Image Caption Generator

Image captioning is the task of generating human-like descriptions or captions for images. Our project aims to leverage the power of deep learning models to automatically analyze visual content and generate accurate and contextually relevant captions. We aim to use two models, VGG16 + LSTM and Inception V3 + Bahdanau Attention, for performance comparison.

The codes (both in oops format and colab code which was developed initially) are provided in the repository with the names BaselineModel.py, AttentionModel.py, Baseline (not in oops), and Attention (not in oops). In order to try the training of our models from Github, Github actions have been utilized. The dataset files (zip) are uploaded in a drive folder that is set to public access, reducing the user's manual work by eliminating the requirement to download them and carry out additional tasks when training from Github. Here is the link for the dataset. To train our models using the GitHub actions, follow these steps:

1. Click on the “Actions tab” in the Github repository
2. Click on the workflow name, ImageCaptionGenerators, available on the left pane
3. Once the workflow page opens, you will see a dropdown menu labeled "Run workflow." Click on it
4. In the dropdown menu, set the respective values for the input parameters:
    a. Model Type: LSTM or Attention
    b. Epoch Number: 2, 5, 10, 20, 50, 70, 100
5. Batch Size: 5, 10, 14, 20, 32, 64, 128
6. The workflow will be triggered manually, and you can refresh the page to see the log of the current workflow run. On the log page, you will find a "Build" tab. Click on the "Build" tab to access the steps of the workflow
7. The log will provide a detailed description of each step being executed, starting from downloading the dataset, preprocessing, training the model for the specified number of epochs and batch size, and displaying the training results
8. Once the workflow job finishes, you will see the build logo in green color, indicating the successful training of the model from the GitHub repository