Problem Statement

To perform instance segmentation on a custom dataset.

Architecture

Mask R-CNN because it is state of the art model for 2D instance segmentation.

Loss Functions

- Classification Loss: Cross-entropy for object classification.
- Bounding Box Regression Loss: Smooth L1 loss to refine box coordinates.
- Mask Loss: Binary cross-entropy to predict pixel-wise masks.

Code

- For this task I chose to work with the 2D information (RGB image, 2D bounding boxes) and implement a model to carry out 2D instance segmentation.
- For evaluation, I chose to compare the model predictions with ground truth using IoU-based metrics i.e. Average Precision (AP) and mean Average Precision (mAP).

Training Script:

- Prepares the dataset by converting 3D bounding boxes into 2D projections.
- Implements the dataloader class to handle dataset loading, including RGB images, masks, and point clouds.
- Utilizes the dataLoader for batching and Mask R-CNN for segmentation.
- Logs training progress, including warnings for invalid targets and loss values.
- Saves the trained model after completing the training loop.

Inference Script:

- Loads the trained model.
- Prepares the test dataset and performs inference.
- Calculates performance metrics such as Average Precision (AP) at multiple IoU thresholds.
- Logs the results, including warnings for missing predictions or errors in processing batches.