YOLOPv2: Better, Faster, Stronger for Panoptic Driving Perception

* Multi-tasking learning approaches have achieved promising results in solving panoptic driving perception problems. This provides both high-precision and high-efficiency performance
* Become a popular paradigm when designing networks for real-time practical autonomous driving systems
* Achieved SOTA performance on BDD100K dataset
* Computer vision based tasks area still challenges in applications of low-cost autonomous driving
* Panoptic driving perception system uses cameras and Lidar
* More powerful network is needed to provide a better high-level and low-level feature fusion for lane detection
* Unpractical to run separate models for each individual task in real-time autonomous driving systems
  + Multi-task learning networks provide a potential solution for saving computational cost
  + Where the network is designed into an encoder-decoder pattern, and the encoder is shared by different tasks
* Also on BDD100K dataset

Real-time traffic object detection

* Object detectors can be divided into one-stage and two-stage detectors
* Two-stage consist of the
  + region proposal component
  + detection refinement component

Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition:

Common resources:  
Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE transactions on pattern analysis and machine intelligence, 37(9):1904–1916, 2015.