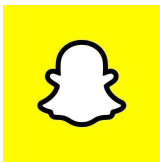




UNIVERSITY of
ROCHESTER



Weakly Supervised Body Part Segmentation with Pose based Part Priors

Zhengyuan Yang, Yuncheng Li, Linjie Yang,
Ning Zhang, Jiebo Luo

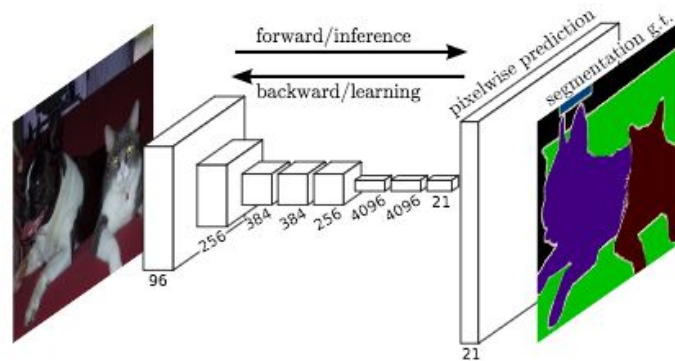
Body Part Segmentation

- Given an image, predict the semantic segmentation mask for each body part



Fully-supervised Method

- Limitations
- Per-pixel classification
- Limitations:
 - High annotation cost
 - Different part definitions



Weakly-supervised Method

- **Motivations**
- Abundant existing weak supervision (e.g., human pose)
- Challenges:
Dense prediction with sparse labels



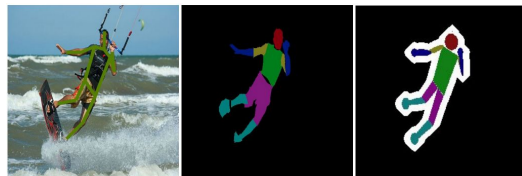
Fully Supervised Training:
Mask supervision



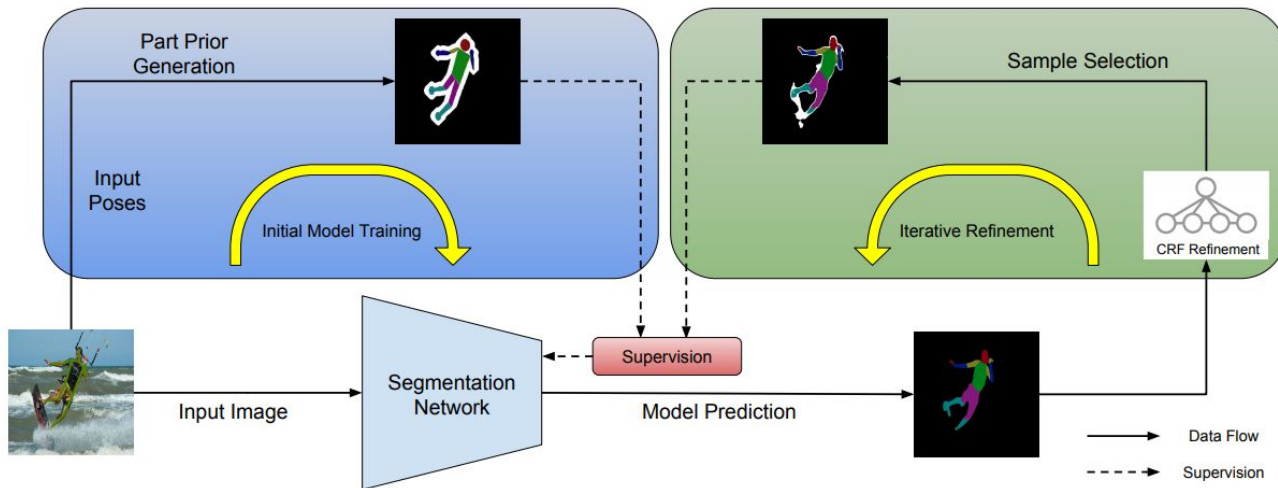
Weakly Supervised Training:
Pose supervision

Method

- Framework overview



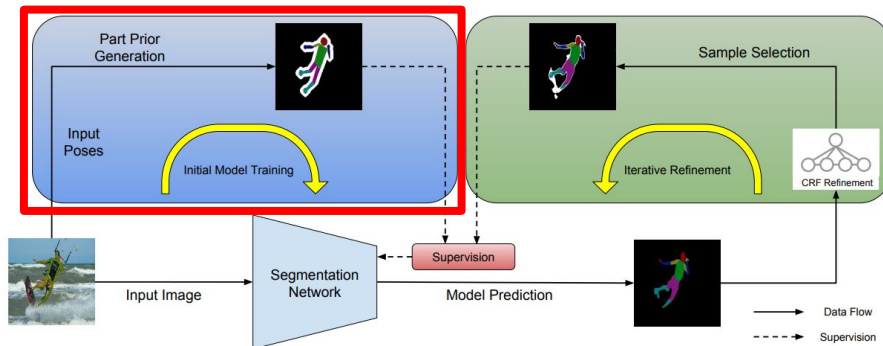
Part prior



- Part prior initialization
- Iterative refinement

Method

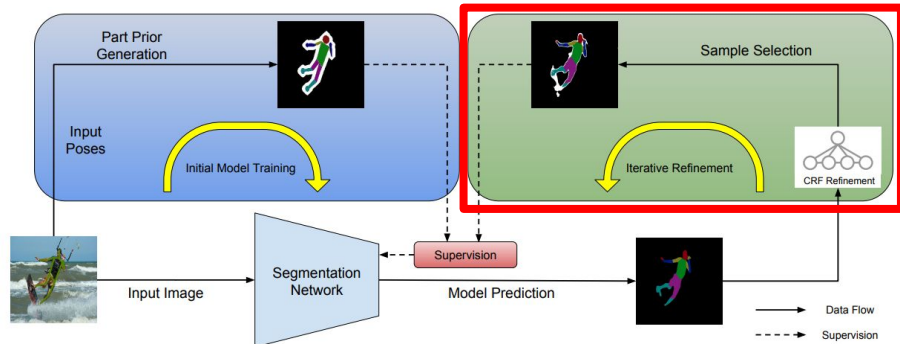
- Part prior initialization



- Projected from sparse poses based on predefined geometric rules
- Uncertain regions in white

Method

- Iterative refinement





- Per-step CRF refinement for the boundaries
- Partial CE loss on both prior and refined masks to get both structure and boundary details

$$L_s = \sum_{i \in F \cup B} \sum_{c=0}^C \ell(f_i(c), y_i(c))$$

Experiment Results

- **Pascal-Person-Part**

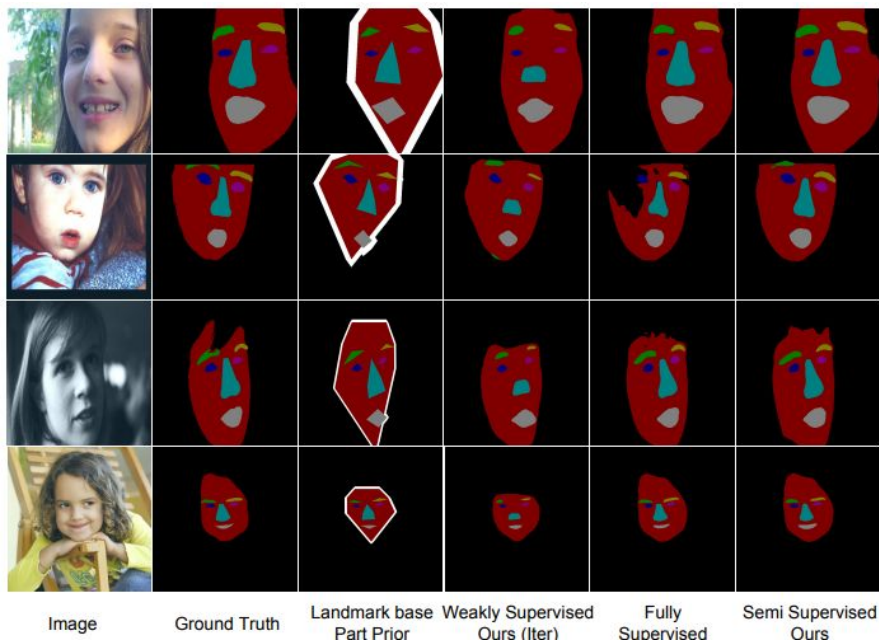
Methods	Supervision	mIoU	Head	Torso	U-Arm	L-Arm	U-Leg	L-Leg	Bkg
LIP [22]	Fully	59.36	83.3	62.4	47.8	45.6	42.3	39.5	94.7
LG-LSTM [40]	Fully	57.97	82.7	61.0	45.4	47.8	42.3	38.0	88.6
Graph LSTM [41]	Fully	60.16	82.7	62.7	46.9	47.7	<u>45.7</u>	40.9	94.6
DeepLab [18]	Fully	 <u>63.64</u>	<u>84.6</u>	<u>66.9</u>	<u>56.0</u>	<u>54.2</u>	45.5	<u>43.4</u>	<u>94.9</u>
Ours (Part prior)	Weakly	40.91	54.9	35.1	34.2	32.4	19.0	24.5	86.2
FastNet [3]	Weakly	42.11	61.6	37.8	32.7	29.2	20.8	25.0	87.1
Ours (Part prior supervision)	Weakly	43.91	50.6	47.0	31.8	29.5	29.3	27.9	87.0
Ours (Base, only PASCAL data)	Weakly	53.54	76.1	54.8	39.6	36.5	37.9	36.1	93.9
Ours (Iter, only PASCAL data)	Weakly	54.72	76.7	55.6	40.2	38.4	38.6	38.9	94.6
Ours (Base, with COCO data)	Weakly	60.35	78.3	59.6	46.6	45.4	46.6	50.0	95.8
Ours (Iter, with COCO data)	Weakly	 62.05	79.6	62.0	48.1	48.5	48.7	51.8	95.8

- With extra data, obtains comparable mIoU to the fully-supervised SOTA

Qualitative Results



Extension to Face Parsing



Methods	Supervision	mIoU
Baseline (Part Prior Supv.)	Weakly	42.54
Ours (Base)	Weakly	47.21
Ours (Iter)	Weakly	54.05
Fully Supervised	Fully	68.48
Ours (Semi)	Semi	73.95

- Extended to other keypoint supervised part segmentation tasks
- Face parsing on the Helen and AFLW datasets

Weakly Supervised Body Part Segmentation with Pose based Part Priors

Poster: #78

Contact:
zyang39@cs.rochester.edu

