Reinforcement Learning for Semantic Segmentation in Indoor Scenes

Md. Alimoor Reza and Jana Kosecka
George Mason University

ROBOTICS SCIENCE AND SYSTEMS
June 18-22, 2016 University of Michigan

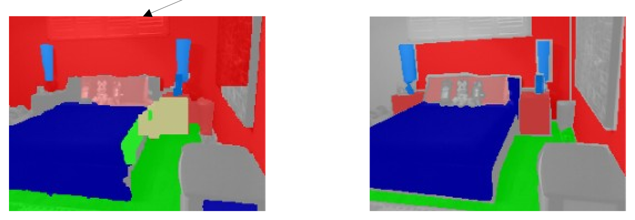
Problem

- Semantic segmentation by combining multiple binary object-background CRF segmentations
- Contrast to a single multi-class CRF
- Combination strategy is learned using Reinforcement Learning
- Modular approach, which can be easily extended to additional categories

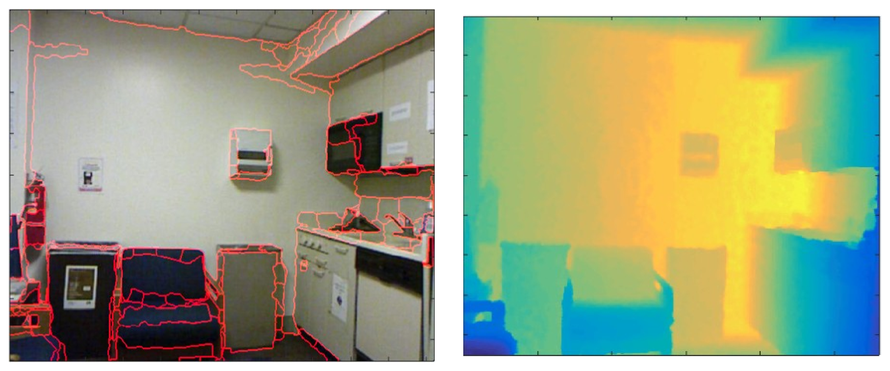

Sample bedroom image


Bed-vs-Background Pillow-vs-Background Lamp-vs-Background

Sequentially combine in a MDP formulation using Reinforcement Learning


Semantic segmentation Ground truth

Approach


RGB with superpixel boundaries and depth

Object-specific classifier (RUSBoost)

Binary CRF Inference for each object category

Combine multiple binary MAP predictions in a Reinforcement Learning framework

Semantically segmented image

Super-pixel and Features

- Superpixels are partitioned into planar-nonplanar regions; non-planar regions are represented by SLIC and planar using robust plane fitting
- Color, texture, geometric, & generic features are computed over the superpixels

Object Specific Binary CRF

- CRF conditional distribution is modeled as follows:

$$p(\mathbf{x}|\mathbf{z}) = \frac{1}{Z(\mathbf{z})}\exp\Big(w_1\sum_{i}^{V}\theta_d(x_i,\mathbf{z}) + w_2\sum_{(i,j)}^{E}\theta_{ps}(x_i,x_j,\mathbf{z})\Big)$$

- **Unary:** output probability of RUSBoost classifier [6] and **Pairwise:** constant cost for label difference
- 4-connected pixel neighborhood graph and solved using Graph Cut

MDP formulation for Sequential Combination

- A model free RL technique, Least Square Policy Iteration (LSPI) [5], is adopted to learn the optimal policy

$$\pi(s) = \arg\max_a Q^{\pi}(s,a) \;=\arg\max_a w_{\pi}^{T}\psi(s,a)$$

- State-action value function $Q^{\pi}(s,a)$ is approximated using a linearly weighted feature function $\psi(s,a)$ of the states and actions. LSPI learns this functional approximation from the samples of the MDP

$$C = C + \psi(s,a)(\psi(s,a) - \gamma\psi(s',\pi(s')))^{T} \qquad b = b + \psi(s,a)r \qquad w = C^{-1}b$$

- Matrix C and vector b are updated from each sample (s,a,s',r)
- **Actions** are selection of the object binary segmentation mask and **reward** is computed based on a *pixel-freq-weighted Jaccard Index (JI)* metric of the current sequentially segmented image

$$R(s,a_k,s') = w_1 * JI_{a_1} + w_2 * JI_{a_2} + \ldots + w_k * JI_{a_k} + r_{bg}$$

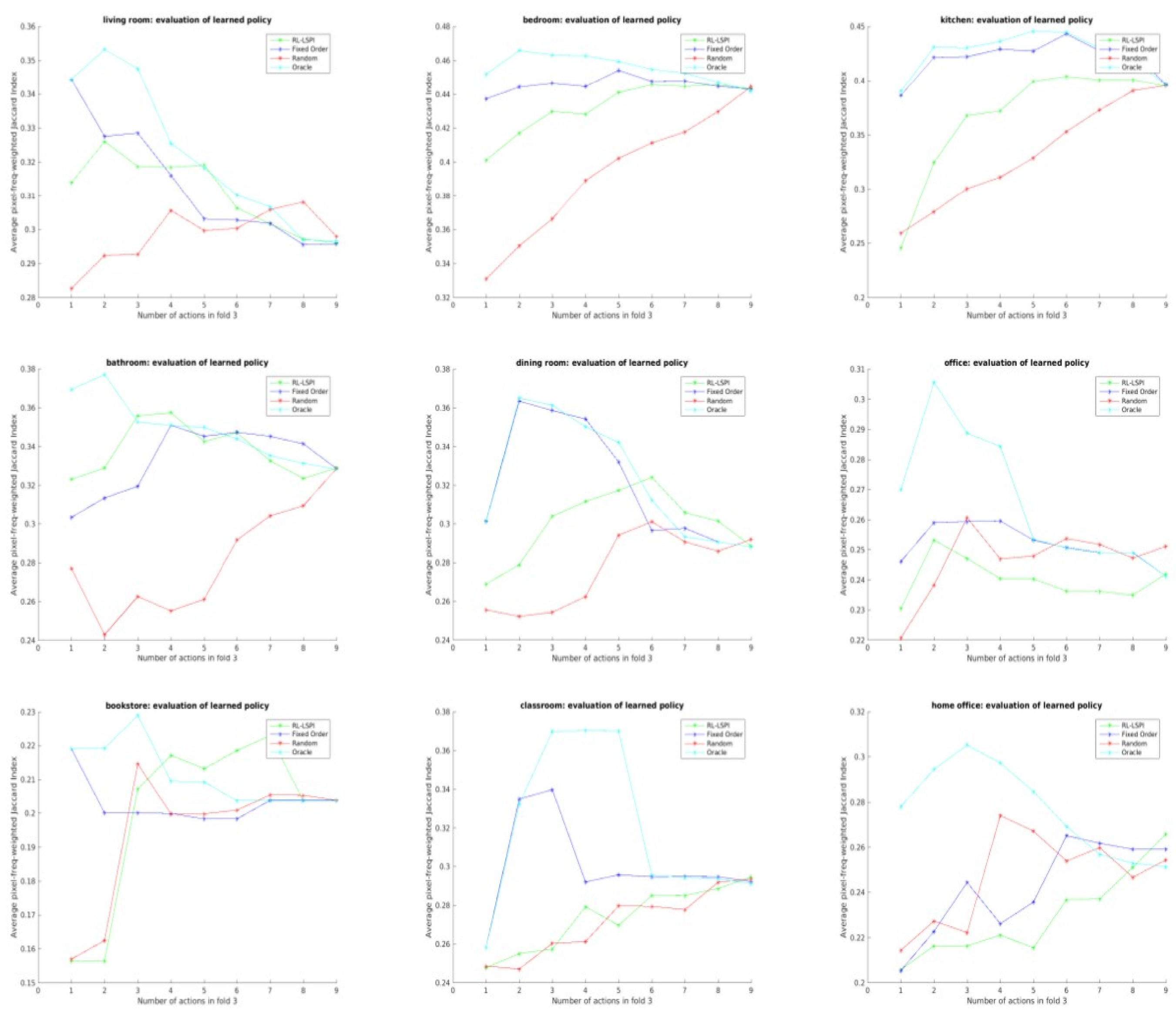


Figure: Performance comparison our LSPI policy against three baselines in NYUD-V2 test set. Each plot shows comparison in a separate scene, where a subset of object are considered.

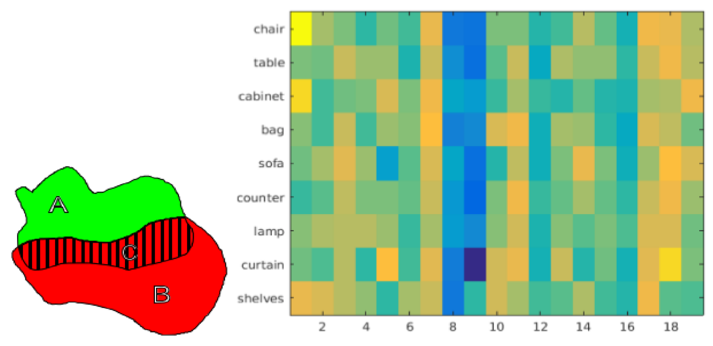


Figure: toy-example for conflict resolution (left) and our leaned weight vector using LSPI in *dining room* scene (right).

Conflict resolution: Considers segment-size ratio heuristic, which prefers small-sized segments that can take place on-top-of-larger segments

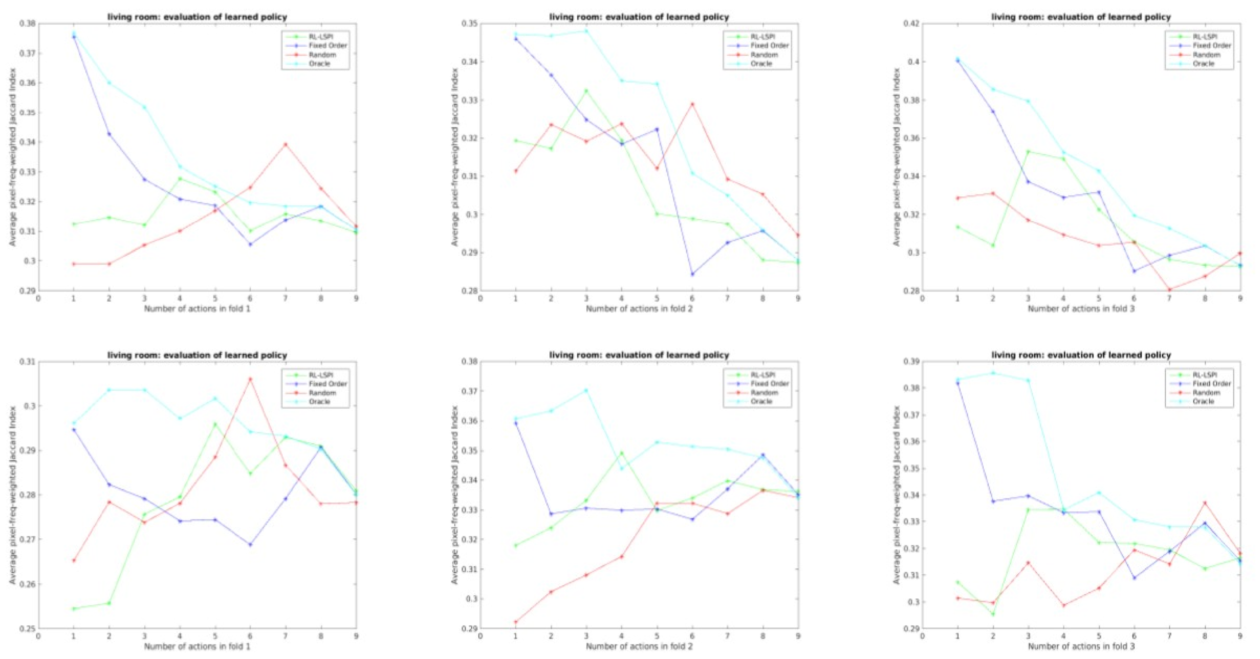


Figure: comparison of our LSPI policy against three baselines in a controlled experiment in *living room* scene

Qualitative Results

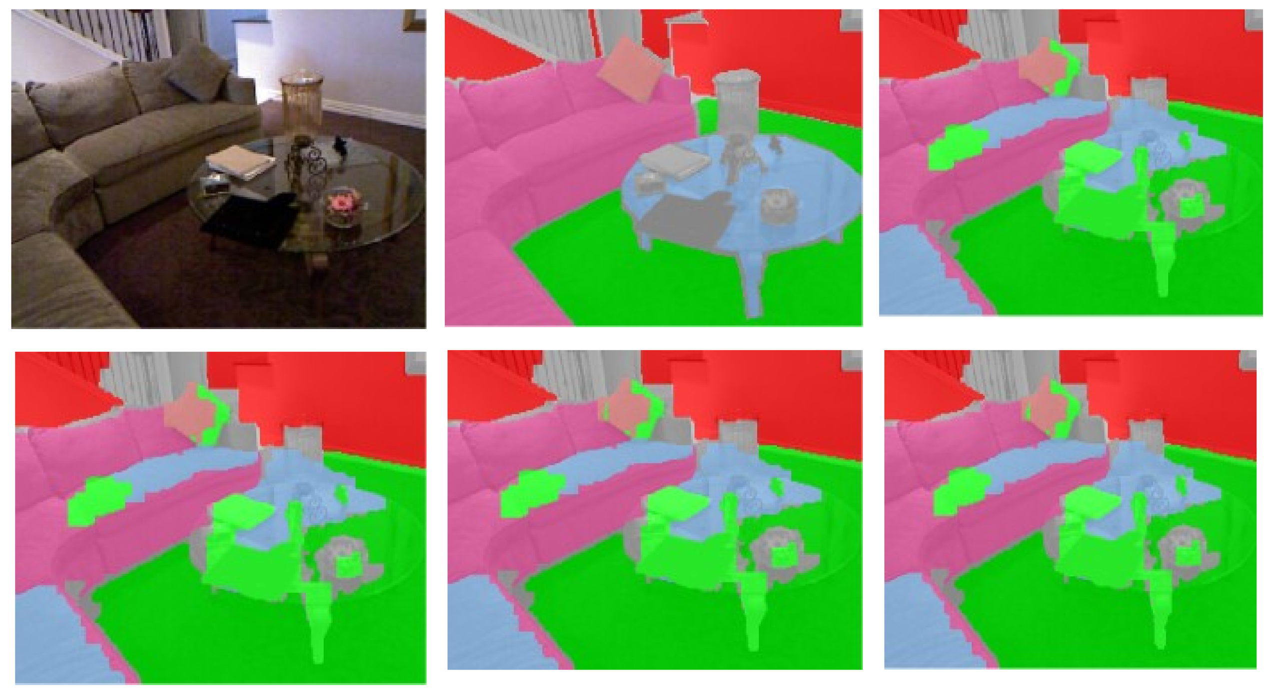


Figure: Qualitative output of our system in *living room* scene. From top-left in clockwise order the images are RGB, ground truth, our LSPI learned, Fixed Order, Random Order and Oracle Policy predictions resp.

Evaluation

- Pixelwise percentage Jaccard Index to evaluate on 9 common scenes in NYU-D V2 dataset [4]

	cabinet	toilet	sink	counter	towel	bathtub	shower c.	bag	shelves	wall	floor	ceiling
Our	25.6	33.3	26.5	**55.6**	21.0	31.8	**16.2**	**0.3**	**6.4**	41.5	61.3	0
[9]	44.8	46.5	35.7	52	25.9	31.1	9.7	0	4.5	67.6	81.2	61.1
[10]	44.9	55.1	37.5	51.3	16.3	38.2	4.2	0.2	3.5	68	81.3	60.5

	bed	pillow	lamp	night s.	dresser	box	clothes	chair	books	wall	floor	ceiling
Our	55.9	26.9	15.5	10.0	28.5	**2.1**	6.9	16.6	**7.3**	61.0	**82.3**	19.6
[9]	57.0	30.3	16.3	21.5	24.3	2.1	4.7	36.7	5.5	67.6	81.2	61.1
[10]	60.5	34.4	34.8	27.2	34.8	2.1	7.4	47.9	6.4	68	81.3	60.5

	cabinet	counter	sink	chair	bag	fridge	table	towel	box	wall	floor	ceiling
Our	**46.1**	36.7	7.9	32.1	**2.1**	**25.9**	12.7	7.9	**2.9**	46.5	**86.9**	50.2
[9]	44.8	52	35.7	36.7	0	16.2	28	25.9	2.1	67.6	81.2	61.1
[10]	44.9	51.3	37.5	47.9	0.2	14.5	29.9	16.3	2.1	68	81.3	60.5

	chair	table	cabinet	desk	shelves	box	person	books	bag	wall	floor	ceiling
Our	27.9	8.2	11.9	**26.2**	**5.4**	0.9	1.2	0	0	35.0	64.4	**63.0**
[9]	36.7	28	44.8	7.1	4.5	2.1	5	5.5	0	67.6	81.2	61.1
[10]	47.9	29.9	44.9	11.3	3.5	2.1	0.2	6.4	0.2	68	81.3	60.5

	chair	table	cabinet	bag	sofa	counter	lamp	curtain	shelves	wall	floor	ceiling
Our	43.9	26.5	23.3	**4.4**	2.4	0.2	8.5	16.1	0	45.8	72.2	**83.3**
[9]	36.7	28	44.8	0	40.8	52	16.3	28.6	4.5	67.6	81.2	61.1
[10]	47.9	29.9	44.9	0.2	47.9	51.3	34.8	29.1	3.5	68	81.3	60.5

	chair	bookshf	table	desk	books	sofa	cabinet	bag	lamp	wall	floor	ceiling
Our	28.0	15.5	5.7	**11.7**	**28.9**	19.7	6.9	**6.3**	4.9	47.9	72.6	0
[9]	36.7	19.5	28	7.1	5.5	40.8	44.8	0	16.3	67.6	81.2	61.1
[10]	47.9	18.1	29.9	11.3	6.4	47.9	44.9	0.2	34.8	68	81.3	60.5

Table: Scene specific performance comparison in NYUD-V2 dataset. From top-to-bottom the scenes are *bathroom, bedroom, kitchen, classroom, dining room, and home-office resp.*

Conclusion

- Demonstrated a task-dependent approach for semantic segmentation where a subset of objects can be sought
- Our binary CRFs have exact solutions
- Reinforcement Learning solution with LSPI affords simplicity

Acknowledgments: This work was supported by NSF IIS-1527208

References

1. S. Gupta, R. Girshick, P. Arbelaez, and J. Malik. Learning rich features from RGB-D images for object detectionand segmentation. In (ECCV). 2014.
2. S. Gupta, P. Arbelaez, and J. Malik. Perceptual organization and recognition of indoor scenes from RGB-D images. (CVPR), 2013.
3. A. Hermans, G. Floros, and B. Leibe. Dense 3d semantic mapping of indoor scenes from RGB-D images. (ICRA), 2014.
4. N. Silberman, D. Hoiem, P. Kohli, and R. Fergus. Indoor segmentation and support inference from rgbd images. (ECCV), 2012
5. Michail G. Lagoudakis and Ronald Parr. Least-squares policy iteration. Journal of Machine Learning Res., 2003.
6. C. Seiffert, T. M. Khoshgoftaar, J. V. Hulse, and A. Napolitano. RUSBoost: A hybrid approach to alleviating class imbalance. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 2010.