

STAT 605 - COCO

Dataset study on

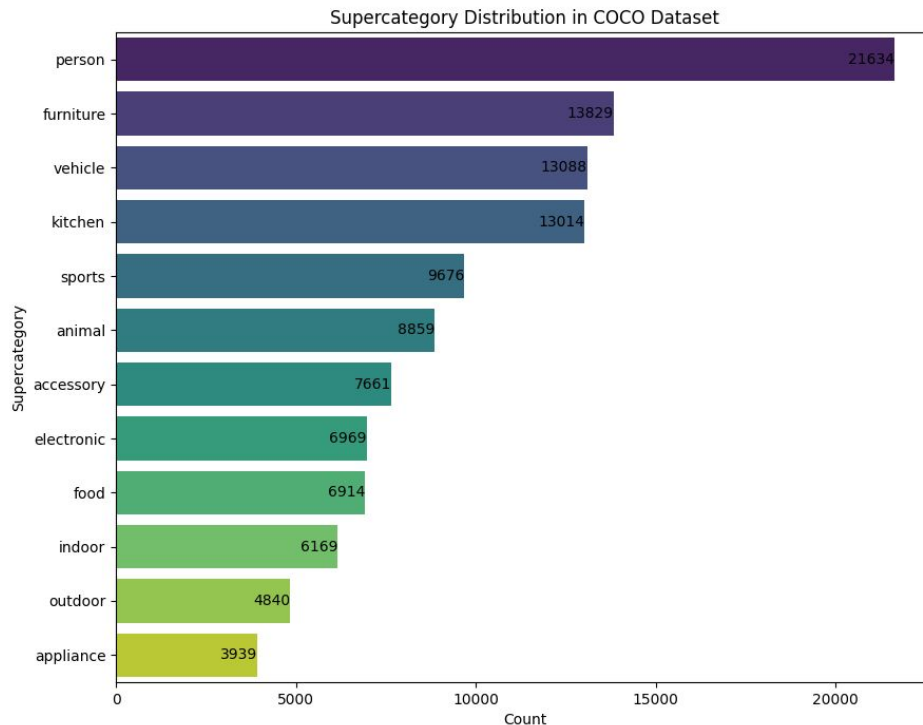
HPC

- Pushpit, Shantam, Kanishk,
Shourya, Dave
- 

Introduction

Leveraging the power of COCO dataset - 27 GB of pure delight.

- **Object Categories:** The dataset consists of 80 distinct object “categories” and 12 distinct “supercategories”, including common objects like "person," "car," "dog," and more. Each object category is associated with a unique ID.
- **Annotations:** For each image in the dataset, precise annotations are provided, including bounding box coordinates, segmentation masks, and category labels. These annotations allow for detailed object localization and understanding.
- **Complexity:** The COCO dataset contains images with varying levels of complexity, including multiple objects, occlusions, and diverse backgrounds. This complexity makes it an excellent benchmark for evaluating computer vision algorithms.



Understanding annotations in terms of masks

Original Image



Main Image



Binary Mask



RGB Mask



Instance Segmentation Mask



Custom data generator

Mask Image



Main Image



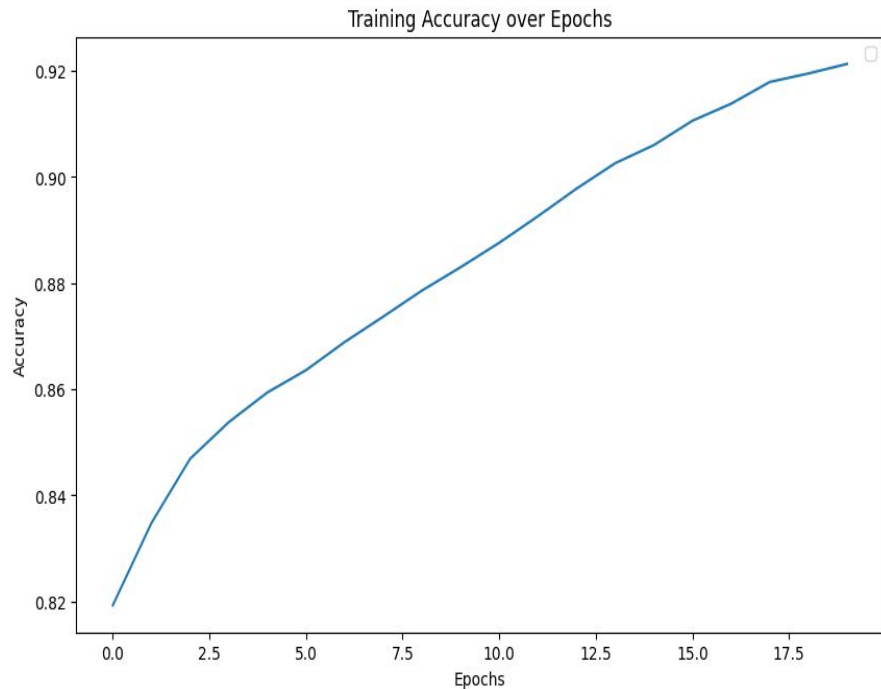
Running a u-net model

The U-Net architecture has several convolutional layers with batch normalization and max pooling to extract features.

It also includes upsampling layers, concatenation of skip connections, and more convolutional layers for the decoder part.

The final layer employs a convolutional layer with softmax activation to generate a probability map for each class.

Epochs training accuracy



Running these 20 epochs took almost 16 hours to run on our local Macbook Air.

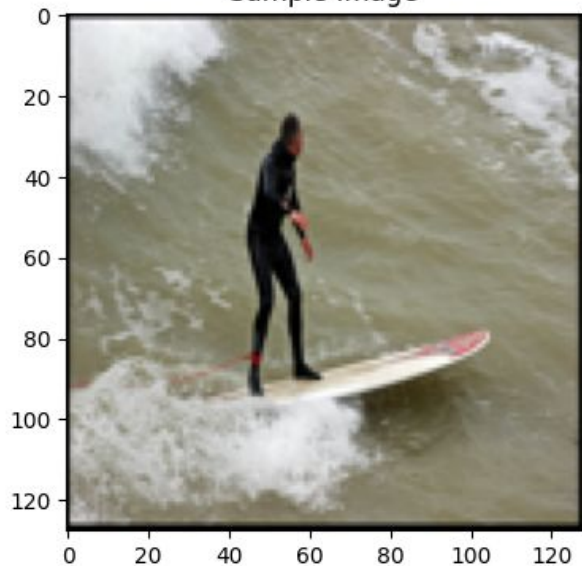


On CHTC, we ran these epochs in ...

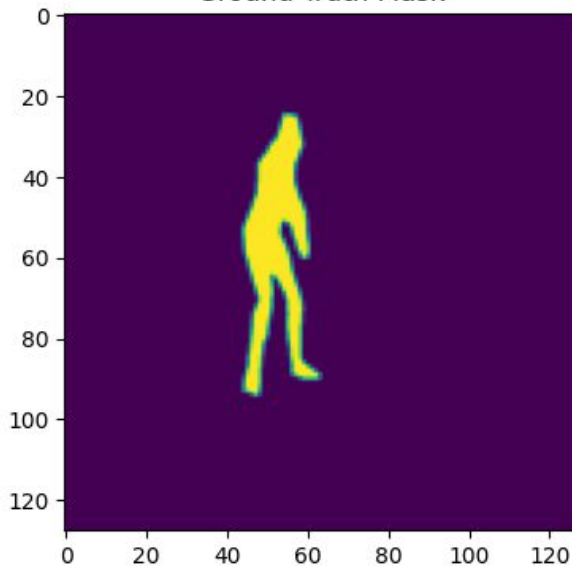
Validation of u-net model

Sample Image, Ground Truth Mask, and Predicted Mask

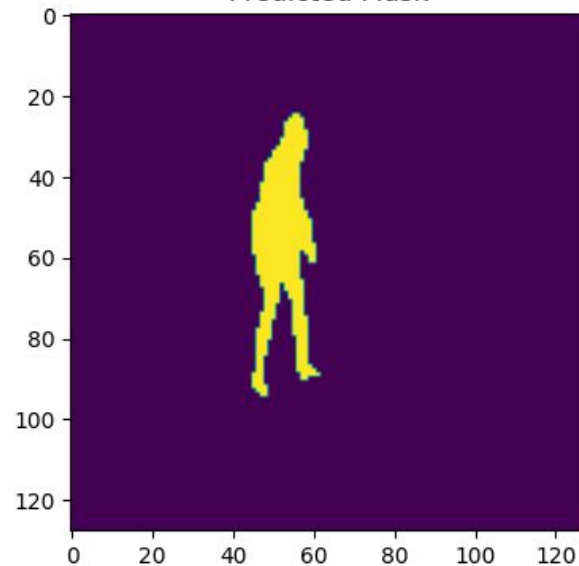
Sample Image



Ground Truth Mask



Predicted Mask



Thank you