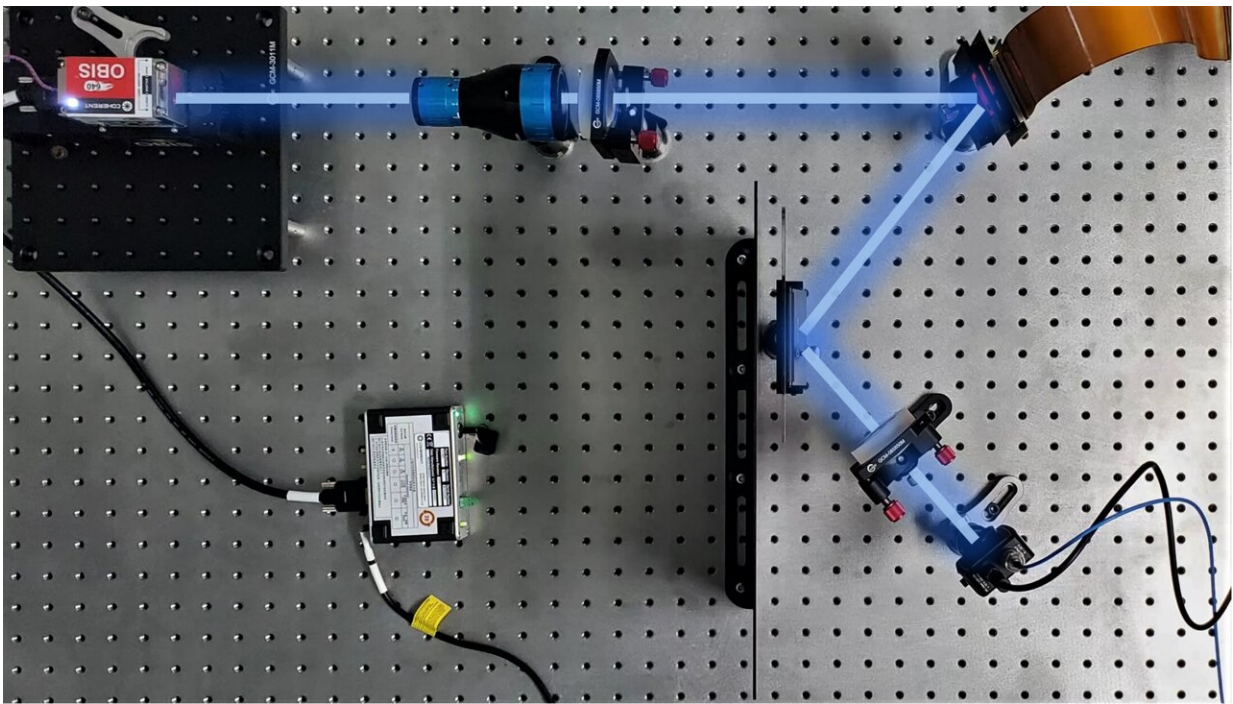# Researchers detect and classify multiple objects without images

May 3 2023

Researchers have developed a new high-speed way to detect the location, size and category of multiple objects without acquiring images or requiring complex scene reconstruction. The proof-of-concept setup for image-free single-pixel object detection (SPOD) is shown. Credit: Lintao Peng, Beijing Institute of Technology

Researchers have developed a new high-speed way to detect the location, size and category of multiple objects without acquiring images or

requiring complex scene reconstruction. Because the new approach greatly decreases the computing power necessary for object detection, it could be useful for identifying hazards while driving.

"Our technique is based on a single-pixel detector, which enables efficient and robust multi-object detection directly from a small number of 2D measurements," said research team leader Liheng Bian from the Beijing Institute of Technology in China. "This type of image-free sensing technology is expected to solve the problems of heavy communication load, high computing overhead and low perception rate of existing visual perception systems."

Today's image-free perception methods can only achieve classification, single object recognition or tracking. To accomplish all three at once, the researchers developed a technique known as image-free single-pixel object detection (SPOD). In *Optics Letters*, they report that SPOD can achieve an object detection accuracy of just over 80%.

The SPOD technique builds on the research group's previous accomplishments in developing imaging-free sensing technology as efficient scene perception technology. Their prior work includes image-free classification, segmentation and character recognition based on a single-pixel detector.

"For autonomous driving, SPOD could be used with lidar to help improve scene reconstruction speed and object detection accuracy," said Bian. "We believe that it has a high enough detection rate and accuracy for autonomous driving while also reducing the transmission bandwidth and computing resource requirements needed for object detection."

Detection without images

Automating advanced visual tasks—whether used to navigate a vehicle

or track a moving plane—usually require detailed [images](#) of a scene to extract the features necessary to identify an object. However, this requires either complex imaging hardware or complicated reconstruction algorithms, which leads to high computational cost, long running time and heavy data transmission load. For this reason, the traditional image first, perceive later approaches may not be best for object detection.
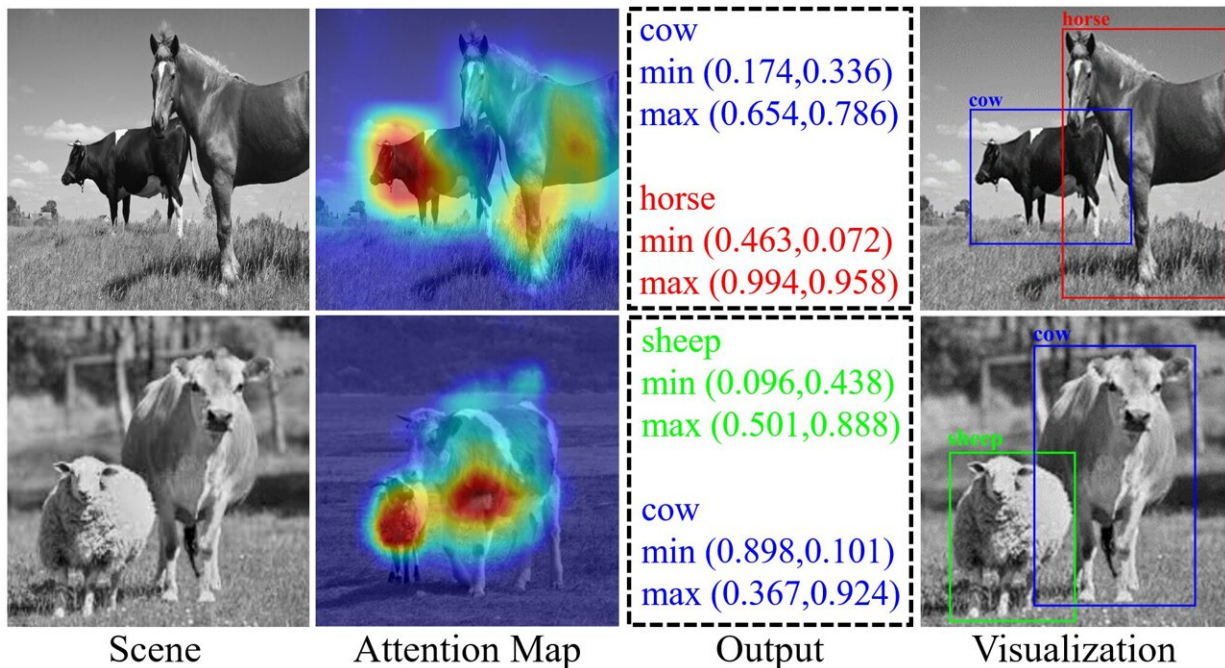
The researchers showed that SPOD achieves a detection accuracy of 82.41% mAP at a sampling rate of 5% with a refresh rate of 63 FPS. Visualization results are pictured. Credit: Lintao Peng, Beijing Institute of Technology

Image-free sensing methods based on single-pixel detectors can cut down on the computational power needed for object detection. Instead of employing a pixelated detector such as a CMOS or CCD, single-pixel imaging illuminates the scene with a sequence of structured light

patterns and then records the transmitted light intensity to acquire the spatial information of objects. This information is then used to computationally reconstruct the object or to calculate its properties.

For SPOD, the researchers used a small but optimized structured light pattern to quickly scan the entire scene and obtain 2D measurements. These measurements are fed into a deep learning model known as a transformer-based encoder to extract the high-dimensional meaningful features in the scene. These features are then fed into a multi-scale attention network-based decoder, which outputs the class, location and size information of all targets in the scene simultaneously.

"Compared to the full-size pattern used by other single-pixel detection methods, the small, optimized pattern produces better image-free sensing performance," said group member Lintao Peng. "Also, the multi-scale attention network in the SPOD decoder reinforces the network's attention to the target area in the scene. This allows more efficient extraction of scene features, enabling state-of-the art object detection performance."

Proof-of-concept demonstration

To experimentally demonstrate SPOD, the researchers built a proof-of-concept setup. Images randomly selected from the Pascal Voc 2012 test dataset were printed on film and used as target scenes. When a sampling rate of 5% was used, the average time to complete spatial light modulation and image-free object detection per scene with SPOD was just 0.016 seconds.

This is much faster than performing scene reconstruction first (0.05 seconds) and then object detection (0.018 seconds. SPOD showed an average detection accuracy of 82.2% for all the object classes included in the test dataset.

"Currently, SPOD cannot detect every possible object category because the existing object detection dataset used to train the model only contains 80 categories," said Peng. "However, when faced with a specific task, the pre-trained model can be fine-tuned to achieve image-free multi-object detection of new target classes for applications such as pedestrian, vehicle or boat detection."

Next, the researchers plan to extend the image-free perception technology to other kinds of detectors and computational acquisition systems to achieve reconstruction-free sensing technology.

Provided by Optica