

ABSTRACT

A method and apparatus for generating a captioning device, and a method and apparatus for outputting a caption. The method for generating a captioning device comprises: acquiring a sample image set; inputting the sample image set into an image encoder of a sentence generator, so as to output an object set; grouping the object set into a first object set and a second object set, wherein the first object set is an object set that is included within a preset object set, and the second object set is an object set that is excluded from the preset object set; inputting, into a sentence decoder of the sentence generator, the object set output by the image encoder, and performing a beam search in a decoding step by taking the first object set and the second object set as constraint conditions, so as to generate a pseudo-image sentence pair set; and training the sentence generator by taking the pseudo-image sentence pair set as a sample set, so as to obtain a captioning device.

Fig. 2