

WHAT IS CLAIMED IS:

1. A goods sensing system based on image monitoring, comprising:
a sample collection unit for collecting a plurality of sets of image samples, wherein
5 each set of the image samples comprise a plurality of sample images of a type of
goods at multiple angles, wherein a set of the image samples of a same type of goods
are provided with a same group identification, and the group identification is the
type of the goods corresponding to the set of image samples;
a model training unit for training a convolutional neural network model according
10 to each sample image in the plurality of sets of the image samples and a group
identification of the sample image to obtain a goods identification model;
a real-time image acquisition unit for continuously acquiring at least one real-time
image of space in front of a shelf, wherein each real-time image comprises part or
all of images of goods; and
15 a goods category acquisition unit for obtaining a type and quantity of the goods
displayed in the real-time image according to the real-time image and the goods
identification model.
2. The goods sensing system based on image monitoring according to claim 1,
further comprising:
20 a first camera connected to the sample collection unit, for shooting images of each
goods at multiple angles and/or multiple distances; and
a second camera connected to the real-time image acquisition unit, for shooting the
real-time image of the space in front of the shelf;
wherein the second camera is provided with a lens, and a field of view of the lens

covers the space in front of the shelf.

3. The goods sensing system based on image monitoring according to claim 2, wherein the number of the second cameras is two or four;

the lens of the second camera faces a middle portion of the space in front of the shelf;

5 the angle between a central axis of the lens of the second camera and a horizontal plane is 30-60 degrees; and/or

a distance between the lens of the second camera and an upper end or a lower end of the shelf is 0.8-1.2 meters; and/or

a distance between the lens of the second camera and one side of the shelf is 0.8-1.2
10 meters.

4. The goods sensing system based on image monitoring according to claim 1, wherein the model training unit comprises:

a grouping model training unit for inputting each sample image of a plurality of sets of training samples and a group identification of each set of the training samples into

15 the convolutional neural network model, and obtaining a grouping model after training;

wherein the training samples are all or part of the plurality of sets of the image samples, and the grouping model is the goods identification model.

5. The goods sensing system based on image monitoring according to claim 4,

20 wherein the model training unit comprises:

a sample classification unit for randomly dividing the plurality of sets of the image samples into training samples and test samples;

a grouping model training unit for inputting each sample image of a plurality of sets of the training samples and a group identification of each set of the training samples

into the convolutional neural network model, and obtaining a grouping model after training;

and a cross-validation unit for validating the grouping model according to each image of a plurality of sets of the test samples and a group identification of each set
5 of the test samples, and calculating accuracy of the model;

wherein when the accuracy of the grouping model is smaller than a preset threshold value, it is returned to the sample classification unit; and when the accuracy of the grouping model is greater than or equal to the preset threshold value, the grouping model is the goods identification model.

10 6. The goods sensing system based on image monitoring according to claim 4 or 5, wherein the grouping model training unit comprises:

a feature extraction unit for inputting an image of each training sample into the convolutional neural network to perform feature extraction and obtain a feature map;

a region proposal generating unit for inputting the image of each training sample
15 into a region proposal network to generate a plurality of region proposals;

a region proposal mapping unit for mapping the region proposals of the image of each training sample to the feature map of a last layer of convolution layers of the convolutional neural network;

and a classifier generating unit for collecting the feature maps and the region
20 proposals of images of the training samples, calculating proposal feature maps, and sending the proposal feature maps into a classifier network to generate a classifier.

7. The goods sensing system based on image monitoring according to claim 6, wherein the cross-validation unit comprises:

a test sample input unit for inputting each image of the plurality of test samples into

the grouping model to obtain test group identifications of the plurality of the test samples;

and an accuracy calculation unit for comparing the test group identifications of the plurality of the test samples with the group identifications of the test samples, and

5 calculating a ratio of the number of the same identifications to the number of the test samples, wherein the ratio is the accuracy of the primary data model.

8. The goods sensing system based on image monitoring according to claim 1, wherein the goods category acquisition unit further comprises:

a group identification acquisition unit for inputting a plurality of real-time images
10 into the goods identification model and acquiring a plurality of group identifications corresponding to the plurality of real-time images as a possibility conclusion;

and an identification credibility calculating unit for calculating a ratio of the number of each group identification in the possibility conclusion to the total number of all group identifications in the possibility conclusion, wherein the ratio is a credibility

15 of each group identification, and a type and/or quantity of the goods corresponding to a group identification with a maximum credibility is the type and/or the quantity of the goods displayed in the real-time images.

9. The goods sensing system based on image monitoring according to claim 8, wherein the group identification acquiring unit comprises:

20 a second feature extraction unit for inputting a real-time image into the convolutional neural network to perform feature extraction and obtain a feature map;

a second region proposal generating unit for inputting the real-time image into a region proposal network, and generating a plurality of region proposals;

a second region proposal mapping unit for mapping the region proposals of the real-

time image to the feature map of a last layer of the convolutional neural network;
and a group identification acquisition unit for collecting the feature maps and region
proposals of the plurality of images, calculating proposal feature maps, sending the
proposal feature maps into a classifier network and acquiring the group
5 identifications corresponding to the real-time images.

10. A method for goods sensing based on image monitoring, comprising:
a sample collection step for collecting a plurality of sets of image samples, wherein
each set of the image samples comprise a plurality of sample images of a type of
goods at multiple angles, wherein a set of image samples of a same type of goods
10 are provided with a same group identification, and the group identification is the
type of the goods corresponding to the set of image samples;
a model training step for training a convolutional neural network model according
to each sample image in the plurality of sets of the image samples and a group
identification of the sample image to obtain a goods identification model;
15 a real-time image acquisition step for continuously acquiring at least one real-time
image of space in front of a shelf, wherein each real-time image comprises a part or
all of images of goods; and
a goods category acquisition step for obtaining a type and quantity of the goods
displayed in the real-time image according to the real-time image and the goods
20 identification model.

11. The method for goods sensing based on image monitoring according to
claim 10, wherein the model training step comprises:
a grouping model training step for inputting each sample image of a plurality of sets
of training samples and a group identification of each set of the training samples into

the convolutional neural network model, and obtaining a grouping model after training;

wherein the training samples are all or part of the plurality of sets of the image samples, and the grouping model is the goods identification model.

5 12. The method for goods sensing based on image monitoring according to claim 10, wherein the model training step comprises:

a sample classification step for randomly dividing the plurality of sets of the image samples into training samples and test samples;

10 a grouping model training step for inputting each sample image of a plurality of sets of the training samples and a group identification of each set of the training samples into the convolutional neural network model, and obtaining a grouping model after training; and

15 a cross-validation step for validating the grouping model according to each image of a plurality of sets of the test samples and a group identification of each set of the test samples, and calculating accuracy of the model;

wherein when the accuracy of the grouping model is smaller than a preset threshold value, returning to the sample classification step; and when the accuracy of the grouping model is greater than or equal to the preset threshold value, the grouping model is the goods identification model.

20 13. The method for goods sensing based on image monitoring according to claim 11 or 12, wherein the grouping model training step comprises:

a feature extraction step for inputting an image of each training sample into the convolutional neural network, and performing feature extraction to obtain a feature map;

a region proposal generation step for inputting the image of each training sample into a region proposal network to generate a plurality of region proposals;

a region proposal mapping step for mapping the region proposals of the image of each training sample to the feature map of a last layer of convolution layers of the

5 convolutional neural network;

and a classifier generation step for collecting the feature maps and the region proposals of images of the training samples, calculating proposal feature maps, and sending the proposal feature maps into a classifier network to generate a classifier.

14. The method for goods sensing based on image monitoring according to
10 claim 12, wherein the cross-validation step comprises:

a test sample input step for inputting each image of the plurality of the test samples into the grouping model to obtain test group identifications of the plurality of the test samples; and

an accuracy calculation step for comparing the test group identification of the
15 plurality of test samples with the group identification of the test samples, and calculating a ratio of the number of the same identifications to the number of the test samples, wherein the ratio is the accuracy of the primary data model.

15. The method for goods sensing based on image monitoring according to claim 10, wherein the goods category acquisition step comprises:

20 a group identification acquisition step for inputting a plurality of real-time images into the goods identification model and obtaining a plurality of group identifications corresponding to the plurality of real-time images as a possibility conclusion; and an identification credibility calculation step for calculating a ratio of the number of each group identification in the possibility conclusion to the total number of all

group identifications in the possibility conclusion, wherein the ratio is a credibility of each group identification, and a type and/or quantity of the goods corresponding to a group identification with a maximum credibility are the type and/or the quantity of the goods displayed in the real-time images.

- 5 16. The method for goods sensing based on image monitoring according to claim 10, wherein the group identification acquisition step comprises:
- a second feature extraction step for inputting a real-time image into the convolutional neural network, and performing feature extraction to obtain a feature map;
- a second region proposal generation step for inputting the real-time image into a
- 10 region proposal network, and generating a plurality of region proposals;
- a second region proposal mapping step for mapping the region proposals of the real-time image to the feature map of a last layer of the convolutional neural network;
- and a group acquisition step for collecting the feature maps and region proposals of the plurality of images, calculating proposal feature maps, sending the proposal
- 15 feature maps into a classifier network and acquiring the group identifications corresponding to the real-time image.