512*
512*
(150 ∼
284)
(0.97 ∼
1.36)*
(0.97 ∼
1.36)*
(2 ∼
2.5)
0.97*
0.97*
2
(−1000 ∼
13000)
(−800 ∼
400)
512*
512
384*
288
$_{w}ithlabel.png8cmSegTHOR:$
$(-5, 5)$
1.1
0.9
1.1
(31 ∼
43)*
(40 ∼
59)*
(24 ∼
47)
48*
64*
48
0.1
p_i
$_{i} =$
$\frac{exp(z_i)}{\Sigma_j exp(z_j)}$
$_{i} =$
$\frac{exp(z_i/T)}{\Sigma_j exp(z_j/T)}$
$_{D}(p(z_t, T), p(z_s, T)) =$
$-\Sigma_i p_i(z_{ti}, T) \log(p_i(z_{si}, T))$
z_t
z_s
$\partial L_D(p(z_t, T), p(z_s, T))$
$\frac{}{\partial z_{si} = \frac{p(z_t, T) - p(z_s, T)}{T} = \frac{1}{T}(\frac{exp(z_{si}/T)}{\Sigma_j exp(z_{sj}/T)} - \frac{exp(z_{ti}/T)}{\Sigma_j exp(z_{tj}/T)})}$
$\partial L_D(p(z_t, T), p(z_s, T))$
$\frac{}{\partial z_{si} = \frac{1}{T}(\frac{1 + \frac{z_{si}}{T}}{N + \Sigma_j \frac{z_{sj}}{T}} - \frac{1 + \frac{z_{ti}}{T}}{N + \Sigma_j \frac{z_{tj}}{T}})}$
$\partial L_D(p(z_t, T), p(z_s, T))$
$\frac{}{\partial z_{si} = \frac{1}{NT^2}(z_{si} - z_{ti})}$
$_{S}(y, p(z_s, T)) =$
$-\Sigma_i y \log(p_i(z_{si}, T))$
α
$_{TotalDistillation} =$
$\alpha L_D(p(z_t, T), p(z_s, T)) +$
$(1 -$
$\alpha) L_S(y, p(z_s, T))$
13
13*
13
3*
3
$\frac{}{N \sum_{i=1}^{N} y_i \log(p_i) + (1 - y_i) \log(1 - p_i)}$
y_i
p_i
0.25
$_{D}(p(z_s), p(z_t)) =$
$L_{KL}(p(z_s), p(z_t)) =$
$\sum p(z_s) \log(\frac{p(z_s)}{p(z_t)})$
$_{total} =$
$\alpha BCE(p(z_s), y) +$
$(1 -$
$\alpha) L_{KL}(p(z_s, T), p(z_t, T))$
α
0.2
0.5
0.71
0.93
0.86
0.82
$0.73 \pm 0.130.95 \pm 0.020.93 \pm 0.030.86 \pm 0.08$
$0.45 \pm 0.110.62 \pm 0.030.81 \pm 0.060.63 \pm 0.17$
$0.62 \pm 0.130.91 \pm 0.040.89 \pm 0.070.79 \pm 0.18$
$0.33 \pm 0.070.13 \pm 0.040.80 \pm 0.050.22 \pm 0.05$
$0.54 \pm 0.130.79 \pm 0.090.88 \pm 0.050.64 \pm 0.17$
$0.64 \pm 0.120.93 \pm 0.030.90 \pm 0.030.82 \pm 0.09$
$0.38 \pm 0.090.60 \pm 0.030.77 \pm 0.070.56 \pm 0.16$
$0.53 \pm 0.110.89 \pm 0.040.86 \pm 0.070.74 \pm 0.19$
$0.29 \pm 0.050.09 \pm 0.030.76 \pm 0.050.22 \pm 0.04$
$0.45 \pm 0.110.77 \pm 0.090.85 \pm 0.060.56 \pm 0.15$
$4.54 \pm 0.61\ 8.11 \pm 1.48\ 4.16 \pm 0.44\ 6.12 \pm 0.61$
$4.74 \pm 0.5310.34 \pm 0.564.52 \pm 0.547.64 \pm 1.16$
$4.60 \pm 0.54\ 8.52 \pm 1.49\ 4.45 \pm 0.28\ 6.02 \pm 0.77$