Image Caption Generator

Purpose:

The purpose of this project is to generate a suitable caption of a given image. The caption should describe the whole scenario of the image in a complete sentence.

Background of Study:

People with weak eyesight or no eyesight have a difficult time to know about whats happening around them. A system that gives a brief description for images are going to be a great help for them and eliminate the accessibility barriers. The VizWiz dataset contains captions with complete description of a scene from everyday life that answers specific needs that might help blind users attain certain tasks.

Dataset:

The VizWiz dataset consists of images with task-oriented information. It contains 39,181 images, each pared with 5 captions. The dataset has 23,431 training images with 117,155 training captions, 7,750 validation images with 38,750 validation captions and 8,000 test images with 40,000 test captions.

Figure: Snapshot of the dataset

Problem Statement:

- Extract feature from the image using CNN and deep learning techniques
- Using those features generate single line caption through NLP

Methodology:

Image processing and deep learning methods will be used to extract features from the image and then the information would be converted to relevant sentence explaining the features and the scenario of the image using NLP with a aim of creating improved assistive image captioning system.

References:

Dataset: Image Captioning – VizWiz