

Table 1: Instance segmentation on COCO AP, Panoptic segmentation on COCO PQ, and Semantic segmentation on ADE20K mIoU.

Method	Params	Backbone	Backbone Pre-training Dataset	Dataset	Detection Dataset
<b>Instance segmentation on COCO AP</b>					
Mask2Former [3]	216M	SwinL	IN-22K-14M	–	50.1
Soft Teacher [36]	284M	SwinL	IN-22K-14M	O365	51.9
SwinV2-G-HTC++ [23]	3.0B	SwinV2-G	IN-22K-ext-70M [23]	O365	53.4
MasK DINO (Ours)	223M	SwinL	IN-22K-14M	–	52.6
MasK DINO (Ours)	223M	SwinL	IN-22K-14M	O365	54.5(+1.9)
<b>Panoptic segmentation on COCO PQ</b>					
Panoptic SegFormer [19]	–M	SwinL	IN-22K-14M	–	55.8
Mask2Former [3]	216M	SwinL	IN-22K-14M	–	57.8
MasK DINO (ours)	223M	SwinL	IN-22K-14M	58.4(+0.6)	–
MasK DINO (ours)	223M	SwinL	IN-22K-14M	O365	59.4(+1.6)
<b>Semantic segmentation on ADE20K mIoU</b>					
Mask2Former [3]	215M	SwinL	IN-22K-14M	–	56.1
SeMask-L MSFaPN-Mask2Former [14]	–M	SwinL-FaPN	IN-22K-14M	–	–
SwinV2-G-UpperNet [23]	3.0B	SwinV2-G	IN-22K-ext-70M [23]	–	59.3
MasK DINO (ours)	223M	SwinL	IN-22K-14M	–	56.6
MasK DINO (ours)	223M	SwinL	IN-22K-14M	O365	59.5

Table 2: Replicated results for instance segmentation on COCO AP, panoptic segmentation on COCO PQ, and semantic segmentation on ADE20K mIoU.

Method	Params	Backbone	Backbone Pre-training Dataset	Dataset	Detection Dataset
<b>Instance segmentation on COCO AP</b>					
Mask2Former [3]	216M	SwinL	IN-22K-14M	–	49.7
Soft Teacher [36]	284M	SwinL	IN-22K-14M	O365	51.5
SwinV2-G-HTC++ [23]	3.0B	SwinV2-G	IN-22K-ext-70M [23]	O365	53.1
MasK DINO (Ours)	223M	SwinL	IN-22K-14M	–	52.3
MasK DINO (Ours)	223M	SwinL	IN-22K-14M	O365	54.2(+0.9)
<b>Panoptic segmentation on COCO PQ</b>					
Panoptic SegFormer [19]	–M	SwinL	IN-22K-14M	–	55.4
Mask2Former [3]	216M	SwinL	IN-22K-14M	–	57.5
MasK DINO (ours)	223M	SwinL	IN-22K-14M	58.1(+0.3)	–
MasK DINO (ours)	223M	SwinL	IN-22K-14M	O365	59.1(+1.6)
<b>Semantic segmentation on ADE20K mIoU</b>					
Mask2Former [3]	215M	SwinL	IN-22K-14M	–	55.8
SeMask-L MSFaPN-Mask2Former [14]	–M	SwinL-FaPN	IN-22K-14M	–	–
SwinV2-G-UpperNet [23]	3.0B	SwinV2-G	IN-22K-ext-70M [23]	–	59.0
MasK DINO (ours)	223M	SwinL	IN-22K-14M	–	56.3
MasK DINO (ours)	223M	SwinL	IN-22K-14M	O365	59.2