

CLAIMS

1. A target object processing method, wherein the method is applied to an edge device, and a first processing module for annotation of first data and a second processing module for adaptation of a scene indicated by second data are deployed on the edge device; and

the method comprises:

inputting the first data into the first processing module to obtain a predicted data annotation result;

inputting the data annotation result into the second processing module, and performing scene-adaptive incremental learning according to the data annotation result to obtain a neural network adapted to the scene of the second data; and

processing a scene corresponding to a target object according to the neural network and data comprising the target object.

2. The method according to claim 1, further comprising:

if the scene is a current first scene, after incremental learning adaptive to the first scene is performed according to the data annotation result to obtain a neural network adapted to the first scene, monitoring that the scene is switched from the first scene to a second scene;

triggering parameter resetting of a parameter in the second processing module; and

performing incremental learning adaptive to the second scene according to the data annotation result to obtain a neural network adapted to the second scene.

3. The method according to claim 1 or 2, further comprising: before obtaining the neural network,

in condition that parameter in the second processing module is to be updated, performing update restriction on updating of the parameter according to a constraint condition.

4. The method according to claim 1, wherein inputting the first data into the first processing module to obtain the predicted data annotation result comprises:

obtaining, according to prediction of scene data by the first processing module, a prediction rule; and

annotating the first data according to the prediction rule to obtain the predicted data annotation result.

5. The method according to any one of claims 1 to 4, wherein the predicted data annotation result is derived from a first output of an output layer, constituting the first processing module, in a data sample training network; and/or

the predicted data annotation result is derived from the first output of output layer constituting the first processing module and a second output of an intermediate layer in the data sample training network.

6. The method according to claim 1, further comprising: before inputting the first data into the first processing module, performing selection on the first data according to a sampling parameter to obtain to-be-processed first data; and

the method further comprising: before obtaining the neural network,

for the to-be-processed first data, if a prediction result output by the first processing module greatly differs from a prediction result output by the second processing module, performing, for one or more frames of first data of which prediction results of the first processing module greatly differ from that of the second processing module, scene-adaptive incremental learning at the edge device.

7. The method according to claim 6, wherein performing, for one or more frames of first data of which prediction results of the first processing module greatly differ from that of the second processing module, scene-adaptive incremental learning at the edge device comprises:

outputting one or more data annotation results obtained by the first processing module for the one or more frames of the first data to the second processing module;

training the second processing module and updating the parameter in the second processing module, according to the one or more data annotation results; and

increasing a value of the sampling parameter by using a preconfigured strategy.

8. The method according to claim 1, further comprising: before inputting the first data into the first processing module, performing selection on the first data according to a sampling parameter to obtain to-be-processed first data; and

the method further comprising: before obtaining the neural network,

for the to-be-processed first data, if the prediction result output by the first processing module slightly differs from the prediction result output by the second processing module, reducing a value of the sampling parameter by using a preconfigured strategy.

9. The method according to claim 1, further comprising: before inputting the first data into the first processing module, using a part of the first data as an online test set; and

the method further comprising: before obtaining the neural network,

for first data in the online test set, if the prediction result output by the first processing module greatly differs from the prediction result output by the second processing module, resetting the parameter in the second processing module.

10. The method according to any one of claims 1 to 9, further comprising: after obtaining the neural network,

monitoring a plurality of edge devices in a plurality of regions to obtain a first edge device in an idle state; and

performing adaptive incremental training adapted to the scene of the second data on the second processing module according to an edge device processing capability corresponding to the first edge device.

11. The method according to any one of claims 1 to 9, further comprising: after obtaining the neural network,

monitoring a plurality of edge devices in the plurality of regions to obtain a plurality of edge device processing capabilities;

selecting, from the plurality of edge devices, a second edge device with a highest edge device processing capability according to the plurality of edge device processing capabilities respectively corresponding to the plurality of edge devices and a current resource consumption; and

performing adaptive incremental training adapted to the scene of the second data on the second processing module according to an edge device processing capability corresponding to the second edge device.

12. A target object processing apparatus, deployed on an edge device, and comprising a first processing module, a second processing module, and a third processing module, wherein

the first processing module is configured to obtain a predicted data annotation result according to input first data;

the second processing module is configured to perform scene-adaptive incremental learning according to the input data annotation result to obtain a neural network adapted to a scene indicated by second data; and

the third processing module is configured to process a scene corresponding to a target object according to a neural network and data comprising the target object.

13. The apparatus according to claim 12, further comprising:

a monitoring module, configured to: if the scene is a current first scene, after incremental learning adaptive to the first scene is performed according to the data annotation result to obtain a neural network adapted to the first scene, monitor that the scene is switched from the first scene to a second scene; and

a resetting trigger module, configured to trigger parameter resetting of a parameter in the second processing module, wherein

the second processing module is further configured to perform incremental learning adaptive to the second scene according to the data annotation result to obtain a neural network adapted to the second scene.

14. The apparatus according to claim 12 or 13, further comprising a parameter update module, configured to:

in condition that the parameter is to be updated, perform update restriction on updating of the parameter according to a constraint condition.

15. The apparatus according to claim 12, wherein the first processing module is further configured to:

obtain a prediction rule according to prediction of scene data; and

annotate the first data according to the prediction rule to obtain the predicted data annotation result.

16. The apparatus according to any one of claims 12 to 15, wherein the predicted data annotation result is derived from a first output of an output layer, constituting the first processing module, in a data sample training network; and/or

the predicted data annotation result is derived from the first output of the output layer constituting the first processing module and a second output of an intermediate layer in the data sample training network.

17. The apparatus according to claim 12, further comprising:

a fourth processing module, configured to: before inputting the first data into the first processing module, perform selection on the first data according to a sampling parameter to obtain to-be-processed first data; and

a fifth processing module, configured to:

for the to-be-processed first data, if a prediction result output by the first processing module greatly differs from a prediction result output by the second processing module, perform scene-adaptive incremental learning at the edge device for one or more frames of the first data of which prediction results of the first processing module greatly differ from that of the second processing module.

18. An electronic device, comprising:

a processor; and

a memory configured to store processor executable instructions, wherein

the processor is configured to perform the method according to any one of claims 1 to 11.

19. A computer storage medium having executable instructions stored therein, wherein the executable instructions, when being executed by a processor, enable the processor to implement the method according to any one of claims 1 to 11.

20. A computer program product, wherein the computer program product, when being executed by a processor, enables the processor to implement the method according to any one of claims 1 to 11.