A Low Cost Solution to the Open Images Instance Segmentation Challenge

Rohit Midha+, Shraddhaa Mohan+

+SSN College of Engineering, Anna University

Summary

1. **Low Cost:** Only 400$ of Google Cloud credits were used during the challenge.
2. **Expert Models:** Varied models were built to tackle class imbalance and class hierarchies.
3. **1 GPU:** Only 1 Nvidia Tesla T4 was used for training with a total training time of 4 days for all expert models.
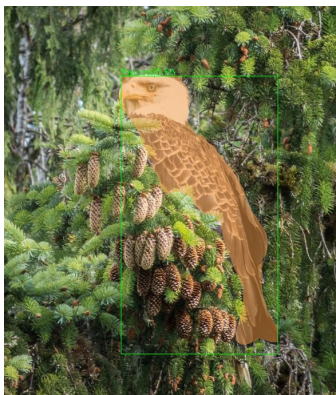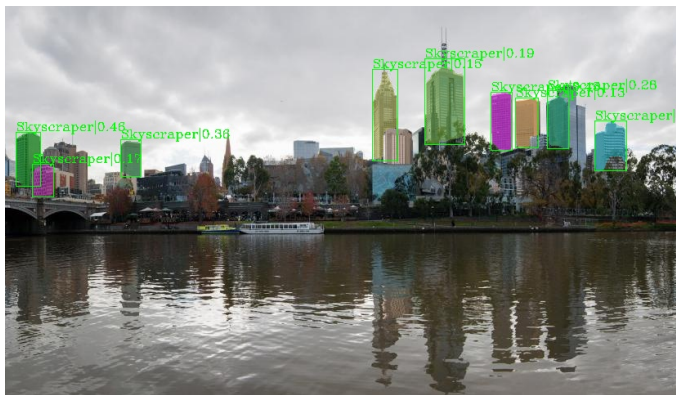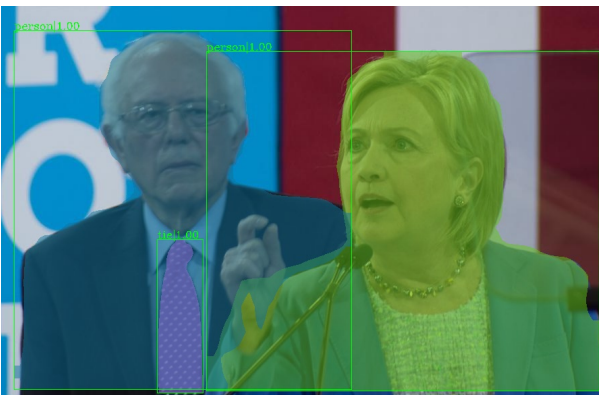
Expert Models

1. **Coco Classes:** Uses pre-trained COCO[1] weights to predict on COCO classes.
2. **Least 40 Classes:** Model that was trained on the classes with least occurrence in train set.
3. **Hierarchical Classes:** Model that was trained on all the classes that are part of a hierarchy further split into expert models with groups of related hierarchies in each expert model.
4. **Remaining Classes:** Model that was trained on all the other classes.

Approach

Our approach used the Hybrid Task Cascade[2] [3] paper as reference, which we slightly modified to work on the Open Images dataset. We had to remove the semantic feature fusion branch for the dataset and we trained a total of 4 models and 4 expert models.

We used the ResNet50[4] architecture and due to lack our hardware limitations we sampled a maximum of 1500 masks for each class for training. We used the original image sizes and images were also augmented during training time using Random Horizontal and Vertical Flips, Varying Brightness and Normalising.

The total training time for all the models on a high memory 8 CPU system and 1 Nvidia Tesla T4 GPU came to about 4 days. We achieved an average train time of 0.9 iterations/second.

Prediction made on (a) COCO classes, (b) Least 40 classes, (c) Hierarchical classes, (d) Remaining classes, (e) Related Hierarchies Expert Model

Results

	Public LB	Private LB
Coco Classes	0.0944	0.0850
Least 40 Classes	0.0280	0.0280
Hierarchical Classes	0.0950	0.0829
Remaining Classes	0.1316	0.1172
Ensemble	0.3450	0.3083
Rank	35	34

Hybrid Task Cascade (semantic feature fusion with box branches is not shown on the figure)
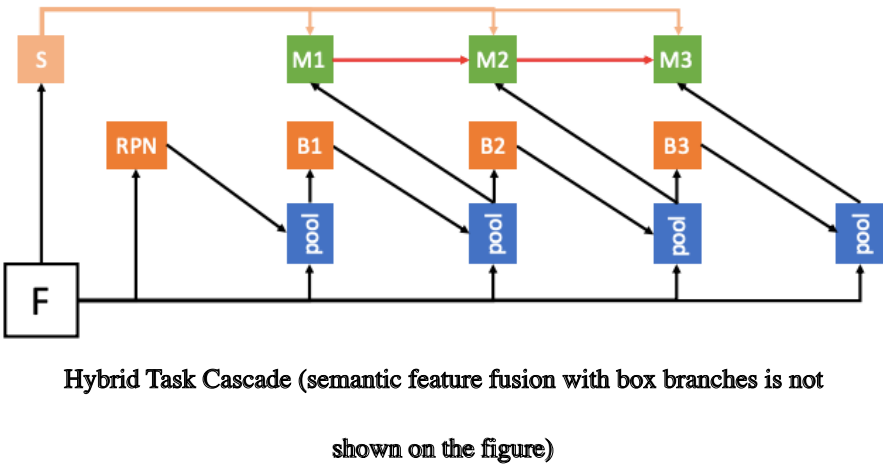
Literature

[1] - Tsung-Yi Lin et al. "Microsoft coco: Common objects in context", 2014.

[2] - Kai Chen et al. "mmdetection", https://github.com/open-mmlab/ mmdetection, 2018.

[3] - Kai Chen et al. "Hybrid task cascade for instance segmentation", 2019.

[4] - He, Kaiming, et al. "Deep residual learning for image recognition.",2015