Weakly-Supervised Semantic Segmentation Network with Deep Seeded Region Growing

Zilong Huang **, Xinggang Wang,** Jiasi Wang**,** Wenyu Liu**,** Jingdong Wang

www.xinggangw.info

Huazhong University of Science and Technology

Weakly-supervised visual learning (WSVL)

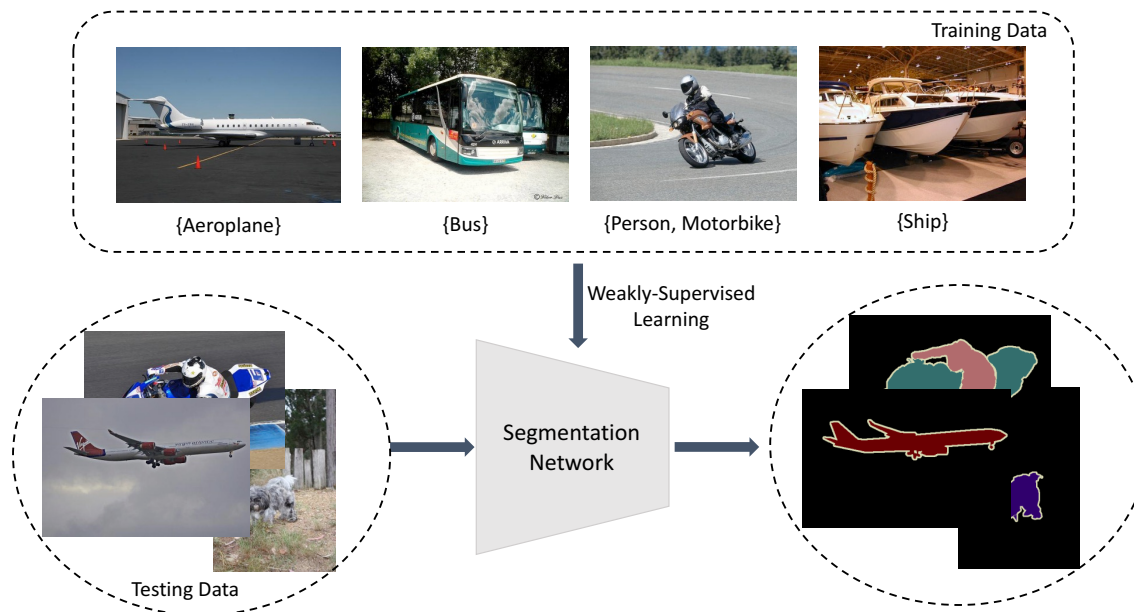☐ **Weakly-supervised visual learning is a new trend in CVPR**

Search keyword "weakly supervised" and "weakly-supervised" in CVPR 17&18

Keyword	Weakly supervised	Weakly-supervised	In total
cvpr17	14	5	19/783
cvpr18	19	10	29/979

Weakly supervised semantic segmentation

☐ ## The task of WSSS

WSSS overcomes the deficiency problem in semantic segmentation labelling.
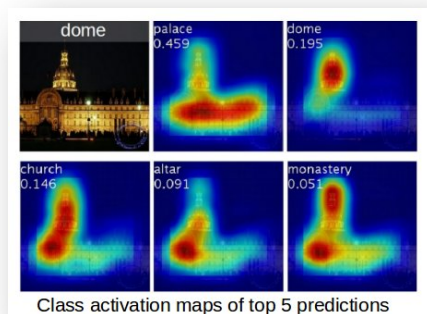
Huazhong University of Science and Technology
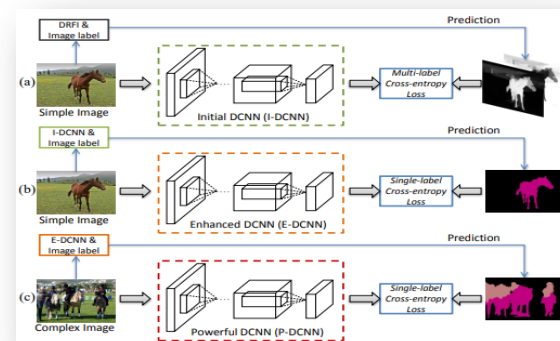
The development of WSSS

$$(x_l, y_l) = \arg \max_{\forall (x,y)} \hat{p}_l(x, y) \qquad \forall l \in \mathcal{L}_I$$

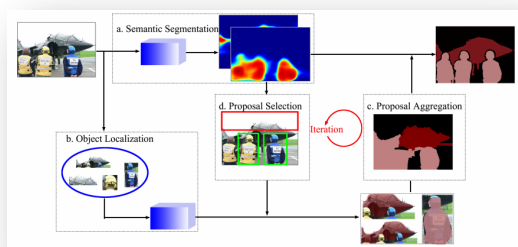$$\text{MIL LOSS} = \frac{-1}{|\mathcal{L}_I|} \sum_{l \in \mathcal{L}_I} \log \hat{p}_l(x_l, y_l)$$

MIL-FCN, Pathak et al, Arxiv 14, ICLRW 15

Class activation maps of top 5 predictions
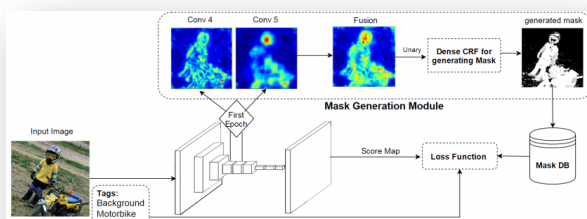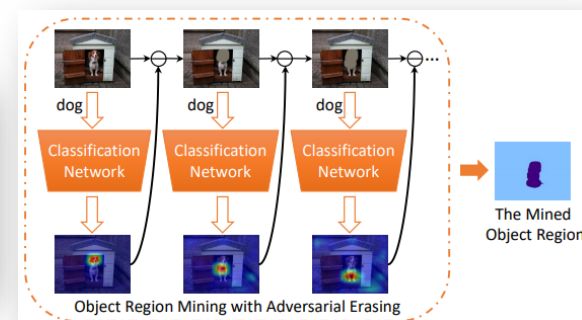
CAM, Zhou et al, CVPR 16

STC, Wei et al, TPAMI 15

Proposal classification, Qi et al, ECCV 16

Built-in FG/BG Model Saleh et al, ECCV 16

Adversarial erasing, Wei et al, CVPR 17

Huazhong University of Science and Technology

Figures are from the original papers
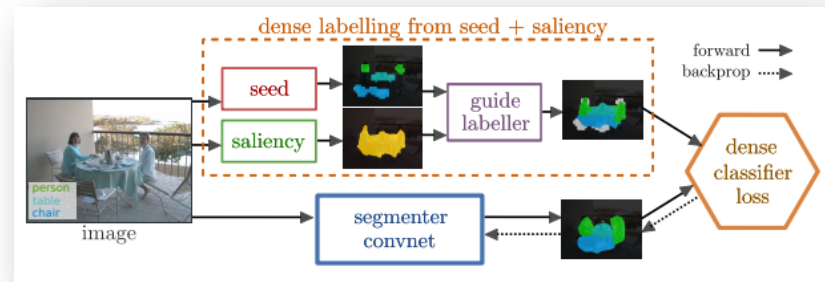
The development of WSSS
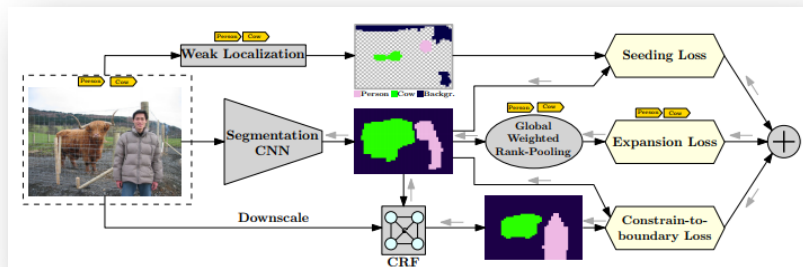
Seeding loss,
Kolesnikov et al, ECCV 17

Saliency guided labler,
Oh et al, CVPR 17

1. **Multi-instance learning**
2. **Saliency guided**
3. **Built-in network information**
4. **Adversarial learning**
5. **Seeding loss**

Figures are from the original papers

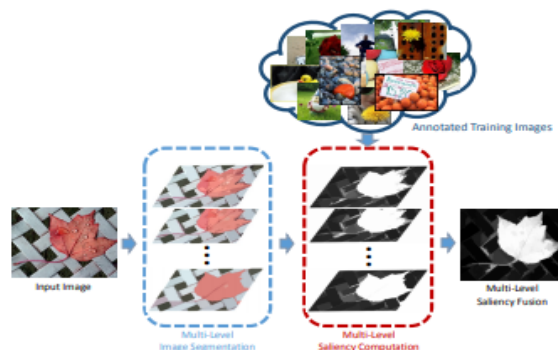# The basic framework in our paper

Step 1: Foreground seeds from CAM
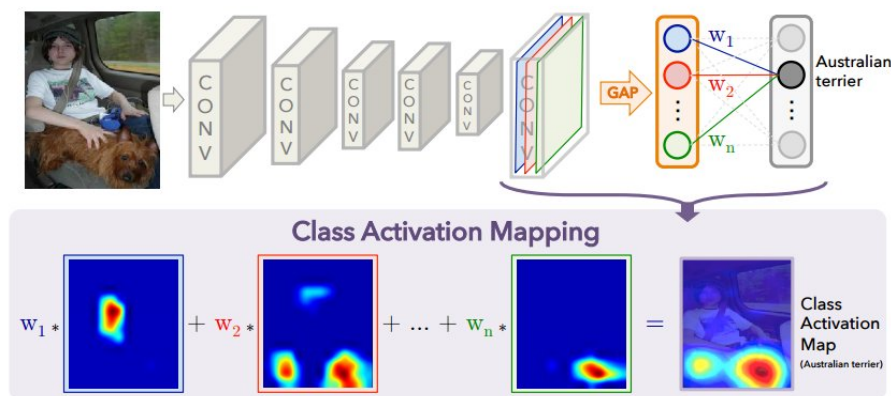
Step 2: Background seeds derived salient region detection [Jiang et al, CVPR13]

Huazhong University of Science and Technology

Figures are from the original papers

The basic framework in our paper

Step 3: FCN with seeding loss

Step 4: Retrain with FCN

Huazhong University of Science and Technology
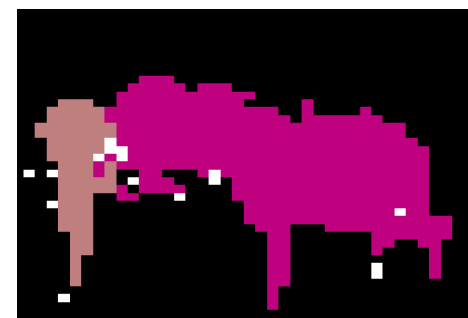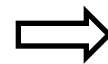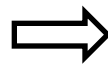
A small trick: balanced seeding loss

Balance the weights between foreground and background

$$\ell_{seed} = -\frac{1}{\sum\limits_{c \in \mathcal{C}} |S_c|} \sum_{c \in \mathcal{C}} \sum_{u \in S_c} \log H_{u,c}$$

$$-\frac{1}{\sum\limits_{c \in \bar{\mathcal{C}}} |S_c|} \sum_{c \in \bar{\mathcal{C}}} \sum_{u \in S_c} \log H_{u,c},$$

However, the seeds are sparse

Image

Seeds

GT
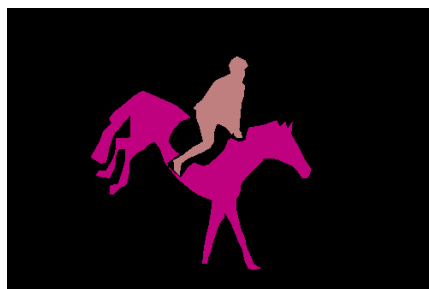
In practice, to retain the precision of seeds, there are about 40% pixels have labels.

How to improve the quality and quantity of seeds

- Better "CAM" network

- Saliency guidance

- Adversarial erasing

- …

- Online seeded region growing

Deep seeded region growing

seeded region growing
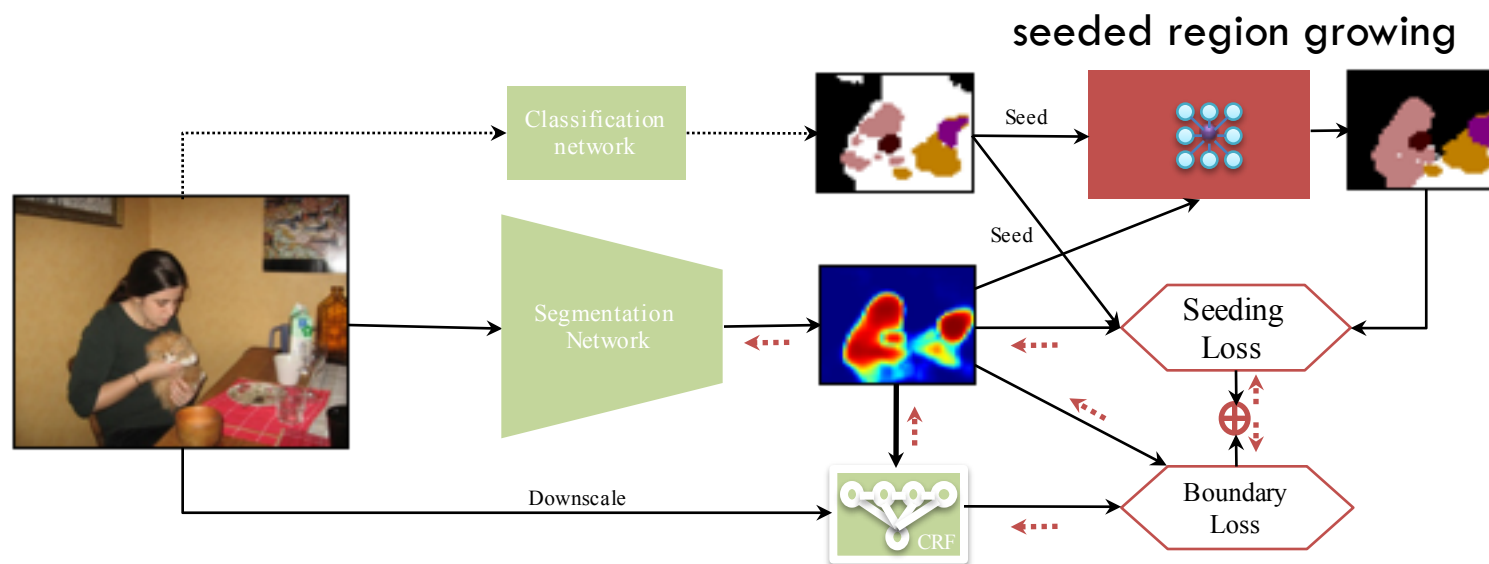
Region growing criteria:

$$P(H_{u,c}, \theta_c) = \begin{cases} \text{TRUE} & H_{u,c} \geq \theta_c \text{ and} \\ & c = \arg\max_{c'} H_{u,c'}, \\ \text{FALSE} & \text{otherwise.} \end{cases}$$

1. **Directly use deep prob features**
2. **Cheap to compute**
3. **Online supervision updating**

Progressively check the neighborhood pixels

Huazhong University of Science and Technology

Deep seeded region growing

| Training Image | Seed Cues | Epoch #1 | Epoch #12 | Ground Truth |
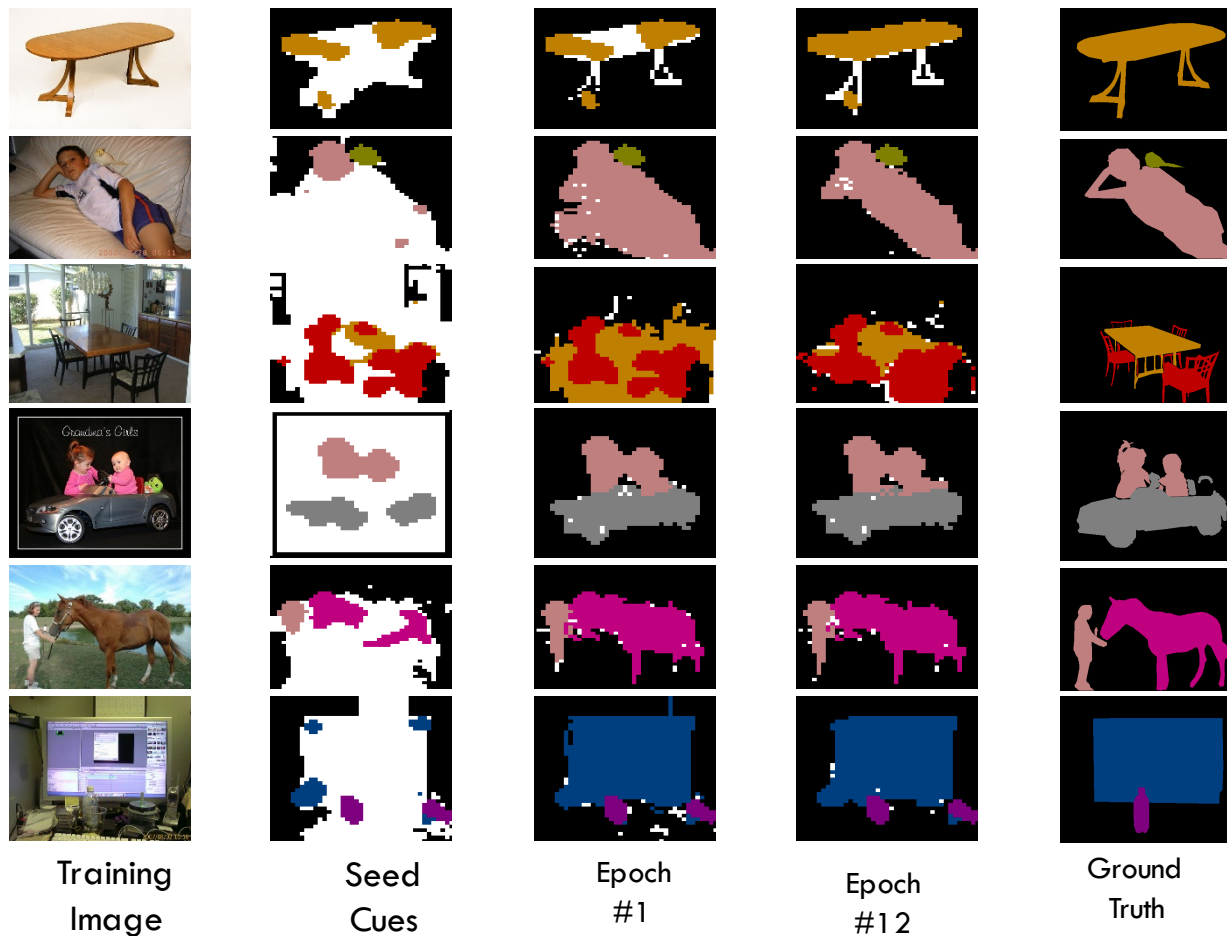
Deep seeded region growing

Algorithm 2 Deep Seeded Region Growing Training

1: **Input:** Training data $D = \{(I_i, S_i)\}_{i=1}^N$.

2: **Initialize:** initialize M_0, $t = 1$.

3: **while** $(t \leq max\_iter)$ **do**

4:     Select a sample $\{I_i, S_i\}$ from input data randomly;

5:     $H_i = M_{t-1}(I_i)$;

6:     Perform $G_i = DSRG(S_i, H_i)$ for seed expansion

7:     Compute the $loss = \ell(G_i, , H_i)$

8:     back propagate the error and update model from M_{t-1} to M_t

9: **end while**

10: **Output:** M

Huazhong University of Science and Technology

Experiments

- Datasets
  - **PASCAL VOC 2012**, 10582 train, 1449 val, 1456 test
  - **COCO**, 80k train, 40k val

- mIoU criterion

- Classification network: VGG-16
- Segmentation network: DeepLab-ASPP

Main Results

PASCAL VOC

Method	Training	Val	Test
DCSM[2]	10k	44.1	45.1
BFBP[3]	10k	46.6	48.0
STC [4]	50k	49.8	51.2
SEC [5]	10k	50.7	51.7
AF-SS [6]	10k	52.6	52.7
Combining Cues [7]	10k	52.8	53.7
AE-PSL [8]	10k	55.0	55.7
DCSP [9]	10k	58.6	59.2
DSRG (VGG16)	10k	**59.0**	**60.4**
DSRG (Resnet101)	10k	**61.4**	**63.2**

COCO

| Method	Val
BFBP[3]	20.4
SEC [5]	22.4
DSRG (Ours)	**26.0**

Ablation studies

The contributions of **Balanced seeding loss, DSRG & Retrain**

Table 2. Comparison of mIoU using different settings of our approach on VOC 2012 val set

| Method	bkg	plane	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	motor	person	plant	sheep	sofa	train	tv	mIoU
baseline	82.5	67.5	23.2	65.7	29.7	47.5	71.8	66.8	76.7	23.3	51.7	26.2	69.7	54.2	63.2	57.2	33.7	64.5	33.5	48.7	46.1	52.5
+BSL	82.4	71.9	29.1	67.7	32.4	49.8	75.5	67.9	74.7	22.8	54.9	26.6	64.3	55.7	64.7	56.0	35.0	67.7	32.7	50.2	45.8	53.6
+DSRG	86.6	70.5	28.8	70.6	34.7	55.7	74.9	70.1	80.2	24.1	63.6	24.8	76.6	64.1	64.9	72.3	38.5	68.7	35.8	51.8	51.9	57.6
+Retrain	87.5	73.1	28.4	75.4	39.5	54.5	78.2	71.3	80.6	25.0	63.3	25.4	77.8	65.4	65.2	72.8	41.2	74.3	34.1	52.1	53.0	59.0

Ablation studies
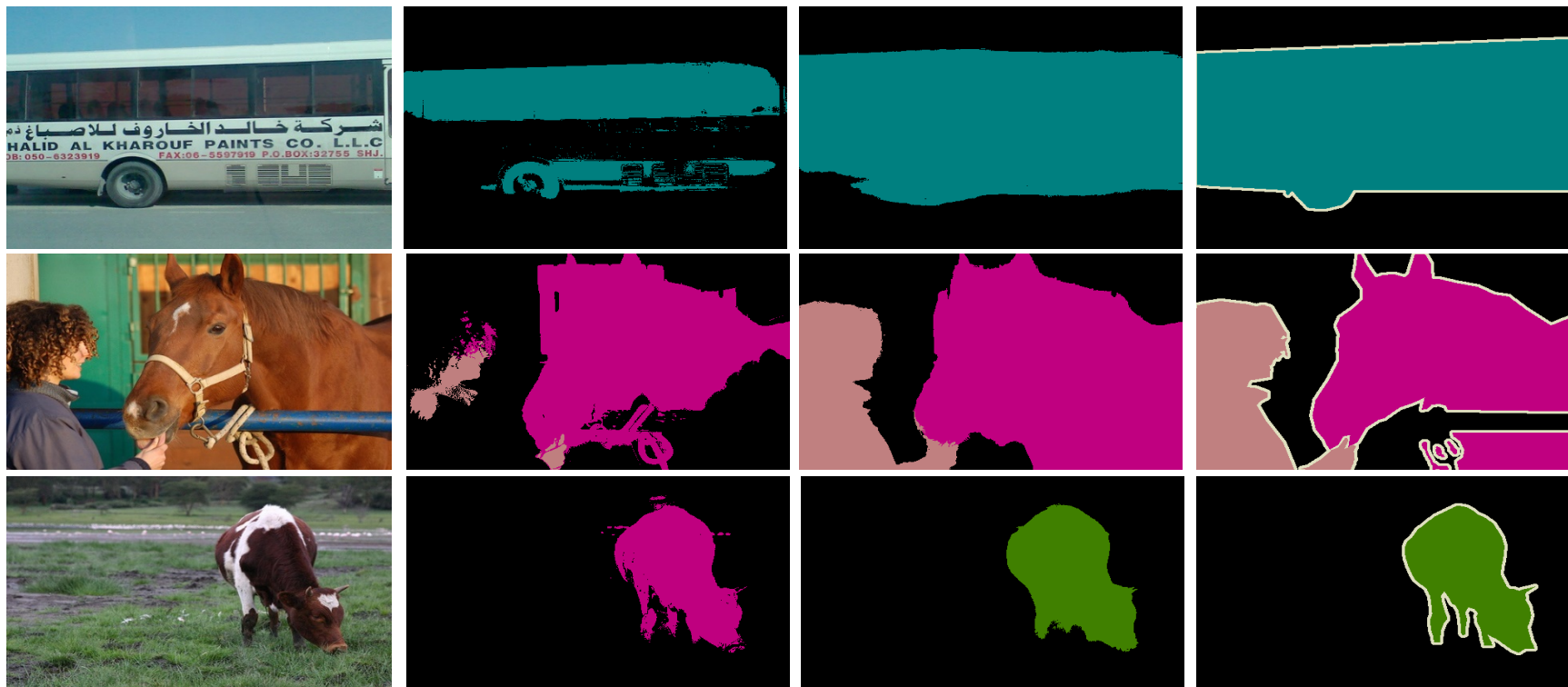
| Image | w/o DSRG | +DSRG | Ground Truth |

Ablation studies
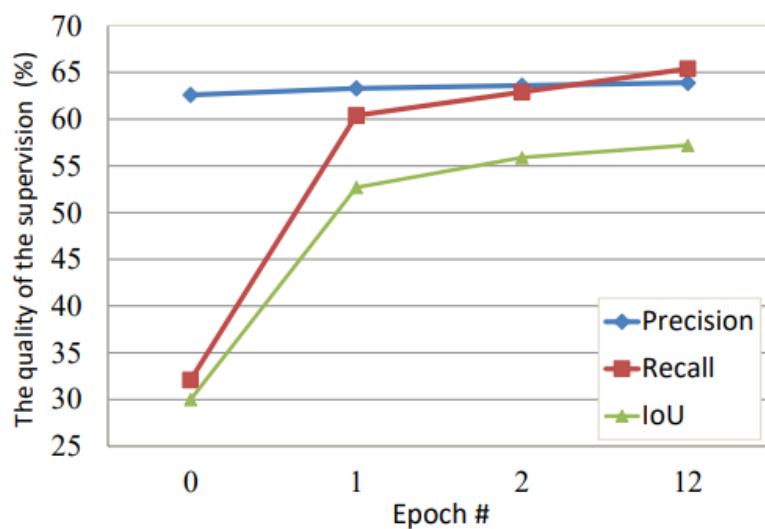
The quality of the dynamic supervision (%)
with respect to the epochs.

θ_b \ θ_f	0.99	0.95	0.90	0.85	0.80
0.99	57.45	57.59	57.63	**57.69**	57.66
0.95	57.43	57.56	57.64	57.67	57.63
0.90	57.23	57.35	57.40	57.44	57.45

Performance on PASCAL val dataset
for different θ

Video demo

Huazhong University of Science and Technology

Discussion

☐ How to interpret DSRG

  ◻ A Neural network generates new label by itself.

  ◻ The inner structure of image/video helps, e.g., [Ahn & Kwak, CVPR 18].

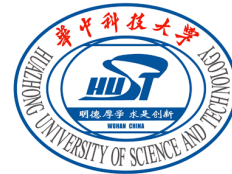 ◻ From the perspective of SSL, pseudo label/supervision [Lee, ICMLw 13, Wang et al, MM 16] works.

Discussion

- Current limitations of WSSS
  - Hard to obtain precise boundaries
  - Does not work well in complex dataset, e.g., COCO & Kitti

- Let deep networks know what is an object, e.g., unsupervised learning from video.
- Weakly and semi-supervised (WASS) visual learning.

Huazhong University of Science and Technology

- The paper is available at
  http://www.xinggangw.info/pubs/cvpr18-dsrg.pdf

- Codes will be available at
  https://github.com/speedinghzl/DSRG

Thanks for your attention!