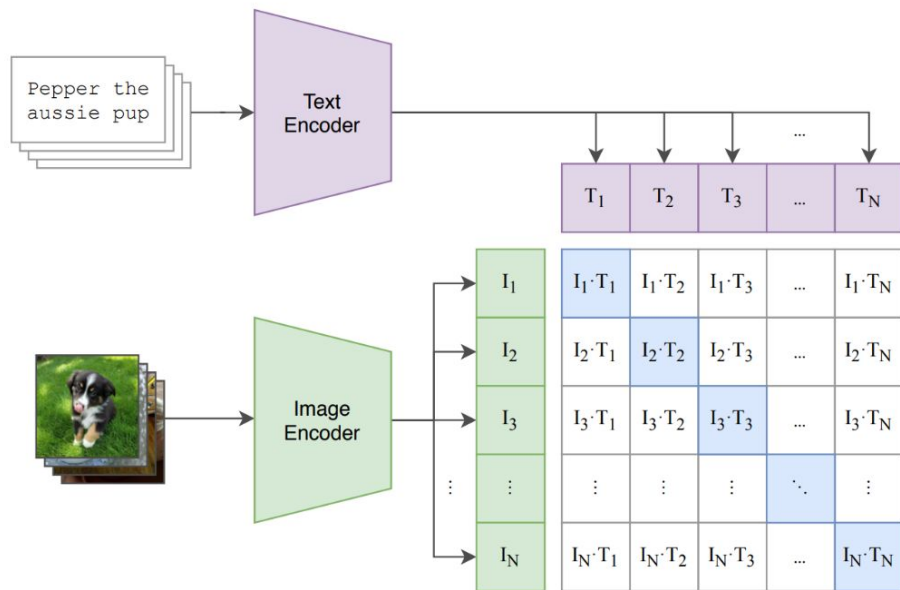


OT3P: Optimal-Transport guided Test-Time Adaptation for Vision-Language Models

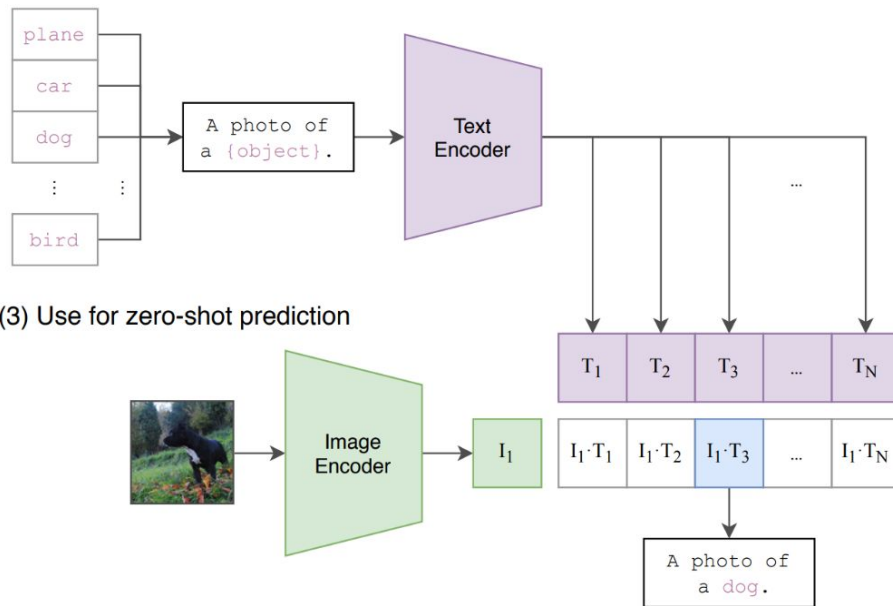
Yunbei Zhang, Janet Wang

Backgrounds

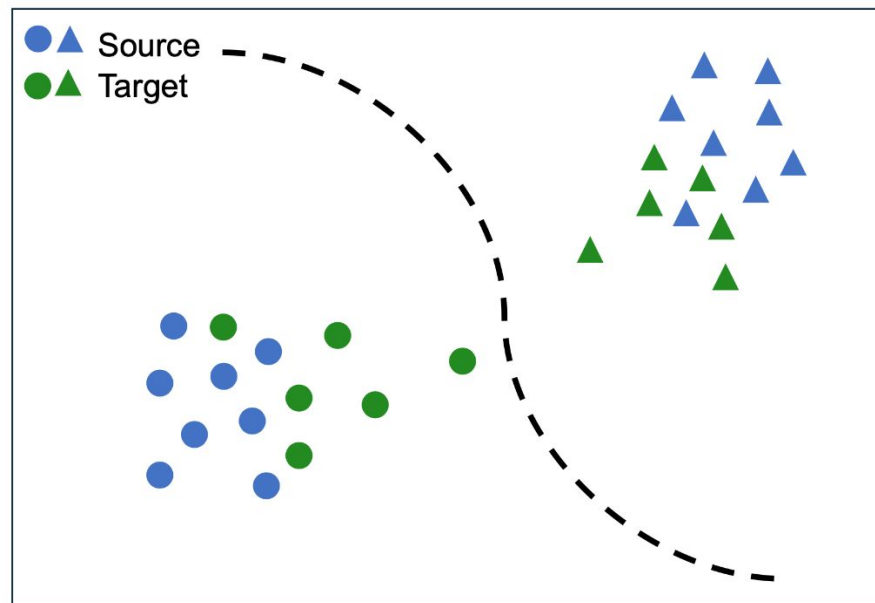
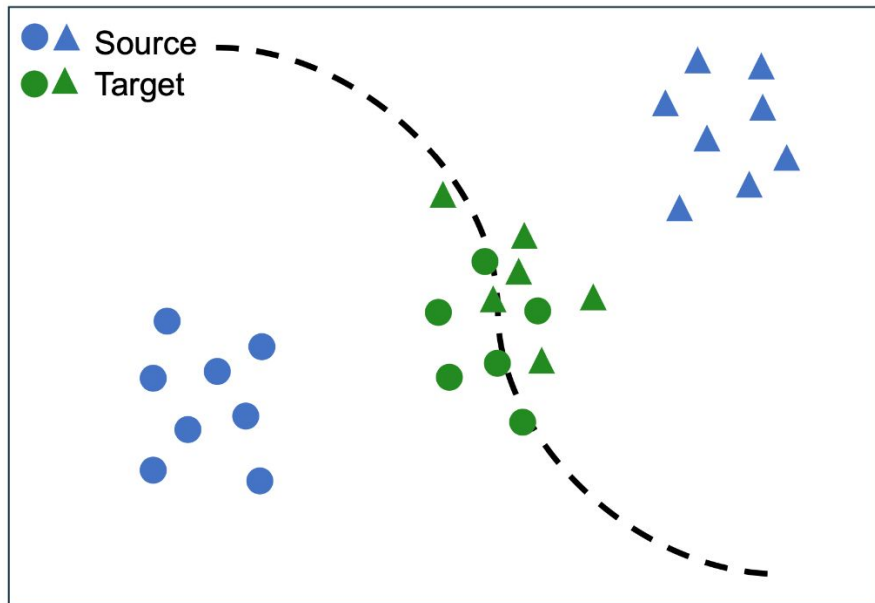
(1) Contrastive pre-training



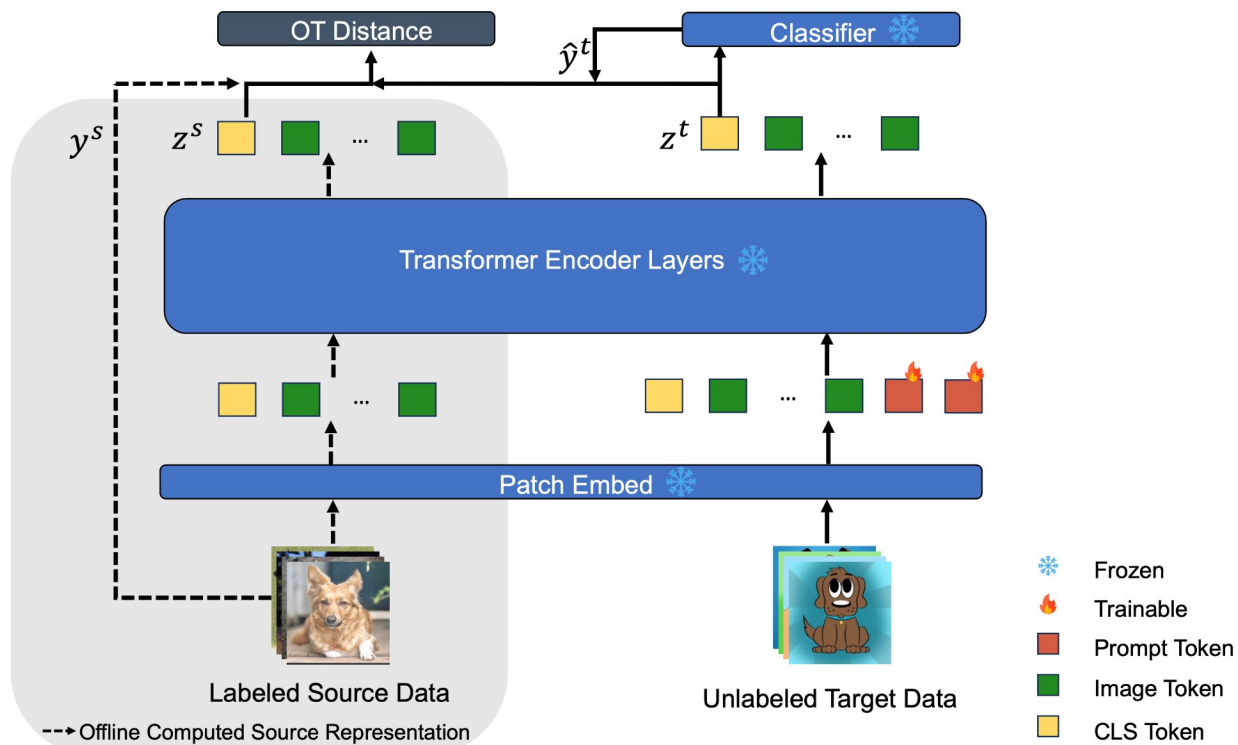
(2) Create dataset classifier from label text



Motivation



Methods



Results for Language Model

Our method improves the ERM performance

Our method outperforms the SOTA

Model	Algorithm	SST	Yelp
BERT	ERM	34.8	37.3
	Tent	41.2	43.5
	OT ³ P	47.2	49.3
RoBERTa	ERM	35.1	36.3
	Tent	46.7	41.5
	OT ³ P	51.3	48.7

Algorithm	ImageNet V2	ImageNet Sketch	ImageNet A	ImageNet R	OOD Avg.
MaPLe	64.07	49.15	50.90	76.98	60.28
MaPLe+TPT	64.87	48.16	58.08	78.12	62.31
OT³P	65.29	50.23	59.37	79.33	63.55