CLAIMS

1. A method for recognizing stacked objects, comprising:

obtaining a to-be-recognized image, wherein the to-be-recognized image comprises an object sequence formed by stacking at least one object;

performing edge detection on the object sequence on the basis of the to-be-recognized image to determine an edge segmentation image of the object sequence, wherein the edge segmentation image comprises edge information of each object forming the object sequence; and

recognizing the category of each object in the object sequence on the basis of the to-be-recognized image and the edge segmentation image;

wherein the edge segmentation image comprises a mask image representing the edge information of each object;

wherein the edge segmentation image is a binarized mask image, pixels of a first pixel value in the edge segmentation image correspond to pixels of the edge of each object in the to-be-recognized image, and pixels of a second pixel value in the edge segmentation image correspond to pixels of the non-edge part of each object in the to-be-recognized image;

wherein the edge segmentation iamge is output from an edge segmentation model into which the to-be-recognized image is input.

2. The method of claim 1, wherein the recognizing the category of each object in the object sequence on the basis of the to-be-recognized image and the edge segmentation image comprises:

segmenting each object in the object sequence of the to-be-recognized image on the basis of the edge segmentation image to obtain a sub-image corresponding to each object; and

performing category recognition on each sub-image to obtain the category of each object.

3. The method of claim 2, wherein the segmenting each object in the object sequence of the to-be-recognized image on the basis of the edge segmentation image to obtain a sub-image corresponding to each object comprises:

determining first position information of each object in the to-be-recognized image on the basis of the edge information of each object forming the object sequence; and

segmenting each object in the to-be-recognized image on the basis of the first position information to obtain each sub-image.

4. The method of claim 2 or 3, wherein the performing category recognition on each sub-image to obtain the category of each object comprises:

performing category recognition on each sub-image to obtain at least two categories and at least two confidences having one-to-one correspondence to the at least two categories; and

in a case where a difference between a highest confidence and a second highest confidence in the at least two confidences is greater than a threshold, determining the category corresponding to the highest confidence as the category of the object corresponding to the sub-image.

5. The method of claim 4, wherein the method further comprises at least one of:

in a case where the difference is less than or equal to the threshold and the category corresponding to the highest confidence is the same as the categories of sub-images corresponding to two second objects adjacent to a first object corresponding to the sub-image in the object sequence, determining the category corresponding to the second highest confidence as the category of the first object corresponding to the sub-image;

in a case where the difference is less than or equal to the threshold and the category corresponding to the second highest confidence is the same as the categories of sub-images corresponding to two second objects adjacent to a first object corresponding to the sub-image

in the object sequence, determining the category corresponding to the highest confidence as the category of the first object corresponding to the sub-image; or

in a case where the difference is less than or equal to the threshold and the category corresponding to the second highest confidence is different from the categories of sub-images corresponding to two second objects adjacent to a first object corresponding to the sub-image in the object sequence, determining the category corresponding to the highest confidence as the category of the first object corresponding to the sub-image.

6. The method of any one of claims 1 to 5, wherein the edge segmentation image has the same size as the to-be-recognized image.

7. The method of any one of claims 2 to 5, wherein the performing edge detection on the object sequence on the basis of the to-be-recognized image to determine an edge segmentation image of the object sequence comprises:

inputting the to-be-recognized image into a trained edge detection model to obtain an edge detection result for each object in the object sequence, wherein the edge detection model is obtained by training based on a sequence object image comprising object edge annotation information; and

generating the edge segmentation image of the object sequence according to the edge detection result.

8. The method of claim 7, wherein the performing category recognition on each sub-image to obtain the category of each object comprises:

inputting each sub-image into a trained object classification model to obtain the category of each corresponding object;

wherein the object classification model is obtained by training based on single-object images, and the single-object images are obtained after segmenting the sequence object image according to the edge detection result for each object.

9. The method of any one of claims 1 to 8, wherein the object has a value attribute corresponding to the category; and the method further comprises:

determining a total value of objects in the object sequence on the basis of the category of each object and the corresponding value attribute.

10. An apparatus for recognizing stacked objects, comprising:

an obtaining unit, configured to obtain a to-be-recognized image, wherein the to-be-recognized image comprises an object sequence formed by stacking at least one object;

a determination unit, configured to perform edge detection on the object sequence on the basis of the to-be-recognized image to determine an edge segmentation image of the object sequence, wherein the edge segmentation image comprises edge information of each object forming the object sequence; and

a recognition unit, configured to recognize the category of each object in the object sequence on the basis of the to-be-recognized image and the edge segmentation image;

wherein the edge segmentation image comprises a mask image representing the edge information of each object;

wherein the edge segmentation image is a binarized mask image, pixels of a first pixel value in the edge segmentation image correspond to pixels of the edge of each object in the to-be-recognized image, and pixels of a second pixel value in the edge segmentation image correspond to pixels of the non-edge part of each object in the to-be-recognized image;

wherein the edge segmentation iamge is output from an edge segmentation model into which the to-be-recognized image is input.

11. A device for recognizing stacked objects, comprising a memory and a processor, wherein the memory stores a computer program capable of running on the processor, and when executing the computer program, the processor is configured to:

obtain a to-be-recognized image, wherein the to-be-recognized image comprises an object sequence formed by stacking at least one object;

perform edge detection on the object sequence on the basis of the to-be-recognized image to determine an edge segmentation image of the object sequence, wherein the edge segmentation image comprises edge information of each object forming the object sequence; and

recognize the category of each object in the object sequence on the basis of the to-be-recognized image and the edge segmentation image.;

wherein the edge segmentation image comprises a mask image representing the edge information of each object;

wherein the edge segmentation image is a binarized mask image, pixels of a first pixel value in the edge segmentation image correspond to pixels of the edge of each object in the to-be-recognized image, and pixels of a second pixel value in the edge segmentation image correspond to pixels of the non-edge part of each object in the to-be-recognized image;

wherein the edge segmentation iamge is output from an edge segmentation model into which the to-be-recognized image is input.

12. The device of claim 11, wherein when recognizing the category of each object in the object sequence on the basis of the to-be-recognized image and the edge segmentation image, the processor is configured to:

segment each object in the object sequence of the to-be-recognized image on the basis of the edge segmentation image to obtain a sub-image corresponding to each object; and

perform category recognition on each sub-image to obtain the category of each object.

13. The device of claim 12, wherein when segmenting each object in the object sequence of the to-be-recognized image on the basis of the edge segmentation image to obtain a sub-image corresponding to each object, the processor is configured to:

determine first position information of each object in the to-be-recognized image on the basis of the edge information of each object forming the object sequence; and

segment each object in the to-be-recognized image on the basis of the first position information to obtain each sub-image.

14. The device of claim 12 or13, wherein when performing category recognition on each sub-image to obtain the category of each object, the processor is configured to:

perform category recognition on each sub-image to obtain at least two categories and at least two confidences having one-to-one correspondence to the at least two categories; and

in a case where a difference between a highest confidence and a second highest confidence in the at least two confidences is greater than a threshold, determine the category corresponding to the highest confidence as the category of the object corresponding to the sub-image.

15. The device of claim 14, wherein the processor is further configured to perform at least one of following operations:

in a case where the difference is less than or equal to the threshold and the category corresponding to the highest confidence is the same as the categories of sub-images corresponding to two second objects adjacent to a first object corresponding to the sub-image in the object sequence, determining the category corresponding to the second highest confidence as the category of the first object corresponding to the sub-image;

in a case where the difference is less than or equal to the threshold and the category corresponding to the second highest confidence is the same as the categories of sub-images corresponding to two second objects adjacent to a first object corresponding to the sub-image in the object sequence, determining the category corresponding to the highest confidence as the category of the first object corresponding to the sub-image; or

in a case where the difference is less than or equal to the threshold and the category corresponding to the second highest confidence is different from the categories of sub-images

corresponding to two second objects adjacent to a first object corresponding to the sub-image in the object sequence, determining the category corresponding to the highest confidence as the category of the first object corresponding to the sub-image.

16. The device of any one of claims 11 to 15, wherein the edge segmentation image has the same size as the to-be-recognized image.

17. A computer storage medium, wherein the computer storage medium stores one or more programs, and when executed by one or more processors, the one or more programs are configured to:

obtain a to-be-recognized image, wherein the to-be-recognized image comprises an object sequence formed by stacking at least one object;

perform edge detection on the object sequence on the basis of the to-be-recognized image to determine an edge segmentation image of the object sequence, wherein the edge segmentation image comprises edge information of each object forming the object sequence; and

recognize the category of each object in the object sequence on the basis of the to-be-recognized image and the edge segmentation image;

wherein the edge segmentation image comprises a mask image representing the edge information of each object;

wherein the edge segmentation image is a binarized mask image, pixels of a first pixel value in the edge segmentation image correspond to pixels of the edge of each object in the to-be-recognized image, and pixels of a second pixel value in the edge segmentation image correspond to pixels of the non-edge part of each object in the to-be-recognized image;

wherein the edge segmentation iamge is output from an edge segmentation model into which the to-be-recognized image is input.

18. A computer program, comprising computer instructions executable by an electronic device, wherein when executed by a processor in the electronic device, the computer instructions are configured to:

obtain a to-be-recognized image, wherein the to-be-recognized image comprises an object sequence formed by stacking at least one object;

perform edge detection on the object sequence on the basis of the to-be-recognized image to determine an edge segmentation image of the object sequence, wherein the edge segmentation image comprises edge information of each object forming the object sequence; and

recognize the category of each object in the object sequence on the basis of the to-be-recognized image and the edge segmentation image;

wherein the edge segmentation image comprises a mask image representing the edge information of each object;

wherein the edge segmentation image is a binarized mask image, pixels of a first pixel value in the edge segmentation image correspond to pixels of the edge of each object in the to-be-recognized image, and pixels of a second pixel value in the edge segmentation image correspond to pixels of the non-edge part of each object in the to-be-recognized image;

wherein the edge segmentation iamge is output from an edge segmentation model into which the to-be-recognized image is input.