The 38th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 20-27, 2024 | VANCOUVER, CANADA

Feature Fusion from Head to Tail for Long-Tailed Visual Recognition

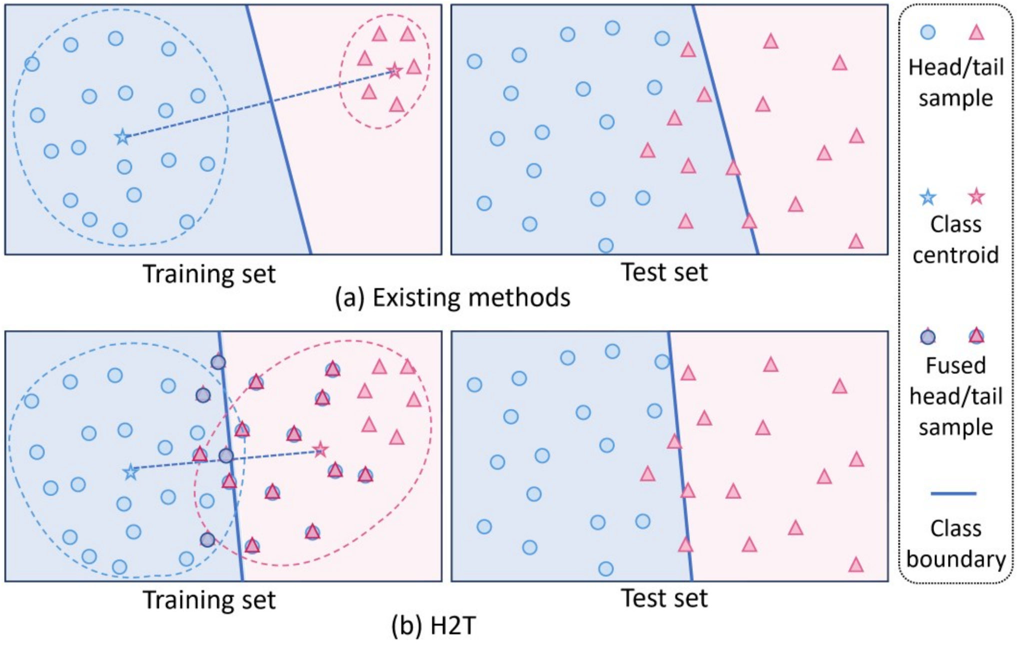
Mengke Li, Zhikai Hu, Yang Lu, Weichao Lan, Yiu-ming Cheung, Hui Huang*

GUANGMING LABORATORY
SHENZHEN UNIVERSITY 1983
HONG KONG BAPTIST UNIVERSITY 1956
UNIVERSITAS AMOIENSIS

Reported by Mengke Li

Introduction

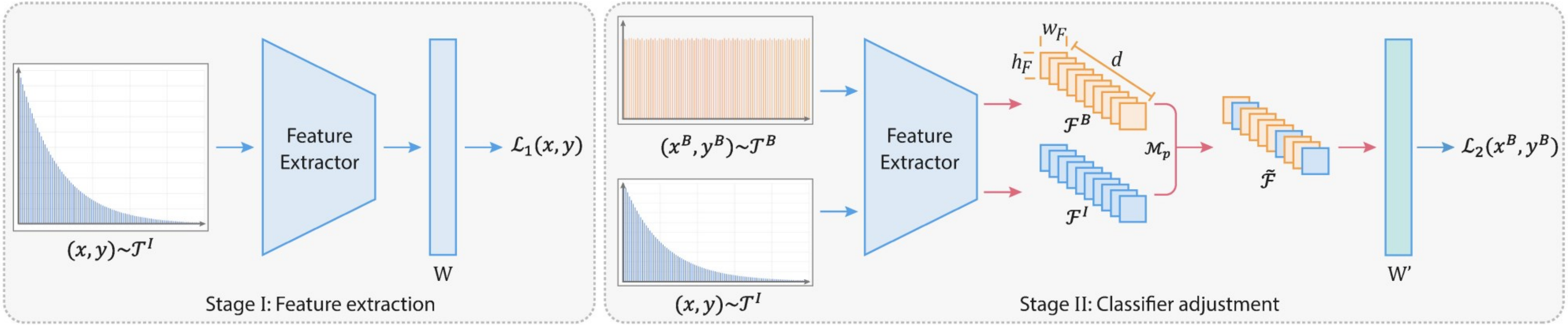
AAAI-24



- The head class has sufficiently sampled and its embedding space is fully occupied.
- The tail class suffers from a scarcity of samples, leading to sparsely distributed semantic regions.

The class boundary is biased.

Method



- Two-stage training scheme (Kang et al. 2020) .
- The second stage involves dual-branch sampling: (1) fused branch -- class-balanced data ; (2) fusing branch -- instance-wise data.
- Fuse the features of the two branches. Only use the labels of the fused branch as ground truth.

Kang, B.; Xie, S.; Rohrbach, M.; Yan, Z.; Gordo, A.; Feng, J.; and Kalantidis, Y. 2020. Decoupling representation and classifier for long-tailed recognition. In *ICLR*.

Rationale

AAAI-24

- $w_t^T = [\dot{w}_t^T, \ddot{w}_t^T], f_t^T = [\dot{f}_t^T, \ddot{f}_t^T]$
- $\ddot{\theta}_*$ (* is h or t) represent the angle between $\ddot{w}_*$ and $\ddot{f}_t - \ddot{f}_h$

Before H2T

$$w_t^T f_t < w_h^T f_t \Rightarrow$$
$$\dot{w}_t^T \dot{f}_t + \ddot{w}_t^T \ddot{f}_t < \dot{w}_h^T \dot{f}_t + \ddot{w}_h^T \ddot{f}_t$$

After H2T

For tail

$$w_t^T \tilde{f}_t > w_h^T \tilde{f} \Rightarrow$$
$$\dot{w}_t^T \dot{f}_t + \ddot{w}_t^T \ddot{f}_t > \dot{w}_h^T \dot{f}_t + \ddot{w}_h^T \ddot{f}_h$$

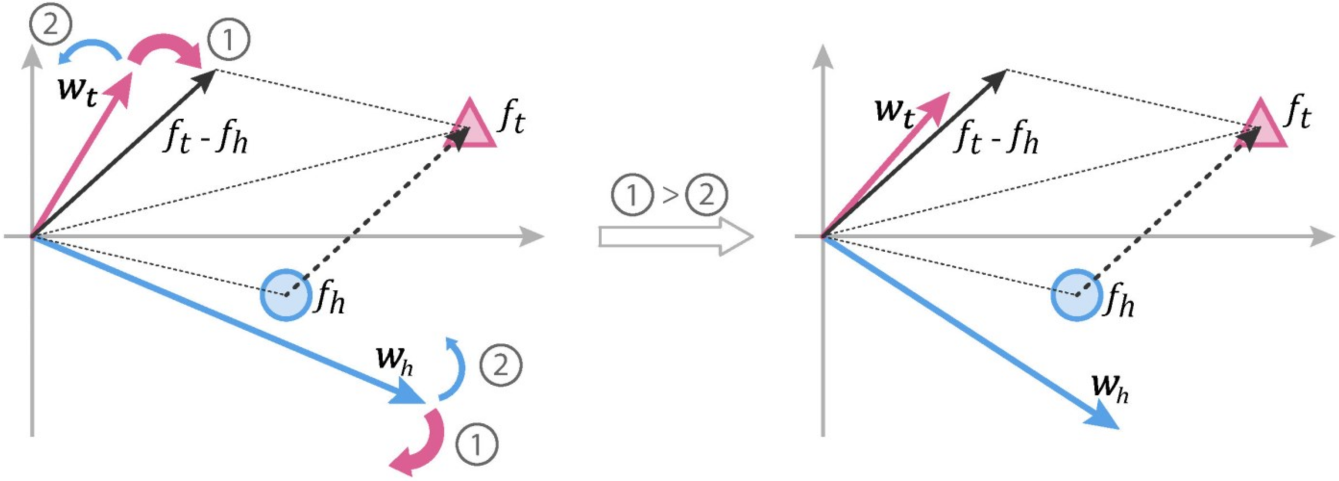
Force ①: $|\ddot{w}_h^T| \cos \ddot{\theta}_h > |\ddot{w}_t^T| \cos \ddot{\theta}_t$

For head

$$w_h^T \tilde{f}_h > w_h^T \tilde{f} \Rightarrow$$
$$\dot{w}_h^T \dot{f}_h + \ddot{w}_t^T \ddot{f}_t > \dot{w}_t^T \dot{f}_h + \ddot{w}_t^T \ddot{f}_h$$

Force ② : $|\dot{w}_t^T| \cos \dot{\theta}_t > |\ddot{w}_h^T| \cos \dot{\theta}_h$

Rationale



Forces ①>② makes the tail sample to "pull" closer to and "push" further away from w_h, leading to the adjustment of decision boundary and enlargement of the tail class space.

Experiments

AAAI-24

Imbalance Ratio	200	100	50
Single Model			
CE loss	35.99	39.55	45.40
LDAM-DRW (Cao et al. 2019)	38.91	42.04	47.62
DR+MU (Kang et al. 2020)	41.73	45.68	50.86
MisLAS (Zhong et al. 2021)	43.45	46.71	52.31
MBJ* (Liu et al. 2022)	-	45.80	52.60
GCL (Li et al. 2022)	**44.76**	48.61	**53.55**
CE+CMO* (Park et al. 2022)	-	43.90	48.30
BSCE+CMO* (Park et al. 2022)	-	46.60	51.40
ABL* (Jin et al. 2023)	-	46.80	52.10
DR+H2T	43.95	**47.73**	52.95
MisLAS+H2T	43.84	47.62	52.73
GCL+H2T	**45.24**	**48.88**	**53.76**
Multi-Expert Model			
BBN (Zhou et al. 2020)	37.21	42.56	47.02
RIDE (Wang et al. 2021b)	45.84	50.37	**54.99**
ACE* (Cai et al. 2021)	-	49.40	50.70
RIDE+CMO* (Park et al. 2022)	-	**50.00**	53.00
ResLT (Cui et al. 2023)	44.26	48.73	53.81
RIDE+H2T	**46.64**	**51.38**	**55.54**
ResLT+H2T	**46.18**	49.60	54.39

Comparison results on CIFAR100-LT

Experiments

AAAI-24

Method	Head	Med	Tail	All
Single Model				
CE loss	64.91	38.10	11.28	44.51
LDAM-DRW (Cao et al. 2019)	58.63	48.95	30.37	49.96
DR (Kang et al. 2020)	62.93	49.77	33.26	52.18
MisLAS (Zhong et al. 2021)	62.53	49.82	34.74	52.29
MBJ⋆ (Liu et al. 2022)	61.60	48.40	39.00	52.10
GCL (Li et al. 2022)	62.24	48.62	52.12	**54.51**
CE+CMO⋆ (Park et al. 2022)	67.00	42.30	20.50	49.10
BSCE+CMO⋆ (Park et al. 2022)	62.00	49.10	36.70	52.30
ABL⋆ (Jin et al. 2023)	62.60	50.30	36.90	53.20
DR+H2T	63.26	50.43	34.11	52.74
MisLAS+H2T	62.42	51.07	35.36	52.90
GCL+H2T	62.36	48.75	52.15	**54.62**
Multi-Expert Model				
BBN (Zhou et al. 2020)	-	-	-	48.30
RIDE (Wang et al. 2021b)	69.59	53.06	30.09	55.72
ACE⋆ (Cai et al. 2021)	-	-	-	54.70
RIDE+CMO⋆ (Park et al. 2022)	66.40	53.90	35.60	**56.20**
ResLT (Cui et al. 2023)	59.39	50.97	41.29	52.66
RIDE+H2T	67.55	54.95	37.08	**56.92**
ResLT+H2T	62.29	52.29	35.31	53.39

Method	Head	Med	Tail	All
Single Model				
CE loss	76.10	69.05	62.44	66.86
LDAM-DRW (Cao et al. 2019)	-	-	-	68.15
DR (Kang et al. 2020)	72.88	71.15	69.24	70.49
MisLAS (Zhong et al. 2021)	72.52	72.08	70.76	71.54
MBJ⋆ (Liu et al. 2022)	-	-	-	70.00
GCL (Li et al. 2022)	66.43	71.66	72.47	71.47
CE+CMO⋆ (Park et al. 2022)	76.90	69.30	66.60	68.90
BSCE+CMO⋆ (Park et al. 2022)	68.80	70.00	72.30	70.90
ABL⋆ (Jin et al. 2023)	-	-	-	71.60
DR+H2T	71.73	72.32	71.30	**71.81**
MisLAS+H2T	69.68	72.49	72.15	**72.05**
GCL+H2T	67.74	71.92	72.22	71.62
Multi-Expert Model				
BBN (Zhou et al. 2020)	-	-	-	69.70
RIDE (Wang et al. 2021b)	76.52	74.23	70.45	72.80
ACE⋆ (Cai et al. 2021)	-	-	-	**72.90**
RIDE+ CMO⋆ (Park et al. 2022)	68.70	72.60	73.10	72.80
ResLT (Cui et al. 2023)⋆	64.85	70.64	72.11	70.69
RIDE+H2T	75.69	74.22	71.36	**73.11**
ResLT with H2T	68.41	72.31	72.09	71.88

Comparison results on imageNet-LT and iNaturalist 2018

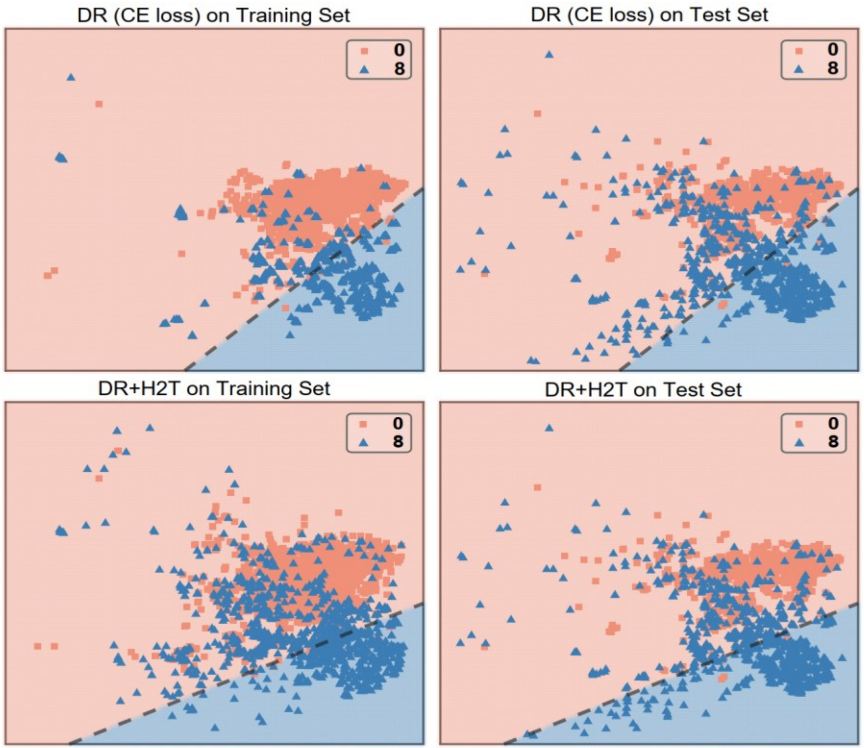
Experiments

AAAI-24

Method	Head	Med	Tail	All
Single Model				
CE loss	46.48	25.66	8.09	29.43
DR (Kang et al. 2020)	41.66	37.79	32.77	37.40
MisLAS (Zhong et al. 2021)	41.95	41.88	34.65	40.38
MBJ* (Liu et al. 2022)	39.50	38.20	35.50	38.10
GCL (Li et al. 2022)	38.64	42.59	38.44	40.30
ABL* (Jin et al. 2023)	41.50	40.80	31.40	39.40
DR+H2T	41.96	42.87	35.33	**40.95**
MisLAS+H2T	41.40	43.04	35.95	<u>**41.03**</u>
GCL+H2T	39.34	42.50	39.46	40.73
Multi-Expert Model				
LFME* (Xiang et al. 2020)	39.30	39.60	24.20	36.20
RIDE (Wang et al. 2021b) (LC)	44.79	40.69	31.97	40.32
RIDE (Wang et al. 2021b) (CC)	44.38	40.59	32.99	40.35
ResLT* (Cui et al. 2023)	40.30	44.40	34.70	41.00
RIDE+H2T (LC)	42.99	42.55	36.25	<u>**41.38**</u>
RIDE+H2T (CC)	42.34	43.21	35.62	**41.30**

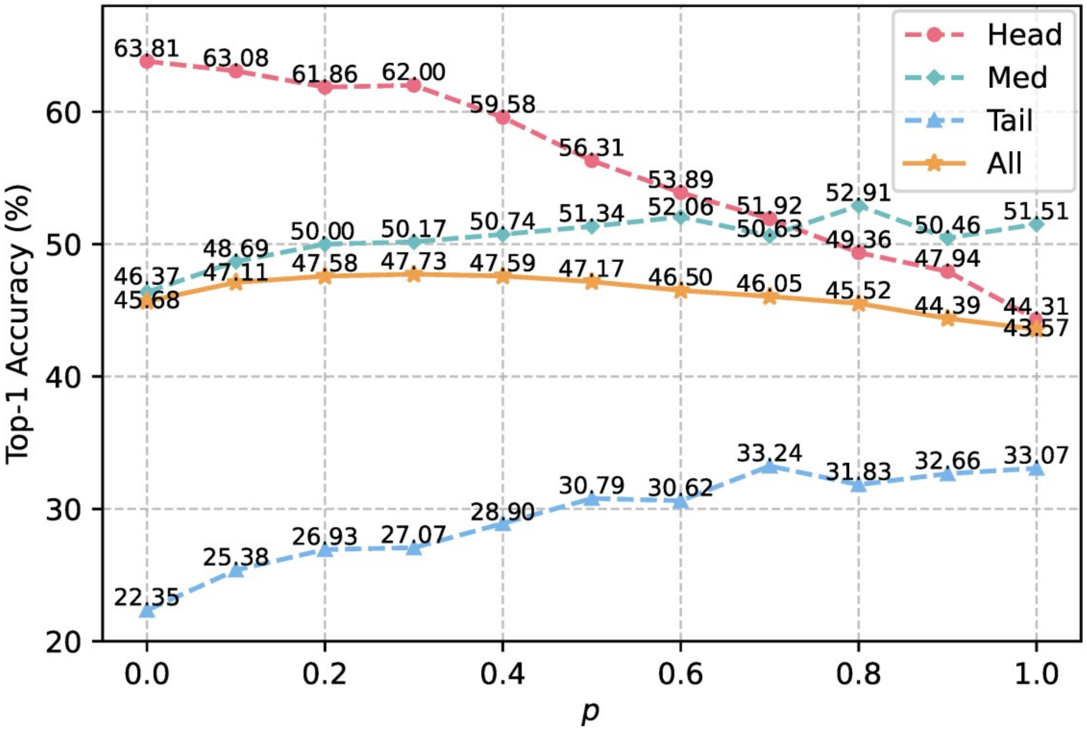
Comparison results on Places-LT. CC, cosine classifier. LC, linear classifier.

Experiments



Decision boundary comparison

Experiments



Change of accuracy (%) on different splits w.r.t. fusing ratio p.

Conclusion

AAAI-24

Pros:

Simple but effective:

- A plug-and-play module;
- H2T produces abundant features to augment tail classes;
- H2T generates two opposing forces that restrain each other, preventing excessive sacrifices in head class accuracy.

Cons:

Sacrifice head class performance:

- H2T only adjusts the decision boundary without altering the feature distribution.

光明实验室 GUANGMING LABORATORY 人工智能与数字经济广东省实验室(深圳)

深圳大学 SHENZHEN UNIVERSITY

香港浸會大學 HONG KONG BAPTIST UNIVERSITY

厦门大学

Thanks

AAAI 2024