WHAT IS CLAIMED IS:

1. A system, comprising:

at least one storage device including a set of instructions; and

at least one processor in communication with the at least one storage device, wherein when executing the set of instructions, the at least one processor is configured to cause the system to:

obtain a video of a target scene that is captured by an image capture device;

determine a target image associated with the target scene based on the video;

determine a static region associated with the target scene in the target image; and

determine a result indicating whether a traffic accident happens in the static region by processing, based on a first model, the static region associated with the target scene in the target image, wherein to determine a static region associated with the target scene, the at least one processor is configured to cause the system to:

determine a plurality of regions associated with one or more static objects of the target scene in the target image by processing the target image based on a second model;.

determine at least two regions of the plurality of regions that satisfy a first condition, the first condition including that a distance between the at least two regions is smaller than or equal to a distance threshold;

generate a combination result by combining the at least two regions, the combination result including a combined bounding box of the at least two regions; and

designate a region corresponding to the combined bounding box as the static region.

2. The system of claim 1, wherein to determine a target image associated with the target scene based on the video, the at least one processor is configured to cause the system to:

obtain, from the video, a first image;

obtain, from the video, one or more second images captured prior to the first image; and

determine the target image by processing a dynamic region associated with the target scene in the first image based on the one or more second images.

3. The system of claim 2, wherein to determine the target image by processing a dynamic region associated with the target scene in the first image based on the one or more second images, the at least one processor is configured to cause the system to:

generate a fusion result by fusing the first image and the one or more second images; and

determine the target image based on the fusion result.

4. The system of claim 2 or claim 3, wherein a time interval between two adjacent images among the one or more second images and the first image relates to a type of the target scene.

5. The system of claim 1, wherein

the second model includes a detection model; and

the plurality of regions associated with the one or more static objects of the target scene in the target image includes one or more bounding boxes of the one or more static objects determined by the detection model.

6. The system of claim 1, wherein to determine the static region associated with the target scene in the target image, the at least one processor is further configured to cause the system to:

determine whether a distance between a region of a remaining portion of the plurality of regions and the combined bounding box is larger than a second distance threshold, wherein the remaining portion excludes the at least two regions constituting the combined bounding box; and

in response to determining that the distance is larger than the second distance threshold, exclude the region of the remaining portion from the static region.

7. The system of claim 1, wherein

the second model includes a segmentation model; and

the plurality of regions associated with the one or more static objects of the target scene in the target image includes one or more connected components of the one or more static objects determined by the segmentation model.

8. The system of claim 2, wherein to determine the target image by processing the dynamic region associated with the target scene in the first image, the at least one processor is configured to cause the system to:

generate a fusion result by fusing the first image and the one or more second images; and

blur the dynamic region in the target image based on the fusion result, such that the target image determined based on the fusion result includes the blurred dynamic region.

9. The system of any one of claims 5-7, wherein to determine a result indicating whether a traffic accident happens in the static region by processing, based on a first model, the static region associated with the target scene in the target image, the at least one processor is configured to cause the system to:

determine a confidence value associated with the traffic accident by processing, based on the first model, the static region associated with the target scene in the target image; and

in response to determining that the confidence value satisfies a second condition, determine that the traffic accident happens in the static region.

10. The system of claim 9, wherein to determine a result indicating whether a traffic accident happens in the static region by processing, based on a first model, the static region associated with the target scene in the target image, the at least one processor is configured to cause the system to:

in response to determining that the confidence value fails to satisfy the second condition, determine whether there are at least two regions of the plurality of regions associated with the one or more static objects in the static region that overlap; and

in response to determining that there are the at least two regions of the plurality of regions associated with the one or more static objects in the static region that overlap, modify the confidence value associated with the traffic accident.

11. The system of claim 10, wherein

the plurality of regions associated with the one or more static objects includes one or more two-dimensional regions; and

to modify the confidence value associated with the traffic accident, the at least one processor is configured to cause the system to:

determine one or more three-dimensional regions associated with the one or more static objects in the static region by performing a three-dimensional transformation on the one or more two-dimensional regions associated with the one or more static objects based on one or more coordinates associated with the one or more two-dimensional regions; and

modify the confidence value associated with the traffic accident based on an overlap between the one or more three-dimensional regions associated with the one or more static objects in the static region.

12. The system of claim 11, wherein to modify the confidence value associated with the traffic accident based on the overlap between the one or more three-dimensional regions, the at least one processor is configured to cause the system to:

determine a ratio between a volume of the overlap between the one or more three-dimensional regions and a minimum volume of the one or more three-dimensional regions; and

modify the confidence value based on the ratio.

13. The system of claim 11, wherein to perform the three-dimensional transformation on the one or more two-dimensional regions based on the one or more coordinates, the at least one processor is configured to cause the system to:

transform the one or more coordinates from a two-dimensional coordinate system into a three-dimensional coordinate system using a projection transformation matrix, an intrinsic parameter of the image capture device, a rotation transformation vector, and a translation transformation vector.

14. A method implemented on a computing device having at least one processor, at least one storage medium, and a communication platform connected to a network, the method comprising:

obtaining a video of a target scene that is captured by an image capture device;

determining a target image associated with the target scene based on the video;

determining a static region associated with the target scene in the target image; and

determining a result indicating whether a traffic accident happens in the static region by processing, based on a first model, the static region associated with the target scene in the target image, wherein the determining a static region associated with the target scene includes:

determining a plurality of regions associated with one or more static objects of the target scene in the target image by processing the target image based on a second model;

determining at least two regions of the plurality of regions that satisfy a first condition, the first condition including that a distance between the at least two regions is smaller than or equal to a distance threshold;

generating a combination result by combining the at least two regions, the combination result including a combined bounding box of the at least two regions; and

designating a region corresponding to the combined bounding box as the static region.

15. The method of claim 14, wherein

the second model includes a detection model; and

the plurality of regions associated with the one or more static objects of the target scene in the target image includes one or more bounding boxes of the one or more static objects determined by the detection model.

16. The method of claim 14, wherein

the second model includes a segmentation model; and

the plurality of regions associated with the one or more static objects of the target scene in the target image includes one or more connected components of the one or more static objects determined by the segmentation model.

17. The method of any one of claims 14-16, wherein the determining a result indicating whether a traffic accident happens in the static region by processing, based on a first model, the static region associated with the target scene in the target image includes:

determining a confidence value associated with the traffic accident by processing, based on the first model, the static region associated with the target scene in the target image; and

in response to determining that the confidence value satisfies a second condition, determining that the traffic accident happens in the static region.

18. The method of claim 17, wherein the determining a result indicating whether a traffic accident happens in the static region by processing, based on a first model, the static region associated with the target scene in the target image includes:

in response to determining that the confidence value fails to satisfy the second condition, determining whether there are at least two regions of the plurality of regions associated with the one or more static objects in the static region that overlap; and

in response to determining that there are the at least two regions of the plurality of regions associated with the one or more static objects in the static region that overlap, modifying the confidence value associated with the traffic accident.

19. The method of claim 18, wherein

the plurality of regions associated with the one or more static objects includes one or more two-dimensional regions; and

the modifying the confidence value associated with the traffic accident includes:

determining one or more three-dimensional regions associated with the one or more static objects in the static region by performing a three-dimensional transformation on the one or more two-dimensional regions associated with the one or more static objects based on one or more coordinates associated with the one or more two-dimensional regions; and

modifying the confidence value associated with the traffic accident based on an overlap between the one or more three-dimensional regions associated with the one or more static objects in the static region.

20. A non-transitory computer readable medium, comprising executable instructions that, when executed by at least one processor, directs the at least one processor to perform a method, the method comprising:

obtaining a video of a target scene that is captured by an image capture device;

determining a target image associated with the target scene based on the video;

determining a static region associated with the target scene in the target image; and

determining a result indicating whether a traffic accident happens in the static region by processing, based on a first model, the static region associated with the target scene in the target image , wherein the determining a static region associated with the target scene includes:

determining a plurality of regions associated with one or more static objects of the target scene in the target image by processing the target image based on a second model;

determining at least two regions of the plurality of regions that satisfy a first condition, the first condition including that a distance between the at least two regions is smaller than or equal to a distance threshold;

generating a combination result by combining the at least two regions, the combination result including a combined bounding box of the at least two regions; and

designating a region corresponding to the combined bounding box as the static region.