An image-based convolutional neural network system for road defects detection

Mohamed Anis Benallal, Mustapha Si Tayeb
Laboratoire SIMPA (Signal-Image-Parole), Département d’informatique, Faculté des Mathématiques et Informatique, Université des Sciences et de la Technologie d’Oran Mohamed Boudiaf (USTO-MB), Oran, Algeria

Article Info
Article history:
Received Feb 25, 2022
Revised Sep 27, 2022
Accepted Oct 27, 2022

Keywords:
Convolutional neural network
Defects detection
Image recognition
Road damages
You only look ones

ABSTRACT
An application of convolutional neural network (CNN) technique for road surface defects detection is presented in this paper. You only look ones (YOLO) algorithm showed its capabilities as an effective object detection technique in many previous works for different problems. Road damages detection and classification is one of the most challenging problems faced by public and private road management agencies. We present here results for a first attempt on applying YOLO to detect cracks and potholes, the most common defects encountered in surface roadways. Image database of the Brazilian highways were used to prepare input data, train the model and test it. Despite considering different types of cracks in one class and a less amount of potholes images, results show that the YOLO algorithm performs well with a global rate of 91% of defect detection. Output results analysis induce us to work on providing a local database for Algerian roadways with a large number of defect images/videos, as well as producing an automatic road-dedicated defects detector device.

This is an open access article under the CC BY-SA license.

Corresponding Author:
Mohamed Anis Benallal
Laboratoire SIMPA, Département d’informatique, Faculté des Mathématiques et Informatique, Université des Sciences et de la Technologie d’Oran Mohamed Boudiaf (USTO-MB)
BP 1505, El M’naouer, Oran, Algeria
Email: mohamed.benallal@univ-usto.dz

1. INTRODUCTION
Countries around the world are likely to encounter the current issues with infrastructure management and maintenance [1]. Traditionally, monitoring the road surface condition was done manually, but this is still the case in our country. There are many ways to build a road. However, the great majority of our roads are constructed of an asphalt concrete layer, followed by one or more layers of angular rock, known as the base course (also called tarmac). Despite being a cost-effective solution, as it provides smooth and durable surface, roads suffer from a set of problems. The most encountered ones are cracks and potholes. Sunlight, rain, snow, sleet, and freezing weather, in addition to cars and trucks traffic contribute to the slow process of altering the road surface.

Since roadways infrastructure is one of the highest value assets owned by a country, maintaining it has a huge prevalence and a direct effect on the economy and the prosperity of people and businesses in a country, that is, to extend the life of an asphalt surface, it needs to be repaired as soon as a defect appears. That means that detecting a damage earlier provides a way of sealing and maintaining a correct road surface quality, and its

Journal homepage: http://ijai.iaescore.com
importance has no longer to be proved.

Object and features detection took a huge place in computer vision these last decades and on of its most boost is the emergence of convolutional neural networks (CNN), mainly by providing a better performance in visualizing and classifying images [2]. They have led to a rapid development of different fields of studies (identification [3], [4], medical [5], [6], automotive [7], and [8]). In the field of object detection, the most recent cutting-edge algorithms are faster region based CNN (Faster R-CNN) [3], [9]–[12], single shot multibox detector (SSD) [11], [13], [14] and you only look once (YOLO) [10], [11], [15]. This last one, as a deep learning-based model, have been considered to be the most efficient approach to detecting objects and characteristics such as surface cracks [16], but also a strong and competitive approach for detecting potholes [11]. In computer vision, identifying objects in images includes detection and localization but also classification. The YOLO algorithm was selected due to its excellent accuracy and speed on achieving these tasks. Although other approaches have demonstrated comparable accuracy on diverse datasets, YOLO exhibits the highest speed in many object detection applications, such as detecting non-certified work on construction sites [17], object detection for disaster response and recovery [18], and soldering defect detection [19]. Along with the recent improvements of GPUs (graphics processing unit), an impressive improvement in accuracy and speed for real-time applications area is observed this last decade.

Automatic road surface monitoring is one of the most challenging problems to manage for road monitoring agencies. Several works in all over the world have been done so far. Techniques differ in terms of input data. In fact, there are three main ways which are investigated:

- Vibrations: as it can be sampled using a very affordable equipment, this technique seems to be very popular, however, being in contact with the damage is a must, for example falling in a pothole to have its signal [20]–[22].
- Laser/ultrasonic: the main goal is characterizing defects and modelling the shape and size of each one. The main disadvantage of these techniques is the fact that equipment is usually very expensive. The exceptional use of them remains rare [23], [24], in fact, research involving such equipment is usually hybridized with images techniques [25]–[27].
- Image analysis and exploitation: which balance between the two above, this is the most common nowadays.

Actually, image-based techniques remain to be the most studied way for detecting and monitoring road defects. The multiplicity of camera types (digital single-lens reflex (DSLR), parking, and smartphone) produces a gigantic source of data. Moreover, capturing images provides a real and visual idea of the entire scene (also analogically similar to the traditional way of monitoring roads). For the maintenance and monitoring of pavement surfaces, automated image analysis methods have been implemented over the years. Images can bring a lot of useful information. In general, cameras are placed in front top or in the rear of the vehicle.

As they are the most encountered, potholes and cracks are the most studied defects [16], [28]. Other works tend to work on detecting defects, but more importantly, classifying them as much as possible [1], [29]. In addition to these two main ways of working on images, other works such as [30] focus on identifying defects by reconstructing them in three-dimensional (3D) models. The most important advantage that this technique has today is the capacity of handling such data in real-time, especially by the increase of computer processing capacity in one hand, but also the use of new generation of algorithms such as deep learning in another hand.

Based on our review on the subject, YOLO applications were previously used in different scenarios. Li et al. [31], ground penetrating radar equipment was used to acquire concealed cracks in asphalt pavement. The use of YOLOv4 as a single class detector on these radar images showed the best overall performance of the whole YOLO series. On the one hand, its detection speed is very rapid, and it can still be investigated to a greater extent utilizing a GPU, thanks to its reduced inference time and higher FPS (frame per second). The higher level of confidence and robustness, on the other hand, indicates that it has a relatively reliable performance in precisely detecting concealed cracks. Zhang et al. [32], YOLOv3 was compared to an improved version (combined with MobileNets and convolutional block attention module) to tackle bridge surface cracks. Results showed that the modified architecture of the algorithm can perform at a higher FPS rate but slightly influence the precision of the model. Du et al. [33], a comparison of the YOLO series was performed against several pavement distress using a local road dataset. Results showed that the YOLO network can be used to forecast potential distress location and category. Moreover, comprehensive detection accuracy reached 73.64%, and the processing speed was just 70% SSD and 9 times faster than Faster R-CNN. This paper presents prelim-
inary results of an experimental application of the YOLO deep-learning technique on tackling the problem of detecting cracks and potholes based on road images, that is, more investigation should be provided by applying YOLO on other regions.

2. METHOD

This section presents the used algorithm and how it can be applied on our study. The details about the used database are also presented to explain how data are used in the algorithm. In fact, raw images of road surface had been prepared and sampled through different steps.

2.1. YOLO algorithm

First published in 2015 by a team of researchers from the University of Washington whose head of this project Joseph Redmon [34], YOLO is an automatic detection algorithm which unlike other methods uses a totally different approach. It represents a CNN for object detection which can be used in real time, and as its name indicates it looks at the processed image only once. In fact, this makes the method extremely faster at the testing phase compared to other approaches [35].

Due to the fact that YOLO is a supervised method, the training dataset must first be manually constructed and contain bounding box information for road defects. Parameters used to describe the bounding box information are: center X and Y, width and height. To perform this task, we used the Yololabel software (an image-labeling tools for object detection), more details are presented in section 2.2. Figure 1 explains the main four steps that YOLO performs to provide the bounding boxes according to the detected objects in the image. In this sequence of images, we can see that the image has been split into several squares (top-left) which generate several random anchor boxes (top-right).

![Figure 1. The YOLO flowchart algorithm](image)

The intersection over union (IoU) is then calculated to group together the anchor boxes that detect the same object (we can see this through the groups of colours in the bottom-right). YOLO uses IoU to provide a box that surrounds the whole targeting object. Basically, it divides the overlap area between the boxes by

An image-based convolutional neural network system for road defects detection (Mohamed Anis Benallal)
the union of those areas. Then finally, thanks to the non-max suppression algorithm included in the YOLO algorithm, we go from thousands of anchor boxes to only 5 in this example (one for cracks and four for potholes in the bottom-left).

2.2. Dataset preparation

To perform our application of the YOLO algorithm, different databases are available online. We choose to use the one provided in [36]. Note that there is no local available database at the time of this paper. However, we are working on providing a new one, especially dedicated to Algerian roadways for future works.

The used dataset includes pictures of defects on Brazilian asphalt roads. This dataset was developed using pictures provided by DNIT (National Department of Transport Infrastructure, Brazil. Available through the access to information law - protocol 50650.003556 / 2017-28). Capturing process was done on highways in the state of Espirito Santo, Rio Grande do Sul and the Federal District, between 2014 and 2017. To use this database, we went through each images and followed these steps:

- Check every one and select images, which contain a set of the studied defects, a total of 665 was obtained.
- Use the YoloLabel software to manually annotate the database. This step is the most time-consuming part of the database preparation. It mainly consists of selecting the appropriate class and drawing bounding boxes around the appropriate defect. The software then provides a text file for each image that contains information about boxes positions and dimensions.

The final step is to upload all files (database and needed YOLO files) into a Google drive repository to be easily used in the processing. Figure 2 presents examples of both classes, which are cracks in Figure 2(a) and potholes in Figure 2(b). Note that these examples represent the sampled defect as it is injected into the YOLO algorithm for the training process (raw images along with text files explained above). As shown, we can assume that the shape, the orientation, the dimension, and even the size are quite different from one defect to another in both classes.

![Figure 2](https://example.com/figure2.png)

Figure 2. Examples of the studied defects (YOLO algorithm inputs); (a) cracks example and (b) potholes examples

3. RESULTS AND DISCUSSION

All the experiment presented here have been done using Colab, the Google Colaboratory platform that offers the possibility of executing Python code directly in the browser. This platform, which is intended for machine learning training and research, enables us to train machine learning models directly in the cloud. As a first attempt, we would like to answer the question of how does the YOLO algorithm work on classifying the two main encountered road defects (cracks and potholes)? For this purpose, we prepared the database by selecting bounding boxes for these two classes. Note that we deliberately choose to put multiple types of cracks (horizontal/vertical/diagonal (an unconnected crack which generally takes a horizontal, vertical, or diagonal line across a pavement), crocodile (interconnected or interlocked fissures that look like a crocodile’s skin or a collection of tiny polygons), transverse (rupture across the pavement) [37]) to understand how it affects the learning process. Potholes are defined as a unique class described as irregularly shaped holes of various sizes in the pavement [37]
Working with YOLO implies defining a set of parameters to perform the learning process. In fact, it is implemented based on the darknet, an open source neural networks. Note that we used the v4 version of the YOLO algorithm [38]. Actually, this version remains the best choice to balance between accuracy and custom configuration. Having two classes, we need to set the configuration file according to, that is: line classes to 2, max\_batches to the minimum of 6,000, line steps to 80\% and 90\% of max\_batches and filters to 21. All of these parameters work in a network sized at 416 by 416. Due to Colab limitations (in its free version), a total of 5,700 iterations were performed. To test the performance of our model, and in addition to the selected images for training, we also selected 650 images from the used database in which 450 visually contain similar learned defects and the rest of 200 images which contain no learned defects (to test false positive cases) or no defects at all. Note that these 650 images have not been used in any step of the learning process. Indeed, we consider that doing so ensures a real and true test to the generalization capacity of the trained network.

To provide this step, we injected all these testing images to the model and then checked manually all the resulting images one by one. For every defect (cracks in all types and potholes), we went through counting the correctness of the detection. Note that we can have many defects in the same image. Figure 3 presents examples of a visual representation for the obtained results. As we can see in its sub-figures, boxes provided by the YOLO algorithm make it capable of: i) avoiding any detection where there is no defect (Figure 3(a)), ii) Detecting multiple defects of the same class (Figure 3(b) for cracks and Figure 3(c) for potholes), and detecting multiple defects of the different classes (Figure 3(d)).

![Image](a)

![Image](b)

![Image](c)

![Image](d)

Figure 3. Examples of the algorithm result detections; (a) no detection, (b) cracks detection, (c) potholes detection, and (d) both crack and pothole detection

Table 1 shows that the testing step, considering the testing dataset as different from the training one and with a comparable number of images, brings us a general detection rate of 91\% with approximately 90\% and 93\% for cracks and potholes respectively. Furthermore, we got an insignificant number of false positive cases for both classes. These results show that, even with taking into consideration multiple types of cracks in the same class, this gives an acceptable detection rate. However, this could be certainly improved with creating multiple classes which anyone correspond to a specific type of cracks. For potholes, even if results are
similar to those of cracks, we consider that we should have a larger number of images to be more consistent. Nevertheless, that lead us to work on other databases, but also create and publish a local one (since local roads significantly suffer from poor road conditions).

<table>
<thead>
<tr>
<th>Table 1. Model testing results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detected</td>
</tr>
<tr>
<td>Number %</td>
</tr>
<tr>
<td>Cracks</td>
</tr>
<tr>
<td>Potholes</td>
</tr>
<tr>
<td>In total</td>
</tr>
</tbody>
</table>

4. CONCLUSION

We presented in the paper an experimental test on detecting the two main classes of road defects (cracks and potholes). The observed lack of studies involving YOLO as a multiple and dedicated road-damages detector induce us to answer the question of, to what extent, the YOLO can be an algorithm to consider using in road surface damages detection and classification. An overall rate of 91% (90% for cracks and 93% for potholes respectively) was obtained using a Brazilian road images database. As a first attempt, we consider these results favourable to more investigations. Furthermore, this comforts us in using the YOLO algorithm as model-based technique to provide a more powerful and versatile system to detect better cracks and potholes, but also work on other kinds of defects. Therefore, this lead us to focus on two main aspects to improve our out-coming results: i) capture videos/images, in order to create a local database of Algerian road defects (work in progress at the time of this paper to provide more images of the studied problem but also to build an effective local detector), and ii) take into consideration the differences between defect features of the same class. In fact, our hypothesis about the observed results of cracks detection rate is mainly due to the choice of combining multiple cracks types, in terms of directions and shape. Overall, results are very motivating to explore the model optimisation in terms of input data, meta-parametrization, and model training.

REFERENCES


**BIOGRAPHIES OF AUTHORS**

**Mohamed Anis Benallal** is a lecturer at the Department of Computer Science at the University of Sciences and Technology of Oran Mohamed Boudiaf (USTO-MB), Algeria. He received his PhD in Computer Science at the University of Perpignan Via Domitia, France in 2016. His research interests are related to machine learning and artificial intelligence, remote sensing, data mining, estimation and interpolation of parameters, natural language processing, environment, and internet of things. He has published research papers international journals, co-author of conference proceedings as well as a chapter of book edited in Springer International Publishing. Besides, he is also one of the founder and a trainer in a non-profit foundation for student transversal trainings. He can be contacted at email: mohamed.benallal@univ-usto.dz.

**Mustapha Si Tayeb** is a lecturer at the Department of Computer Science at the University of Sciences and Technology of Oran Mohamed Boudiaf (USTO-MB), Algeria. He received his PhD in Computer Science at the same university in 2018. His research interests include issues related to artificial intelligence, remote sensing, image processing, data collection, analysis and classification, internet of things, and pattern extraction. He is author of a set of research studies published at national and international conference, and international journals. He can be contacted at email: mustapha.sitayeb@univ-usto.dz.