CLAIMS

1. A system comprising:

one or more processors; and

memory containing instructions to control the one or more processors to:

receive a plurality of images representing an interior of a physical environment;

generate a unified three-dimensional (3D) coordinate system based on the plurality of images;

identify, using a neural network for object recognition, a first object in a first image of the plurality of images, the first object being associated with first object location information, the first object location information comprising 3D coordinates within the unified 3D coordinate system;

identify, using the neural network for object recognition, a second object in a second image of the plurality of images;

map the first object and the second object to specific 3D coordinates within the unified 3D coordinate system;

compare the specific 3D coordinates of the first object and the second object within the unified coordinate system to determine if the first object and the second object represent a same physical object;

if the comparison of the specific 3D coordinates indicates the first object and the second object represent the same physical object instance, then the first object and the second object are an instance of a single object;

store the specific 3D coordinates associated with the single object;

display an interactive walkthrough visualization of a three-dimensional (3D) model of the physical environment, the 3D model including the single object;

receive an object information request that queries for object instances through the interactive walkthrough visualization; and

provide the specific 3D coordinates of the single object for display in the interactive walkthrough visualization.

2. The system of claim 1, wherein the instructions further control the one or more processors to determine if the first object and the second object are of a same object type.

3. The system of claim 2, wherein the first object and the second object are the instance of the single object if the first object and the second object are located near or at a similar location and the first object is the same object type as the second object.

4. The system of claim 2, wherein identify the first object using the neural network for object recognition further comprises categorizing the first object using the neural network to identify an object type of the first object.

5. The system of claim 1, wherein the instructions further control the one or more processors to determine the first object location information and the second object location information based on a mesh of the 3D model.

6. The system of claim 1, wherein the first image includes position and orientation of an image capture device that captured the first image in the physical environment, the instructions further control the one or more processors to determine the first object location information based on the position and orientation of the image capture device that captured the first image.

7. The system of claim 6, wherein the second image includes position and orientation of an image capture device that captured the second image in the physical environment, the instructions further control the one or more processors to determine the second object location information based on the position and orientation of the image capture device that captured the second image.

8. The system of claim 1, wherein the first image is associated with depth data of the physical environment, the depth data being generated by a depth data device, the first object location being determined based on the first object depicted in the first image and associated relevant depth data.

9. The system of claim 8, wherein the depth data device is a LiDAR.

10. A non-transitory computer-readable medium comprising executable instructions, the executable instructions being executable by one or more processors to perform a method, the method comprising:

receiving a plurality of images representing an interior of a physical environment;

generating a unified three-dimensional (3D) coordinate system based on the plurality of images;

identifying, using a neural network for object recognition, a first object in a first image of the plurality of images, the first object being associated with first object location information, the first object location information comprising 3D coordinates within the unified 3D coordinate system;

identifying, using the neural network for object recognition, a second object in a second image of a plurality of images;

mapping the first object and the second object to specific 3D coordinates within the unified 3D coordinate system;

comparing the specific 3D coordinates of the first object and the second object within the unified coordinate system to determine if the first object and the second object represent a same physical object instance;

if the comparison of the specific 3D coordinates indicates the first object and the second object represent the same physical object instance, then the first object and the second object are an instance of a single object;

storing the specific 3D coordinates associated with the single object;

displaying an interactive walkthrough visualization of a three-dimensional (3D) model of the physical environment, the 3D model including the single object;

receiving an object information request that queries for object instances through the interactive walkthrough visualization; and

providing the specific 3D coordinates of the single object for display in the interactive walkthrough visualization.

11. The non-transitory computer-readable medium of claim 10, the executable instructions that are executable by the one or more processors to further:

determining if the first object and the second object are of a same object type.

12. The non-transitory computer-readable medium of claim 11, wherein the first object and the second object are the instance of the single object if the first object and the second object are located near or at a similar location and the first object is the same object type as the second object.

13. The non-transitory computer-readable medium of claim 11, wherein identifying the first object using the neural network for object recognition further comprises categorizing the first object using the neural network to identify the object type of the first object.

14. The non-transitory computer-readable medium of claim 10, the executable instructions that are executable by the one or more processors to further:

controlling the one or more processors to determine the first object location information and the second object location information based on a mesh of the 3D model.

15. The non-transitory computer-readable medium of claim 10, wherein the first image includes position and orientation of an image capture device that captured the first image in the physical environment, the executable instructions that are executable by the one or more processors to further controlling the one or more processors to determine the first object location information based on the position and orientation of the image capture device that captured the first image.

16. The non-transitory computer-readable medium of claim 15, wherein the second image includes position and orientation of an image capture device that captured the second image in the physical environment, the executable instructions that are executable by the one or more processors to further determining the second object location information based on the position and orientation of the image capture device that captured the second image.

17. The non-transitory computer-readable medium of claim 10, wherein the first image is associated with depth data of the physical environment, the depth data being generated by a depth data device, the first object location being determined based on the first object depicted in the first image and associated relevant depth data.

18. The non-transitory computer-readable medium of claim 17, wherein the depth data device is a LiDAR.

19. A method comprising:

receiving a plurality of images representing an interior of a physical environment;

generating a unified three-dimensional (3D) coordinate system based on the plurality of images;

identifying, using a neural network for object recognition, a first object in a first image of the plurality of images, the first object being associated with first object location information, the first object location information comprising 3D coordinates within the unified 3D coordinate system;

identifying, using the neural network for object recognition, a second object in a second image of a plurality of images;

mapping the first object and the second object to specific 3D coordinates within the unified 3D coordinate system;

comparing the specific 3D coordinates of the first object and the second object within the unified coordinate system to determine if the first object and the second object represent a same physical object instance;

if the comparison of the specific 3D coordinates indicates the first object and the second object represent the same physical object instance, then the first object and the second object are an instance of a single object;

storing the specific 3D coordinates associated with the single object;

displaying an interactive walkthrough visualization of a three-dimensional (3D) model of the physical environment, the 3D model including the single object;

receiving an object information request that queries for object instances through the interactive walkthrough visualization; and

providing the specific 3D coordinates of the single object for display in the interactive walkthrough visualization.