ICML International Conference On Machine Learning
2026 · SEOUL

Code

WeChat

Von Mises-Fisher Mixture Model with Dynamic Shrinkage for Realistic Test-Time Transduction

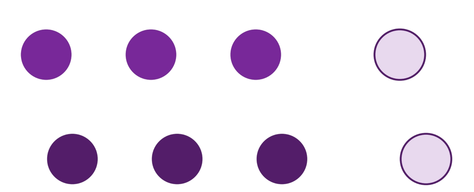
Jiazhen Huang[1] · Zhiming Liu[1] · Changhu Wang[2] · Wei Ju[3,✉] · Ziyue Qiao[4,✉] · Xiao Luo[5]

[1]Tsinghua University · [2]Fred Hutchinson Cancer Center · [3]Peking University · [4]Great Bay University · [5]University of Wisconsin–Madison

Overview & Background

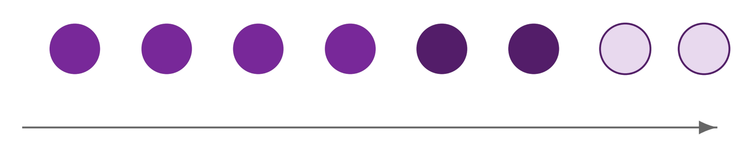
Realistic Class Imbalance

Batch adaptation



Limited effective classes; batches processed independently

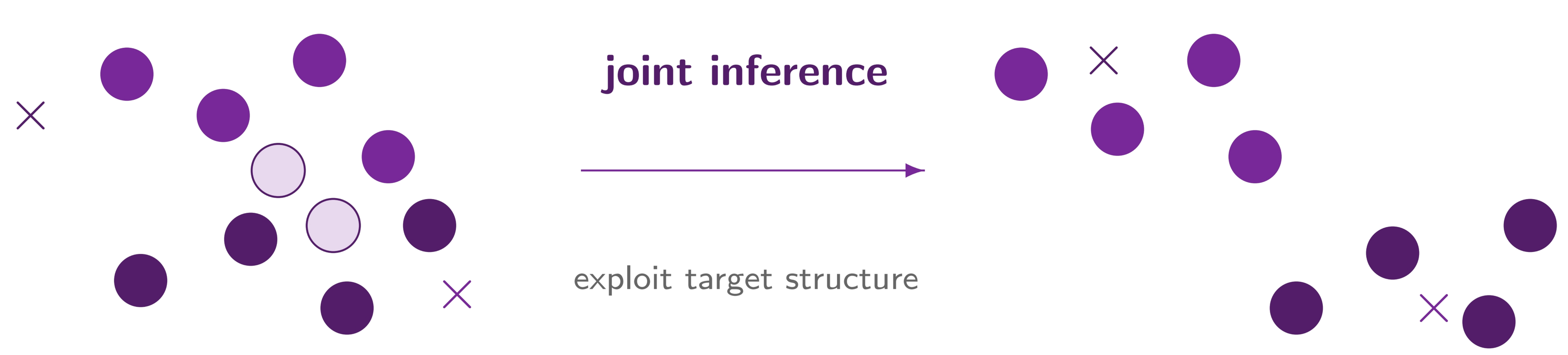
Online adaptation



non-i.i.d. stream; historical batches accessible

Transduction Paradigm

Zero-shot predictions **Adapted predictions**



- Perform **joint prediction** by exploiting all available samples
- Being **training-free, model-agnostic** and **fast**
- Novel paradigm for Test-Time Adaptation (TTA)

PLE View: Shrinkage Toward a KL Anchor

For any regular minimal exponential family, **a KL-anchored PLE objective** yields an **adaptive shrinkage behavior** for transduction.

$$p(x\mid\eta)=h(x)\exp\big(\eta^\top T(x)-A(\eta)\big).$$

$$\min_{\{\eta_k\}}-\sum_{i=1}^{N}\sum_{k=1}^{K}z_{i,k}\log p(x_i\mid\eta_k)+\alpha\sum_{k=1}^{K}\mathrm{KL}\big(q_k\,\|\,p(\cdot\mid\eta_k)\big),$$

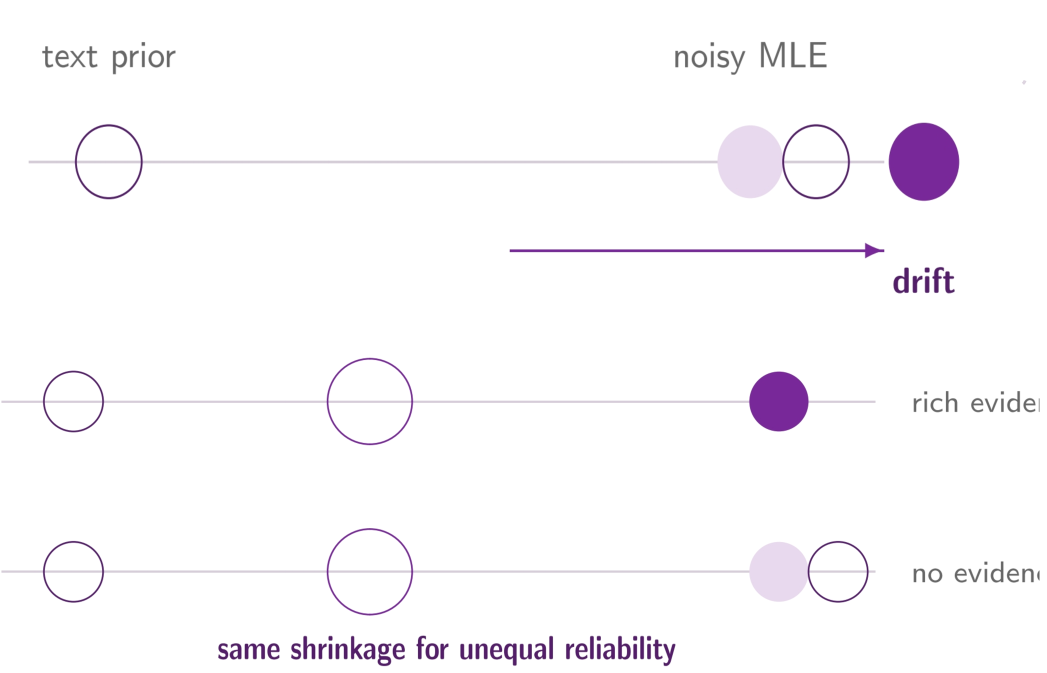
PLE objective

$$\nabla A(\eta_k^*)=\frac{S_k+\alpha\mu'_k}{n_k+\alpha}=\beta_k\hat\mu_k+(1-\beta_k)\mu'_k,\qquad \beta_k=\frac{n_k}{n_k+\alpha}.$$

Data-driven convex combination in mean-parameter space.

Limitation 1: no anchor
With α = 0, PLE reduces to MLE and can overfit locally biased statistics.

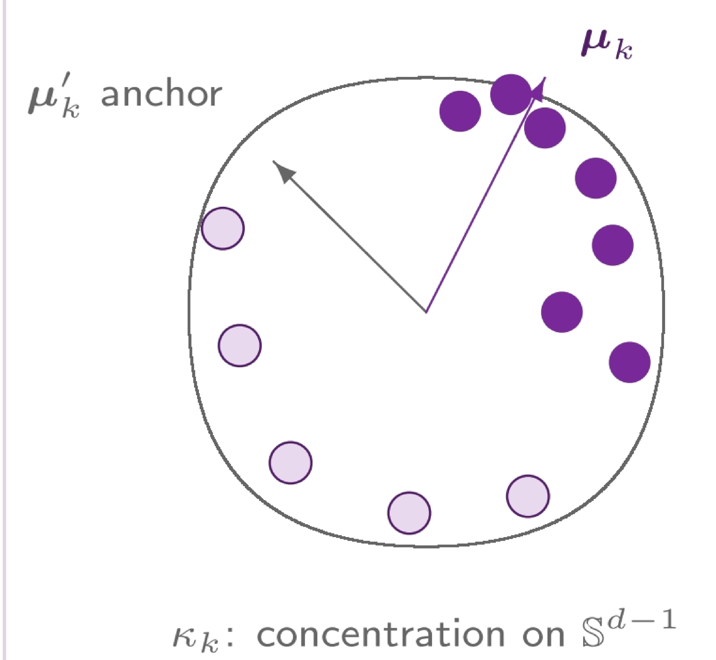
Limitation 2: bounded but static shrinkage
A fixed anchor weight α assumes equal reliability across instances and classes.



Method

MOON: Training-free Test-Time Tranduction

vMF Mixture Model



κ_k: concentration on $\mathbb{S}^{d-1}$

$$p^{\text{vMF}}_{i,k}\propto C_d(\kappa_k)\exp(\kappa_k\boldsymbol{\mu}_k^\top\mathbf{f}_i)$$

MOON Objective

Dynamic weights from zero-shot priors!

$$\mathcal{L}_{\text{PLE}}(\mathbf{z};\boldsymbol{\mu},\boldsymbol{\kappa})=\gamma\Big[\underbrace{-\sum_i\mathbf{z}_i^\top\log\mathbf{p}_i^{\text{vMF}}}_{\text{Log-Likelihood}}+\mathcal{R}(\mathbf{z})\Big]+\underbrace{\alpha\sum_k\mathrm{KL}(\mathcal{V}'_k\|\mathcal{V}_k)}_{\text{KL Statistical anchor}},$$

$$\mathcal{R}(\mathbf{z})=\underbrace{-\sum_{i,j}\omega_{ij}\mathbf{z}_i^\top\mathbf{z}_j}_{\text{Laplacian reg.}}+\underbrace{\sum_i\mathrm{KL}(\mathbf{z}_i\|\hat{\mathbf{y}}_i)}_{\text{Text supervision}}.$$

MOON models VLM embeddings **on the unit hypersphere**, and perform a **KL-anchored PLE objective** with **dynamic shrinkge** for adaptation.

TL; DR

MOON is a training-free, model-agnostic, and efficient method for test-time adaptation (TTA) of VLMs (e.g., CLIP), designed for realistic class imbalanced scenarios.

Dynamic Shrinkage

Instance-level

$$\gamma_i=1-\frac{H(\hat{\mathbf{y}}_i)}{\log K},$$

entropy-guided

$$H(\hat{\mathbf{y}}_i)=-\sum_{k=1}^{K}\hat{y}_{i,k}\log\hat{y}_{i,k}$$

High-entropy samples contribute less to parameter estimation.

Class-level

$$\alpha_k=\frac{1}{\lambda_k},$$

confidence (max prob)-guided

$$\lambda_k=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\hat{y}_{i,k}\odot\max_i\hat{y}_{i,k}}.$$

Low-confidence outlier classes shrink more towards the anchor.

Closed-Form Parameter Update

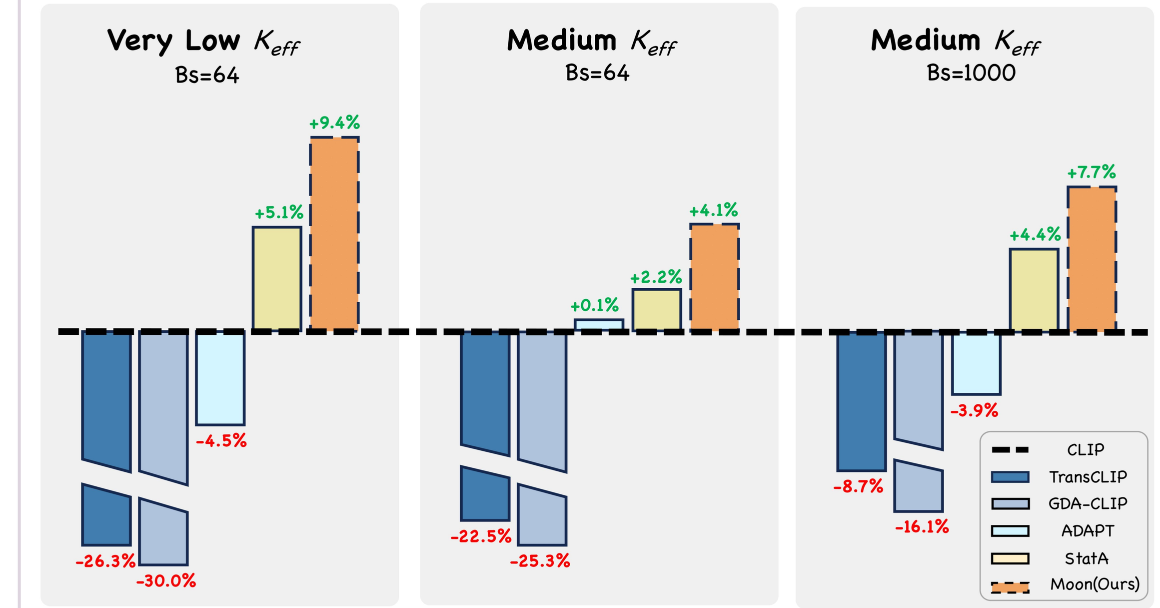
$$n_k=\sum_i\gamma_i z_{i,k},\qquad \beta_k=\frac{n_k}{n_k+\alpha_k},$$

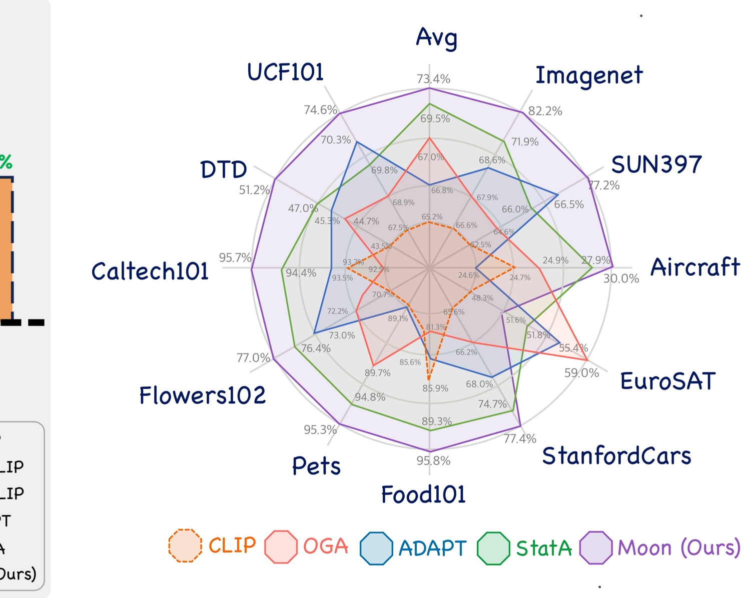
$$\boldsymbol{\mu}_k=\frac{\beta_k\boldsymbol{v}_k+(1-\beta_k)\boldsymbol{\mu}'_k}{\|\beta_k\boldsymbol{v}_k+(1-\beta_k)\boldsymbol{\mu}'_k\|},\quad A_d(\kappa_k)=\|\beta_k\boldsymbol{v}_k+(1-\beta_k)\boldsymbol{\mu}'_k\|,$$

BSUM iteration, until convergence

Experiment

Main Results under Realistic Imbalance





K_{eff}	Method	ImageNet	SUN397	Aircraft	EuroSAT	StanfordCars	Food101	Pets	Flowers102	Caltech101	DTD	UCF101	Avg.
	CLIP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	67.5	65.2
	MTA	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	69.0$_{+1.5}$	66.6$_{+1.3}$
Very Low (1~4)	Dirichlet	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	69.0$_{+1.5}$	68.5$_{+3.3}$
	ZLaP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	29.2$_{-38.3}$	27.5$_{-37.8}$
	GDA-CLIP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	35.2$_{-32.3}$	35.3$_{-30.0}$
	TransCLIP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	37.7$_{-29.8}$	38.9$_{-26.3}$
	ADAPT	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	61.4$_{-6.1}$	60.7$_{-4.5}$
	StatA	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	70.2$_{+2.7}$	70.4$_{+5.1}$
	MOON	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	75.7$_{+8.2}$	**74.6**$_{+9.4}$
Low (2~10)	Dirichlet	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	70.9$_{+3.4}$	70.3$_{+5.1}$
	ZLaP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	38.7$_{-28.8}$	35.2$_{-30.0}$
	GDA-CLIP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	37.2$_{-30.3}$	35.5$_{-29.8}$
	TransCLIP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	40.9$_{-26.6}$	40.4$_{-24.8}$
	ADAPT	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	63.8$_{-3.7}$	63.7$_{-1.5}$
	StatA	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	69.6$_{+2.1}$	69.3$_{+4.1}$
	MOON	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	74.2$_{+6.7}$	**73.1**$_{+7.9}$
Medium (5~25)	Dirichlet	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	67.8$_{+0.3}$	67.5$_{+2.2}$
	ZLaP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	50.3$_{-17.2}$	44.7$_{-20.6}$
	GDA-CLIP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	43.7$_{-23.8}$	40.0$_{-25.3}$
	TransCLIP	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	46.5$_{-21.0}$	42.7$_{-22.5}$
	ADAPT	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	66.7$_{-0.8}$	65.3$_{+0.1}$
	StatA	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	69.1$_{+1.6}$	67.4$_{+2.2}$
	MOON	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	71.5$_{+4.0}$	**69.4**$_{+4.1}$

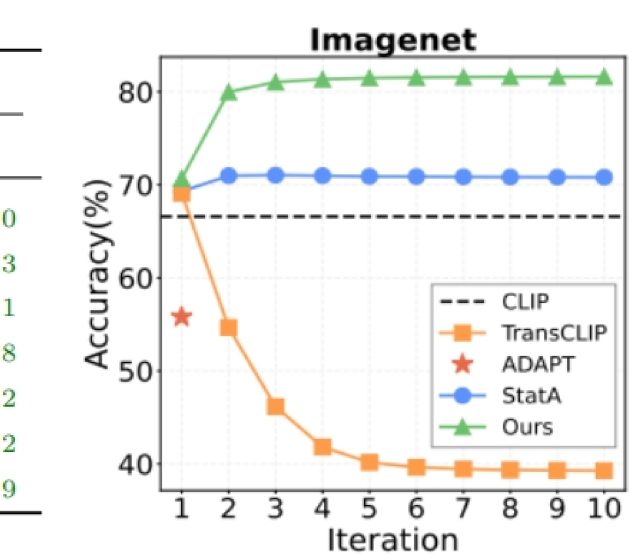
Scenario	Method	ImageNet	SUN397	Aircraft	EuroSAT	StanfordCars	Food101	Pets	Flowers102	Caltech101	DTD	UCF101	Avg.
	CLIP	66.6	62.5	24.7	48.3	65.6	85.9	89.1	70.7	93.2	43.5	67.5	65.2
	MTA	69.3$_{+2.7}$	64.8$_{+2.3}$	27.4$_{+2.7}$	46.9$_{-1.4}$	68.0$_{+2.4}$	87.2$_{+1.3}$	89.4$_{+0.3}$	71.7$_{+1.0}$	94.0$_{+0.8}$	44.4$_{+0.9}$	69.0$_{+1.5}$	66.6$_{+1.3}$
Low ($\xi=0.1$)	TENT	66.6$_{\pm0.0}$	64.5$_{+2.0}$	24.6$_{-0.1}$	51.8$_{+3.5}$	65.7$_{+0.1}$	85.9	89.3$_{+0.2}$	70.6$_{-0.1}$	93.4$_{+0.2}$	44.0$_{+0.5}$	67.8$_{+0.3}$	65.8$_{+0.6}$
	TDA	68.3$_{+1.7}$	66.0$_{+3.5}$	25.4$_{+0.7}$	60.6$_{+12.3}$	66.9$_{+1.3}$	86.1$_{+0.2}$	89.6$_{+0.5}$	72.5$_{+1.8}$	93.4$_{+0.2}$	45.5$_{+2.0}$	71.0$_{+3.5}$	**67.7**$_{+2.5}$
	DMN	68.0$_{+1.4}$	64.8$_{+2.3}$	24.9$_{+0.2}$	59.8$_{+11.5}$	67.0$_{+1.4}$	84.2$_{-1.7}$	89.9$_{+0.8}$	73.3$_{+2.6}$	92.2$_{-1.0}$	44.8$_{+1.3}$	70.3$_{+2.8}$	67.2$_{+2.0}$
	OGA	68.8$_{+2.2}$	66.1$_{+3.6}$	24.9$_{+0.2}$	61.8$_{+13.5}$	66.2$_{+0.6}$	86.2$_{+0.3}$	90.2$_{+1.1}$	72.0$_{+1.3}$	93.4$_{+0.2}$	44.9$_{+1.4}$	69.4$_{+1.9}$	67.6$_{+2.4}$
	ADAPT	69.2$_{+2.6}$	65.4$_{+2.9}$	24.4$_{-0.3}$	51.9$_{+3.6}$	67.6$_{+2.0}$	77.4$_{-8.5}$	88.5$_{-0.6}$	72.9$_{+2.2}$	92.4$_{-0.8}$	45.2$_{+1.7}$	69.9$_{+2.4}$	65.9$_{+0.7}$
	StatA	66.2$_{-0.4}$	63.6$_{+1.1}$	24.3$_{-0.4}$	52.3$_{+4.0}$	67.4$_{+1.8}$	88.0$_{+2.1}$	92.5$_{+3.4}$	72.7$_{+2.0}$	94.2$_{+1.0}$	46.8$_{+3.3}$	68.8$_{+1.3}$	67.0$_{+1.7}$
	MOON	67.0$_{+0.4}$	63.5$_{+1.0}$	23.2$_{-1.5}$	50.1$_{+1.8}$	66.5$_{+0.9}$	91.2$_{+5.3}$	91.6$_{+2.5}$	71.4$_{+0.7}$	94.0$_{+0.8}$	45.3$_{+1.8}$	68.2$_{+0.7}$	66.5$_{+1.3}$
Medium ($\xi=0.01$)	TENT	66.7$_{+0.1}$	64.3$_{+1.8}$	24.6$_{-0.1}$	47.9$_{-0.4}$	65.6$_{\pm0.0}$	85.9$_{\pm0.0}$	89.4$_{+0.3}$	70.6$_{-0.1}$	93.3$_{+0.1}$	44.0$_{+0.5}$	67.8$_{+0.3}$	65.5$_{+0.2}$
	TDA	68.2$_{+1.6}$	65.6$_{+3.1}$	25.2$_{+0.5}$	56.5$_{+8.2}$	66.5$_{+0.9}$	85.8$_{-0.1}$	89.3$_{+0.2}$	72.6$_{+1.9}$	93.5$_{+0.3}$	45.2$_{+1.7}$	70.1$_{+2.6}$	67.1$_{+1.9}$
	DMN	68.0$_{+1.4}$	64.8$_{+2.3}$	24.9$_{+0.2}$	56.2$_{+7.9}$	66.8$_{+1.2}$	81.9$_{-4.0}$	89.0$_{-0.1}$	73.0$_{+2.3}$	92.1$_{-1.1}$	44.9$_{+1.4}$	69.6$_{+2.1}$	66.5$_{+1.2}$
	OGA	68.6$_{+2.0}$	65.4$_{+2.9}$	24.9$_{+0.2}$	58.6$_{+10.3}$	66.2$_{+0.6}$	85.9$_{\pm0.0}$	89.8$_{+0.7}$	72.2$_{+1.5}$	93.4$_{+0.2}$	44.8$_{+1.3}$	69.1$_{+1.6}$	67.2$_{+1.9}$
	ADAPT	69.3$_{+2.7}$	67.2$_{+4.7}$	24.6$_{-0.1}$	58.5$_{+10.2}$	68.6$_{+3.0}$	83.9$_{-2.0}$	90.2$_{+1.1}$	73.2$_{+2.5}$	92.5$_{-0.7}$	45.3$_{+1.8}$	71.3$_{+3.8}$	67.7$_{+2.5}$
	StatA	69.6$_{+3.0}$	65.9$_{+3.4}$	27.3$_{+2.6}$	52.3$_{+4.0}$	73.2$_{+7.6}$	89.1$_{+3.2}$	94.6$_{+5.5}$	75.6$_{+4.9}$	94.3$_{+1.1}$	46.8$_{+3.3}$	69.7$_{+2.2}$	68.9$_{+3.7}$
	MOON	76.5$_{+9.9}$	73.9$_{+11.4}$	28.3$_{+3.6}$	51.5$_{+3.2}$	75.1$_{+9.5}$	95.3$_{+9.4}$	94.8$_{+5.7}$	75.7$_{+5.0}$	95.3$_{+2.1}$	50.0$_{+6.5}$	73.4$_{+5.9}$	**71.8**$_{+6.6}$
High ($\xi=0.001$)	TENT	66.8$_{+0.2}$	64.3$_{+1.8}$	24.8$_{+0.1}$	45.6$_{-2.7}$	65.6$_{\pm0.0}$	86.1$_{+0.2}$	89.4$_{+0.3}$	70.5$_{-0.2}$	93.4$_{+0.2}$	44.0$_{+0.5}$	67.9$_{+0.4}$	65.3$_{+0.1}$
	TDA	67.9$_{+1.3}$	65.1$_{+2.6}$	25.1$_{+0.4}$	55.3$_{+7.0}$	66.3$_{+0.7}$	85.5$_{-0.4}$	89.0$_{-0.1}$	72.5$_{+1.8}$	93.6$_{+0.4}$	45.1$_{+1.6}$	69.7$_{+2.2}$	66.8$_{+1.6}$
	DMN	67.9$_{+1.3}$	64.8$_{+2.3}$	24.9$_{+0.2}$	56.3$_{+8.0}$	66.8$_{+1.2}$	79.9$_{-6.0}$	88.9$_{-0.2}$	72.9$_{+2.2}$	92.1$_{-1.1}$	44.8$_{+1.3}$	69.4$_{+1.9}$	66.3$_{+1.0}$
	OGA	67.9$_{+1.3}$	64.6$_{+2.1}$	24.9$_{+0.2}$	59.0$_{+10.7}$	66.2$_{+0.6}$	85.6$_{-0.3}$	89.7$_{+0.6}$	72.2$_{+1.5}$	93.5$_{+0.3}$	44.7$_{+1.2}$	68.9$_{+1.4}$	67.0$_{+1.8}$
	ADAPT	68.6$_{+2.0}$	66.5$_{+4.0}$	24.6$_{-0.1}$	55.4$_{+7.1}$	68.0$_{+2.4}$	81.3$_{-4.6}$	89.1$_{\pm0.0}$	73.0$_{+2.3}$	92.5$_{-0.7}$	45.3$_{+1.8}$	70.3$_{+2.8}$	66.8$_{+1.5}$
	StatA	71.9$_{+5.3}$	66.0$_{+3.5}$	27.9$_{+3.2}$	51.8$_{+3.5}$	74.7$_{+9.1}$	89.3$_{+3.4}$	94.8$_{+5.7}$	76.4$_{+5.7}$	94.4$_{+1.2}$	47.0$_{+3.5}$	69.8$_{+2.3}$	69.5$_{+4.2}$
	MOON	82.2$_{+15.6}$	77.2$_{+14.7}$	30.0$_{+5.3}$	51.6$_{+3.3}$	77.4$_{+11.8}$	95.8$_{+9.9}$	95.3$_{+6.2}$	77.0$_{+6.3}$	95.7$_{+2.5}$	51.2$_{+7.7}$	74.6$_{+7.1}$	**73.4**$_{+8.2}$
Separate	TENT	66.7$_{+0.1}$	64.2$_{+1.7}$	24.7$_{\pm0.0}$	37.0$_{-11.3}$	65.6$_{\pm0.0}$	86.1$_{+0.2}$	89.3$_{+0.2}$	70.8$_{+0.1}$	93.4$_{+0.2}$	43.9$_{+0.4}$	67.9$_{+0.4}$	64.5$_{-0.7}$
	TDA	67.4$_{+0.8}$	64.6$_{+2.1}$	24.9$_{+0.2}$	55.3$_{+7.0}$	65.9$_{+0.3}$	85.2$_{-0.7}$	88.9$_{-0.2}$	72.3$_{+1.6}$	93.6$_{+0.4}$	45.0$_{+1.5}$	69.6$_{+2.1}$	66.6$_{+1.4}$
	DMN	67.7$_{+1.1}$	64.7$_{+2.2}$	24.9$_{+0.2}$	55.1$_{+6.8}$	66.7$_{+1.1}$	78.5$_{-7.4}$	88.0$_{-1.1}$	72.8$_{+2.1}$	91.9$_{-1.3}$	44.8$_{+1.3}$	69.0$_{+1.5}$	65.8$_{+0.6}$
	OGA	67.2$_{+0.6}$	64.1$_{+1.6}$	24.9$_{+0.2}$	58.0$_{+9.7}$	66.1$_{+0.5}$	85.4$_{-0.5}$	89.5$_{+0.4}$	72.1$_{+1.4}$	93.4$_{+0.2}$	44.6$_{+1.1}$	68.7$_{+1.2}$	66.7$_{+1.5}$
	ADAPT	68.2$_{+1.6}$	65.7$_{+3.2}$	25.0$_{+0.3}$	53.7$_{+5.4}$	67.4$_{+1.8}$	78.2$_{-7.7}$	88.2$_{-0.9}$	72.4$_{+1.7}$	92.5$_{-0.7}$	44.8$_{+1.3}$	69.5$_{+2.0}$	66.0$_{+0.7}$
	StatA	71.7$_{+5.1}$	64.9$_{+2.4}$	28.9$_{+4.2}$	48.2$_{-0.1}$	75.2$_{+9.6}$	88.9$_{+3.0}$	95.2$_{+6.1}$	77.6$_{+6.9}$	94.3$_{+1.1}$	45.8$_{+2.3}$	69.0$_{+1.5}$	69.1$_{+3.8}$
	MOON	82.4$_{+15.8}$	76.7$_{+14.2}$	32.1$_{+7.4}$	51.2$_{+2.9}$	77.8$_{+12.2}$	95.5$_{+9.6}$	95.1$_{+6.0}$	77.9$_{+7.2}$	95.9$_{+2.7}$	53.6$_{+10.1}$	74.6$_{+7.1}$	**73.9**$_{+8.6}$

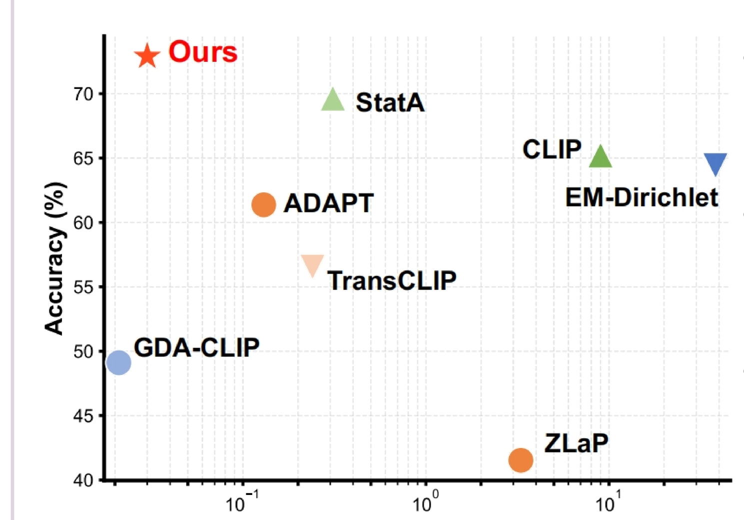
Further Discussion

α	γ	Update $\boldsymbol{\mu}$	Update $\boldsymbol{\kappa}$	Bs=64	Bs=1,000
✗	✓	✓	✓	69.0	68.9
✓	✗	✓	✓	72.3	70.4
✓	✓	✗	✓	66.8	66.5
✓	✓	✓	✗	70.8	68.9
✓	✓	✓	✓	**72.4**	**71.1**

Batch size	CLIP forward	TransCLIP	StatA	ADAPT	MOON
		10	10	1	10
128	8.47	0.08	0.11	0.06	**0.03**
1,000	8.99	0.24	0.31	0.13	**0.03**
50,000	36.90	0.47	0.59	0.21	**0.05**

	#Params	ImageNet		Average	
		Zero-shot	Ours	Zero-shot	Ours
CLIP RN50	102M	58.2	76.8 +18.6	58.7	65.7 +7.0
CLIP RN101	120M	61.3	77.1 +15.8	59.5	64.8 +5.3
CLIP ViT-B/32	151M	62.0	78.5 +16.5	61.9	68.0 +6.1
CLIP ViT-L/14	428M	73.5	84.9 +11.4	72.6	77.4 +4.8
OpenCLIP	150M	73.0	83.4 +10.4	72.5	76.7 +4.2
SigLIP	878M	82.3	91.3 +9.0	81.8	85.1 +3.2
EVA-CLIP	1.14B	78.0	81.2 +3.2	76.5	78.4 +1.9





- **Effective**: Improving CLIP by **13.2%** on ImageNet, outperforming SoTA by **8.8%**
- **Efficient**: Being **training-free**, **3.3%** of CLIP forward latency and **2.5x faster** than SoTA
- **Practical**: Being **model-agnostic**, requiring **no task-specific hyperparameter tuning**
- **General:** Applicable for **any VLMs**