Defectra: Road Defect Severity Assessment and Classification

A Special Problem Proposal
Presented to
the Faculty of the Division of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Visayas
Miag-ao, Iloilo

In Partial Fulfillment
of the Requirements for the Degree of
Bachelor of Science in Computer Science by

BELEBER, Benz Vrianne
CATALAN, Perserose
SENCIL, Kristian Lyle

Francis DIMZON
Adviser
Jumar CADONDON
Co-Adviser

November 18, 2024

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

According to the National Road Length by Classification, Surface Type, and Condition of the Department of Public Works and Highways (DPWH), as of October 2022 approximately 98.97

In an interview with the Road Board of DPWH Region 6 it was stated that road condition assessments are mostly done manually with heavy reliance on engineering judgment. In addition, manual assessment of roads is also time consuming which leaves maintenance operations to wait for lengthy assessments (J. Chua, Personal Interview. 16 September 2024). In a study conducted by Ramos, Dacanay, and Bronuela-Ambrocio (2022), it was found that the Philippines' current method of manual pavement surveying is considered as a gap since it takes an average of 2-3 months to cover a 250 km road as opposed to a 1 day duration in the Australian Road Research Board for the same road length. Ramos et al. (2022) recommended that to significantly improve efficiency of surveying methods and data gathering processes, automated survey tools are to be employed. It was also added that use of such automated surveying tools can also guarantee the safety of road surveyors (Ramos et al., 2022).

If the process of assessment on the severity of road defects can be automated then the whole process of assessing the quality of roads can be hastened up which can also enable maintenance operations to commence as soon as possible if necessary. If not automated, the delay of assessments will continue and roads that are supposedly needing maintenance may not be properly maintained which can affect the general public that is utilizing public roads daily.

1.2 Problem Statement

Roads support almost every aspect of daily life, from providing a way to transport goods and services to allowing people to stay connected with their communities. However, road defects such as cracks and potholes damage roads over time, and they can increase accident risks and affect the overall transportation. The current way of inspecting the roads for maintenance is often slow as it is done manually, which makes it harder to detect and fix defects early. The delay in addressing these problems can lead to even worse road conditions (J. Chua, Personal Interview. 16 September 2024). There are several research studies into automated road defect classification that have advanced in recent years but most of them focus on identifying the types of defects rather than assessing their severity or characteristics like depth. Without reliable data on the depth of the defect, road maintenance authorities may underestimate the severity of certain defects. To address these challenges, advancements are needed across various areas. An effective solution should not only detect and classify road defects but also measure their severity to better prioritize repairs. Failing to address this problem will require more extensive repairs for damaged roads, which raises the cost and strains the budget. Additionally, road maintenance would still be slow and cause disruptions in daily activities. Using an automated system that accurately detects, classifies, and assess the severity of road defects by incorporating depth are necessary to efficiently monitor road quality.

1.3 Research Objectives

1.3.1 General Objective

This special problem aims to develop an automated system that will accurately detect, classify, and assess the severity of the different types of road defects by using image analysis, depth measurement technologies, and combination of machine learning and computer vision techniques.

1.3.2 Specific Objectives

Specifically, this special problem aims:

1. To collect high-quality images of road surfaces that capture different types of

road defects including their depth in various lighting and weather conditions.

2. To develop and train a machine learning model to detect, classify, and assess the severity of road defects from images.

3. To measure the accuracy of the system by comparing the depth measurements against ground truth data collected from actual road inspections

4. To develop a prototype system that can detect and measure road defects from image input, analyze their depth, and assess their severity.

1.4 Scope and Limitations of the Research

This system will include a collection of images of different road defects, such as potholes and cracks, using cameras and depth-sensing tools. The images will be captured under various lighting and weather conditions to ensure that the data has variations. The scope is limited to visual and depth data. High-quality and diverse image data sets are essential for training an efficient model, and by focusing on capturing the depth, it will allow a more accurate assessment of severity of the road defects.

Depth measurement tools, such as LiDAR drones or stereo cameras will be used to record the depth of the road defect. Only accessible defects will be measured, any cracks and potholes filled with water may not be accurately assessed.

A machine learning model will be used to identify, classify, and assess the severity of road defects. It will use the image dataset to classify and assess the road defect types accurately, however, the effectiveness will depend on the quality and quantity of the training dataset. There can be a limited variety of images or inaccuracies due to environmental factors. The model will allow consistent and automated assessment of road defects which is more efficient than manual inspection.

The accuracy of the system will be evaluated by comparing the depth measurement it produces against data collected from the field through manual inspections. However, the comparisons could be limited to selected sample sites because collecting field data across a wide area can be time-consuming. Comparing the data is important to validate the reliability of the system. It ensures that the data that the system produces is accurate so it increases confidence in using it for road maintenance.

1.5 Significance of the Research

This special problem aims to be significant to the following:

Computer Science Community. This system can contribute to advancements in computer vision and machine learning by using both visual and depth data to assess the severity of road defects. It introduces a more comprehensive approach compared to the usual image-only or manual inspection methods. This combination can be applied to other fields that need both visual and depth analysis like medical imaging.

Concerned Government Agencies. This system offers a valuable tool for road safety and maintenance. Not only can this detect and classify anomalies, it can also assess the defect's severity which allows them to prioritize repairs, optimal project expenditures, and better overall road safety and quality.

Field Engineers. In the scorching heat, field engineers are no longer required to be on foot unless it requires its engineering judgement when surveying a road segment. It can hasten the overall assessment process.

Future Researchers. The special problem can serve as a baseline and guide of researchers with the aim to pursue special problems similar or related to this.

Chapter 2

Review of Related Literature

2.1 Related Literature

This section of the chapter presents related literature that is considered essential for the development of this special problem.

2.1.1 Deep Learning

Kelleher (2019) states that deep learning is inclined on making large-scale neural networks geared towards creating data-driven decisions. Furthermore, it was also argued that deep learning is oriented towards large-scale, complex data.

2.1.2 YOLOv5

According to Solawetz (2024), YOLOv5 is a model from a family of computer vision models used for object detection. YOLOv5 is reported to perform comparably to state-of-the-art techniques. It is designed to extract features from raw input images, used primarily in training object detection models alongside various data augmentation techniques.

2.1.3 Image and Video Processing

Kumar (2024) defines image processing as a process of turning an image into its digital form and extracting data from it through certain functions and operations. Usual processes are considered to treat images as 2D signals wherein different processing methods utilize these signals. Like image processing, Riches Resources (2020) defines video processing as being able to extract information and data from video footage through signal processing methods. However, in video processing due to the diversity of video formats, compression and decompression methods are often expected to be performed on videos before processing methods to either increase or decrease bitrate.

2.1.4 LiDAR

Wasser (2024) describes LiDAR as a technology utilized to measure the depth of a point from a certain height through its active remote sensing. During this process, a LiDAR measures the distance traveled through the time an emitted light takes to travel to the ground and back. Wasser (2024) states that this measured distance is converted into elevation.

2.2 Related Studies

This section of the chapter presents related studies conducted by other researchers wherein the methodology and technologies used may serve as basis in the development of this special problem.

2.2.1 Automated Detection and Classification of Road Anomalies in VANET Using Deep Learning

In the study of Bibi et al. (2021) it was noted that identification of active road defects are critical in maintaining smooth and safe flow of traffic. Detection and subsequent repair of such defects in roads are crucial in keeping vehicles using such roads away from mechanical failures. The study also emphasized the growth in use of autonomous vehicles in research data gathering which is what the researchers utilized in data gathering procedures. With the presence of autonomous vehicles, this allowed the researchers to use a combination of sensors and deep neural networks in deploying artificial intelligence. The study aimed to allow autonomous vehicles to avoid critical road defects that can possibly lead to dangerous situations. Researchers used Resnet-18 and VGG-11 in automatic detection and classification of road defects. Researchers concluded that the trained model was able to perform better than other techniques for road defect detection (Bibi et al., 2021). The study is able to provide the effectiveness of using deep learning models in training artificial intelligence for road defect detection and classification. However, the study lacks findings regarding the severity of detected defects which is crucial in automating manual procedures of road surveying in the Philippines.

2.2.2 Smartphones as Sensors for Road Surface Monitoring

In their study, Chapman, Li, and Sattar (2018) noted the rise of sensing capabilities of smartphones which they utilized in monitoring road surface to detect and identify anomalies. The researchers considered different approaches in detecting road surface anomalies using smartphone sensors. One of which are threshold-based approaches which was determined to be quite difficult due to several factors that are affecting the process of determining the interval length of a window function in spectral analysis (Chapman et al., 2018). The researchers also utilized a machine learning approach adapted from another study. It was stated that k-means was used in classifying sensor data and in training the SVM algorithm. Due to the requirement of training a supervised algorithm using a labeled sample data was required before classifying data from sensors, the approach was considered to be impractical for real-time situations (Chapman et al., 2018). In addition, Chapman et al. (2018) also noted various challenges when utilizing smartphones as sensors for data gathering such as sensors being dependent on the device's placement and orientation, smoothness of captured data, and the speed of the vehicle it is being mounted on. Lastly, it was also concluded that the accuracy and performance of using smartphone sensors is challenging to compare due to the limited data sets and reported algorithms.

2.2.3 Road Anomaly Detection through Deep Learning Approaches

The study of Guo, Luo, and Lu (2020) aimed to utilize deep learning models in classifying road anomalies. The researchers used three deep learning approaches namely Convolutional Neural Network, Deep Feedforward Network, and Recurrent Neural Network. In comparing the performance of the three deep learning approaches, the researchers fixed some hyperparameters. Results revealed that the RNN model was the most stable among the three and in the case of the CNN and DFN models, the researchers suggested the use of wheel speed signals to ensure accuracy. And lastly, the researchers concluded that the RNN model was best due to high prediction performance with small set parameters (Guo et al., 2020).

2.2.4 Road Surface Quality Monitoring Using Machine Learning Algorithms

The study of Bansal et al. (2021) aimed to utilize machine learning algorithms in classifying road defects as well as predict their locations. Another implication of the study was to provide useful information to commuters and maintenance data for authorities regarding road conditions. The researchers gathered data using various methods such as smartphone GPS, gyroscopes, and accelerometers. Bansal et al. (2021) also argued that early existing road monitoring models are unable to predict locations of road defects and are dependent on fixed roads and static vehicle speed. Neural and deep neural networks were utilized in the classification of anomalies which was concluded by the researchers to yield accurate results and are applicable on a larger scale of data (Bansal et al., 2021). The study of Bansal et al. (2021) can be considered as an effective method in gathering data about road conditions. However, it was stated in the study that relevant authorities will be provided with maintenance operation and there is no presence of any severity assessment in the study. This may cause confusion due to a lack of assessment on what is the road condition that will require extensive maintenance or repair.

2.2.5 Assessing Severity of Road Cracks Using Deep Learning-Based Segmentation and Detection

In the study of Ha, Kim, and Kim (2022), it was argued that the detection, classification, and severity assessment of road cracks should be automated due to the bottleneck it causes during the entire process of surveying. For the study, the researchers utilized SqueezNet, U-Net, and MobileNet-SSD models for crack classification and severity assessment. Furthermore, the researchers also employed separate U-nets for linear and area cracking cases. For crack detection, the researchers followed the process of pre-processing, detection, classification. During preprocessing images were smoothed out using image processing techniques. The researchers also utilized YOLOv5 object detection models for classification of pavement cracking wherein the YOLOv51 model recorded the highest accuracy.

The researchers however stated images used for the study are only 2D images which may have allowed higher accuracy rates. Furthermore, the researchers suggest incorporating depth information in the models to further enhance results.

2.2.6 Roadway pavement anomaly classification utilizing smartphones and artificial intelligence

The study of Christodoulou,Dimitrio, and Kyriakou (2016) presented what is considered as a low-cost technology which was the use of Artificial Neural Networks in training a model for road anomaly detection from data gathered by smartphone sensors. The researchers were able to collect case study data using two-dimensional indicators of the smartphone's roll and pitch values. In the study's discussion, the data collected displayed some complexity due to acceleration and vehicle speed which lead to detected anomalies being not as conclusive as planned. The researchers also added that the plots are unable to show parameters that could verify the data's correctness and accuracy. Despite the setbacks, the researchers still fed the data into the Artificial Neural Network that was expected to produce two outputs which were "no defect" and "defect". The method still yielded above 90% accuracy but due to the limited number of possible outcomes in the data processing the researchers still needed to test the methodology with larger data sets and roads with higher volumes of anomalies.

2.2.7 Pothole Mapping and Patching Quantity Estimates using LiDAR-Based Mobile Mapping Systems

In the study of Ravi, Habib, and Bullock (2020) utilized LiDAR technology in order to propose pothole mapping methods for public agencies which was argued to be laborious and time-consuming manual classification and quantity estimation. The researchers of the study made use of a wheel-based mobile LiDAR system driven at speeds of 40 - 50 mph during data gathering. In order to ensure accuracy of collected data, the researchers made multiple drive-runs which allowed the comparison of scanned data between two sensors. In a given 3D point cloud, the researchers also presented a process of pinpointing identified pothole points and these pothole points are then classified into different clusters through a distance-based growing strategy. Analysis procedures are then done by clusters. The researchers however established a minimum volume threshold for potholes in a given bounding box. In classifying pothole points, points inside a segment or tile are used along with plane-fitting to determine planar position and orientation parameters. Individual points are then again used to extract information like signed normal distance from the best-fitting plane. Distance collected is then stated to be the depth beneath the road surface and the researchers classified potential potholes to have a depth that is greater than 1 cm.

Chapter 3

Methodology

This chapter outlines the systematic approach taken to address the problem of classifying and assessing road defects using artificial intelligence. The methodology is divided into key phases: data collection, data preprocessing, algorithm selection, system implementation, and ss. Each phase is essential to accurately classify and assess road defects.

3.1 Research Activities

3.1.1 Inquiry

The team conducted initial inquiries to understand the problem domain and existing road maintenance practices. This phase included consulting the engineers under the Road Maintenance Department of the government agency Department of Public Works and Highways (DPWH). This allowed te team to gather information about road defects classification and maintenance practices. The inquiries were conducted during the first week and involved an interview with Engr. Jane Chua of the said department, who provided us a copy of the official road maintenance manual. The manual allows this project to be aligned with the established standards and practices of the DPWH.

3.1.2 Brainstorming

Potential solutions, algorithms, and system architectures were discussed by the team and the special problem adviser in this phase. These sessions, conducted in-class and virtually via Zoom, helped narrow down the overview of the system, leading to the selection of the main architecture YOLOv5 for defect detection and Epipolar Spatio-Temporal Networks (ESTN) for depth estimation. YOLO

3.1.3 Algorithm Selection

Potential solutions, algorithms, and system architectures were discussed by the team and the special problem adviser in this phase. These sessions, conducted in-class and virtually via Zoom, helped narrow down the overview of the system, leading to the selection of the main architecture YOLOv5 for defect detection and Epipolar Spatio-Temporal Networks (ESTN) for depth estimation. YOLO

3.1.3.1 Defect Detection

YOLOv5 was selected for its balance of real-time processing capability and accuracy, essential for detecting road defects in dynamic environments.

3.1.3.2 Severity Assessment

The Multi-view Depth Estimation using Epipolar Spatio-Temporal Networks was selected due to the high cost and limited accessibility of LiDAR technology. By applying epipolar geometry and temporal consistency across sequential frames, this approach provides an accurate depth estimation from a standard video footage (Long et al., 2021).

3.1.4 Observation and Experimentation

3.1.5 Design and Testing

The system is tested using data gathered from ground truthing which involves manual inspection and measuring of road defects to verify the type, shape, and dimensions of the defect. These manual observations serve as a baseline reference to measure the system's accuracy in detecting, classifying, and severity assessment of road defects.

3.1.5.1 Challenges and Limitations

One major limitation is the availability of local labeled datasets, which affects the model's training, as most datasets available are those captured from foreign countries only.

3.1.5.2 Documentation

Documentation was conducted throughout the project, ensuring a detailed record of methods, results, and challenges. This documentation not only served as a basis for the final SP report but also provided transparency and reproducibility for future studies.

3.2 Calendar of Activities

Table 1 shows a Gantt chart of the activities. Each bullet represents approximately one week worth of activity.

Activities	**Aug**	**Sep**	**Nov**	**Dec**
Pre-proposal Preparation	· · · ·			
Literature Review	· · ·	·		
Data Collection	· ·	· ·		
Algorithm Selection		· ·		
System Design		·	· ·	
Prelimenary Testing			· ·	·
Documentation and SP Writing	· · · ·	· · · ·	· · · ·	· ·

Table 3.1: Timetable of Activities

References

Fedkiw, R., Stam, J., & Jensen, H. W. (2001). Visual simulation of smoke. In E. Fiume (Ed.), *Proceedings of siggraph 2001* (pp. 15–22). ACM Press / ACM SIGGRAPH.

Jobson, D. J., Rahman, Z., & Woodell, G. A. (1995). Retinex image processing: Improved fidelity to direct visual observation. In *Proceedings of the is&t fourth color imaging conference: Color science, systems, and applications* (Vol. 4, pp. 124–125).

Kartch, D. (2000). *Efficient rendering and compression for full-parallax computer-generated holographic stereograms* (Unpublished doctoral dissertation). Cornell University.

Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., ... Fulk, D. (2000). The digital michelangelo project. In K. Akeley (Ed.), *Proceedings of siggraph 2000* (pp. 131–144). New York: ACM Press / ACM SIGGRAPH.

Park, S. W., Linsen, L., Kreylos, O., Owens, J. D., & Hamann, B. (2006, March/April). Discrete sibson interpolation. *IEEE Transactions on Visualization and Computer Graphics*, *12*(2), 243–253.

Parke, F. I., & Waters, K. (1996). *Computer facial animation*. A. K. Peters.

Pellacini, F., Vidimče, K., Lefohn, A., Mohr, A., Leone, M., & Warren, J. (2005, August). Lpics: a hybrid hardware-accelerated relighting engine for computer cinematography. *ACM Transactions on Graphics*, *24*(3), 464–470.

Sako, Y., & Fujimura, K. (2000). Shape similarity by homotropic deformation. *The Visual Computer*, *16*(1), 47–61.

Yee, Y. L. H. (2000). *Spatiotemporal sensistivity and visual attention for efficient rendering of dynamic environments* (Unpublished master's thesis). Cornell University.