Abstract—Object detection is to identify objects from images. In autonomous driving systems, object detection serves as an intermediate module, which is used as the input of autonomous decisions for vehicles. That is, the accuracy of autonomous decisions relies on the object detection. The state-of-the-art object detection modules are designed based on Deep Neural Networks (DNNs). It is difficult to employ white-box testing on DNNs since the output of a single neuron is inexplicable. Existing work conducted metamorphic testing for object detection via image synthesis: the detected object in the original image should be detected in the new synthetic image. However, a synthetic image may not look real from humans’ perspective. Even the object detection module fails in detecting such synthetic image, the failure may not reflect the ability of object detection. In this paper, we propose an automatic approach to testing object detection via 3D reconstruction of vehicles in real photos. The 3D reconstruction is developed via vanishing point estimation in photos and heuristic based image insertion. Our approach adds new objects to blank spaces in photos to synthesize images. For example, a new vehicle can be added to a photo of a road and vehicles. In this approach, the output synthetic images are expected to be more natural-looking than randomly synthesizing images. The experiment is conducting on 500 driving photos from the Apollo autonomous driving dataset.

Index Terms—metamorphic testing, object detection system, vanishing point, image processing

I. INTRODUCTION

Autonomous driving systems are widely developed in automobile manufacturers and software companies. For example, Tesla [1] and Baidu Apollo [2] have released their research tools for third-party researchers and users. In an autonomous driving system, object detection is a key module that is used to identify objects from images, such as persons or vehicles on the road. Object detection is an intermediate module, whose output is the input of autonomous decision. According to the detected objects, the autonomous driving system can automatically decide its next action in driving, such as turning right or breaking. Thus, the accuracy of autonomous decisions highly relies on the object detection module.

It is necessary to test the object detection modules ahead of autonomous decisions to evaluate its reliability. Most of effective object detection tools are developed based on Deep Neural Networks (DNNs) [3]. As a kind of deep neural network (DNN), Convolution Neural Networks (CNNs) are used to solve the object detection problem, such as Fast-RCNN [4]. The performance of DNNs depends on the network architecture. The input and output of DNNs are not directly correlated. Executing an undesigned test may not reveal a
failure of DNNs [5]. This makes testing DNNs or DNN-based object detection difficult.

**Related work.** Object detection is widely used in image processing. Existing work have checked the ability of object detection [6]–[8]. Recent work by Wang and Su [9] conducted image synthesis via metamorphic testing [10] to examine failures of object detection. Their work added one object to a photo to synthesize a new image. This image is used to test object detection tools and check whether the metamorphic relation is violated.

**Motivation.** In autonomous driving systems, object detection is used to detect objects on the road. The input photo of object detection is taken from on-board cameras. That is, if a synthetic image is different from natural-looking images, the image may be not used to test the object detection in autonomous driving systems. We reimplemented the state-of-the-art approach in [9] and applied it to a driving photo from the Apollo dataset [11]. Then an image with a running man on the roadside tree is synthesized. Fig. 1 is a running example of our work, where Fig. 1(a) is the above synthetic image. This synthetic image can cause a failure of the yolov5 model, a tool of object detection [12]. However, this testing result is not reliable because the synthetic image does not exist in the real world. Instead, we aim to synthesize new images that are similar to photos from on-board cameras. Fig. 1(c) is an example of an synthetic image with an added vehicle via 3D reconstruction.

**Research problem.** This motivates our work: the goal is to design a new approach to generate synthetic images, which can be natural-looking and can trigger failures of object detection. As shown by Patki et al. [13], artificial data can be used to partially replace the real data. The research problem of this work is *how to generate synthetic natural-looking images for testing object detection in autonomous driving systems.*

**Contribution.** The major contribution of this work is to propose a new approach to testing object detection in autonomous driving systems via image synthesis based on 3D reconstruction.

III. EXPERIMENTAL SETUP

**Implementation.** The approach in this paper is implemented on the top of the one-stage instance segmentation tool YOLACT [15] implemented via Pytorch. The 3D reconstruction is conducted via vanishing point estimation [14].

**Setup.** Our testing tool is specially designed for autonomous driving systems. In the experiment, we chose the Apollo driving dataset [11]. We randomly selected 500 original images from the Apollo driving dataset and synthesized new images for each original one. The yolov5 model serves as the object detection system under test. We loaded a pre-trained yolov5 model from the PyTorch Hub and its pre-trained weight. We randomly select 500 original images from the Apollo space dataset.
<table>
<thead>
<tr>
<th>Category</th>
<th>Recognition</th>
<th>Classification</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Preliminary evaluation. We selected 25 pairs of images from the Apollo space dataset. In each pair, one image is used as the background; the other is used to extract objects. Then 25 synthetic images are generated based on the 3D reconstruction. In these 25 images, 11 natural-looking synthetic images are finally selected to evaluate the detection results. We found six detection failures in these images. We divided the detection failures into three categories: recognition failure, classification failure, and location failure. These failures are shown in Table I. We illustrated one of the detection failures in Fig. 1(d). After a vehicle is inserted to the background image, the yolov5 mistakenly identified a part of railings as another vehicle.

IV. CONCLUSIONS

This paper proposes an automatic approach to testing object detection via 3D reconstruction of vehicles in real photos. The original photo is reconstructed into 3D models via vanishing point estimation. Then a vehicle image is inserted to the photo as a synthetic image. Our work is to check whether the original detected objects can be detected in the synthetic images.

REFERENCES


