CLAIMS

1. A method for target detection, comprising:

obtaining feature data of an input image;

determining multiple candidate bounding boxes of the input image according to the feature data;

obtaining a foreground segmentation result of the input image according to the feature data, the foreground segmentation result comprising indication information for indicating whether each of multiple pixels of the input image belongs to a foreground; and

obtaining a target detection result of the input image according to the multiple candidate bounding boxes and the foreground segmentation result.

2. The method of claim 1, wherein obtaining the target detection result of the input image according to the multiple candidate bounding boxes and the foreground segmentation result comprises:

selecting at least one target bounding box from the multiple candidate bounding boxes according to an overlapping area between each candidate bounding box in the multiple candidate bounding boxes and a foreground image region corresponding to the foreground segmentation result; and

obtaining the target detection result of the input image based on the at least one target bounding box.

3. The method of claim 2, wherein selecting the at least one target bounding box from the multiple candidate bounding boxes according to the overlapping area between each candidate bounding box in the multiple candidate bounding boxes and the foreground image region

corresponding to the foreground segmentation result comprises:

for each candidate bounding box in the multiple candidate bounding boxes, if a ratio of an overlapping area between the candidate bounding box and the foreground image region to an area of the candidate bounding box is greater than a first threshold, taking the candidate bounding box as the target bounding box.

4. The method of claim 2 or 3, wherein the at least one target bounding box comprises a first bounding box and a second bounding box, and obtaining the target detection result of the input image based on the at least one target bounding box comprises:

determining an overlapping parameter between the first bounding box and the second bounding box based on an angle between the first bounding box and the second bounding box; and

determining a target object position corresponding to the first bounding box and the second bounding box based on the overlapping parameter between the first bounding box and the second bounding box.

5. The method of claim 4, wherein determining the overlapping parameter between the first bounding box and the second bounding box based on the angle between the first bounding box and the second bounding box comprises:

obtaining an angle factor according to the angle between the first bounding box and the second bounding box; and

obtaining the overlapping parameter according to an Intersection over Union (IoU) between the first bounding box and the second bounding box and to the angle factor.

6. The method of claim 5, wherein the overlapping parameter between the first bounding box and the second bounding box is a product of the IoU and the angle factor; and the angle factor

increases with an increase of the angle between the first bounding box and the second bounding box,

or, wherein in a case where the IoU keeps fixed, the overlapping parameter between the first bounding box and the second bounding box increases with the increase of the angle between the first bounding box and the second bounding box.

7. The method of any one of claims 4 to 6, wherein determining the target object position corresponding to the first bounding box and the second bounding box based on the overlapping parameter between the first bounding box and the second bounding box comprises:

in a case where the overlapping parameter between the first bounding box and the second bounding box is greater than a second threshold, taking one of the first bounding box and the second bounding box as the target object position.

8. The method of claim 7, wherein taking the one of the first bounding box and the second bounding box as the target object position comprises:

determining an overlapping parameter between the first bounding box and the foreground image region corresponding to the foreground segmentation result, and determining an overlapping parameter between the second bounding box and the foreground image region; and

taking one of the first bounding box and the second bounding box, of which the overlapping parameter with the foreground image region is larger than that of another, as the target object position.

9. The method of any one of claims 4 to 8, wherein determining the target object position corresponding to the first bounding box and the second bounding box based on the

overlapping parameter between the first bounding box and the second bounding box comprises:

in a case where the overlapping parameter between the first bounding box and the second bounding box is smaller than or equal to the second threshold, taking each of the first bounding box and the second bounding box as a target object position.

10. The method of any one of claims 1 to 9, wherein a length-width ratio of a to-be-detected target object in the input image is greater than a specific value.

11. A training method for a target detection network, wherein the target detection network comprises a feature extraction network, a target prediction network and a foreground segmentation network, and the method comprises:

performing feature extraction processing on a sample image through the feature extraction network to obtain feature data of the sample image;

obtaining, according to the feature data, multiple sample candidate bounding boxes through the target prediction network;

obtaining, according to the feature data, a sample foreground segmentation result of the sample image through the foreground segmentation network, the sample foreground segmentation result comprising indication information for indicating whether each of multiple pixels of the sample image belongs to a foreground;

determining a network loss value according to the multiple sample candidate bounding boxes, the sample foreground segmentation result and labeling information of the sample image; and

adjusting a network parameter of the target detection network based on the network loss value.

12. The method of claim 11, wherein the labeling information comprises at least one ground-truth bounding box of at least one target object comprised in the sample image, and determining the network loss value according to the multiple sample candidate bounding boxes, the sample foreground segmentation result and the labeling information of the sample image comprises:

for each candidate bounding box in the multiple candidate bounding boxes, determining an Intersection over Union (IoU) between the candidate bounding box and each of at least one ground-truth target bounding box labeled in the sample image; and

determining a first network loss value according to the determined IoU for each of the multiple candidate bounding boxes.

13. The method of claim 12, wherein the IoU between the candidate bounding box and the ground-truth target bounding box is obtained based on a circumcircle comprising the candidate bounding box and the ground-truth target bounding box.

14. The method of any one of claims 11 to 13, wherein in a process of determining the network loss value, a weight corresponding to a width of the candidate bounding box is higher than a weight corresponding to a length of the candidate bounding box.

15. The method of any one of claims 11 to 14, wherein obtaining, according to the feature data, the sample foreground segmentation result of the sample image through the foreground segmentation network comprises:

performing upsampling processing on the feature data, so as to make a size of the processed feature data to be same as that of the sample image; and

performing pixel segmentation based on the processed feature data to obtain the sample foreground segmentation result of the sample image.

16. The method of any one of claims 11 to 15, wherein a length-width ratio of a target object comprised in the sample image is higher than a set value.

17. A device for target detection, comprising a memory and a processor, wherein the memory is configured to store computer instructions capable of running on the processor, and the processor is configured to execute the computer instructions to implement the method of any one of claims 1 to 10.

18. A training device for a target detection network, comprising a memory and a processor, wherein the memory is configured to store computer instructions capable of running on the processor, and the processor is configured to execute the computer instructions to implement the method of any one of claims 11 to 16.

19. A non-volatile computer-readable storage medium, storing computer programs thereon, wherein the computer programs are executed by a processor to cause the processor to implement the method of any one of claims 1 to 10, or to implement the method of any one of claims 11 to 16.

20. A computer program product, which is stored in computer-readable storage medium, comprising computer readable codes, wherein when the computer readable codes are executed on a device, a processor of the device is configured to implement the method of any one of claims 1 to 10, or to implement the method of any one of claims 11 to 16.