DA-RNN: Semantic Mapping with Data Associated Recurrent Neural Networks

Yu Xiang and Dieter Fox

Paul G. Allen School of Computer Science & Engineering

University of Washington

RSS 2017

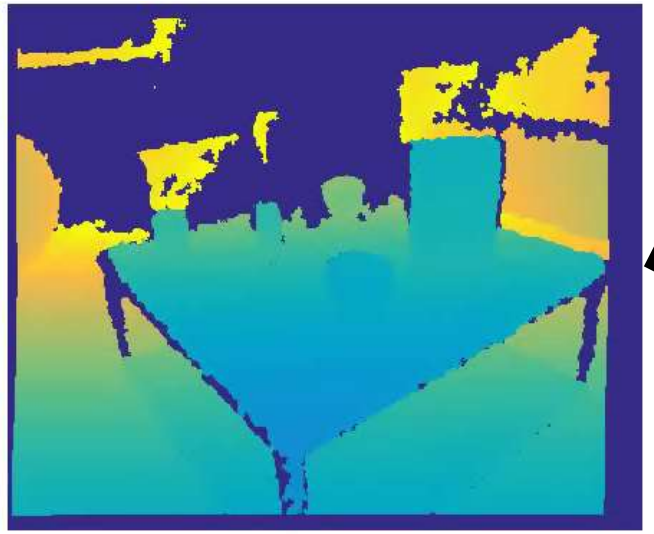
Robotics and State Estimation Lab

3D Scene Understanding

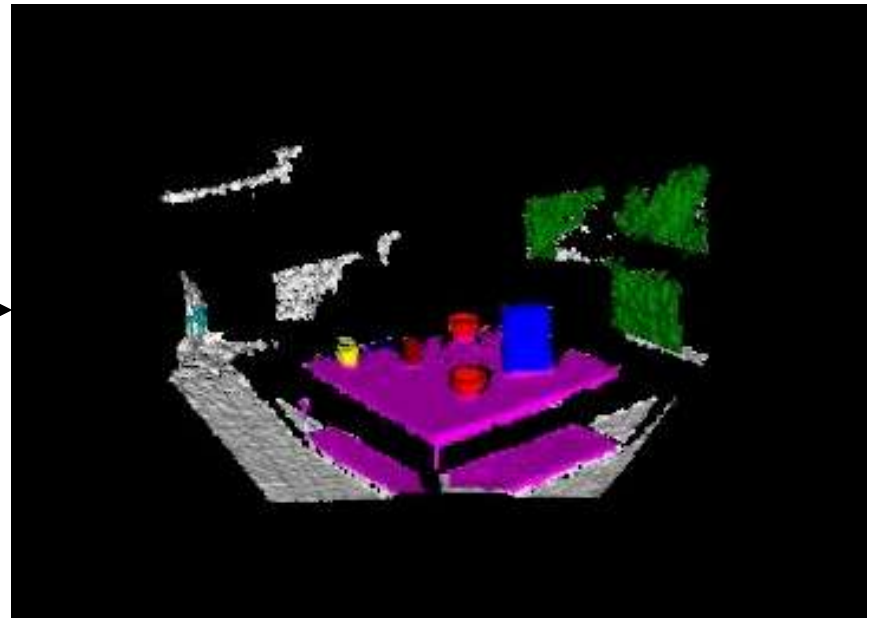
- Navigation
- Manipulation
- …

- Geometry
  - ✓ Free space
  - ✓ Surface
- Semantics
  - ✓ Objects
  - ✓ Affordances

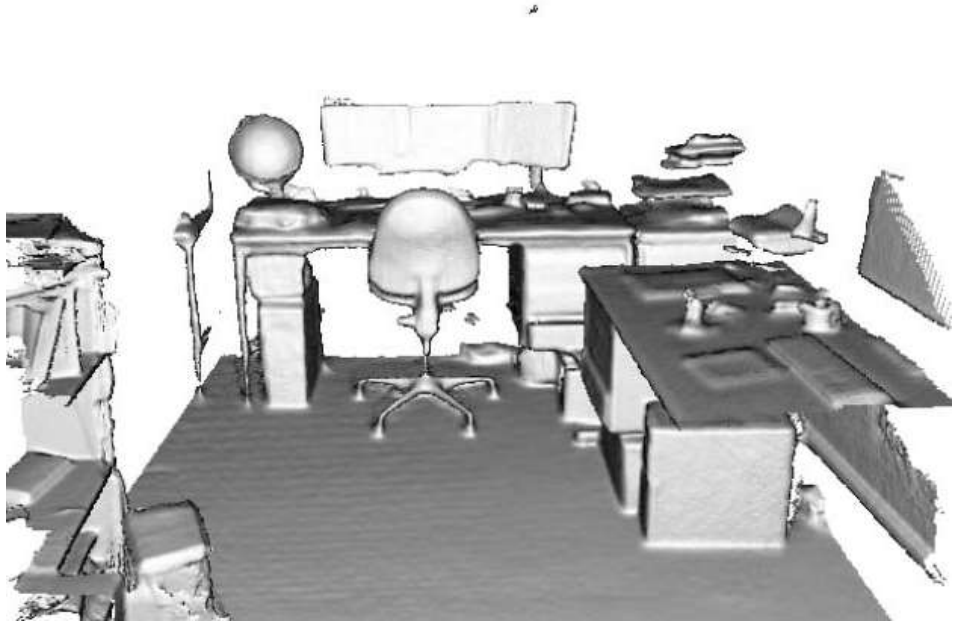
Semantic Mapping with Data Associated Recurrent Neural Networks (DA-RNNs)



DA-RNN



Related Work: 3D Scene Reconstruction



KinectFusion

- ✓ Geometry
- ✓ Data Association
- ✗ Semantics

- Newcombe et al., ISMAR'11
- Henry et al., IJRR'12, 3DV'13
- Whelan et al., RSS Workshop'12, RSS'15
- Keller et al., 3DV'13

Related Work: Semantic Labeling

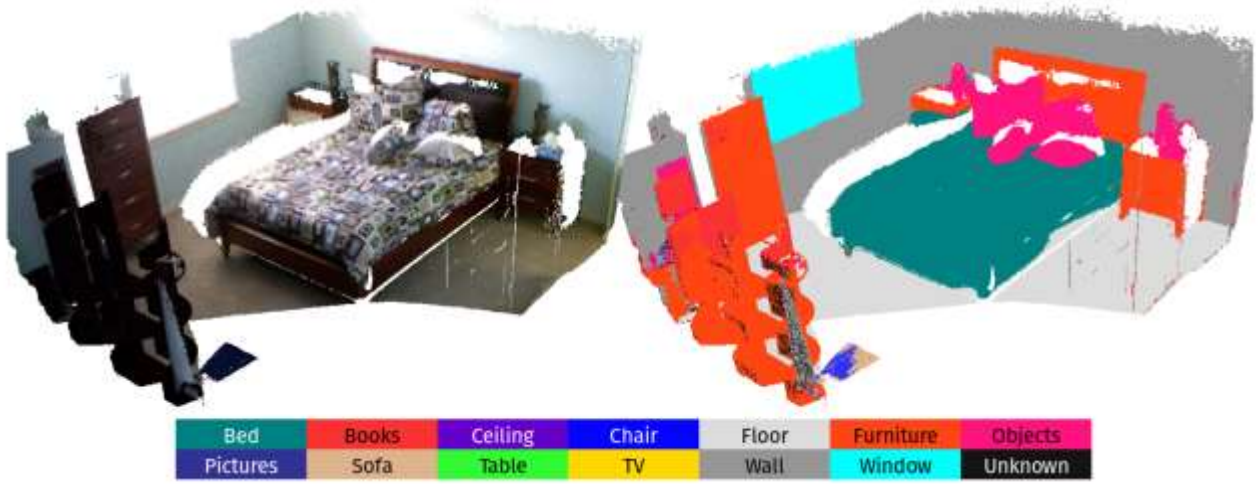
✘ Geometry

✘ Data Association

✓ Semantics

- Long et al., CVPR'12
- Zheng et al., ICCV'15
- Chen et al., ICLR'15
- Badrinarayanan et al., CVPR'15

Related Work: Semantic Mapping

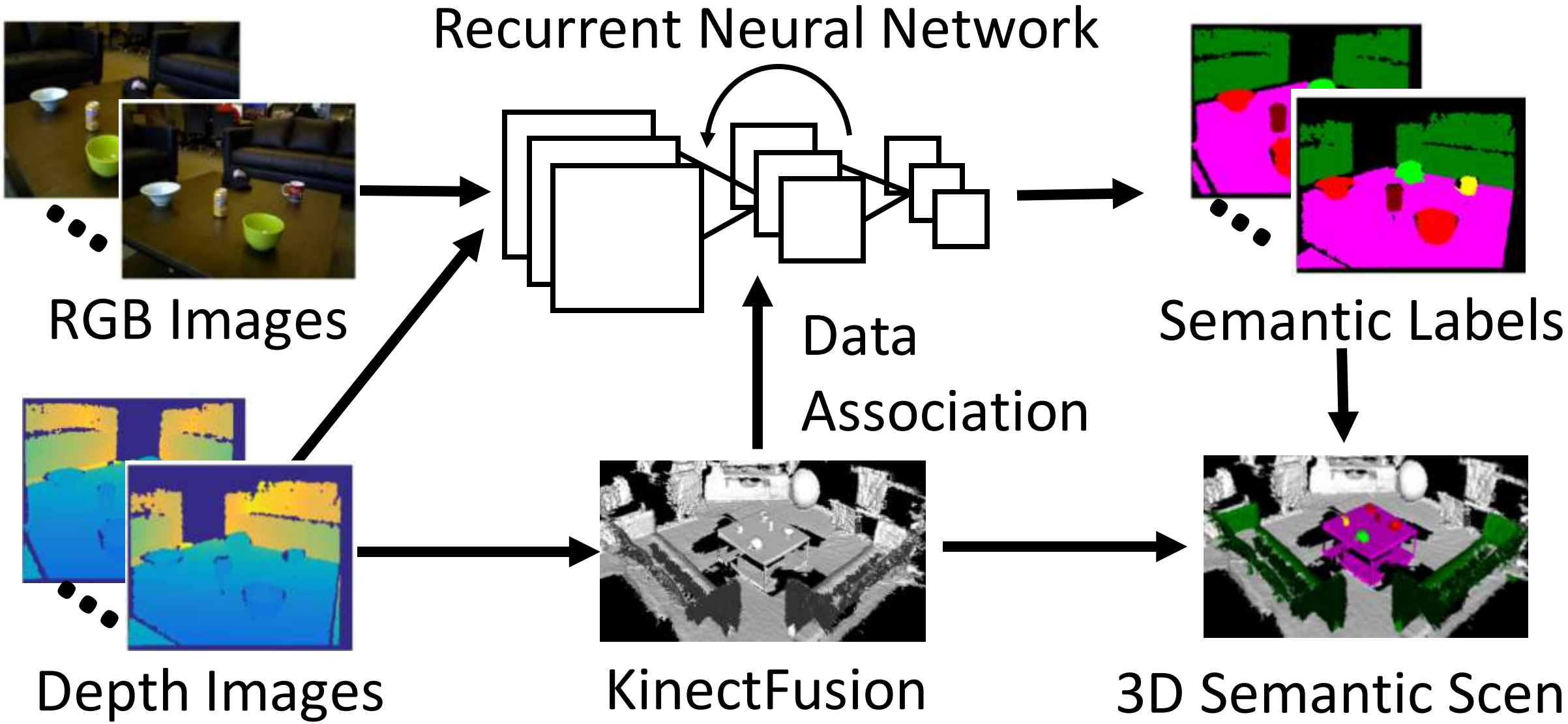


SemanticFusion

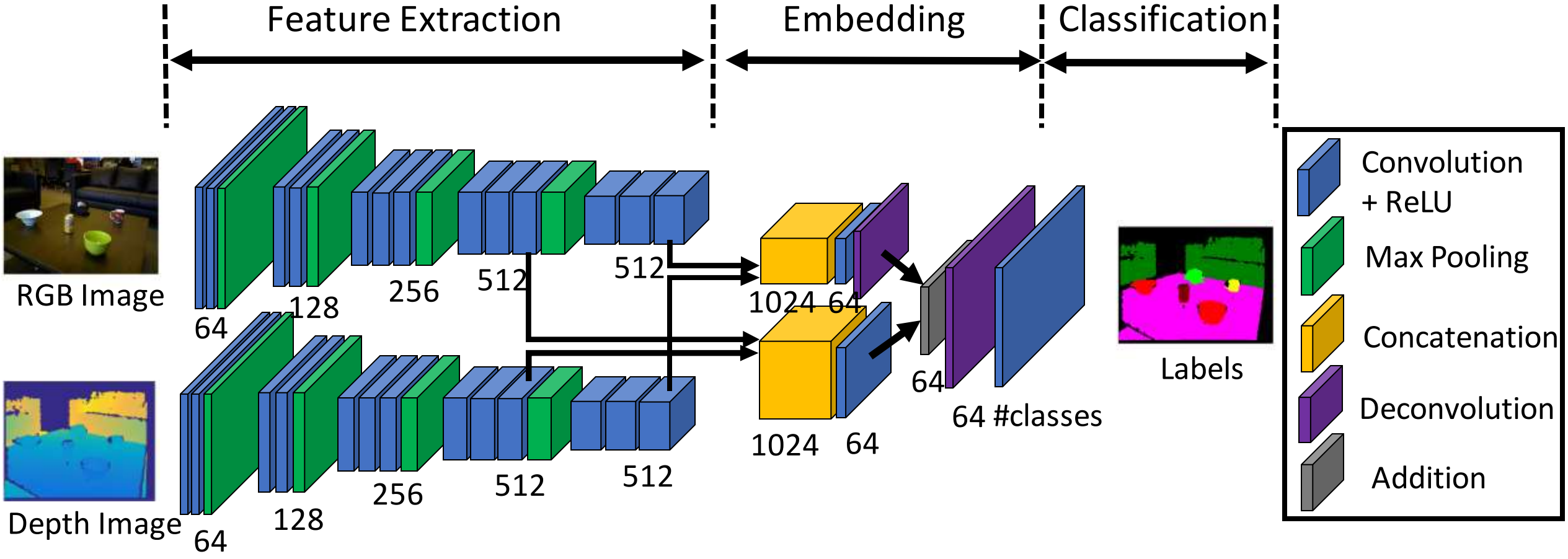
- ✓ Geometry
- ✓ Data Association
- ✓ Semantics

- Salas-Moreno et al., CVPR'13
- McCormac et al., ICRA'17

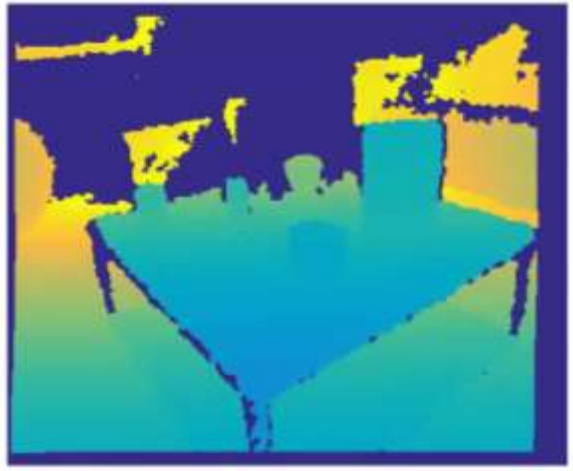
Our Contribution: DA-RNN

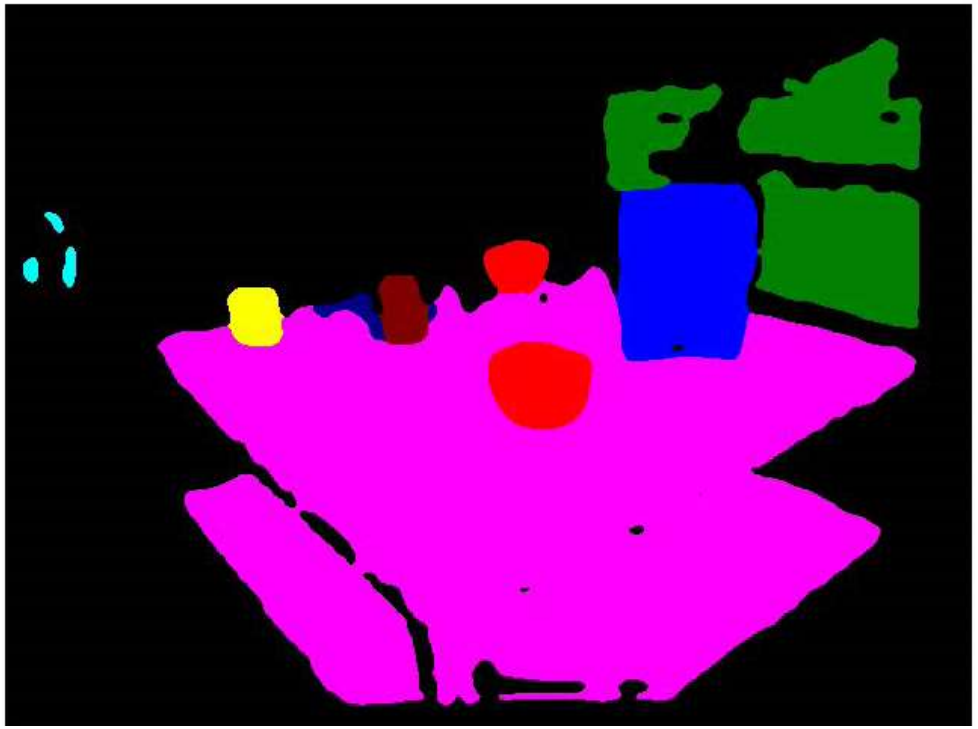


Single Frame Labeling with FCNs



Experiments: Results on RGB-D Scene Dataset [1]



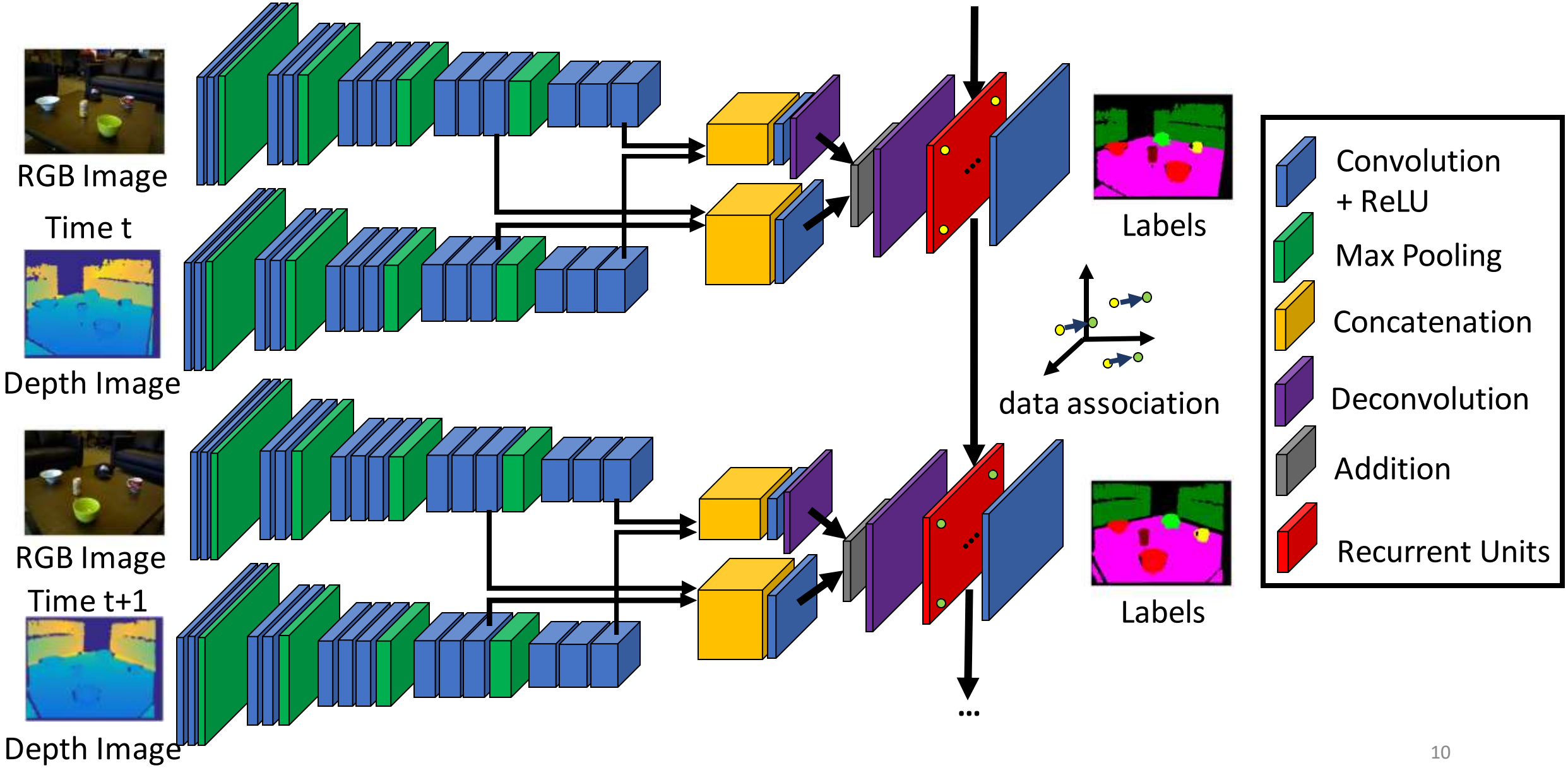


Methods	FCN [2]	Our FCN
Background	94.3	**96.1**
Bowl	78.6	**87.0**
Cap	61.2	**79.0**
Cereal Box	80.4	**87.5**
Coffee Mug	62.7	**75.7**
Coffee Table	93.6	**95.2**
Office Chair	67.3	**71.6**
Soda Can	73.5	**82.9**
Sofa	90.8	**92.9**
Table	84.2	**89.8**
MEAN	78.7	**85.8**

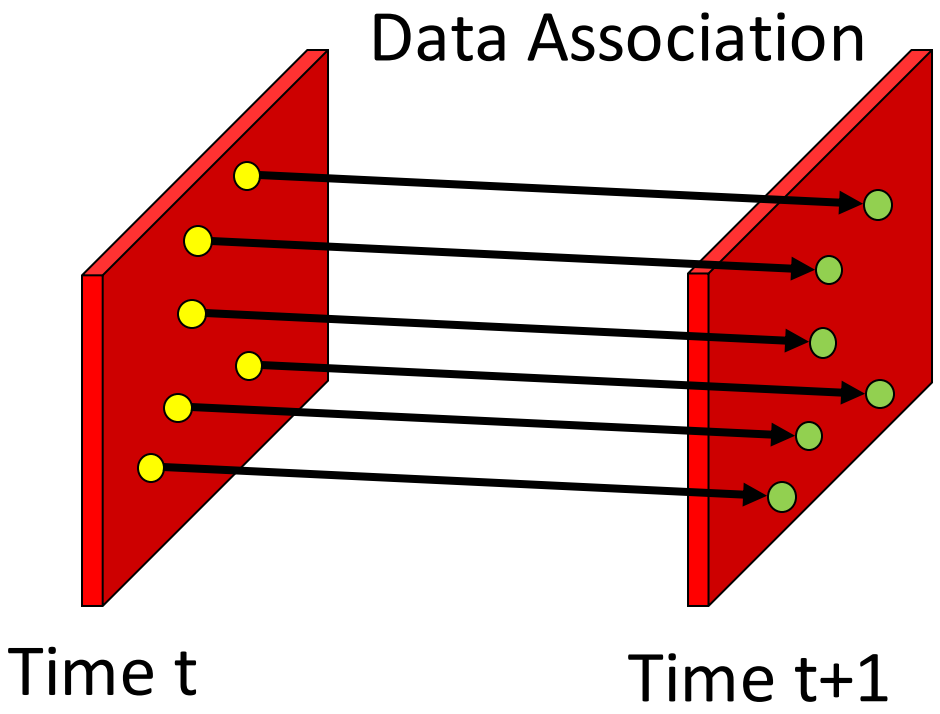
segmentation intersection over union (IoU)

[1] K. Lai, L. Bo and D. Fox. Unsupervised feature learning for 3D scene labeling. In ICRA'14.
[2] J. Long, E. Shelhamer and T. Darrell. Fully convolutional networks for semantic segmentation. In CVPR'15.

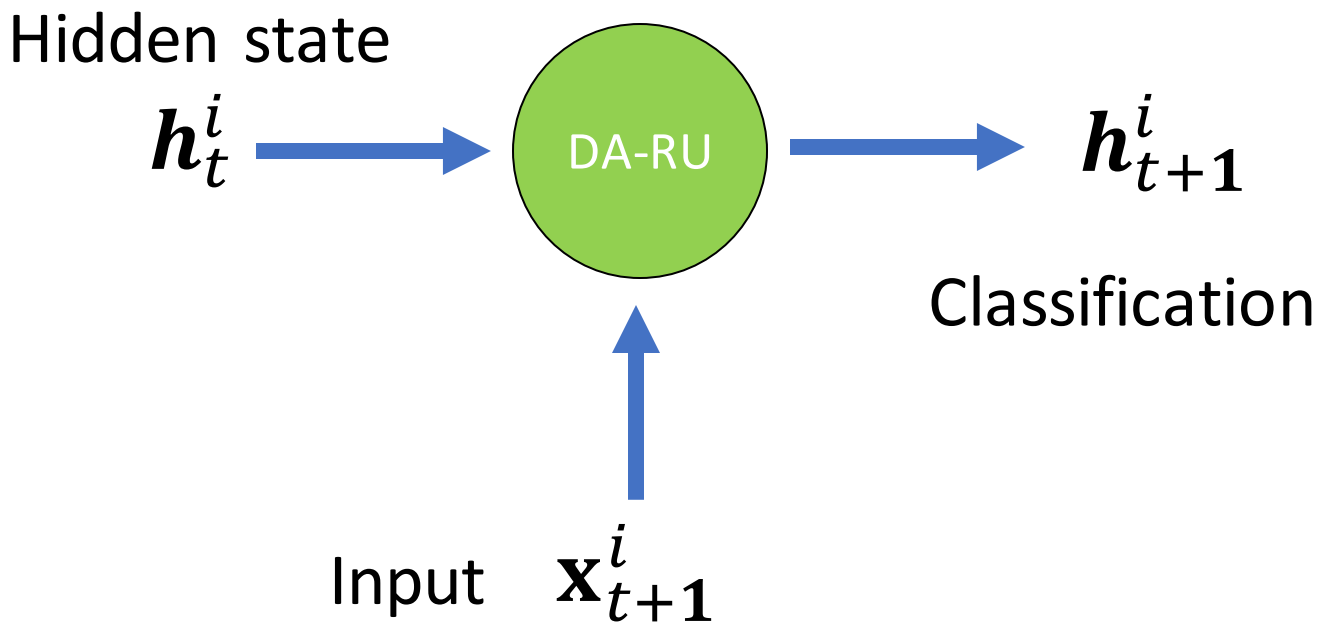
Video Semantic Labeling with DA-RNNs



Data Associated Recurrent Units (DA-RUs)

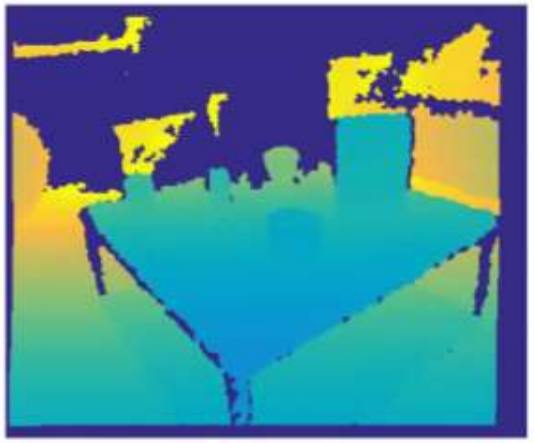


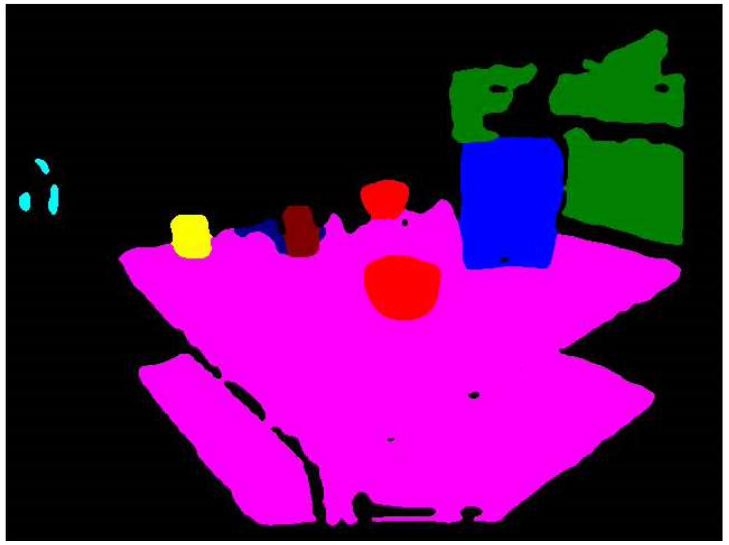
Recurrent layer

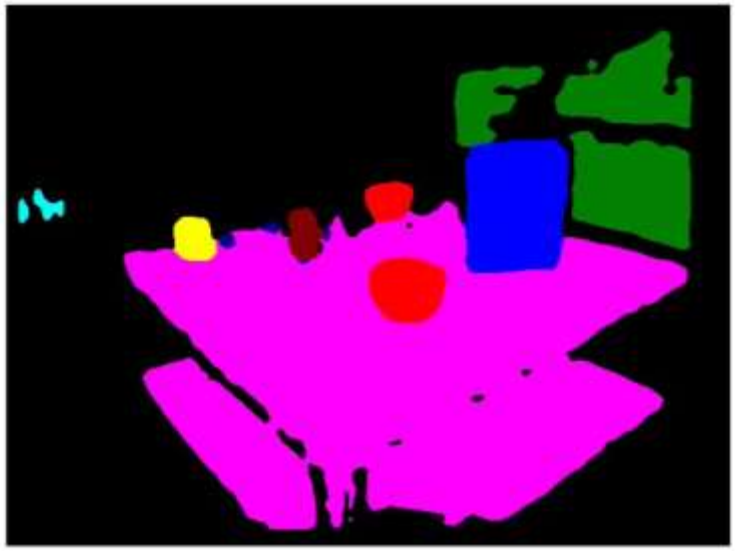


Weighted Moving Averaging
with learnable parameters

Experiments: Results on RGB-D Scene Dataset [1]







FCN

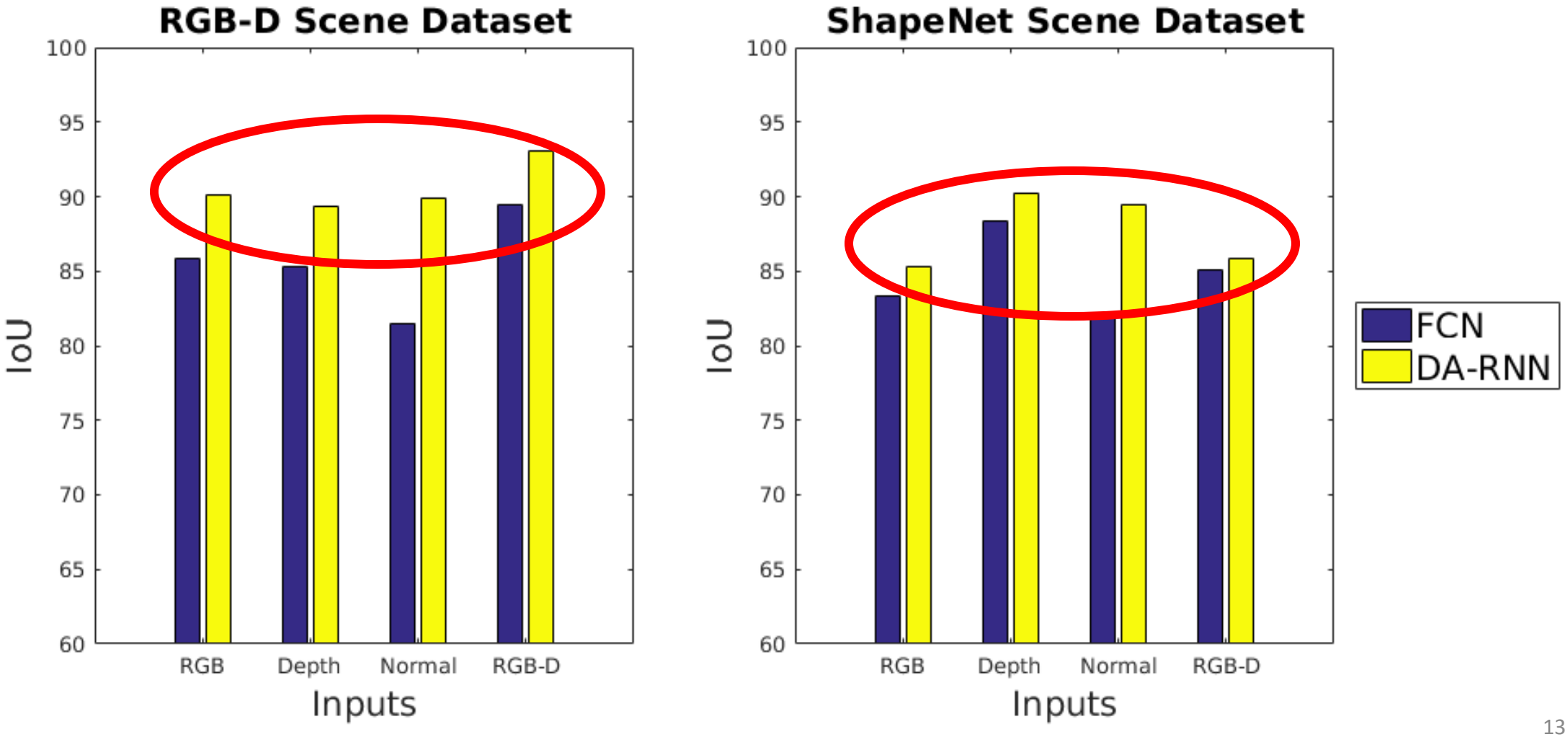
DA-RNN

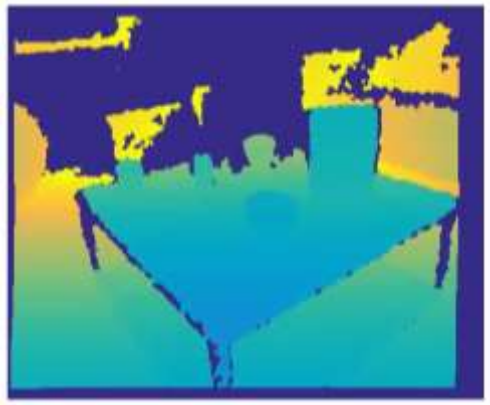
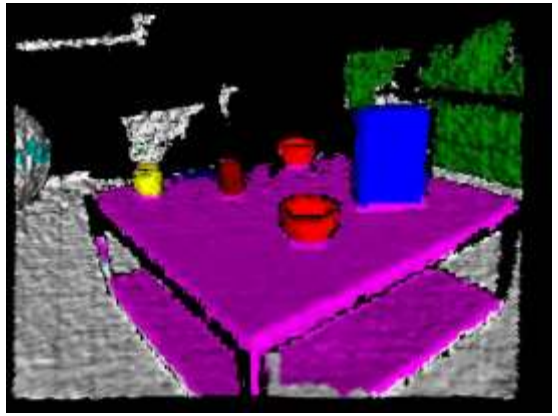
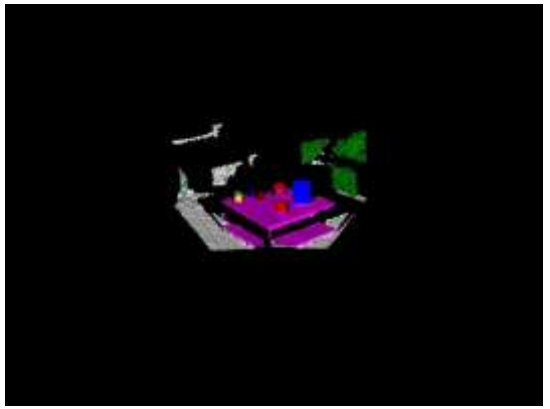
Methods	Our FCN	Our DA-RNN
Background	96.1	**97.6**
Bowl	87.0	**92.7**
Cap	79.0	**84.4**
Cereal Box	87.5	**88.3**
Coffee Mug	75.7	**86.3**
Coffee Table	95.2	**97.3**
Office Chair	71.6	**77.0**
Soda Can	82.9	**88.7**
Sofa	92.9	**95.6**
Table	89.8	**92.8**
MEAN	85.8	**90.1**

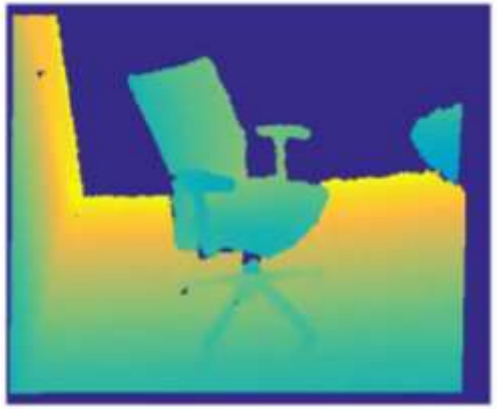
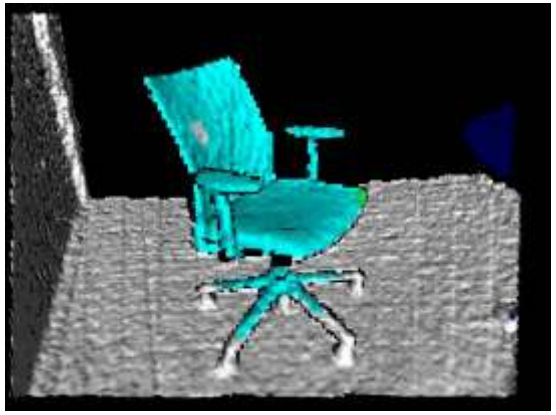
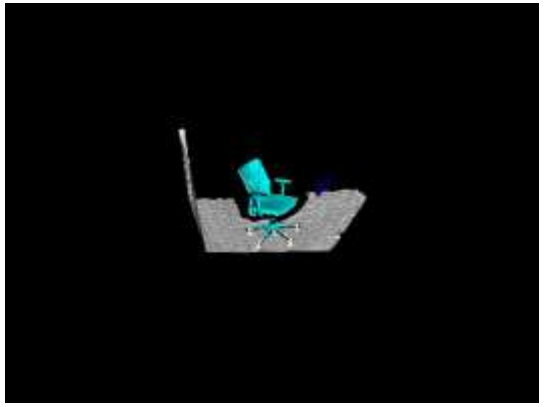
segmentation intersection over union (IoU)

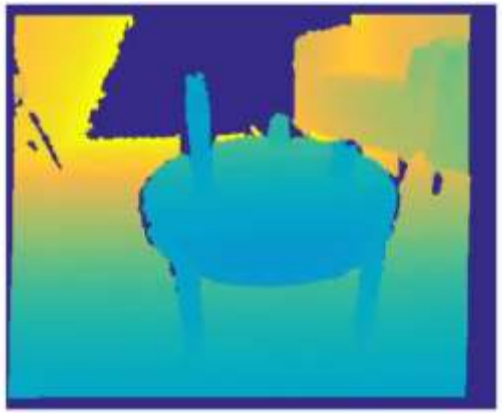
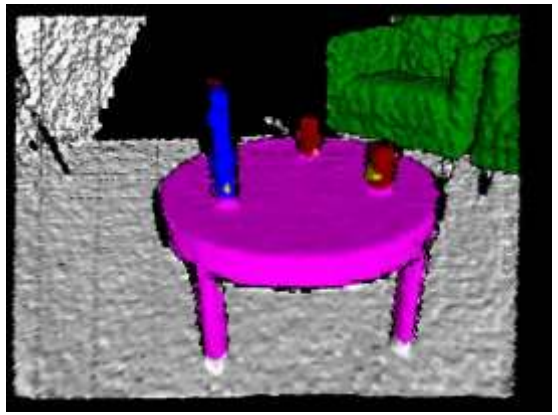
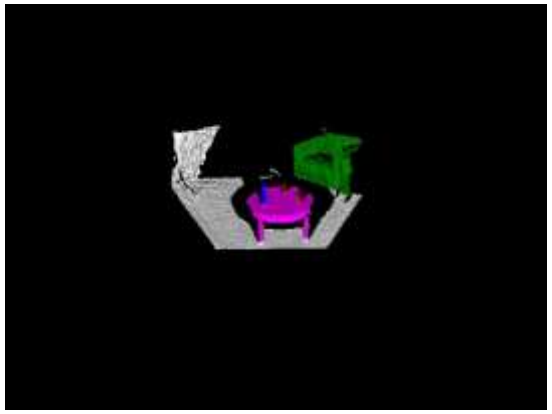
[1] K. Lai, L. Bo and D. Fox. Unsupervised feature learning for 3D scene labeling. In ICRA'14.

Experiments: Analysis on Network Inputs



RGB Images

Depth Images

Semantic Mapping

Conclusion

- DA-RNN, A novel framework for joint 3D mapping and semantic labeling
- A new recurrent neural network with data associated recurrent units for video semantic labeling
- Code and datasets are available online

Thank you!