Enhancing Pothole Detection Robustness to Blurred Images and Real-Time Implementation on Embedded Boards

Youngjae Sim[1], Taehee Kim[2]

***Abstract*— Potholes, formed by the deterioration of pavement due to environmental stressors and wear, pose a growing challenge to road infrastructure, especially as climate change accelerates temperature extremes. Existing machine learning approaches to pothole detection often underperform due to limited datasets, which lack the diversity and volume necessary to account for real-world conditions such as blurred images from fast-moving vehicles. This research proposes a novel pothole detection model that integrates an image deblurring network with a pothole detection network. The experimental results demonstrate the effectiveness of our method, showing significant improvements in average precision and detection accuracy. Implemented on an embedded board, it successfully detected potholes in images captured by fast-moving vehicles.**

I. INTRODUCTION

Potholes are large depressions on the road surface and cause numerous road accidents. In South Korea from 2018 to 2023, there were nearly 6,000 pothole accidents, costing 9.1 billion KRW in compensation [1]. Even globally, especially in the U.S., pothole damages cost drivers $3 billion annually, or $306 per driver on average [2]. As climate change and extreme fluctuations in temperature increase pothole frequencies on the road, the problem of detecting and repairing potholes is becoming more vital [3].

As pothole accidents are increasing, there is an urgent need for building an accurate algorithm that detects potholes under real-world conditions. Researchers have developed various machine learning models for pothole detection [4, 5, 6]; however, previous models have failed to be efficiently implemented in real life because they were trained by a small dataset with a lack of diversity, which resulted in bias. Moreover, as images in the dataset were clear images of potholes on roads, the models failed to detect potholes in blurry road images that are more realistic today with the norm of fast driving.

[1] Y. S. is with Korea International School, 27, Daewangpangyo-ro, 385 beon-gil Bundang-gu, Seongnam-si, Gyeonggi-do South Korea 13543 (corresponding author to email: jimmysim2026@gmail.com).
[2] T. K. is with St. Johnsbury Academy Jeju, 10, Global edu-ro, 304 beon-gil, Daejeong-eup, Seogwipo-si, Jeju Special Self-Governing Province South Korea 63644 (corresponding author to email: seconds19012926@gmail.com).

In this research, we introduce a novel pothole detection system capable of detecting potholes in blurry images. We have successfully implemented the method on an embedded board and validated its utility in real-world applications.

II. BACKGROUND KNOWLEDGE

Potholes form through a process known as the freeze-thaw process, where water seeps into the road and collects underneath. During winter, the water freezes, expands, and pushes up the road surface. As sunlight softens the water below the cracked surface, an empty pocket forms; under the weight of traffic, the pocket collapses and leaves a pothole. Driving over potholes at high speeds can cause fatal accidents for the driver and nearby drivers. Potholes are especially dangerous in winter, as driving over ice can result in black ice, a coating of glaze ice on a surface. Potholes are also a cause of heavy financial burdens for both drivers and local governments, as pothole accidents require expensive repairs. Because of these threats, a system for automating pothole detection is crucially needed for road regulation.

Object detection is a major machine learning technique used to locate instances of objects in images or videos. Classification is a type of machine learning used to identify and categorize objects. A prominent classification network is a Convolutional Neural Network (CNN). CNNs take images as input and output specific categories that the input images most likely belong to. Object detection is an essential part of this research, as it involves detecting potholes from visual inputs with state-of-the-art models. However, applying these models to real-world conditions introduces new challenges.

Recording clear, high-quality videos from vehicles traveling at high speeds is a significant technical challenge. One issue is motion blur, which occurs when the camera's speed is not fast enough to freeze and capture the rapid movement of the landscape, resulting in details appearing obscured. Another problem is the limitations of vehicle camera frame rates (FPS): as general video cameras usually offer 30-60 FPS, most videos become blurry in high-speed environments like vehicles driving up to 80 mph (137 km/h) on highways. Additional factors like vehicle vibrations and weather conditions (rain, snow, etc.) further affect the clarity of captured videos. With these difficulties, recording clear videos for data has been a challenge in previous research. This research aims to tackle this issue, detecting potholes at high accuracy even with blurry images.

III. PROPOSED METHOD

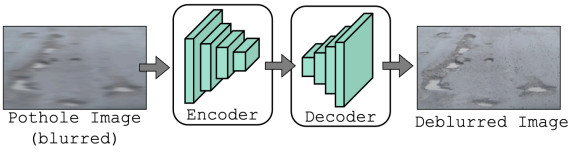


Figure 1. Proposed image deblur network

The proposed image deblur network deblurs pothole images from the dataset, ultimately improving pothole detection. The network is composed of an encoder and a decoder. The encoder processes the blurry image as input and extracts key features by generating a feature map, which encodes the structural information of the pothole in the image. The decoder then reconstructs a deblurred image from the feature map, providing a clearer version of the original input.

$$L_1 = \frac{1}{WH}\sum_x\sum_y\left|I(x,y) - \widehat{I(x,y)}\right|$$

Equation 1: L1 Loss Function

To train the network, we employed the L1 loss function, which evaluates the loss of the proposed image deblur network in deblurring the entire image. W and H are the width and height of the image, respectively. $I(x, y)$ is the actual pixel value of the blurry image at position (x, y), whereas $\widehat{I(x, y)}$ is the deblurred pixel value at position (x, y). The absolute difference in the actual and deblurred pixel values is added for all position values x and y.

$$L_p = \left|\left|VGG19(I) - VGG19\left(\widehat{I}\right)\right|\right|^2$$

Equation 2: Perceptual Loss Function

Additionally, to evaluate the perceptual discrepancy of the deblur network, we used the perceptual loss function. I—the actual image's pixels—and $\widehat{I}$—the deblurred image's pixels—are inputted into an ImageNet pre-trained VGG19 CNN [7, 8]. Then, the resulting activation maps are compared and squared.

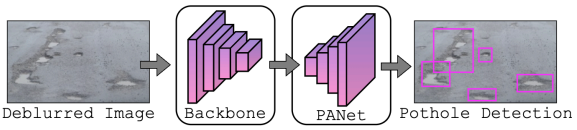


Figure 2. Proposed pothole detection network

The proposed pothole detection network detects potholes in deblurred images. Given an image as input, it outputs the original image with bounding boxes around predicted potholes. The backbone outputs a feature map from the inputted image, which is taken by PANet. PANet outputs the original image with predicted bounding boxes around potholes. Three loss functions are employed to evaluate the performance of the network: the confidence loss function, classification loss function, and regression loss function.

$$L_{confidence} = \sum_{s\in S^2}\sum_{b\in B} I_{s,b}^{obj}\left(C_s - \widehat{C}_s\right)^2 + \lambda_{noobj}\sum_{s\in S^2}\sum_{b\in B} I_{s,b}^{noobj}\left(C_s - \widehat{C}_s\right)^2$$

Equation 3: Confidence Loss Function

The confidence loss function evaluates the pothole detection network's detection of potholes. C_s is either 1 or 0, for whether the inputted image has a pothole at pixel s. $\widehat{C}_s$ is the predicted probability of the image having a pothole at pixel s. The value of λ is less than 1 and serves to reduce any bias that $noobj$—cells without objects—has on the value of the function. The indicator function obj ensures that the first squared difference is only considered for cells with potholes, whereas $noobj$ ensures that the second squared difference is only considered for cells without potholes.

$$L_{classification} = \sum_{s\in S^2} I_s^{obj}\sum_{c\in classes}\left(p_s(c) - \widehat{p}_s(c)\right)^2$$

Equation 4: Classification Loss Function

The classification loss function evaluates the model's ability to assign correct class probabilities to different objects detected in the inputted image. $p_s(c)$ is the ground truth probability (1 or 0) of class c at cell s, whereas $\widehat{p}_s(c)$ is the predicted probability of the class at that cell. The squared error between the predicted and ground truth probabilities for each class is only calculated for cells with potholes.

$$\begin{aligned} L_{regression} &= \sum_{s\in S^2}\sum_{b\in B} I_{s,b}^{obj}\left[\left(x_s - \widehat{x}_s\right)^2 + \left(y_s - \widehat{y}_s\right)\right] \\ &+ \sum_{s\in S^2}\sum_{b\in B} I_{s,b}^{obj}\left[\left(\sqrt{w_s} - \sqrt{\widehat{w}_s}\right)^2 + \left(\sqrt{h_s} - \sqrt{\widehat{h}_s}\right)^2\right] \end{aligned}$$

Equation 5: Regression Loss Function

The regression loss function evaluates the model's ability to define the position and size of pothole bounding boxes. x_s and y_s are the actual positions of the potholes in the image, and $\widehat{x}_s$ and $\widehat{y}_s$ are the predicted positions. w_s and h_s are the actual width and height of the pothole bounding box, whereas $\widehat{w}_s$ and $\widehat{h}_s$ are the predicted width and height. The four values are put inside a square root to reduce any potential bias from disproportionate dimensions.

IV. EVALUATION

To train the model for diverse conditions, we combined datasets from Roboflow and Google Bangkit 2021 [9, 10]. The Roboflow dataset consists of 665 images captured from long distances and the Google Bangkit 2021 dataset consists of 3,125 training samples and 843 test samples captured from short distances. We divided the combined dataset into 80% training samples and 20% test samples. By combining two datasets of pictures taken from various distances, we allowed the model to learn diverse road patterns, thereby enhancing its versatile detection of potholes.

	Normal	Blurred	Blur Removed
EfficientDet	65.4	48.1	54.4
RetinaNet	66.7	49.5	54.8
FocalNet	70.1	53.7	57.9
YOLOv5	74.7	54.7	60.1
DETR	77.5	57.9	65.2

Table 1. Compassion between AP scores of five different CNN architectures in three different conditions: normal (clear) images, blurry images, and blurry images deblurred with our deblur network

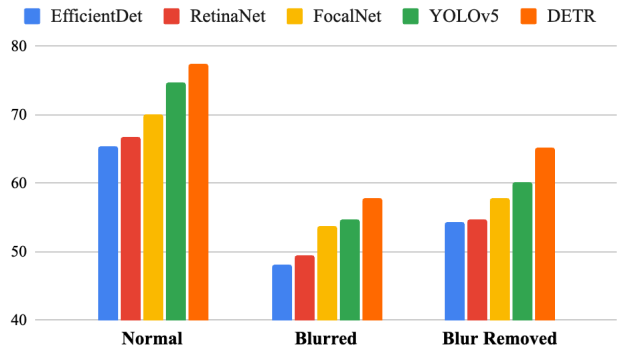
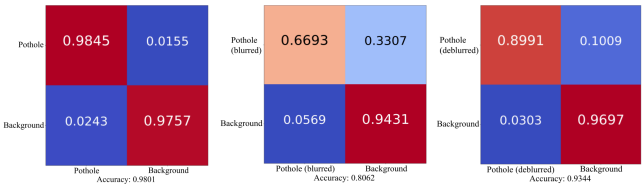


Figure 3. Graph of Table 1

Table 1 includes five different CNN architectures' average precision (AP) scores with normal (clear) images, blurry images, and blurry images deblurred with our proposed image deblur network. As seen in the table, the proposed method has been continuously effective for all five architectures compared to blurry images. Among the five, DETR has shown the highest AP score with blur-removed images. However, DETR's high computation amount and memory consumption make it implausible to be implemented in an embedded board for practical purposes. YOLOv5, with the second highest AP score, was thus determined as the most efficient and plausible architecture to maximize our proposed methods' performance.



(a) (b) (c)

Figure 4. Confusion matrix evaluation
(a): normal samples; (b): blurred samples; (c): deblurred samples (proposed deblur process applied)

Figure 7 illustrates confusion matrices of the model's accuracies when trained with normal images, blurred images, and blurred images applied to the proposed image deblur network. As seen in confusion matrix (a), when trained with normal, clear samples, the model achieved high accuracies for detecting both potholes (0.9845) and backgrounds (0.9757). In contrast, when trained with blurred samples (b), the model had a significantly lower accuracy in detecting potholes (0.6693) and a slightly lower accuracy in detecting backgrounds (0.9431). When the deblur process was applied on blurred samples (c), the model had a significant increase in accuracy for detecting potholes (0.8991) compared to (b) and had a background detection accuracy very close to that when trained with normal samples (0.9697). The data demonstrates the effectiveness of the proposed deblur network in removing blurry artifacts and thereby allowing the model to accurately detect potholes.

V. Conclusion

This research proposed a novel pothole detection system capable of accurately functioning even with blurry images. The image deblur network deblurs blurry images from the dataset, which helps improve the accuracy of pothole detection. The experimental results clearly demonstrate the effectiveness of our deblur method, as it has significantly improved both AP scores and accuracies in detecting potholes compared to training with just blurry samples. From the data, YOLOv5 has been determined appropriate for practical usage. Implemented into the NVIDIA Jetson Board, the proposed model was able to detect potholes from images taken from fast-driving vehicles. For further research, our image deblur network can be improved through incorporating more, diverse training data. We also believe integrating additional sensors like accelerometers will enhance detection accuracy under various conditions.

References

[1] Kim H. S., "Compensation for Damage from Highway Damage: KRW 9.1 Billion Over 5 Years," *Urban Maintenance News*, Oct. 12, 2023. https://www.dosijeongbi.com/4758
[2] A. Vigderman, "The State of Potholes in the U.S.," Autoinsurance.com, Nov. 27, 2023. https://www.autoinsurance.com/research/pothole-damage/
[3] K. Ravilious, "Weatherwatch: pothole problem to deepen amid climate crisis," *The Guardian*, Mar. 07, 2024. Available: https://www.theguardian.com/world/2024/mar/07/weatherwatch-pothole-problem-deepen-climate-crisis
[4] C. Wu et al., "An Automated Machine-Learning Approach for Road Pothole Detection Using Smartphone Sensor Data," MDPI, vol. 20, no. 19, Sep. 2020, doi: https://doi.org/10.3390/s20195564.
[5] O. A. Egaji, G. Evans, M. G. Griffiths, and G. Islas, "Real-time machine learning-based approach for pothole detection," *Expert Systems with Applications*, vol. 184, Dec. 2021, doi: https://doi.org/10.1016/j.eswa.2021.115562.
[6] C. Saisree and Dr. K. U, "Pothole Detection Using Deep Learning Classification Method," ScienceDirect, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.190.
[7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," ScienceDirect, Jun. 2009, doi: https://doi.org/10.1109/cvpr.2009.5206848.
[8] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv*, Sep. 2014, doi: https://doi.org/10.48550/arXiv.1409.1556.
[9] A. R. Chitholian, "Pothole Dataset," Roboflow, Nov. 2020. https://public.roboflow.com/object-detection/pothole
[10] A. A. Praramadhan, "Potholes Detection for YOLOV4," Kaggle, 2021. https://www.kaggle.com/datasets/anugrahakbar/potholes-detection-for-yolo v4