Breadcrumbs: Adversarial Class-Balanced Sampling for Long-tailed Recognition

Bo Liu[1], Haoxiang Li[1], Hao Kang[1], Gang Hua[1], and Nuno Vasconcelos[2]

[1] Wormpex AI Research
{richardboliu,lhxustcer,haokheseri,ganghua}@gmail.com
[2] University of California San Diego
nuno@ece.ucsd.edu

1 Proof of Lemma 1

Lemma 1 *Consider the augmentation of $\mathcal{Z}_y^e$ with the snapshot transferred from epoch $e' < e$ by EMANATE, i.e. $\mathcal{A}_y^e = \mathcal{Z}_y^e \cup \mathcal{Z}_y^{e' \to e}$, where $\mathcal{Z}_y^{e' \to e}$ is as defined in (6, paper). Then*

$$L_y(\mathcal{A}_y^e, \mathbf{W}^e, \mathbf{b}^e) - L_y(\mathcal{Z}_y^e, \mathbf{W}^e, \mathbf{b}^e) \geq \frac{L_y(\mathcal{Z}_y^{e'}, \mathbf{W}^{e'}, \mathbf{b}^{e'}) - L_y(\mathcal{Z}_y^e, \mathbf{W}^e, \mathbf{b}^e)}{2}, \quad (1)$$

where $(\mathbf{W}^e, \mathbf{b}^e)$ is the classifier of (10, paper).

Proof. From (9, paper),

$$L_y(\mathcal{A}_y^e, \mathbf{W}^e, \mathbf{b}^e) = = \frac{1}{2} L_y(\mathcal{Z}_y^e, \mathbf{W}^e, \mathbf{b}^e) + \frac{1}{2} L_y(\mathcal{Z}_y^{e' \to e}, \mathbf{W}^e, \mathbf{b}^e)$$

and since

$$\begin{aligned} L_y(\mathcal{Z}_y^{e' \to e}, \mathbf{W}^e, \mathbf{b}^e) &= L_y(\{\mathbf{z}_i^{e'} - \bar{\mathbf{z}}^{e'} + \bar{\mathbf{z}}^e\}, \mathbf{W}^e, \mathbf{b}^e) \\ &= -\frac{1}{|\mathcal{Z}_y^{e'}|} \Sigma_i \log \nu_y(\mathbf{W}^e \mathbf{z}_i^{e'} - \mathbf{W}^e \bar{\mathbf{z}}^{e'} + \mathbf{W}^e \bar{\mathbf{z}}^{e'} + \mathbf{b}^e) \\ &= L(\mathcal{Z}_y^{e'}, \mathbf{W}^e, \mathbf{b}^e - \mathbf{W}^e \bar{\mathbf{z}}^{e'} + \mathbf{W}^e \bar{\mathbf{z}}^e) \end{aligned}$$

it follows from (10, paper) that

$$L_y(\mathcal{Z}_y^{e' \to e}, \mathbf{W}^e, \mathbf{b}^e) \geq L(\mathcal{Z}_y^{e'}, \mathbf{W}^{e'}, \mathbf{b}^{e'})$$

and

$$L_y(\mathcal{A}_y^e, \mathbf{W}^e, \mathbf{b}^e) \geq \frac{L_y(\mathcal{Z}_y^e, \mathbf{W}^e, \mathbf{b}^e) + L(\mathcal{Z}_y^{e'}, \mathbf{W}^{e'}, \mathbf{b}^{e'})}{2}$$

from which the lemma follows.