

Data Collection and Preprocessing Phase

Date	11 November 2024
Team ID	739849
Project Title	Pixelprose - crafting visual stories with intelligent image captioning
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 step involves understanding the characteristics of the dataset, including the number of images, their dimensions, formats, and labels. It also includes analyzing class distributions and identifying potential imbalances or missing data to ensure the preprocessing steps are tailored to the dataset's needs.
Resizing	Images in the dataset are resized to a consistent dimension to ensure uniformity across all inputs. This is essential for deep learning models, which require fixed-size inputs. Resizing ensures compatibility with the model architecture and reduces computational complexity.
Normalization	Pixel values of images are scaled to a specific range (e.g., 0-1 or -1 to 1) to improve the model's convergence during training. Normalization reduces the impact of varying pixel intensity scales and helps the model learn faster and more effectively.
Data Augmentation	Data augmentation artificially increases the diversity of the dataset by applying transformations like rotation, flipping, scaling, cropping, or adding noise. This step helps prevent overfitting and improves the model's ability to generalize to unseen data.
Denoising	Unwanted noise in images is removed using filters or algorithms like Gaussian blur or median filtering. This step improves the clarity of image features and enhances the model's ability to focus on relevant visual information.

Edge Detection	Edge detection techniques, such as Sobel, Canny, or Laplacian filters, are used to highlight the boundaries of objects within images. These features can provide additional insights into the structure and shapes within the image, aiding in feature extraction.
Color Space Conversion	Images are converted between different color spaces (e.g., RGB to grayscale or HSV) depending on the task requirements. This step can simplify the model's processing by reducing redundant information or emphasizing specific features.
Image Cropping	Cropping removes unnecessary parts of an image to focus on the region of interest. This step ensures the model trains on relevant data, improving performance by eliminating irrelevant background information.
Batch Normalization	Batch normalization will be applied during training to normalize the activations within layers, improving convergence and generalization of the model.