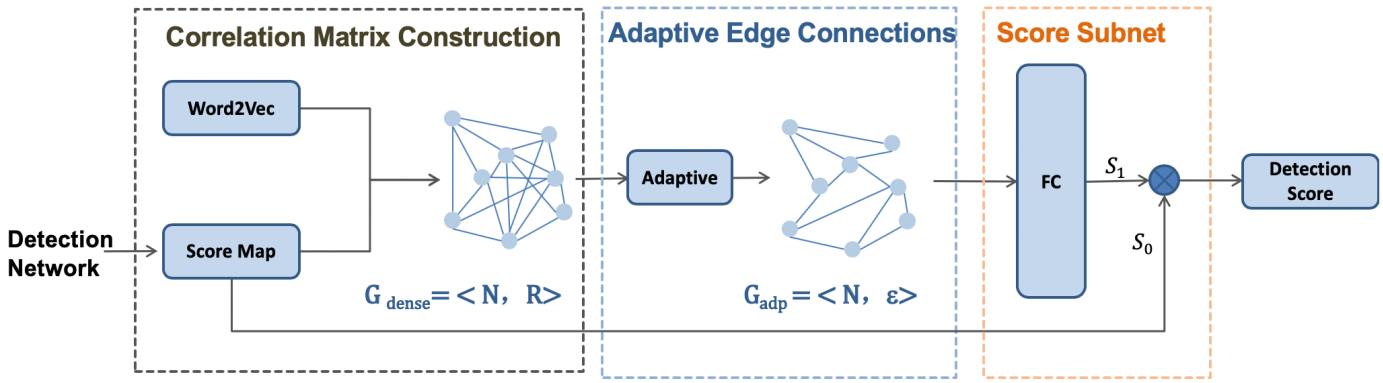
Adaptive Word Embedding Module for Semantic Reasoning in Large-scale Detection

Yu Zhang, Xiaoyu Wu*, Ruolin Zhu
Communication University of China

ABSTRACT

In recent years, convolutional neural networks have achieved rapid development in the field of object detection. However, due to the imbalance of data, high costs in labor and uneven level of data labeling, the overall performance of the previous detection network has dropped sharply when dataset extended to the large-scale with hundreds and thousands of categories. Therefore, one approach is to use the idea of transferring knowledge to reduce the cost of labeling tasks.

METHOD



The schema of Adaptive Word Embedding Module. Correlation Matrix Construction can obtain the dense region-to-region graph $G_{dense} = < N, R >$ according to the classification score of detection network and word2vec pre-trained model. Adaptive Edge Connections emphasize semantic features and suppress less useful ones to obtain an adaptive region-to-region graph $G_{adp} = < N, \varepsilon >$. Score Subnet performs score recalibration.

RESULTS

Experimental Results on COCO2017

Method	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L	AR	AR_{10}	AR_{100}	AR_S	AR_M	AR_L
Faster RCNN	28.7	49.5	29.8	10.7	32.7	44.1	27.2	39.4	40.3	18.2	45.9	60.0
Without adaption	23.0	36.2	25.1	8.5	26.0	32.2	25.1	38.1	38.9	16.5	44.7	58.6
Ours	**32.8**	**53.2**	**34.8**	**13.3**	**37.3**	**49.7**	**29.7**	**43.6**	**44.6**	**22.4**	**50.5**	**64.1**

Experimental Results of Large-scale Detection

Dataset	Method	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L	AR	AR_{10}	AR_{100}	AR_S	AR_M	AR_L
VG_{1000}	Faster RCNN	5.8	10.7	5.7	1.9	5.8	10.0	13.7	17.2	17.2	4.9	15.7	25.3
	HKRM	7.8	13.4	8.1	4.1	8.1	12.7	18.1	22.7	22.7	9.6	20.8	31.4
	Ours	**6.7**	**12.2**	**6.6**	**2.8**	**7.0**	**11.3**	**15.8**	**20.0**	**20.1**	**7.5**	**18.9**	**27.7**
ADE	Faster RCNN	7.9	14.7	7.5	2.1	5.8	13.2	10.6	14.2	14.4	4.5	11.9	22.4
	HKRM	10.3	18.0	10.4	7.9	16.8	16.8	13.6	18.3	18.5	7.1	15.5	28.4
	Ours	**10.0**	**18.0**	**9.9**	**3.9**	**8.0**	**16.7**	**13.7**	**18.0**	**18.2**	**7.0**	**15.5**	**28.1**

Visualization of Experimental Results

CONLUSIONS

We presented a novel Adaptive Word Embedding Module for global semantic reasoning. The experiments indicated the superior performance of our approach over the baseline detection network on COCO, VG and ADE. We demonstrate that Adaptive Word Embedding Module can alleviate the problems of large-scale object detection.