

1 DEFECTRA: ROAD DEFECT SEVERITY ASSESSMENT
2 AND CLASSIFICATION

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

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

12 BELEBER, Benz Vrianne
13 CATALAN, Perserose
14 SENCIL, Kristian Lyle

15 Francis DIMZON
16 Adviser
17 Jumar CADONDON
18 Co-Adviser

19 April 18, 2025

Abstract

21 From 150 to 200 words of short, direct and complete sentences, the abstract should
22 be informative enough to serve as a substitute for reading the entire SP document
23 itself. It states the rationale and the objectives of the research. In the final Special
24 Problem document (i.e., the document you'll submit for your final defense), the
25 abstract should also contain a description of your research results, findings, and
26 contribution(s).

27 Suggested keywords based on ACM Computing Classification system can be
28 found at https://dl.acm.org/ccs/ccs_flat.cfm

29 **Keywords:** Keyword 1, keyword 2, keyword 3, keyword 4, etc.

Contents

31	1 Introduction	1
32	1.1 Overview	1
33	1.2 Problem Statement	2
34	1.3 Research Objectives	3
35	1.3.1 General Objective	3
36	1.3.2 Specific Objectives	3
37	1.4 Scope and Limitations of the Research	4
38	1.5 Significance of the Research	4
39	2 Review of Related Literature	6
40	2.1 Frameworks	6
41	2.1.1 Deep Learning	6
42	2.1.2 YOLOv5	6
43	2.1.3 Image and Video Processing	7
44	2.1.4 Stereo Vision	7
45	2.2 Related Studies	7
46	2.2.1 Deep Learning Studies	7
47	2.2.2 Machine Learning Studies	9

48	2.2.3 Computer Vision Studies	10
49	2.3 Chapter Summary	12
50	3 Methodology	13
51	3.1 Research Activities	13
52	3.1.1 Data Collection	13
53	3.1.2 Algorithm Selection	14
54	3.1.3 Design, Testing, and Experimentation	15
55	3.1.4 Challenges and Limitations	17
56	3.2 Calendar of Activities	18
57	4 Preliminary Results/System Prototype	19
58	References	20
59	A Appendix Title	22
60	B Resource Persons	23

⁶¹ List of Figures

62 List of Tables

<small>63</small>	2.1	Comparison of Related Studies on Road Anomaly Detection using	
<small>64</small>		Deep Learning Techniques and Stereo Vision	12
<small>65</small>	3.1	Timetable of Activities for 2024	18
<small>66</small>	3.2	Timetable of Activities for 2025	18

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

92 can also enable maintenance operations to commence as soon as possible if nec-
93 essary. If not automated, the delay of assessments will continue and roads that
94 are supposedly needing maintenance may not be properly maintained which can
95 affect the general public that is utilizing public roads daily.

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

108 1.2 Problem Statement

109 Roads support almost every aspect of daily life, from providing a way to transport
110 goods and services to allowing people to stay connected with their communities.
111 However, road defects such as cracks and potholes damage roads over time, and
112 they can increase accident risks and affect the overall transportation. The current
113 way of inspecting the roads for maintenance is often slow as it is done manually,
114 which makes it harder to detect and fix defects early. The delay in addressing
115 these problems can lead to even worse road conditions (J. Chua, Personal Inter-
116 view. 16 September 2024). There are several research studies into automated
117 road defect classification that have advanced in recent years but most of them
118 focus on identifying the types of defects rather than assessing their severity or
119 characteristics like depth. Without reliable data on the depth of the defect, road
120 maintenance authorities may underestimate the severity of certain defects. To ad-
121 dress these challenges, advancements are needed across various areas. An effective
122 solution should not only detect and classify road defects but also measure their
123 severity to better prioritize repairs. Failing to address this problem will require
124 more extensive repairs for damaged roads, which raises the cost and strains the
125 budget. Additionally, road maintenance would still be slow and cause disruptions
126 in daily activities. Using an automated system that accurately detects, classifies,
127 and assess the severity of road defects by incorporating depth are necessary to
128 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 and assess the severity 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, precision and recall will be used in order to determine the number of true positives and true positive rate and also the number of false positives and negatives detected by the system. In addition, in order to calculate the average precision and recall of the system the F-1 Score will also be used. Lastly, the Mean Absolute Error will be utilized in order to provide a straightforward measure of average error magnitude and Root Mean Square Error as a measure for performance since data is continuous.

1.3.2 Specific Objectives

Specifically, this special problem aims:

1. To collect high-quality images of road surfaces that capture potholes including their depth in various lighting, camera distance and orientation.
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 precision and recall, F1-score, root mean square error, and mean absolute error as metrics 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 that will be deployed through stereo camera mounted vehicles used for road surveying.

159 1.4 Scope and Limitations of the Research

160 This system will focus solely on detecting and assessing the severity of potholes
161 through image analysis and depth measurement technologies. The scope includes
162 the collection of pothole images using cameras and depth-sensing tools under
163 various lighting and weather conditions, ensuring the data captures real-world
164 variations.

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

170 The machine learning model developed will focus exclusively on detecting pot-
171 holes and assessing their severity through depth measurement. The accuracy of
172 the model's depth measurements will be evaluated by comparing them against
173 data collected from actual field inspections. However, this comparison will be
174 limited to selected sample sites, as collecting field data over a large area can be
175 time-consuming and resource-intensive.

176 Environmental factors such as lighting, road surface texture, and weather con-
177 ditions may impact the model's performance. The accuracy and reliability of the
178 model will depend on the quality and variety of the training dataset. Its ability
179 to generalize to unseen pothole images will need to be carefully validated.

180 1.5 Significance of the Research

181 This special problem aims to be significant to the following:

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

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

191 project expenditures, and better overall road safety and quality.

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

195 *Future Researchers.* The special problem can serve as a baseline and guide of
196 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 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, 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 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

240 in use of autonomous vehicles in research data gathering which is what the re-
241 searchers utilized in data gathering procedures. With the presence of autonomous
242 vehicles, this allowed the researchers to use a combination of sensors and deep
243 neural networks in deploying artificial intelligence. The study aimed to allow au-
244 tonomous vehicles to avoid critical road defects that can possibly lead to dangerous
245 situations. Researchers used Resnet-18 and VGG-11 in automatic detection and
246 classification of road defects. Researchers concluded that the trained model was
247 able to perform better than other techniques for road defect detection (Bibi et al.,
248 2021). The study is able to provide the effectiveness of using deep learning models
249 in training artificial intelligence for road defect detection and classification. How-
250 ever, the study lacks findings regarding the severity of detected defects which is
251 crucial in automating manual procedures of road surveying in the Philippines.

252 **2.2.1.2 Road Anomaly Detection through Deep Learning Approaches**

253 The study of Luo, Lu, and Guo (2020) aimed to utilize deep learning models in
254 classifying road anomalies. The researchers used three deep learning approaches
255 namely Convolutional Neural Network, Deep Feedforward Network, and Recurrent
256 Neural Network from data collected through the sensors in the vehicle’s suspension
257 system. In comparing the performance of the three deep learning approaches, the
258 researchers fixed some hyperparameters. Results revealed that the RNN model
259 was the most stable among the three and in the case of the CNN and DFN
260 models, the researchers suggested the use of wheel speed signals to ensure accuracy.
261 And lastly, the researchers concluded that the RNN model was best due to high
262 prediction performance with small set parameters (Luo et al., 2020).

263 **2.2.1.3 Assessing Severity of Road Cracks Using Deep Learning-Based** 264 **Segmentation and Detection**

265 In the study of Ha et al. (2022), it was argued that the detection, classification,
266 and severity assessment of road cracks should be automated due to the bottleneck
267 it causes during the entire process of surveying. For the study, the researchers
268 utilized SqueezeNet, U-Net, and MobileNet-SSD models for crack classification and
269 severity assessment. Furthermore, the researchers also employed separate U-nets
270 for linear and area cracking cases. For crack detection, the researchers followed
271 the process of pre-processing, detection, classification. During preprocessing im-
272 ages were smoothed out using image processing techniques. The researchers also
273 utilized YOLOv5 object detection models for classification of pavement cracking
274 wherein the YOLOv5l model recorded the highest accuracy. The researchers how-
275 ever stated images used for the study are only 2D images which may have allowed

276 higher accuracy rates. Furthermore, the researchers suggest incorporating depth
277 information in the models to further enhance results.

278 **2.2.1.4 Roadway pavement anomaly classification utilizing smartphones** 279 **and artificial intelligence**

280 The study of Kyriakou, Christodoulou, and Dimitriou (2016) presented what is
281 considered as a low-cost technology which was the use of Artificial Neural Net-
282 works in training a model for road anomaly detection from data gathered by
283 smartphone sensors. The researchers were able to collect case study data us-
284 ing two-dimensional indicators of the smartphone’s roll and pitch values. In the
285 study’s discussion, the data collected displayed some complexity due to accelera-
286 tion and vehicle speed which lead to detected anomalies being not as conclusive as
287 planned. The researchers also added that the plots are unable to show parameters
288 that could verify the data’s correctness and accuracy. Despite the setbacks, the
289 researchers still fed the data into the Artificial Neural Network that was expected
290 to produce two outputs which were “no defect” and “defect.” The method still
291 yielded above 90% accuracy but due to the limited number of possible outcomes
292 in the data processing the researchers still needed to test the methodology with
293 larger data sets and roads with higher volumes of anomalies.

294 **2.2.2 Machine Learning Studies**

295 **2.2.2.1 Smartphones as Sensors for Road Surface Monitoring**

296 In their study, Sattar, Li, and Chapman (2018) noted the rise of sensing capabil-
297 ities of smartphones which they utilized in monitoring road surface to detect and
298 identify anomalies. The researchers considered different approaches in detecting
299 road surface anomalies using smartphone sensors. One of which are threshold-
300 based approaches which was determined to be quite difficult due to several factors
301 that are affecting the process of determining the interval length of a window
302 function in spectral analysis (Sattar et al., 2018). The researchers also utilized
303 a machine learning approach adapted from another study. It was stated that k-
304 means was used in classifying sensor data and in training the SVM algorithm. Due
305 to the requirement of training a supervised algorithm using a labeled sample data
306 was required before classifying data from sensors, the approach was considered to
307 be impractical for real-time situations (Sattar et al., 2018). In addition, Sattar
308 et al. (2018) also noted various challenges when utilizing smartphones as sensors
309 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

344 able to detect potholes that are deeper than 5 cm. The researchers concluded
345 that such system can be utilized in commercial applications. However, it is also
346 worth noting that despite the system being able to detect potholes and measure
347 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	SqueezeNet, 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 will be taken to address the problem of classifying and assessing road defects using artificial intelligence. The methodology will be 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 will be essential for accurately classifying and assessing road defects.

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. Pothole severity will be classified based on a combination of area and depth. 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

375 used in the United States. The LTPP categorizes potholes based on depth thresh-
376 olds, which will be integrated with DPWH’s volume-based assessment to provide
377 a more comprehensive severity classification framework. The data collection will
378 involve capturing at least 500 images of potholes from various locations within
379 the UP Visayas Campus and the Province of Iloilo. These locations were selected
380 based on reports of road deterioration and input from the DPWH to ensure the
381 dataset represents real-world conditions. Ground truth data of pothole depth
382 will be collected by the researchers by measuring the depth of different points
383 in an individual pothole and then solving for its average depth. The aforemen-
384 tioned process was validated by Engr. Benjamin Javellana, Assistant Director of
385 DPWH Region 6. In order to individually locate or determine each pothole where
386 the ground truth data is collected, images taken will be labeled with their corre-
387 sponding coordinates, street names, and nearby landmarks. In addition to locally
388 collected data, open-source datasets such as the Dataset by Eric Tam from the
389 Crowdsensing-based Road Damage Detection Challenge focusing on road defects
390 and the Dataset by Atikur Rahman Chitholian, featuring 665 labeled pothole im-
391 ages from urban streets will be reviewed to supplement the model training and
392 improve generalization.

393 **3.1.2 Algorithm Selection**

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

399 **3.1.2.1 Pothole Detection**

400 YOLOv5 was selected due to its high accuracy and ability to process images in
401 real-time, making it suitable for detecting road defects in dynamic environments.
402 Its architecture is optimized for speed and performance, which is crucial for large-
403 scale deployment in road inspections.

404 **3.1.2.2 Severity Assessment**

405 The Multi-view Depth Estimation using Epipolar Spatio-Temporal Networks was
406 selected due to the high cost and limited accessibility of LiDAR technology. By

407 applying epipolar geometry and temporal consistency across sequential frames,
408 this approach provides an accurate depth estimation from standard video footage
409 (Long, Wang, Zhang, Mei, & Shen, 2021).

410 **3.1.3 Design, Testing, and Experimentation**

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

413 **3.1.3.1 Model Design**

414 The system was designed to operate with two core components: YOLOv5 for
415 pothole detection and ESTN for depth estimation. The model architecture was
416 chosen based on the real-time processing capabilities and the need for accurate
417 depth estimation from standard video footage. The design ensures that the system
418 can detect defects and provide severity assessments in a seamless workflow.

419 **3.1.3.2 Data Set**

420 The YOLOv5 model was trained using two datasets from Universe Roboflow. One
421 of the data sets was posted by a user named Eric Tam. It was also stated that
422 the images from the dataset are sourced from a Crowdsensing-based Road Damage
423 Detection Challenge from 2022 in Japan. The challenge involves contestants being
424 required to submit road damage datasets, shortlist their data set, and use the data
425 set for road damage detection and classification models. The use of this data set
426 in training models for road damage detection and classification ensures that the
427 data is viable for training the YOLOv5 model. The dataset contains various road
428 defects in Japan. Another data set used in training the YOLOv5 model was also
429 uploaded in Universe Roboflow by a user named Atikur Rahman Chitholian which
430 was stated to be part of his undergraduate thesis. The dataset is comprised of 665
431 images with potholes being labeled. It was also stated that the data set can be
432 utilized in automatically detecting and categorizing potholes found in the streets
433 of cities. Data preprocessing techniques were applied to both datasets to improve
434 model accuracy and generalization. These included resizing images to a uniform
435 size, applying augmentation techniques (flipping, rotation, and color adjustment)
436 to increase dataset variability, and normalizing pixel values to ensure consistency
437 across images.

438 3.1.3.3 Prototype Development

439 A prototype system will be developed in order to test the effectiveness of the
440 model. The prototype will involve the StereoPi V2 Kit which was acquired through
441 an official international distributor. After assembling the camera, it was further
442 modified to address the it's heating and make it suitable for outdoor use.

443 3.1.3.4 Performance Metrics

444 The performance of the YOLOv5 model will be evaluated using mean Average
445 Precision (mAP). mAP is a widely used metric in object detection tasks and is
446 particularly useful for assessing models that need to detect and classify multiple
447 object categories. In this case, mAP will provide a comprehensive evaluation of the
448 model's ability to detect and classify potholes, offering an aggregated score across
449 the relevant detection thresholds. This ensures a balanced assessment of both
450 detection accuracy and classification performance, which is essential for accurately
451 identifying potholes across varying conditions. The effectiveness of mAP for this
452 task is well-established in object detection literature (Everingham et al., 2015; Lin
453 et al., 2014).

454 For the accuracy of depth estimation using the Epipolar Spatio-Temporal Net-
455 works (ESTN), Root Mean Squared Error (RMSE) and Mean Absolute Error
456 (MAE) will be used. RMSE is chosen for its ability to penalize larger errors more
457 heavily, making it suitable for assessing depth estimation performance where larger
458 deviations from the ground truth are more significant (Zhang et al., 2018). MAE is
459 also employed to provide a straightforward measure of average error magnitude,
460 offering a complementary evaluation of depth estimation without emphasizing
461 larger errors as much (Zhang et al., 2020).

462 3.1.3.5 Testing and Validation

463 The testing process will begin with a detailed testing plan that includes both
464 simulated and real-world testing scenarios. Initially, the model will be tested in
465 controlled environments to ensure it can detect and assess road defects accurately.
466 Following this, real-world testing will be conducted using the StereoPi kit on local
467 roads, specifically at the University of the Philippines Visayas Miagao Campus.
468 The system's performance will be validated by comparing its predictions with
469 ground-truth data collected from manual inspections.

470 **3.1.3.6 Documentation**

471 Throughout the research activities, thorough documentation will be maintained.
472 This documentation will capture all methods, results, challenges, and adjustments
473 made during the experimentation phases. It ensures the reproducibility of the
474 work and provides transparency for future research endeavors.

475 **3.1.4 Challenges and Limitations**

476 **3.1.4.1 Availability of Local Datasets**

477 The lack of locally labeled datasets for road defects has posed a challenge in
478 training accurate models. The majority of available datasets are sourced from
479 international locations, which may not fully represent the road conditions found
480 in the study area. To address the lack of locally labeled datasets, the researchers
481 will create a pilot dataset from local roads within the University of the Philippines
482 Visayas Miagao Campus. This dataset will be manually annotated according to
483 DPWH’s classification standards, ensuring local relevance.

484 **3.1.4.2 Data Quality and Variability**

485 Variations in the quality and resolution of the data collected from different sources
486 may impact the performance of the trained models. In particular, images captured
487 under varying weather conditions or lighting may affect the accuracy of pothole
488 detection. To address this, the researchers plan to use the StereoPi kit to capture
489 images under optimal weather and lighting conditions, such as mid-morning or
490 early afternoon on clear days, ensuring consistent image quality for stereo vision
491 analysis. The kit’s stereo cameras will be calibrated for uniform resolution and
492 focus. Data augmentation techniques will also be applied to simulate varying con-
493 ditions, and pre-processing steps like noise reduction and contrast enhancement
494 will be used to improve the quality of the captured data. This approach aims
495 to minimize the impact of environmental factors on the accuracy of road pothole
496 detection and depth estimation.

497

3.2 Calendar of Activities

498

Table 1 shows a Gantt chart of the activities. Each bullet represents approximately

499

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

500 Chapter 4

501 Preliminary Results/System 502 Prototype

503 This chapter presents the preliminary results or the system prototype of your SP.
504 Include screenshots, tables, or graphs and provide the discussion of results.

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
- Kelleher, J. (2019). *Deep learning*. Retrieved from https://books.google.com.ph/books?hl=en&lr=&id=b06qDwAAQBAJ&oi=fnd&pg=PP9&dq=what+is+deep+learning&ots=_pCSXIk_WN&sig=EoGHTk7LjEBuR_OtFNX87LY0YU4&redir_esc=y#v=onepage&q=what%20is%20deep%20learning&f=false (In Google Books)
- 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>
- Long, Y., Wang, J., Zhang, H., Mei, T., & Shen, X. (2021). Multi-view depth estimation using epipolar spatio-temporal networks. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition (cvpr)*. doi: 10.1109/CVPR46437.2021.01567
- 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

537 Ramos, J. A., Dacanay, J. P., & Bronuela-Ambrocio, L. (2023). *A re-*
538 *view of the current practices in the pavement surface monitoring in the*
539 *philippines* (Doctoral dissertation, University of the Philippines Diliman).
540 Retrieved from [https://ncts.upd.edu.ph/tssp/wp-content/uploads/](https://ncts.upd.edu.ph/tssp/wp-content/uploads/2023/01/TSSP2022_09.pdf)
541 [2023/01/TSSP2022_09.pdf](https://ncts.upd.edu.ph/tssp/wp-content/uploads/2023/01/TSSP2022_09.pdf)

542 Resources, R. (2020). Video processing. *Riches Project EU*. Re-
543 trieved from [https://resources.riches-project.eu/glossary/video](https://resources.riches-project.eu/glossary/video-processing/)
544 [-processing/](https://resources.riches-project.eu/glossary/video-processing/) (Accessed: January 1, 2025)

545 Sattar, S., Li, S., & Chapman, M. (2018). Road surface monitoring us-
546 ing smartphone sensors: A review. *Sensors*, 18(11), 3845–3845. doi:
547 10.3390/s18113845

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

551 Solawetz, J. (2024, April). What is yolov5? a guide for beginners. *Roboflow Blog*.
552 Retrieved from [https://blog.roboflow.com/yolov5-improvements-and](https://blog.roboflow.com/yolov5-improvements-and-evaluation/)
553 [-evaluation/](https://blog.roboflow.com/yolov5-improvements-and-evaluation/)

554 **Appendix A**

555 **Appendix Title**

556 **Appendix B**

557 **Resource Persons**

558 **Prof. Jumar Cadondon**

559 Assistant Professor

560 Division of Physical Sciences and Mathematics

561 University of the Philippines Visayas

562 jgcadondon@up.edu.ph

563 **Engr. Jane Chua**

564 Engineer

565 DPWH Region 6

566 chua.jane@dpwh.gov.ph

567

568 **Engr. Marilou Zamora**

569 Chief

570 Planning and Design

571 DPWH Region 6

572 zamora.marilou@dpwh.gov.ph