SAIDL-CV

Generating Binary Masks

The initial task involved generating binary masks by extracting the coordinates of the polygon from a JSON file and matching them with corresponding images stored in a separate folder. To accomplish this, I uploaded the entire dataset, including the images and JSON file, to Google Drive. Next, I mounted the Drive in Colab and iterated through each image in the dataset. I extracted the relevant phrases associated with each image and renamed the corresponding mask using the format 'img_id_phrase', effectively creating the binary masks from the available data.

During this process, I encountered errors with some images due to the presence of forward slashes ('/') in their phrases. These slashes caused file path issues and resulted in "file not found" errors. I handled these errors separately to ensure the smooth progression of the task.

Creating Datasets

After generating the binary masks, I proceeded to create a dataloader using the images stored in the mask folder. These binary masks represented the segmentation data required for further analysis and model training.

Building Model

To build the model, I imported the necessary requirements from the transformers model, specifically ClipSeg. I utilized the same architecture as Clip and instantiated the model using CLIPSegForImageSegmentation. Subsequently, I trained the model using a combination of BCE Loss and Dice Loss while optimizing it with Stochastic Gradient Descent. Upon evaluation, I observed that the Dice Loss outperformed other loss functions on the dataset.