcon.2019.102882.

El-Wakeel, A. S., A. Osman, A. Noureldin, and H. S. Hassanein. 2017. "Road test experiments and statistical analysis for real-time monitoring of road surface conditions." In *Proc., GLOBECOM 2017-2017 IEEE Global Communications Conf.* New York: IEEE.

Eriksson, J., L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan. 2008. "The pothole patrol: Using a mobile sensor network for road surface monitoring." In *Proc., 6th Int. Conf. on Mobile Systems, Applications, and Services*. New York: Association for Computing Machinery.

Etten, A. V., 2019. "Satellite imagery multiscale rapid detection with windowed networks." In *Proc., 2019 IEEE Winter Conf. on Applications of Computer Vision (WACV)*, 735–743. New York: IEEE.

Gong, L., L. An, M. Liu, and J. Zhang. 2012. "Road damage detection from high-resolution RS image." In *Proc., 2012 IEEE Int. Geoscience and Remote Sensing Symp.* New York: IEEE.

Harikrishnan, P., and V. P. Gopi. 2017. "Vehicle vibration signal processing for road surface monitoring." *IEEE Sens. J.* 17 (16): 5192–5197. https://doi.org/10.1109/JSEN.2017.2719865.

He, K., X. Zhang, S. Ren, and J. Sun. 2016. "Deep residual learning for image recognition." In *Proc., IEEE Conf. on Computer Vision and Pattern Recognition*, 770–778. New York: IEEE.

Katsuma, R., and S. Yoshida. 2018. "Dynamic routing for emergency vehicle by collecting real-time road conditions." *Int. J. Commun. Network Syst. Sci.* 11 (2): 27–44. https://doi.org/10.4236/ijcns.2018.112003.

Kumar, R., A. Mukherjee, and V. Singh. 2017. "Community sensor network for monitoring road roughness using smartphones." *J. Comput. Civ. Eng.* 31 (3): 04016059. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000624.

Li, J., Q. Qin, H. Ma, and W. Yuan. 2010. Study on road damage assessment based on RS and GIS." In *Proc., 2010 IEEE Int. Geoscience and Remote Sensing Symp.* New York: IEEE.

Li, P., H. Xu, and B. Song. 2011. "A novel method for urban road damage detection using very high resolution satellite imagery and road map." *Photogramm. Eng. Remote Sens.* 77 (10): 1057–1066. https://doi.org/10.14358/PERS.77.10.1057.

Lin, T.-Y., P. Goyal, R. Girshick, K. He, and P. Dollár. 2017. Focal loss for dense object detection." In *Proc., IEEE Int. Conf. on Computer Vision*, 2980–2988. New York: IEEE.

Maeda, H., Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata. 2018. "Road damage detection and classification using deep neural networks with smartphone images." *Computer-Aided Civ. Infrastruct. Eng.* 33 (12): 1127–1141. https://doi.org/10.1111/mice.12387.

Malin, F., I. Norros, and S. Innamaa. 2019. "Accident risk of road and weather conditions on different road types." *Accid. Anal. Prev.* 122 (Jan): 181–188. https://doi.org/10.1016/j.aap.2018.10.014.

Mertz, C. 2011. "Continuous road damage detection using regular service vehicles." In *Proc., ITS World Congress*. Washington, DC: ITS America.

Mohan, P., V. N. Padmanabhan, and R. Ramjee. 2008. Nericell: Rich monitoring of road and traffic conditions using mobile smartphones." In *Proc., 6th ACM Conf. on Embedded Network Sensor Systems*. New York: Association for Computing Machinery.

Moon, S., N. Ham, S. Kim, L. Hou, J.-H. Kim, and J.-J. Kim. 2020. "Fourth industrialization-oriented offsite construction: Case study of an application to an irregular commercial building." *Eng. Constr. Architect. Manage.* 27 (9): 2271–2286.

Mundhenk, T. N., G. Konjevod, W. A. Sakla, and K. Boakye. 2016. "A large contextual dataset for classification, detection and counting of cars with deep learning." In *Proc., European Conf. on Computer Vision*. Amsterdam, Netherlands: Springer.

Park, J., K. Im, and J. H. Lee. 2019. "PotholeEye—How can we effectively maintain the pavement distress?" In *Proc., 17th Annual Int. Conf. on Mobile Systems, Applications, and Services*. New York: Association for Computing Machinery.

Saleh, S. M., S. Sugiarto, A. Hilal, and D. Ariansyah. 2017. "A study on the traffic impact of the road corridors due to flyover construction at Surabaya intersection." In *Proc., AIP Conf.* College Park, MD: AIP Publishing LLC.

Sammut, C., and G. I. Webb. 2011. *Encyclopedia of machine learning*. New York: Springer.

Singh, R. K., and S. Suman. 2012. "Accident analysis and prediction of model on national highways." *Int. J. Adv. Technol. Civ. Eng.* 1 (2): 25–30.

Strazdins, G., A. Mednis, G. Kanonirs, R. Zviedris, and L. Selavo. 2011. "Towards vehicular sensor networks with android smartphones for road surface monitoring." In *Proc., 2nd Int. Workshop on Networks of Cooperating Objects (CONET'11), Electronic Proc. CPS Week*. Riga, Latvia: Institute of Electronics and Computer Science.

Suda, Y., K. Nakano, H. Sugiyama, R. Hayashi, and S. Yamabe. 2010. "Estimation of road surface conditions from tire vibration." In *Proc., ASME Int. Mechanical Engineering Congress and Exposition*. New York: ASME.

SZ DJI Technology Co., Ltd. 2022a. "DJI Phantom 4 Pro - Product Information." Accessed January 2022. https://www.dji.com/au/phantom-4-pro/info#specs.

SZ DJI Technology Co., Ltd. 2022b. "DJI 9 Mavic 2 Pro - Product Information." Accessed January 2022. https://www.dji.com/au/mavic-2/info#specs.

Tang, T., J. Li, H. Huang, and X. Yang. 2014. "A car-following model with real-time road conditions and numerical tests." *Measurement* 48 (Feb): 63–76. https://doi.org/10.1016/j.measurement.2013.10.035.

Varona, B., A. Monteserin, and A. Teyseyre. 2019. "A deep learning approach to automatic road surface monitoring and pothole detection." *Personal Ubiquitous Comput.* 24: 519–534. https://doi.org/10.1007/s00779-019-01234-z.

Vittorio, A., V. Rosolino, I. Teresa, C. M. Vittoria, and P. G. Vincenzo. 2014. "Automated sensing system for monitoring of road surface quality by mobile devices." *Procedia-Social Behav. Sci.* 111 (Feb): 242–251. https://doi.org/10.1016/j.sbspro.2014.01.057.

Wang, J., Q. Qin, J. Zhao, X. Ye, X. Qin, X. Yang, J. Wang, X. Zheng, and Y. Sun. 2015. A knowledge-based method for road damage detection using high-resolution remote sensing image." In *Proc., 2015 IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*. New York: IEEE.

Yulianandha Mabrur, A. 2019. "Analisis pemanfaatan opensource dronedeploy dalam proses mozaik foto udara (UAV)." *Pawon: Jurnal Arsitektur* 3 (2): 79–92. https://doi.org/10.36040/pawon.v3i02.891.

Zekavat, P. R., S. Moon, and L. E. Bernold. 2014. "Performance of short and long range wireless communication technologies in construction." *Autom. Constr.* 47 (Nov): 50–61. https://doi.org/10.1016/j.autcon.2014.07.008.

Zekavat, P. R., S. Moon, and L. E. Bernold. 2015. "Holonic construction management: Unified framework for ICT-supported process control." *J. Manage. Eng.* 31 (1): A4014008. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000316.

Zhang, C. 2008. "An UAV-based photogrammetric mapping system for road condition assessment." *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* 37: 627–632.

Zhang, C., and A. Elaksher. 2012. "An unmanned aerial vehicle-based imaging system for 3D measurement of unpaved road surface distresses 1." *Comput.-Aided Civ. Infrastruct. Eng.* 27 (2): 118–129. https://doi.org/10.1111/j.1467-8667.2011.00727.x.

Zhang, C., N. R. Sabar, E. Chung, A. Bhaskar, and X. Guo. 2019. "Optimisation of lane-changing advisory at the motorway lane drop bottleneck." *Transp. Res. Part C: Emerging Technol.* 106 (Sep): 303–316. https://doi.org/10.1016/j.trc.2019.07.016.