World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:02, 2023

Deep-Learning to Generation of Weights for Image Captioning Using Part-of-Speech Approach

Authors : Tiago do Carmo Nogueira, Cássio Dener Noronha Vinhal, Gélson da Cruz Júnior, Matheus Rudolfo Diedrich Ullmann
Abstract : Generating automatic image descriptions through natural language is a challenging task. Image captioning is a task that consistently describes an image by combining computer vision and natural language processing techniques. To accomplish this task, cutting-edge models use encoder-decoder structures. Thus, Convolutional Neural Networks (CNN) are used to extract the characteristics of the images, and Recurrent Neural Networks (RNN) generate the descriptive sentences of the images. However, cutting-edge approaches still suffer from problems of generating incorrect captions and accumulating errors in the decoders. To solve this problem, we propose a model based on the encoder-decoder structure, introducing a module that generates the weights according to the importance of the word to form the sentence, using the part-of-speech (PoS). Thus, the results demonstrate that our model surpasses state-of-the-art models.
Keywords : gated recurrent units, caption generation, convolutional neural network, part-of-speech
Conference Title : ICIAPR 2023 : International Conference on Image Analysis and Pattern Recognition
Conference Location : Lisbon, Portugal
Conference Dates : February 06-07, 2023