Paper: Automatic Image Caption Generation

Name: Milon Hossain

Id: 1610497042

**Abstract:** Image captioning involves creating a caption for an image automatically. As a research field that has recently arisen, it draws more and more attractiveness. Semantic image details must be captured and convoyed in photographs to accomplish the purpose of image captioning. Image captioning is a very complex activity that links both computer vision study groups and natural language processing. In order to solve this problem, different approaches have been proposed. Representative methods in each category are summarized, and their strengths and limitations are talked about. This paper focused their main attention on neural network-based methods, which give state of the art results. Neural network-based methods are further divided into subcategories based on the specific framework they use. Each subcategory of neural network-based methods is discussed in detail.

**Dataset:** In this paper various types of datasets are used for different methods of solving this project. Datasets be like: one for the image and another for the caption. Method comparison on Microsoft COCO Caption dataset under the commonly used protocol. Method comparison on datasets Flcikr8k and Flick30k are also used here.

**Complexity:** In this paper various types of critical mathematical and statistical terms have been used. Here used visual models to performs detections in images for extracting semantic information including objects, attributes and spatial relationships.

**Improvement:** There may be some other way or method to improve the complexity. Along with the quick Deep neural network development, using more powerful as expression models or visual models and network architecture could be built. Then the quality of image definition generation would certainly increase. Then it will be easier for the system to generate appropriate caption for each image.

**Conclusion:** This paper represents a system that automatically generate image caption using deep neural network and various types of datasets used in this project. In this paper their main attention was focused on neural network-based methods, which gives state of the art result. Because different frameworks are used in neural network-based methods, further divided into subcategories and discussed each subcategory, respectively. After that, state of the art methods is compared on benchmark datasets. Representative methods in each category are summarized, and strengths and limitations of each type of work are talked about in this paper.

This work was Supported by National Natural Science Foundation in China.

**Reference:** https://www.researchgate.net/publication/325384133