# CLAIMS

1. An object prediction method, applied to a neural network and comprising:

performing feature extraction processing on a to-be-predicted object to obtain feature information of the to-be-predicted object;

determining multiple intermediate prediction results for the to-be-predicted object according to the feature information;

performing fusion processing on the multiple intermediate prediction results to obtain fusion information; and

determining multiple target prediction results for the to-be-predicted object according to the fusion information.

2. The method according to claim 1, wherein performing feature extraction processing on the to-be-predicted object to obtain feature information of the to-be-predicted object comprises:

performing feature extraction processing on a to-be-predicted object to obtain multiple levels of features; and

performing aggregation processing on the multiple levels of features to obtain feature information of the to-be-predicted object.

3. The method according to claim 1, wherein determining multiple intermediate prediction results for the to-be-predicted object according to the feature information comprises:

reconstructing the feature information to obtain multiple reconstitution features; and

determining multiple intermediate prediction results for the to-be-predicted object according to the multiple reconstruction features.

4. The method according to claim 1, wherein performing fusion processing on the multiple intermediate prediction results to obtain fusion information comprises:

re-processing the multiple intermediate prediction results to obtain re-processing results of the multiple intermediate prediction results; and

performing fusion processing on the re-processing results of the multiple intermediate prediction results to obtain fusion information.

5. The method according to claim 4, wherein performing fusion processing on the re-processing results of the multiple intermediate prediction results to obtain fusion information comprises:

performing superposition processing on the re-processing results of the multiple intermediate prediction results to obtain fusion information.

6. The method according to claim 4, wherein the multiple intermediate prediction results comprise a first intermediate prediction result and second intermediate prediction results, wherein the first intermediate prediction result has the highest correlation degree with the target prediction result; and

performing fusion processing on the re-processing results of the multiple intermediate prediction results to obtain fusion information comprises:

processing the re-processing results of the second intermediate prediction results to obtain reference results, and performing superposition processing on the re-processing result of the first intermediate prediction result and the reference results to obtain fusion information for the target prediction results;

or

determining an attention coefficient according to the re-processing result of the first intermediate prediction result, wherein the attention coefficient is a reference coefficient determined according to an attention mechanism, processing the re-processing results of the second intermediate prediction results to obtain reference results, performing dot product processing on the reference results and the attention coefficient to obtain attention content, and performing superposition processing on the re-processing result of the first intermediate prediction result and the attention content to obtain fusion information for the target prediction results.

7. The method according to claim 1, wherein determining multiple target prediction results for the to-be-predicted object according to the fusion information comprises:

determining fusion information for multiple target prediction results;

processing the fusion information to obtain target features; and

determining the multiple target prediction results according to the target features.

8. The method according to claim 1, wherein the neural network is trained with a to-be-predicted object,

the steps for training the neural network with the to-be-predicted object comprise:

inputting the to-be-predicted object into a feature extraction network in the neural network for feature extraction processing to obtain feature information of the to-be-predicted object;

determining multiple intermediate prediction results for the to-be-predicted object by inputting the feature information into a first prediction network in the neural network for processing;

inputting the intermediate prediction results into a fusion network in the neural network for fusion processing to obtain fusion information;

determining multiple target prediction results for the to-be-predicted object by inputting the fusion information into multiple second prediction networks in the neural network for processing respectively;

determining a model loss of the neural network according to the multiple intermediate prediction results, annotation information of the multiple intermediate prediction results, the multiple target prediction results, and annotation information of the multiple target prediction results; and

adjusting a network parameter value of the neural network according to the model loss.

9. The method according to claim 8, wherein the steps for training the neural network with the to-be-predicted object further comprise:

determining the annotation information of the multiple target prediction results before inputting the to-be-predicted object into the feature extraction network in the neural network for feature extraction processing to obtain the feature information of the to-be-predicted object; and

determining the annotation information of the multiple intermediate prediction results

according to the annotation information of the multiple target prediction results.

10. The method according to claim 1, wherein determining multiple intermediate prediction results for the to-be-predicted object according to the feature information comprises:

determining a depth estimation intermediate prediction result, a curved surface normal intermediate prediction result, a contour intermediate prediction result, and a semantic segmentation intermediate prediction result for the to-be-predicted object according to the feature information;

performing fusion processing on the multiple intermediate prediction results to obtain fusion information comprises:

performing fusion processing on the depth estimation intermediate prediction result, the curved surface normal intermediate prediction result, the contour intermediate prediction result, and the semantic segmentation intermediate prediction result to obtain fusion information; and

determining multiple target prediction results for the to-be-predicted object according to the fusion information comprises:

determining a depth estimation result and a scene segmentation result for the to-be-predicted object according to the fusion information.

11. An object prediction apparatus, applied to a neural network and comprising:

a feature extraction module, configured to perform feature extraction processing on a to-be-predicted object to obtain feature information of the to-be-predicted object;

an intermediate prediction result determination module, configured to determine multiple intermediate prediction results for the to-be-predicted object according to the feature information;

a fusion module, configured to perform fusion processing on the multiple intermediate prediction results to obtain fusion information; and

a target prediction result determination module, configured to determine multiple target prediction results for the to-be-predicted object according to the fusion information.

12. The apparatus according to claim 11, wherein the feature extraction module comprises:

a feature obtaining submodule, configured to perform feature extraction processing on a

to-be-predicted object to obtain multiple levels of features; and

a feature information obtaining submodule, configured to perform aggregation processing on the multiple levels of features to obtain feature information of the to-be-predicted object.

13. The apparatus according to claim 11, wherein the intermediate prediction result determination module comprises:

a reconstruction feature obtaining submodule, configured to reconstruct the feature information to obtain multiple reconstitution features; and

an intermediate prediction result obtaining submodule, configured to determine multiple intermediate prediction results for the to-be-predicted object according to the multiple reconstitution features.

14. The apparatus according to claim 11, wherein the fusion module comprises:

a re-processing result obtaining submodule, configured to re-process the multiple intermediate prediction results to obtain re-processing results of the multiple intermediate prediction results; and

a fusion information obtaining submodule, configured to perform fusion processing on the re-processing results of the multiple intermediate prediction results to obtain fusion information.

15. The apparatus according to claim 14, wherein the fusion information obtaining submodule is configured to:

perform superposition processing on the re-processing results of the multiple intermediate prediction results to obtain fusion information.

16. The apparatus according to claim 14, wherein the multiple intermediate prediction results comprise a first intermediate prediction result and second intermediate prediction results, wherein the first intermediate prediction result has the highest correlation degree with the target prediction result; and

the fusion information obtaining submodule is configured to:

process the re-processing results of the second intermediate prediction results to obtain

reference results, and perform superposition processing on the re-processing result of the first intermediate prediction result and the reference results to obtain fusion information for the target prediction results;

or

determine an attention coefficient according to the re-processing result of the first intermediate prediction result, wherein the attention coefficient is a reference coefficient determined according to an attention mechanism, process the re-processing results of the second intermediate prediction results to obtain reference results, perform dot product processing on the reference results and the attention coefficient to obtain attention content, and perform superposition processing on the re-processing result of the first intermediate prediction result and the attention content to obtain fusion information for the target prediction results.

17. The apparatus according to claim 11, wherein the target prediction result determination module comprises:

a fusion information determination submodule, configured to determine fusion information for multiple target prediction results;

a target feature obtaining submodule, configured to process the fusion information to obtain target features; and

a target prediction result determination submodule, configured to determine the multiple target prediction results according to the target features.

18. The apparatus according to claim 11, wherein the neural network is trained with a to-be-predicted object,

the apparatus further comprises:

a first obtaining module, configured to input the to-be-predicted object into a feature extraction network in the neural network for feature extraction processing to obtain feature information of the to-be-predicted object;

a first determination module, configured to determine multiple intermediate prediction results for the to-be-predicted object by inputting the feature information into a first prediction network in the neural network for processing, and;

a second obtaining module, configured to input the intermediate prediction results into a fusion network in the neural network for fusion processing to obtain the fusion information;

a second determination module, configured to determine multiple target prediction results for the to-be-predicted object by inputting the fusion information into multiple second prediction networks in the neural network for processing respectively;

a third determination module, configured to determine a model loss of the neural network according to the multiple intermediate prediction results, annotation information of the multiple intermediate prediction results, the multiple target prediction results, and annotation information of the multiple target prediction results; and

a network parameter value adjustment module, configured to adjust a network parameter value of the neural network according to the model loss.

19. The apparatus according to claim 18, further comprising:

an annotation information determination module, configured to determine the annotation information of the multiple target prediction results before inputting the to-be-predicted object into the feature extraction network in the neural network for feature extraction processing to obtain the feature information of the to-be-predicted object; and

an intermediate annotation information determination module, configured to determine the annotation information of the multiple intermediate prediction results according to the annotation information of the multiple target prediction results.

20. A computer readable storage medium having computer program instructions stored thereon, wherein when the computer program instructions are executed by a processor, the method according to any one of claims 1 to 10 is implemented.