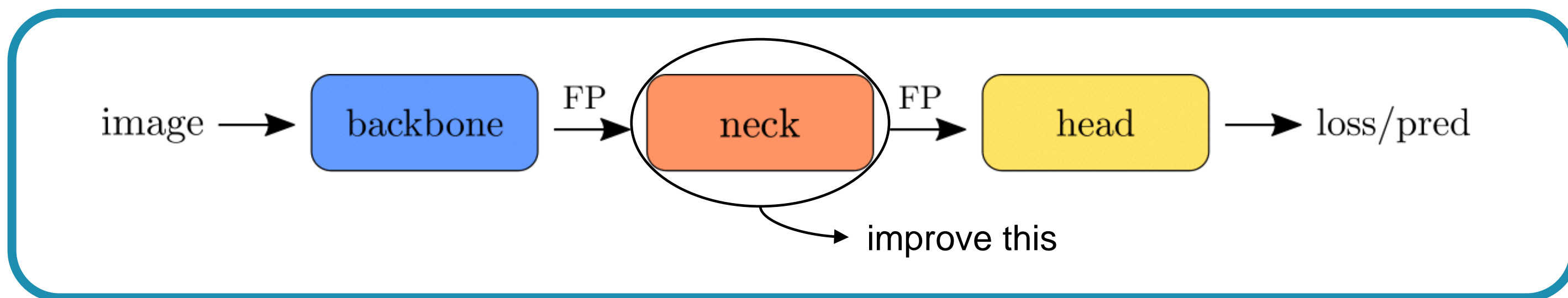


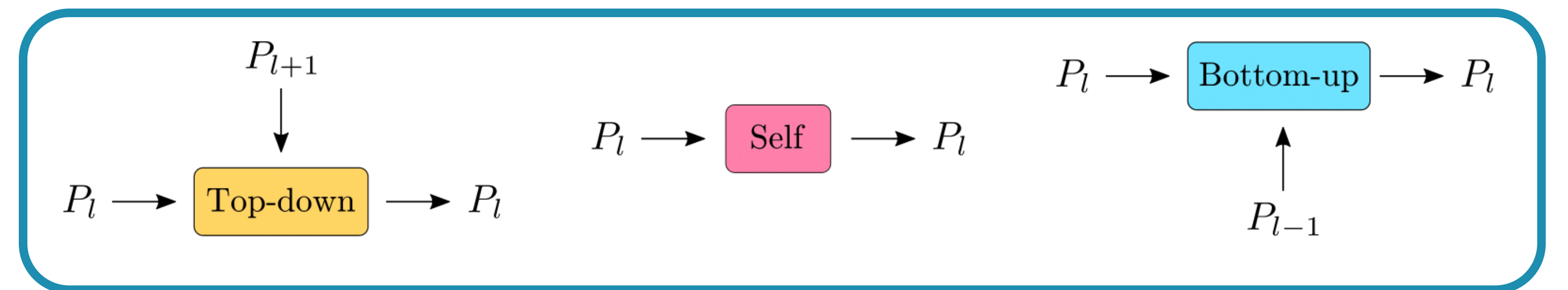
# Trident Pyramid Networks for Object Detection



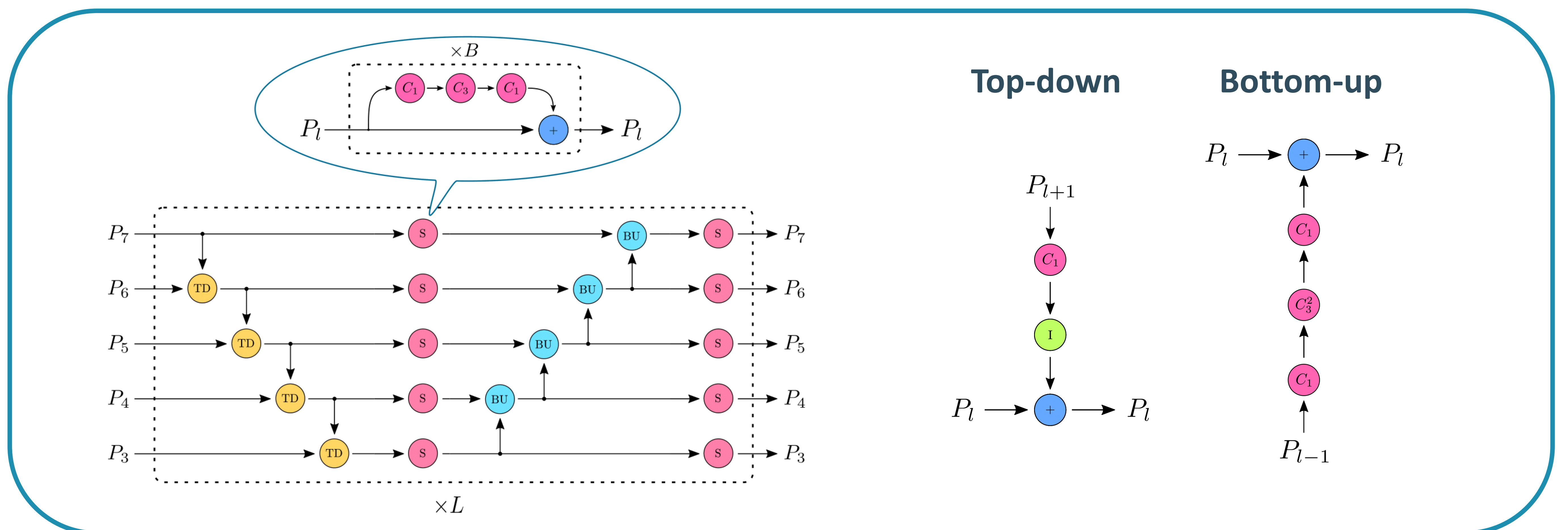
## Problem setting



## Building blocks



## Trident Pyramid Network (TPN)



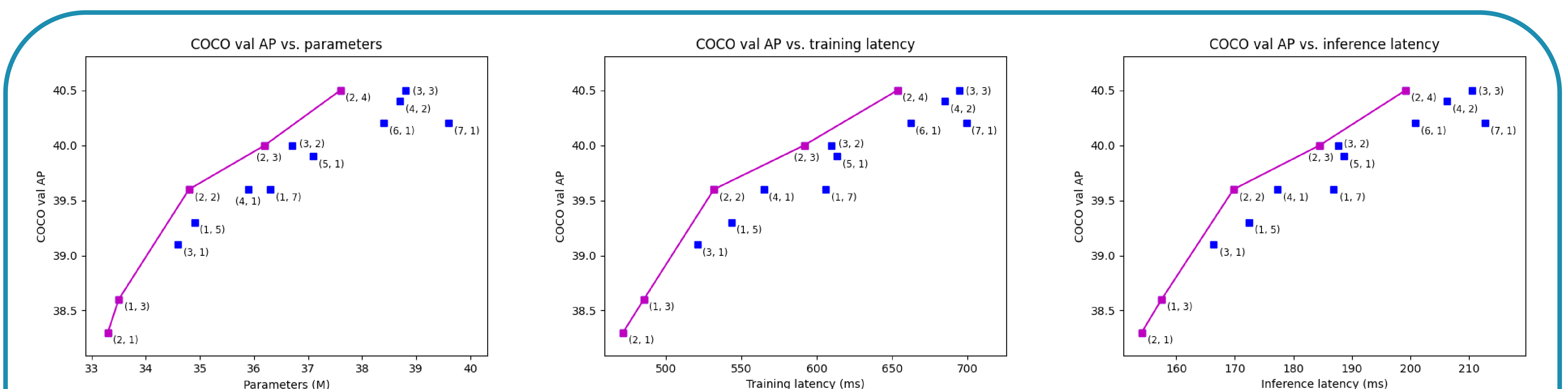
## Main results

Back	Neck	L	B	C	AP	AP <sub>50</sub>	AP <sub>75</sub>	AP <sub>S</sub>	AP <sub>M</sub>	AP <sub>L</sub>	Params	tFPS	tMem	iFPS	iMem
R50	TPN	1	7	1	41.3	60.5	44.2	26.3	45.9	52.5	36.3 M	1.7	3.31 GB	5.3	0.50 GB
R50	TPN	2	3	1	41.6	60.9	44.6	26.4	45.8	53.2	36.2 M	1.7	3.21 GB	5.5	0.50 GB
R50	TPN	3	2	1	<b>41.8</b>	61.1	44.4	26.2	46.1	53.7	36.7 M	1.6	3.27 GB	5.3	0.50 GB
R50	TPN	5	1	1	<b>41.8</b>	<b>61.2</b>	<b>45.0</b>	26.0	<b>46.3</b>	53.4	37.1 M	1.6	3.22 GB	5.3	0.50 GB
R50	BiFPN	7	-	1	41.3	<b>61.2</b>	43.7	<b>27.1</b>	45.2	<b>53.8</b>	34.7 M	1.8	3.22 GB	6.0	0.49 GB
R50	bFPN	1	14	1	39.6	60.3	42.4	24.2	43.5	51.3	36.1 M	1.7	3.26 GB	5.4	0.49 GB
R50	hFPN	1	14	1	40.0	60.2	43.0	25.6	43.9	51.1	36.1 M	1.7	3.26 GB	5.4	0.49 GB
R101	FPN	1	-	4	40.1	60.1	42.8	24.0	44.0	52.7	55.1 M	1.4	3.20 GB	4.1	0.57 GB
R101	TPN	1	2	1	40.9	61.0	44.2	25.0	45.3	52.6	51.7 M	1.6	3.20 GB	4.6	0.55 GB

**Contribution 1:** Our TPN neck outperforms the BiFPN neck (main baseline) by 0.5 AP on COCO.

**Contribution 2:** It is more beneficial to put additional computation into the neck than into the backbone or head.

## Optimal neck configurations



**Contribution 3:** Balanced configurations with  $L, B \geq 2$  outperform the unbalanced configurations  $(L, 1)$  and  $(1, B)$  under similar computation budgets. Existing works (e.g. BiFPN) mostly use  $(L, 1)$ .