



Text Perceptron: Towards End-to-End Arbitrary-Shaped Text Spotting

AAAI-2020

February 7 - 12 2020, New York, New York USA

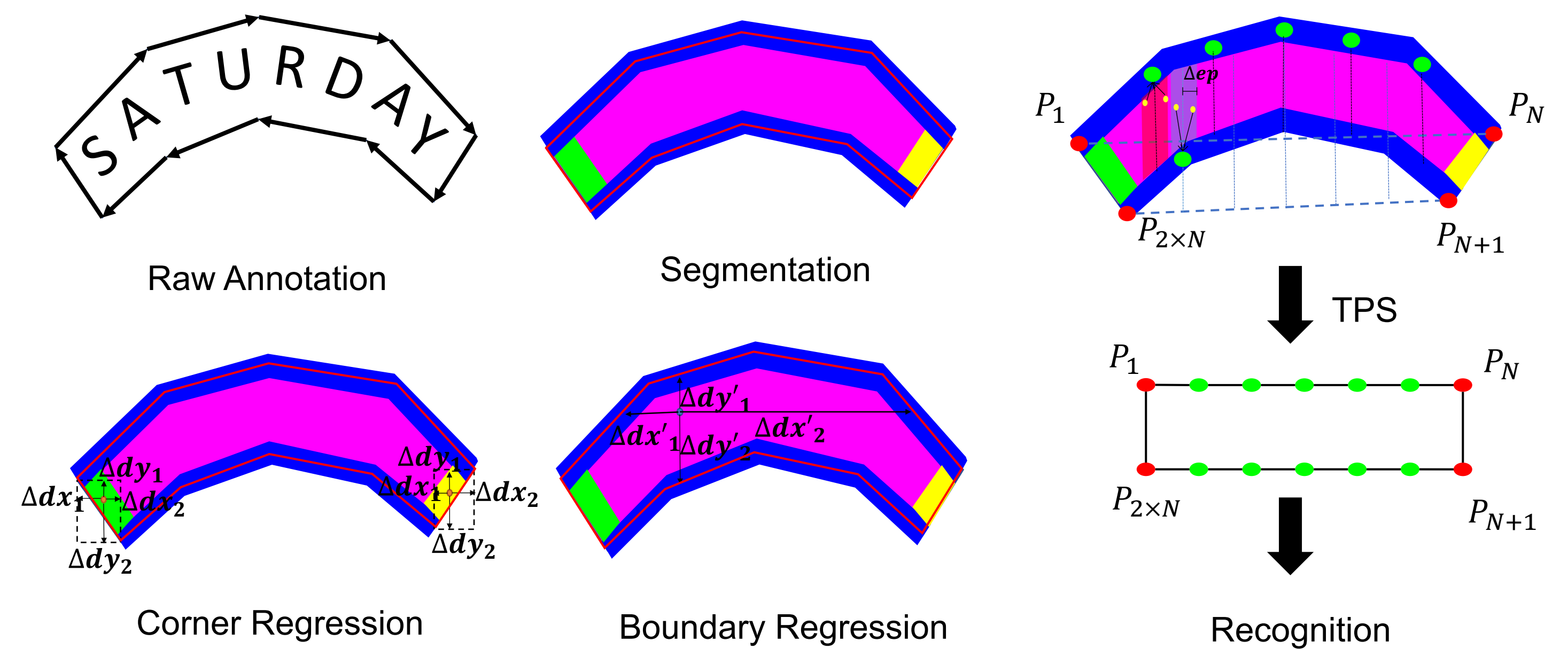
Liang Qiao¹ Sanli Tang¹ Zhazhan Cheng^{2,1};
Yunlu Xu¹ Yi Niu¹ Shiliang Pu¹ Fei Wu²
¹ Hikvision Research Institute ² Zhejiang University



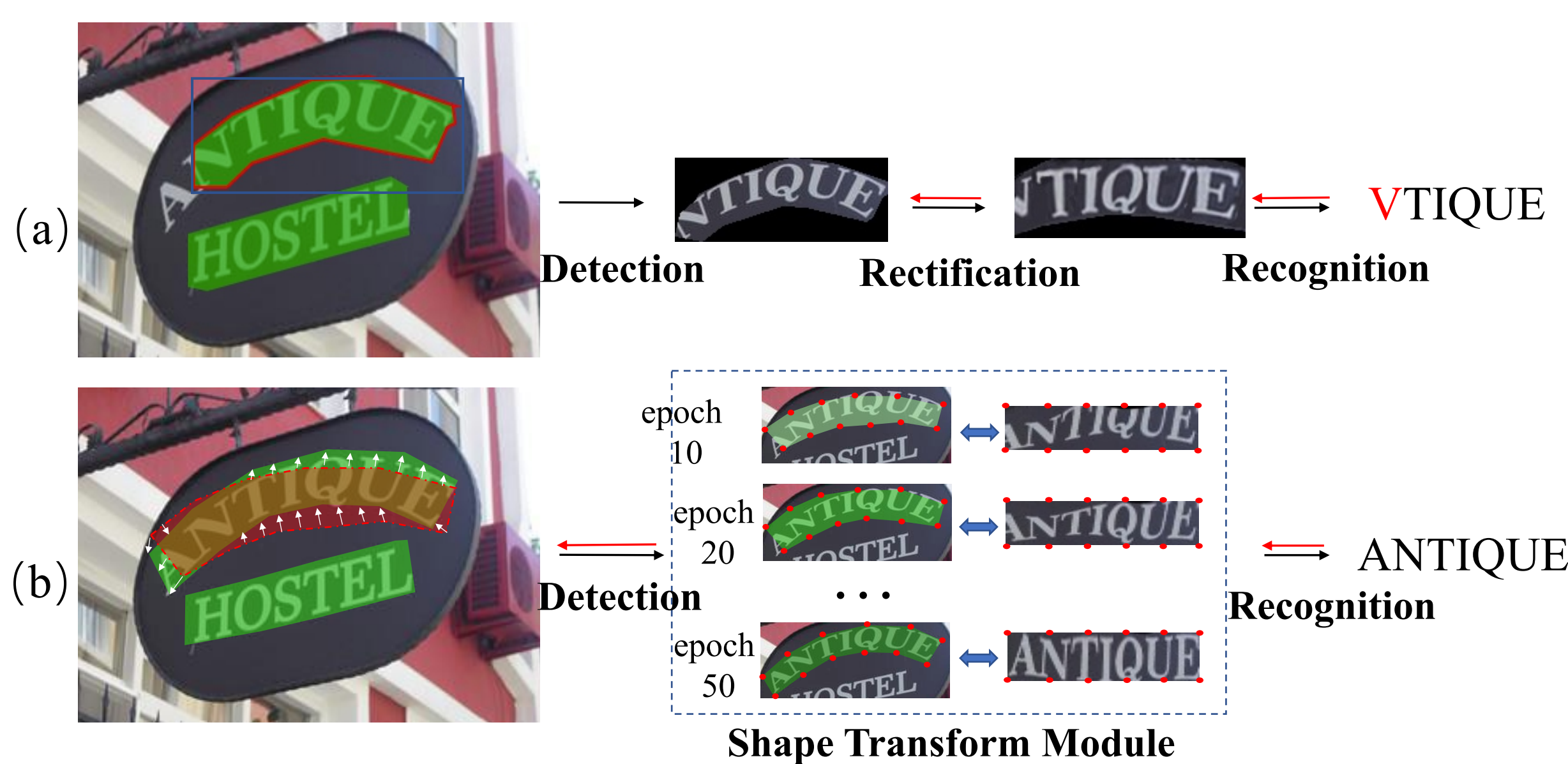
Paper ID: 893 Contact Mail: qiaoliang6@hikvision.com

MOTIVATION

- Pipelined text spotting methods result in **suboptimal performance**. Is the GT of detection for recognition?
- Contour-based detection results requires **complicated rectification** network in handling curved text. How detection provides more information?



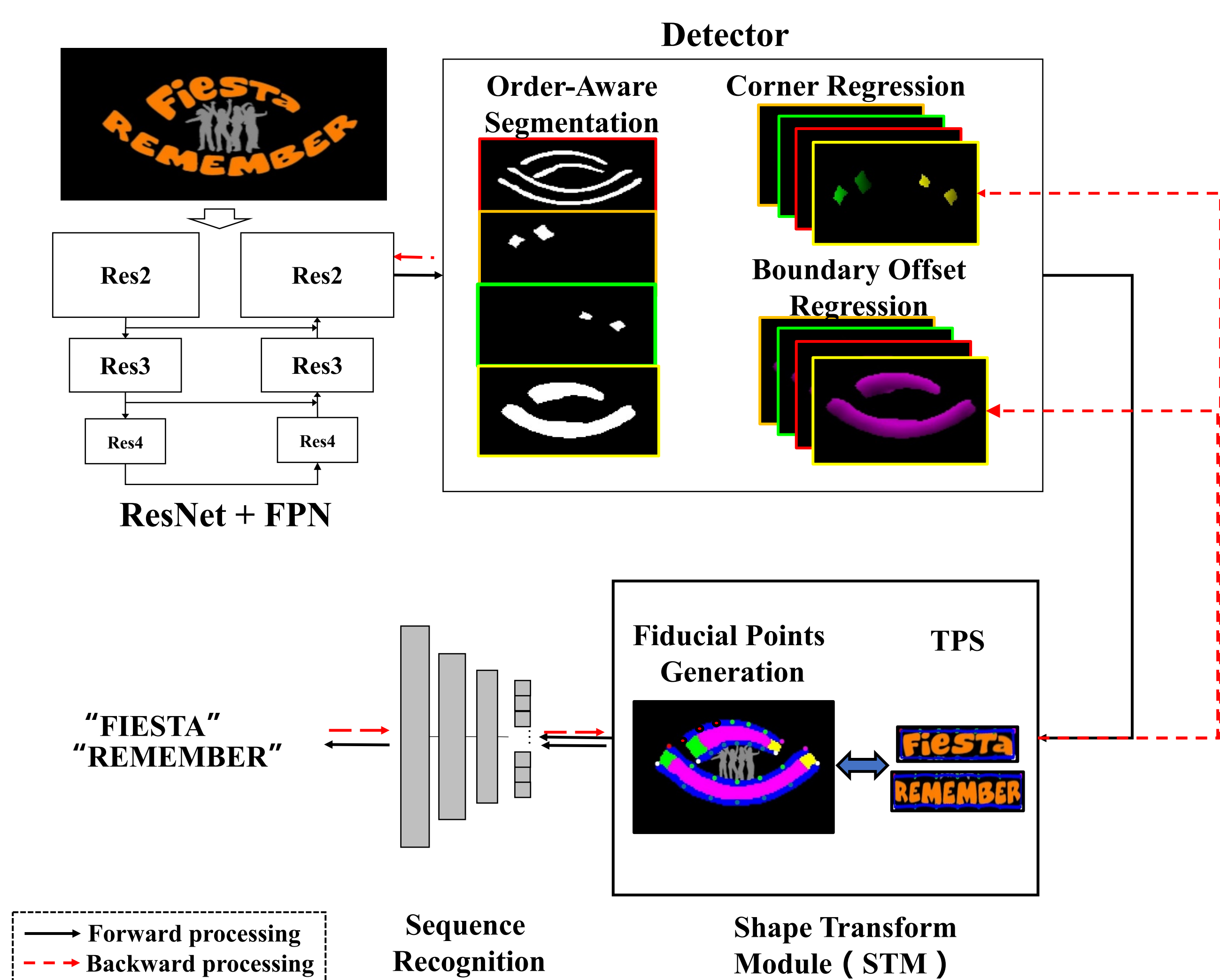
$$\mathcal{L} = \mathcal{L}_{cls} + \lambda_b \mathcal{L}_{corner} + \lambda_c \mathcal{L}_{boundary} + \lambda_r \mathcal{L}_{recog}$$



EXPERIMENT



METHOD



Results on Total-Text Dataset

Method	Detection			End-to-End	
	P	R	F	None	Full
TextSnake (2018)	82.7	74.5	78.4	-	-
FTSN (2018)	84.7	78.0	81.3	-	-
TextField (2019)	81.2	79.9	80.6	-	-
SPCNet (2019)	83.0	82.8	82.9	-	-
CSE (2019b)	81.4	79.1	80.2	-	-
PSENet-1s (2019)	84.0	78.0	80.9	-	-
LOMO (2019)	75.7	88.6	81.6	-	-
Mask TextSpotter (2018)	69.0	55.0	61.3	52.9	71.8
TextNet (2018)	68.2	59.5	63.5	54.0	-
Ours (2-stage)	88.1	78.9	83.3	63.3	73.9
Ours (End-to-end)	88.8	81.8	85.2	69.7	78.3

- We design an efficient **order-aware text detector** to extract arbitrary-shaped text. It learns three tasks simultaneously: order-aware segmentation, corner regression, and boundary offset regression.
- We develop the **differentiable Shape Transform Module (STM)** devoting to optimizing both detection and recognition in an end-to-end trainable manner.
- Extensive experiments show that our method achieves competitive results on two regular text benchmarks, and also **significantly surpasses previous methods** on two irregular text benchmarks.

Results on SCUT-CTW1500 Dataset

Method	Detection			End-to-End
	P	R	F	None
TextSnake (2018)	69.7	85.3	75.6	-
TextField (2019)	83.0	79.8	81.4	-
CSE (2019b)	81.1	76.0	78.4	-
PSENet-1s (2019)	84.8	79.7	82.2	-
LOMO (2019)	69.6	89.2	78.4	-
Ours (2-stage)	88.7	78.2	83.1	48.6
Ours (End-to-end)	87.5	81.9	84.6	57.0