Paper reading-Jing Liu-5.23

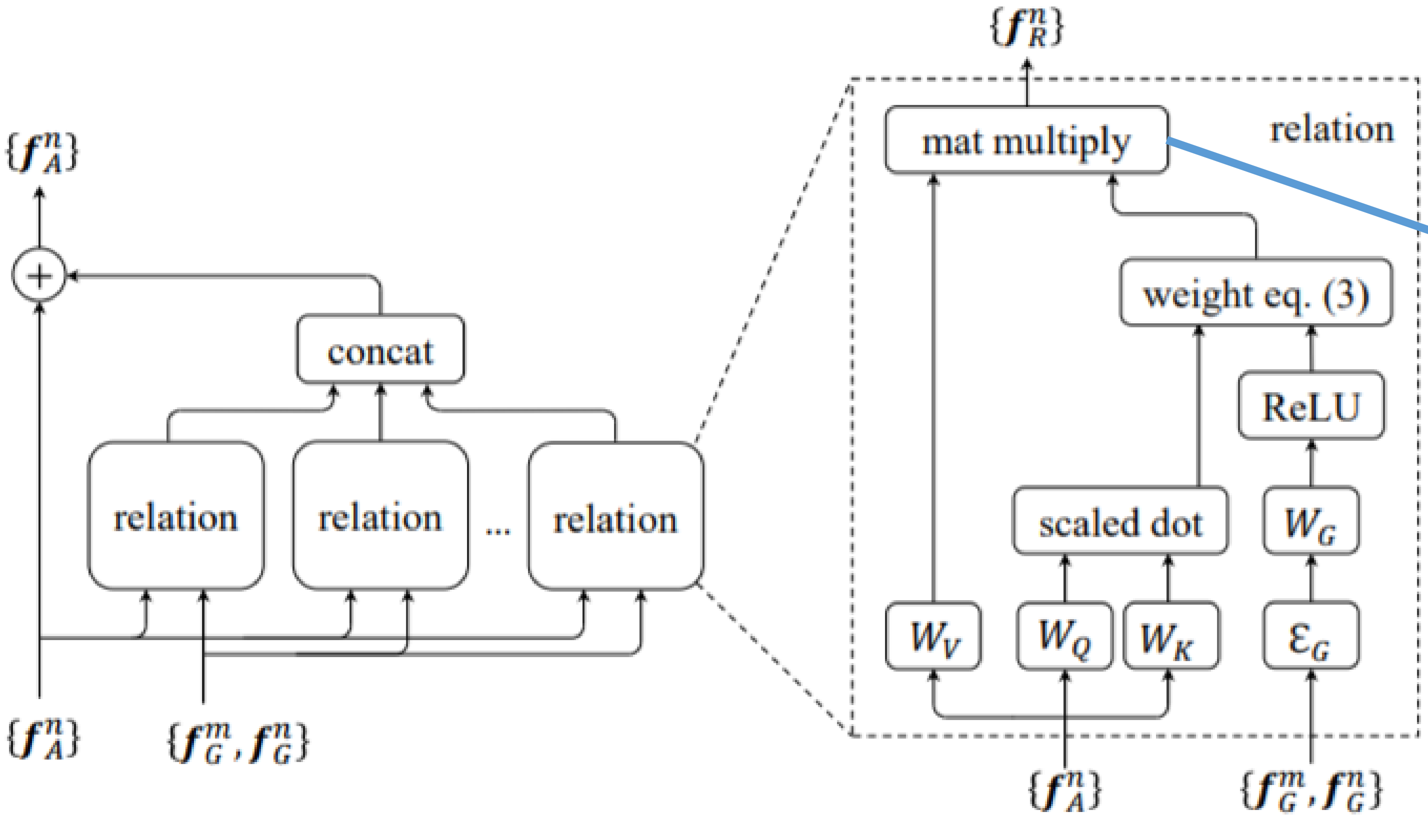
Relation Networks for Object Detection

Han Hu* Jiayuan Gu*,† Zheng Zhang* Jifeng Dai Yichen Wei

Microsoft Research Asia

https://arxiv.org/abs/1711.11575

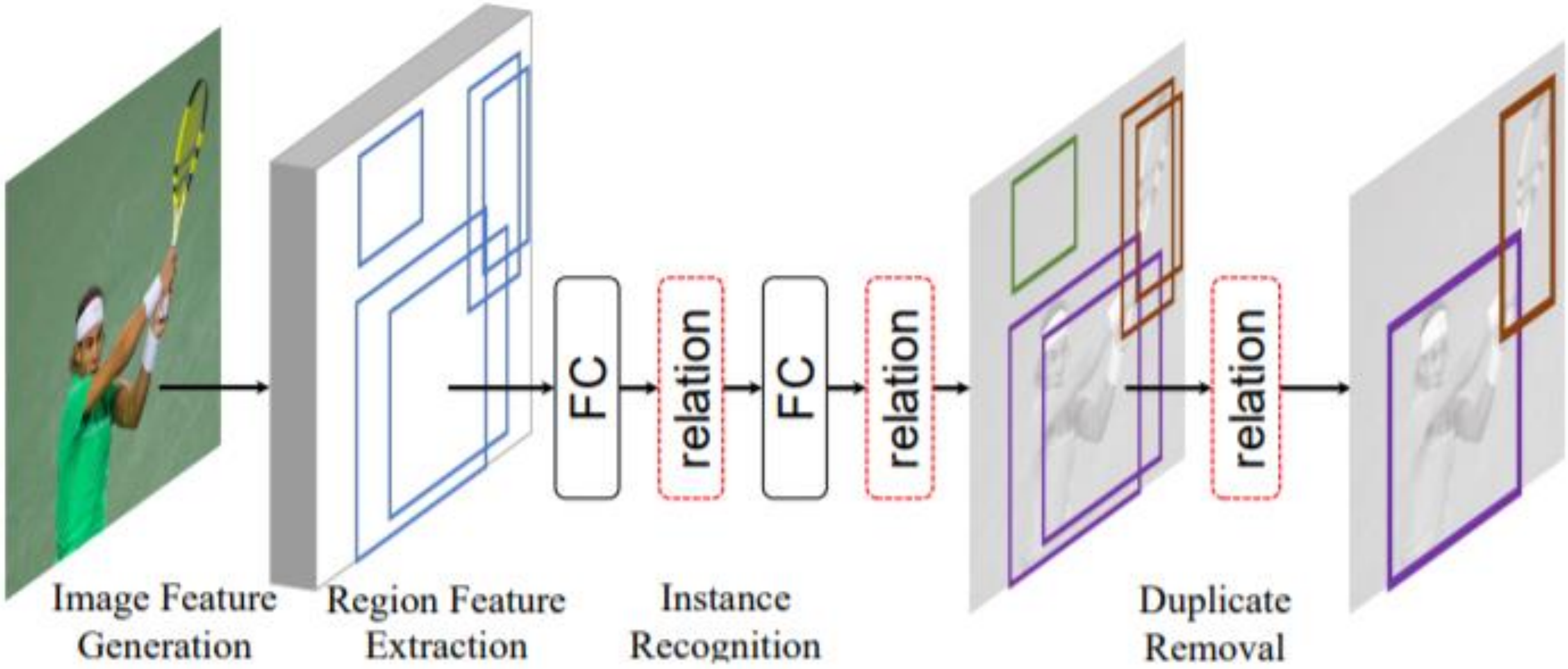
Object Relation Module



$$\mathbf{f}_R(n) = \sum_m \omega^{mn} \cdot (W_V \cdot \mathbf{f}_A^m).$$

Hu H, Gu J, Zhang Z, et al. Relation Networks for Object Detection[J]. arXiv preprint arXiv:1711.11575, 2017.

Object relation module (illustrated as red dashed boxes) can be conveniently adopted to improve both instance recognition and duplicate removal steps, resulting in an end-to-end object detector.



Hu H, Gu J, Zhang Z, et al. Relation Networks for Object Detection[J]. arXiv preprint arXiv:1711.11575, 2017.

1： generates full image features
2： generates regional features
3： performs instance recognition
4： performs duplicate removal

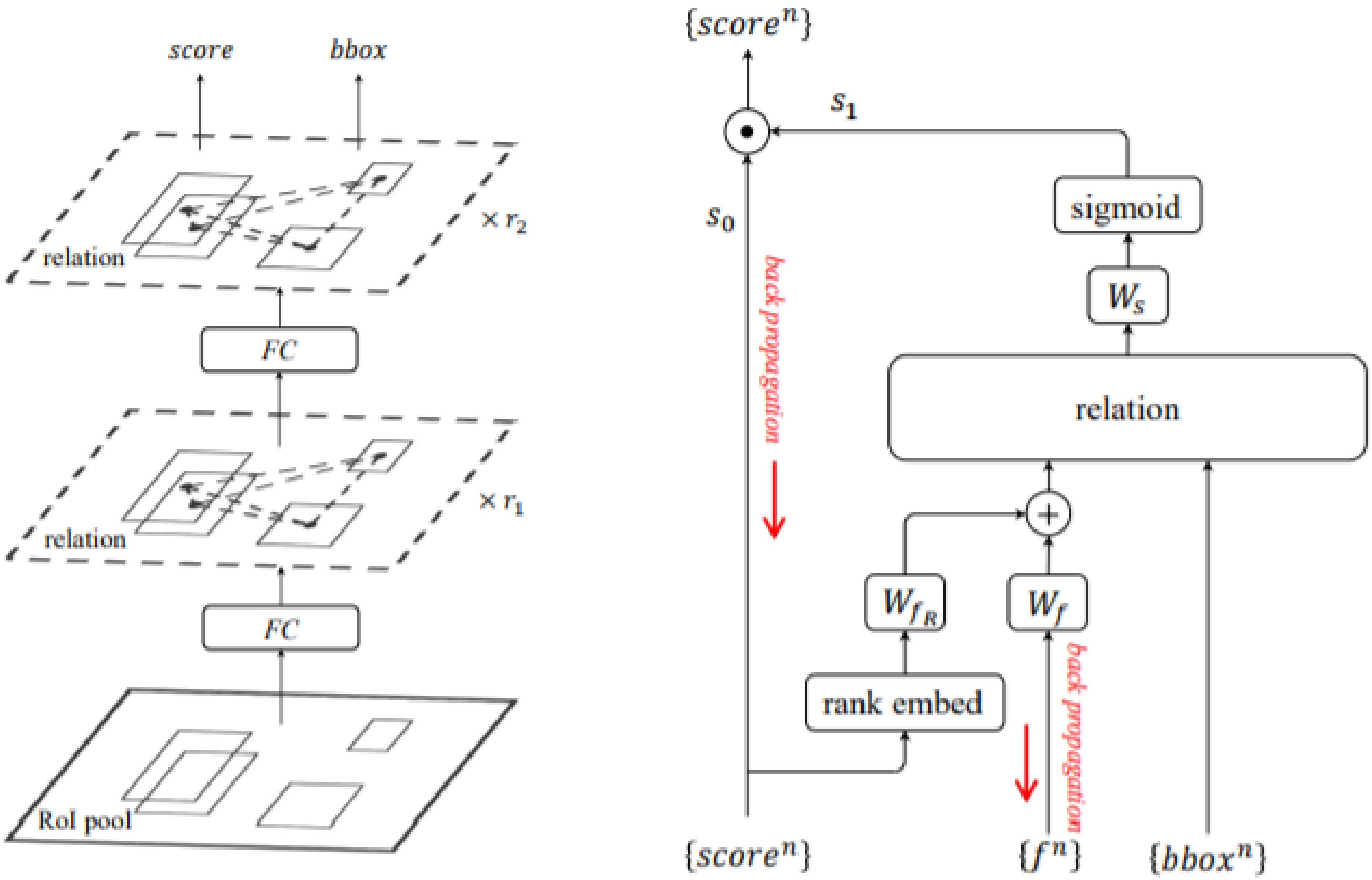
Instance Recongnition

$$
\begin{aligned}
RoI\_Feat_n &\xrightarrow{FC} \quad 1024 \\
&\xrightarrow{FC} \quad 1024 \\
&\xrightarrow{LINEAR} (score_n, bbox_n)
\end{aligned}
$$

$$
\begin{aligned}
\{RoI\_Feat_n\}_{n=1}^{N} &\xrightarrow{FC} \quad 1024 \cdot N \xrightarrow{\{RM\}^{r_1}} 1024 \cdot N \\
&\xrightarrow{FC} \quad 1024 \cdot N \xrightarrow{\{RM\}^{r_2}} 1024 \cdot N \\
&\xrightarrow{LINEAR} \{(score_n, bbox_n)\}_{n=1}^{N}
\end{aligned}
$$

Hu H, Gu J, Zhang Z, et al. Relation Networks for Object Detection[J]. arXiv preprint arXiv:1711.11575, 2017.

Duplicate Remove



Hu H, Gu J, Zhang Z, et al. Relation Networks for Object Detection[J]. arXiv preprint arXiv:1711.11575, 2017.

Experiments

1 Ablation study on relation module structure and parameter

2fc baseline	(a): usage of geometric feature			(b): number of relations N_r						(c): number of relation modules $\{r_1, r_2\}$				
	none	unary	ours*	1	2	4	8	16*	32	$\{1,0\}$	$\{0,1\}$	$\{1,1\}$*	$\{2,2\}$	$\{4,4\}$
29.6	30.3	31.1	**31.9**	30.5	30.6	31.3	31.7	**31.9**	31.7	31.7	31.4	31.9	32.5	**32.8**

Table 1. Ablation study of relation module structure and parameters (* for default). mAP@all is reported.

Hu H, Gu J, Zhang Z, et al. Relation Networks for Object Detection[J]. arXiv preprint arXiv:1711.11575, 2017.

2 relation module improve the instance recognition performance

head	mAP	mAP_{50}	mAP_{75}	# params	# FLOPS
(a) 2fc (1024)	29.6	50.9	30.1	38.0M	80.2B
(b) 2fc (1432)	29.7	50.3	30.2	44.1M	82.0B
(c) 3fc (1024)	29.0	49.4	29.6	39.0M	80.5B
(d) 2fc+res $\{r_1, r_2\}$=$\{1, 1\}$	29.9	50.6	30.5	44.0M	82.1B
(e) 2fc (1024) + global	29.6	50.3	30.8	38.2M	82.2B
(f) 2fc+RM $\{r_1, r_2\}$=$\{1, 1\}$	**31.9**	53.7	33.1	44.0M	82.6B
(g) 2fc+res $\{r_1, r_2\}$=$\{2, 2\}$	29.8	50.5	30.5	50.0M	84.0B
(h) 2fc+RM $\{r_1, r_2\}$=$\{2, 2\}$	**32.5**	54.0	33.8	50.0M	84.9B

Table 2. Comparison of various heads with similar complexity.

Hu H, Gu J, Zhang Z, et al. Relation Networks for Object Detection[J]. arXiv preprint arXiv:1711.11575, 2017.

3 relation module for duplicate removal

NMS	ours	rank f_R		appearance f	geometric $bbox$	
	$\{f_R, f, bbox\}$	*none*	s_0	*none*	*none*	*unary*
29.6	**30.3**	26.6	28.3	29.9	28.1	28.2

Table 3. Ablation study of input features for duplicate removal network (*none* indicates without such feature).

Hu H, Gu J, Zhang Z, et al. Relation Networks for Object Detection[J]. arXiv preprint arXiv:1711.11575, 2017.

4 What is learnt in relation module

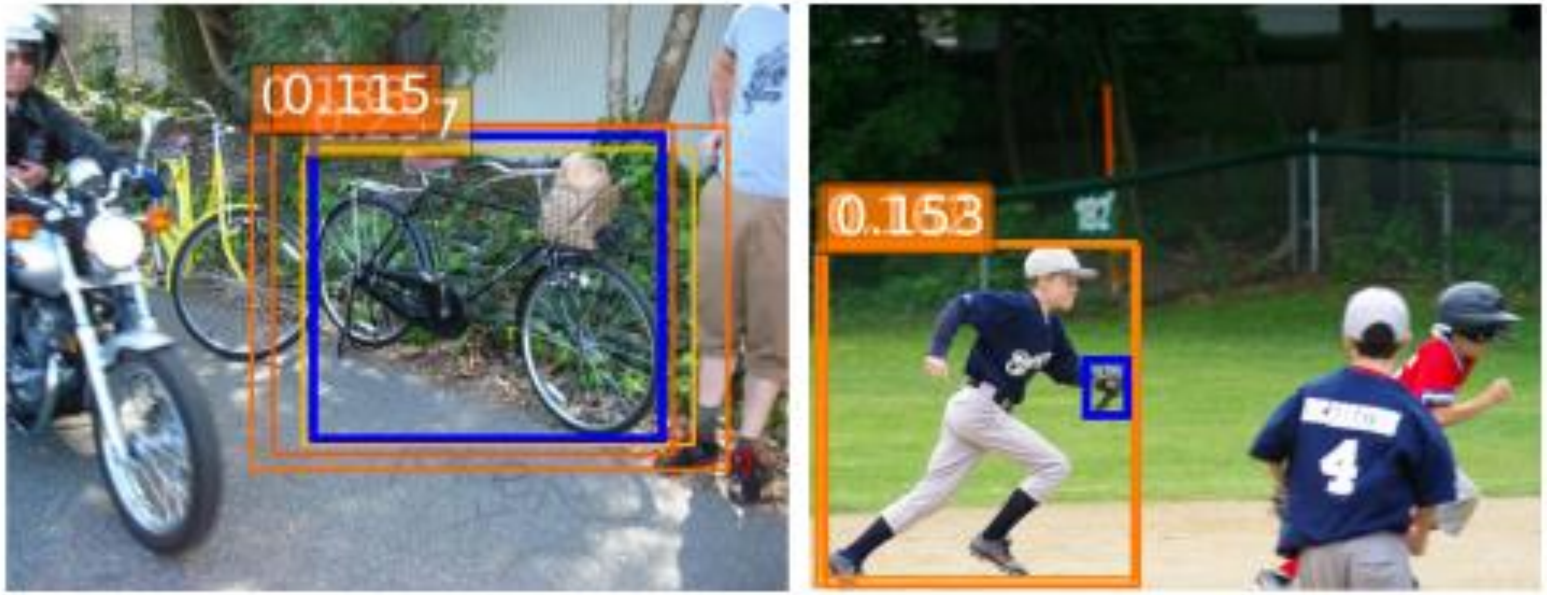


Figure 4. Representative examples with high relation weights in Eq. (3). The reference object n is blue. The other objects contributing a high weight (shown on the top-left) are yellow.

Hu H, Gu J, Zhang Z, et al. Relation Networks for Object Detection[J]. arXiv preprint arXiv:1711.11575, 2017.