Abstract

Automated driving systems (ADSs) are complex entities comprising numerous components, and traditional testing methods often struggle to ensure their safety, primarily due to the diversity driving environments. Among the core components of ADSs, object detection using deep neural networks (DNNs) has shown remarkable effectiveness in recognizing surrounding vehicles, pedestrians, and obstacles. However, ensuring the safety of object detection in such systems requires more than just high detection performance—it requires the ability to verify whether the positions of detected objects comply with safety-relevant specifications that govern system responses. Unfortunately, current testing practices are often grounded in informal specifications, which lack the precision and rigor needed to validate object-detection behaviors, particularly in safety-critical contexts. These informal descriptions make it difficult to define expected behavior and judge whether a system has passed or failed a given test. To address this issue, this paper first propose the bounding box specification language (BBSL), a framework capable of mathematically articulating the specifications for object and event detection and response (OEDR) tasks in ADSs. BBSL enables stakeholders such as developers and testers to mathematically express requirements concerning object positions and their relationships in a highly interpretable and verifiable manner. Building on this language, we further propose a functional testing approach for object-detection modules in ADSs. This approach leverages BBSL to define the expected behaviors in a testable form and allows us to determine whether the system ' s object recognition results satisfy these function. Notably, our approach can identify safety-critical defects that may remain undetected by conventional testing methods that rely solely on performance metrics such as accuracy or Intersection over Union (IoU). Furthermore, our proposed approach can identify safety-critical defects that conventional tests, which focus solely on performance evaluation, might overlook. Furthermore, we propose two sets of test criteria. The first set reflects the diversity of object positions and sizes within an image, while the second set includes coverage metrics that determine whether the test cases cover all conditions outlined by the BBSL specifications. Overall, our contributions facilitate the implementation of functional testing for object-detection systems using DNNs, a challenge previously considered formidable.

Keywords— Automated driving, coverage, formal specification, object detection, testing