Class-Balanced Loss Based on Effective Number of Samples

Yin Cui[1,2], Menglin Jia[1], Tsung-Yi Lin[3], Yang Song[4], Serge Belongie[1,2]

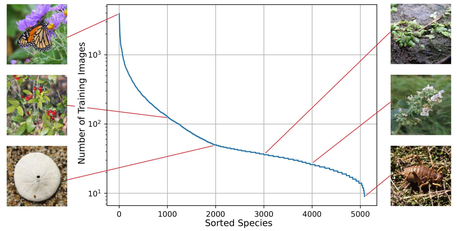
[1] Cornell University [2] Cornell Tech [3] Google Brain [4] Alphabet Inc.

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

Motivation

The World is Long-Tailed

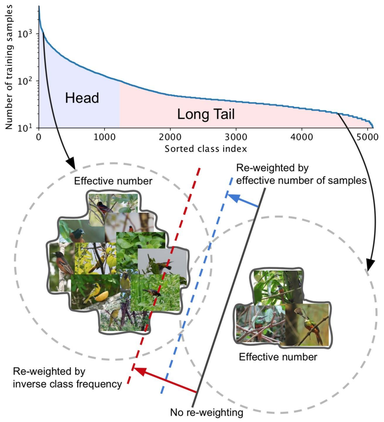
- Distribution of images per category in iNaturalist 2017:



Data Overlap

- The more data, the better. However, as the number of samples increases, the marginal benefit a model can extract from the data diminishes.

Overview



Effective Number of Samples:

$$E_n = (1-\beta^n)/(1-\beta)$$

$$\beta \in [0, 1)$$

n: number of samples.

Class-Balanced Loss:

$$\mathrm{CB}(\mathbf{p}, y) = \frac{1}{E_{n_y}}\mathcal{L}(\mathbf{p}, y) = \frac{1-\beta}{1-\beta^{n_y}}\mathcal{L}(\mathbf{p}, y)$$

Effective Number of Samples

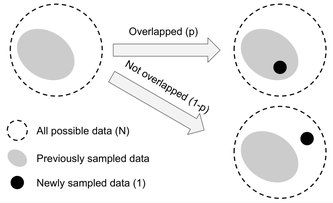
Data Sampling as Random Covering

- To measure data overlap, we associate each sample with a small neighboring region of unit volume 1. Assume the volume of all possible data is N.

Definition (Effective Number). The *effective number* of samples is the expected volume of samples.

Proposition (Effective Number). $E_n = (1-\beta^n)/(1-\beta)$, where $\beta = (N-1)/N$.

Implication (Asymptotic Properties). $E_n = 1$ if $\beta = 0$ $(N = 1)$. $E_n \to n$ as $\beta \to 1$ $(N \to \infty)$.



Class-Balanced Loss

- Class-Balanced Softmax Loss:

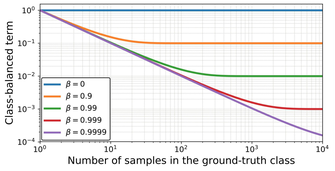
$$\mathrm{CB}_{\mathrm{softmax}}(\mathbf{z}, y) = -\frac{1-\beta}{1-\beta^{n_y}}\log\left(\frac{\exp(z_y)}{\sum_{j=1}^{C}\exp(z_j)}\right)$$

- Class-Balanced Sigmoid Loss:

$$\mathrm{CB}_{\mathrm{sigmoid}}(\mathbf{z}, y) = -\frac{1-\beta}{1-\beta^{n_y}}\sum_{i=1}^{C}\log\left(\frac{1}{1+\exp(-z_i^t)}\right)$$

- Class-Balanced Focal Loss:

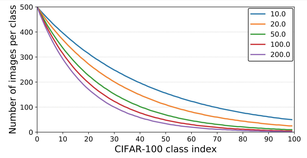
$$\mathrm{CB}_{\mathrm{focal}}(\mathbf{z}, y) = -\frac{1-\beta}{1-\beta^{n_y}}\sum_{i=1}^{C}(1-p_i^t)^{\gamma}\log(p_i^t)$$



$$z_i^t = \begin{cases} z_i, & \text{if } i = y. \\ -z_i, & \text{otherwise.} \end{cases} \qquad p_i^t = \mathrm{sigmoid}(z_i^t)$$

Dataset

Dataset Name	# Classes	Imbalance
Long-Tailed CIFAR-10	10	10.00 - 200.00
Long-Tailed CIFAR-100	100	10.00 - 200.00
iNaturalist 2017	5,089	435.44
iNaturalist 2018	8,142	500.00
ILSVRC 2012	1,000	1.78

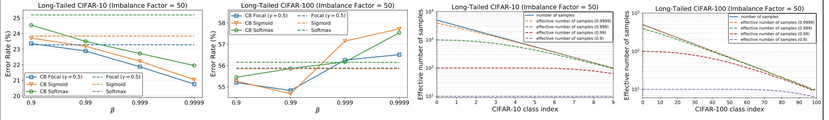


- Imbalance: the number of training samples in the largest class divided by the smallest class.

Experiments

Classification Error Rate of ResNet-32 on CIFAR

Dataset Name	Long-Tailed CIFAR-10						Long-Tailed CIFAR-100					
Imbalance	200	100	50	20	10	1	200	100	50	20	10	1
Softmax	**34.32**	29.64	25.19	17.77	13.61	6.61	65.16	61.68	56.15	48.86	44.29	29.07
Sigmoid	34.51	**29.55**	23.84	**16.40**	**12.97**	**6.36**	64.39	**61.22**	55.85	48.57	44.73	**28.39**
Focal ($\gamma = 0.5$)	36.00	29.77	**23.28**	17.11	13.19	6.75	65.00	61.31	55.88	48.90	44.30	28.55
Focal ($\gamma = 1.0$)	34.71	29.62	23.29	17.24	13.34	6.60	**64.38**	61.59	**55.68**	**48.05**	**44.22**	28.85
Focal ($\gamma = 2.0$)	35.12	30.41	23.48	16.77	13.68	6.61	65.25	61.61	56.30	48.98	45.00	28.52
Class-Balanced	**31.11**	**25.43**	**20.73**	**15.64**	**12.51**	**6.36***	**63.77**	**60.40**	**54.68**	**47.41**	**42.01**	**28.39***
Loss Type	SM	Focal	Focal	SM	SGM	SGM	Focal	Focal	SGM	Focal	Focal	SGM
β	0.9999	0.9999	0.9999	0.9999	0.9999	-	0.9	0.9	0.99	0.99	0.999	-
γ	-	1.0	2.0	-	-	-	1.0	1.0	-	0.5	0.5	-



Classification Error Rate on ImageNet and iNaturalist

					iNaturalist 2017		iNaturalist 2018		ILSVRC 2012	
Network	Loss	β	γ	Input Size	Top-1	Top-5	Top-1	Top-5	Top-1	Top-5
ResNet-50	Softmax	-	-	224 × 224	45.38	22.67	42.86	21.31	23.92	7.03
ResNet-101	Softmax	-	-	224 × 224	42.57	20.42	39.47	18.86	22.65	6.47
ResNet-152	Softmax	-	-	224 × 224	41.42	19.47	38.61	18.07	21.68	5.92
ResNet-50	CB Focal	0.999	0.5	224 × 224	41.92	20.92	38.88	18.97	22.71	6.72
ResNet-101	CB Focal	0.999	0.5	224 × 224	39.06	18.96	36.12	17.18	21.57	5.91
ResNet-152	CB Focal	0.999	0.5	224 × 224	38.06	18.42	35.21	16.34	20.87	5.61
ResNet-50	CB Focal	0.999	0.5	320 × 320	38.16	18.28	35.84	16.85	21.99	6.27
ResNet-101	CB Focal	0.999	0.5	320 × 320	34.96	15.90	32.02	14.27	20.25	5.34
ResNet-152	CB Focal	0.999	0.5	320 × 320	33.73	14.96	30.95	13.54	19.72	4.97

ResNet-50 Training Curves