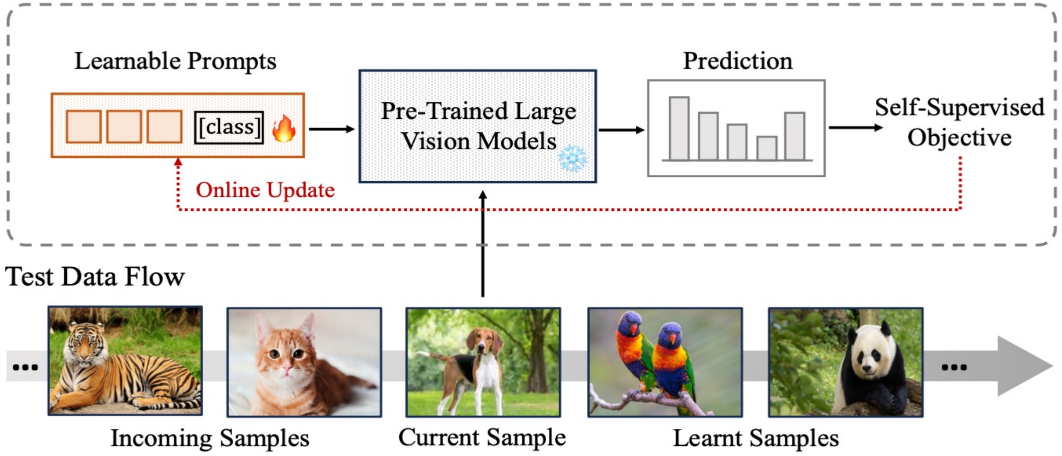
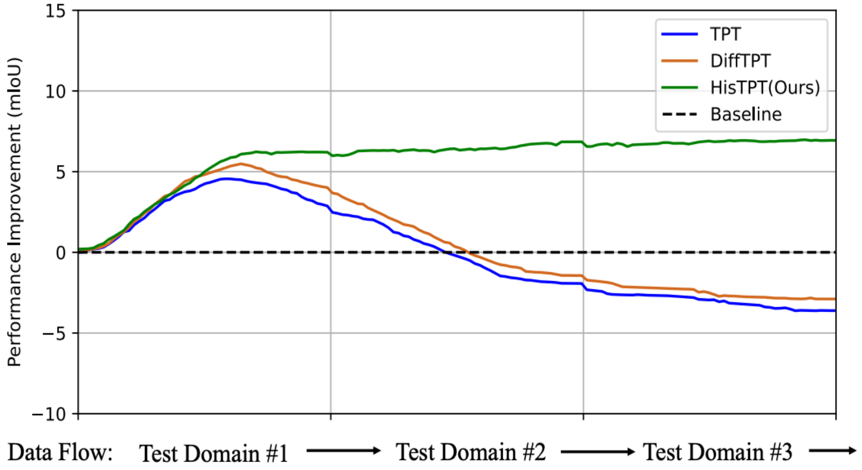
Historical Test-time Prompt Tuning for Vision Foundation Models

Jingyi Zhang[1], Jiaxing Huang[1], Xiaoqin Zhang[2], Ling Shao[3], Shijian Lu[1]

[1] Nanyang Technological University [2] Zhejiang University of Technology [3] General Terminus Technologies

1. Introduction: Motivation

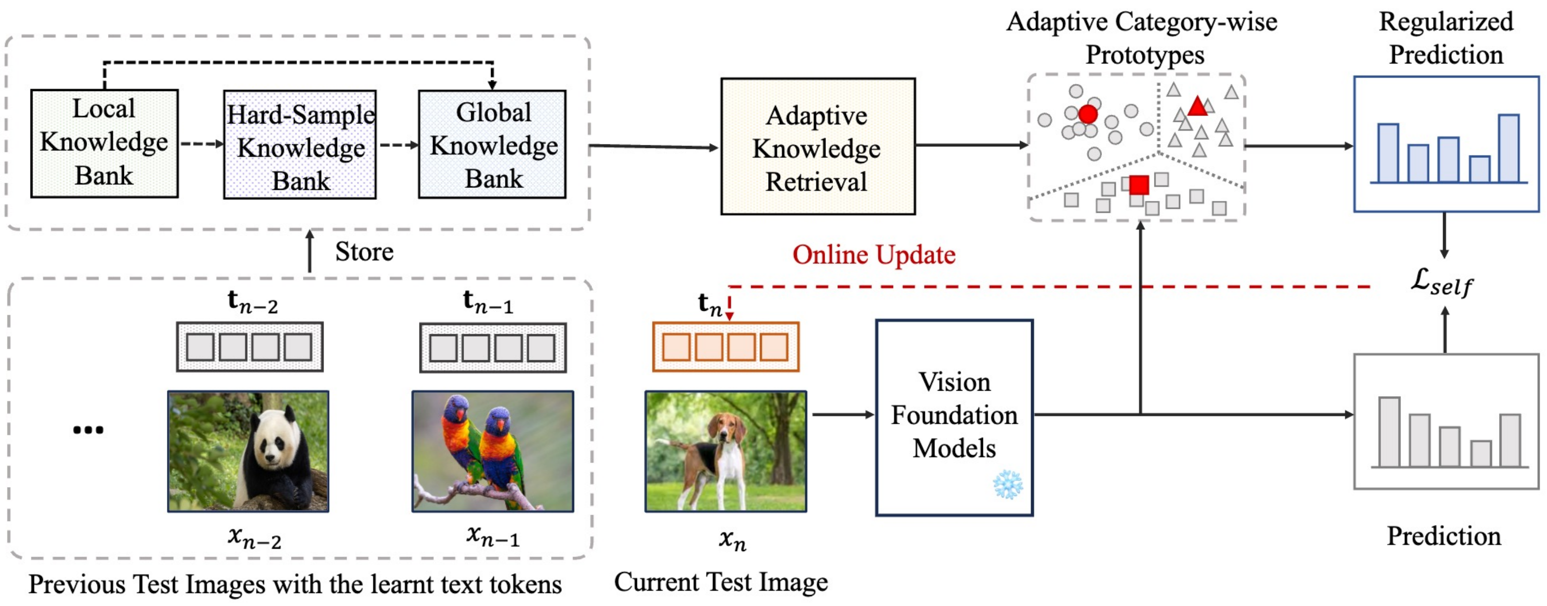




Test-time Prompt Tuning learns and optimizes prompts from a continuous flow of unlabelled test samples during the inference stage.

Existing test-time prompt tuning methods tend to **'forget'** the historical knowledge learnt from previous test samples when the prompts are continuously updated.

2. Method: HisTPT - Overview



Knowledge Bank Construction:
HisTPT features three types of knowledge banks, i.e., *Local knowledge bank, Hard-sample knowledge bank, Global knowledge bank,* for learning and memorizing up-to-date, difficult and representative knowledge from previous test samples.

Adaptive Knowledge Retrieval Mechanism:
HisTPT introduces an Adaptive Knowledge Retrieval Mechanism that enables adaptive retrieval of memorized knowledge for prediction regularization and prompt optimization of each test sample.

3. Method: HisTPT – Quantitative Results

Table 1: Test-time prompt tuning on semantic segmentation over 6 widely adopted datasets. mIoU is reported.

Method	Cityscapes	BDD	Mapillary	ADE	Pascal	ACDC_{Fog}	ACDC_{Night}	ACDC_{Rain}	ACDC_{Snow}	Mean
SEEM-Tiny	39.2	37.4	14.7	14.6	45.1	34.6	20.7	33.1	35.8	30.5
TPT [7]	42.3	38.9	15.4	16.1	46.8	35.2	21.4	34.9	36.5	31.9
TPT [7] + HisTPT	45.1	41.8	17.5	17.6	49.4	37.2	22.9	37.2	37.8	**34.0**
DiffTPT [8]	42.9	39.6	15.8	16.3	47.1	35.7	21.6	35.3	36.6	32.3
DiffTPT [8] + HisTPT	45.4	42.1	16.7	17.9	49.2	47.6	22.7	37.7	38.1	**35.2**
HisTPT	44.7	41.2	17.2	17.3	48.7	36.8	22.1	36.7	37.1	**33.5**
SEEM-Large	49.3	44.6	18.7	15.2	37.1	48.1	32.0	47.4	45.0	37.4
TPT [7]	50.1	45.2	19.1	15.7	40.2	48.7	32.4	47.9	45.7	38.3
TPT [7] + HisTPT	52.1	47.4	21.3	17.1	45.8	52.1	33.4	49.4	48.8	**40.8**
DiffTPT [8]	50.4	45.7	19.3	16.1	41.2	49.1	32.2	48.2	46.3	38.7
DiffTPT [8] + HisTPT	52.4	47.8	21.1	17.4	46.3	52.4	33.6	49.7	49.1	**41.0**
HisTPT	51.9	47.3	20.1	16.9	45.7	51.6	33.1	49.1	48.5	**40.4**

4. Method: HisTPT – Ablation Studies

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Table 4: Ablation study of the proposed HisTPT over Cityscapes semantic segmentation task.

Method	Histrocial Knowledge Banks			Adaptive knowledge retrieval	mIoU
	local knowledge bank	hard-sample knowledge bank	global knowledge bank		
SEEM-Tiny					39.2
	✓				41.1
		✓			40.9
			✓		41.7
	✓	✓			42.2
	✓		✓		42.8
		✓	✓		42.5
	✓	✓	✓		43.6
HisTPT	✓	✓	✓	✓	**44.7**

5. Method: HisTPT – Discussion

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Table 6: Test-time prompt tuning on semantic segmentation across continuously changing test domains. mIoU is reported.

Test Order (→)	Normal	Fog	Night	Rain	Snow
SEEM-Tiny	39.2	34.6	20.7	33.1	35.8
TPT	42.3(+3.1)	34.8(+0.2)	20.1(-0.6)	31.7(-1.4)	30.6(-5.2)
DiffTPT	42.9(+3.7)	35.2(+0.6)	20.3(-0.4)	32.0(-1.1)	31.4(-4.4)
HisTPT	44.7(+5.5)	36.9(+2.3)	23.6(+2.9)	37.3(+4.2)	38.1(+2.3)

(a)

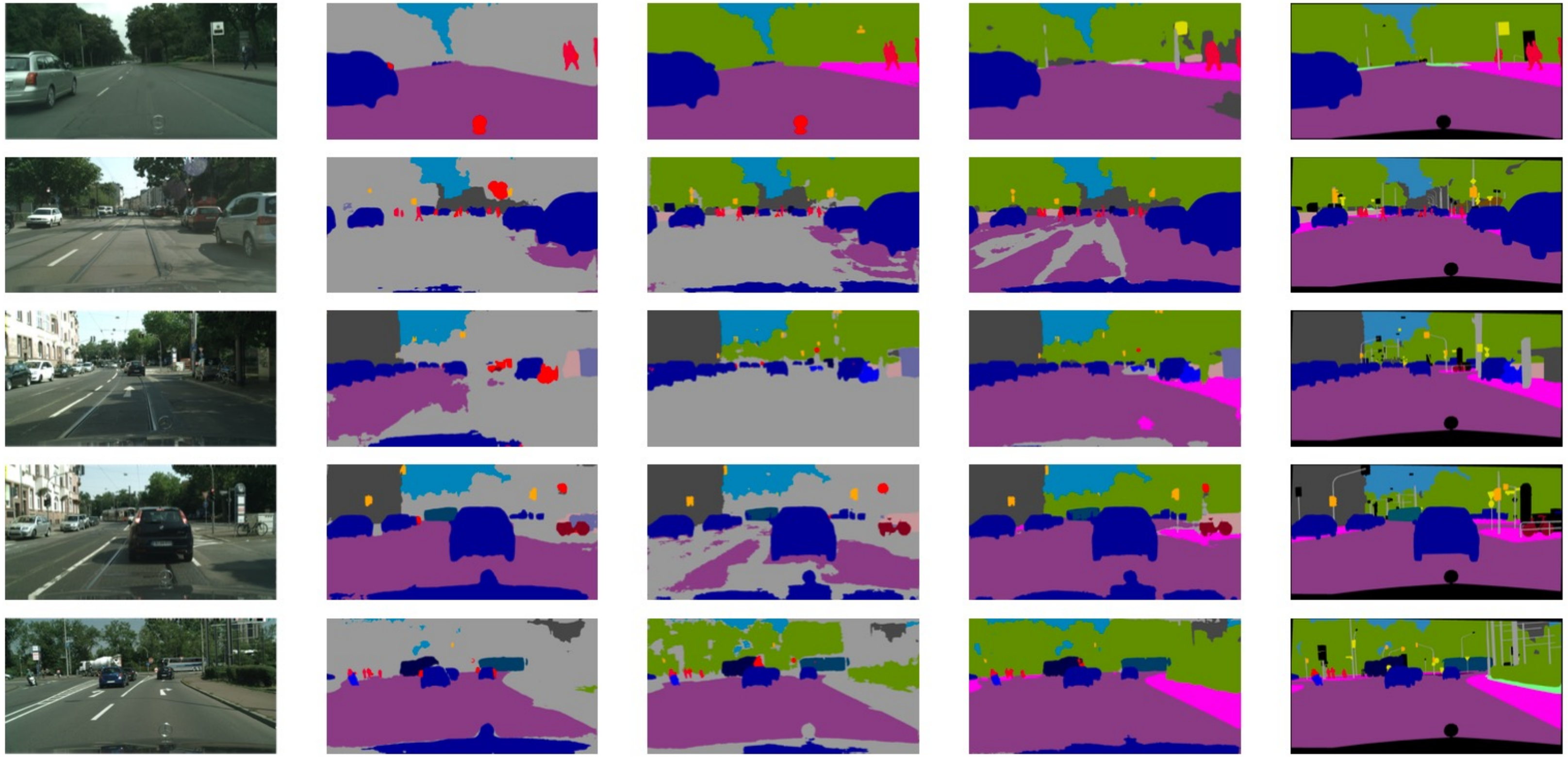
Test Order (→)	Snow	Rain	Night	Fog	Normal
SEEM-Tiny	35.8	33.1	20.7	34.6	39.2
TPT	36.5(+0.7)	34.1(+1.0)	20.1(-0.6)	32.7(-1.9)	35.8(-3.4)
DiffTPT	36.6(+0.8)	34.7(+1.6)	20.5(-0.2)	32.9(-1.7)	36.1(-3.1)
HisTPT	37.1(+1.3)	36.8(+3.7)	22.1(+1.4)	37.0(+2.4)	44.9(+5.7)

(b)

HisTPT improves the performance consistently across different weathers.

6. Method: HisTPT – Qualitative Results

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE



Original Image SEEM-Tiny [3] TPT [7] HisTPT(Ours) Ground Truth

Figure 6: Qualitative comparison of HisTPT with the baseline model (SEEM-Tiny) [3] and TPT [7] over semantic segmentation task on Cityscapes.

Historical Test-time Prompt Tuning for Vision Foundation Models

Thank you!

Jingyi Zhang[1], Jiaxing Huang[1], Xiaoqin Zhang[2], Ling Shao[3], Shijian Lu[1]

[1] Nanyang Technological University [2] Zhejiang University of Technology [3] General Terminus Technologies