Poster Session
THU-PM-326

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

HyperMatch: Noise-Tolerant Semi-Supervised Learning via Relaxed Contrastive Constraint

Beitong Zhou*, Jing Lu*, Kerui Liu, Yunlu Xu, Zhanzhan Cheng†, Yi Niu

Hikvision Research Institute

HIKVISION®

Overviews

In this paper:

- We proposed an **enhanced contrastive learning** method, HyperMatch, to handle the effective separation and exploitation of clean and noisy pseudo labels.

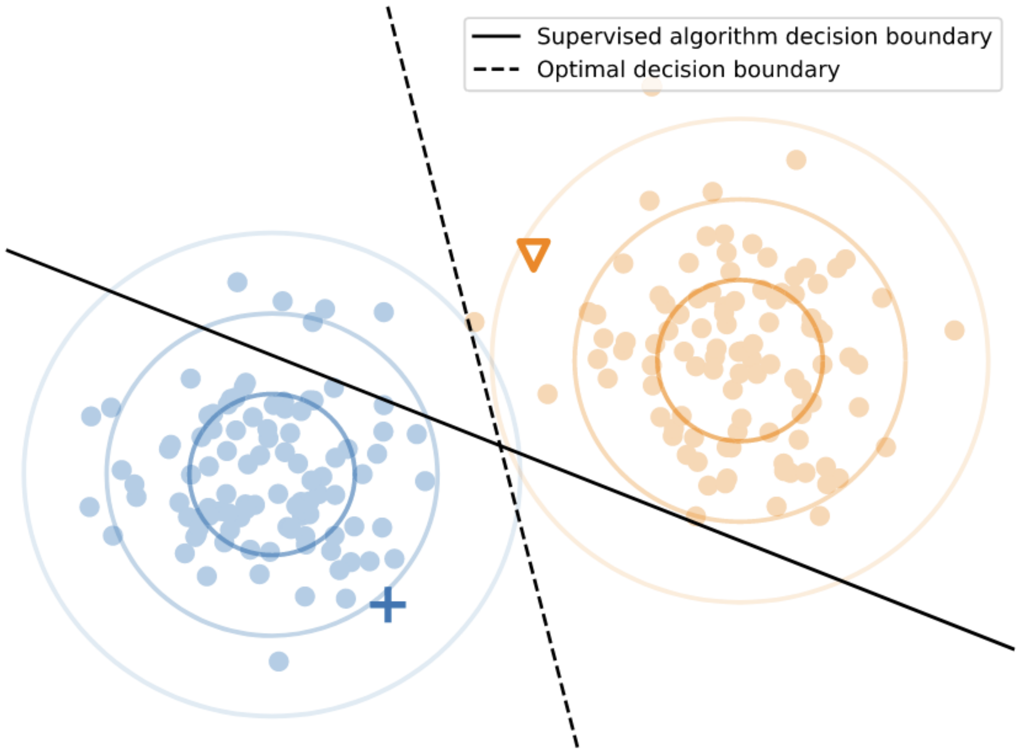
- We relax the assignment be categorizing the noisy sample into a **hyper-class** (a union of top-K nearest classes), followed by the proposed **Relaxed Contrastive Loss** to mitigate the confirmation bias.

- HyperMatch achieves superior performance on SSL benchmarks.

Semi-Supervised Learning

Semi-supervised learning methods attempt to utilize **unlabeled data** to construct a classifier whose performance exceeds the performance of classifiers obtained using only **labeled data**.

The cluster from the unlabeled data helps us considerably in placing the decision boundary.



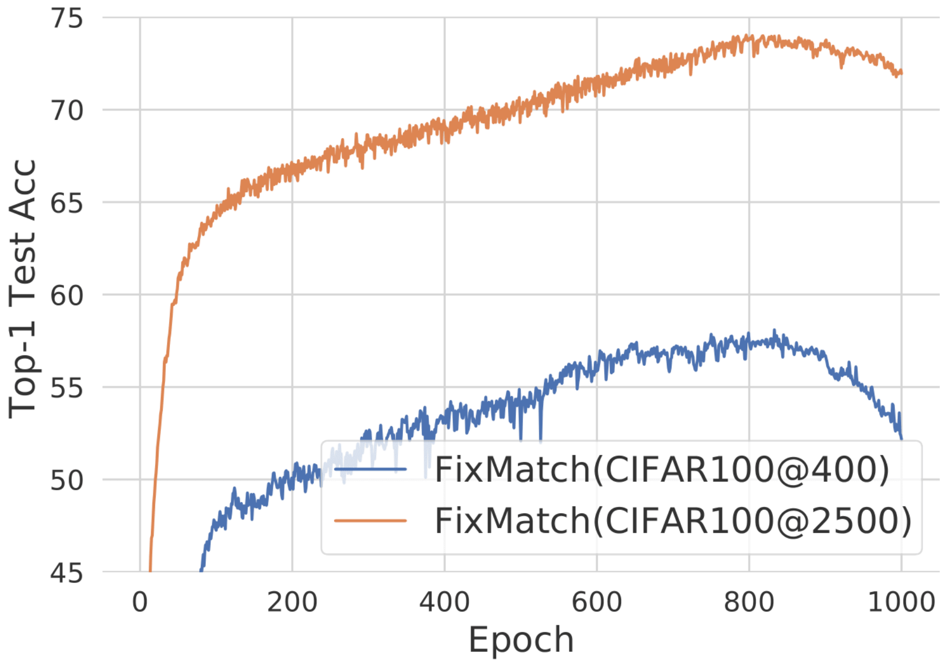
A basic example of binary classification in the presence of unlabeled data.

van Engelen, J.E., & Hoos, H.H. (2019). A survey on semi-supervised learning. Machine Learning, 109, 373-440.

Confirmation Bias

The model makes over-confident errors and reinforces those misclassified patterns while ignoring or underweighting evidence that contradicts them.

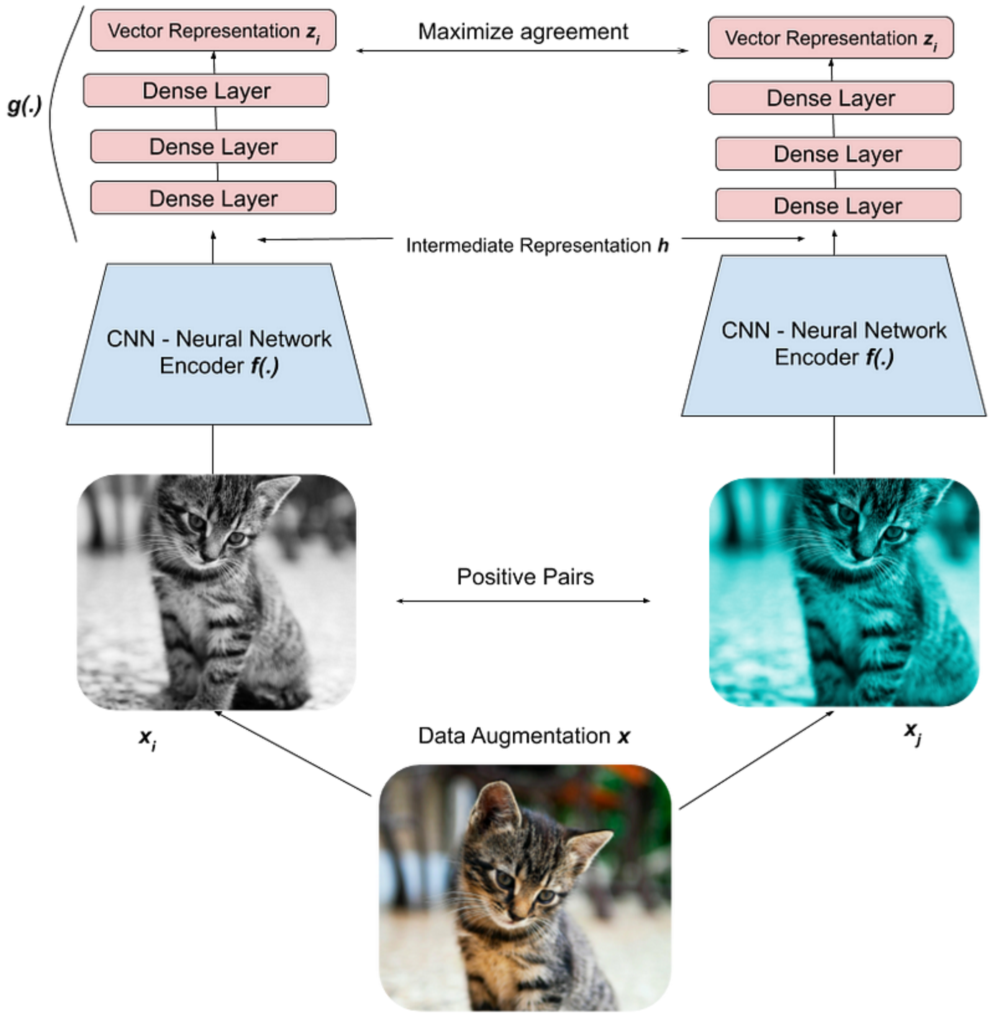
This problematic phenomenon is named **confirmation bias**, which prevents the model from further improving its performance.



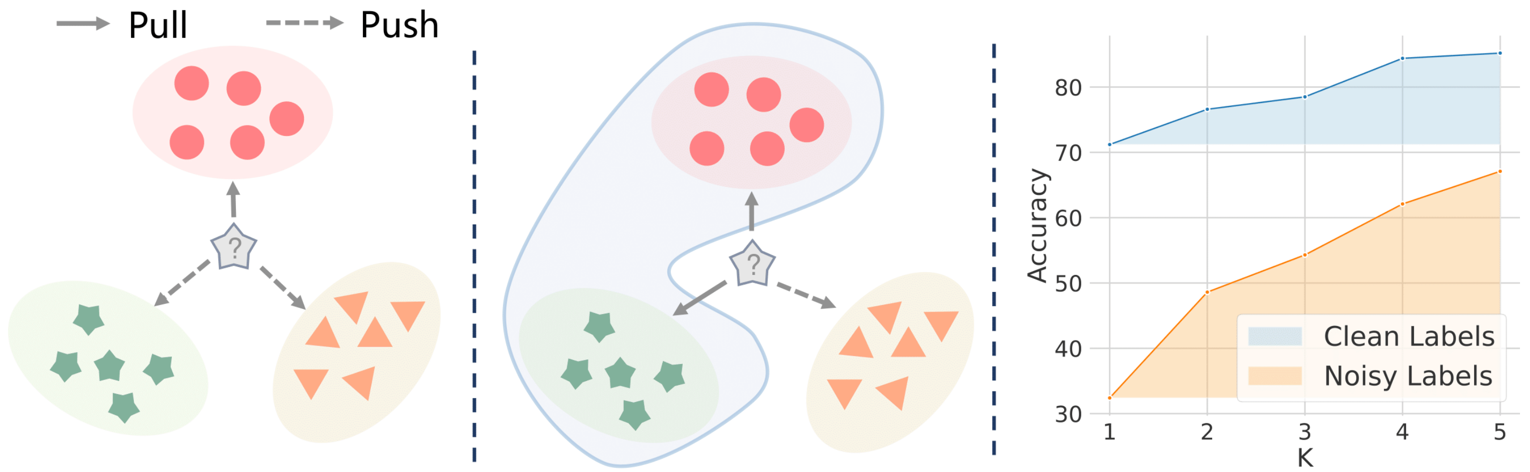
The test accuracy of FixMatch drops after a long period of training due to confirmation bias.

Contrastive Learning

Contrastive learning is a machine learning technique used to learn the general features of a dataset without labels by teaching the model which data points are similar or different.



Motivation



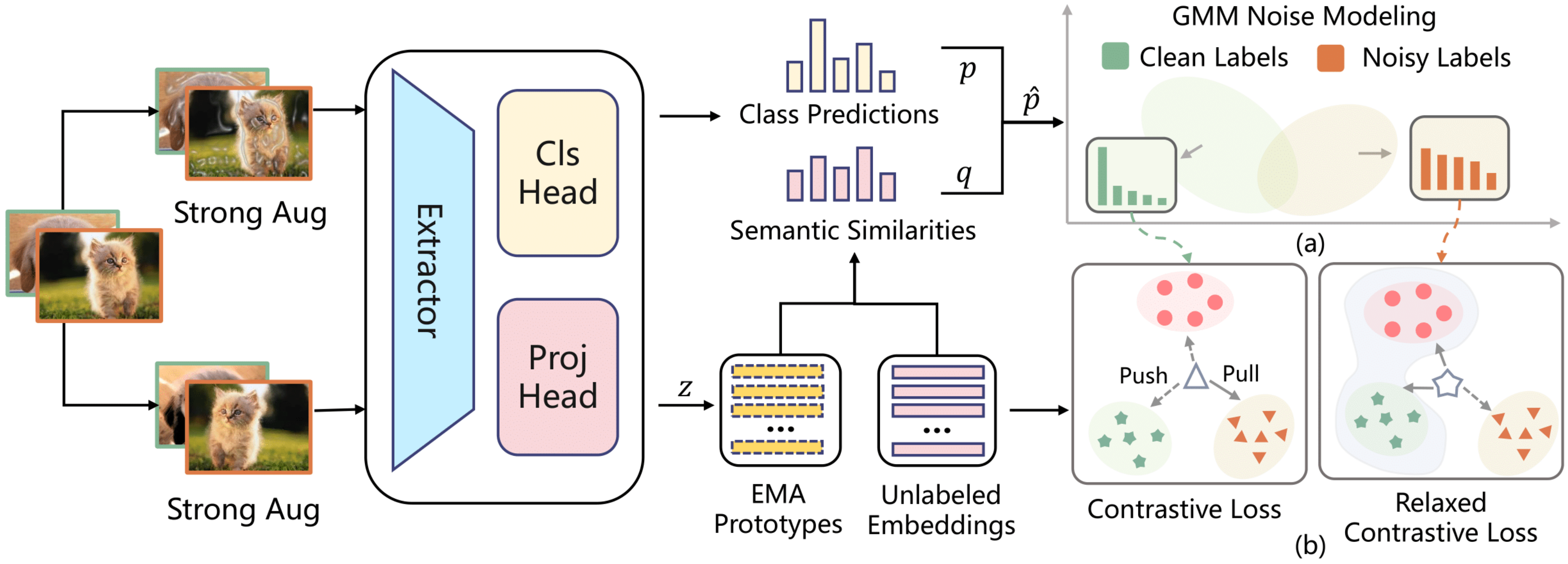
a) Pseudo-Label-based Class Assignment
b) Hyper-Class Assignment
c) Top-K Accuracy of Pseudo Labels

a) **Class assignment**: the instance is pulled close to the wrong pseudo label class and pushed away from ground-truth class;

b) **Hyper-Class assignment**: the instance is assigned to hyper-class (the union of top-K nearest classes) with ground-truth class included;

c) As K grows, noisy labels benefit much more than clean labels;

Method



a) Pseudo label partition

Class distribution calibration

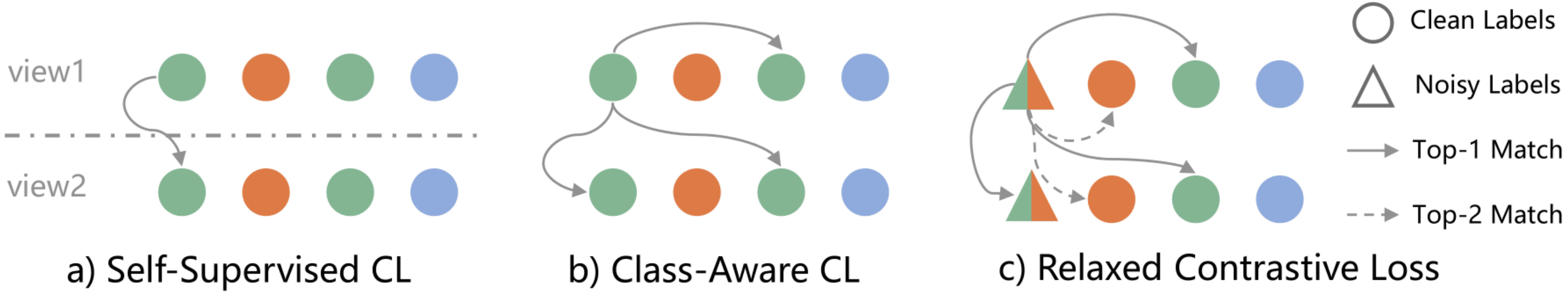
GMM noise modeling

a) Relaxed Contrastive Loss

Clean samples – contrastive loss

Noisy samples – relaxed contrastive loss

Method



$$w_{ij}^{\text{self}} = \begin{cases} 1, & \text{if } z_i \text{ and } z_j \text{ are from the same sample} \\ 0, & \text{otherwise} \end{cases}$$

$$w_{ij}^{\text{class}} = \begin{cases} w_{i,j}^{\text{re}}, & \text{if } \hat{y}_i = \hat{y}_j \\ 0, & \text{otherwise} \end{cases}$$

$$w_{ij}^{\text{relax}} = \begin{cases} \hat{p}_i^{c_{i,k}} * \hat{p}_j^{c_{i,k}}, & \text{if } u_j \in \mathcal{HS}_i \\ 0, & \text{otherwise} \end{cases}$$

Experiments

Method	CIFAR10			CIFAR100			STL10
	40	250	4000	400	2500	10000	
MixMatch [3]	52.46 ± 11.5	88.95 ± 0.86	93.58 ± 0.10	32.39 ± 1.32	60.06 ± 0.37	71.69 ± 0.33	38.02 ± 8.29
ReMixMatch [2]	80.90 ± 9.64	94.56 ± 0.05	95.28 ± 0.13	55.72 ± 2.06	72.57 ± 0.31	76.97 ± 0.56	-
SSWPL [33]	-	-	-	-	73.48 ± 0.45	79.12 ± 0.85	-
LaplaceNet [28]	-	-	95.35 ± 0.07	-	68.36 ± 0.02	73.40 ± 0.23	-
FixMatch(RA) [29]	86.19 ± 3.37	94.93 ± 0.65	95.74 ± 0.05	51.15 ± 1.75	71.71 ± 0.11	77.40 ± 0.12	65.38 ± 0.42
CoMatch [24]	93.09 ± 1.39	95.09 ± 0.33	95.44 ± 0.20	58.11 ± 2.34	71.63 ± 0.35	79.14 ± 0.36	79.80 ± 0.38
SimMatch [42]	$\mathbf{94.40 \pm 1.37}$	$\mathbf{95.16 \pm 0.39}$	96.04 ± 0.01	62.19 ± 2.21	74.93 ± 0.32	79.42 ± 0.11	-
CCSSL [39]	90.83 ± 2.78	94.86 ± 0.55	95.54 ± 0.20	61.19 ± 1.65	75.7 ± 0.63	80.68 ± 0.16	80.01 ± 1.39
HyperMatch	93.92 ± 1.10	95.01 ± 0.23	$\mathbf{96.05 \pm 0.12}$	$\mathbf{63.01 \pm 0.57}$	$\mathbf{76.45 \pm 0.35}$	$\mathbf{81.09 \pm 0.28}$	$\mathbf{82.98 \pm 0.37}$

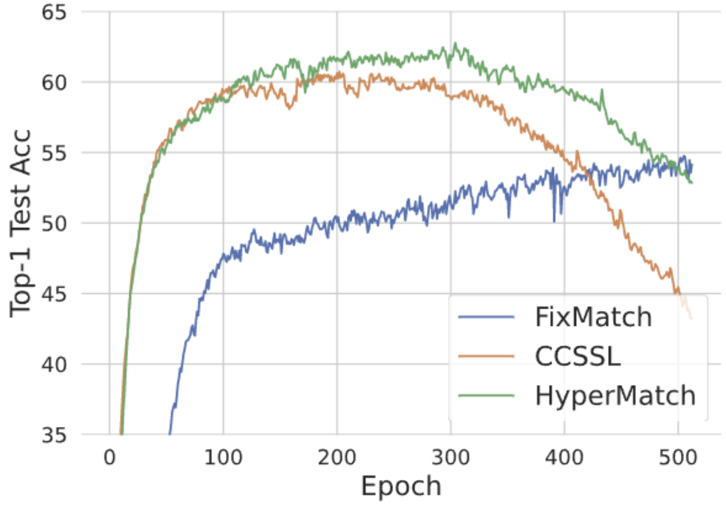
(1) CIFAR10 / CIFAR100 / STL10 image classification

HyperMatch achieves SOTA performance on several SSL benchmarks and also steady performance improvement on a more complex Semi-iNat 2021 dataset with OOD samples.

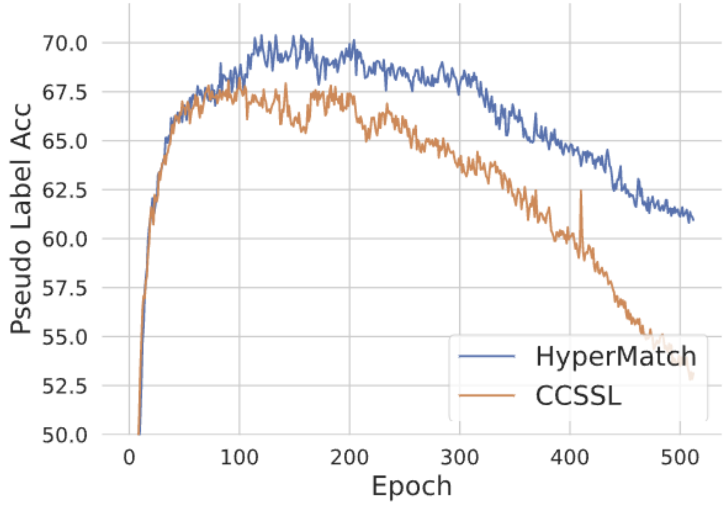
Method	Semi-iNat 2021	
	From Scratch	Moco Pretrain
Supervised	19.09	34.96
CoMatch [24]	20.94	38.94
FixMatch [29]	21.41	40.3
CCSSL (CoMatch) [39]	24.12	39.85
CCSSL (FixMatch) [39]	31.21	41.28
HyperMatch (FixMatch)	**33.47**	**42.57**

(2) Semi-iNat 2021

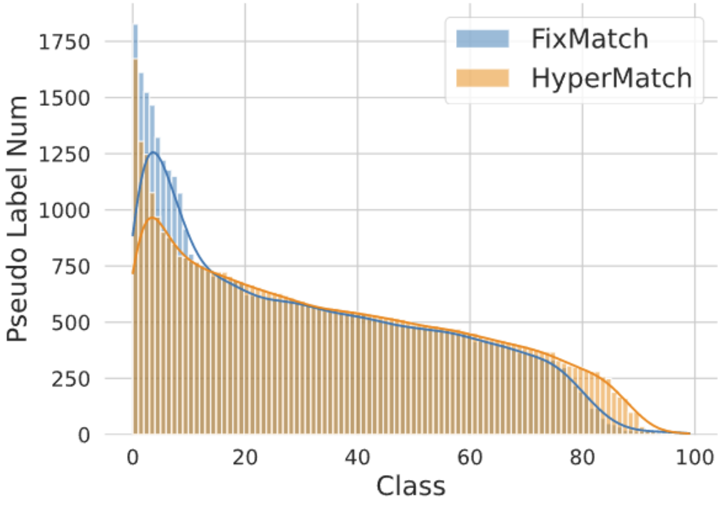
Experiments



a) Test Accuracy



b) Pseudo Label Accuracy



c) Pseudo Label Class Distributions

HyperMatch shows fast convergence speed and better pseudo label accuracy compared with other methods. The mitigated imbalanced distributions of pseudo label class indicates HyperMatch's ability to alleviate confirmation bias.

Poster Session
THU-PM-326

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Thanks!

Beitong Zhou*, Jing Lu*, Kerui Liu, Yunlu Xu, Zhanzhan Cheng†, Yi Niu
Hikvision Research Institute
zhoubt@hust.edu.cn, {lujing6, liukerui, xuyunlu, chengzhanzhan, niuyi}@hikvision.com

HIKVISION®