

Automatic Image Captioning using Deep Learning

A. T. V Sai Sony¹, N. Sailaja², B. Roshini³, K. Prasad⁴, Mr. S. Syam Kumar⁵

Students, Department of Computer Science and Engineering (Data Science)^{1,2,3,4}

Assistant Professor, Department of Computer Science and Engineering (Data Science)^{1,2,3,4}

Dadi Institute of Engineering & Technology (Autonomous), Anakapalle, India

Abstract: Image captioning aims to detect this information by describing the image content through image and text processing techniques. Now-a-days image captioning is one of the recent and growing research problem. Day by day various solutions are being introduced for solving the problem. Even though, many solutions are already available, a lot of attention is still required for getting better and precise results. So, we came up with the idea of developing a image captioning model using different combinations of Convolutional Neural Network architecture along with Long Short Term Memory to get better results. We have used three combination of CNN and LSTM for developing the model. The proposed model is trained with three Convolutional Neural Network architecture such as Inception-v3, Xception, ResNet50 for feature extraction and Long ShortTerm Memory for generating the relevant captions. Among these models, the best combination based on the accuracy of the model. It is trained using the Flicker8k dataset.

Keywords: Object detection, Image captioning, Deep neural networks, Semantic-instance segmentation