Explanation and Documentation for Building Count Task

Objective:

The primary goal of this task is to develop a pipeline that accurately counts the number of buildings in a provided drone image.

Task Overview:

The task involves processing a drone image to detect and count all the buildings present. The pipeline needs to implement either classical computer vision techniques or leverage machine learning models like YOLO (You Only Look Once) for object detection. The final output should include the total count of buildings in the image, along with an annotated image where the detected buildings are labeled.

Approach:

1. **Model Selection:**
    - I chose to use pre-trained YOLOv8 models for building segmentation and detection due to their efficiency in object detection tasks and their proven success in handling complex images like those captured by drones.
    - The models considered include:
        - spacenetv2.pt: A medium model fine-tuned on SpaceNetV2 for instance segmentation.
        - buildingn-flkennedy.pt: A small model designed specifically for building segmentation.
    - After evaluating the models, I found that spacenetv2.pt provided the most accurate building count, detecting 1258 buildings, while buildingn-flkennedy.pt detected 606 buildings.
2. **Data Processing:**
    - **Image Preprocessing:**

        - The provided image was first read using OpenCV and converted from BGR to RGB format for accurate processing.
    - **Model Inference:**
        - I utilized the AutoDetectionModel class from the SAHI (Slicing Aided Hyper Inference) library to load the chosen model (spacenetv2.pt).
        - The image was sliced into smaller patches (400x400 pixels) with overlapping regions to ensure that buildings near the edges of the slices were not missed during detection.
        - Predictions were made on these slices, and the results were aggregated to count the total number of buildings in the entire image.

3. **Post-processing:**
    - **Filtering Predictions:**
        - Only predictions with a confidence score greater than 0.45 were considered to minimize false positives.
    - **Annotating the Image:**
        - Bounding boxes were drawn around each detected building, and masks were applied to highlight the buildings on the image.
4. **Output:**
    - The final output consists of the total count of buildings in the image (1258 buildings detected using spacenetv2.pt).
    - An annotated image was generated where each building is labeled with a bounding box and highlighted using a mask.

Challenges and Solutions:

- **Model Selection:**

    - Selecting the appropriate model for this task was critical. Initially, several models were considered, but spacenetv2.pt and buildingn-flkennedy.pt provided the best results. Despite this, there was a significant difference in the building counts between the models, highlighting the importance of model choice.

- **Image Slicing:**
    - Slicing the image was necessary due to the large size of the drone image. Overlapping slices helped in ensuring that buildings on the edges of slices were accurately detected.
- **Confidence Thresholding:**
    - Adjusting the confidence threshold was key to balancing between detecting all buildings and minimizing false positives.

Recommendations for Improvement:

- **Custom Model Training:**
    - For more accurate results, it is recommended to train a custom model specifically for this task. A model fine-tuned on a dataset of drone images with annotated buildings would likely yield better performance in terms of both detection accuracy and consistency.

Submission Guidelines:

- **Source Code:** The provided code includes all necessary steps for loading the image, processing it using the chosen YOLOv8 model, and generating the annotated image.
- **Explanation:** The approach, techniques used, and challenges encountered have been documented in detail.
- **Annotated Image and Building Count:** The final output includes the annotated image with buildings labeled and the total building count.

This documentation should provide a clear understanding of the steps taken and the reasoning behind the choices made in this task.