AON: Towards Arbitrarily-Oriented Text Recognition

Zhanzhan Cheng; Yi Niu; Shiliang Pu — *Hikvision Research Institute*

Yangliu Xu — *Tongji University*

Fan Bai; Shuigeng Zhou — *Fudan University*

Paper Download

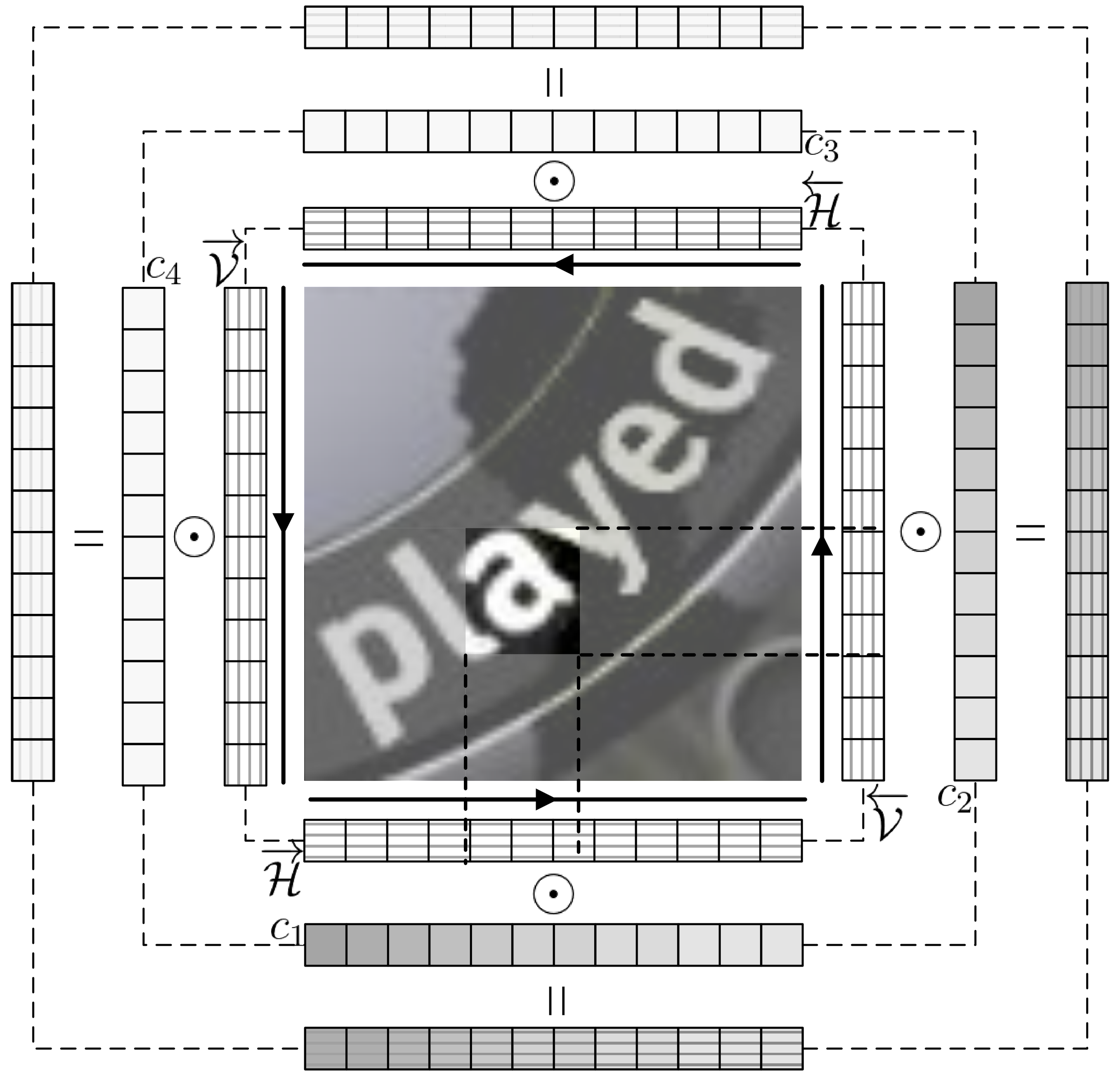
Zhanzhan Cheng

Fan Bai

2018 CVPR SALT LAKE CITY • JUNE 18-22

- Existing methods on text recognition mainly work with regular (horizontal and frontal) texts.
- In real-world applications, many scene texts ae in **irregular arrangements** (e.g. arbitrarily-oriented, curved, slant and perspective etc.)

- We propose the arbitrary orientation network (**AON**) to extract scene text features in four directions and the character placement clues.
- AON obtains state-of-the-art performance in irregular benchmarks, and is comparable to major existing methods in regular benchmarks.
- The learned character placement clues can be used to generate placement trends of character sequences by positioning each character and drawing text orientations in the original images.



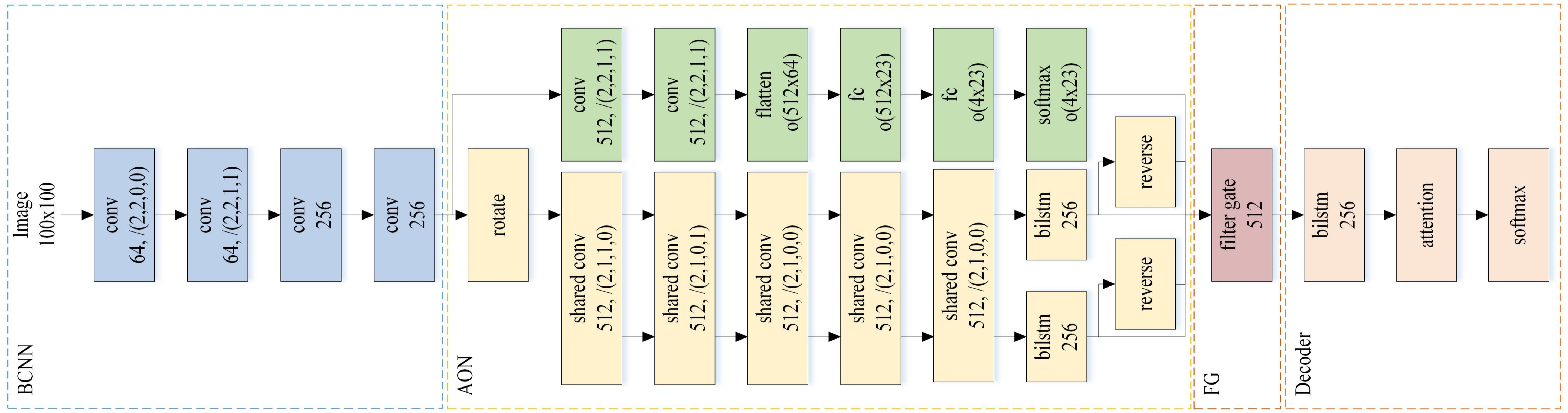
$$① \ \mathcal{H} = \begin{cases} \overrightarrow{\mathcal{H}} : (h_1, ..., h_L)^T, & left \rightarrow right \\ \overleftarrow{\mathcal{H}} : (h_L, ..., h_1)^T, & right \rightarrow left \end{cases}$$

$$② \ \mathcal{V} = \begin{cases} \overrightarrow{\mathcal{V}} : (v_1, ..., v_L)^T, & top \rightarrow bottom \\ \overleftarrow{\mathcal{V}} : (v_L, ..., v_1)^T. & bottom \rightarrow top \end{cases}$$

$$③ \ \mathcal{C} = (c_1, ..., c_L)^T.$$

$$④ \ \hat{h}'_i = [\overrightarrow{\mathcal{H}}_i \ \overleftarrow{\mathcal{H}}_i \ \overrightarrow{\mathcal{V}}_i \ \overleftarrow{\mathcal{V}}_i]c_i.$$

$$⑤ \ \hat{h}_i = \tanh(\hat{h}'_i).$$



Performance on Irregular Datasets

Method	SVT-Perspective			CT80	IC15
	50	Full	None	None	None
ABBYY[35]	40.5	26.1	–	–	–
Mishra *et al.*[11]	45.7	24.7	–	–	–
Wang *et al.*[37]	40.2	32.4	–	–	–
Phan *et al.*[28]	75.6	67.0	–	–	–
Shi *et al.*[31]	92.6	72.6	66.8	54.9	–
Shi *et al.*[32]	91.2	77.4	71.8	59.2	–
Yang *et al.*[39]	93.0	80.2	**75.8**	69.3	–
Cheng *et al.*[6]	92.6	81.6	71.5	63.9	66.2
Naive_base	92.4	83.3	70.5	75.4	67.8
STN_base	**94.6**	82.8	68.5	73.7	67.5
Ours	94.0	**83.7**	73.0	**76.8**	**68.2**

Performance on Regular Datasets

Method	IIIT5k			SVT		IC03		
	50	1k	None	50	None	50	Full	None
ABBYY[35]	24.3	–	–	35.0	–	56.0	55.0	–
Jaderberg *et al.*[17]	97.1	92.7	–	95.4	80.7	98.7	**98.6**	93.1
Jaderberg *et al.*[16]	95.5	89.6	–	93.2	71.7	97.8	97.0	89.6
Shi *et al.*[31]	97.6	94.4	78.2	96.4	80.8	98.7	97.6	89.4
Shi *et al.*[32]	96.2	93.8	81.9	95.5	81.9	98.3	96.2	90.1
Lee *et al.*[22]	96.8	94.4	78.4	96.3	80.7	97.9	97.0	88.7
Yang *et al.*[39]	97.8	96.1	–	95.2	–	–	97.7	–
Cheng's baseline[6]	98.9	96.8	83.7	95.7	82.2	98.5	96.7	91.5
Cheng *et al.*[6]	99.3	97.5	**87.4**	**97.1**	**85.9**	**99.2**	97.3	**94.2**
Naive_base	99.5	**98.1**	86.0	96.9	81.9	98.5	96.5	90.5
STN_base	99.5	97.8	85.9	96.3	80.7	98.5	96.2	89.2
Ours	**99.6**	**98.1**	87.0	96.0	82.8	98.5	97.1	91.5