

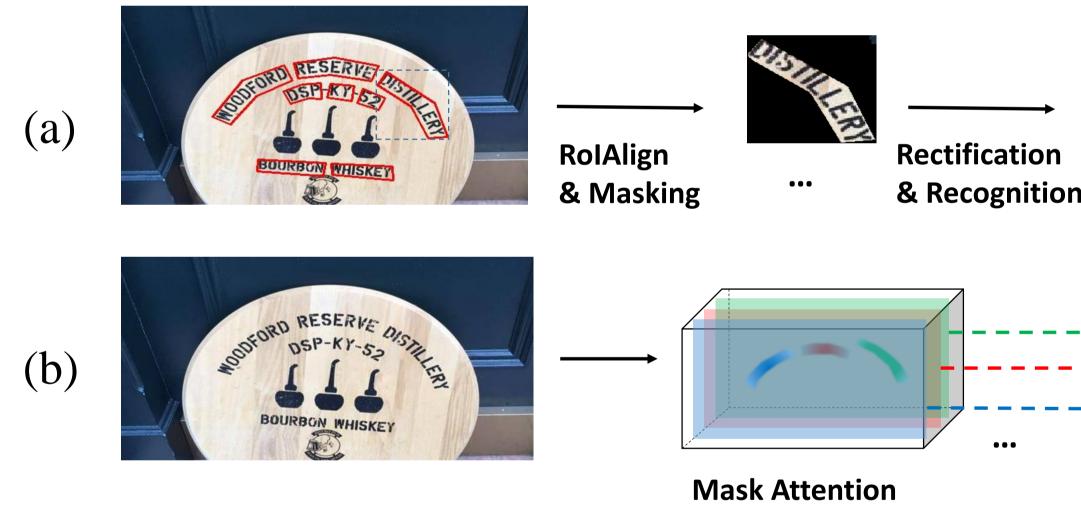






MOTIVATION & CONTRIBUTIONS

- Existing two-staged end-to-end text spotting methods adopt complicated Rol operations with time-consuming NMS.
- Recognition highly relies on detection results and its annotations. Accurate boundaries need to be depicted during detection.

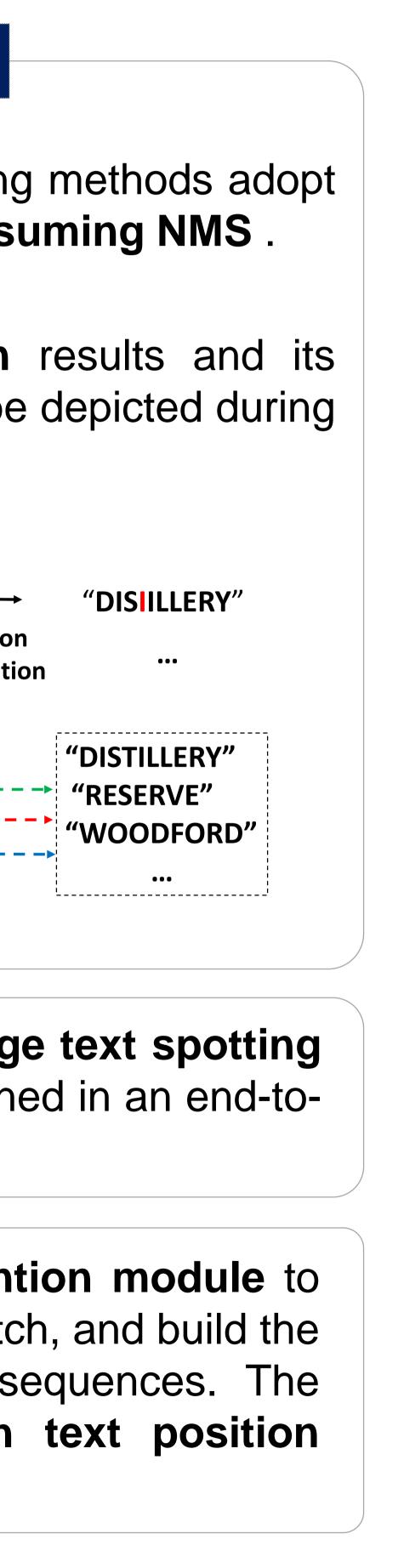


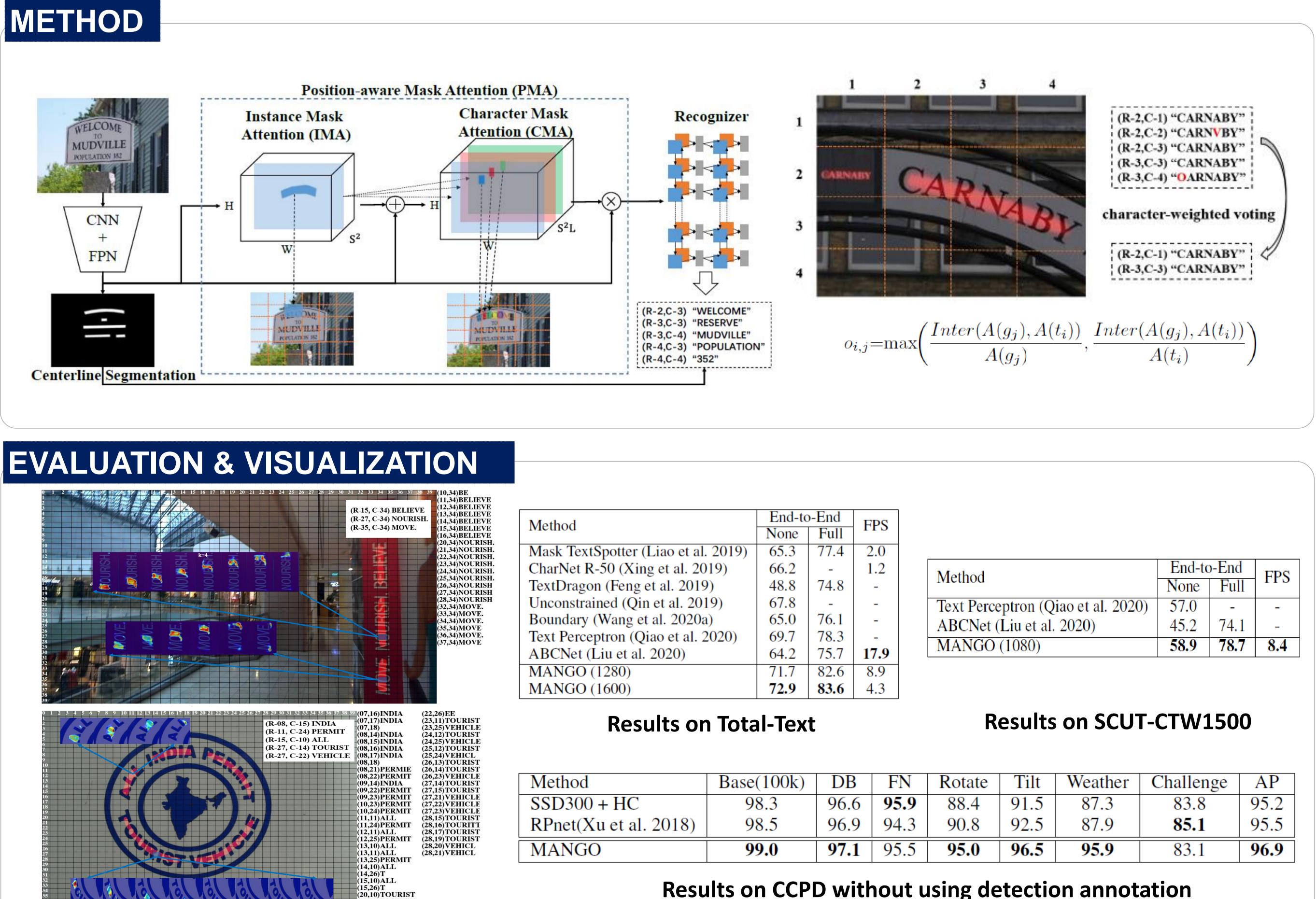
- We propose a compact and robust **one-stage text spotting framework** named MANGO that can be trained in an end-toend manner.
- We develop the position-aware mask attention module to generate the text instance features into a batch, and build the one-to-one mapping with final character sequences. The module can be trained with only rough text position information and text annotations.



MANGO: A Mask Attention Guided One-Stage Scene Text Spotting

Zhanzhan Cheng³¹ Shiliang Pu¹ Ying Cheng² Yi Niu ¹ Fei Wu³ Liang Qiao¹ Yunlu Xu¹ ¹ Hikvision Research Institute ³Zhejiang University ²TongJi University





(F	32 33 34 35 36 R-15, C-34) BELI R-27, C-34) NOU R-35, C-34) MOV	IEVE RISH. 7E. (12,34 (13,34 (14,34 (15,34 (16,34 (20,34 (21,34 (22,34 (23,34 (24,34 (25,34 (26,34 (33,34 (34,34 (35,34 (36,34)	4)BE 4)BELIEVE 4)BELIEVE 4)BELIEVE 4)BELIEVE 4)BELIEVE 4)BELIEVE 4)NOURISH. 4)NOURISH. 4)NOURISH. 4)NOURISH. 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOURISH 4)NOVE. 4)MOVE. 4)
S 36 37 38 39 NDIA ERMIT LL OURIST EHICLE	(07,16)INDIA (07,17)INDIA (07,17)INDIA (08,14)INDIA (08,15)INDIA (08,16)INDIA (08,17)INDIA (08,17)INDIA (08,21)PERMIT (08,22)PERMIT (09,22)PERMIT (09,23)PERMIT (10,23)PERMIT (10,23)PERMIT (10,24)PERMIT (11,11)ALL (11,24)PERMIT (12,11)ALL (12,25)PERMIT (13,10)ALL (13,25)PERMIT (14,10)ALL (13,25)PERMIT (14,10)ALL (15,26)T (20,10)TOURIS (21,26)EE (22,11)TOURIS	Γ (26,23)VE (27,14)TO Γ (27,15)TO Γ (27,21)VE Γ (27,22)VE Γ (27,23)VE (28,15)TO Γ (28,16)TO (28,17)TO Γ (28,19)TO (28,20)VE (28,21)VE Γ	URIST HICLE URIST HICLE URIST URIST URIST URIST URIST HICLE HICLE HICLE URIST URIST URIST URIST URIST URIST URIST

Method	End-to	FPS	
Method	None	Full	115
Mask TextSpotter (Liao et al. 2019)	65.3	77.4	2.0
CharNet R-50 (Xing et al. 2019)	66.2	-	1.2
TextDragon (Feng et al. 2019)	48.8	74.8	-
Unconstrained (Qin et al. 2019)	67.8	-	-
Boundary (Wang et al. 2020a)	65.0	76.1	-
Text Perceptron (Qiao et al. 2020)	69.7	78.3	-
ABCNet (Liu et al. 2020)	64.2	75.7	17.9
MANGO (1280)	71.7	82.6	8.9
MANGO (1600)	72.9	83.6	4.3
•			

Method	Base(100k)	DB	FN	Rotate	Tilt	Weather	Challenge	AP
SSD300 + HC	98.3	96.6	95.9	88.4	91.5	87.3	83.8	95.2
RPnet(Xu et al. 2018)	98.5	96.9	94.3	90.8	92.5	87.9	85.1	95.5
MANGO	99.0	97.1	95.5	95.0	96.5	95.9	83.1	96.9

Results on CCPD without using detection annotation



Method	End-to	FPS		
Ivietilou	None	Full	ггэ	
Text Perceptron (Qiao et al. 2020)	57.0	-	-	
ABCNet (Liu et al. 2020)	45.2	74.1	-	
MANGO (1080)	58.9	78.7	8.4	