|  |  |  |  |
| --- | --- | --- | --- |
| Title | A Survey on Bengali Image Captioning: Architectures, Challenges, and Directions | | |
| Author(s) Name | Shamima Sukhi, Abu Quwsar Ohi, Md Saifur Rahman, MF Mridha | | |
| Contact Email(s) | firoz.mridha@aiub.edu | | |
| Published Journal Name | International Conference on Science & Contemporary Technologies (ICSCT) | | |
| Type of Publication | Conference | | |
| Volume |  | Issue |  |
| Publisher | IEEE | | |
| Publication Date | 2021/8/5 | | |
| ISSN |  | | |
| DOI | 10.1109/ICSCT53883.2021.9642554. | | |
| URL | https://ieeexplore.ieee.org/abstract/document/9642554 | | |
| Other Related Info. |  | | |
|  | | | |

|  |  |
| --- | --- |
| Abstract |  |
| The image captioning process generates human-like captions for an image. In recent years of an appeared research area, automatic image captioning is acquiring more attention. Despite being a challenging problem, several approaches have been proposed where deep learning methods can handle these problems for image captioning and proved to be state-of-the-art. Most of the methods use the convolutional neural network (CNN) and recurrent neural network (RNN) variants. In this survey, we examine various approaches to Bangla image captioning. We investigate image captioning datasets, classify all image captioning systems, review, and represent all of the existing models’ methodologies. We give an overview of the evaluation metrics and provide a comparison of the present methods. Moreover, we discuss their strengths and limitations. Finally, for automatic image captioning, we investigate additional possibilities and future directions. | |