CLAIMS

1. A stacked object recognition method, comprising:

acquiring an image to be recognized, the image to be recognized comprising an object sequence formed by stacking at least one object;

performing edge detection and semantic segmentation on the object sequence based on the image to be recognized to determine an edge segmentation image of the object sequence and a semantic segmentation image of the object sequence, the edge segmentation image comprising edge information of each object of the object sequence and each pixel in the semantic segmentation image representing a class of the object to which the pixel belongs;

determining number information of the at least one object in the object sequence based on the edge segmentation image; and

determining the class of each object in the object sequence based on the edge segmentation image and the semantic segmentation image.

2. The method of claim 1, wherein the determining the class of each object in the object sequence based on the edge segmentation image and the semantic segmentation image comprises:

determining a boundary position of each object in the object sequence in the image to be recognized based on the edge segmentation image; and

determining the class of each object in the object sequence based on pixel values of pixels in a region corresponding to the boundary position of each object in the semantic segmentation image, the pixel value of the pixel representing a class identifier of the object to which the pixel belongs.

3. The method of claim 2, wherein the determining the class of each object in the object sequence based on pixel values of pixels in a region corresponding to the boundary position of each object in the semantic segmentation image comprises:

for each object,

statistically obtaining the pixel values of the pixels in the region corresponding to the boundary position of the object in the semantic segmentation image;

determining the pixel value corresponding to a maximum number of pixels in the region according to a statistical result; and

determining a class identifier represented by the pixel value corresponding to the maximum number of pixels as a class identifier of the object.

4. The method of any one of claims 1-3, wherein the performing edge detection and semantic segmentation on the object sequence based on the image to be recognized to determine an edge segmentation image of the object sequence and a semantic segmentation image of the object sequence comprises:

sequentially performing convolution processing one time and pooling processing one time on the image to be recognized to obtain a first pooled image;

performing at least one first operation based on the first pooled image, the first operation comprising sequentially performing convolution processing one time and pooling processing one time based on an image obtained from latest pooling processing to obtain a first intermediate image;

performing merging processing and down-sampling processing on the first pooled image and each first intermediate image to obtain the edge segmentation image;

performing at least one second operation based on a first intermediate image obtained from a last first operation, the second operation comprising sequentially performing convolution processing one time and pooling processing one time based on an image obtained from latest pooling processing to obtain a second intermediate image; and

performing merging processing and down-sampling processing on the first intermediate image obtained from the last first operation and each second intermediate image to obtain the semantic segmentation image.

5. The method of any one of claims 1-4, wherein the edge segmentation image comprises a mask image representing the edge information of each object, and/or, the edge segmentation image is the same as the image to be recognized in size;

the semantic segmentation image comprises a mask image representing semantic information of each pixel, and/or, the semantic segmentation image is the same as the image to be recognized in size.

6. The method of claim 5, wherein the edge segmentation image is a binarized mask image, a pixel with a first pixel value in the edge segmentation image corresponds to an edge pixel of each object in the image to be recognized, and a pixel with a second pixel value in the edge segmentation image corresponds to a non-edge pixel of each object in the image to be recognized.

7. The method of any one of claims 1-6, wherein the performing edge detection and semantic segmentation on the object sequence based on the image to be recognized to determine an edge segmentation image of the object sequence and a semantic segmentation image of the object sequence comprises:

inputting the image to be recognized to a trained edge detection model to obtain an edge detection result of each object in the object sequence, the edge detection model being obtained by training based on a sequence object image comprising object edge labeling information;

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

inputting the image to be recognized to a trained semantic segmentation model to obtain a semantic segmentation result of each object in the object sequence, the semantic segmentation model being obtained by training based on a sequence object image comprising object semantic segmentation labeling information; and

generating the semantic segmentation image of the object sequence according to the semantic segmentation result.

8. The method of any one of claims 1-7, wherein the determining the class of each object in the object sequence based on the edge segmentation image and the semantic segmentation image comprises:

fusing the edge segmentation image and the semantic segmentation image to obtain a fusion image, the fusion image comprising the semantic segmentation image and the edge information of each object displayed in the semantic segmentation image;

determining a pixel value corresponding to a maximum number of pixels in a region corresponding to the edge information of each object in the fusion image; and

determining a class represented by the pixel value corresponding to the maximum number of pixels as the class of each object.

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

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

10. A stacked object recognition device, comprising a memory and a processor,

wherein the memory stores a computer program capable of running in the processor;

wherein when executing the computer program, the processor is configured to:

acquire an image to be recognized, the image to be recognized comprising an object sequence formed by stacking at least one object;

perform edge detection and semantic segmentation on the object sequence based on

the image to be recognized to determine an edge segmentation image of the object sequence and a semantic segmentation image of the object sequence, the edge segmentation image comprising edge information of each object of the object sequence and each pixel in the semantic segmentation image representing a class of the object to which the pixel belongs;

determine number information of the at least one object in the object sequence based on the edge segmentation image; and

determine the class of each object in the object sequence based on the edge segmentation image and the semantic segmentation image.

11. The device of claim 10, wherein when determining the class of each object in the object sequence based on the edge segmentation image and the semantic segmentation image, the processor is configured to:

determine a boundary position of each object in the object sequence in the image to be recognized based on the edge segmentation image; and

determine the class of each object in the object sequence based on pixel values of pixels in a region corresponding to the boundary position of each object in the semantic segmentation image, the pixel value of the pixel representing a class identifier of the object to which the pixel belongs.

12. The device of claim 11, wherein when determining the class of each object in the object sequence based on the pixel values of pixels in the region corresponding to the boundary position of each object in the semantic segmentation image, the processor is configured to:

for each object,

statistically obtain the pixel values of the pixels in the region corresponding to the boundary position of the object in the semantic segmentation image;

determine the pixel value corresponding to a maximum number of pixels in the region according to a statistical result; and

determine a class identifier represented by the pixel value corresponding to the maximum number of pixels as a class identifier of the object.

13. The device of any one of claims 10-12, wherein when performing the edge detection and the semantic segmentation on the object sequence based on the image to be recognized to determine the edge segmentation image of the object sequence and the semantic segmentation image of the object sequence, the processor is configured to:

sequentially perform convolution processing one time and pooling processing one time on the image to be recognized to obtain a first pooled image;

perform at least one first operation based on the first pooled image, the first operation comprising sequentially performing convolution processing one time and pooling processing one time based on an image obtained from latest pooling processing to obtain a first intermediate image;

perform merging processing and down-sampling processing on the first pooled image and each first intermediate image to obtain the edge segmentation image;

perform at least one second operation based on a first intermediate image obtained from a last first operation, the second operation comprising sequentially performing convolution processing one time and pooling processing one time based on an image obtained from latest pooling processing to obtain a second intermediate image; and

perform merging processing and down-sampling processing on the first intermediate image obtained from the last first operation and each second intermediate image to obtain the semantic segmentation image.

14. The device of any one of claims 10-13, wherein the edge segmentation image comprises a mask image representing the edge information of each object, and/or, the edge segmentation image is the same as the image to be recognized in size;

the semantic segmentation image comprises a mask image representing semantic information of each pixel, and/or, the semantic segmentation image is the same as the image to be

recognized in size.

15. The device of claim 14, wherein the edge segmentation image is a binarized mask image, a pixel with a first pixel value in the edge segmentation image corresponds to an edge pixel of each object in the image to be recognized, and a pixel with a second pixel value in the edge segmentation image corresponds to a non-edge pixel of each object in the image to be recognized.

16. The device of any one of claims 10-15, wherein when performing the edge detection and the semantic segmentation on the object sequence based on the image to be recognized to determine the edge segmentation image of the object sequence and the semantic segmentation image of the object sequence, the processor is configured to:

input the image to be recognized to a trained edge detection model to obtain an edge detection result of each object in the object sequence, the edge detection model being obtained by training based on a sequence object image comprising object edge labeling information;

generate the edge segmentation image of the object sequence according to the edge detection result;

input the image to be recognized to a trained semantic segmentation model to obtain a semantic segmentation result of each object in the object sequence, the semantic segmentation model being obtained by training based on a sequence object image comprising object semantic segmentation labeling information; and

generate the semantic segmentation image of the object sequence according to the semantic segmentation result.

17. The device of any one of claims 10-16, wherein when determining the class of each object in the object sequence based on the edge segmentation image and the semantic segmentation image, the processor is configured to:

fuse the edge segmentation image and the semantic segmentation image to obtain a fusion image, the fusion image comprising the semantic segmentation image and the edge information of each object displayed in the semantic segmentation image;

determine a pixel value corresponding to a maximum number of pixels in a region corresponding to the edge information of each object in the fusion image; and

determine a class represented by the pixel value corresponding to the maximum number of pixels as the class of each object.

18. The device of any one of claims 10-17, wherein the object has a value attribute corresponding to the class; and the processor is further configured to:

determine a total value of objects in the object sequence based on the class of each object and the corresponding value attribute.

19. A computer storage medium, storing at least one program, wherein when executed by at least one processor, the at least one program is configured to:

acquire an image to be recognized, the image to be recognized comprising an object sequence formed by stacking at least one object;

perform edge detection and semantic segmentation on the object sequence based on the image to be recognized to determine an edge segmentation image of the object sequence and a semantic segmentation image of the object sequence, the edge segmentation image comprising edge information of each object of the object sequence and each pixel in the semantic segmentation image representing a class of the object to which the pixel belongs;

determine number information of the at least one object in the object sequence based on the edge segmentation image; and

determine the class of each object in the object sequence based on the edge segmentation image and the semantic segmentation image.

20. 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:

acquire an image to be recognized, the image to be recognized comprising an object sequence formed by stacking at least one object;

perform edge detection and semantic segmentation on the object sequence based on the image to be recognized to determine an edge segmentation image of the object sequence and a semantic segmentation image of the object sequence, the edge segmentation image comprising edge information of each object of the object sequence and each pixel in the semantic segmentation image representing a class of the object to which the pixel belongs;

determine number information of the at least one object in the object sequence based on the edge segmentation image; and

determine the class of each object in the object sequence based on the edge segmentation image and the semantic segmentation image.