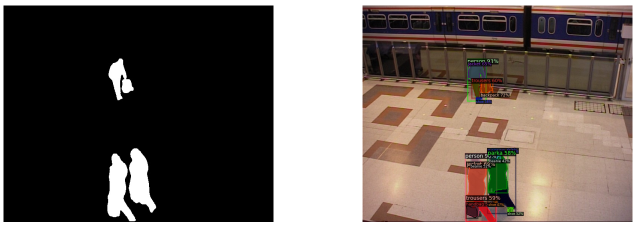
A. Instance-level foreground detection

Unlike previous methods, our method builds an instance-level background model. Therefore, ZBS can achieve instance-level foreground detection. Figure 1 shows the difference between binary foreground detection and instance-level foreground detection. Figure 1b shows that ZBS can detect moving foreground of different granularities, including person, backpack, shoe, beanie, *etc*., and can correctly classify stationary subway and crossbar as background.



(a) Binary foreground detection. (b) Instance-level foreground detection.

Figure 1. The binary and instance-level foreground detection of ZBS. Our method can detect the moving foreground of different granularities.

B. Abandoned object detection

Abandoned object detection in video surveillance is critical for ensuring public safety and is a crucial component of Intelligent Monitoring. This task presents a challenge, as the categories of abandoned objects are highly diverse and difficult to learn through traditional supervised training methods. Traditional background subtraction techniques often prove insufficient in addressing this issue. Our proposed method, however, offers a solution by incorporating a stronger semantic discernment and instance-level background model, resulting in effective detection of abandoned objects.

To adapt to new tasks, we have added a new rule that considers both motion information and the relationships between instances. If an object exhibits isolated, static behavior or moves independently after previously moving in sync with categories such as a person or car, the instance is deemed to be an abandoned object. This straightforward semantic rule has proven to be effective in diverse environments. We have conducted thorough experiments on the public datasets PETS2006 and ABODA, as well as a non-public traffic abandoned object detection dataset known as TADA.

B.1. PETS2006

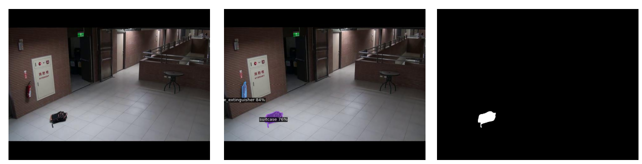
The PETS2006 dataset includes sequences from seven different scenes. Each sequence contains an abandonment event except for the third event. We evaluates all seven sequences and our method successfully detectes the abandoned objects for the entire PETS2006 dataset without any false alarms. As shown in Figure 2, the results from the PETS2006 dataset demonstrate the efficacy of our approach.

(a) (b) (c)

Figure 2. The detection results of the PETS2006. (a) is the original frame. (b) is the all-instance detection results. (c) is the abandoned object detection results of our method.

B.2. ABODA

The **AB**andoned **O**bjects **DA**taset (ABODA) [1] contains 11 sequences that present a range of challenging scenarios for abandoned object detection, including crowded scenes, changes in illumination, night-time detection, and both indoor and outdoor environments. Figure 3 displays the results from the $video1.avi$ sequence in ABODA.



(a) (b) (c)

Figure 3. The detection results of the ABODA. (a) is the original frame in the video. (b) is the all-instance detection results. (c) is the abandoned object detection results of our method.

B.3. TADA

The **T**raffic **A**bandoned object detection **DA**taset (TADA) is a household traffic abandoned object detection dataset that comprises 20 sequences, 14 of which contain traffic abandoned objects. These objects typically consist of various types of traffic litter, such as plastic bags, which have diverse appearances and shapes and are usually carried by the wind. This presents a formidable challenge for abandoned object detection. Figure 4 displays the results from the TADA dataset.

(a) (b) (c)

Figure 4. The detection results of the TADA. (a) is the original frame in the video. (b) is the all-instance detection results. (c) is the abandoned object detection results of our method.

References

[1] Kevin Lin, Shen-Chi Chen, Chu-Song Chen, Daw-Tung Lin, and Yi-Ping Hung. Abandoned object detection via temporal consistency modeling and back-tracing verification for visual surveillance. *IEEE Transactions on Information Forensics and Security*, 2015. 1