Motivation: Autonomous driving needs real-time analysis of a driving environment. One approach is image analysis, the AI for which has moved swiftly. These algorithms perform semantic image segmentation of a driving environment, on objects relevant to a driver (e.g. car, pedestrian). This project tests one of the most recent developments, Mask R-CNN (2017), fine-tuned on a dataset for the 2018 CVPR WAD Video Segmentation Challenge. I built a Mask R-CNN implementation that trains on Collaboratory, pulling data from Google Drive, with different training parameters. Results: little to no measurable improvement over baseline COCO. Inputs: images of driving scenes; outputs: the same driving scene, with “mask” applied pixel per pixel.

Dataset

CVPR 2018 WAD Video Dataset; COCO Dataset:
- source: colored video images from car cameras, of vehicle driving environments
- description: these environments are labeled with 36 classes, pixel by pixel, where an output mask is generated for the input image.
- While the COCO dataset has 80+ categories, and the CVPR 30+, I defined 6 categories of interest: car, motorcycle, bicycle, pedestrian/person, truck, and bus.
- This dataset was comprised of 720p annotated images of 30+ classes from Berkeley Deep Drive and Apollo. 93.2 GB of these images were used for a training set, with an additional 4 GB for a dev set and a final private test set.
- Each pixel in the image contains information about object instance, and class; i.e. int(PixelValue/1000) is the labelled class, and PixelValue % 1000 is the instance id.

Implementation

Per Matterport:
- ResNet101 + FPN for CNN "backbone"
- L = L_class + L_box + L_mask
- pre-trained from MS COCO dataset
- fine-tuned head layers with WAD
- Image dimensions: 1024x1024

Configurable Hyperparameters

- learning rate = .001 - .01
- learning momentum: .9 - .8
- Mask Threshold: 32x32, 64x64, 128x128

Discussion

- data management
- compute options
- models, implementations, and choices
- teams, personal failures

Future

- implement the Data class
- confirm every step with other people
- play with other implementations

References:
https://github.com/ Matterport/Mask R-CNN