Enhanced Image Caption Using Scene-Graph Generation

Zhu, Feiyu (School: The High School Affiliated to Renmin University of China)

As deep learning is being implemented into multiple areas, the technique of describing the content of the image automatically has receive public attention and is being broadly used in circumstances like image retrieval, writing image-based stories and so on. However, present approaches often suffers from laking detailed information, and the sentences generated are too vague. This paper makes use of a scene-graph generation network to generate structured information, mainly the relationship of the objects, of the given image to make up for the vagueness of the base sentence generated by a CNN-LSTM neuron network. As a result, the sentences generated by the new method contain more detailed and structured information than using the past methods, and receive a higher BLEU-4 score. In addition, scene-graph also serves as a medium to make use of knowledge graph to introduce common sense and rules into image caption in the future.

Awards Won: