METHOD AND DEVICE FOR TRAINING DOCUMENT REPRESENTATION MODEL, AND STORAGE MEDIUM

ABSTRACT

Embodiments of this specification provide a method and a device for training a document representation model, and a storage medium. In the method, a text feature corresponding to a document image can be generated in a text dimension, and a visual feature corresponding to the document image can be generated in a visual dimension; and a plurality of training tasks can be executed on the document representation model based on the text feature and the visual feature, to train the document representation model, where the visual feature is obtained by visually marking areas corresponding to prediction objects of the plurality of training tasks in the document image and performing visual feature extraction on a visually marked image, and prediction objects of different training tasks correspond to different visual marking manners.

Fig. 3