Review Paper Image Caption Generator Using an Artificial Intelligence

Vijeta Sawant, Samrudhi Mahadik, Prof. D. S. Sisodiya

Department of Artificial Intelligence & Data Science

ISBM College of Engineering, Nande, Pune, India

sawantvijeta02@gmail.com and samrudhimahadik2002@gmail.com

samrudhimahadik2002@gmail.comand dharamveersisodiya4@gmail.com

Abstract: *An extremely fascinating piece of technology is an image caption generator, which automatically creates insightful explanations for photos using sophisticated algorithms and machine learning. It's similar to having a personal AI photographer for your images. This is how it operates. The photograph is first sent into the algorithm, which examines the items, people, and scenes in the picture as well as its visual content. The system then creates a caption that explains what's happening in the picture based on this analysis.*

An image caption generator seeks to precisely convey the main idea of the picture while offering a succinct and insightful explanation. It can be quite beneficial for improving accessibility for those with visual impairments as well as for organizing and classifying sizable image collections. Regardless of your intention if you're seeking to explore the possibilities of this technology or add context to your own images, an image caption generator might be a fun tool to experiment.

Keywords: CNN, RNN, LSTM, Artificial Intelligence

REFERENCES

[1] P. Anderson, X. He, C. Buehler, D. Teney, M. Johnson, S. Gould, and L. Zhang - Presented aat CVPR 2018

[2] A. Aker and R. Gaizauskas. Generating image descriptions using dependency relational patterns. In ACL, 2010.

[3] D. Bahdanau, K. Cho, and Y. Bengio. Neural machine translation by jointly learning to align and translate. arXiv:1409.0473, 2014. 1, 2

[4] K. Cho, B. van Merrienboer, C. Gulcehre, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In EMNLP, 214. 1, 2, 3

[5] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. In ICML, 2014. 3

[6] D. Elliott and F. Keller. Image description using visual dependency representations. In EMNLP, 20. 2