Article │ Received 6 June 2023; Accepted 20 July 2023; Published 30 November 2023
https://doi.org/10.55092/pcs2023020002

Deep learning for road defect detection from aerial imagery

Amit Hasan Sadhin[1,*], Siti Zaiton Mohd Hashim[1], Reshad Rayhan[2]

[1] Faculty of Computing, University of Technology Malaysia, Johor, Malaysia
[2] Faculty of Electrical Engineering, University of Technology Malaysia, Johor, Malaysia

* Correspondence author; E-mail: samithasan7866@gmail.com.

Abstract: This research highlights the efficiency of YOLO-NAS and Detectron2 algorithms in detecting road defects like cracks and potholes using drone-captured aerial images. YOLO-NAS demonstrates significant accuracy, achieving a mAP score of 71.23% and F1 score of 70.04%, outperforming previous YOLO versions. Detectron2 exhibits an AP score of 55%, surpassing state-of-the-art experiments in coco instance segmentation. Both models display confidence values close to 100%, ensuring reliable object detection. The results show the potential of integrating drone-based inspection systems with deep learning algorithms to improve road safety, reduce manual efforts, and enhance infrastructure management. This approach can contribute to a country's economic and social progress by facilitating efficient road maintenance and defect detection. Future implementations may involve real-time detection using drones for timely road defect assessment and decision-making.

Keywords: YOLO-NAS; neural network; Detectron2; UAVs; quantization; AutoNAC

1. Introduction

Since thousands of years ago, roads have played a significant part in human civilisation. A well-maintained road may do all these responsibilities for the general economic and social growth of a nation, including transporting the sick to the hospital, students to education, and the nation to development. But maintaining the standard and upkeep of the road infrastructure has never been easy for the authorities. Road imperfections are a typical occurrence in practically every nation on earth. The most dangerous flaws that result in risky driving are cracks and potholes. But regrettably, due to a lack of human resources, maintaining roads has always been difficult, particularly when trying to identify problem areas [1]. Maintaining the integrity and safety of road networks depends on the prompt detection and repair of such flaws. The time required for inspections and general maintenance can be greatly decreased by combining drone technology with deep learning techniques.

For smooth and safe travel, to reduce vehicle maintenance costs, and to boost transportation effectiveness, roadways must be free of defects. By identifying road flaws like cracks and potholes, one can reduce the time required for general road repair and increase road safety [2]. Road flaws that are promptly inspected and fixed can greatly lower the accidents brought on by defective roads, thereby saving lives and lessening the financial burden involved with maintaining road infrastructure.

Due to significant road repair programs and superior infrastructure management, accidents caused by road defects may be reduced in industrialized nations. However, according to the United Nations Road Safety Collaboration's 2021 report, road infrastructure issues are still at blame for 22% of all incidents. Inaccurate road inspection, human mistake, and a lack of comprehension are the main causes of this. An automated drone base inspection system may greatly eliminate human inspection mistakes, improving life safety.

Drones are crucial to inspections because they enable access to challenging locations, capture sharp photographs, and gather information quickly. Compared to manual inspection, they increase safety while saving time and money [4]. Real-time data enables quicker decision-making, better management of the infrastructure, and more successful maintenance interventions. A stable setting like a UAV-based inspection system for road safety will greatly benefit.

In recent times, deep learning algorithms are showing remarkable influences in inspection in safety-critical systems [5] and the ability of decision-making and object detection is being well functioned every second. This study will use three algorithms such as YOLO-NAS, and Detectron2, to determine the best-performing algorithms from drone-based road defect inspection systems.

2. Literature review

In recent times, the integration of deep learning algorithms and drone technology has had a profound impact on various industries, particularly in terms of improving robustness, precision, and stability, which will advance social progress, economic growth, and life safety. Notably, different classification and detection algorithms have been applied to safety-critical inspections, such as road defects, bridge inspections, and pavement cracks [6]. Among these, detecting road defects, including potholes and road cracks, holds paramount importance for ensuring human safety, as these imperfections can lead to significant damage and accidents. As a result, accurately identifying road defects while reducing inspection time remains a top priority in the pursuit of road safety. The advancements in deep learning and drone-based inspections are expected to play a pivotal role in achieving these objectives effectively and efficiently.

The recent version of YOLO family network YOLO-NAS shows remarkable accuracy and fast processing, which employs Neural Architecture Search (NAS) to maintain accuracy during post-training quantization, making it most suitable for detecting undersized objects from critical background, such as road defects images from drone [7].

In addition to YOLO-NAS, notable progress has been made in defect detection methodologies. For instance, [8] employed Adaptive Pixel Segmentation to detect road cracks by leveraging Gaussian Cumulative Density. Their deep learning approach demonstrated acceptable accuracy and robustness in various environments. On the other hand, crack detection at the pixel level using encoding-decoding methods, as explored by [9] involves the use of black-box images, enabling pixel-by-pixel analysis. These distinct techniques showcase advancements in defect detection, offering valuable insights into detecting road cracks effectively and handling diverse imaging scenarios.

Furthermore, Detectron2 brings a new dimension to object detection with its versatile and modular design, revolutionizing the field. Several key research has highlighted its capabilities and contributions to the world of computer vision: The detectron2's two-stage approach, incorporating with region proposal and refinement stages helps it to be prices and efficient for object detection and its modular design helps for adapting for different background [10].

The drone-based inspection system is revolutionizing in recent times, [11] presents the advantages of employing Unmanned Aerial Vehicles (UAVs) for crack damage detection in civil infrastructure. The study highlights how aerial inspections offer a comprehensive and holistic view of large areas, leading to a reduction in inspection time while ensuring the safety of inspectors. The use of UAVs in inspections accelerates the process, facilitating better decision-making by providing detailed imagery and data from various perspectives.

3. Methods

YOLO-NAS and Detectron2 algorithms were used in this study to identify road faults including cracks and potholes.

3.1. Dataset

Images captured by drones operated by the University of Technology Malaysia were used in this study for model training, testing, and validation. Without augmentation, the acquired road defect datasets comprise a total of 1000 000 images; after augmentation, the number of total images was 2700+. consists of two different kinds of potholes and road fissures. During model training, both categories are used to identify road faults.

The pictures were taken at a distance of 25 to 35 feet. Each image in the collection is distinct and has never been used in any previous research. Except for the identical cracks and potholes photographs that were taken at various angles and heights, there are no duplicate images in the dataset. Both photos with several cracks and potholes and photographs with just one crack and pothole are shown. The pictures were taken under various actual-world conditions. There are a total of 617 cracks and 592 potholes according to the data gathered.

3.2. YOLO-NAS

YOLO-NAS is the most recent version of YOLO networks, which shows significant acceptance accuracy in different object detection works due to its remarkable accuracy and fast processing. YOLO-NAS use QSP and QCI blocks which allows NAS to maintain accuracy during post training quantization, and this is the reason for a stable accuracy in YOLO-NAS—figure 1: below shows the overall YOLO-NAS architecture and features extraction techniques throughout YOLO-NAS layers.

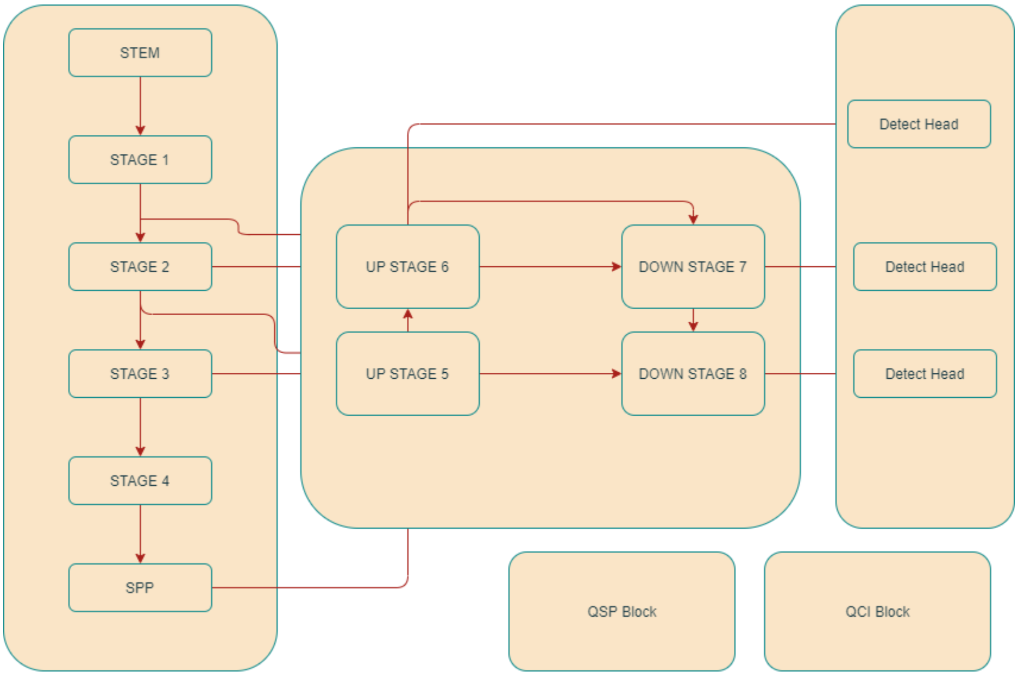

Figure 1. YOLO-NAS Achitecture, NAS or Neural Architecture Search used AutoNAC for the latency balancing adopted from [7].

YOLO-NAS follow a hybrid quantization method developed by Deci, a company that focuses on creating and deploying deep learning models. The primary purpose of YOLO-NAS is to detect undersized objects from critical backgrounds, as in the dataset collected using drones. The hybrid approach of YOLO-NAS introduces the flexibility of detecting comparatively small objects with high accuracy. In YOLO-NAS, Re-parameterization for 8-bit quantization is used with quantization-aware modules known as QSP and QCI to reduce accuracy loss during post-training quantization, as shown in Figure1.

The used AutoNAC system was essential in the development of YOLO-NAS, and it is adaptable to various tasks, data specifics, inference environment, and performance objectives. It enables users to tailor an appropriate structure to their unique demands, achieving an optimal trade-off between accuracy and inference time. This method considers data, hardware, and other elements relevant to the inference process.

3.3. Detectron2

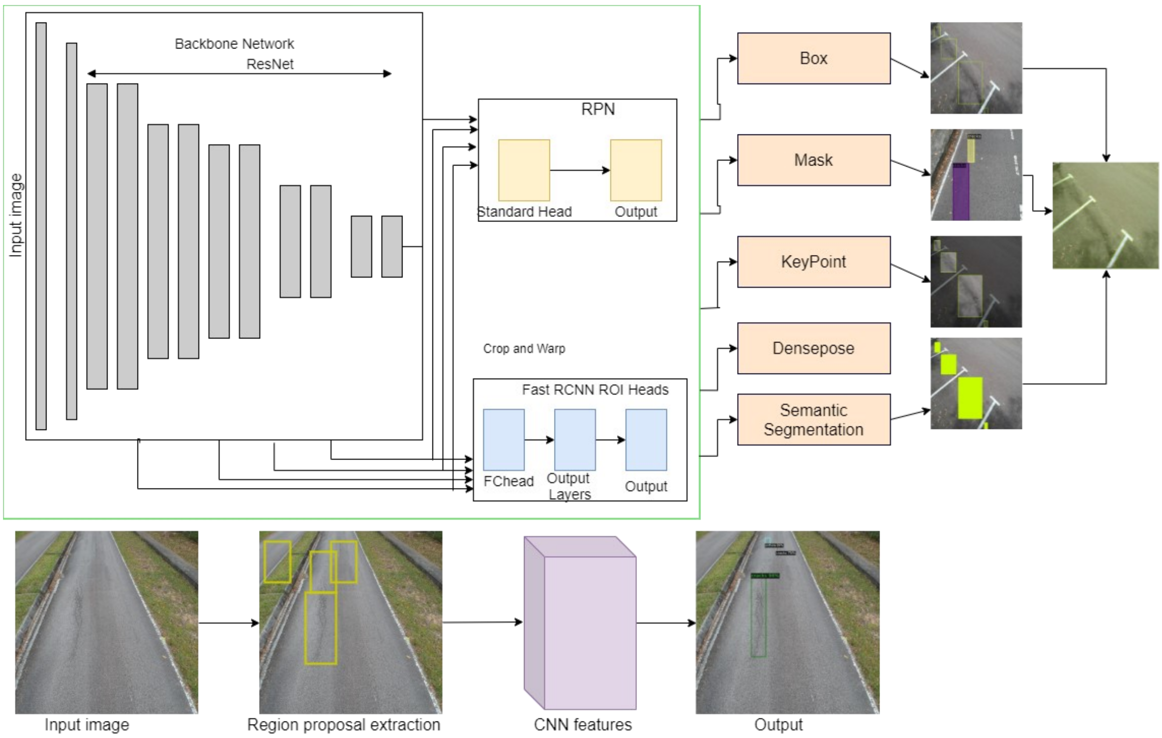

Figure 2. Detectron2 high level architectural overview on road defect images (adopted from meta-AI research).

Detectron2, as depicted in Figure-2, is a state-of-the-art object detection framework built on the PyTorch library. Its implementation in PyTorch offers unparalleled flexibility and extensibility, enabling fast processing on both multiple GPU and single GPU servers. Detectron2 is known for its high-quality implementations of various models, including DensePose and Mask R-CNN.

In this study, the region-based convolutional neural network (Faster R-CNN) is being integrated into the Detectron2 framework. An addition to the original Fast R-CNN called Faster R-CNN was created to increase the efficiency and precision of object detection tasks. The introduction of the Region Proposal Network (RPN), as shown in figure 2, which significantly accelerates the speed of region proposal generation, constitutes the fundamental innovation of Faster R-CNN.

The meta-architecture of the base RCNN (Region-based Convolutional Neural Network) with Feature Pyramid Network (FPN) in Figure 2 showcases a high-level overview of the backbone network's functionality. The backbone network efficiently extracts feature maps from the input images, which are subsequently fed into the Region Proposal Network for detecting object regions. The multi-scaled features enable the RPN to generate proposal boxes, and through cropping and warping, feature maps are transformed into fixed-size features. These features are then used to fine-tune the box positions and classify the objects using fully-connected layers.

Detectron2's robust architecture and flexible design make it a popular choice for object detection tasks. The framework's ability to efficiently implement various models, including Faster R-CNN, and its seamless integration with PyTorch allow researchers and practitioners to achieve state-of-the-art performance in object detection tasks, such as road defect detection from images acquired by drones.

3.4. Performance evaluation

The F1-score, also referred to as the F-measure, is used to calculate the typical mean of recall and accuracy. A balanced model with the appropriate combination of recall and precision can be created by maximizing the F1 score [12].

$$F1 = 2 \times \frac{\text{Recall x Precision}}{\text{Recall+Precision}} \tag{1}$$

Most object identification models use mAP (Mean Average Precision) and AP (Average precision) to calculate their overall accuracy. The trade-off between precision and recall is represented by the AP's consideration of false positives (FP) and false negatives (FN). For most detecting activities, AP is therefore a suitable measurement [13].

$$mAP = \frac{1}{N}\sum_{i=0}^{N} APi \tag{1}$$

4. Result and discussion

In this research, two prominent object detection algorithms are being used to analyses the performance comparatively into this new dataset. In recent times, YOLO-NAS shows significant advancement among the YOLO family networks, whereas the Detectron2 framework shows extraordinary performance in object detection.

The YOLO-NAS shows a mAP score of 71.23% and F1 score of 70.04% which shows significant accuracy than the previous experiment performed on YOLOv5 by Arya *et al.*, 2020, while the Detectron2 shows an AP score of 55%, which is better than the existing state-of-the-art experiments of Detectron2 with coco instance segmentation. In this particular research, faster RCNN X101-FPN has been used, which gives a total of 51.4% F1 scores in global road defect detections challenge 2020's experiments [15], while in this experiment using drone images, the F1 score is 55.17%, which shows the remarkable performance of Detectron2 in road defects detection using aerial imagery. In terms of overall training time, YOLO-NAS took 5.43 Hours for a dataset containing 2700 images of 800 Pixels while for the same dataset Detectron2 took 3 hour 32 minutes in NVIDIA A100 Tensor Core GPU.

Table 1. Performance comparison of object detection algorithms for road defects using aerial imagery.

Algorithm	mAP Score	F1 Score	AP Score	Training Time (hrs)
YOLO-NAS	71.23%	70.04%	-	5.43
Detectron2	-	55.17%	55%	3.53

The confidence values of YOLO-NAS and Detectron2 shows a significant percentage. The confidence value is crucial to determining correctly detected objects' class probability. It functions as a probability score, reflecting the model's confidence in accurately detecting a specific object class. False positives can be filtered out by setting a threshold, assuring precision. In this research, both models show confidence values near to 100% in most cases as shown in Figure 3.

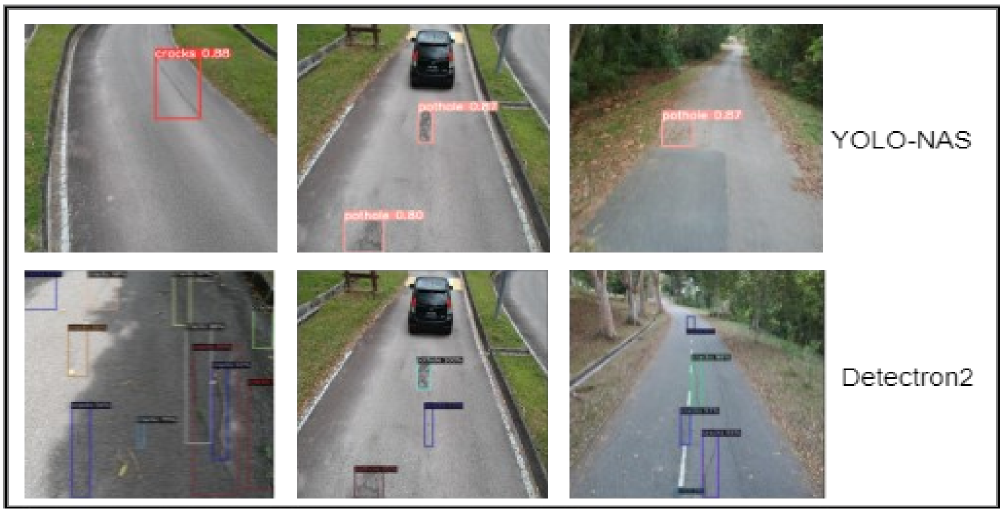

Figure 2. Inference YOLO-NAS and Detectron2 for input size 800x800 pixels.

The performance analysis highlights YOLO-NAS's substantial advancement in accuracy and Detectron2's exceptional performance in road defect detection using aerial imagery. Each algorithm possesses its unique strengths, making them valuable tools for different use cases and environments. Researchers and practitioners can leverage this analytical comparison to make informed decisions when selecting the most suitable algorithm for their specific object detection tasks.

5. Conclusions

In this research, the used models show significant accuracy on the road defects dataset collected using drones. From this research, we can clearly state that the algorithms perform better with aerial images, the reason is the suitable angles and deceptive of defects. Moreover, it shows a wide range of coverage of the affected area, which gives the accessibility of better detection. While training the algorithms achieved multiple perspectives of each class. Due to this consistency, the algorithms performed well in this dataset and showed a suitable accuracy for road defect detection.

Although, the two used models' architecture performed differently. For optimal performance, the YOLO-NAS architecture employs quantization-aware blocks and selective quantization. The design of the model includes adaptive quantization, which skips quantization in specific levels based on the trade-off between latency/throughput improvement and accuracy loss. On the other hand, Detectron2 (With faster RCNN) is a versatile object detection framework built on PyTorch that employs a two-stage approach that includes region proposal and refining stages. In future research, the detection model can be implemented into drones for real-time detection using a based station from which an operator can control.

Conflicts of interest

The authors declare no conflict of interest.

References

[1] Fan R, Bocus MJ, Zhu Y, Jiao J, Wang L, *et al.* Road crack detection using deep convolutional neural network and adaptive thresholding. In *2019 IEEE Intelligent Vehicles Symposium (IV),* 2019 Jun 9, pp. 474-479

[2] Shi Y, Cui L, Qi Z, Meng F, Chen Z. Automatic Road crack detection using random structured forests. *IEEE Trans Intell Transp Syst*. 2016 May 18;17(12):3434-45.

[3] WHO. United Nations Road Safety Collaboration. Available: https://www.who.int/groups/united-nations-road-safety-collaboration (accessed Jun. 01, 2023).

[4] Lesiak P. Drones in the inspection of railway engineering facilities. *Transp Overv Prz Komun*, 2021, pp. 1–10.

[5] Sari Y, Prakoso PB, Baskara AR. Road crack detection using support vector machine (SVM) and OTSU algorithm. In *2019 6th International Conference on Electric Vehicular Technology (ICEVT)*, 2019 Nov 18, pp. 349-354.

[6] Al-Yassin H, Mousa JI, Fadhel MA, Al-Shamma O, Alzubaidi L. Statistical accuracy analysis of different detecting algorithms for surveillance system in smart city. *Indonesian Journal of Electrical Engineering and Computer Science*. 2020;18(2):979-86.

[7] Terven J, Cordova-Esparza D. A comprehensive review of YOLO: From YOLOv1 and beyond. *arXiv* 2023. arXiv preprint arXiv:2304.00501.

[8] Safaei N, Smadi O, Safaei B, Masoud A. Efficient road crack detection based on an adaptive pixel-level segmentation algorithm. *Transp Res Record*. 2021 Sep;2675(9):370-81.

[9] Bang S, Park S, Kim H, Kim H. Encoder–decoder network for pixel-level road crack detection in black-box images. *Comput Aided Civil Infrastruct Eng*. 2019 Aug;34(8):713-27.

[10] Wen H, Huang C, Guo S. The application of convolutional neural networks (CNNs) to recognize defects in 3D-printed parts. *Materials*. 2021 Jan;14(10):2575.

[11] Kao SP, Chang YC, Wang FL. Combining the YOLOv4 deep learning model with UAV imagery processing technology in the extraction and quantization of cracks in bridges. *Sensors*. 2023 Feb 25;23(5):2572.

[12] Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS one*. 2015 Mar 4;10(3):e0118432.

[13] Redmon J, Farhadi A. Yolov3: An incremental improvement. *arXiv* preprint arXiv:1804.02767. 2018 Apr 8.

[14] Arya D, Maeda H, Ghosh SK, Toshniwal D, Omata H, Kashiyama T, Sekimoto Y. Global road damage detection: State-of-the-art solutions. In *2020 IEEE International Conference on Big Data (Big Data)*, 2020 Dec 10, pp. 5533-5539

[15] Pham V, Pham C, Dang T. Road damage detection and classification with detectron2 and faster r-cnn. In *2020 IEEE International Conference on Big Data (Big Data)*, 2020 Dec 10, pp. 5592-5601.