

ABSTRACT

An image question answering method, apparatus and system, and a storage medium are provided. The image question answering method includes: extracting a question feature representing a semantic meaning of a question, a global feature of an image, and a detection frame feature of a detection frame encircling an object in the image; obtaining a first weight of each of at least one area of the image and a second weight of each of at least one detection frame of the image according to question feature, the global feature, and the detection frame feature; performing weighting processing on the global feature by using the first weight to obtain an area attention feature of the image; performing weighting processing on the detection frame feature by using the second weight to obtain a detection frame attention feature of the image; and predicting an answer to the question on the basis of the question feature, the area attention feature, and the detection frame attention feature. The technical solution provided by the present application may comprehensively consider the global feature of the image and the detection frame feature of the detection frame including the object, thereby improving the efficiency and accuracy of image question answering.

[Figure 1]