

Supplementary Material for “Scene Text Retrieval via Joint Text Detection and Similarity Learning”

This document contains the supplementary material for “Scene Text Retrieval via Joint Text Detection and Similarity Learning”. The algorithm details of word augmentation strategy are given in Sec. 1. The qualitative results of Chinese scene text retrieval are given in Sec. 2.

only obtains the correct ranking lists but also localizes precise the bounding boxes of text instances.

1. The details of word augmentation strategy

Algorithm 1 Word Augmentation Strategy

Require:

Character set: $B = \{b_0, b_1, \dots, b_{L-1}\}$.
// MD stands for multinomial distribution.
Multinomial distribution: $MD_4(4 : p_1, \dots, p_4)$.
// N is the character number of input word M .
Input: Word $M = [m_0, m_1, \dots, m_{N-1}]$, $m_i \in B$.

Ensure:

Output: Word $W = []$.

- 1: Generate random sequence $A = (a_0, \dots, a_{N-1})$,
 $A \sim MD_4(4 : p_1, \dots, p_4)$,
 $a_i \in \{insert, delete, replace, keep\}$.
 - 2: **for** $i=0$; $i < N$; $i++$ **do**
 - 3: **if** $a_i = insert$ **then**
 - 4: $W.append(m_i)$
 - 5: $W.append(random.choice(B))$
 - 6: **end if**
 - 7: **if** $a_i = delete$ **then**
 - 8: continue
 - 9: **end if**
 - 10: **if** $a_i = replace$ **then**
 - 11: $W.append(random.choice(B))$
 - 12: **end if**
 - 13: **if** $a_i = keep$ **then**
 - 14: $W.append(m_i)$
 - 15: **end if**
 - 16: **end for**
-

2. Examples of Chinese scene text retrieval results

Some qualitative results of Chinese scene text retrieval are shown in Fig. 1. We can observe that our method not



(a)



(b)

Figure 1. Retrieval results on CSVTR. Only top-8 retrieved images are shown for query words “Home Inn” in (a) and “Auntie Jenny” in (b).