Claims

1. A question answering method, comprising:

   obtaining question information to be answered; and

   inputting the question information into a question answering model, to obtain answer information corresponding to the question information;

   wherein the question answering model is trained based on predicted answer information, the predicted answer information is obtained by predicting, by using a pre-trained question answering model, answer information corresponding to a question sample, the pre-trained question answering model is trained based on a predicted image-text matching result, the predicted image-text matching result is obtained by matching an image to be trained with target semantic information using the pre-trained question answering model, and the target semantic information is generated based on an image description text corresponding to the image to be trained and a dialog text associated with the image to be trained.

2. The question answering method according to claim 1, wherein before inputting the question information into the question answering model, to obtain the answer information corresponding to the question information, the question answering method further comprises:

   obtaining a pre-training sample, wherein the pre-training sample comprises an image to be trained, an image description text corresponding to the image to be trained, and a dialog text associated with the image to be trained;

   training the pre-trained question answering model based on the predicted image-text matching result corresponding to the pre-training sample;

   obtaining a question sample, wherein the question sample comprises an image to be trained and the dialog text associated with the image to be trained; and

   inputting the question sample into the pre-trained question answering model, to output the predicted answer information, and training the pre-trained question answering model based on the predicted answer information, to obtain the question answering model.

3. The question answering method according to claim 2, wherein training the pre-trained question answering model based on the predicted answer information, to obtain the question answering model comprises:

   calculating an autoregressive loss value based on the predicted answer information; and

training the pre-trained question answering model based on the autoregressive loss value, to obtain the question answering model.

4. The question answering method according to claim 2 or 3, wherein training the pre-trained question answering model based on the predicted image-text matching result corresponding to the pre-training sample comprises:

performing, for the image to be trained in the pre-training sample, image processing on the image to be trained by using the pre-trained question answering model, to obtain an image vector;

performing, for the dialog text in the pre-training sample and associated with the image to be trained, text processing on the dialog text by using the pre-trained question answering model, to obtain a dialog vector;

generating the predicted image-text matching result based on the image vector and the dialog vector; and

determining an image-text alignment loss value based on the predicted image-text matching result, and training the pre-trained question answering model based on the image-text alignment loss value.

5. The question answering method according to claim 4, wherein training the pre-trained question answering model based on the image-text alignment loss value comprises:

calculating a dialog annotation alignment loss value based on the image vector, and calculating an image dialog loss value based on the dialog vector; and

training the pre-trained question answering model based on the dialog annotation alignment loss value, the image dialog loss value, and the image-text alignment loss value.

6. The question answering method according to claim 2, wherein before obtaining the pre-training sample, the method further comprises:

obtaining an initial training sample, wherein the initial training sample comprises the image to be trained and an image annotation text corresponding to the image to be trained;

performing, for the image to be trained in the initial training sample, image processing on the image to be trained by using an initial question answering model, to obtain an image vector;

performing, for the image annotation text corresponding to the image to be trained in the initial training sample, text processing on the image annotation text by using the initial question answering model, to obtain a text annotation vector;

determining a predicted image description text based on the image vector and the text annotation vector; and training the initial question answering model based on the predicted image description text.

7. The question answering method according to claim 6, wherein training the initial question answering model based on the predicted image description text comprises:

calculating an image-text annotation loss value based on the image vector and the text annotation vector;

calculating an image-text alignment loss value based on the predicted image description text; and

training the initial question answering model based on the image-text annotation loss value and the image-text alignment loss value.

8. The question answering method according to claim 1, wherein after obtaining the answer information corresponding to the question information, the method further comprises:

sending the answer information to a user;

receiving a feedback result of the user for the answer information; and

adjusting model parameters in the question answering model based on the feedback result.

9. The question answering method according to claim 1, wherein the predicted image-text matching result is obtained by matching the image to be trained with the target semantic information based on an information matching layer in the pre-trained question answering model, and the target semantic information is obtained through vector matching between the image description text and the dialog text.

10. A question answering method, applied to a cloud side device, wherein the question answering method comprises:

receiving a question answering request sent by a terminal side device, wherein the question answering request carries question information;

inputting the question information into a question answering model, to obtain answer information corresponding to the question information;

wherein the question answering model is trained based on predicted answer information, the predicted answer information is obtained by predicting, by using a pre-trained question answering model, answer information corresponding to a question sample, the pre-trained question answering model is trained based on a predicted image-text matching result, the predicted image-text matching result is obtained by matching an image to be trained with target semantic information using the pre-trained question answering model, and the target semantic information is generated based on an image description text corresponding to the image to be trained and a dialog text associated with the image to be trained; and

sending the answer information corresponding to the question information to the terminal side device.

11. A question answering model training method, applied to a cloud side device, and the question answering model training method comprises:

obtaining a pre-training sample, wherein the pre-training sample comprises an image to be trained, an image description text corresponding to the image to be trained, and a dialog text associated with the image to be trained;

training the pre-trained question answering model based on a predicted image-text matching result corresponding to the pre-training sample, wherein the predicted image-text matching result is obtained by matching the image to be trained with target semantic information using the pre-trained question answering model, and the target semantic information is generated based on the image description text and the dialog text;

obtaining a question sample, wherein the question sample comprises an image to be trained and dialog text associated with the image to be trained;

inputting the question sample into the pre-trained question answering model, to output the predicted answer information, and training the pre-trained question answering model based on the predicted answer information, to obtain a question answering model; and

sending model parameters of the question answering model to a terminal side device.

12. A textual question answering method, applied to a terminal side device, and the textual question answering method comprises:

receiving textual question information;

inputting the textual question information into a question answering model, to obtain textual answer information corresponding to the textual question information, wherein the question answering model is obtained according to the method as claimed in claim 10; and displaying the textual answer information.

13. The textual question answering method according to claim 12, wherein before inputting the textual question information into the question answering model, to obtain the textual answer information corresponding to the textual question information, the method further comprises:

obtaining model parameters of a question answering model in a cloud side device;

obtaining a textual question sample; and

constructing a question answering model to be trained based on the model parameters, and training the question answering model to be trained based on the textual question sample, to obtain the question answering model.

14. A computing device, comprising:

a memory and a processor;

wherein the memory is configured to store computer-executable instructions, the processor is configured to execute the computer-executable instructions, and when the computer-executable instructions are executed by the processor, the steps in the method as claimed in any one of claims 1 to 18 are implemented.

15. A computer-readable storage medium, storing computer-executable instructions, wherein when the computer-executable instructions are executed by a processor, the steps in the method as claimed in any one of claims 1 to 13 are implemented