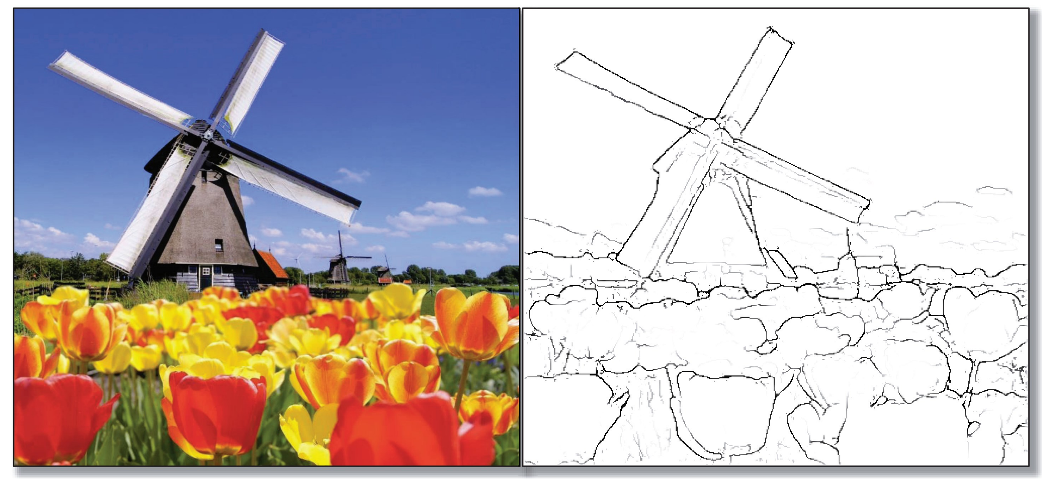
Learning Relaxed Deep Supervision for Better Edge Detection

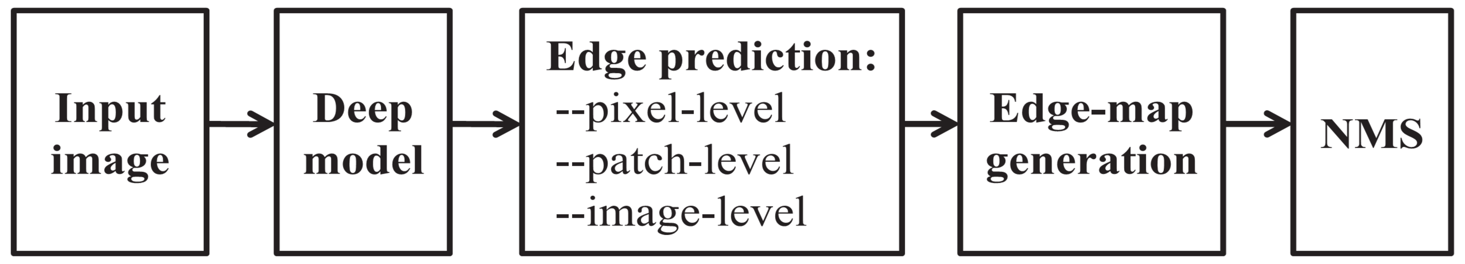
Yu Liu and Michael S. Lew LIACS Media Lab, Leiden University, Netherlands

Universiteit Leiden CVPR 2016



Introduction

Background: recent advances in the design of edge features are moving from carefully-engineered descriptors to hierarchical deep features.



** NMS: non-maximal suppression

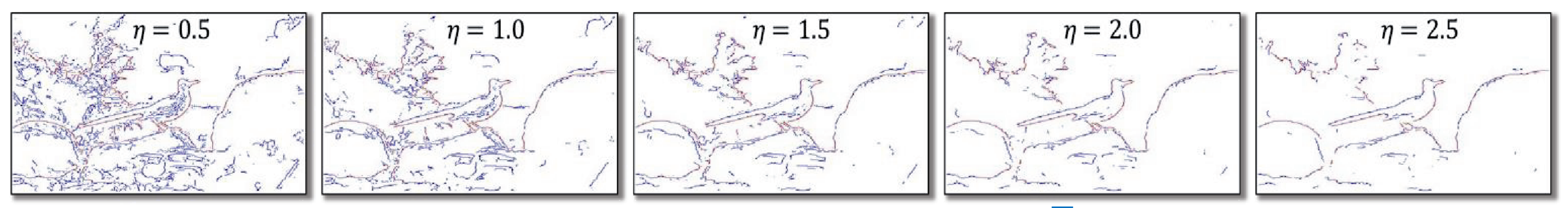
Key motivation:

1. In contrast to using a general ground-truth to guide intermediate predictions, we propose a relaxed deep supervision (RDS) method to consider the diversities in deep neural networks.
2. The relaxed labels in RDS can be seen as some false positives that are difficult to be classified in edge detection. We intend to consider these false positives in the supervision.
3. To alleviate the deficiency of the training images, we capture coarse edge annotations from segmentations.

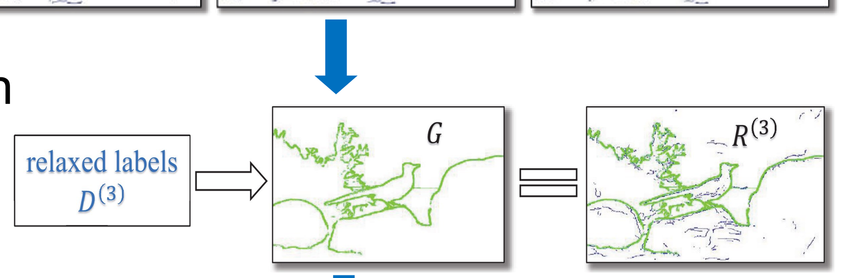
Method

RDS generation: we present two efficient ways to generate relaxed deep supervision based on Canny [1] or SE [2].

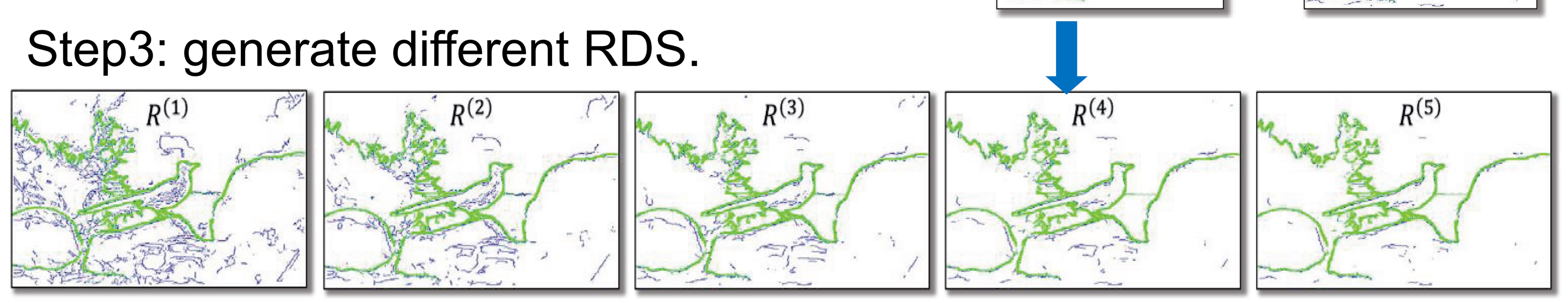
Step1: use SE edge maps and capture relaxed labels.



Step2: merge relaxed labels with the original ground-truth.



Step3: generate different RDS.

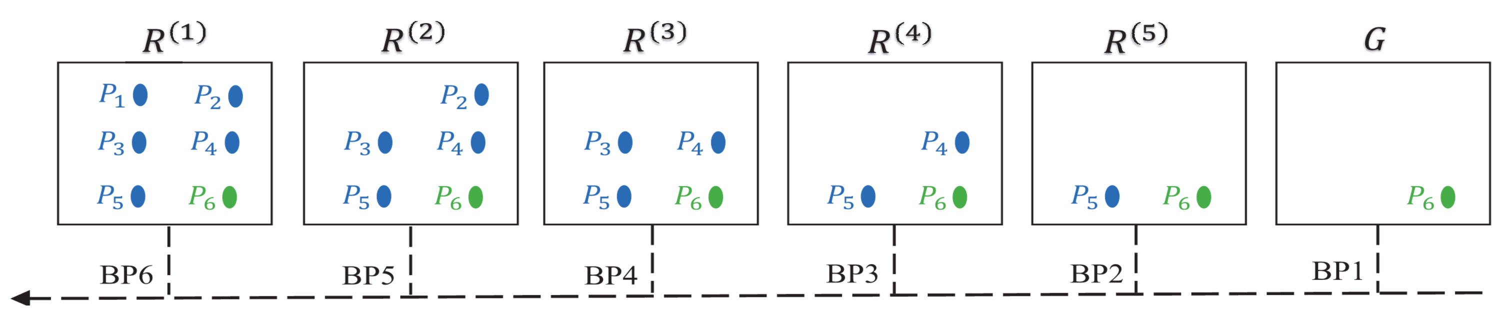


the overlap edges between the SE response and GT.

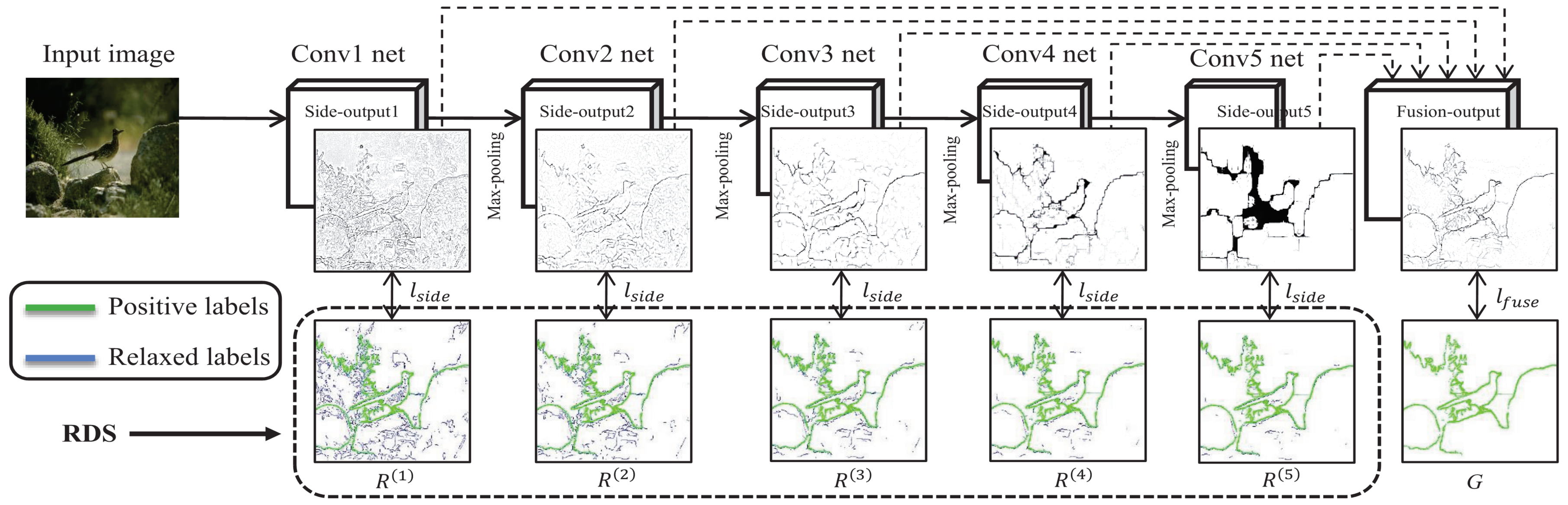
the positives labels in GT

the relaxed labels

RDS employs a coarse-to-fine supervision strategy.



Training with RDS: Our RDS method is designed to improve many vision algorithms. In these experiments, we chose the HED network [3] as a representative edge algorithm to improve.



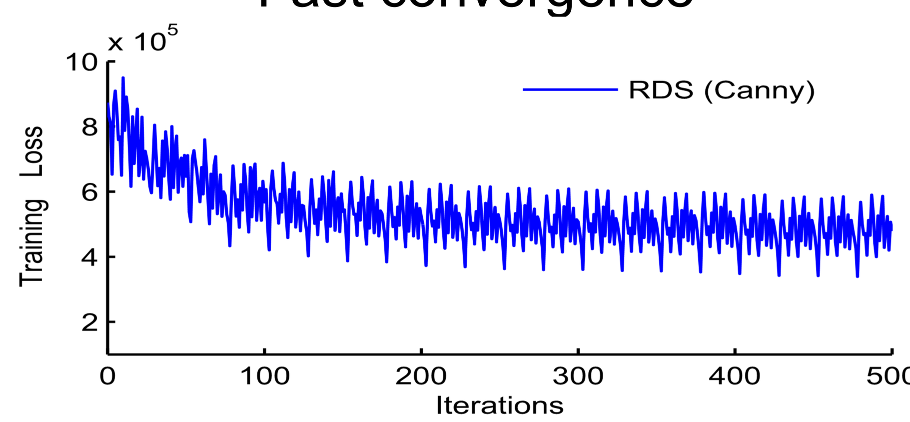
$$l_{side}(\widehat{G}^{(k)}_{i,j}, R^{(k)}_{i,j}) = \begin{cases} \alpha \cdot \log \mathrm{P}(\widehat{G}^{(k)}_{i,j}), & R^{(k)}_{i,j} = 1 \\ \beta \cdot \log(1 - \mathrm{P}(\widehat{G}^{(k)}_{i,j})), & R^{(k)}_{i,j} = 0 \\ 0, & R^{(k)}_{i,j} = 2 \end{cases}$$

1. positive labels (edges)

0: negative labels (non-edges)

2: relaxed labels

Fast convergence



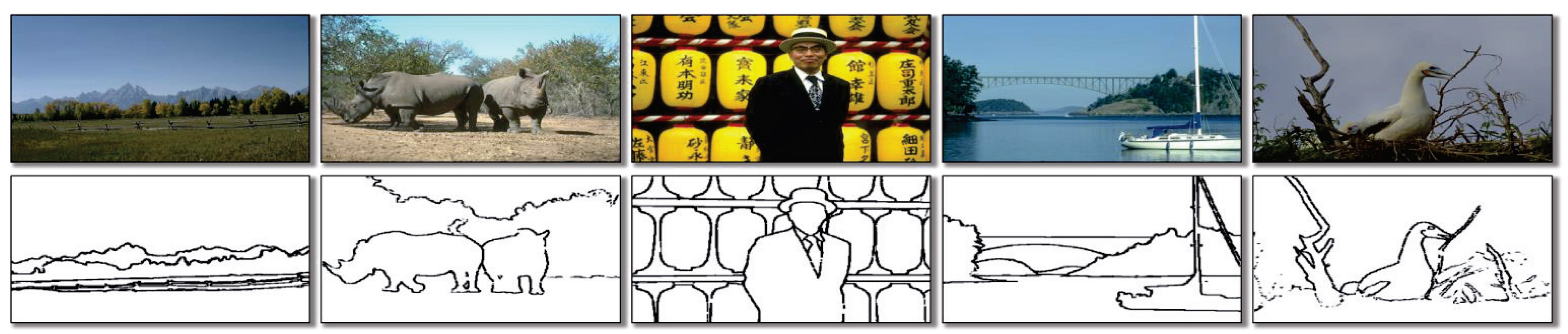
Pre-training with coarse edge annotations

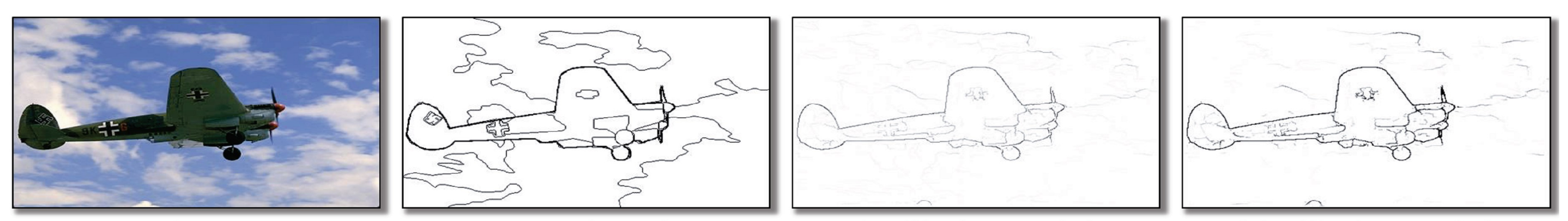
BSDS 500 PASCAL Contexts: 10103 images

(a) Fine edge annotations

(b) Coarse edge annotations

Ground-truth dilation: we dilate the positive labels in the ground-truth of train set using traditional morphologic dilation operator.





test image ground truth w/o dilation training with dilation training

	RDS(Canny) w/o dilation training	RDS(Canny) with dilation training
ODS, OIS, AP	.768, .787, .802	.785, .803, .813

Results

Table 1. Results on BSDS500 testing set

	ODS	OIS	AP
Baseline 1	.762	.782	.766
Baseline 2	.780	.802	.786
RDS(Canny)	.785	.803	.813
RDS(SE)	.787	.804	.817
RDS(gPb)	.786	.803	.814
CEA	.765	.785	.724
RDS(Canny) + CEA	.790	.809	**.819**
RDS(SE) + CEA	**.792**	**.810**	.818

Three measures: fixed contour threshold (ODS), per-image best threshold (OIS) and average precision (AP)

Baseline 1: only supervises the fusion-output prediction with the general supervision (i.e. GT).

Baseline 2: imposes the general supervision to not only the fusion-output prediction, but also five side-output predictions.

Table 2. Comparison on BSDS500.

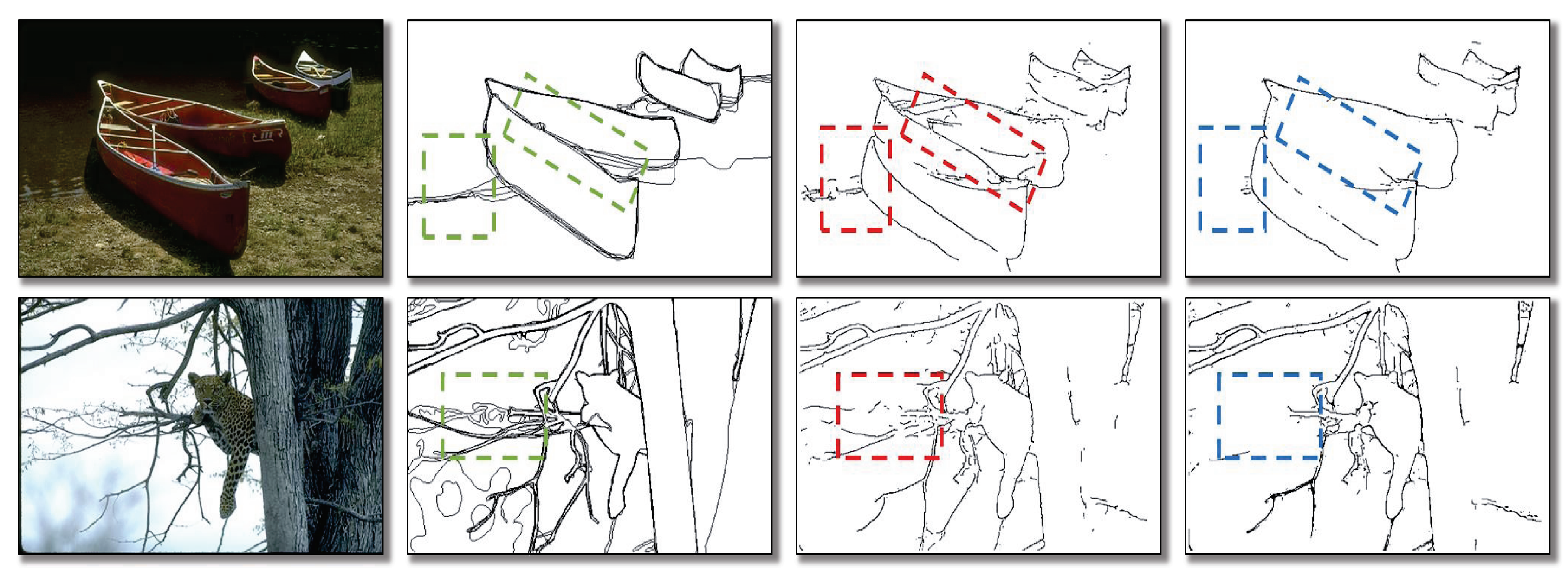
	ODS	OIS	AP
Human	.80	.80	-
DeepNet	.738	.759	.758
N4-Fields	.753	.769	.784
DeepEdge	.753	.772	.807
MSC	.756	.776	.787
CSCNN	.756	775	.798
DeepContour	.757	.776	.790
HFL	.767	.788	.795
HED	.782	.802	.787
HED-merge	.782	.804	**.833**
RDS (ours)	**.792**	**.810**	.818

Table 3. Cross-dataset generalization results. The model is trained on BSDS500 and is used to evaluate NYU test set.

	maxDist=.0075	maxDist=.011
DeepContour	.55	-
SE	.55	.64
RDS(SE)	.611	.627
RDS(SE) + CEA	**.655**	**.674**

** We do not perform late-merging and data augmentation on BSDS500.

Image Ground-truth RDS HED



Conclusion

(1) This work provides promising insights into efficiently exploiting diverse deep supervision. (2) RDS has achieved top performance on the BSDS500 dataset and may even exceed human levels. (3) It is feasible to apply this relaxation strategy to other visual recognition tasks such as image classification and segmentation.

References

1. J. Canny. A computational approach to edge detection. TPAMI, 1986.
2. P. Dollar and C. L. Zitnick. Structured forests for fast edge detection. ICCV, 2013.
3. S. Xie and Z. Tu. Holistically-nested edge detection. ICCV, 2015.