



Not All Pixels are Equal: Difficulty-aware Semantic Segmentation via Deep Layer Cascade

Xiaoxiao Li, Ziwei Liu, Ping Luo, Chen Change Loy, Xiaoou Tang



Multimedia Lab, The Chinese University of Hong Kong

Problem



Input Video



State-of-the-art Method (4 FPS)



Deep Layer Cascade (17 FPS)

State-of-the-art



State-of-the-art Method (4 FPS)

- Why Slow?

- Very Deep Backbone Network
- High Resolution Feature Map



Fully Convolutional Network

Motivation



Image



Easy Region

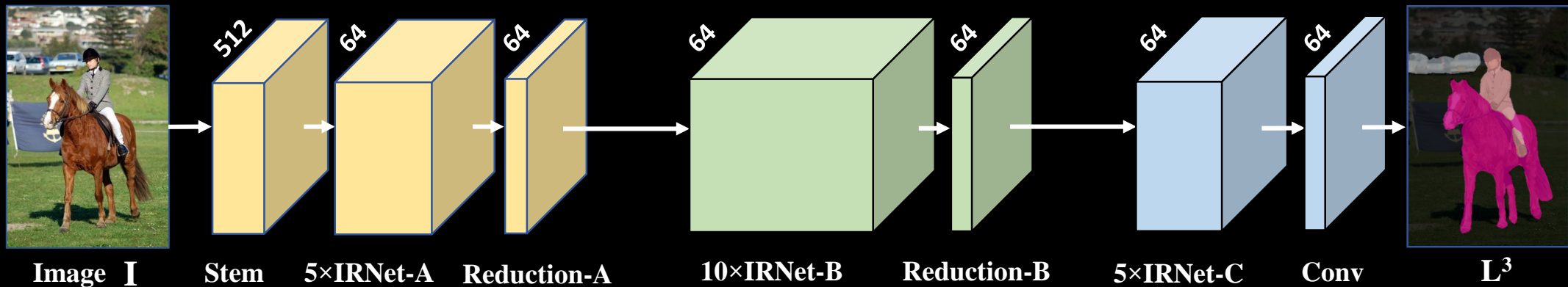
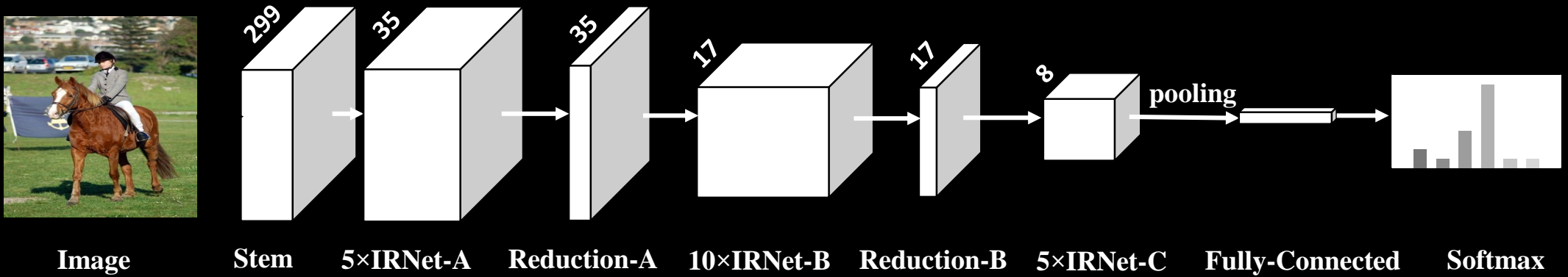


Moderate Region

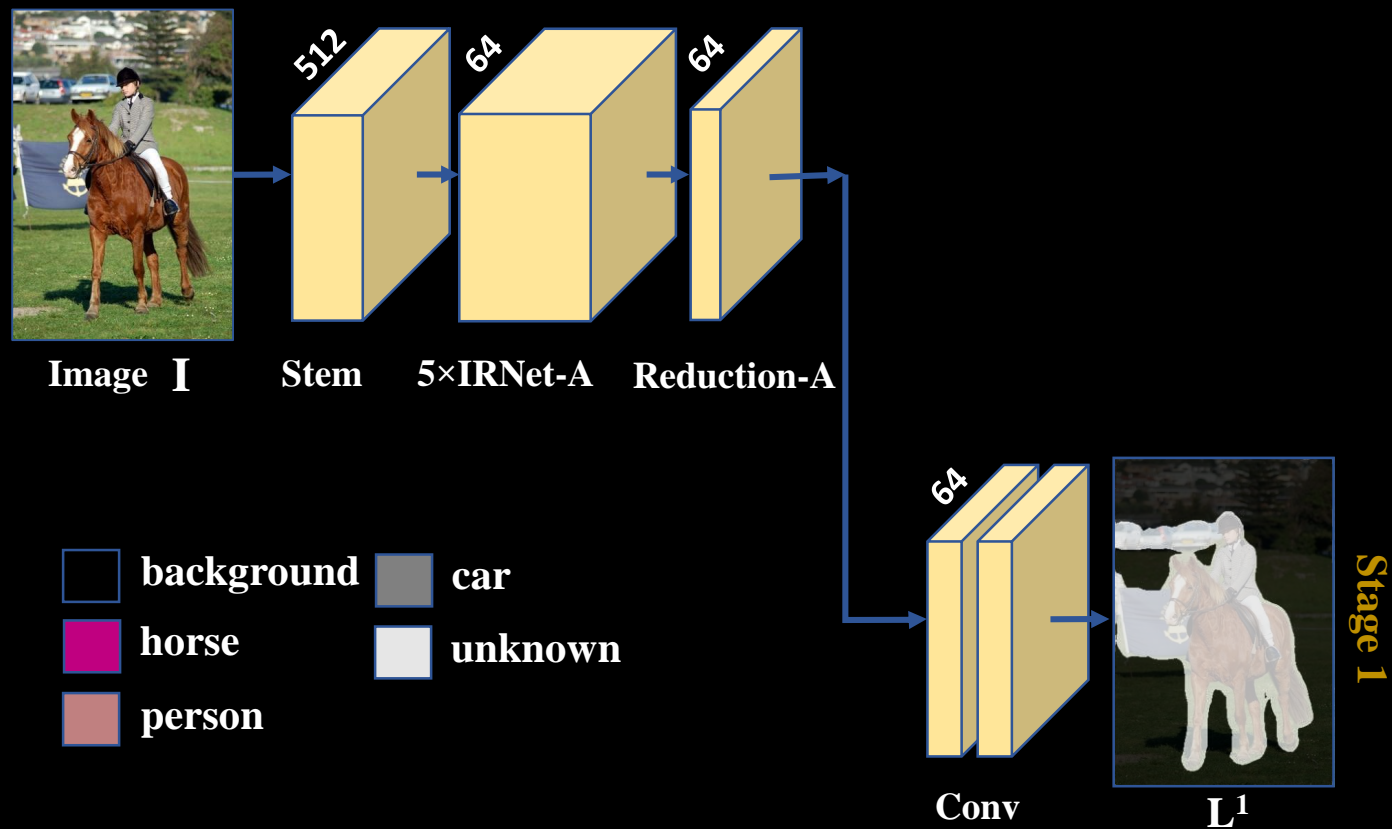


Hard Region

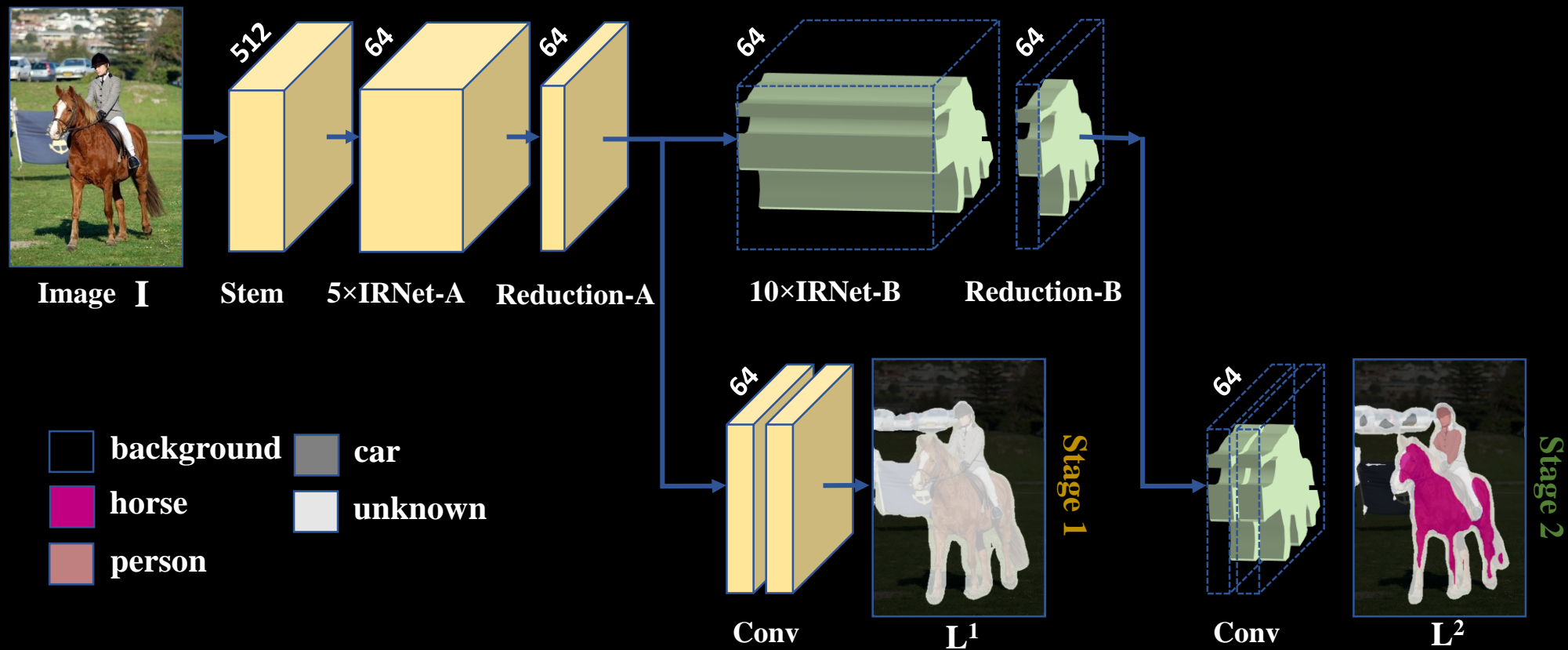
Contemporary Model



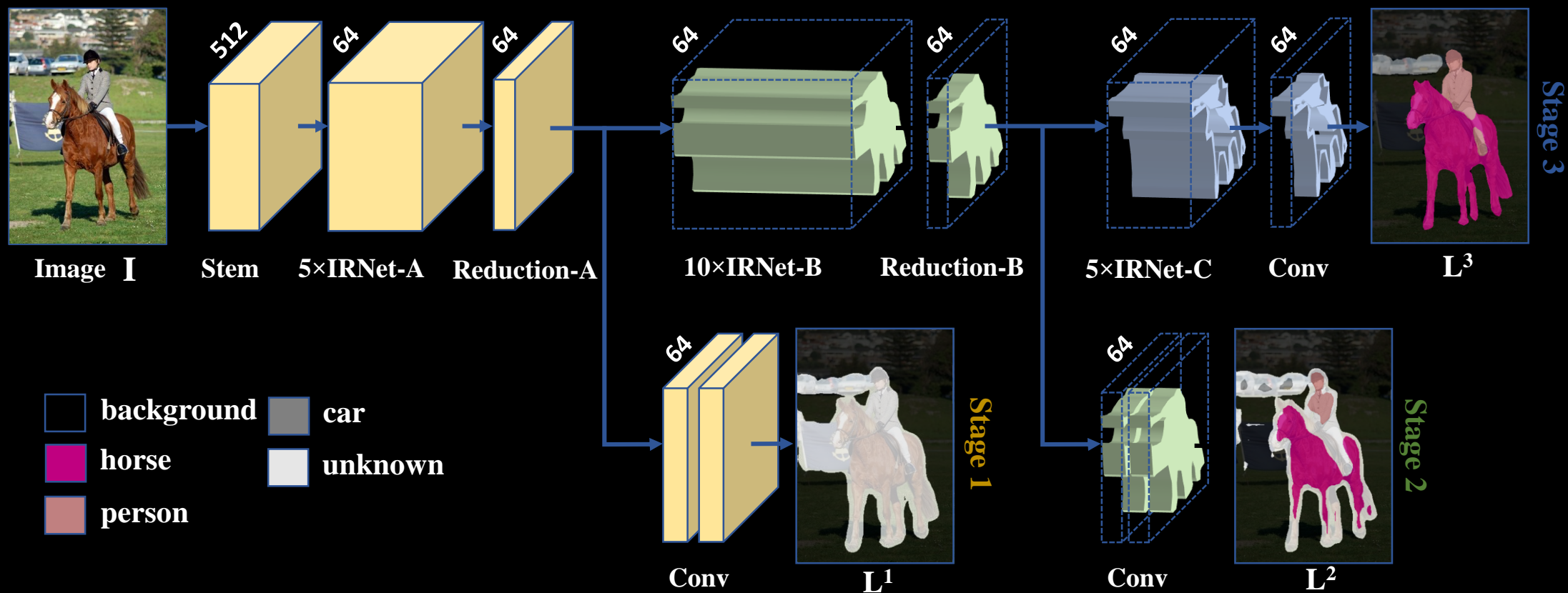
Deep Layer Cascade



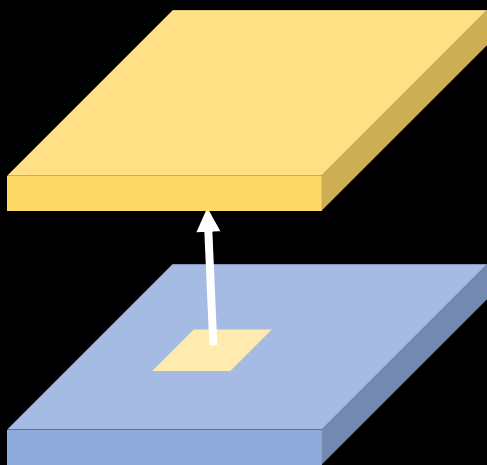
Deep Layer Cascade



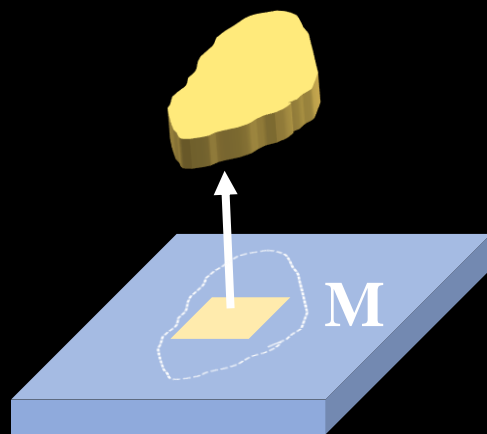
Deep Layer Cascade



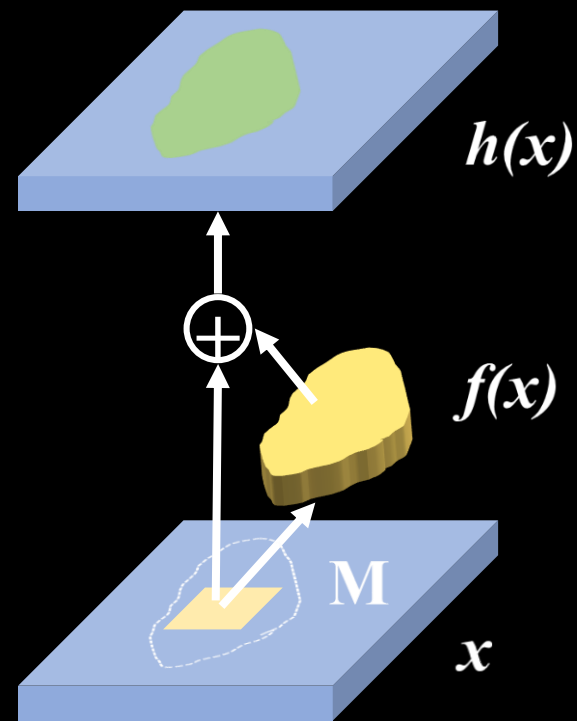
Region Convolution



Convolution



Region Convolution



Region Convolution with Residual

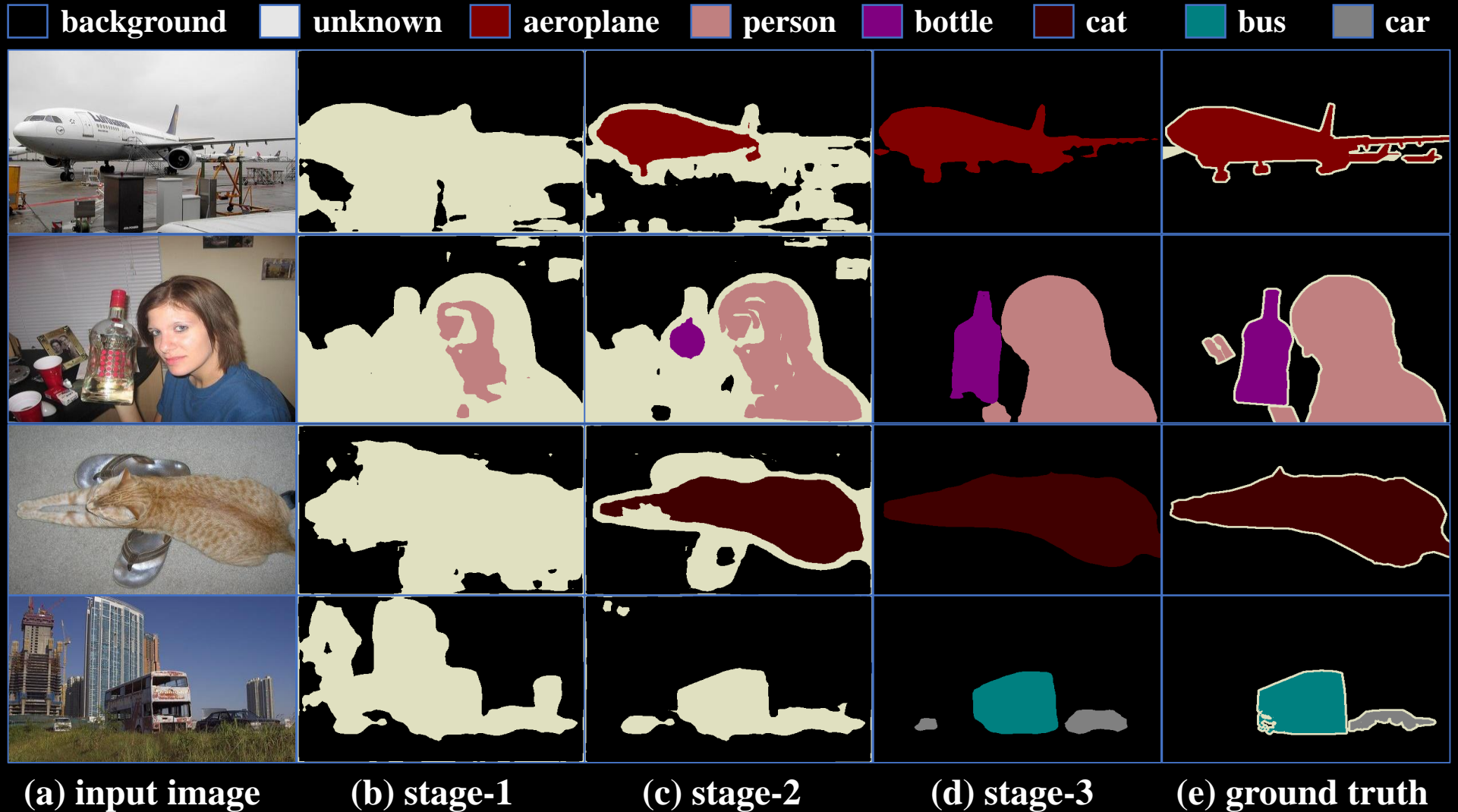
Performance

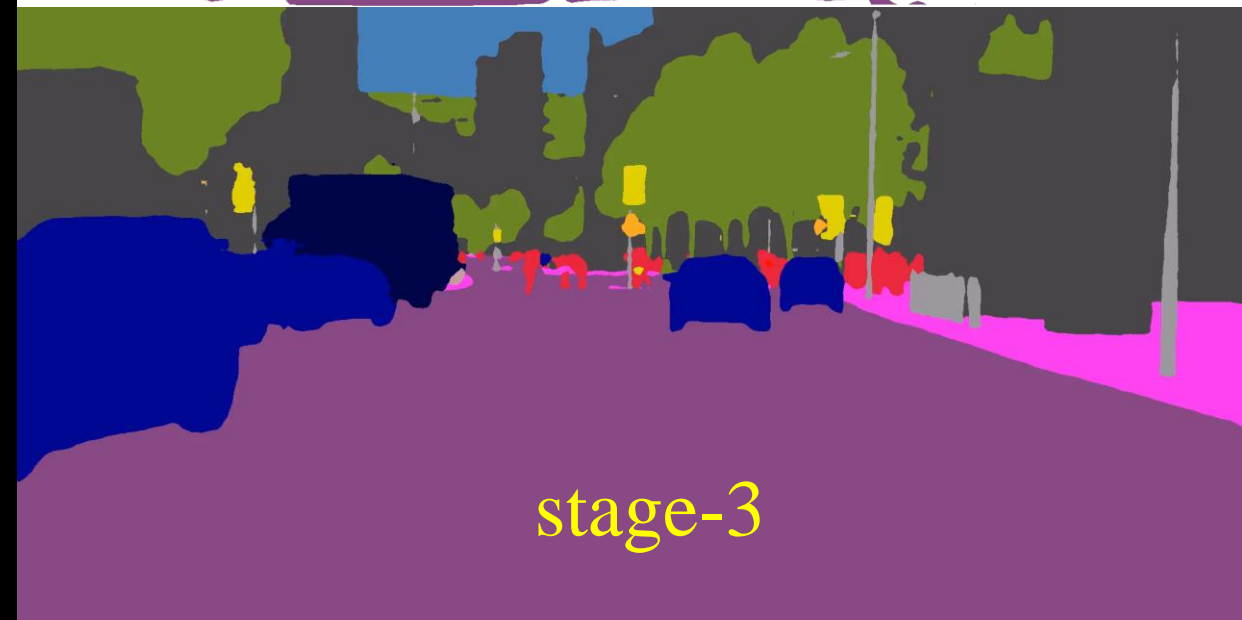
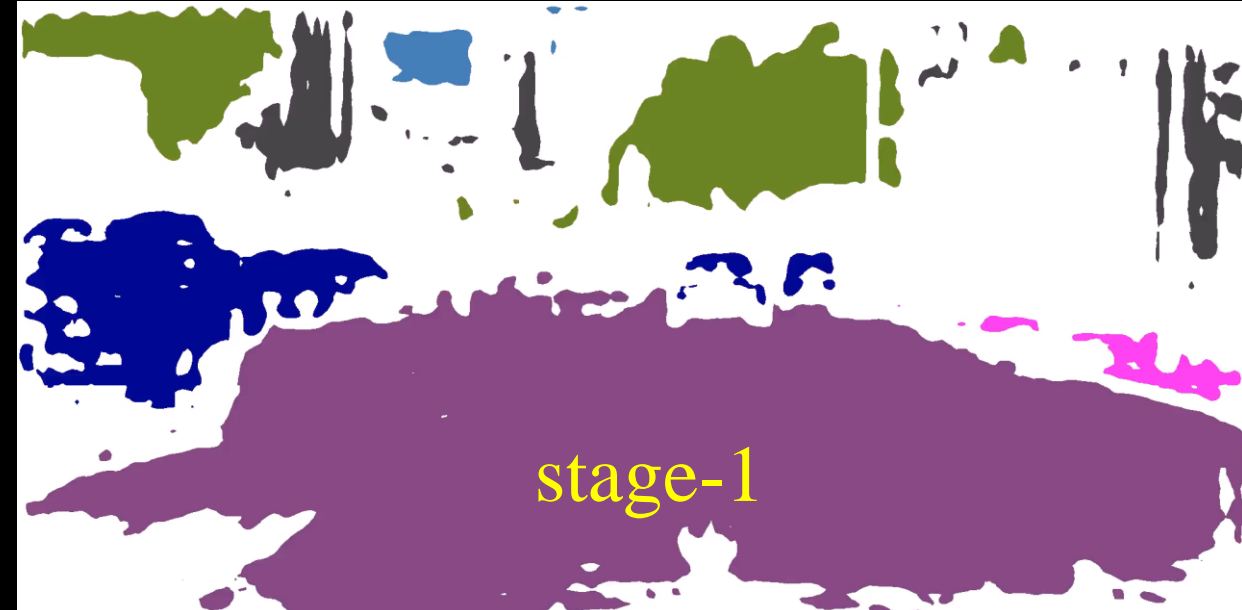
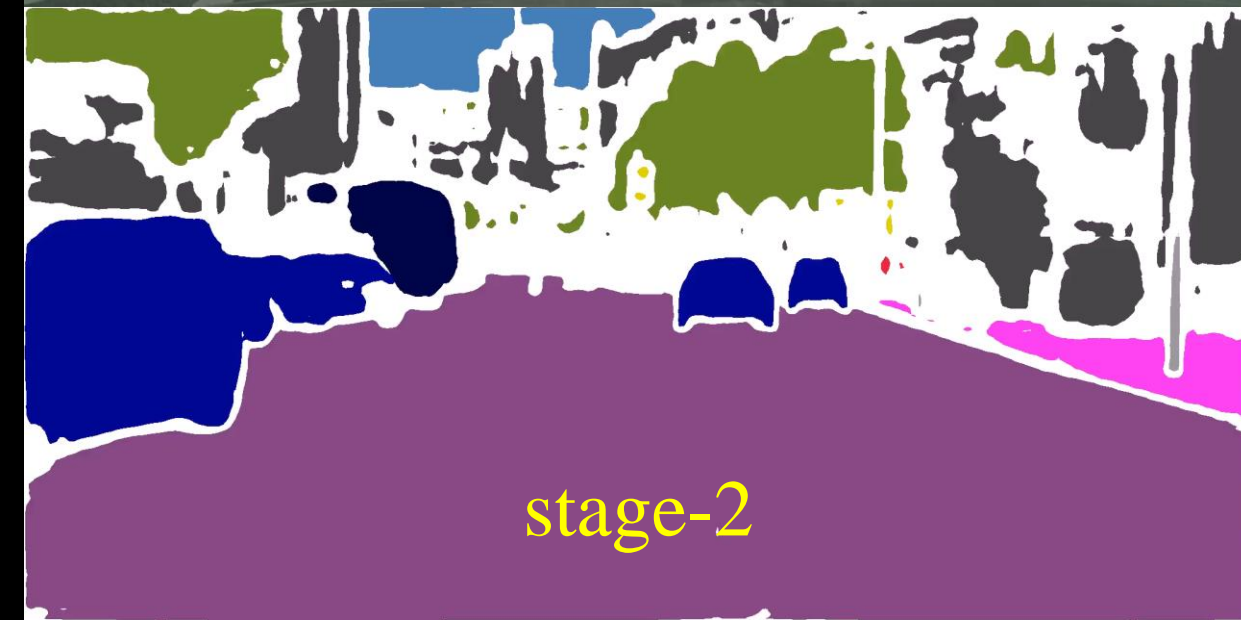
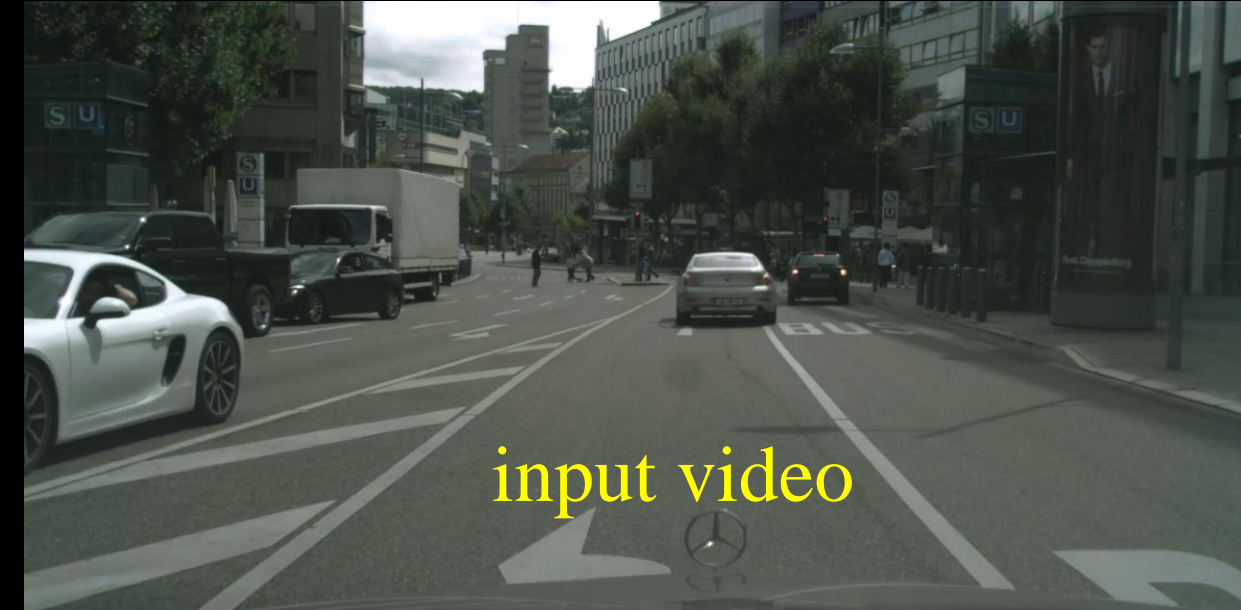
PASCAL VOC 2012

	mIoU	FPS (Backbone Network)
DPN	77.5	5.7
Adelaide	79.1	-
Deeplab-v2	79.7	7.1
LC(w/o COCO)	80.3	14.7
LC(with COCO)	82.7	

(PASCAL VOC 2012 Challenge test set)

Stage Visualization





- Difficulty-Aware Learning Paradigm

- End-To-End Trainable Framework

- Region Convolution → Real-Time

Thanks!

Not All Pixels are Equal:
Difficulty-aware Semantic Segmentation via Deep Layer Cascade

Code and models are available @

Project Page: <http://personal.ie.cuhk.edu.hk/~lz013/projects/LayerCascade.html>