Table 7: Comparisons with state-of-the-arts on validation set: The evaluation metric is MR and the input scale is 1x. The top three results are highlighted in red, green and blue, respectively.

Method	Backbone	Reasonable	Heavy	Partial	Bare
FRCNN [49]	VGG-16	15.4%	-	-	-
FRCNN+Seg [49]	VGG-16	14.8%	-	-	-
TLL [37]	ResNet-50	15.5%	53.6%	17.2%	10.0%
TLL+MRF [37]	ResNet-50	14.4%	52.0%	15.9%	9.2%
Repulsion Loss [43]	ResNet-50	13.2%	56.9%	16.8%	7.6%
OR-CNN [52]	VGG-16	12.8%	55.7%	15.3%	6.7%
HBAN [25]	VGG-16	12.5%	48.1%	-	-
ALF [23]	ResNet-50	12.0%	51.9%	11.4%	8.4%
Adaptive NMS [21]	ResNet-50	11.9%	54.0%	11.4%	6.2%
CSP [24]	ResNet-50	11.0%	49.3%	10.4%	7.3%
MGAN [28]	VGG-16	10.5%	47.2%	-	-
PSC-Net [45]	VGG-16	10.4%	39.7%	-	-
APD [47]	ResNet-50	10.6%	49.8%	9.5%	7.1%
APD [47]	DLA-34	8.8%	46.6%	8.3%	5.8%
ACSP(Smooth L1)	ResNet-101	10.0%	46.1%	8.8%	6.7%
ACSP(Vanilla L1)	ResNet-101	9.3%	46.3%	8.7%	5.6%