PHILIPPINE ENGINEERING JOURNAL
PEJ 2023; Vol. 44, No. 1: 65-88

Received: 08 February 2023
Accepted: 22 June 2023

Assessment of Longitudinal Pavement Marking Conditions Based on Deep Learning

Lea B. Bronuela-Ambrocio, Emilson Ryan D. Antes, and Mellaine Denise S. De Leon
Institute of Civil Engineering, College of Engineering,
University of the Philippines Diliman, Quezon City, Philippines

Abstract *- Pavement markings play a vital role in managing regulations and the safety of road users. Markings at best conditions effectively convey warning messages and information to the drivers without diverting their attention off the road. Due to traffic wear, regular evaluation and proper maintenance are being conducted by authorized agencies. The Department of Public Works and Highways (DPWH) is the legal authority that applies, removes, and maintains pavement markings on national roads in the Philippines. They used retroreflectivity, a property of pavement markings to reflect the light from car headlamps, to assess the condition of the pavement markings. This needs equipment laid on the pavement markings, interrupting the traffic flow. Another evaluation method employed in the country is a manual inspection in which the surveyors assess the remains of the markings on the road. Markings that are less than 50% based on the perception of the surveyor are subject to repainting. This method is subjected to human error or high subjectivity of measurement.*

This study proposes a more time-efficient, more economical, and standardized alternative method to evaluate the condition of pavement markings. This approach uses the images captured from the existing pavement markings and a wear index adopted from foreign standards to rate the condition of the remaining marks. It aims to produce a program that employs an object detection algorithm called You Only Look Once (YOLO). The results of this program were validated using image processing and retroreflectivity test. Based on the findings, the proposed method is aligned with the current DPWH practice. It gives a relatively high goodness of fit equal to 0.83 with the present method of 50% pavement marking judgment.

Keywords: pavement markings, YOLOv5, image processing, object detection algorithm

I. INTRODUCTION

1.1 Background of the Study

Clear and visible pavement markings effectively convey warning messages and information to drivers without taking their eyes off the road. It improves road visibility during the night or in harsh weather conditions like fog or rain. It is designed to be easily understood, standardized, and conformant to national or local specifications. It is painted on the pavement surface in the form of lines, symbols, messages, or numerals [1]. Moreover, it comes with distinct colors and patterns such as longitudinal lines, transverse lines, and other special markings for various purposes such as intersections, road diversions, stops, etc. These are commonly made up of retroreflective paint and thermoplastic materials [2].

The pavement marking deterioration is impacted by age and traffic [3]. Poorly maintained markings may fail to impart information to road users. It can be evaluated based on retroreflectivity, color, and wear. The DPWH conducts reftroreflectivity tests or visual assessment methods to assess the performance of painted markings along the national roads. This test is conducted in accordance with DPWH Department Order No. 130 series of 2013 [4] and ASTM Standard [5], in which a piece of equipment needs to be laid on the markings to assess the amount of light reflected toward the driver's eyes). The retroreflectivity test takes a longer time compared to the visual assessment. Furthermore, traffic must be interrupted to laid the equipment on the ground. The visual assessment does not require equipment but is prone to human error. The criteria for repair for this method depends on the surveyor's judgment if the remaining marking subject for inspection is less than 50%.

This study aims to propose a new method that can be used locally to assess pavement marking conditions through a computer program that employs an object detection algorithm with a promising result called YOLOv5. The longitudinal pavement markings were considered for this study, and the results were verified and validated using image processing and a retroreflectivity tester, respectively.

1.2 Automated Evaluation of Pavement Markings

Foreign countries like China, Japan, India, Malaysia, the Czech Republic, and the United States of America (USA) used computer programs to assess the conditions of pavement markings. As early as 2000, Burrow et al. [6] were able to measure the erosion of pavement markings through digital video image analysis. Image digitization and image segmentation were used as post-processing techniques to evaluate the conditions of the markings further. Collado et al. [7] performed image filtering to delineate the road markings from the road surface. Noda et al. [8] used image processing to recognize the pavement marking images collected from an in-vehicle camera. They created a program that recognizes pavement markings in different environmental conditions.

Moreover, the Iowa Department of Transportation [9] reported that using image processing to assess pavement markings enhances the capacity to detect the presence of pavement markings. They use a program that calculates the percentage of the remaining white and yellow markings on the concrete roadways, which eventually expanded to other pavement surface types. Another study by Zhang and Ge (2012) [10] presented a systematical approach that automatically detects the present state of the pavement markings on Washington, Mississippi roads. They utilized Matlab for detecting color, image segmentation, and enhancement. A vision-based framework for pavement marking detection and condition assessment for Australian roads is the most recent study by Xu et al. [11]. It generates instructive information on the distribution of different levels of wear and tear on road markings. Azmi et al. [12] studied detecting missing road lane markings using YOLOv5, which implies a promising preliminary result of a mean average precision (mAP) equal to 0.995.

You Look Only Once (YOLO) is one of the latest object detection models introduced in 2015. According to Gui et al. [13] and Azmi et al. [12], this is one of the most popular and growing algorithms in machine learning. It uses convolutional neural networks (CNN) to distinguish specific objects and predict class probabilities. The core of the YOLO algorithm is straightforward and can be used for real-time detection. It considers detection time and reduces the error of detecting the background as the object. However, the earlier models had to improve accurate positioning, small object identification, and objects very close to each other and in groups indicating that it was not able to consider faded and small-scale pavement markings [14]. Hence, the shortcomings of the algorithm were improved by developing various versions. Currently, it has five versions: YOLO, YOLOv2, YOLOv3, YOLOv4, and YOLOv5. YOLOv5 is the most up-to-date and accurate version. It is 90% more modest than the preceding version, YOLOv4. Furthermore, it is more flexible in the control of model size, has a faster detection speed, and is very impressive in detecting small targets [15]. Overall, YOLO is known to have an architecture that allows for end-to-end training and real-time performance while maintaining high average precision and accuracy. With this, several studies, especially transportation-related application studies, have used this algorithm for automatic object detection. Table 1 summarizes the existing automated pavement marking evaluations, methods, and algorithms.

Table 1. Summary of existing pavement marking evaluation, method, and algorithm.

Author	Pavement Marking Evaluation	Performance of Method	Algorithm	Remarks
Burrow et al. (2000)	Wear (Erosion)	Analysis of pixels and Image enhancement	Image segmentation	
Collado et. al (2004)	Tracking and Classification of Road Lanes using longitudinal road markings	birds-eye view images pixels are analyzed to delineate the lane boundaries	Image segmentation Hugh Transformation	linear transformation for detecting straight lines, which has a good tolerance for image noise
Noda et al. (2009)	Accurate recognition of road markings	enhances the appearance changes of road marking in in-vehicle camera images	Generative Learning Method	
Smadi et al. (2007)	Remaining pavement marking paint on highways	Converts images to binary image conversion and then compared to an empirically defined threshold	Image segmentation	
Ge, (2011)	Determination of the present condition of pavement markings	Detect edge lines of pavement markings on an image and estimates intersection points in the polar sections	Image processing Hugh Transformation	linear transformation for detecting straight lines, which has a good tolerance for image noise
Xu et al. (2011)	detection and assessment of the condition of pavement marking by measuring the worn percentage	Analyze the pixels of video images and detects the lines of longitudinal marking	Image segmentation Cany and Hugh Transformation	linear transformation for detecting straight lines, which has a good tolerance for image noise
Kang (2020)	Visibility Analysis for Pavement Markings	Use annotated images to detect pavement markings and analyzes the mask and its intensity	.YOLOv3	Use muti-scale detection, comprised entirely of Convolutional Neural Networks.
Dewi (2023)	Road surface Traffic Signs Recognition	Use annotated images to detect road markings	YOLO v2 YOLOv3 YOLOv4 YOLOv4-tiny	YOLO v4-tiny redesigned for best speed or accuracy trade-off
Azmi et al. (2022)	detection of missing road lane markings	Use training images for the deep learning model.	YOLOv5	Most super in speed for detection performance among YOLO algorithm that can capture the amount of data for long stretch of road
Guo (2022)	pavement distresses identification	Uses labeled iPhone images as training images for the deep learning model.	YOLOv5	

A study by Zhang et al. [16] used YOLOv5 to detect vehicles from pictures, videos, and real-time surveillance even in non-ideal conditions such as nighttime, heavy traffic, and vehicle overcrowding. The training results presented a mean average precision of 0.70. Another study by Qui et al. [17] used the YOLOv5 algorithm to detect road traffic elements such as zebra crossings, bus stations, and roadside parking spaces from an aerial view. This was done to address issues in the detection of multi-scale objects and small objects. The training results presented a mean average precision of 0.931, which was 19.2% higher than the original and unimproved YOLOv5 model. This means that the detection errors for the small and medium objects were decreased. Guo et al. [13] use deep learning and YOLOv5 to identify pavement distress. His generated model showed a precision, recall, and mean average precision of 95.5%, 94.3%, and 95%, respectively, using a confidence level of 0.4 for model interferences. Gonal and Baliga [18] conducted a similar study that automates the detection and classification of road damage utilizing an Android device. They used YOLOv5 to detect pavement cracks, potholes, and blurred pavement markings to train data from the Czech Republic, India, and Japan. Results show a relatively high value of the predicted percentages for each label class.

II. METHODOLOGY

2.1 Development of Pavement Marking Detection Program

The conceptual framework of this study is shown in Figure 1. The wear or durability performance of the pavement markings was used as a criterion for this study. It considers the overall percentage of pavement marking left on the road surface and the retained retroreflectivity over time [19]. The wear index was adopted to serve as a rating of the condition of the existing pavement markings. Multiple images of these markings, taken using similar setups, were captured along the national road to train the object detection program. This was done to consider different pavement marking conditions across multiple areas. The data were collected for the development of the program in which it was divided into training, testing, and validation sets (see Figure 1). The training datasets were manually annotated with appropriate wear indices based on the visual condition of the pavement markings. The other datasets were used for testing and validation.

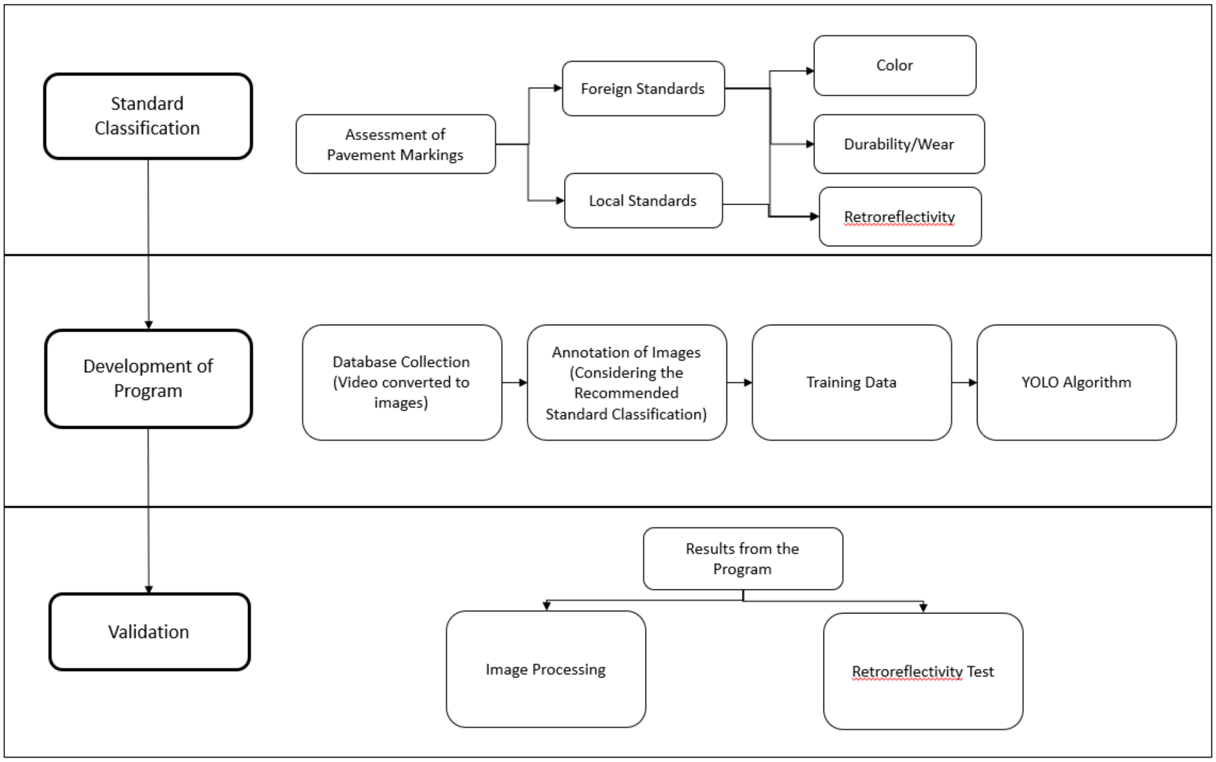


Figure 1. Conceptual Framework

2.1.1 Adoption of Wear Standard

The wear experienced by the pavement markings or their durability is evaluated through visual inspection followed by assessing the percentage of the remaining paint on the road surface. A well-trained evaluator will rate the percentage of retained markings based on their perception. Pull-out test is another durability assessment method conducted by applying a tensile force to the markings and pavement surface until they separate from each other [20]. However, the second method is time, energy-consuming, and impractical as the evaluator needs to impose destruction on the existing markings.

The wear of pavement markings in the United Kingdom (UK) is assessed by visual inspection [21]. Table 2 shows the wear index score of the UK for different percentages of the remaining paint.

Table 2. Wear Standards of the United Kingdom

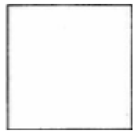
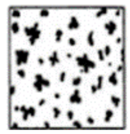
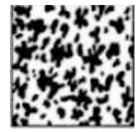
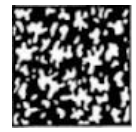
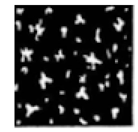
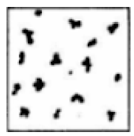
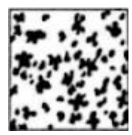
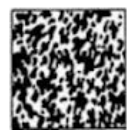
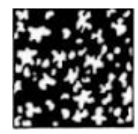
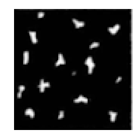
Assessment	Wear index/score	Defect type
Non-existent, residue only	0	Critical defect
Barely Visible	10	Critical defect
Visible, but has randomly spaced small bare spots	20	Potentially critical defect - judgment required taking into account location and function
Marginal - some visible wear, larger bare spots	30	Non-critical defect
Very little wear	40	Non-critical defect
No obvious wear	50	Not a defect

Ohio Department of Transportation (ODOT) evaluates the durability of pavement markings with a similar method. A numerical rating that ranges from 1 to 10 with an increment of 1 is used as a rating scale [22]. A rating of 1 indicates 10% of the remaining marking, while 10 represents 100% remaining marking. A graphical scale is used to compare existing pavement markings visually (see Table 3), and a score of less than eight is considered critical and must be repainted [23][24].

Similarly, Singapore also uses the wear parameter to evaluate the condition of pavement markings. Unlike other methods that are purely based on the visual judgment of the evaluator, this method involves an onsite test with a wire mesh grid to assess the degree of wear. The grid is 500 mm x 100 mm divided into 20 squares. After that, a photograph is taken at a distance of at least 1 meter above the marking. Each square will be assessed according to the percentage of remaining marks through visual judgment and evaluated using the letter grades A to D. Squares with the same grades will be added and multiplied by a certain weight factor depending on their letter grade. The squares' scores will then be added to determine the total wear index. If the score or wear index results to greater than 20, the strip of pavement marking is classified as critical and must be repainted [25]. However, this method was deemed to be impractical and time-consuming, thus, was not considered in the adopted wear index.

The first two (2) published manuals were used to decide on the most suitable wear standards that can be implemented for developing the detection program. The graphical representations established by ODOT for assessing pavement markings (see Table 3) were found to be practical, economical, and time efficient. However, having ten classifications can be too confusing for inspectors. Hence, researchers decided to merge two adjacent classifications so that the new consolidated pavement marking wear index will only correspond to five (5) grades, namely Grades A, B, C, D, and E (see Table 3).

Table 3. Visual Basis of the Consolidated Wear Index based on ODOT Manual.

Proposed Wear Index Classification	Grade A	Grade B	Grade C	Grade D	Grade E
					
Classification Based on ODOT Manual	10	8	6	4	2
					
Classification Based on ODOT Manual	9	7	5	3	1

Moreover, the 80% paint remaining threshold of ODOT needs to be lowered because it may result in high maintenance costs. On this note, the researchers adopted the 60% threshold paint remaining in the UK. This is relatively high with the 50% threshold established by DPWH. However, it can provide a buffer for the marking to reach a critical condition. It ensures adequate visibility, addressing road safety issues and concerns.

2.1.2 Object Detection Program

Figure 2 shows the camera mounted on a vehicle set-up used to capture the lane markings, while Figure 3 shows the process of developing the object detection program. An attachable piece of equipment to a pick-up truck was used to collect data on existing pavement markings. The camera, at a depressed angle, recorded videos of the pavement surface while traversing a particular road. The angle of camera must be angled down to reduce the area of the camera sensor to see further distances which are not resolvable. The camera's field of view must cover at least a 4-meter-wide span of road. This width is approximately equal to one typical pavement lane. The survey vehicle traversed the national roads along Manila, Rizal Province Nueva Ecija, and Tarlac Province to capture various road surface conditions. With this, bulk data of images were captured to ensure that the dataset can represent most of the instances of wear that pavement markings can experience. These images have a uniform size of 1920 pixels x 1080 pixels. All images with longitudinal pavement markings were extracted, filtered to avoid duplications, and annotated with their appropriate grades based on the established wear index for the program.

Figure 2. Camera mounted on a vehicle.

2.1.3 Annotation of Data

Annotation or labeling of images was done through an online tool called hasty.ai [31]. This application creates ground-truth bounding box annotations and machine learning to aid in finding, creating, and labeling images through its built-in artificial intelligence (AI) assistance tools, such as object detection. These AI assistants were trained based on the provided data to deliver a better-annotated dataset. Proper classification of pavement marking conditions is critical at this stage to increase the accuracy of the object detection model. A total number of 2925 images were annotated with their corresponding grade classification. Eighty percent of these images were used as training images, while ten percent were used for testing and another ten percent for validation (see Figure 4). Multiple instances can be detected in a single image, and Table 4 shows the distribution of the instances obtained from the images per grade. Missing pavement markings were not considered for the model development because they might cause confusion and complex results for the object detection program.

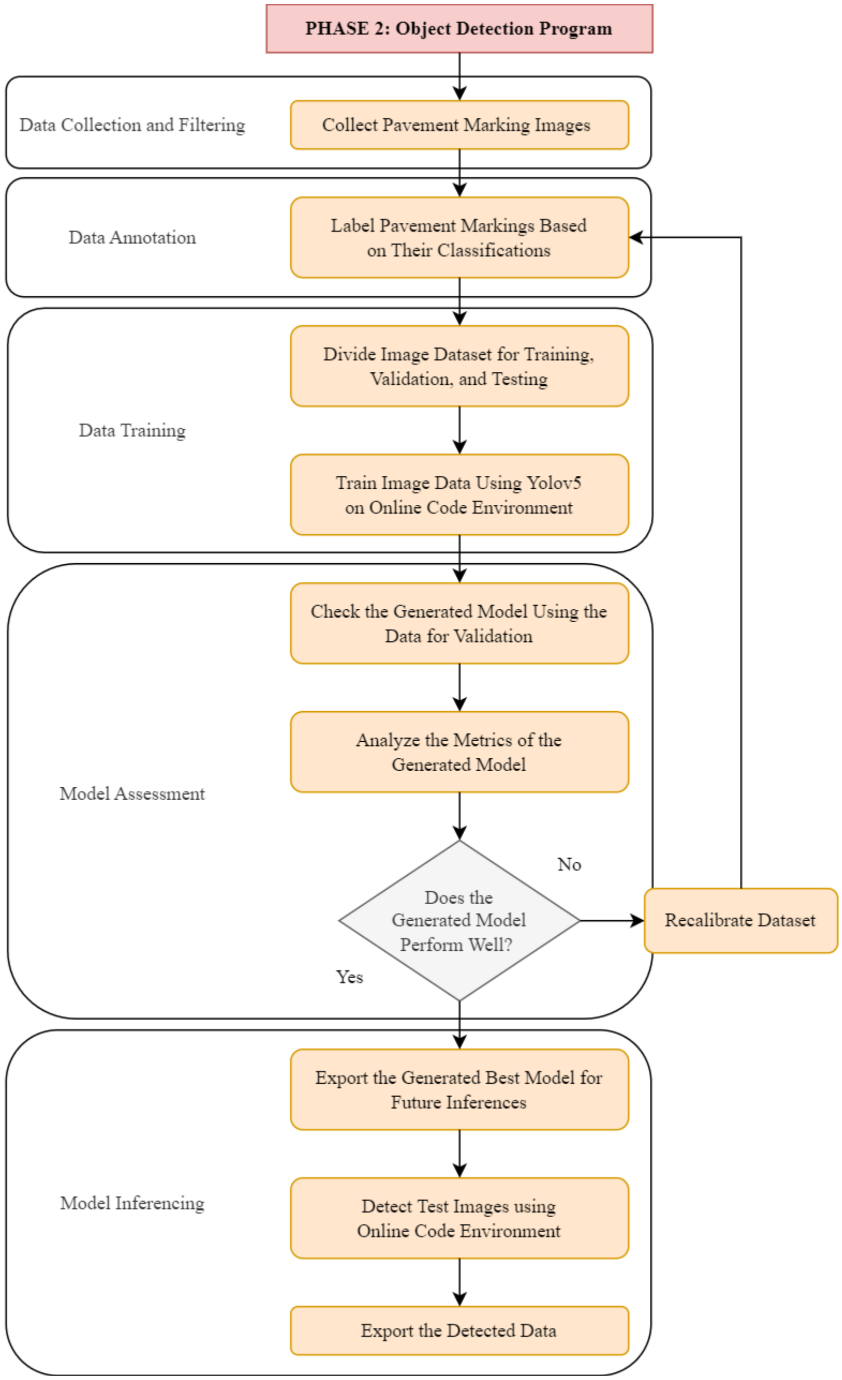


Figure 3. Detailed Methodological Framework of the Object Detection Program

Table 4. Number of Instances per Pavement Condition Grade

Grades	Instances
A	1284
B	962
C	997
D	900
E	818

2.1.4 Training Data

The annotated images were converted to an acceptable format for the YOLOv5 code. YOLO algorithm requires a single forward propagation across a neural network to detect an object, and it uses the imported dataset of images with labeled objects. It also uses techniques such as residual blocks, bounding box regression, and intersection over union to detect objects accurately. It is based on regression, wherein it predicts classes and bounding boxes for the whole image in a single execution instead of selecting the interesting area of a picture.

The code from the Ultralytics YOLOv5 GitHub repository [26] was copied and cloned to a new Google Colab programming environment to train the pavement markings images. The resolution of the images was set to the default value of 640, and the batch size was set to 16. Multiple test runs were done considering the number of epochs and randomness to avoid overfitting and underfitting the model. An epoch refers to a cycle of training the entire dataset. Multiple trials were made to attain the highest mean average precision, which appears to be between the 1st and 100th epoch. Google Colab offers a limited time for a graphical training model, and epochs higher than 200 may run infinitely. On that note, the number of epochs was set to 100 because this was deemed optimal for the data training based on the multiple test runs. The model training was then run, trying all possible values of epochs. The best model generated was assessed based on its performance and quality.

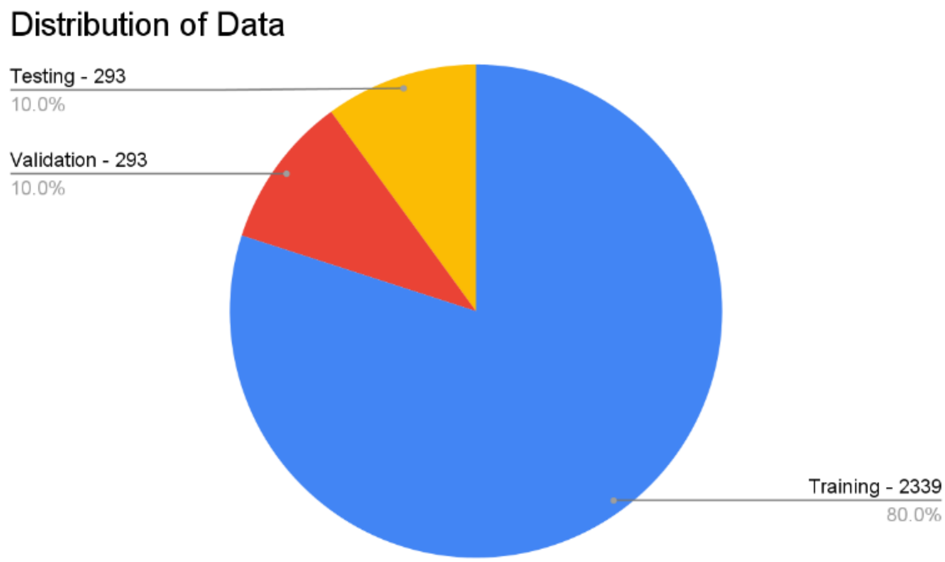


Figure 4. Distribution of Pavement Marking Data

2.1.5 Assessment of Model

The performance and quality of the generated model were evaluated based on the following metrics;

a. Precision (P) refers to the number of true predictions out of all the detected instances or the percentage of all detection results that are correctly detected.

$$P = \frac{TP}{(TP+FP)} = \frac{TP}{All\ Detection} \quad (1)$$

where TP: True Positive - correct detections
FP: False Positive - incorrect detections

b. Recall (R) refers to the number of true predictions out of all the actual existing possible instances. It indicates how well a positive prediction is made from a given set of positive inputs.

$$R = \frac{TP}{(TP+FN)} = \frac{TP}{All\ Ground\ Truths} \quad (2)$$

where TP: True Positive - correct detections
FN: False Negative - additional detections that should have occurred

c. Confusion Matrix is used to exhibit the performance of an algorithm that classifies objects. This is usually used when there are more than two classifications [27]. Table 5 shows the sample confusion matrix.

Table 5. Sample Confusion Matrix

	Actual		
Predicted		Positive	Negative
	Positive	TP	FP
	Negative	FN	TN

True Positive (TP) means an object was detected and correctly classified. False Positive (FP) means that a particular object was detected when it should not. False Negative (FN) represents an object that should have been detected but was not detected, and True Negative (TN) means that an object that should not be detected was not detected.

d. Mean Average Precision at 0.5 thresholds (mAP@0.5) classifies the detected instance as positive if the Intersection over Union (IoU) is equal to or greater than 0.5. The value for IoU measures the overlap between the predicted bounding box coordinates and the ground truth box coordinates. This is done by dividing the area of overlap by the area of union, as the name suggests [28]. Figure 5 shows the occurrences and differences between the two (2) kinds of bounding boxes. The values for IoU will be used to compute the entries that will complete a confusion matrix. Then, the precision and recall metrics per class will be plotted [28]. The area under a certain precision-recall curve is equal to the value of the Average Precision (AP) of a particular class.

$$mAP = 1/N \ \Sigma APi \tag{3}$$

where mAP: mean Average Precision
APi: Average Precision per Class
N: number of classes

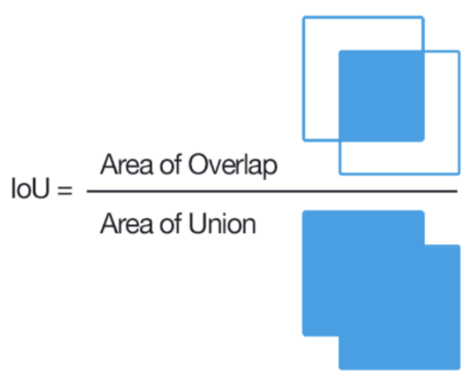


Figure 5. Illustration of How IoU is Computed [27]

The developed program for detecting and classifying the conditions of pavement markings can be accessed through this link:
https://colab.research.google.com/drive/14xrqBI98t0ZX9Mi0Mr3MaCVryTw4yjuB.

Please note that the program must be copied to create a separate version where it can be freely used. This must be done to avoid any significant changes to the official version or the master copy.

2.2 Verification of the Model

Photographs, taken from smartphones, of pavement markings with various conditions were taken within the vicinity of the University of the Philippines Diliman. A total of 96 instances of pavement markings were collected. Each marking was photographed twice; the first shot was taken at a depressed angled view (ranging from 30 to 60 degrees from the horizontal), like the angle taken during the data collection (refer to Figure 6). The second shot was taken with a bird's eye view. These photos were used to detect the classification of a specimen. The images with a bird's eye view were subjected to a simple program from MatLab that uses the r2bgray, and imbinarize functions were used to verify results. These functions convert images to grayscale and create binary images (see Figure 7). The threshold values of 0.85 and 0.87 account for lighting variations in the images. The number of white pixels and total pixels (white plus black pixels) were computed to determine the percentage of the remaining paint. The verification was done by comparing the classification grade obtained from the object detection program and the percentage of remaining paint calculated from MatLab.

(a) (b)

Figure 6. Sample photos taken at (a) angled view and (b) top view.

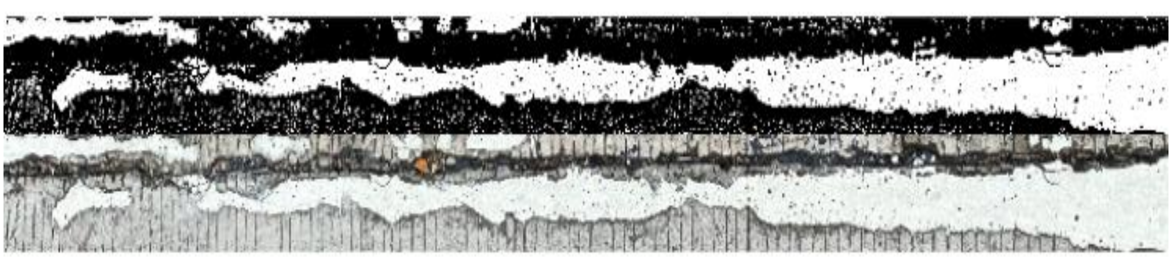


Figure 7. Sample Binarization of a Pavement Marking

2.3 Validation of the Object Detection Program

A total of seven (7) pavement markings inside the compound of the DPWH Bureau of Research and Standards (BRS) were tested based on their retroreflectivity performance. The mcd/m$^{\wedge 2}$/lux of these markings were measured by the equipment and recorded for further analysis. Furthermore, the retroreflectometer was also used on the part of the pavement without any marking to simulate the retroreflectance value of almost faded pavement markings. Concurrently, each pavement marking was photographed to be run through the developed object detection program and its counterpart image processing program so that the exact numerical percent of paint remaining in pavement markings could be determined.

Results from the object detection program and the image processing program were compared for each pavement marking. The percentages of paint remaining computed by the image processing program were plotted on a scatter plot against the retroreflectivity performance shown by the retroreflectometer. Afterward, an appropriate equation was fitted to determine the relationship between the two data and analyze their accuracy. Based on this, the percent of paint remaining corresponding to the DPWH critical retroreflectivity value for white pavement markings of 150 mcd/m^2/lux would be calculated.

III. RESULTS AND DISCUSSION

3.1 Wear Index

The researcher adopted a similar method to ODOT, where a visual assessment of pavement markings is done based on the remaining marking percentage. Table 6 summarizes the grade classification of pavement markings, their description, and the appropriate action. As mentioned in the previous section, these were used as a guide for annotating the training data.

3.2 Detection and Classification of Pavement Markings

The developed object detection program was able to detect and classify pavement markings. Figures 8 (a) and (b) show a set of annotated images used for training and a similar set of pictures subjected to the object detection program, respectively. The latter shows the predicted category and the confidence level labeled on the top of the bounding boxes. The two pictures indicate that the program could properly detect and classify pavement markings with respect to the annotations made by the researchers. It is noteworthy that object detection is mainly focused on broken centerlines type of pavement markings, which explains why some pavement markings are not detected in all instances.

Table 6. Pavement Marking Wear Index Grade and Description

Grades	Description	Decision
A	No obvious wear 80-100% pavement markings remaining	Does not need repainting
B	Very little to some visible wear 60-80% of pavement markings remaining	Does not need repainting
C	Visible Wear/Randomly spaced small bare spots 40-60 % pavement markings remaining	Needs repainting
D	Barely Visible 20-40 % pavement markings remaining	Needs repainting
E	Non-existent/Residue Only 0-20 % pavement markings remaining	Needs repainting

Moreover, it can also be observed that there are a few differences across some images. In one instance, a tile in row 2, column 1, pavement markings was annotated as Grade E. Still, the program classified it as Grade D. Furthermore, in another instance in row 1, column 4 tile, a pavement marking was classified as Grade C when it was supposed to be Grade B. As shown, the confidence level of predictions at these boxes was as low as 0.5 and 0.4. This means that the program was also not that highly certain of its performance in those given contexts. This happens on pavement markings that are on the borderline of two adjacent classification grades.

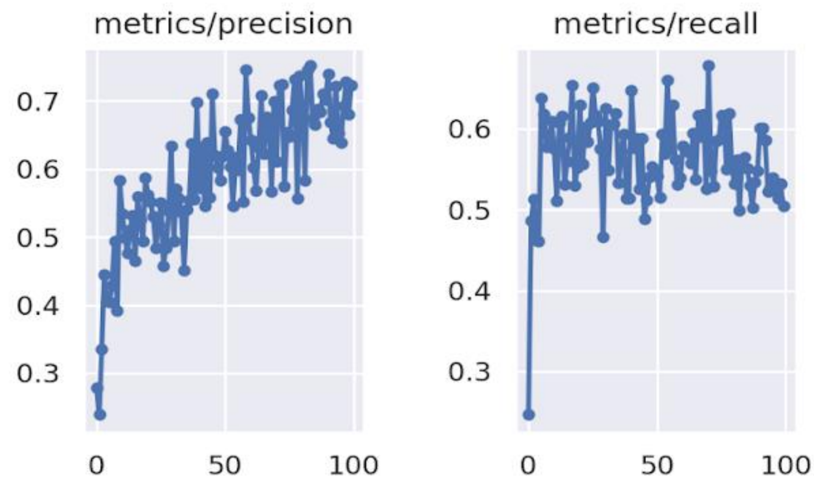
(a)

(b)

Figure 8. Sample Annotated Images (a) and Sample Predicted Images (b)

3.3 Assessment of Data Training Metrics

The effectiveness of the object detection program was measured by the YOLOv5 metrics mentioned in the previous sections. Figure 9 shows the overall behavior of the best-generated training model based on its precision and recall percentage per epoch obtained after the training process. The graphs show an increase in precision as the epoch rises, which eventually plateaued at epoch 60. Moreover, the average precision value lies between 0.6 and 0.7, while the recall value lies between 0.5 to 0.6. This means that more than 60% of the detected pavement markings can be correctly classified, and it can correctly predict more than 50% of actual existing possible instances.



(a) (b)

Figure 9. (a) Precision and (b) Recall Percentage per No. of Epoch

The Intersection over Union (IoU) threshold of average precision was set at a value of 0.5. This means that the program will classify the detected instance as positive or correct if the IoU is equal to or greater than 0.5. Using these computed IoU values, the confusion matrix of various classes was completed for all generated models per epoch. Figure 10 shows the mean average precision at 0.5 thresholds (mAP@0.5) across the 100 epochs.

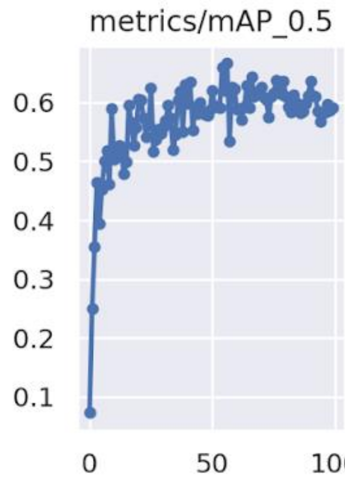


Figure 10. mAP@0.5 per Epoch

The best-generated model with the best detection performance appears to have the highest mAP@0.5. A value of 0.667 was obtained from the detection program. This means that the program could correctly predict the classification of markings at an average rate of 66.7%. On the other hand, a precision value (P) of 0.672 was obtained, indicating that 67.2% of the predictions generated by the generated best model were correct. Lastly, a recall (R) value of 0.63 was also obtained for the whole program. This implies that the developed best model could correctly predict 63% of the instances it should detect. These values were produced at epoch 56, the best epoch, shown in Figure 8. Table 6 shows the summary of P, R, and mAP@0.5 values generated per class.

Table 6. Model Summary

Class	P	R	mAP@0.5
All	0.672	0.63	0.667
Grade A	1.000	0.555	0.747
Grade B	0.747	0.715	0.773
Grade C	0.663	0.667	0.681
Grade D	0.356	0.607	0.484
Grade E	0.596	0.607	0.651

The generated model has a mAP@0.5 of 66.7%, which is lower than other existing studies. This may be caused by the longer objects needed to be detected by this particular program compared to other developed programs that detect people, cars, road signs, etc. This means that the annotations done by the researchers and the predictions done by the program might be cut along the length of the pavement marking. With that, the intersection area will be much lower if the boxes do not match very well, thus lowering the computed values for precision, recall, confusion matrix, and, eventually, the mAP@0.5. This is evident in Figure 11, where the bounding box size of the annotated pictures differs from the size of the bounding box size of the predicted instances for the same pavement marking. Given this circumstance, a mAP@0.5 of 0.67 still can give an effective and reliable result for assessing pavement marking conditions.

Figure 11. Differences between the Annotated Box (Left) and the Predicted Box (Right) of the Same Pavement Marking

Lastly, Figure 12 shows the confusion matrix generated by the model. The main diagonal contains the TP entries for each class. This means that grades A, B, C, D, and E correctly predicted 73%, 69%, 0.58%, 0.61%, and 0.62%, respectively, for the data that was processed. On the other hand, the values on the left or right of a certain diagonal are False Positives (FP). For example, grade C has a TP of 0.58, the values that can be seen on its left and right are 0.14 and 0.07, respectively. This implies that the program incorrectly predicts 14% of grade B instances as Grade C and incorrectly predicts 7% of grade D instances as grade C.

Furthermore, the values above or below the TP's are FN. Looking at the same diagonal (0.58), the values 0.06 and 0.25 can be seen above and below, respectively. This implies that the program incorrectly predicts 6% of grade C instances as grade B and 25% as grade D. In the classification of grade E, there is a 30% prediction that grade E is considered as background. In contrast, there is a 39% prediction of it and vice-versa. This considerable percentage thinks that Grade E is almost faded, with practically no to little pavement marking left.

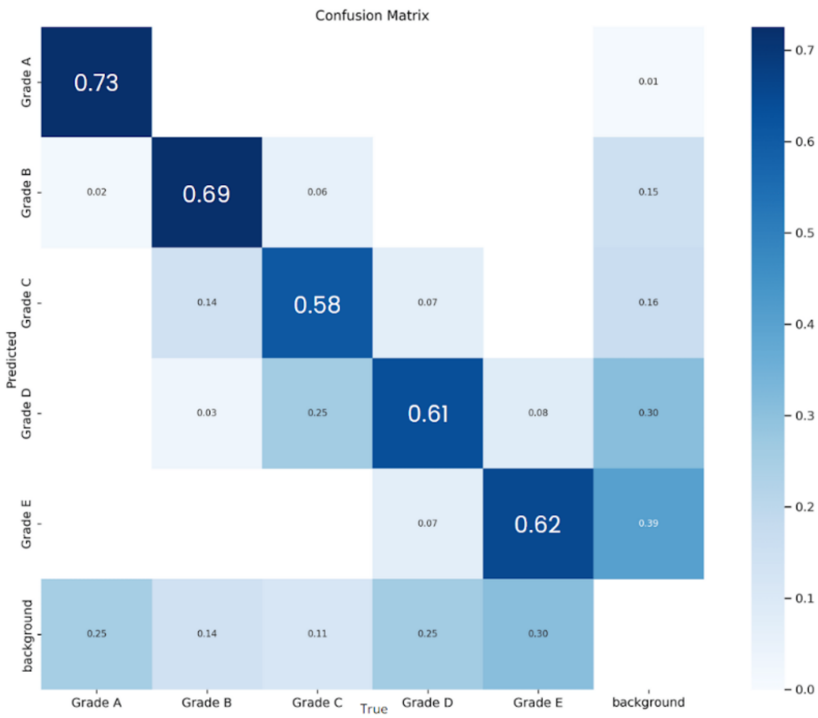


Figure 12. Confusion Matrix of the Best Generated Model

To conclude, the program has a prediction rate of around 60% for each classification grade. The closeness of the results of each grade implies that the program is having difficulty detecting and classifying the markings since it appears similar, and the distinction between them is not much unlike detecting objects with different shapes and kinds. Given this type of similarity, it should be noted that the program may not correctly classify pavement markings in some instances, especially in adjacent grades.

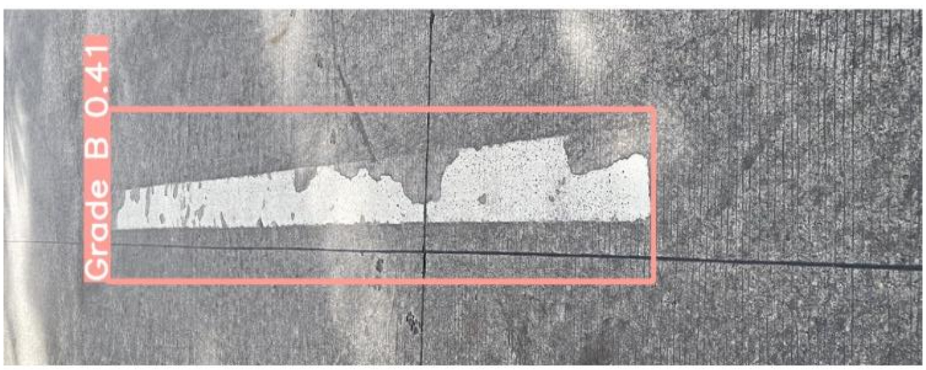
3.4 Verification of Results

A total of 96 pavement markings were photographed around the University. As mentioned in the methodology, it was photographed with an angled and bird's eye view. Using the generated model, the images from the angled views detect 23, 18, 24, 22, and 9 images for grades A, B, C, D, and E, respectively. The bird's eye view of these images was processed in Matlab, where the percentage of remaining paint on the surface was computed using the white and black pixels. This computation considers the percent of paint remaining following the wear index grade. As a result, out of all 23 pavement markings classified by the object detection program as Grade A, 19 were validated by the image processing program to be from the same grade. At the same time, 15 out of 18 pictures matched Grade B. And the matchings for other grades are 20/24, 18/22, and 8/9 for Grades C, D, and E, respectively. A similar confusion matrix to compare the results from the object detection program and the results from the image processing through Matlab was prepared in Table 7. Furthermore, all the processed images were placed side by side and compared to each other, as seen in Figure 13.

Overall, the two programs had produced 80/96 similar grade instances or a match rate of 83.33%, with all classes having a match rate higher than 80%. This implies that the object detection program can be reliable in estimating the percentage of paint remaining on a pavement.

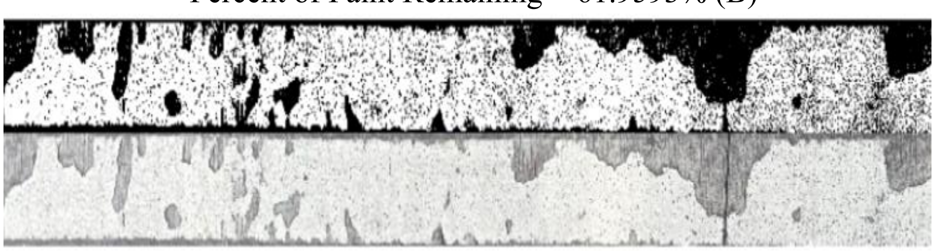
Table 7. Object Detection Program versus Image Processing Program

		Done by the Image Processing Program				
		A	**B**	**C**	**D**	**E**
Done by the Object Detection Program	**A**	0.83	0.17			
	B		0.83	0.17		
	C		0.17	0.83		
	D			0.18	0.82	
	E				0.11	0.89



(a)

Percent of Paint Remaining = 61.9593% (B)



(b)

Figure 13. Detected Pavement Marking (a) with its Image Processed Version (b)

3.5 Validation of Results

Seven pavement markings were subjected to Retroreflectivity Test in the DPWH-Bureau of Research and Standards compound. These markings were also photographed to process in the proposed object detection program. The relationship of the retroreflectivity performance of the seven (7) pavement markings against their corresponding wear grade and its equivalent percent paint remaining are shown in Table 8 and Figure 14. It can be seen that the retroreflectivity values range from 138-165 millicandela per square meter per lux (mcd/m^2/lux). In addition, the retroreflectivity value of a portion of the pavement without any pavement marking was probed to be equal to 100 mcd/m^2/lux. Note that the results of these retroreflectivity values were based on a condition of pavement markings that are not applied under a controlled environment indicating the possibility of the inconsistent composition of reflective glass beads.

Table 8. Results from the Image Processing Program, Retroreflectometer, and Object Detection Program

Paint Remaining (%)	Retroreflectivity (mcd/m^2/ lux)	Classification
0.0000	100	E
27.6688	138	D
30.3158	149	D
54.8845	154	C
54.5980	157	C
64.8205	160	B
87.1103	165	A
80.8978	165	A

A high level of correlation between the two values, with an R^2 equal to 0.83, indicates a positive linear relationship. As the retroreflectivity value increases, the percent paint remaining and its corresponding wear grade index also increase. This is also aligned with the notion that high values of retroreflectivity indicate a good grade of pavement marking condition and vice versa. Furthermore, it was also estimated that the failed pavement marking in terms of wear in relation to the existing standard of DPWH with a minimum retroreflectivity value of 150 mcd/m^2/lux must have a grade condition within Grade C (40-60%). As indicated in Section 2.1.1 and Table 5, the classification Grade C is recommended to be subjected to repainting to provide adequate visibility of pavement markings. However, with the results of the retroreflectivity tests, a conditional Grade C can also be observed, which suggests room for engineering judgment if the pavement marking will be repainted or not.

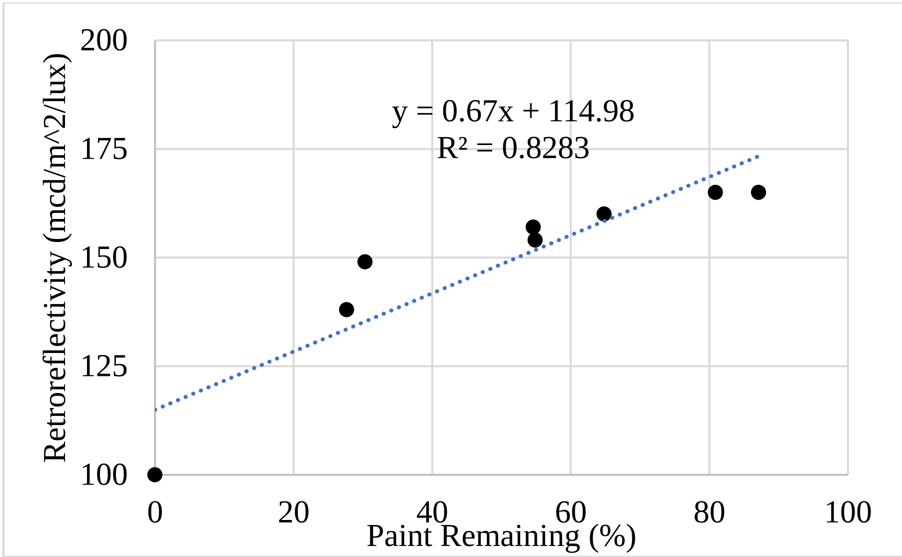


Figure 14. Relationship Between Retroreflectivity and Percent of Paint Remaining

IV. CONCLUSION AND RECOMMENDATION

This study proposes a new, considerably efficient methodology for assessing the wear conditions of pavement markings that the authorized agency can employ. The following conclusions were derived from this study.

a. The wear index implemented by UK and ODOT was adopted in this research considering practicality of their method in time and cost. The two methods were combined and simplified in five-grade classification conditions of the markings.
b. An object detection algorithm called YOLOv5 was used to create a program that can detect and classify the condition of pavement markings using images. The program could detect and predict the classification of markings with a prediction rate of around 60%. Each classification has prediction rates of 73%, 69%, 58%, 61%, and 62% for grades A, B, C, D, and E, respectively.
c. The performance of the program was validated by conducting retroreflective tests, and image processing using Matlab. The retroreflectivity of the pavement markings that was classified by the proposed method under Class C were closed to the value of 150mcd/m^2/lux. The proposed classification that should be subjected for repainting was aligned with the reflectivity value assigned by the current standards of DPWH. Furthermore, the program performances showed a goodness of fit of 83% with the image processing method.

Hence, this study was able to develop an assessment method that can be used to evaluate longitudinal pavement markings. A more extensive data gathering of pavement marking images can improve the training data of the object detection program. The number of instances of pavement markings can give higher accuracy.

V. ACKNOWLEDGEMENTS

This study was funded through the University of the Philippine Alumni Association of Los Angeles (UPAAGLA). This was also conducted with the assistance of the project PAVE-Prototype Automated Visual Survey Equipment funded by the Department of Science and Technology (DOST), which is monitored by the Philippine Council for Industry, Energy, and Emerging Technology Research and Development, and the Department of Public Works and Highways – Bureau of Research and Standards – Technical Services Department.

References

[1] Sigua R. 2016. Fundamentals of traffic engineering. The University of the Philippines Press.

[2] Department of Public Works and Highway. 2012. Retrieved from https://www.dpwh.gov.ph/DPWH/references/guidelines_manuals/highway_safety_design_standards_manual.

[3] Malyuta D. 2015. Analysis of factors affecting pavement markings and pavement marking Retroreflectivity in Tennessee Highways [masters]. University of Tennessee at Chattanooga.

[4] Department of Public Works and Highway. 2013. Retrieved from https://www.dpwh.gov.ph/dpwh/sites/default/files/issuances/DO_103_S2013.pdf.

[5] ASTM E1710-18. Standard test method for measurement of retroreflective pavement marking materials with CEN-prescribed geometry using a portable retroreflectometer. doi: 10.1520/E1710-18

[6] Burrow MPN, Evdorides HT, Snaith MS. 2000. Road marking assessment using digital image analysis. Proceedings of the Institution of Civil Engineers Transport. 141(2):107-112.

[7] Collado, JM, HilarioC, De La Escalera A, Armingol JM. 2006. Adaptive road lanes detection and classification. Proceedings of Lecture Notes in Computer Science. 4179:1151-1162.

[8] Noda M, Takahashi T, Deguchi D, Ide I, Murase H, Kojima Y, Naito T. 2008. Recognition of road markings from in-vehicle camera images by a generative learning method. Proceedings of the Institute of Electronics, Information and Communication Engineers (IEICE). 108(263):31-36.

[9] Iowa State University Institute for Transportation. Retrieved from https://intrans.iastate.edu/app/uploads/2018/03/pvmt_markings_presence_tool_w_cvr. pdf.

[10] Ge H. 2010. Determination of the presence conditions of pavement markings using image processing [masters]. Texas A&M University.

[11] Xu S, Wang J, Wu P, Shou W, Wang X, Chen M. 2021. Vision-based pavement marking detection and condition assessment—A case study. Applied Sciences. 11(7): 3152. http://dx.doi.org/10.3390/app11073152

[12] Azmi N, Sophian A, Bawono A. 2022. Deep-learning-based detection of missing road lane markings using YOLOv5 algorithm. IOP Conference Series: Materials Science and Engineering. doi: 10.1088/1757-899X/1244/1/012021

[13] Guo, K, HeC, Yang M, Wang S. 2022. A pavement distresses identification method optimized for YOLOv5s. Sci Rep 12, 3542. https://doi.org/10.1038/s41598-022-07527-3.

[14] Jiang P, Ergu D, Liu F, Cai Y, Ma B. 2022. A review of Yolo algorithm developments. Procedia Computer Science. 199: 1066-1073. https://doi.org/10.1016/j.procs.2022.01.135

[15] Karthi M, Muthulakshmi V, Priscilla R, Praveen P, Vanisri K. 2021. Evolution of YOLO-V5 algorithm for object detection: Automated detection of library books and performance validation of dataset. 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES). 1-6. doi:10.1109/ICSES52305.2021.9633834

[16] Zhang K, Wang C, Yu X, Zheng A, Gao M, Pan Z, Chen G, Shen Z. 2021. Research on mine vehicle tracking and detection technology based on YOLOv5. Systems Science & Control Engineering. 10(1): 347-366. doi: 10.1080/21642583.2022.2057370

[17] Qui M, Huang L, Tang B. 2022. ASFF-YOLOv5: Multielement detection method for road traffic in UAV images based on multiscale feature fusion. Remote Sensing. 14(14):3498.
[18] Gonal A. Baliga B. 2022. Road damage detection and classification using YOLOv5 on an android device. Ijraset Journal for Research in Applied Science and Engineering Technology.10(8):333-339. https://doi.org/10.22214/ijraset.2022.46175
[19] Zhang Y, Ge H. 2012. Assessment of presence conditions of pavement markings with Image Processing. Transportation Research Record. 2272(1):94–102. https://doi.org/10.3141/2272-11
[20] US Department of Transportation Federal Highway Administration. 1994. Retrieved from https://safety.fhwa.dot.gov/ped_bike/docs/rdwydelin.pdf.
[21] Highways England. 2020. Road Layout Inspection & Assessment: CS 126 - Inspection and assessment of road markings and road studs. Retrieved from: https://www.standardsforhighways.co.uk/prod/attachments/3bbc1e06-b5be-4b5d-a2cd-2e6d93eae24f.
[22] Ohio Department of Transportation. 2008. Construction and Material Specifications (C&MS) Manual.
[23] NDSU Upper Great Plains Transportation Institute. 2017. Retrieved from https://www.ugpti.org/resources/reports/downloads/mpc17-341.pdf.
[24] Mohi A. 2009. Performance evaluation of pavement markings on Portland cement concrete bridge decks [masters]. University of Akron.
[25] Singapore Standard Council. 2013. Specification for hot-applied thermoplastic road marking materials- Materials, performance and application. Retrieved from http://www.biaoxian.org.cn/uploadfile/2021/0610/20210610082422191.pdf.
[26] Jocher G. 2022. Ultralytrics / yolov5. Retrieved from https://github.com/ultralytics/yolov5?fbclid=IwAR0TVQo_ndvUJVqJR-klsB3GA_dvGFR5BfOZZdhApkG95LvSU6MkQdrV0a0.
[27] Machine Learning Mastery. 2020. Retrieved from https://machinelearningmastery.com/confusion-matrix-machine-learning/.
[28] Shah D. 2022. Retrieved from https://www.v7labs.com/blog/mean-average-precision#:~:text=Mean%20Average%20Precision(mAP)%20is%20a%20metric%20used%20to%20evaluate,values%20from%200%20to%201.
[29] Dewi C, Chen R, Zhuang Y, Jiang X, Yu H. 2023. Recognizing road surface traffic signs based on Yolo models considering image flips. Big data and Cognitive Computing Journal. 7(1). doi:10.3390/bdcc7010054
[30] Kang K, Chen D, Peng C, Koo D, Kang T, Kim J. 2020. Development of an automated visibility analysis framework for pavement markings based on the deep learning Approach. Remote Sensing Journal, 12:3837. doi:10.3390/rs12223837
[31] Adhikari B, Huttunen H. 2020. Iterative bounding box annotation for object detection. 25th International Conference on Pattern Recognition (ICPR2020); Milan, Italy. 2021. pp. 4040-4046. doi: 10.1109/ICPR48806.2021.9412956