Review of: "CNN-Based Road Damage Detection"

Prasenjit Dhar

Potential competing interests: No potential competing interests to declare.

Recommendation: REJECT

The paper proposes a CNN-based robot system for detecting different types of road damage such as cracks, potholes, and bumps using the YOLOv5 object detection algorithm. The system is trained on a dataset of road damage images and can detect and classify road damage in real time. Overall, the presentation quality of the paper is poor, with inappropriate choice of words at certain places. Tables and figures are not placed as per the standards. There are lots of grammatical mistakes in the paper. A very small number of references are used.

1. The title of the paper is not specific. It should contain the main key terms of the methods used in the paper.

2. In the abstract, the authors state, “Our suggested method is extremely accurate, around 90%...” How do they claim this result without any state-of-the-art comparison? The keywords used after the abstract are very poor.

3. In the Introduction, the authors talk about previous works in the literature. However, the lack of citations in the Introduction makes the paper less comprehensive. The contributions and specific goals of the paper are missing.

4. In the literature survey, the limitations of the existing works have not been presented. Further, the research gaps have not been identified in the paper.

5. There has to be some lines of content between section 3 and 3.1.

Many sentences in section 3.1 do not relate to the model overview. They talk about road conditions only.

What do the authors mean by the line “Our experiments suggest that our technique can achieve a Mean F1-Score” ? The statement is very hypothetical.

What exactly do Figures 1 and 2 represent ? There is no description of these figures. Further, both the figures are quite conceptual and generalized, which do not provide any insights into the proposed framework.

The sources of data collection have not been mentioned. Annotation details are also missing.

What exact data pre-processing techniques have the authors used in the paper ?

The YOLOv5 architecture is missing.

The training, hyperparameter, and experimental configuration details are missing.

The hardware implementation and integration details are not discussed in-depth.

The scalability and computational efficiency of the proposed system for large-scale deployment are not discussed.

Overall, the methodology section does not provide much insight into deploying the framework in real time.

6. There has to be some lines of content between section 4 and 4.1.

No sample images of the input dataset have been provided in the paper, which does not provide insights on the different classes used.

Detailed information about the dataset, such as image size, format, resolution, stats, etc., is missing.

The actual results should be presented in a separate sub-section.

The performance metrics, such as precision, recall, F1-score, and mAP, are not clearly discussed.

The precision, recall, F1-score, accuracy (or mAP) values in tabular form are missing, which makes the paper incomprehensive.

The paper does not compare the proposed approach with existing road damage detection methods, which makes the paper less competitive.

Figure 6 is unnecessary. Rather, testing images should be included as outputs, which is missing in the paper and raises a question about the existence of the hardware model.

Figure 8 is totally not understood.

What do the authors infer from Figures 3, 4, 5, 8, 9, and 10? No analysis has been presented for these results.

Detailed discussion on the proposed approach is missing.

Overall, the result and discussion section is poorly presented.

7. The conclusion section is very shallow.

There is no need to put future scope as a separate sub-section.

The last statement in the future scope is not clear. To the best of my knowledge, most of the future prospects mentioned do exist in the current scenario.