上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

22-27 MAY 2022
ICASSP 2022
SINGAPORE
Paper ID:4772

SOLVING THE LONG-TAILED PROBLEM VIA INTRA- AND INTER-CATEGORY BALANCE

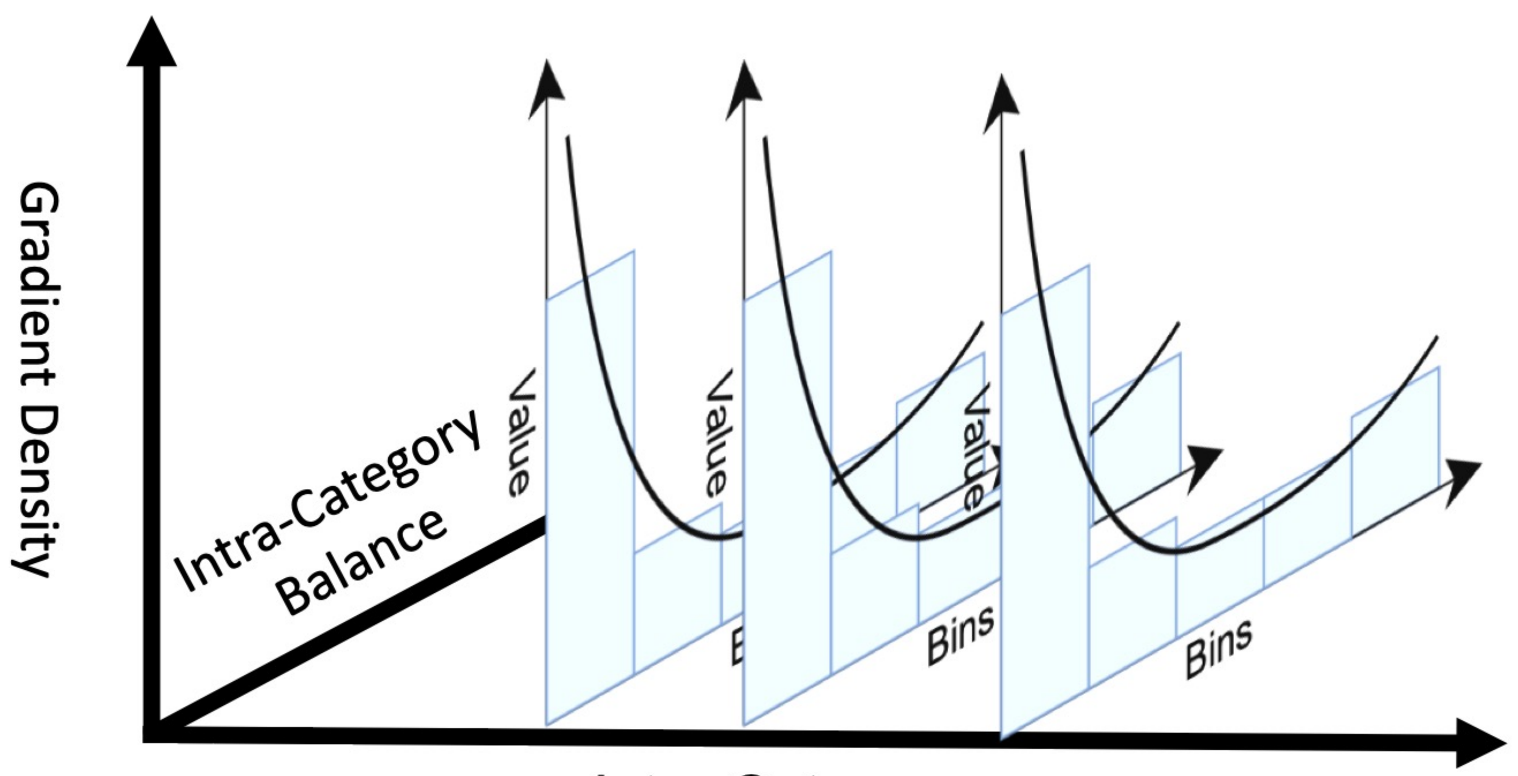
Renhui Zhang, Tiancheng Lin, Rui Zhang, Yi Xu

Shanghai Jiao Tong University

Introduction

- Benchmark datasets for visual recognition assume that data is uniformly distributed, while real-world datasets obey long-tailed distribution.
- Current approaches handle the long-tailed problem only focus on inter-category balance.
- We propose a novel gradient harmonized mechanism with category-wise adaptive precision decouple the difficulty and sample size imbalance, which are correspondingly solved via intra- and inter-category balance strategies.

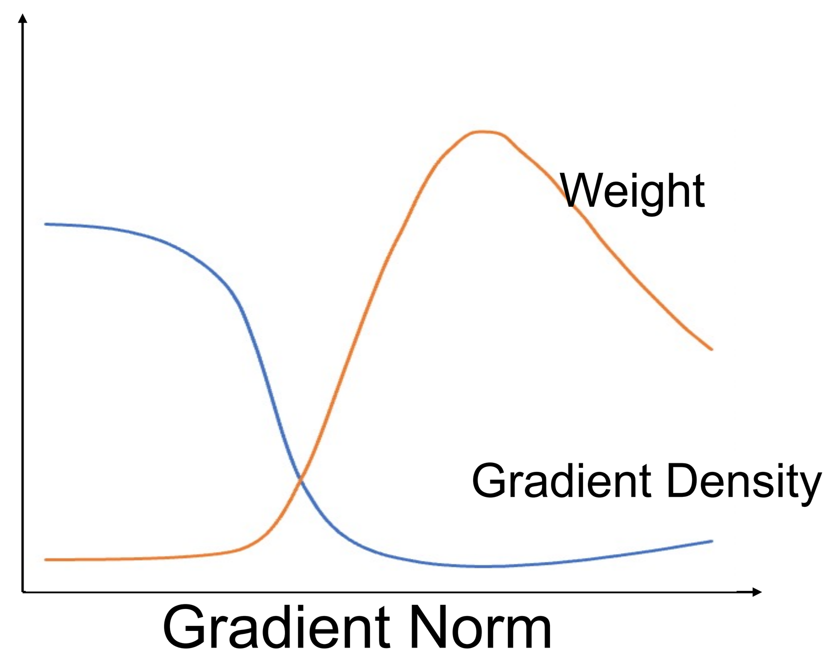
Category-wise GHM



Category-wise GHM dependently count gradient density for each category, thus we can obtain the exact difficulty distribution for each category. Besides, it decouples difficulty and sample size imbalance.

Intra- and Inter-Category Balance

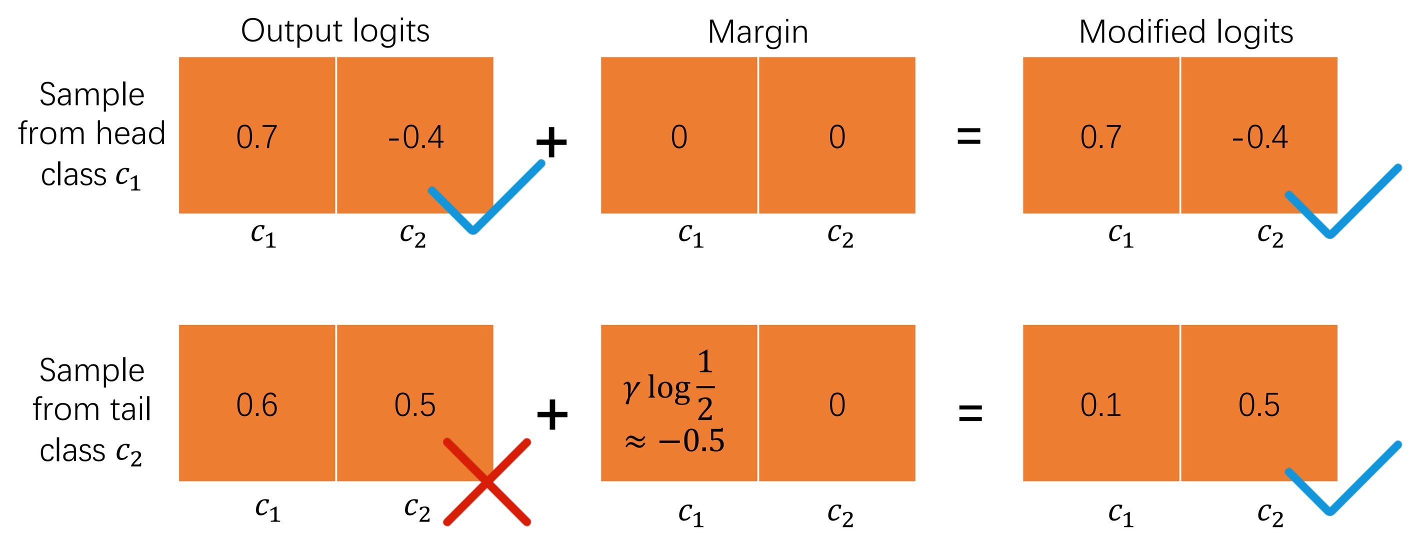
- Intra-category balance emphasize hard samples while reducing the weight of simple samples and outliers in each category according to the gradient density.



$$w = \frac{1}{GD_{c,u}}$$

- Inter-category balance correct the classification boundary shift by adding margins to the logits.

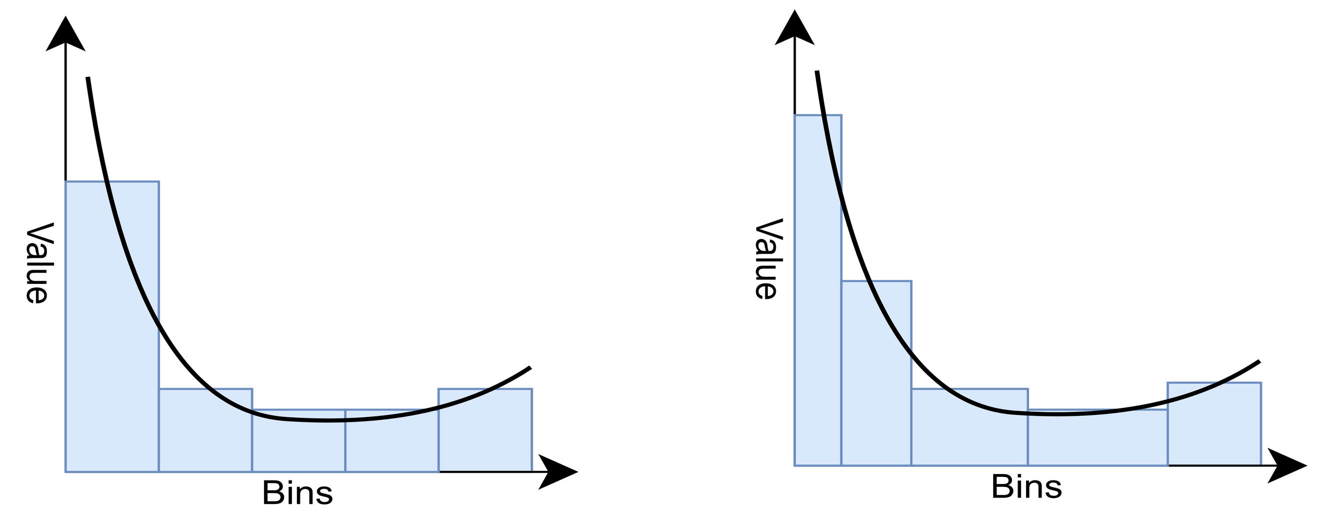
$$M_{c_1,c_2} = \gamma \log[\min(1, S_{c_2}/S_{c_1})]$$



- The final loss function combines intra- and inter-category balance.

$$\mathcal{L} = -\sum_{m \in C} \left[W_{m,i} \cdot y_m \log \left(\frac{e^{z_m}}{\sum_{n \in C} e^{z_n + M_{m,n}}} \right) \right]$$

Adaptive Precision



Adaptively adjust the precision of different bins to achieve higher approximation accuracy, thus improve the performance.

Results

Experiments with different imbalance factors

Table 1. Top-1 accuracy (%) of ResNet-32 with various loss function on long-tailed CIFAR-10/100 and TinyImageNet. Imbalance facotr means the ratio of sample size of head classes to tail classes.

Dataset	Long-Tailed CIFAR-10				Long-Tailed CIFAR-100				Long-Tailed TinyImagenet			
Imbalance factor	500	100	10	1	500	100	10	1	500	100	10	1
Softmax	59.76	71.74	86.69	93.00	30.90	38.77	57.21	71.00	31.42	39.06	54.51	**63.32**
Class Balanced	58.50	73.82	87.18	92.71	31.23	39.12	56.45	70.75	30.90	39.48	54.17	63.06
Focal [23]	57.72	72.43	86.70	92.66	30.31	38.12	56.12	70.13	30.93	39.23	54.00	63.22
GHM [24]	41.50	51.66	67.87	79.35	12.85	14.02	24.49	34.04	8.00	8.03	9.92	10.13
Effective Number [18]	57.97	71.47	87.13	92.98	30.40	39.32	57.37	70.75	30.90	39.48	54.17	63.06
Equalization [19]	58.29	72.61	87.08	**93.04**	30.82	39.84	58.14	**71.66**	29.63	37.39	52.75	62.84
Seesaw [20]	64.48	75.18	**87.76**	93.03	33.63	40.87	**57.83**	71.35	34.07	40.98	54.60	63.05
ours	**66.13**	**75.39**	87.29	92.75	**33.85**	**41.59**	57.81	71.41	**34.79**	**42.66**	**55.35**	63.21

Experiments on large-scale real-world datasets

Table 2. Top-1 accuracy (%) of ResNet-50 on long-tailed ImageNet and iNaturalist2017 dataset. Categories are divided into four groups and sample size gradually decreases from "Many" to "Rare".

	Long-Tailed ImageNet(imbalance factor=100)					iNaturalist2017(imbalance factor=435)				
	Many	Medium	Few	Rare	Over All	Many	Medium	Few	Rare	Over All
Softmax	**63.91**	48.07	22.15	12.55	36.67	84.92	69.88	50.21	40.63	54.64
ClassBalance	63.30	47.71	22.26	12.63	36.48	85.29	70.66	50.88	41.31	55.31
Focal[23]	62.70	47.70	23.42	12.86	36.67	84.41	70.01	51.96	41.13	55.15
EffictiveNumber[18]	63.86	47.91	23.42	13.70	37.22	87.77	72.96	54.45	44.49	58.25
Seesaw[20]	63.51	48.69	24.37	15.52	38.02	87.48	**73.60**	55.64	46.78	59.15
ours	63.61	**48.73**	**25.78**	**16.95**	**38.77**	**87.84**	73.59	**56.75**	**48.02**	**59.85**

Ablations

Table 3. Top-1 accuracy (%) of ResNet-50 on long-tailed ImageNet. "A" and "R" represent intra-category balance and inter-category balance respectively. Besides, "U" and "AU" represent URA and AURA respectively.

	ImageNet-LT(imbalance factor=100)				
	Many	Medium	Few	Rare	Over All
Softmax	63.91	48.07	22.15	12.55	36.67
R	63.51	48.69	24.37	15.52	38.02
A+U	63.93	48.32	22.60	13.44	37.07
R+A+U	63.55	48.67	25.16	16.30	38.42
R+A+AU	63.61	48.73	25.78	16.95	**38.77**