

# Pavement Distress Classification Using Deep Learning Method Based on Digital Image

Dwi Ratna Sulistyaningrum  
*Mathematics Department*  
*Sepuluh November Technology*  
*Institute*  
 Surabaya, Indonesia  
 dwiratna@matematika.its.ac.id

Daniel Oranova  
*Informatics Engineering Department*  
*Sepuluh November Technology*  
*Institute*  
 Surabaya, Indonesia  
 daniel@if.its.ac.id

Ravy Hayu Pramestya  
*Mathematics Department*  
*Sepuluh November Technology*  
*Institute*  
 Surabaya, Indonesia  
 ravyhayyu@gmail.com

Imam Mukhlash  
*Mathematics Department*  
*Sepuluh November Technology*  
*Institute*  
 Surabaya, Indonesia  
 imamm@matematika.its.ac.id

Budi Setiyono  
*Mathematics Department*  
*Sepuluh November Technology*  
*Institute*  
 Surabaya, Indonesia  
 budi@matematika.its.ac.id

Ervina Ahyudanari  
*Civil Engineering Department*  
*Sepuluh November Technology*  
*Institute*  
 Surabaya, Indonesia  
 ervina@ce.its.ac.id

**Abstract**—Maintaining the road regularly is a necessity, because the road is a vital infrastructure. One of automatic road maintenance steps is the detection of road distress type. Several methods have been used to detect and classify road distress automatically. This research determines the existence and classifies the deterioration of pavement using the Deep Learning method. The type of road distress detected are potholes, line-cracks, and non-line cracks. In this study, the deep learning method implemented is You Only Look Once (YOLO). The YOLO method uses Convolutional Neural Network (CNN) in its architecture and has given good results in object detection both on images and videos. YOLO has been tested in various datasets and given faster and accurate results. In this research, the pre-processing steps are cropping and resizing images then annotating the data. After that, the training is done by fine-tuning YOLO network process. The YOLO architecture uses 9-layer convolution and six layer maxpool. The testing results for datasets show that the highest accuracy is 99% and the highest average IoU is 75,1%. The run time of classification is 0.883 seconds per image.

**Keywords:** *distress classification, deep learning, image processing, You Only Look Once*

## I. INTRODUCTION

Road Maintenance Manual of Directorate General of Highways [1] explained that road defect is divided into 19 types, some of which are alligators cracking, bleeding, block cracking, pothole, and so on. Based on data obtained from the department of public works [2] dated 15 January 2007 shows that the total length of national road has reached 34506.45 kilometers. Along 2529 kilometer of the full length is heavy damage, and 3828 kilometer light damaged. The road condition affects promoting economic growth, safety, and environmental sustainability. Massive damage road may cause high-cost vehicle maintenance and lead to an accident. The infrastructure condition drives productivities and guarantees in reducing price disparity. Additionally, due to the road condition, vehicle speed may be unable to reach the required speed to minimize the emission. Therefore,

maintaining the sound of public infrastructure is essential to ensure sustainable development.

In the road maintenance management system, accuracy and real-time information could prevent the reducing serviceability of the road. The Directorate General of Highways in Indonesia uses various parameters in identifying the performance of the road and highways pavements. The methods available are pavement distress evaluation using pavement condition index, riding quality using international roughness index, and pavement structural condition using falling weight reflectometers (FWD). However, these methods require extensive and experienced labor, and high-cost associated to the vast area to be covered. Besides, the processing time to analyze the data needs quite long and low accuracy. This research attempts to provide an alternative method to reduce any obstacle in collecting and analyzing data of road performance.

Automatically assessment uses the tools that can capture the defect of road image, automatically distinguish and classify the type of road defect, then calculate the level of road defect in accordance with the type of road defect. Therefore, the method needs to be able to classify the type of road defect since the assessment of each road defect can be different. This method provides more effective, objective, and secure in conducting road surveys. The result of the assessment also can be used as a reference in delivering appropriate action for road maintenance.

Automatically identification of road distress has been investigated in a study entitled Road Crack Detection using Deep Convolutional Neural Network (CNN) [3]. This study compares CNN with SVM (support vector machine) and Boosting methods. The result is the CNN method gives the highest precision, which is 86%. In a study entitled Automatic Pavement Crack Detection Based on Structured Prediction with the Convolutional Neural Network [4] use CNN. This study detects cracked roads image and standard road images. By using three types of datasets and several partitions of training and testing data, CNN gives the highest

precision. The precision achieves 96%. This study shows that CNN provides excellent results in detecting and classifying image objects.

In the paper entitled "Classification of Road Defects Using Grey Level Co-Occurrence Matrix (GLCM) and Radial Basis Function (RBF)" has classified the types of damage into three classes namely holes, cracks and other defects, with accuracy reaches 93% [5]. In the study, the dataset was a road image containing one type of damage. In this paper, the classification of road type damage has been carried out using image datasets that contain one defect.

One of the YOLO (You Only Look Once) studies proved that this method, provides high accuracy n Object Recognition in Aerial Images Using Convolutional Neural Network, [6]. This study uses the YOLO method to detect aircraft from satellite imagery. The results show that the accuracy reaches 97.8%. Besides, YOLO has also been used in research which the title is Real-Time Object Detection for Autonomous Driving Based on Deep Learning, [7]. This study uses the YOLO method to detect real-time objects. This study also compares YOLO method with other region-based object detection methods. YOLO gets 84.89% in orientation accuracy and 0.031 seconds per image in processing speed.

In this study, the detects and classifies the types of road damage used deep learning methods. The classes are holes, cracked lines, and non-line cracks. The deep learning method implemented is YOLO (You Only Look Once). YOLO uses a Single Convolutional Neural Network (CNN) to classify and detect the object position using boundary boxes.

II. METHOD

The overall design of our proposed method is shown in Fig. 1. Three main stages in this study are (i) pre-processing, (ii) image classification, and detection using YOLO and (iii) accuracy testing.

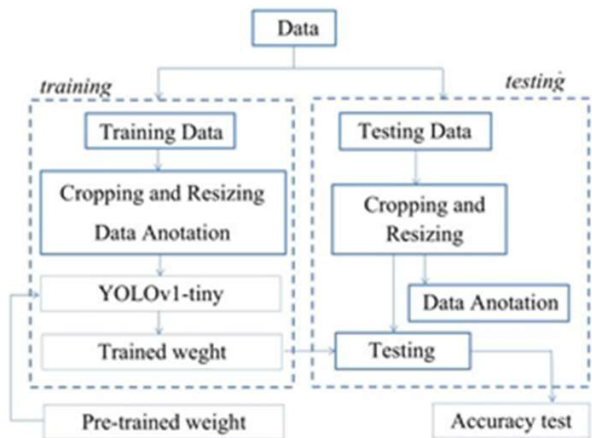


Fig. 1. Flow Diagram of Proposed Method.

A. Pre-processing

Pre-processing is done before detecting and classifying data image. The pre-processing stage is cropping and

resizing the image. Cropping is cutting the initial image data into a square that is adjusted to the location of the object in the image. Resizing is resetting the image size. Then building the data annotation. The process of building data annotation is the process of providing the note as object information in each image. The object information serves as a target or reference to gain weight during the training process. This process is called data annotation.

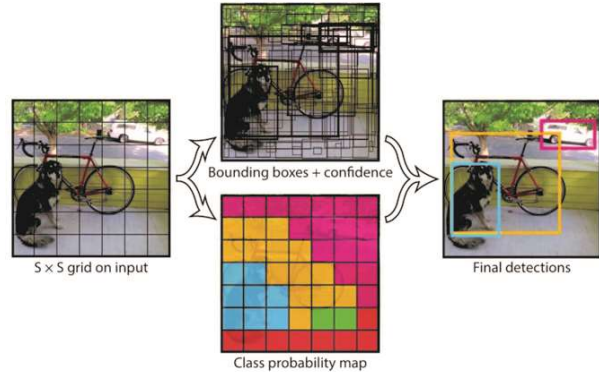


Fig. 2. Model YOLO.

B. Deep learning – You Only Look Once (YOLO)

Deep Learning is a specific sub-field of Machine Learning and is a new way to study the representations of data that emphasize the learning of layers serially and become more meaningful [8]. The latest method in object detection is YOLO (You Only Look Once) method. YOLO makes the object detection process as a single regression problem, which processes directly from image to boundary box coordinates and class probabilities [9].

YOLO uses features of the entire image to predict each boundary box. YOLO predicts all boundary boxes for all object classes in an image simultaneously. YOLO divides the input image into an  $S \times S$  grid. If the center of the object is in a grid, the grid cell is responsible for detecting the object. Each grid cell B predicts the bounding box and confidence values for the boxes. The confidence value reflects how confident the box contains objects and also how accurately the box is predicted.

Each boundary box consists of 5 predictions: x, y, w, h, and value of confidence (p). The coordinates (x, y) represent the center of the box relative to the boundary of the grid cell, width (w), and height (h). Finally, the prediction of the confidence value expresses the IoU between the predicted box and the ground truth box.

Each grid cell also predicts the probability of a conditional class C. This probability is conditioned in the grid cell containing the object. YOLO predicts a set of class probabilities per grid cell, whatever the number of boxes B. The model illustration can be seen in Fig. 2. YOLO applies the model to CNN. The convolutional layer of the network extracts features from the image while the fully-connected layer predicts the probability and coordinates of the output.

The basic version of YOLO has 24 convolutional layers, followed by two fully-connected layers. This architecture

uses a  $1 \times 1$  reduction layer followed by a  $3 \times 3$  convolution layer.

C. Accuracy Test

The performance of the road damage classification method using deep learning is determined through the accuracy of object classifications and the IoU (the Intersection over Union). The equation for classification accuracy is shown in (1).

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

where TP (True Positive) is a class of objects that are correctly classified, TN (true negative) is another class of objects that are correctly classified, FP (false positive) is a class of objects that are not correctly classified, while FN (false negative) is another class of objects that are not properly classified [10].

The accuracy of the object boundary box detection is calculated using the Intersection over Union (IoU) boundary box. The IoU calculates the intersection area between the object prediction box and the ground truth box and divides it by the area of their union [11]. The IoU formulation is shown in (2).

$$IoU = Area\ of\ overlap / Area\ of\ the\ union \quad (2)$$

III. DESIGN AND EXPERIMENT

A. Method Design

The purpose of this study was to detect and classify road distress. So, the expected output is the location and type of each road distress that is contained in the image data. The detection is presented on the boundary box. The kind of road distress is classified into three classes. The three classes are the pothole class, which contains hole type road damage, the crack class which contains the type of cracked line (longitudinal and transversal) and the distressed class which contains non-line crack types.

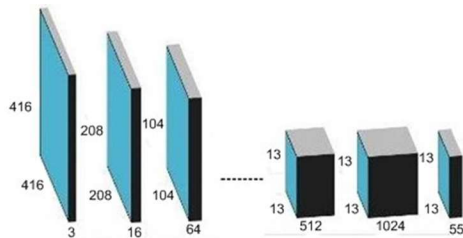


Fig. 3. YOLO v1-tiny Network Architecture.



Fig. 4. Examples of an image to experiment.

The process starts with pre-processing data. The first pre-processing is cropping image data becomes square and resizing to  $416 \times 416$  pixels. The second pre-processing is data annotating. The annotation data file contains values  $[(x_1, y_1), (x_2, y_2), Conf]$  respectively are the coordinates of the beginning of the object boundary box, the coordinate end of the object boundary box, the confidence value of the object boundary box and object class.

The next process is to detect and classify objects using the YOLO method. This study uses YOLI v1-tiny. YOLO v1-tiny network has six convolution layers followed by Max pool, and three convolution layers. YOLO has two main stages, namely, training and testing. In YOLO v1-tiny training stages, the method uses Fig. 3. YOLO v1-tiny pre-trained weight to fine-tuning the data.

B. Experiment

Units The data used in this study amounted to 300 images consisting of images with various types of road damage, namely holes, longitudinal cracks, transversal cracks, and non-line cracks. The data is divided into 240 images for the training process and 60 images for the testing process. Fig. 4 shows an example of some test images.

In the testing process, it is also necessary to adjust the threshold value as the limit of the object confidence value of the boundary box is shown. Some threshold values tested were 0.3, 0.2, 0.1 and 0.05.

IV. RESULTS AND DISCUSSION

In this study, the performance of the detection and classification is influenced by the threshold and epoch values. The method performance is determined by the accuracy of object classification and the detection of the boundary box, as indicated by the IoU average value. The total object for testing is 40 objects in pothole class, 30 objects in the crack class, and 30 objects in non-line crack class or distress class. In this study, indicators of the high accuracy of classification value and IoU average values are 70%.

TABLE I shows the accuracy of object classification. That is the accuracy of an object that is classified in its actual class. TABLE I shows the average value of the IoU bounding box. That is the accuracy of detecting the bounding box on the object.

TABLE I. THE ACCURACY RESULT OF DATASET

Epoch	Threshold	Object Classified			Accuracy
		Pothole	Crack	Distress	
100	0.1	40	28	30	99%
	0.05	40	28	30	99%
300	0.3	40	22	29	91%
	0.2	40	28	29	97%
	0.1	40	29	30	99%
	0.05	40	29	30	99%

We can see in TABLE I, and a high threshold value provides a low accuracy classification value. Threshold value

0.1 and 0.05 gives 99% in the value of object classification accuracy. In the table, it can be seen that many classification errors occur in crack classes.

TABLE II. THE IOU AVERAGE OF DATASET

Epoch	Threshold	IoU Average	Time (s/img)
100	0.1	68.4%	0.850
	0.05	71.4%	0.883
300	0.3	51.4%	0.882
	0.2	66.1%	0.882
	0.1	75.1%	0.883
	0.05	75.1%	0.883

A low threshold value can provide a high average IoU value. We can see in TABLE II, that epoch value 300 and threshold 0.1 and 0.05 produce a high average IoU value. Also, the greater the threshold value is, the smaller IoU boundary box is obtained. Only the boundary box with high confidence will come out as a result. On the other hand, the amount of data used during the training process is small. So, some boundary box has low confidence value.

Fig. 5 is a sample pair of input and output data without interference that has been identified successfully. We can see in the picture; all objects are successfully recognized correctly. While Fig. 6 is a sample pair of input and output data with the presence of white paint on the road as an object interference. The object interference in the image is completely undetectable and unclassified as any type of road distress. Because on the process of data annotation, the road paint object is not trained as any type of road distress.

V. CONCLUSIONS

This research has successfully built a road distress object detection and classification process using the YOLO v1-tiny method with pre-trained weight. The result of the experiment shows the highest accuracy is 99%, and the highest average IoU is 75,1%. The performance is obtained when the epoch is 300, and the threshold is 0.1 and 0.05. The run time of classification is 0.883 seconds per image.

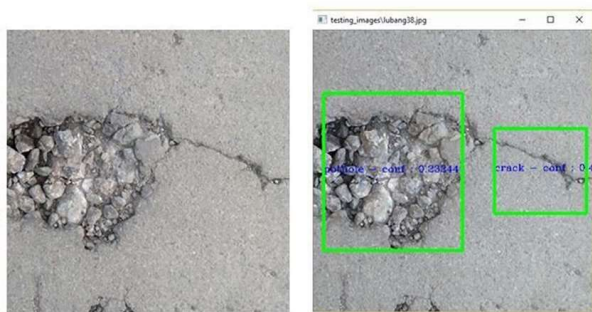


Fig. 5. Bounding box in an image without interference. (a) input image (b) output image.



Fig. 6. Bounding box in an image with object interference. (a) input image (b) output image.

ACKNOWLEDGMENT

This research is supported by the Institute for Research and Community Development (LPPM) Sepuluh Nopember Technology Institute Surabaya with Funding Agreement Letter number 961/PKS/ITS/2019. We thank LPPM-ITS for giving us a chance to present this paper in the International Conference.

REFERENCES

- [1] Department of Public Works, Routine Maintenance Manual for National and Provincial Roads, vol. 2: Standard Repair Method (in Indonesian), 1995.
- [2] The National Transportation Policy Coordinating and Evaluation Team of the Indonesian Ministry of Economic Affairs, Policies and Strategies to improve sustainable management of road infrastructure (in Indonesian), 2009.
- [3] L. Zhang, F. Yang, Y.D. Zhang, and Y.J. Zhu, Road crack detection using deep convolutional neural network, in Image Processing (ICIP), 2016 IEEE International Conference on, IEEE, pp. 37083712, 2016.
- [4] Z. Fan, Y. Wu, J. Lu, and W. Li, Automatic pavement crack detection based on structured prediction with the convolutional neural network, arXiv preprint arXiv:1802.02208, 2018.
- [5] R. H. Pramestya, D. R. Sulistyanningrum, B. Setiyono, I. Mukhlash and Z. Firdaus, "Road Defect Classification Using Gray Level Co-Occurrence Matrix (GLCM) and Radial Basis Function (RBF)," 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), pp. 285-289, 2018.
- [6] M. Radovic, O. Adarkwa, and Q. Wang, Object recognition in aerial images using convolutional neural networks, Journal of Imaging 3(2), 21, 2017.
- [7] G. Liu, Real-time object detection for autonomous driving-based on deep learning, PhD thesis, 2017.
- [8] G. Mooij, I. Bagulho, and H. Huisman, Automatic segmentation of prostate zones, arXiv preprint arXiv:1806.07146, 2018.
- [9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, , You only look once: Unified, real-time object detection, in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779788, 2016.
- [10] S.K. Ryu, T. Kim, and Y.R. Kim, Image-based pothole detection system for its service and road management system, Mathematical Problems in Engineering, 2015.
- [11] M. Everingham, L. Van Gool, C.K. Williams, J. Winn, J. dan A. Zisserman, The pascal visual object classes (voc) challenge, International journal of computer vision 88(2), 303338, 2010.