

## CLAIMS

---

1. A target detecting method, comprising:

separately extracting, by means of a neural network, characteristics of a template frame and a detection frame, wherein the template frame is a detection box image of a target object, and the template frame is smaller than the detection frame in image size;

obtaining a classification weight and a regression weight of a local region detector based on the characteristic of the template frame;

inputting the characteristic of the detection frame into the local region detector to obtain classification results and regression results of multiple alternative boxes output by the local region detector; and

obtaining a detection box for the target object in the detection frame according to the classification results and regression results of the multiple alternative boxes output by the local region detector.

2. The method according to claim 1, further comprising:

extracting, by means of the neural network, characteristics of at least one other detection frame which is after the detection frame in timing in a video sequence;

sequentially inputting the characteristics of the at least one other detection frame into the local region detector to sequentially obtain multiple alternative boxes in the at least one other detection frame and classification results and regression results of the alternative boxes output by the local region detector; and

sequentially obtaining a detection box for the target object in the at least one other detection frame according to the classification results and regression results of the multiple alternative boxes of the at least one other detection frame.

3. The method according to claim 1 or 2, wherein separately extracting, by means of the neural network, characteristics of the template frame and the detection frame comprises:

separately extracting, by means of the same neural network, the characteristics of the template frame and the detection frame; or

respectively extracting, by means of different neural networks of the same structure,

the characteristics of the template frame and the detection frame.

4. The method according to any one of claims 1 to 3, wherein the template frame is a frame which is before the detection frame in detection timing in the video sequence and determined by the detection box for the target object.

5. The method according to any one of claims 1 to 4, wherein the detection frame is a current frame on which target object detection needs to be performed, or a regional image which may comprise the target object in the current frame.

6. The method according to claim 5, wherein when the detection frame is the regional image which may comprise the target object in the current frame on which the target object detection needs to be performed, the method further comprises:

taking a center point of the template frame as the center point, capturing the regional image with a length and/or width correspondingly greater than the image length and/or width of the template frame in the current frame as the detection frame.

7. The method according to any one of claims 1 to 6, wherein obtaining the classification weight of the local region detector based on the characteristic of the template frame comprises:

performing a convolution operation on the characteristic of the template frame by means of a first convolutional layer, wherein a first characteristic obtained by the convolution operation is used as the classification weight of the local region detector.

8. The method according to any one of claims 1 to 7, wherein obtaining the regression weight of the local region detector based on the characteristic of the template frame comprises:

performing the convolution operation on the characteristic of the template frame by means of a second convolutional layer, wherein a second characteristic obtained by the convolution operation is used as the regression weight of the local region detector.

9. The method according to any one of claims 1 to 8, wherein inputting the characteristic of the detection frame into the local region detector to obtain classification results and regression results of multiple alternative boxes output by the local region detector comprises:

using the classification weight to perform the convolution operation on the characteristic of the detection frame to obtain the classification results of the multiple alternative boxes; and using

the regression weight to perform the convolution operation on the characteristic of the detection frame to obtain the regression results of the multiple alternative boxes.

10. The method according to claim 9, wherein after the characteristic of the detection frame is extracted, the method further comprises: performing the convolution operation on the characteristic of the detection frame by means of a third convolutional layer to obtain a third characteristic, the number of channels of the third characteristic being identical to that of the characteristic of the detection frame; and

using the classification weight to perform the convolution operation on the characteristic of the detection frame to obtain the classification results of the multiple alternative boxes comprises: using the classification weight to perform the convolution operation on the third characteristic to obtain the classification results of the multiple alternative boxes.

11. The method according to claim 9 or 10, wherein after the characteristic of the template frame is extracted, the method further comprises: performing the convolution operation on the characteristic of the detection frame by means of a fourth convolutional layer to obtain a fourth characteristic, the number of channels of the fourth characteristic being identical to that of the characteristic of the detection frame; and

using the regression weight to perform the convolution operation on the characteristic of the detection frame to obtain the regression results of the multiple alternative boxes comprises: using the regression weight to perform the convolution operation on the fourth characteristic to obtain the regression results of the multiple alternative boxes.

12. The method according to any one of claims 1 to 11, wherein obtaining the detection box for the target object in the detection frame according to the classification results and regression results of the multiple alternative boxes output by the local region detector comprises:

selecting one alternative box from the multiple alternative boxes according to the classification results and the regression results, and performing regression on the selected alternative box according to an offset of the selected alternative box to obtain the detection box for the target object in the detection frame.

13. The method according to claim 12, wherein selecting one alternative box from the multiple alternative boxes according to the classification results and the regression results comprises:

selecting one alternative box from the multiple alternative boxes according to weight coefficients of the classification results and the regression results.

14. The method according to claim 12, wherein after the regression results are obtained, the method further comprises: adjusting the classification results according to the regression results; and

selecting one alternative box from the multiple alternative boxes according to the classification results and the regression results comprises: selecting one alternative box from the multiple alternative boxes according to the adjusted classification results.

15. A target detecting network training method, comprising:

separately extracting, by means of a neural network, characteristics of a template frame and a detection frame, wherein the template frame is a detection box image of a target object, and the template frame is smaller than the detection frame in image size;

performing a convolution operation on the characteristic of the template frame by means of a first convolutional layer, wherein a first characteristic obtained by the convolution operation is used as a classification weight of a local region detector; and performing the convolution operation on the characteristic of the template frame by means of a second convolutional layer, wherein a second characteristic obtained by the convolution operation is used as a regression weight of the local region detector;

inputting the characteristic of the detection frame into the local region detector to obtain classification results and regression results of multiple alternative boxes output by the local region detector;

obtaining a detection box for the target object in the detection frame according to the classification results and regression results of the multiple alternative boxes output by the local region detector; and

taking the obtained detection box for the target object in the detection frame as a predictive

detection box, training the neural network, the first convolutional layer, and the second convolutional layer based on label information of the detection frame and the predictive detection box.

16. A target detecting apparatus, comprising:

a neural network, configured to separately extract characteristics of a template frame and a detection frame, wherein the template frame is a detection box image of a target object, and the template frame is smaller than the detection frame in image size;

a first convolutional layer, configured to increase channels of the characteristic of the template frame to obtain a first characteristic as a classification weight of a local region detector;

a second convolutional layer, configured to increase channels of the characteristic of the template frame to obtain a second characteristic as a regression weight of the local region detector;

the local region detector, configured to output classification results and regression results of multiple alternative boxes according to the characteristic of detection frame; and

an obtaining unit, configured to obtain a detection box for the target object in the detection frame according to the classification results and regression results of the multiple alternative boxes output by the local region detector.

17. An electronic device, comprising the target detecting apparatus according claim 16.

18. An electronic device, comprising:

a memory, configured to store executable instructions; and

a processor, configured to communicate with the memory to execute the executable instructions so as to complete the operations of the method according to any one of claims 1 to 15.

19. A computer storage medium, configured to store computer-readable instructions, wherein when the instructions are executed, the operations of the method according to any one of claims 1 to 15 are implemented.

20. A computer program, wherein the computer program comprises computer readable

instructions, when the computer readable instructions run in an electronic apparatus, a processor of the electronic apparatus executes the method according to any one of claims 1 to 15.