WHAT IS CLAIMED IS:

1. A training method of three-dimensional point cloud processing model, comprising:

processing three-dimensional point cloud data using a three-dimensional point cloud processing model to acquire information of points in the three-dimensional point cloud data;

acquiring information of pixels in two-dimensional image data, wherein the two-dimensional image data and the three-dimensional point cloud data cover a same area;

determining an alignment constraint loss between the three-dimensional point cloud data and the two-dimensional image data using the information of the points and the information of the pixels;

determining a classification loss of the three-dimensional point cloud processing model using classification information of the points in the three-dimensional point cloud data predicted by the three-dimensional point cloud processing model and pre-labeled scene classification information; and

training the three-dimensional point cloud processing model using the alignment constraint loss and the classification loss.

2. The training method according to claim 1, wherein:

the information of the points comprises at least one of

classification information of the points or feature information of the points; and

the information of the pixels comprises at least one of classification information of the pixels or feature information of the pixels.

3. The training method according to claim 1 or 2, wherein the determining the alignment constraint loss between the three-dimensional point cloud data and the two-dimensional image data using the information of the points and the information of the pixels comprises:

determining a matching degree between the three-dimensional point cloud data and the two-dimensional image data using the information of the points and the information of the pixels; and

determining the alignment constraint loss according to the matching degree.

4. The training method according to claim 3, wherein the matching degree comprises at least one of a matching degree of features between the points and the pixels, a matching degree of global features between the three-dimensional point cloud data and the two-dimensional image data, or a matching degree of semantic between the points and the pixels;

optionally, the matching degree of features between the points

and the pixels is determined according to a similarity between the points and corresponding pixels;

optionally, the matching degree of global features is determined according to a similarity between any two points in the three-dimensional point cloud data and a similarity between any two pixels in the two-dimensional image data; or

optionally, the matching degree of semantic is determined according to first classification information of each of the points and second classification information of each of the pixels.

5. The training method according to claim 4, wherein the matching degree comprises the matching degree of features between the points and the pixels, and the determining the alignment constraint loss according to the matching degree comprises:

mapping features of the points and features of the pixels to a same dimension to obtain a first point cloud feature of each of the points and a first pixel feature of each of the pixels respectively;

randomly acquiring a first number of points and pixels corresponding to the first number of points;

for each of the first number of points, determining a contrastive learning loss between the point and a pixel corresponding to the point according to a similarity between the first point cloud feature of the point and the first pixel feature

of the pixel corresponding to the point and a sum of similarities between the first point cloud feature of the point and the first pixel features of the pixels; and

determining the alignment constraint loss according to a sum of contrastive learning losses of the first number of points;

optionally, for each of the first number of points, the contrastive learning loss between the point and the pixel corresponding to the point is negatively correlated with the similarity between the first point cloud feature of the point and the first pixel feature of the pixel corresponding to the point, and positively correlated with the sum of the similarities between the first point cloud feature of the point and the first pixel features of the pixels.

6. The training method according to claim 4, wherein the matching degree comprises the matching degree of global features, and the determining the alignment constraint loss according to the matching degree comprises:

mapping features of the points and features of the pixels to a same dimension to respectively obtain a second point cloud feature of each of the points and a second pixel feature of each of the pixels;

randomly acquiring a second number of points and pixels corresponding to the second number of the points;

for any two points of the second number of points, determining first similarity information between second point cloud features of the two points and second similarity information between second pixel features of two pixels corresponding to the two points of a second number of pixels, and determining a similarity loss corresponding to the two points according to a difference between the first similarity information and the second similarity information; and

determining the alignment constraint loss according to a sum of similarity losses of points in the three-dimensional point cloud data;

optionally, the first similarity information and the second similarity information are similarity matrices.

7. The training method according to claim 4, wherein the second classification information is obtained by processing the each of the pixels using a two-dimensional image processing model, the matching degree comprises the matching degree of semantic, and the determining the alignment constraint loss according to the matching degree comprises:

mapping the three-dimensional point cloud data and the two-dimensional image data to a same dimension to obtain mapped three-dimensional point cloud data;

processing the mapped three-dimensional point cloud data using

the two-dimensional image processing model to obtain third classification information of the each of the points; and

determining a semantic matching loss using the second classification information and the third classification information, and determining the alignment constraint loss according to the semantic matching loss;

optionally, the alignment constraint loss is positively correlated with the second classification information and negatively correlated with the third classification information.

8. The training method according to any one of claims 1 to 7, wherein the acquiring the information of the pixels in the two-dimensional image data comprises:

processing the two-dimensional image data using a two-dimensional image processing model to acquire the information of the pixels in the two-dimensional image data;

optionally, the two-dimensional image processing model is implemented using a two-dimensional residual network.

9. The training method according to claim 8, wherein the determining the classification loss of the three-dimensional point cloud processing model using the classification information of the points in the three-dimensional point cloud data predicted by the three-dimensional point cloud processing model and the pre-labeled

scene classification information comprises:

determining a first classification loss according to the classification information of the points in the three-dimensional point cloud data predicted by the three-dimensional point cloud processing model and the pre-labeled scene classification information;

determining a second classification loss according to the classification information of the pixels in the two-dimensional image data predicted by the two-dimensional image processing model and the scene classification information; and

determining the classification loss of the three-dimensional point cloud processing model according to the first classification loss and the second classification loss;

optionally, the first classification loss is determined according to a pooling processing result of the classification information of the three-dimensional point cloud data predicted by the three-dimensional point cloud processing model and the scene classification information; and

the second classification loss is determined according to a pooling processing result of the classification information of the pixels in the two-dimensional image data predicted by the two-dimensional image processing model and the scene classification information.

10. The training method according to any one of claims 1 to 9, wherein the three-dimensional point cloud processing model is trained using a weighted result of the alignment constraint loss and the classification loss.

11. The training method according to any one of claims 1 to 10, wherein the three-dimensional point cloud processing model is implemented using a three-dimensional residual network.

12. A training method of semantic segmentation model, comprising:

predicting classification information of points in three-dimensional point cloud data to be processed using a three-dimensional point cloud processing model, wherein the three-dimensional point cloud processing model is trained using the training method of a three-dimensional point cloud processing model according to any one of claims 1 to 19;

generating pseudo labels of the points in the three-dimensional point cloud data to be processed according to the classification information; and

training a semantic segmentation model of three-dimensional point cloud using the pseudo labels.

13. A training device of three-dimensional point cloud

processing model, comprising:

a memory; and

a processor coupled to the memory, the processor being configured to, based on instructions stored in the memory, perform the training method of three-dimensional point cloud processing model according to any one of claims 1 to 11.

14. A training system for a semantic segmentation model, comprising:

the training device of a three-dimensional point cloud processing model according to claim 13; and

a training device of semantic segmentation model, comprising:

a memory; and

a processor coupled to the memory, the processor being configured to, based on instructions stored in the memory, perform the training method of semantic segmentation model according to claim 12.

15. A computer-readable storage medium having stored thereon a computer program, which, when executed by a processor, implements the training method of three-dimensional point cloud processing model according to any one of claims 1 to 11.