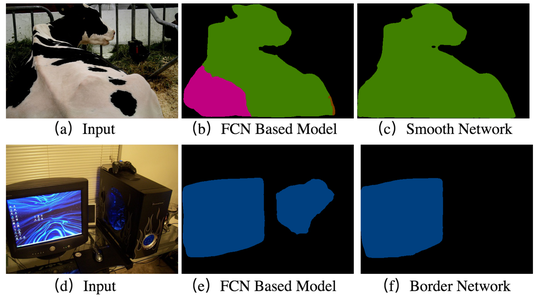
Learning a Discriminative Feature Network for Semantic Segmentation

Changqian Yu[1], Jingbo Wang[2], Chao Peng[3], Changxin Gao[1], Gang Yu[3], and Nong Sang[1]
[1]Huazhong University of Science and Technology [2]Peking University [3]Megvii Inc. (Face++)

CVPR 2018 SALT LAKE CITY • JUNE 18-22

Motivation

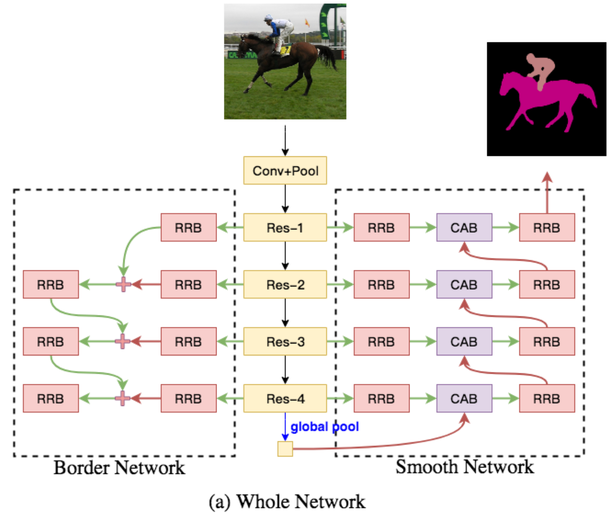


- **Intra-class Inconsistency**: the patches which share the same semantic label but different appearances
- **Inter-class Indistinction**: the two adjacent patches which have different semantic labels but with similar appearances

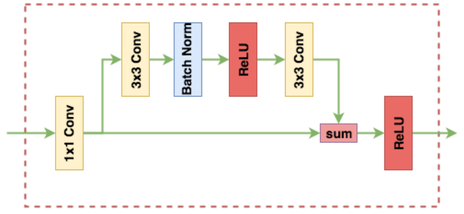
Contributions

- Rethink the task from a macroscopic point of view: to regard the semantic segmentation as a task to assign a consistent semantic label to one category of things, not just at the pixel level
- Smooth Network to enhance the intra-class consistency with the global context and the Channel Attention Block
- Border Network with deep supervision to enlarge the variation of features on both sides of the semantic boundary. This can also refine the semantic boundary of prediction.

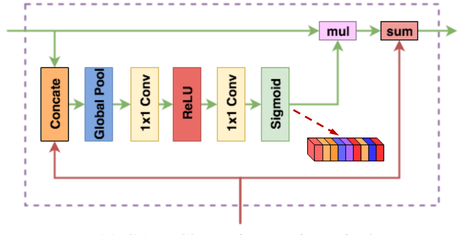
Discriminative Feature Network



(a) Whole Network



(b) RRB: Refinement Residual Block



(c) CAB: Channel Attention Block

Experimental Results

Ablation Study

- **Comparison with base model**

Method	Mean IOU(%)
Res-101$^-$	69.26
Res-101	72.86
Res-101+RRB	76.65
Res-101+RRB+GP	78.20
Res-101+RRB+GP+CAB	79.31
Res-101+RRB+DS	77.08
Res-101+RRB+GP+DS	78.51
Res-101+RRB+GP+CAB+DS	79.54

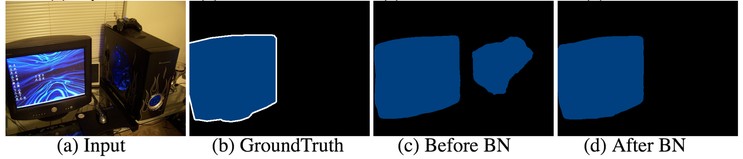
RRB: Refinement Residual Block GP: global pooling
CAB: Channel Attention Block DS: deep supervision.

- **Comparison between Smooth Network and Border Network**

Method	Mean IOU(%)
Res-101+SN	79.54
Res-101+SN+BN	79.67
Res-101+SN+MS_Flip	79.90
Res-101+SN+BN+MS_Flip	80.01

SN: Smooth Network
BN: Border Network
MS_Flip: adding multi-scale inputs and left-right flipped inputs.

Visualization of Smooth Network and Border Network



Visualization of Boundary

Results on PASCAL VOC 2012

Method	Mean IOU(%)
FCN [27]	62.2
Zoom-out [29]	69.6
ParseNet [24]	69.8
Deeplab v2-CRF [5]	71.6
DPN [26]	74.1
Piecewise [20]	75.3
LRR-CRF [11]	75.9
PSPNet [40]	82.6
Ours	**82.7**
DLC$^+$ [18]	82.7
DUC$^+$ [34]	83.1
GCN$^+$ [30]	83.6
RefineNet$^+$ [19]	84.2
ResNet-38$^+$ [35]	84.9
PSPNet$^+$ [40]	85.4
Deeplab v3$^+$ [6]	85.7
Ours$^+$	**86.2**

Results on Cityscapes

Method	Mean IOU(%) w/o coarse	Mean IOU(%) w/ coarse
CRF-RNN [41]	62.5	-
FCN [27]	65.3	-
DPN [26]	66.8	59.1
LRR [11]	69.7	71.8
Deeplab v2-CRF [5]	70.4	-
Piecewise [20]	71.6	-
RefineNet [19]	73.6	-
SegModel [10]	78.5	79.2
DUC [34]	77.6	80.1
PSPNet [40]	78.4	80.2
Ours	**79.3**	**80.3**