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

April 28, 2025

Abstract

Road surveying is a crucial part of the maintenance processes of roads in the Philippines that is carried out by the Department of Public Works and Highways. However, the current process of road surveying is time consuming which delays much needed maintenance operations. Existing studies involving automated pothole detection lack integration of the pothole's depth in assessing its severity which is essential for automating road surveying procedures. A system that incorporates estimated depth information in assessing pothole severity is developed in order to automate the manual process of depth measurement and severity assessment in road surveying. For depth estimation, stereo vision is favorable in this context as depth may be estimated through the disparity generated by a stereo pair. In obtaining a stereo view of the potholes, the StereoPi V2 is utilized along with some modifications that would make it eligible for outdoor use. To address camera imperfections, a fitted inverse model was applied to improve the accuracy of depth estimates. Linear regression analysis revealed a strong positive correlation between estimated and actual depths, with the system measuring pothole depths mostly within 2 cm of the true values.

Keywords: pothole, depth estimation, stereo vision, StereoPi V2

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% of roads in the Philippines is paved which is either made of concrete or asphalt (DPWH, 2022). Since the DPWH is an institution under the government, it is paramount to maintain such roads in order to avoid accidents and congested traffic situations especially in heavily urbanized areas where there are a lot of vehicles.

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, & Bronuela-Ambrocio, 2023), 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., 2023).

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.

Existing studies involving road defects such as potholes mainly focus on the detection of potholes using deep learning models and almost not considering the severity of detected potholes or did not incorporate any depth information from potholes (Ha, Kim, & Kim, 2022; Kumar, 2024; Bibi et al., 2021). In addition, for studies that include severity assessment on potholes, the main goal of the study is not directed towards road maintenance automation but other factors such as improvement of ride quality for the vehicle. Another issue found in existing solutions is the lack of incorporation to the context of Philippine roads. With these issues in mind, the study aims to utilize stereo vision from StereoPi V2 in order to obtain multi-perspective views of detected potholes to be used in severity asssessment by focusing on estimating the depth of individual potholes for automated road condition monitoring.

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 assess the severity of road defects by incorporating depth is necessary to efficiently monitor road quality.

1.3 Research Objectives

1.3.1 General Objective

This special problem aims to develop a system that accurately estimates the depth of potholes on road surfaces by using image analysis, depth measurement technologies, and computer vision techniques. The system will focus specifically on measuring the depth of potholes to assess their severity, enabling faster and more accurate road maintenance decisions, and there are no current practices in the Philippines involving depth information of potholes in assessing their severity. In accordance with the Department of Public Works and Highways Region 6's manual for road maintenance, the study will classify potholes into different severity levels such as low, medium, and high, which will be primarily based on their depth. In order to measure the system's accuracy, linear regression in order to represent the difference between the depth calculated from the disparity and the actual depth of the pothole from ground truth data.

1.3.2 Specific Objectives

Specifically, this special problem aims:

1. To collect high-quality stereo images of road surfaces that capture potholes including their depth in favorable conditions
2. To measure the accuracy of the system by comparing the depth measurements against ground truth data collected from actual road inspections and to utilize linear regression, root mean square error, and mean absolute error as a metric for evaluation.
3. To develop a prototype system that can detect and measure road potholes from image input, analyze their depth, and assess their severity.

1.4 Scope and Limitations of the Research

This system focuses solely on detecting and assessing the severity of potholes through image analysis and depth measurement technologies. The scope includes the collection of pothole images using cameras and depth-sensing tools under a favorable weather condition.

Depth-sensing tools, such as stereo cameras, will be used to record the depth of potholes specifically. The system will not address other road defects like cracks or other surface deformations; therefore, it will detect and analyze only potholes. Additionally, only accessible potholes will be measured, meaning those that are filled with water or obscured by debris may not be accurately assessed.

The system developed focuses exclusively on detecting potholes and assessing their severity through depth measurement. The accuracy of the system's depth measurements is evaluated by comparing them against data collected from actual field inspections. However, this comparison is limited to selected sample sites, as collecting field data over a large area can be time-consuming and resource-intensive.

Environmental factors such as lighting, road surface texture, and weather conditions may impact the system's performance. The accuracy and reliability of the system will depend on the quality of camera calibration and disparity map finetuning. Its ability to measure the depth of pothole images needs careful validation.

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 their engineering judgement when surveying a road segment. It can hasten the overall assessment process.

Future Researchers. The special problem may 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 Frameworks

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

2.1.1 Depth Estimation

Depth estimation as defined by Sanz, Mezcua, and Pena (2012) as a set of processes that aims to extract a representation of a certain scene's spatial composition. Stereo vision is stated to be among the depth estimation strategies (Sanz et al., 2012).

2.1.2 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, 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.

2.1.3 Stereo Vision

MathWorks (n.d.) defines stereo vision as a process of utilizing multiple 2D perspectives in order to extract information in 3D. In addition, most uses of stereo vision involve estimating an objects distance from an observer or camera. The 3D information is stated to be extracted with stereo pairs or pair of images through estimation of relative depth of points in a scene which are then represented through a stereo map that is made through the matching of the pair's corresponding points.

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 Deep Learning Studies

2.2.1.1 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 and incorporation of pothole depth in their model which are both crucial in automating manual procedures of road surveying in the Philippines.

2.2.1.2 Road Anomaly Detection through Deep Learning Approaches

The study of Luo, Lu, and Guo (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 from data collected through the sensors in the vehicle's suspension system. 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 (Luo et al., 2020). However, proper severity assessment through depth information was not stated to be utilized in any of the three approaches used in the study.

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

In the study of Ha et al. (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.

2.2.1.4 Roadway pavement anomaly classification utilizing smartphones and artificial intelligence

The study of Kyriakou, Christodoulou, and Dimitriou (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.

2.2.2 Machine Learning Studies

2.2.2.1 Smartphones as Sensors for Road Surface Monitoring

In their study, Sattar, Li, and Chapman (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 (Sattar 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 (Sattar et al., 2018). In addition, Sattar 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.

2.2.2.2 Road Surface Quality Monitoring Using Machine Learning Algorithms

The study of Singh, Bansal, Kamal, and Kumar (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. (Singh 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 (Singh et al., 2021). The study of Singh 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.

2.2.3 Computer Vision Studies

2.2.3.1 Stereo Vision Based Pothole Detection System for Improved Ride Quality

In the study of Ramaiah and Kundu (2021) it was stated that stereo vision has been earning attention due to its reliable obstacle detection and recognition. Furthermore, the study also discussed that such technology would be useful in improving ride quality in automated vehicles by integrating it in a predictive suspension control system. The proposed study was to develop a novel stereo vision based pothole detection system which also calculates the depth accurately. However, the study focused on improving ride quality by using the 3D information from detected potholes in controlling the damping coefficient of the suspension system. Overall, the pothole detection system was able to achieve 84% accuracy and is able to detect potholes that are deeper than 5 cm. The researchers concluded that such system can be utilized in commercial applications. However, it is also worth noting that despite the system being able to detect potholes and measure its depth, the overall severity of the pothole and road condition was not addressed.

2.3 Chapter Summary

The reviewed literature involved various techniques and approaches in road anomaly detection and classification. These approaches are discussed and summarized below along with their limitations and research gaps.

Study	Technology/ Techniques Used	Key Findings	Limitations
Automated Detection and Classification of Road Anomalies in VANET Using Deep Learning	Resnet-18 and VGG-11	Trained model is able to provide the effectiveness of using deep learning models in training artificial intelligence for road defect detection and classification.	Lacks findings regarding the severity of detected defects.
Smartphones as sensors for Road surface monitoring	Machine Learning, Smartphones	Approach was considered impractical for real-life applications.	Sensors are dependent on device's placement and orientation, smoothness of data, and speed of vehicle it is mounted on. Accuracy of results is difficult to compare.
Road Anomaly Detection through Deep Learning Approaches	Convolutional Neural Network, Deep Feedforward Network, and Recurrent Neural Network	Identified that RNN was the best deep learning approach due to high prediction performance.	Data collection is considered too difficult and complicated to execute due to sensors being mounted on an integral part of the vehicle.
Assessing Severity of Road Cracks Using Deep Learning-Based Segmentation and Detection	SqueezNet, U-Net, YOLOv5, and MobileNet-SSD models	YOLOv51 model recorded the highest accuracy.	Only 2D images are used for the study which may have allowed higher accuracy rates, and the study also lacked depth information.
Stereo Vision Based Pothole Detection System for Improved Ride Quality	Pair of stereo images captured by a stereo camera	System was able to achieve 84% accuracy and is able to detect potholes that are deeper than 5 cm.	Overall severity of the pothole and road condition was not addressed.

Table 2.1: Comparison of Related Studies on Road Anomaly Detection using Deep Learning Techniques and Stereo Vision

Chapter 3

Methodology

This chapter outlines the systematic approach that were taken to address the problem of pothole depth estimation using StereoPi V2. The methodology is divided into key phases: data collection, algorithm selection, design, testing and experimentation, and challenges and limitations. Each phase will play a crucial role in accurately classifying and assessing road defects. Each phase is essential for accurately estimating the depth of potholes using StereoPi V2.

3.1 Research Activities

3.1.1 Data Collection

The researchers 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). An interview with Engr. Jane Chua provided a comprehensive overview of the DPWH's road maintenance manual, which was crucial in aligning this project with existing standards. This collaboration with DPWH provided insights into road pothole classification standards, ensuring that the collected data will align with industry standards. The DPWH manual primarily focuses on the volume of detected potholes within a road segment as a measure of severity. However, since depth is not explicitly measured in their current procedures, the study will supplement this by referencing international standards such as the Long-Term Pavement Performance (LTPP) classification used in the United States. The LTPP categorizes potholes based on depth thresholds, which will be integrated with DPWH's volume-based assessment to provide a more comprehensive severity classification framework. The data collection involved capturing around 130 images of potholes from various locations within the UP Visayas Campus. Ground truth data of pothole depth were collected by the researchers by measuring the depth of different points in an individual pothole and then solving for its average depth. The aforementioned process was validated by Engr. Benjamin Javellana, Assistant Director of the DPWH Regional Office 6 Maintenance Division. In order to individually locate or determine each pothole where the ground truth data is collected, images taken were labeled with their corresponding coordinates, street names, and nearby landmarks.

3.1.1.1 Data Collection (Ground Truth Data)

The researchers collected depth information from 130 potholes around the University of the Philippines Visayas Miagao Campus. During data collection, the researchers are equipped with safety vests and an early warning device to give caution to incoming vehicles. To measure the depth of each pothole, the researchers recorded four depth points within the pothole and calculated their average.

3.1.2 Algorithm Selection

Potential solutions, algorithms, and system architectures were discussed by the researchers 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 Epipolar Spatio-Temporal Networks (ESTN) for depth estimation.

3.1.3 Design, Testing, and Experimentation

This section outlines both the design and testing of the system, as well as the experimentation process to validate the selected methodologies.

3.1.3.1 Materials and Equipment

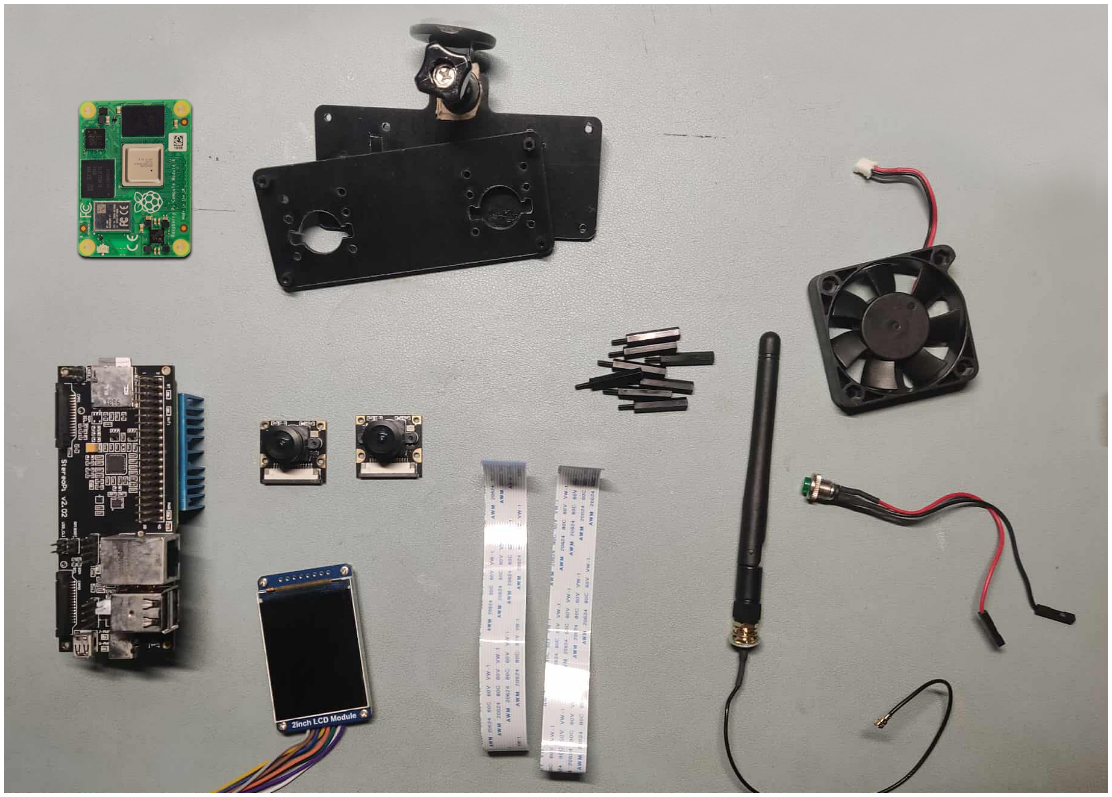

Figure 3.1: Components used in the prototype development.

The prototype system was constructed using the following materials and components:

- StereoPi V2 Board
- Raspberry Pi Compute Module 4 (CM4)
- Dual RaspberryPi Camera Modules with Fisheye Lens
- 3D Printed Custom Housing
- 2-inch LCD Module
- Micro SD Card
- Antenna
- Momentary Push Button

3.1.3.2 Prototype Building

The prototype involved the StereoPi V2 Kit which was acquired through an official international distributer. After assembling the camera, it was further modified to address the it's heating by incorporating a heat sink and a small computer fan to make it suitable for outdoor use.

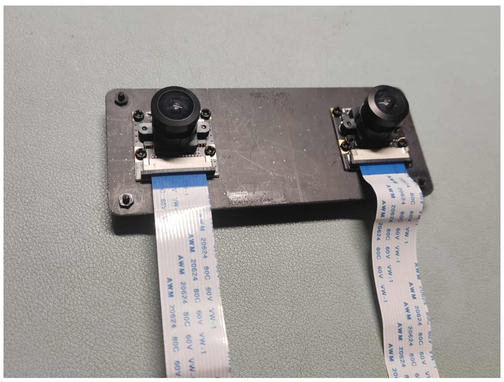

Figure 3.2: Dual RPi Camera Modules attached to the custom housing.

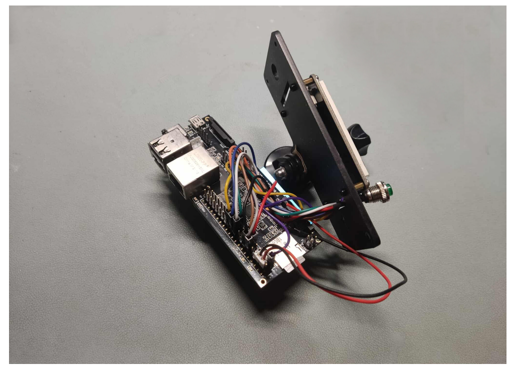

Figure 3.3: LCD Module connected to the StereoPi board.

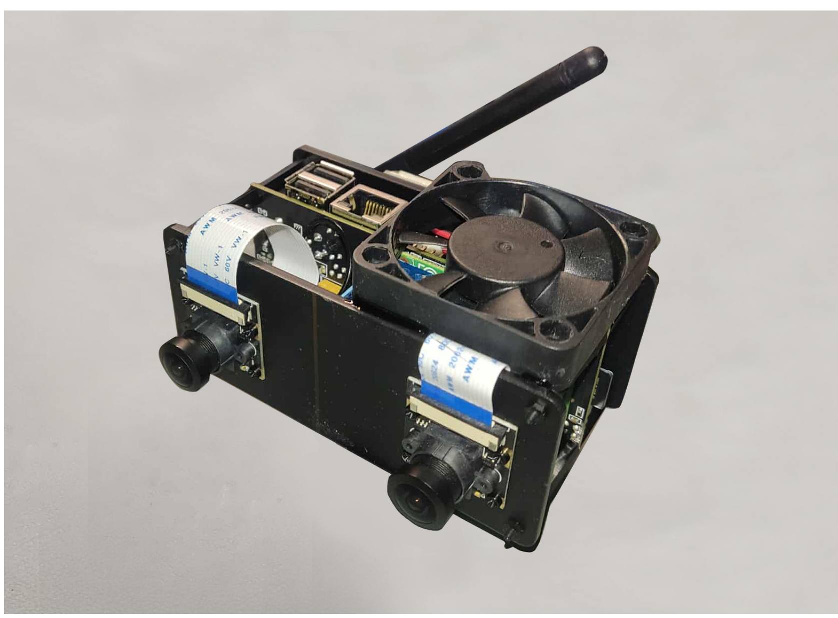

Figure 3.4: The finished prototype.

3.1.3.3 Camera Calibration (Fisheye Distortion)

The StereoPi V2 is first calibrated using a 9 by 6 checkerboard, with a checker size of 55mm, from different angles through calibration scripts that came with the package. This process ensured that the camera is working properly in capturing stereo imagery. This removed distortion from captured images allowing depth estimation with more accuracy.

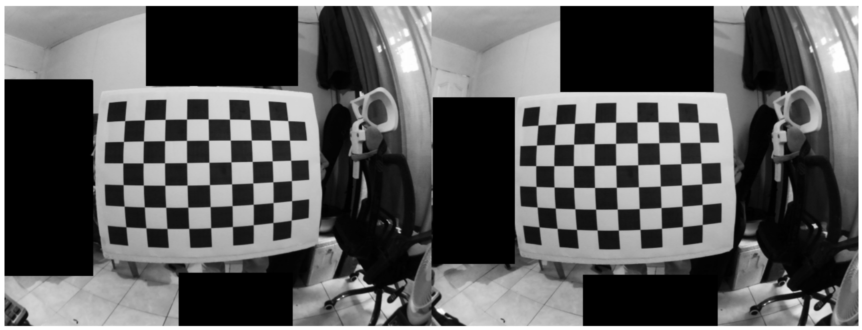

Figure 3.5: Calibration process with a checkerboard to correct fisheye lens distortion.

3.1.3.4 Camera Calibration (Disparity Map Fine-tuning)

The stereo image pairs captured by the system were first rectified to ensure proper alignment of corresponding features. Block matching parameters were then fine-tuned to produce clearer and more accurate disparity maps. It was observed that the effective operational range of the stereo camera system extends from approximately 30 to 80 cm. At distances closer than 30 cm, the disparity maps exhibited significant noise, while at distances beyond 80 cm, disparity information became sparse or blank.

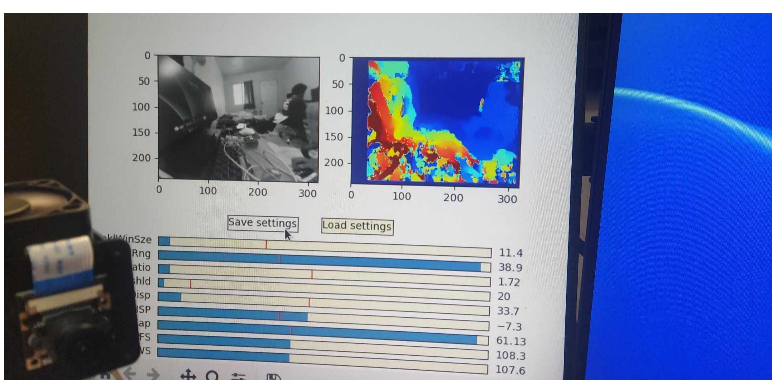

Figure 3.6: Parameter tuning process to achieve cleaner and more accurate disparity maps.

3.1.3.5 Initial Testing

Initial testing was conducted to verify the functionality and basic accuracy of the stereoscopic camera system in a controlled environment. Artificial potholes with known depths were created to simulate varying real-world scenarios. The system captured disparity maps, and estimated depths were computed using the standard stereo camera depth formula.

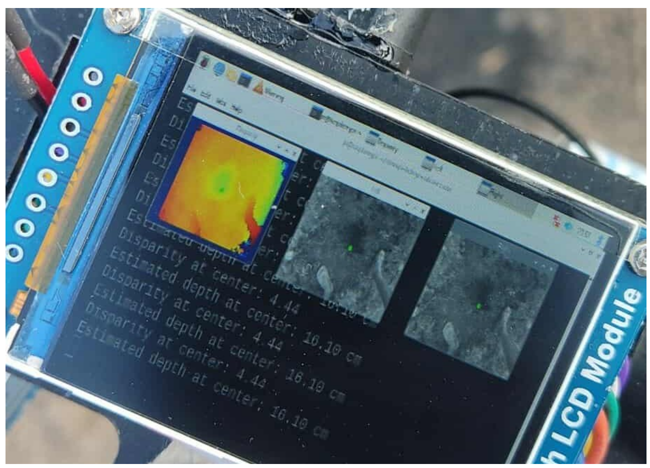

Figure 3.7: The system tested on a simulated pothole.

However, the results revealed a non-linear relationship between the computed disparity values and the actual distances. This discrepancy indicated that the traditional depth estimation method was insufficient for the current setup. To address this, the researchers collected multiple data points and correlating known distances to their respective disparity readings and fitted an inverse model to better represent the system's behavior (see Figure 4.4). This updated disparity-to-depth model was subsequently used in the final testing phase.

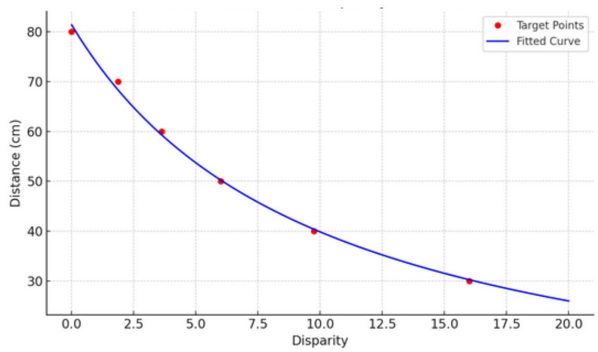

Figure 3.8: Inverse Model Fit to Disparity vs. Distance.

3.1.3.6 Performance Metrics

The accuracy of the pothole depth estimated by StereoPi V2 was analyzed using Linear Regression in order to model the difference between the disparity and distance. The lower the disparity indicates that the pothole is deeper.

3.1.3.7 Final Testing and Validation

The testing process began with a detailed testing plan that includes both simulated and real-world testing scenarios. Initially, the system is tested in controlled environments to ensure it can estimate pothole depth effectively. Following this, real-world testing was conducted using the StereoPi kit on previously located potholes, specifically at the University of the Philippines Visayas Miagao Campus. The system's performance was validated by comparing its predictions with ground-truth data collected from manual inspections.

3.1.3.8 Documentation

Throughout the research activities, thorough documentation was maintained. This documentation captured all methods, results, challenges, and adjustments made during the experimentation phases. It ensured the reproducibility of the work and provided transparency for future research endeavors.

3.1.4 Challenges and Limitations

3.1.4.1 Camera Limitations

During the data collection process, the researchers were faced with various issues involving the StereoPi V2 kit. Issues like inaccuracies in the rectified image pair and generated disparity map were very apparent in the early stages of data collection due to limited related studies and literature involving the camera. In addition, the camera also yielded some inaccurate depth estimation and over reliance on controlled environments which prompted the researchers to further improve its tuning and calibration.

3.2 Calendar of Activities

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

Table 3.1: Timetable of Activities for 2024

Activities (2024)	Aug	Sept	Oct	Nov	Dec
Pre-proposal Preparation	**W4**				
Literature Review	**W3**	**W1**			
Data Collection	**W2**	**W2**			
Algorithm Selection		**W2**			
System Design		**W1**	**W2**	**W2**	
Preliminary Testing				**W2**	**W1**
Documentation and SP Writing	**W4**	**W4**	**W4**	**W4**	**W2**

Table 3.2: Timetable of Activities for 2025

Activities (2025)	Jan	Feb	Mar	Apr	May	Jun
Data Collection	**W4**					
System Design	**W3**	**W2**	**W2**			
Model testing	**W3**	**W4**	**W4**			
Results Analysis			**W2**	**W4**		
Conclusion Formulation				**W2**	**W3**	
Documentation and SP Writing	**W4**	**W4**	**W4**	**W4**	**W4**	**W2**

Chapter 4

Preliminary Results/System Prototype

This chapter presents the results on estimating the depth of potholes using the StereoPi system. It details the prototype construction, calibration of the system, and the application of regression analysis to improve depth estimation. It also contains the measurements taken during the testing phases, comparing the ground truth depths with the value estimated by the camera. Findings are presented systematically, supported by tables showing the collected data, images of the outputs, and discussion on the analysis of results.

4.1 System Calibration and Model Refinement

After the initial testing, the system was calibrated using a controlled setup, where artificial potholes with known depths were created. The stereo camera system captured disparity maps, from which depth was calculated using the standard stereo vision formula:

$$\text{Depth} = \frac{f \times B}{d}$$

where:

- f is the focal length in pixels,
- B is the baseline distance between the two cameras,

- d is the disparity.

However, preliminary observations revealed that the relationship between measured disparity and depth was shifted from the ideal. Their relationship is inherently nonlinear, specifically an inverse relationship (of the form y=1/xy=1/x). As disparity decreases, depth increases rapidly and nonlinearly. However, due to real-world factors such as lens distortion, imperfect calibration, stereo matching errors, and pixel quantization, the actual relationship between measured disparity and true depth often deviates from the theoretical ideal (Scharstein Szeliski, 2002). cite here

To address the nonlinear behavior, a curve fitting approach was introduced. Specifically, an inverse model was fitted to the collected data points, relating disparity and ground-truth distance measurements.

An inverse function of the form:

$$y = a + \frac{b}{x}$$

where:

- y is the estimated distance (in cm),
- x is the measured disparity,
- a and b are coefficients obtained through regression analysis.

4.2 Model Refinement Using Regression

The regression analysis produced the following model parameters:

- a = . . .
- b = ...

The model achieved the following performance on the test data:

Metric	Value
Mean Absolute Error (MAE)	X cm
Root Mean Square Error (RMSE)	X cm

Table 4.1: Performance Metrics for the Regression Model

The relatively low MAE and RMSE indicate that the fitted model significantly improved the accuracy of depth estimation compared to the original stereo formula.

4.3 Error Analysis

Despite the improvements, minor estimation errors remained. These errors were primarily attributed to:

- Low-light imaging conditions affecting disparity computation,
- Shallow potholes with depths less than 1 cm, where disparity resolution becomes a limiting factor,
- Perspective distortion when the stereo camera was not parallel to the ground plane.

4.4 Testing Results

Following calibration, actual potholes located around the University of the Philippines Visayas (UPV) campus were tested. The ground truth depths of the potholes were measured manually and compared with the depths estimated by the camera. Based on the results, the StereoPi camera was able to estimate the depths fairly close to the ground truth values.The smallest difference was seen in Pothole 5, where the estimated depth was only 0.24 cm away from the ground truth. The largest difference was found in Pothole 1, where the error was 3.45 cm. For the other potholes, the differences were 0.67 cm for Pothole 2, 2.07 cm for Pothole 3, and 2.66 cm for Pothole 4. Most of the time, the camera's estimated depths were only off by about one to three centimeters. Table 4.2 shows the comparison between the manually measured ground truth depths and the depths estimated by the StereoPi camera for each simulated pothole.

Table 4.2: Ground Truth and StereoPi Depth Measurements

Pothole	**Ground Truth 1 (cm)**	**Ground Truth 2 (cm)**	**Ground Truth Avg (cm)**	**Est Depth 1(cm)**	**Est Depth 2 (cm)**	**Est Depth Avg(cm)**	**Diff(cm)**
1	14.6	14.4	14.5	11.16	10.94	11.05	3.45
2	12.0	12.1	12.05	12.36	10.4	11.38	0.67
3	6.4	6.5	6.45	4.76	4.0	4.38	2.07
4	9.8	9.3	9.55	6.16	7.62	6.89	2.66
5	13.9	14.3	14.1	13.04	14.68	13.86	0.24

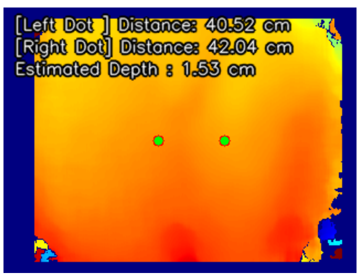

Figure 4.1: Disparity Map

Figure 4.2: Left Stereo Image

Figure 4.3: Right Stereo Image

4.5 Discussion

The Linear Regression test on the collected data revealed a strong positive linear relationship between the estimated and ground truth depths ($R = 0.937$). The coefficient of determination ($R^2 = 0.878$) also indicates that 87.8% of the differences in the estimated depth are correctly predicted based on the ground truth data. After calculating for the Mean Absolute Error, it was also found that estimated pothole depths differ from the actual ground truth data by around 1.82 cm. In addition, the Root Mean Square Error also revealed that the typical error size is at 1.19 cm.

R	**R^2**	**Root Mean Square Error**	**Mean Absolute Error**
0.937	0.878	1.19	1.82

Table 4.3: Linear Regression Model for Pothole Depth Estimation

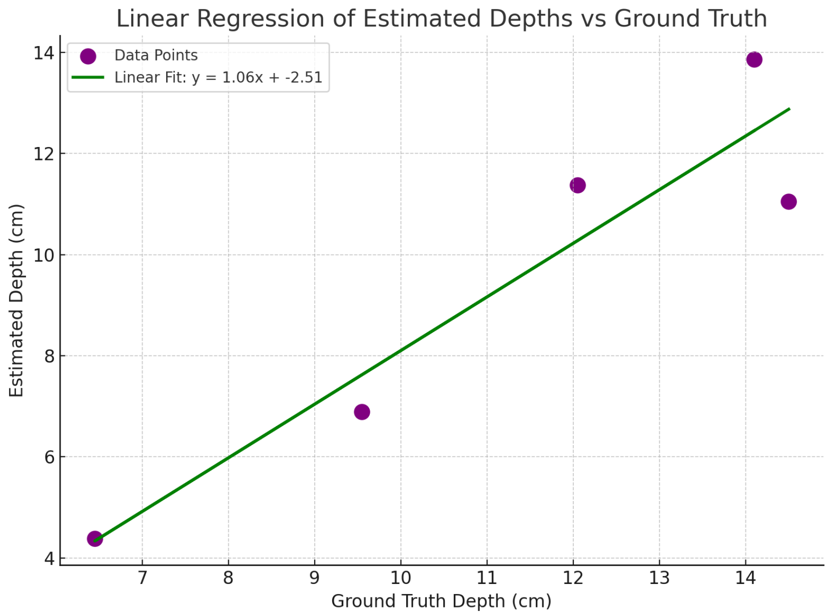

Figure 4.4: Inverse Model Fit to Disparity vs. Distance.

Chapter 5

Summary, Conclusions, Discussion, and Recommendations

This chapter provides conclusions based on the research findings from data collected on the development of a pothole depth estimation system using stereo vision technology. It also presents a discussion and recommendations for future research. This chapter reviews the purpose of the study, research questions, related literature, methodology, and findings. It then presents the conclusions, a discussion of the results, recommendations for practice, suggestions for further research, and the final conclusion of the study.

5.1 Summary

This special project addressed the critical issue of road maintenance by developing a system capable of estimating the depth of potholes to help prioritize repairs. The purpose of the project was to create an automated method that not only detects potholes but also assesses their severity based on depth, responding to the current manual and slow road inspection practices. The researchers aimed to collect high-quality images of potholes under varying conditions, to validate the system's depth estimation accuracy using ground truth measurements and linear regression analysis, and to build a working prototype using stereo vision that can detect, measure, and assess potholes.

To achieve these objectives, a hardware prototype was built using the StereoPi

V2 board and Raspberry Pi Compute Module 4, equipped with dual fisheye-lens cameras. Camera calibration was performed using a 9x6 checkerboard pattern with known square sizes to correct for fisheye lens distortion and ensure proper alignment of the stereo pair. After calibration, disparity map generation was fine-tuned by adjusting block matching parameters to produce clearer and more reliable disparity maps. Initial testing was conducted using simulated potholes with known depths to verify the functionality of the system and identify the non-linear behavior present in stereo vision depth measurements. It was observed that using the standard stereo depth formula led to inaccuracies, particularly at greater distances.

The calibrated system and fitted regression model were validated by comparing the estimated depths with the manually measured depths. The findings showed that the system was able to estimate pothole depths within approximately ± 2 cm of actual measurements, achieving a Mean Absolute Error (MAE) of 1.82 cm and a Root Mean Square Error (RMSE) of 1.19 cm. A strong positive linear relationship was observed between the estimated and actual depths (R = 0.937, $R^2 = 0.878$).

5.2 Conclusions

The researchers conclude the following based on the findings:

- The system effectively captures and analyzes depth information from stereo images, providing a viable method for automated pothole severity assessment.
- Incorporating depth measurements significantly improves pothole repair prioritization compared to traditional visual-only inspections, allowing maintenance decisions to be based on objective, measurable data.
- The system achieved an acceptable regression model fit, with a strong positive correlation (R = 0.937) and a coefficient of determination ($R^2 = 0.878$), confirming that the depth estimates closely align with the ground truth measurements.The system obtained satisfactory error metrics, with a Mean Absolute Error (MAE) of 1.82 cm and a Root Mean Square Error (RMSE) of 1.19 cm, indicating reliable performance for both pothole detection and depth estimation tasks.
- The proposed approach fills a critical gap in current road maintenance practices, especially within the Philippine context where depth-based severity

classification is not yet systematically implemented.

5.3 Discussion

The study found that stereo vision works effectively in helping estimate the depth of road potholes. The system built using the StereoPi V2 camera was able to measure pothole depths with results mostly within ±2 cm of the actual ground truth values. This matches the general observation in earlier studies (e.g., Ramaiah and Kundu, 2021), which showed that stereo vision can provide useful 3D information for road obstacle detection. However, this study advances previous work by focusing not just on detection, but on depth-based severity classification, which was largely missing in earlier research.

The outputs of the system were generally positive, showing that with proper calibration and tuning, consistent and reliable depth estimates can be produced. Calibration using checkerboards and tuning block matching parameters were crucial steps in achieving these results. Similar to the findings of Sanz et al. (2012), proper stereo camera calibration was found to be critical to achieving acceptable disparity maps. This reinforces the importance of calibration techniques, especially in real-world outdoor conditions where environmental factors introduce noise.

It was also observed that incorporating depth measurements into pothole detection greatly improves how potholes are prioritized for repairs compared to traditional visual-only inspections. This insight fills a notable gap in current practices, especially in the Philippine context where depth measurements are not typically part of road surveys (Ramos et al., 2023). Depth-based severity classification enables road maintenance teams to make more informed and objective decisions on which potholes to prioritize for immediate repair, helping to optimize resource allocation and improve public road safety.

The system achieved a strong positive regression model fit (R = 0.937, R^2 = 0.878) and satisfactory error measures (MAE = 1.82 cm, RMSE = 1.19 cm). These results confirm that stereo vision, when combined with simple regression modeling, can reliably estimate pothole depths. This finding is significant because earlier machine learning-based road detection studies (such as Bibi et al., 2021) focused mostly on classifying the existence of defects, not measuring their severity.

However, the study also highlighted limitations affecting system performance, including sensitivity to camera calibration quality, lighting conditions, road surface texture, and the camera's vertical positioning during image capture. Outdoor testing revealed that low lighting and shallow potholes made it difficult to generate clean disparity maps, sometimes causing minor estimation errors. These observations are consistent with Sattar et al. (2018), who reported that mobile road sensing systems often struggle in low-light or highly variable surface conditions. Understanding these challenges is important because it points to practical improvements, such as using better cameras, adding lighting support, or applying more robust image enhancement methods in future versions of the system.

5.4 Recommendations for Practice

Based on the findings of this special project, the following recommendations are proposed for future researchers, engineers, and road maintenance agencies:

*Use stereo vision systems for road surveys.*Road maintenance agencies should consider using calibrated stereo vision systems to estimate pothole depth, allowing for better prioritization of road repairs compared to visual inspections alone.

*Incorporate depth-based severity classification in maintenance procedures.*Authorities should update road inspection protocols to include depth measurements, making pothole severity assessment more objective and standardized.

5.5 Suggestions for further research

Based on the limitations encountered and the results obtained,the researchers have observed that there are lapses and possible improvements to further better this system.

Better camera. While the StereoPi V2 camera was effective for basic depth estimation, its performance is limited by its resolution, sensitivity to lighting, and depth range. Future researchers could consider using higher-quality stereo cameras or depth sensors with better image resolution and low-light capabilities to achieve more accurate and consistent disparity maps.

Improve camera calibration and tuning. While the StereoPi system produced good depth estimates, the results still varied depending on the precision of the camera calibration. Future researchers can explore better calibration techniques and finer parameter adjustments to minimize errors, especially in challenging environments.

5.6 Conclusion

This special project has successfully developed a system that addresses the problem of pothole severity assessment using depth measurement.The research shows that stereo vision, even using accessible and affordable technology, holds strong potential for future development in road maintenance automation. By building upon the foundation laid by this project, future systems can become even more accurate, efficient, and practical for real-world deployment.

References

Bibi, R., Saeed, Y., Zeb, A., Ghazal, T. M., Rahman, T., Said, R. A., ... Khan, M. A. (2021). Edge ai-based automated detection and classification of road anomalies in vanet using deep learning. *Computational Intelligence and Neuroscience*, *2021*(1). doi: 10.1155/2021/6262194

Ha, J., Kim, D., & Kim, M. (2022). Assessing severity of road cracks using deep learning-based segmentation and detection. *The Journal of Supercomputing*, *78*(16), 17721–17735. doi: 10.1007/s11227-022-04560-x

Kumar, A. (2024, October). What is image processing: Overview, applications, benefits, and more. *AI and Machine Learning*. Retrieved from https://www.simplilearn.com/image-processing-article (Accessed: January 1, 2025)

Kyriakou, C., Christodoulou, S. E., & Dimitriou, L. (2016, April). Roadway pavement anomaly classification utilizing smartphones and artificial intelligence. In *Proceedings of the ieee conference.* Retrieved from https://ieeexplore.ieee.org/abstract/document/7495459

Luo, D., Lu, J., & Guo, G. (2020, June). Road anomaly detection through deep learning approaches. *IEEE Journals and Magazine*. (https://ieeexplore.ieee.org/document/9123753/)

Ramaiah, N. K. B., & Kundu, S. (2021). Stereo vision based pothole detection system for improved ride quality. *SAE International Journal of Advances and Current Practices in Mobility*, *3*(5), 2603–2610. doi: 10.4271/2021-01-0085

Ramos, J. A., Dacanay, J. P., & Bronuela-Ambrocio, L. (2023). *A review of the current practices in the pavement surface monitoring in the philippines* (Doctoral dissertation, University of the Philippines Diliman). Retrieved from https://ncts.upd.edu.ph/tssp/wp-content/uploads/2023/01/TSSP2022_09.pdf

Resources, R. (2020). Video processing. *Riches Project EU*. Retrieved from https://resources.riches-project.eu/glossary/video-processing/ (Accessed: January 1, 2025)

Sanz, P., Mezcua, B., & Pena, J. (2012). Depth estimation: An introduction.

Current Advancements in Stereo Vision. Retrieved from http://dx.doi.org/10.5772/45904 doi: 10.5772/45904

Sattar, S., Li, S., & Chapman, M. (2018). Road surface monitoring using smartphone sensors: A review. *Sensors*, *18*(11), 3845–3845. doi: 10.3390/s18113845

Singh, P., Bansal, A., Kamal, A. E., & Kumar, S. (2021). Road surface quality monitoring using machine learning algorithm. In *Smart innovation, systems and technologies* (pp. 423–432). doi: 10.1007/978-981-16-6482-3_42

Appendix A

Appendix Title

Appendix B

Resource Persons

Prof. Jumar Cadondon
Assistant Professor
Division of Physical Sciences and Mathematics
UniversityofthePhilippinesVisayas

jgcadondon@up.edu.ph

Engr. Jane Chua
Engineer
DPWH Region 6
chua.jane@dpwh.gov.ph

Engr. Marilou Zamora
Chief
Planning and Design
DPWHRegion6

zamora.marilou@dpwh.gov.ph

Engr. Benjamin Javellana
Assistant Director
Maintenance
DPWHRegion6

javellana.benjamin@dpwh.gov.ph