Advanced Computer Vision

Практический курс

Савельева Юлия Олеговна, 3 семестр, 22.09.2022

Text Recognition
Целевая функция

Реализовать CTC Loss:

https://www.cs.toronto.edu/ ~graves/icml_2006.pdf

Алгоритм вычисления alpha ->

Реализовать также вычисление beta по аналогии, только в обратном направлении

Algorithm 1: $ComputeAlpha$ (for single element in the batch)

Data:

$out_{|\hat{A}| \times T}$ (result of softmax), where $|\hat{A}|$ (alphabet size), $$T = \bar{W}_{unpadded}/4;$$

$label$ (encoded by alphabet);

$bl = 0$ (blank index)

begin

$\quad S = len(label) * 2 + 1$

$\quad T = out.shape[1]$

$\quad a = zeros(S, T)$

$\quad a[0][0] = out[bl][0]$

$\quad a[1][0] = out[label[0]][0]$

$\quad c = a[0][0] + a[1][0]$

$\quad$ **if** $c > 0$ **then**

$\quad\quad a[0][0] = \frac{a[0][0]}{c}$

$\quad\quad a[1][0] = \frac{a[1][0]}{c}$

$\quad$ **for** $t := 1$ **to** T **do**

$\quad\quad start = \max(0, S - 2 \times (T - t))$

$\quad\quad end = \min(2 \times t + 2, S)$

$\quad\quad$ **for** $s := start$ **to** end **do**

$\quad\quad\quad i = max(0, floor(\frac{s-1}{2}))$

$\quad\quad\quad$ **if** $s \bmod 2 = 0$ **then**

$\quad\quad\quad\quad$ **if** $s = 0$ **then**

$\quad\quad\quad\quad\quad a[s][t] = a[s][t-1] \times out[bl][t]$

$\quad\quad\quad\quad$ **else**

$\quad\quad\quad\quad\quad a[s][t] = (a[s][t-1] + a[s-1][t-1]) \times out[bl][t]$

$\quad\quad\quad$ **else if** $s = 1$ **or** $label[i] = label[i-1]$ **then**

$\quad\quad\quad\quad a[s][t] = (a[s][t-1] + a[s-1][t-1]) \times out[label[i]][t]$

$\quad\quad\quad$ **else**

$\quad\quad\quad\quad a[s][t] =$

$\quad\quad\quad\quad (a[s][t-1] + a[s-1][t-1] + a[s-2][t-1]) \times out[label[i]][t]$

$\quad\quad c = \sum_{i=start}^{end} a[i][t]$

$\quad\quad$ **if** $c > 0$ **then**

$\quad\quad\quad$ **for** $s := start$ **to** end **do**

$\quad\quad\quad\quad a[s][t] = \frac{a[s][t]}{c}$

return a

Дедлайн 06.10.2022 00:00