

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 January 22, 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 Related Literature	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	17
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 such defects and almost not considering the severity of detected pot-
98 holes. In addition, for studies that include severity assessment on potholes, the
99 main goal of the study is not directed towards road maintenance automation but
100 other factors such as improvement of ride quality for the vehicle. Another issue
101 found in existing solutions is the lack of incorporation to the context of Philippine
102 roads. With these issues in mind, the study aims to utilize stereo vision from
103 StereoPi V2 in order to obtain multi-perspective views of detected potholes to be
104 used in severity assessment for automated road condition monitoring.

105 1.2 Problem Statement

106 Roads support almost every aspect of daily life, from providing a way to transport
107 goods and services to allowing people to stay connected with their communities.
108 However, road defects such as cracks and potholes damage roads over time, and
109 they can increase accident risks and affect the overall transportation. The current
110 way of inspecting the roads for maintenance is often slow as it is done manually,
111 which makes it harder to detect and fix defects early. The delay in addressing
112 these problems can lead to even worse road conditions (J. Chua, Personal Inter-
113 view. 16 September 2024). There are several research studies into automated
114 road defect classification that have advanced in recent years but most of them
115 focus on identifying the types of defects rather than assessing their severity or
116 characteristics like depth. Without reliable data on the depth of the defect, road
117 maintenance authorities may underestimate the severity of certain defects. To ad-
118 dress these challenges, advancements are needed across various areas. An effective
119 solution should not only detect and classify road defects but also measure their
120 severity to better prioritize repairs. Failing to address this problem will require
121 more extensive repairs for damaged roads, which raises the cost and strains the
122 budget. Additionally, road maintenance would still be slow and cause disruptions
123 in daily activities. Using an automated system that accurately detects, classifies,
124 and assess the severity of road defects by incorporating depth are necessary to
125 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 a combination of machine learning and computer vision techniques. The system will focus on measuring the depth of potholes to assess their severity, enabling faster and more accurate road maintenance decisions. In accordance with, the Department of Public Works and Highways Region 6's manual for road maintenance, the study will include volume of detected potholes in a certain segment as a metric for assessing severity. Furthermore, using depth estimation, the study will also include the depth of potholes as an additional metric for assessing pothole severity. 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.

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 and weather conditions.
2. To develop and train a machine learning model to detect and assess the severity of potholes 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 potholes from image input, analyze their depth, and assess their severity that will be deployed through stereo camera mounted vehicles used for road surveying.

1.4 Scope and Limitations of the Research

This system will focus 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 various lighting and weather conditions, ensuring the data captures real-world variations. High-quality and diverse image datasets will be crucial for training the model to accurately assess pothole severity based on depth.

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 machine learning model developed will focus exclusively on detecting potholes and assessing their severity through depth measurement. The accuracy of the model's depth measurements will be evaluated by comparing them against data collected from actual field inspections. However, this comparison will be 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 model's performance. The accuracy and reliability of the model will depend on the quality and variety of the training dataset. Its ability to generalize to unseen pothole images will need to be carefully validated.

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

187 also assess the defect's severity which allows them to prioritize repairs, optimal
188 project expenditures, and better overall road safety and quality.

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

192 *Future Researchers.* The special problem can serve as a baseline and guide of
193 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.

209 **2.1.3 Image and Video Processing**

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

219 **2.1.4 Stereo Vision**

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

226 **2.2 Related Studies**

227 This section of the chapter presents related studies conducted by other researchers
228 wherein the methodology and technologies used may serve as basis in the devel-
229 opment of this special problem.

230 **2.2.1 Deep Learning Studies**

231 **2.2.1.1 Automated Detection and Classification of Road Anomalies in** 232 **VANET Using Deep Learning**

233 In the study of Bibi et al. (2021) it was noted that identification of active road
234 defects are critical in maintaining smooth and safe flow of traffic. Detection and
235 subsequent repair of such defects in roads are crucial in keeping vehicles using
236 such roads away from mechanical failures. The study also emphasized the growth

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

249 **2.2.1.2 Road Anomaly Detection through Deep Learning Approaches**

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

260 **2.2.1.3 Assessing Severity of Road Cracks Using Deep Learning-Based** 261 **Segmentation and Detection**

262 In the study of (Ha, Kim, & Kim, 2022), it was argued that the detection, classi-
263 fication, and severity assessment of road cracks should be automated due to the
264 bottleneck it causes during the entire process of surveying. For the study, the
265 researchers utilized SqueezeNet, U-Net, and MobileNet-SSD models for crack clas-
266 sification and severity assessment. Furthermore, the researchers also employed
267 separate U-nets for linear and area cracking cases. For crack detection, the re-
268 searchers followed the process of pre-processing, detection, classification. Dur-
269 ing preprocessing images were smoothed out using image processing techniques.
270 The researchers also utilized YOLOv5 object detection models for classification of
271 pavement cracking wherein the YOLOv51 model recorded the highest accuracy.
272 The researchers however stated images used for the study are only 2D images

273 which may have allowed higher accuracy rates. Furthermore, the researchers sug-
274 gest incorporating depth information in the models to further enhance results.

275 **2.2.1.4 Roadway pavement anomaly classification utilizing smartphones** 276 **and artificial intelligence**

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

291 **2.2.2 Machine Learning Studies**

292 **2.2.2.1 Smartphones as Sensors for Road Surface Monitoring**

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

341 able to detect potholes that are deeper than 5 cm. The researchers concluded
342 that such system can be utilized in commercial applications. However, it is also
343 worth noting that despite the system being able to detect potholes and measure
344 its depth, the overall severity of the pothole and road condition was not addressed.

345

2.3 Chapter Summary

346

347

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

348

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. The researchers will also manually annotate the pilot dataset based on these standards, ensuring local relevance. In order to classify the severity of detected potholes using their depth, pothole severity standards used in foreign countries may be used such as the LTTP standards used in the United States of America where pothole severity is based on the depth of the pothole. In addition, the DPWH manual also has

372 their own measure of severity which is the volume of potholes detected in a certain
373 segment of road which will also be used as a metric in determining the severity
374 of potholes in this study. The model to be used for the study will be YOLOv5
375 due to it being able to work in high inference speeds which makes it suitable for
376 real-time application. Images from a normal camera are gathered from ground
377 truthing potholes around the area where the dataset will be collected which will be
378 then used to train the model in pothole detection for local relevance. Images from
379 stereo camera will be used in order to train the model in estimating the depth of
380 detected potholes wherein multiple images of the same pothole will be taken. The
381 Department Of Public Works and Highways Region 6 currently does not have any
382 procedure in measuring pothole depth as they are only concerned with pothole
383 area. In order to individually locate or determine each pothole where the ground
384 truth data is collected, images taken will be labeled with their corresponding
385 coordinates, street names, and nearby landmarks.

386 **3.1.2 Algorithm Selection**

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

392 **3.1.2.1 Pothole Detection**

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

397 **3.1.2.2 Severity Assessment**

398 The Multi-view Depth Estimation using Epipolar Spatio-Temporal Networks was
399 selected due to the high cost and limited accessibility of LiDAR technology. By
400 applying epipolar geometry and temporal consistency across sequential frames,
401 this approach provides an accurate depth estimation from standard video footage
402 (Long, Wang, Zhang, Mei, & Shen, 2021).

403 **3.1.3 Design, Testing, and Experimentation**

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

406 **3.1.3.1 Model Design**

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

412 **3.1.3.2 Data Set**

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

431 **3.1.3.3 Performance Metrics**

432 The performance of the YOLOv5 model will be evaluated using mean Average
433 Precision (mAP). mAP is a widely used metric in object detection tasks and is

434 particularly useful for assessing models that need to detect and classify multiple
435 object categories. In this case, mAP will provide a comprehensive evaluation of the
436 model’s ability to detect and classify potholes, offering an aggregated score across
437 the relevant detection thresholds. This ensures a balanced assessment of both
438 detection accuracy and classification performance, which is essential for accurately
439 identifying potholes across varying conditions. The effectiveness of mAP for this
440 task is well-established in object detection literature (Everingham et al., 2015; Lin
441 et al., 2014).

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

450 **3.1.3.4 Testing and Validation**

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

458 **3.1.3.5 Documentation**

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

463 **3.1.4 Challenges and Limitations**

464 **3.1.4.1 Availability of Local Datasets**

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

472 **3.1.4.2 Data Quality and Variability**

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

485 **3.2 Calendar of Activities**

486 Table 1 shows a Gantt chart of the activities. Each bullet represents approximately
487 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

488 Chapter 4

489 Preliminary Results/System 490 Prototype

491 This chapter presents the preliminary results or the system prototype of your SP.
492 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

- 525 Ramos, J. A., Dacanay, J. P., & Bronuela-Ambrocio, L. (2023). *A re-*
526 *view of the current practices in the pavement surface monitoring in the*
527 *philippines* (Doctoral dissertation, University of the Philippines Diliman).
528 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)
529 [2023/01/TSSP2022_09.pdf](https://ncts.upd.edu.ph/tssp/wp-content/uploads/2023/01/TSSP2022_09.pdf)
- 530 Resources, R. (2020). Video processing. *Riches Project EU*. Re-
531 trieved from [https://resources.riches-project.eu/glossary/video](https://resources.riches-project.eu/glossary/video-processing/)
532 [-processing/](https://resources.riches-project.eu/glossary/video-processing/) (Accessed: January 1, 2025)
- 533 Sattar, S., Li, S., & Chapman, M. (2018). Road surface monitoring us-
534 ing smartphone sensors: A review. *Sensors*, 18(11), 3845–3845. doi:
535 10.3390/s18113845
- 536 Singh, P., Bansal, A., Kamal, A. E., & Kumar, S. (2021). Road surface quality
537 monitoring using machine learning algorithm. In *Smart innovation, systems*
538 *and technologies* (pp. 423–432). doi: 10.1007/978-981-16-6482-3_42
- 539 Solawetz, J. (2024, April). What is yolov5? a guide for beginners. *Roboflow Blog*.
540 Retrieved from [https://blog.roboflow.com/yolov5-improvements-and](https://blog.roboflow.com/yolov5-improvements-and-evaluation/)
541 [-evaluation/](https://blog.roboflow.com/yolov5-improvements-and-evaluation/)

542 **Appendix A**

543 **Appendix Title**

544 **Appendix B**

545 **Resource Persons**

546 **Prof. Jumar Cadondon**

547 Assistant Professor

548 Division of Physical Sciences and Mathematics

549 University of the Philippines Visayas

550 jgcadondon@up.edu.ph

551 **Engr. Jane Chua**

552 Engineer

553 DPWH Region 6

554 chua.jane@dpwh.gov.ph

555

556 **Engr. Marilou Zamora**

557 Chief

558 Planning and Design

559 DPWH Region 6

560 zamora.marilou@dpwh.gov.ph