

研究の目的

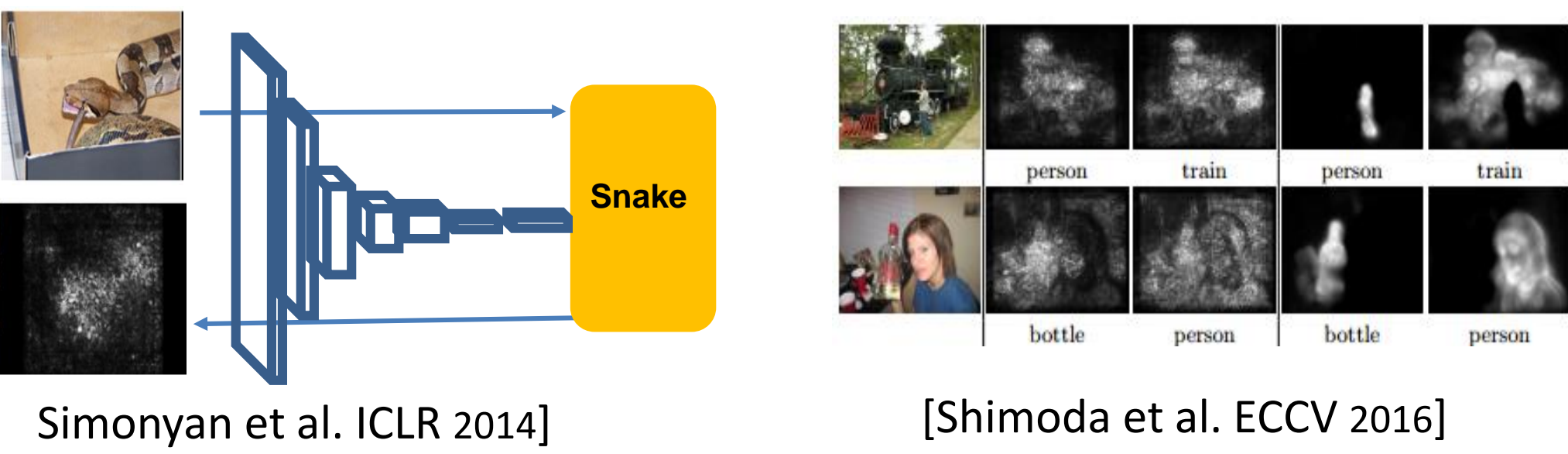
弱教師あり領域分割
画像ラベルのみから学習し物体の領域を推定



研究の背景

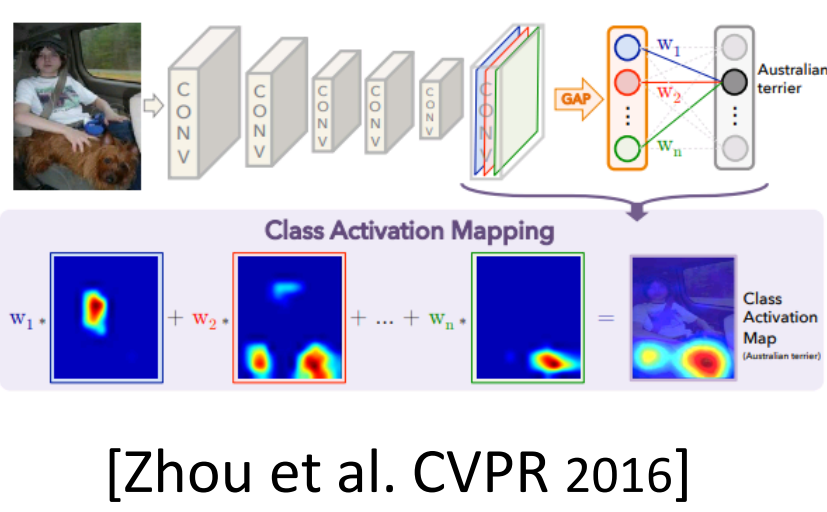
逆伝搬による認識結果の可視化

- 認識結果を可視化し、領域分割
- 学習時のプロセスを活用
- 誤差を逆伝搬すると学習に重要な領域が反応



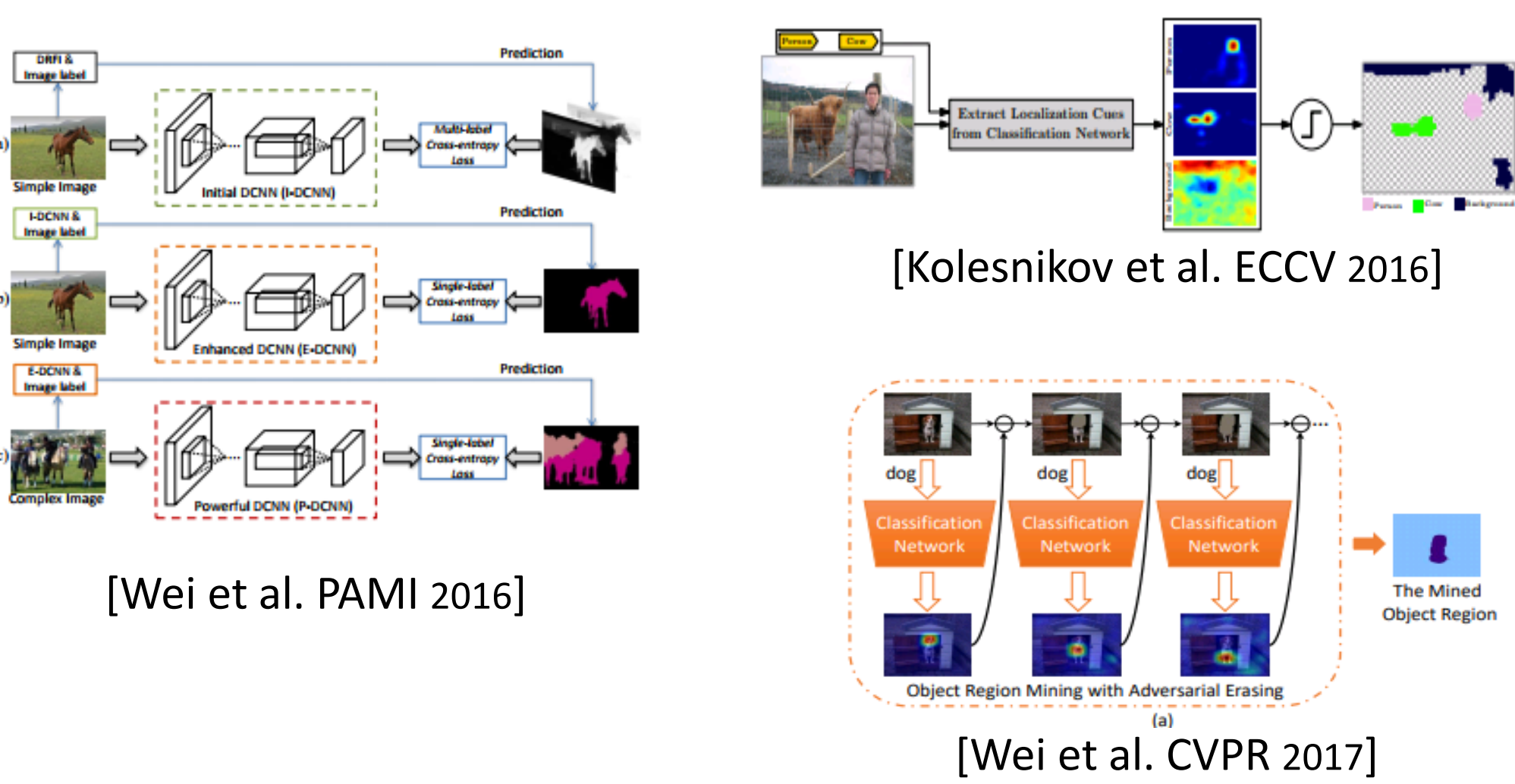
CAMIによる認識結果の可視化

- GAPでクラス分類を学習
- テスト時にGAPを使わずにWeightをかける



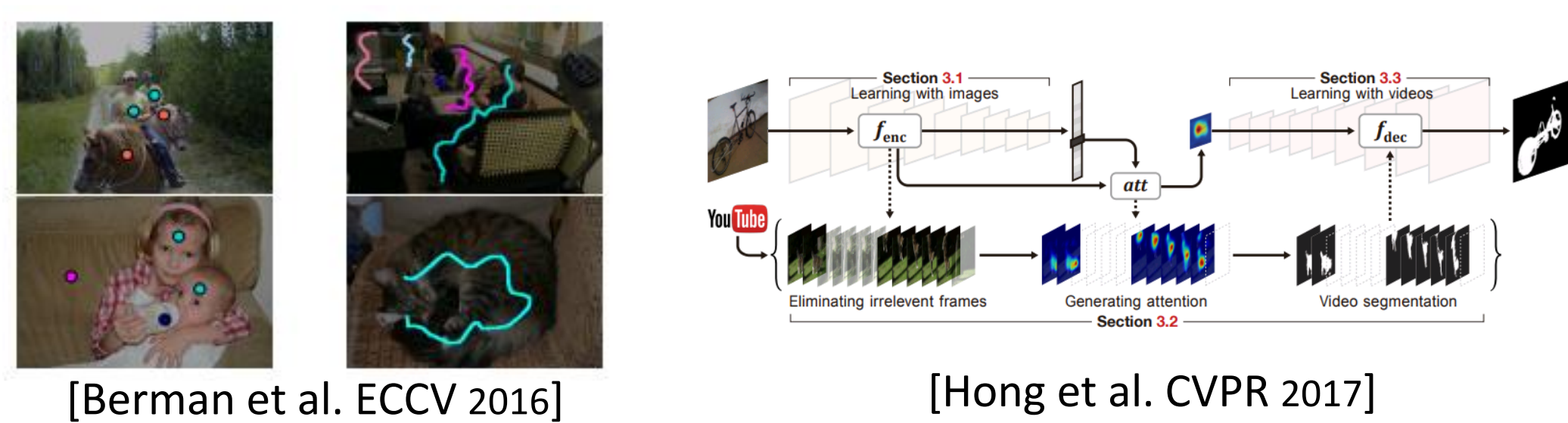
Full supervised modelの学習

- 領域シードを事前に算出
- 完全教師ありのフレームワークで学習
- 弱教師あり領域分割で高精度を達成



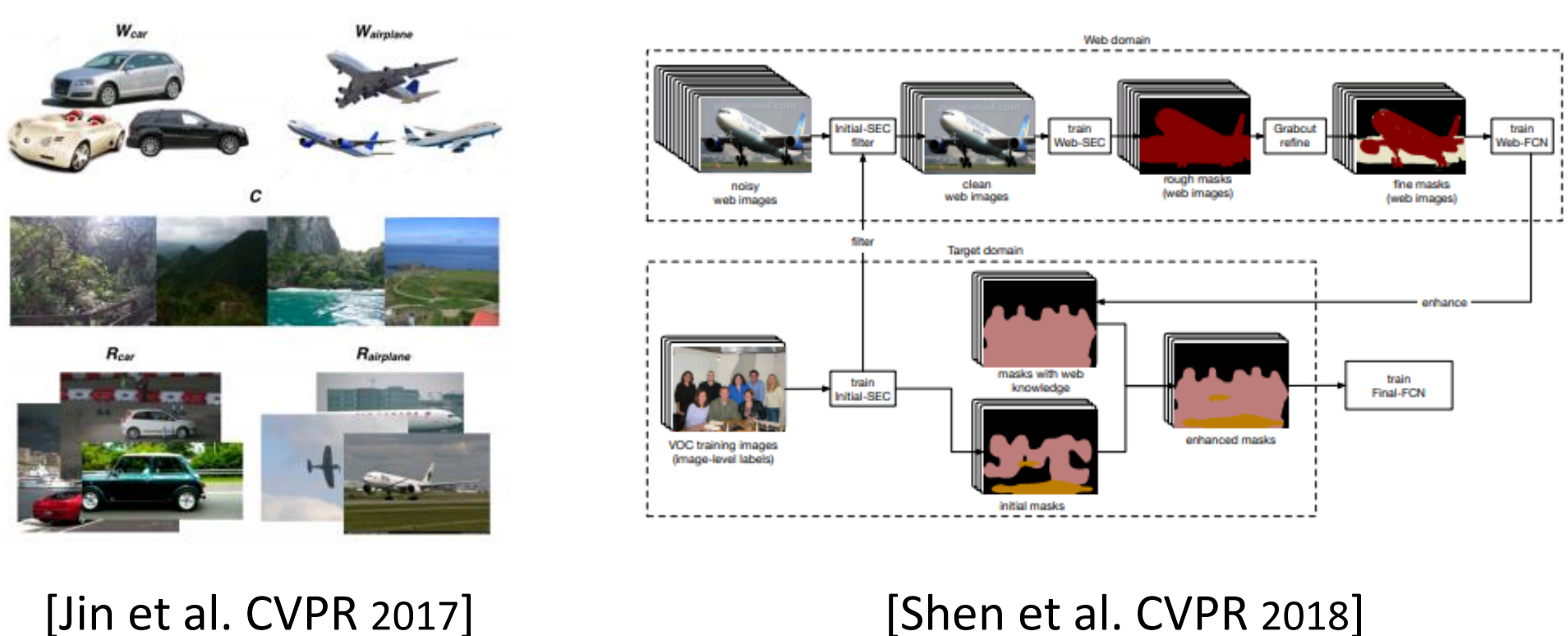
教師情報の制限の緩和

- ポイントアノテーション
- Supervised saliency (物体のマスク、カテゴリ情報なし)
- Webビデオの活用



Webly supervised segmentation

- Web画像を用いて弱教師あり領域分割の精度向上
- Web画像のDomainとScene画像のDomainの違いを利用



提案手法のモチベーション

現在の弱教師あり領域分割の課題

- 可視化結果と実際の物体の領域のギャップを埋める
- 共起の強い物体の領域の認識
- 物体領域の境界の探索
- ノイズの多い教師情報から学習
- 少ないサンプルから学習
- 物体の大きさ、角度、遮蔽など見え方の変化に弱くなる

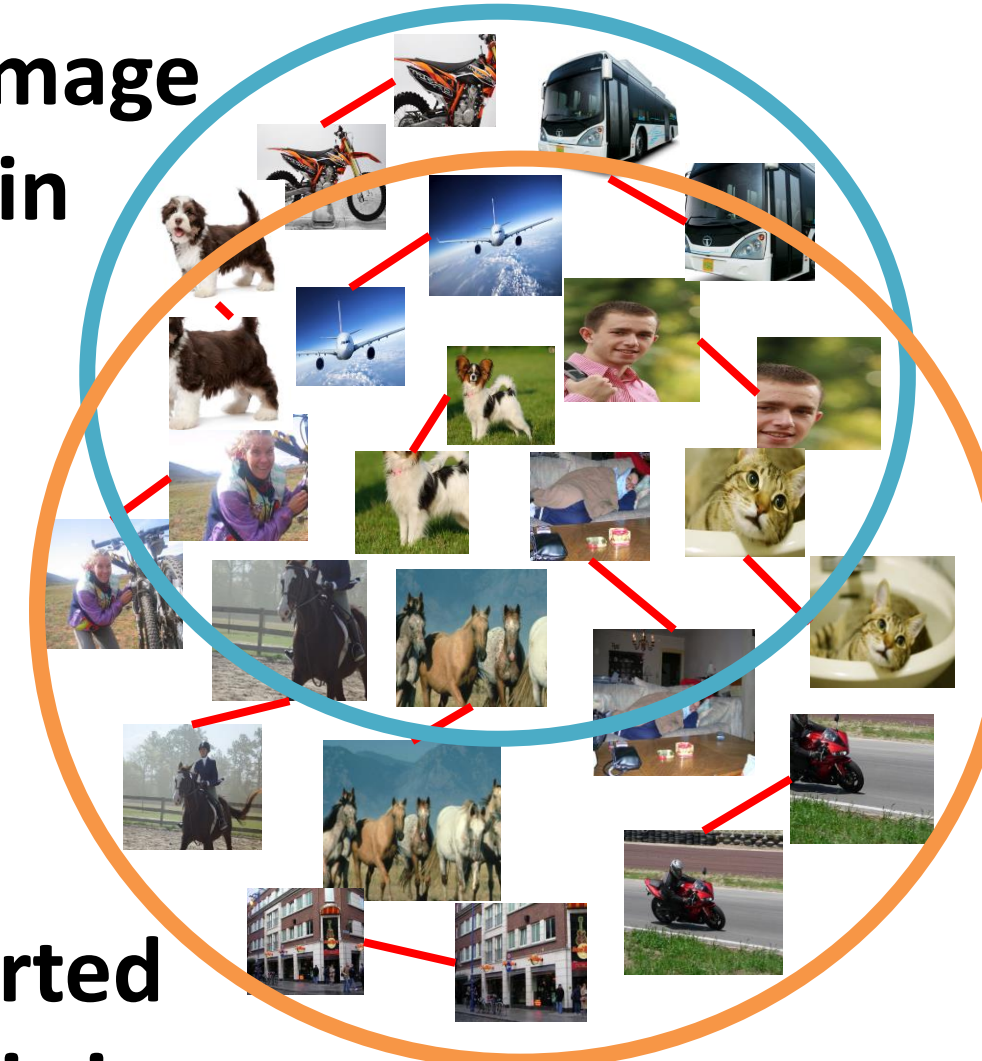
Web画像を用いた教師データの水増し

- Web画像
- 検索クエリ->ラベル
- シングルラベル (Presence)
- Scene画像 (Pascal VOC)
- マルチラベル (Presence + absence)

Web image domain



Converted web image domain



Realistic image domain



Converted realistic image domain



異なるドメインの分布を近づけて学習に活用したい

- 適切な領域を切り取れば近い分布に
- 難しいのでランダムに切り取り
- 学習が不安定に->変換前の画像と紐づける

提案手法

Seedの生成

- Adversarial erasing
- CAM+オクルージョン+学習+繰り返し
- CAM+オクルージョン+繰り返しに変更
- Supervised Saliency (DRFI [Jiang et al. CVPR 2013])
- Unsupervised saliencyに変更 [Zhu et al. CVPR 2014]

大域領域と局所領域の一貫性の学習

- 大域領域 x^g 、局所領域 x^l とする
- それぞれの推定結果の共通領域の一貫性を学習
- Consistency loss

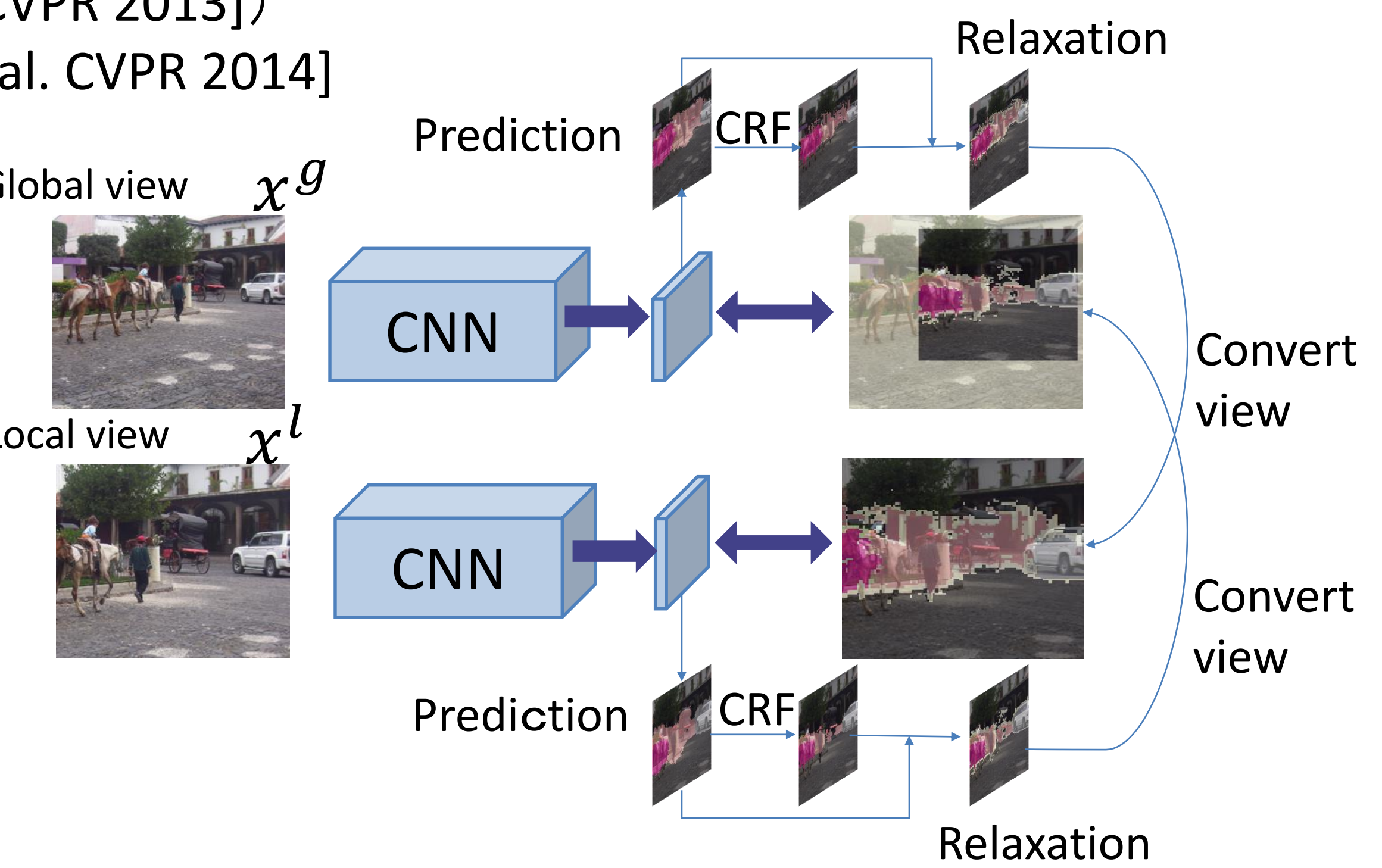


最適化

- Loss = Seed loss + GAP loss + CRF loss + Consistency loss

動的なSeedの評価領域の制限

- Seedはノイズを多く含んでいる
- 認識結果の確信度を用いてSeedの評価領域を減らす
- エポックごとに減らす領域を増加



実験結果

Pascal VOC 2012 dataset
Training image 10582枚
Validation 1449枚 Testset 1456枚

Ablation study (val set)

Seed loss + gap loss	CRF loss	Consistency loss	Seedの修正	MIoU
✓	—	—	—	45.4
✓	✓	—	—	46.0
✓	✓	✓	—	54.4
✓	—	—	✓	38.2
✓	✓	✓	✓	44.3
✓	✓	✓	✓	59.2

Seedの評価領域の削減と精度の変化 (val set)

Rate	0.0	0.1	0.2	0.3	0.4	0.5
mIoU	54.4	3.7	25.9	54.7	51.2	60.3
Rate	0.6	0.7	0.8	0.9	1.0	
mIoU	56.2	59.2	58.1	52.7	3.6	

Pascal VOC 2012 Mean IoU

弱教師あり領域分割	publish	Val	Test
追加情報を活用			
Point annotation	ECCV 2016	46.1	-
Video	CVPR 2017	58.1	58.7
画像ラベル情報のみ			
Global pooling	ICLR 2015	25.7	24.9
Ours	ECCV 2016	44.1	45.1
SEC	ECCV 2016	50.7	51.7
Adversarial Erasing	CVPR 2017	55.0	55.7
Iterative mining	CVPR 2018	60.3	61.2
Affinity Net	CVPR 2018	61.7	63.7
Web画像を使用			
STC	PAMI 2017	49.8	51.2
WebS	CVPR2017	52.6	55.3
Web SEC	CVPR2018	63.0	63.9
Ours	-	60.3	61.3

Web SECは7万6000枚のWeb画像を使用
本手法は6000枚のWeb画像を使用

Input, Seedの修正なし, Seed loss + GAP loss (mIoU 45.4%), Seedの修正あり, Seed loss + GAP loss (mIoU 38.2%), Seedの修正あり, Seed loss + GAP loss + CRF loss (mIoU 44.3%), Seedの修正あり, Seed loss + GAP loss + CRF loss + Consistency loss (mIoU 59.2%), Ground truth

