**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 7 November 2024 |
| Team ID | 739889 |
| Project Title | Image Caption Generator |
| Maximum Marks | 6 Marks |

**Preprocessing Template**

The images will be preprocessed by resizing, normalizing, augmenting, denoising, adjusting contrast, detecting edges, converting color space, cropping, batch normalizing, and whitening data. These steps will enhance data quality, promote model generalization, and improve convergence during neural network training, ensuring robust and efficient performance across various computer vision tasks.

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | This stage would involve providing a description of the image dataset that the captioning model will be trained and evaluated on. |
| Resizing | Images in the dataset would be resized to a uniform target size to ensure consistent input for the model. |
| Normalization | Pixel values of the images would be normalized to a specific range, which can help in stabilizing the training process and improving model performance. |
| Data Augmentation | Techniques such as flipping, rotation, shifting, zooming, or shearing would be applied to artificially increase the size and diversity of the training data. |
| Denoising | Filters might be applied to reduce noise within the images, potentially leading to cleaner feature extraction. |
| Edge Detection | Algorithms would be used to identify and highlight significant edges in the images, which can provide important structural information. |
| Color Space Conversion | Images could be converted from one color space to another, possibly to emphasize certain color characteristics or simplify the input. |
| Image Cropping | Crop images to focus on the regions containing objects of interest. |
| Batch Normalization | Apply batch normalization to the input of each layer in the neural network. |
| **Data Preprocessing Code Screenshots** | |
| Loading Data |  |
| Extracting Images |  |
| Mapping Image to Captiona |  |
| Preprocessing the Text |  |
| Tokenizer |  |
| Model Architecture |  |
| Generate Cption |  |
| Extracting the model |  |