Review Report

Image caption generation

Kusa Anjana and M. Swetha

Corresponding Author:
anjanakusa4@gmail.com

ABSTRACT
Image caption generation is a multidisciplinary task at the intersection of computer vision and natural language processing, which aims to automatically produce descriptive and coherent textual descriptions for given images. This process involves extracting meaningful visual features from images using techniques such as convolutional neural networks (CNNs), followed by generating relevant captions through language models, often utilizing recurrent neural networks (RNNs) or transformer architectures. Image captioning has significant applications in accessibility for visually impaired individuals, image retrieval, and content summarization. Recent advances leverage attention mechanisms and large-scale datasets to improve the accuracy and contextual relevance of generated captions, making the technology increasingly effective in understanding and describing complex visual scenes..

[1]Pursuing - MCA, [2]Assistant Professor- Department of CSE,
[1,2] Vaagdevi Engineering College, Warangal, Telangana, India.