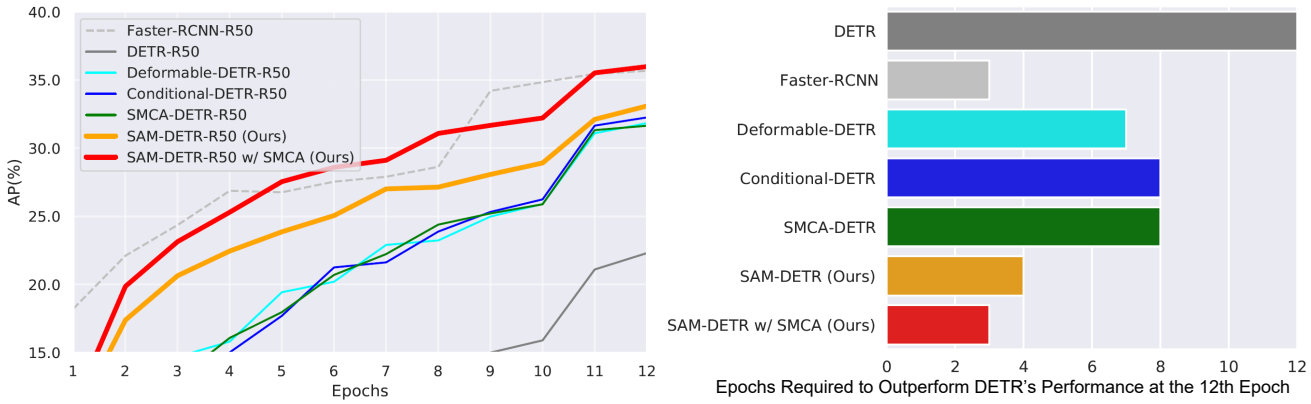
Supplementary Materials
Accelerating DETR Convergence via Semantic-Aligned Matching

Gongjie Zhang[1] Zhipeng Luo[1,2] Yingchen Yu[1] Kaiwen Cui[1] Shijian Lu[*1]
[1]Nanyang Technological University, Singapore [2]SenseTime Research
{gongjiezhang, shijian.lu}@ntu.edu.sg {zhipeng001, yingchen001, kaiwen001}@e.ntu.edu.sg

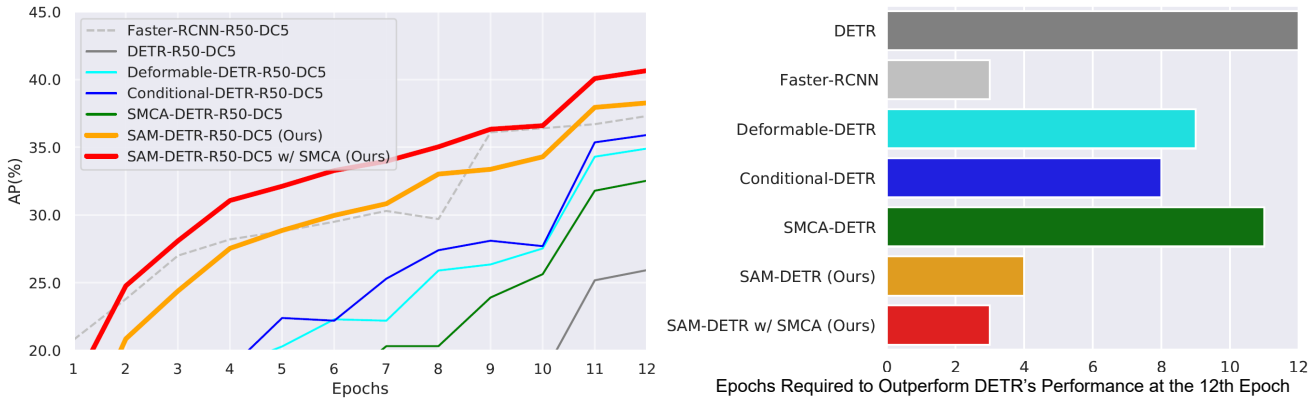
A. Additional Results on Convergence Speed

We further demonstrate SAM-DETR's superior convergence by providing convergence curves and comparing the number of training epochs to reach the baseline accuracy under various setups.

12-Epoch, ResNet-50:



12-Epoch, ResNet-50-DC5:



It can be observed that our proposed SAM-DETR achieves the best convergence speed among all DETR-like detectors. Compared with the original DETR baseline, our approach reduces the number of training epochs by three times to reach its performance. In contrast, it takes the original DETR baseline ∼150 epochs and ∼350 epochs with ResNet-50 and ResNet-50-DC5, respectively, to achieve the performance of SAM-DETR w/ SMCA trained for just 12 epochs. The results demonstrate SAM-DETR's significant improvement in terms of convergence speed.

* Corresponding author.