Interpretable Structure-Evolving LSTM

Xiaodan Liang[1], Liang Lin[2], Xiaohui Shen[3], Jiashi Feng[3], Shuicheng Yan[3] , Eric P. Xing[1]
[1]Carnegie Mellon University, [2]Sun Yat-sen University, [3]National University of Singapore, [4]Adobe Research

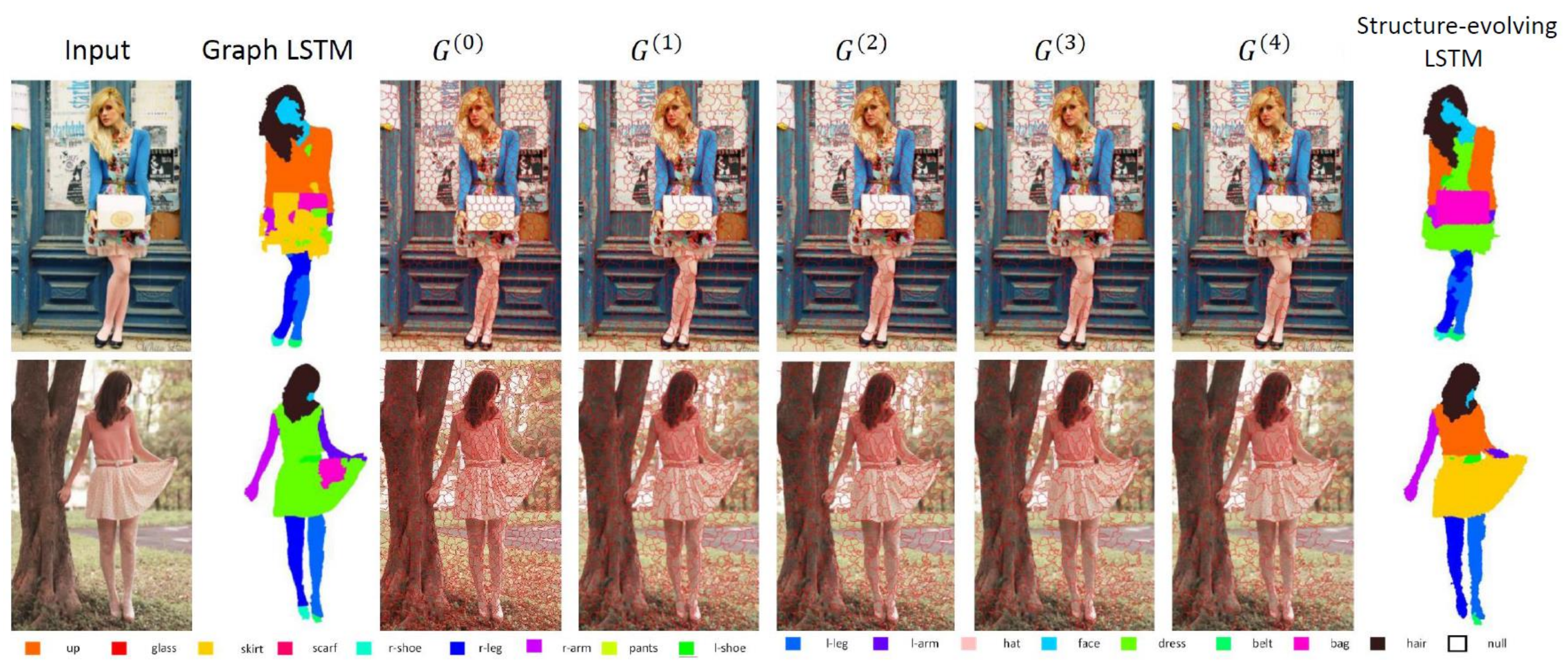
Carnegie Mellon University

IEEE 2017 Conference on Computer Vision and Pattern Recognition
CVPR July 21-26 HONOLULU 2017

Contributions of our structure-evolving LSTM

- A general framework for learning interpretable data representation via LSTM over hierarchal graph structures.
- It learns the intermediate interpretable multi-level graph structures in a progressive and stochastic way from data during the LSTM network optimization.
- It evolves the multi-level graph representations by stochastically merging the graph nodes with high compatibilities along the stacked LSTM layers.

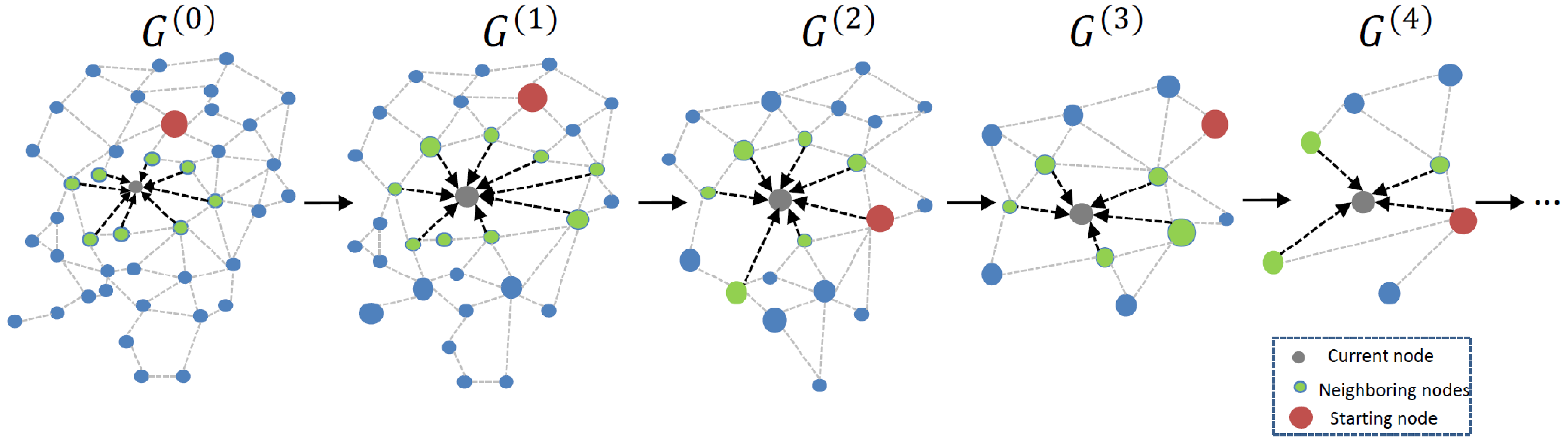
Result comparisons:



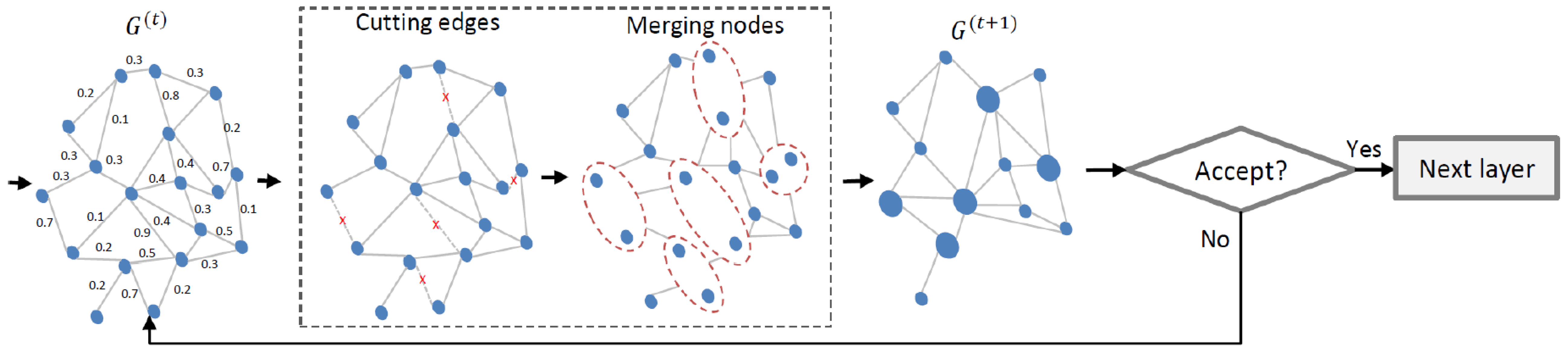
Method	head	torso	u-arms	l-arms	u-legs	l-legs	Bkg	Avg
DeepLab-LargeFOV [5]	78.09	54.02	37.29	36.85	33.73	29.61	92.85	51.78
DeepLab-LargeFOV-CRF [5]	80.13	55.56	36.43	38.72	35.50	30.82	93.52	52.95
HAZN [32]	80.79	59.11	43.05	42.76	38.99	34.46	93.59	56.11
Attention [6]	-	-	-	-	-	-	-	56.39
Grid LSTM [15]	81.85	58.85	43.10	46.87	40.07	34.59	85.97	55.90
Row LSTM [29]	82.60	60.13	44.29	47.22	40.83	35.51	87.07	56.80
Diagonal BiLSTM [29]	82.67	60.64	45.02	47.59	41.95	37.32	88.16	57.62
LG-LSTM [19]	82.72	60.99	45.40	47.76	42.33	37.96	88.63	57.97
Graph LSTM [18]	82.69	62.68	46.88	47.71	45.66	40.93	94.59	60.16
Graph LSTM (multi-scale superpixel maps) [18]	83.93	64.67	48.79	**49.44**	46.57	41.38	92.36	61.02
Structure-evolving LSTM (deterministic 0.5)	82.93	62.59	46.91	48.06	44.73	40.39	91.77	59.63
Structure-evolving LSTM (deterministic 0.7)	**84.16**	66.16	49.90	48.24	48.29	44.13	94.53	62.20
Structure-evolving LSTM (deterministic 0.9)	83.52	64.17	48.39	49.02	46.26	42.20	93.36	60.99
Structure-evolving LSTM	82.89	**67.15**	**51.42**	48.72	**51.72**	**45.91**	**97.18**	**63.57**

Interpretable Structure-evolving LSTM:

- ✓ Compared with existing LSTM, the structure-evolving LSTM has the capability of modeling long-range interactions using the dynamically evolved hierarchical graph topologies to capture the multi-level inherent correlations embedded in the data.
- ✓ It evolves the hierarchical graph structures with a stochastic and bottom-up node merging process.



- ✓ Stochastic node-merging process by a Metropolis-Hastings method:



- ✓ Example structure on semantic object parsing task: