



NEURAL INFORMATION
PROCESSING SYSTEMS

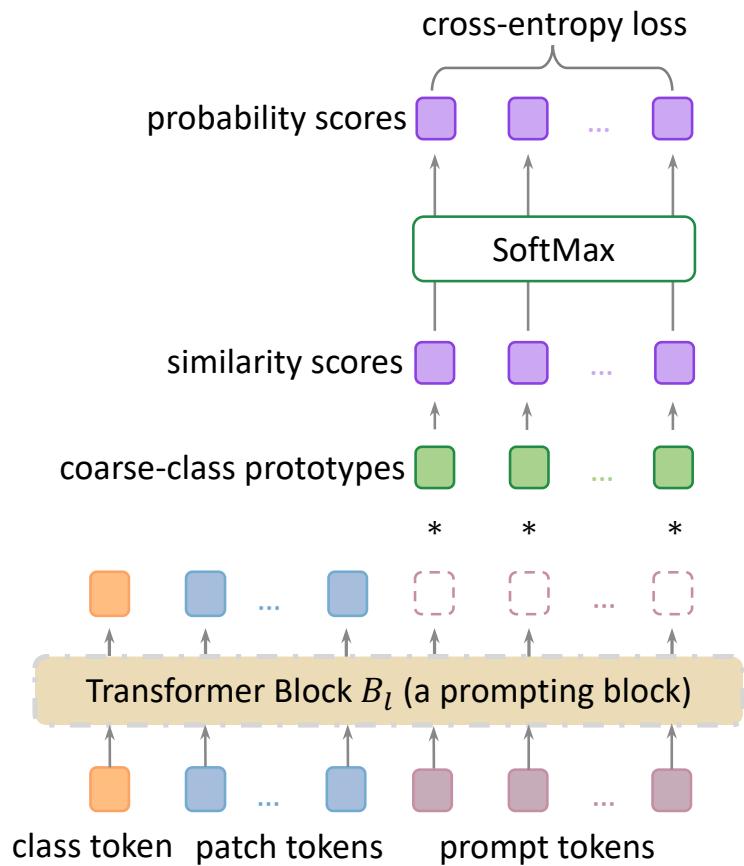
TransHP: Image Classification with Hierarchical Prompting

Speaker: Wenhao Wang

Authors: Wenhao Wang, Yifan Sun, Wei Li, and Yi Yang

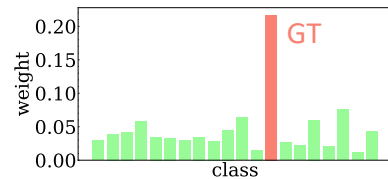
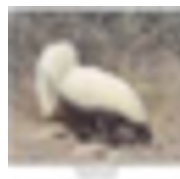
University of Technology Sydney, Baidu Inc., Zhejiang University

A prompting block in TransHP

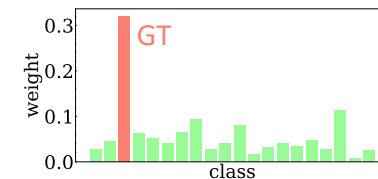
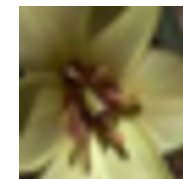


(i)

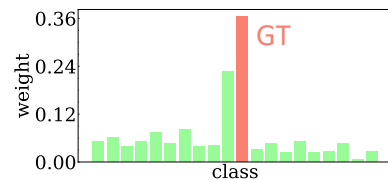
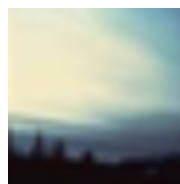
Autonomous prompt selection.



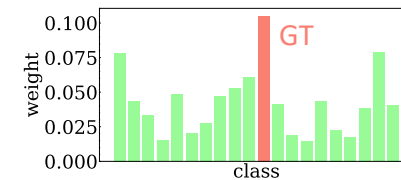
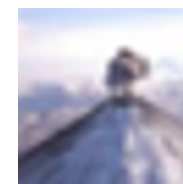
(a)



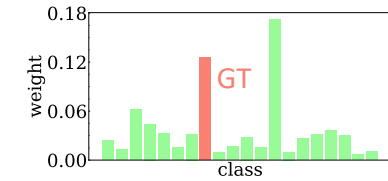
(b)



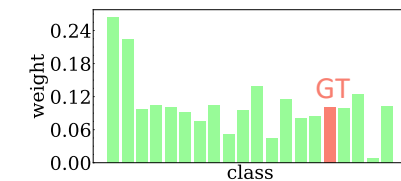
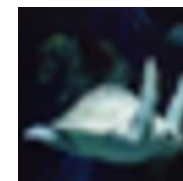
(c)



(d)



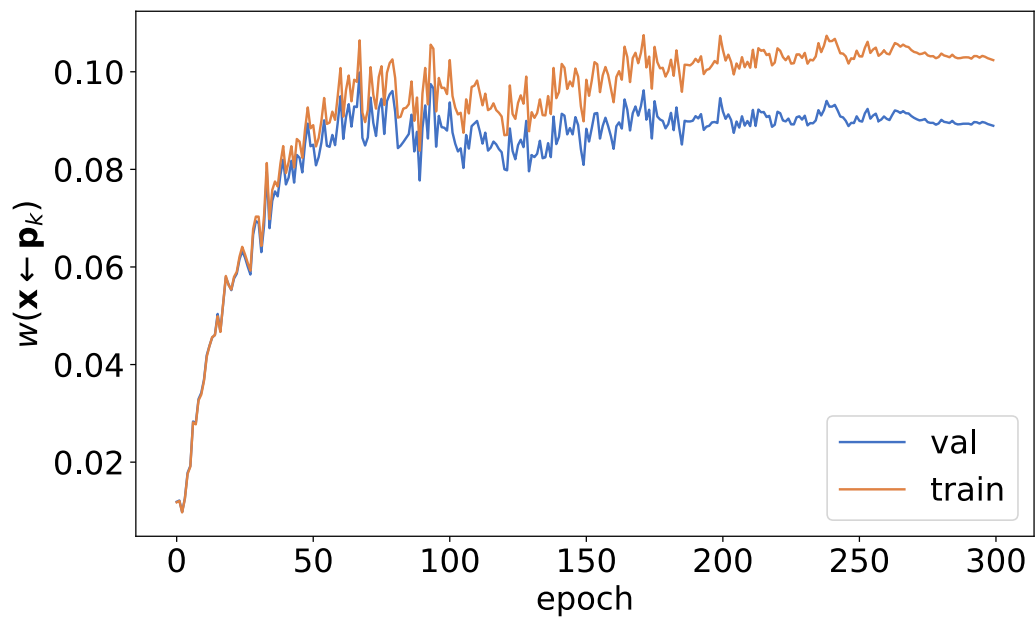
(e)



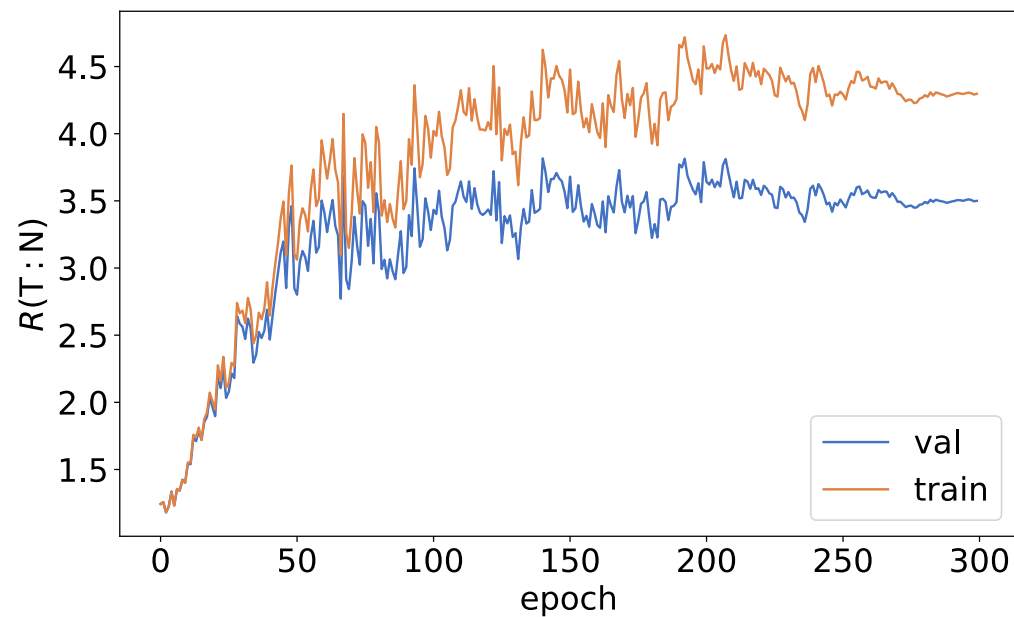
(f)

(ii)

TransHP gradually focuses on the predicted coarse class



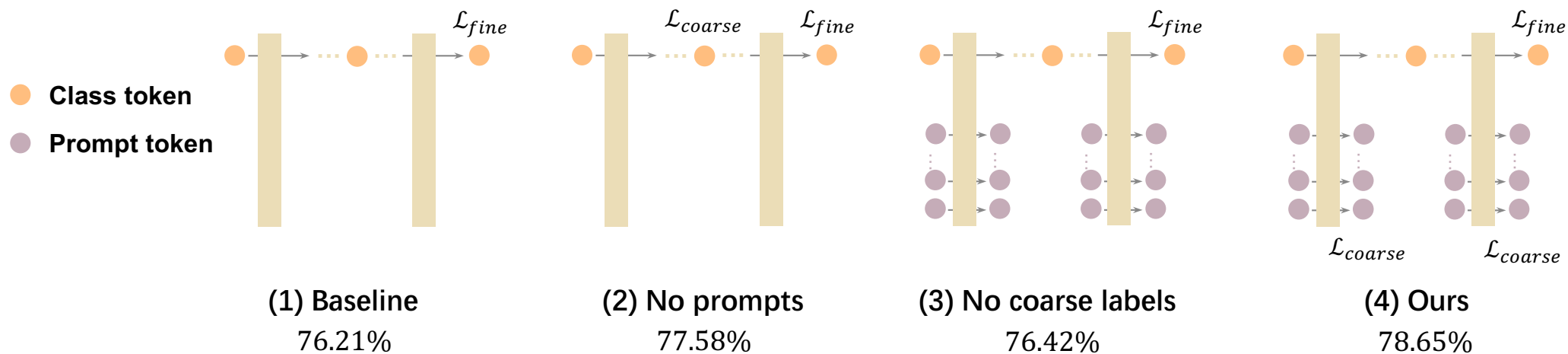
(a)



(b)

TransHP Improves the Accuracy

Improvement on ImageNet and the ablation study.



TransHP gains consistent improvements on more datasets.

Accuracy (%)	iNaturalist-2018	iNaturalist-2019	CIFAR-100	DeepFashion
Baseline (w/o Pre)	51.07	57.33	61.77	83.42
TransHP (w/o Pre)	53.22	59.24	67.09	85.72
Baseline (w Pre)	63.01	69.31	84.98	88.54
TransHP (w Pre)	64.21	71.62	86.85	89.93

TransHP Improves the Accuracy

TransHP improves various backbones.

Accuracy (%)	ViT-B/16	ViT-L/16	DeiT-S	DeiT-B
Baseline	76.68*	76.37*	79.82	81.80
TransHP	79.51	78.80	80.55	82.35

Comparison with state-of-the-art hierarchical classification methods.

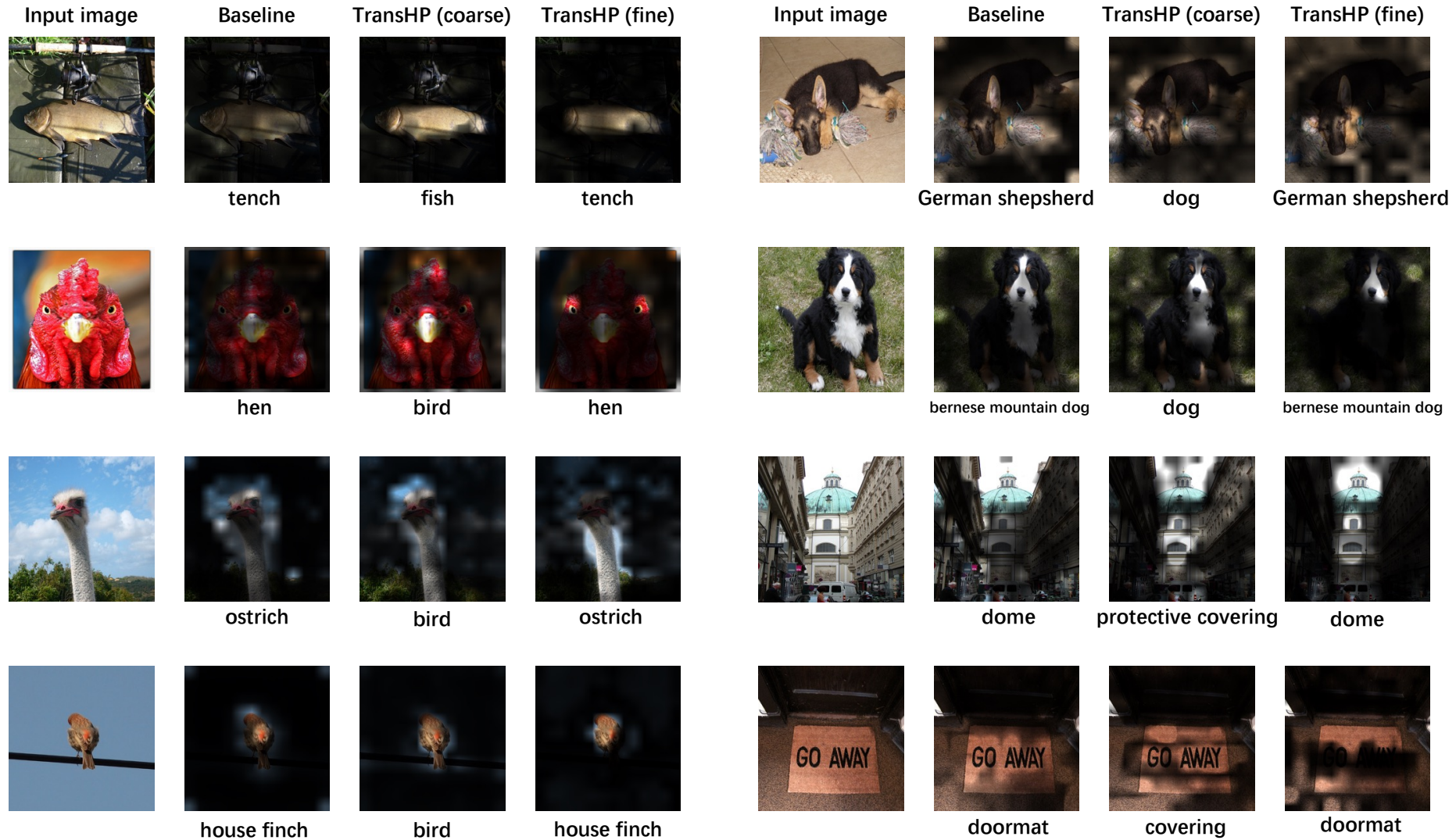
Accuracy (%)	ImageNet	iNat-2018	iNat-2019	CIFAR-100	DeepFashion
Baseline	76.21	63.01	69.31	84.98	88.54
Guided	76.05	63.11	69.66	85.10	88.32
HiMulConE	77.52	63.46	70.87	85.43	88.87
TransHP	78.65	64.21	71.62	86.85	89.93

TransHP under the data-scarce scenario.

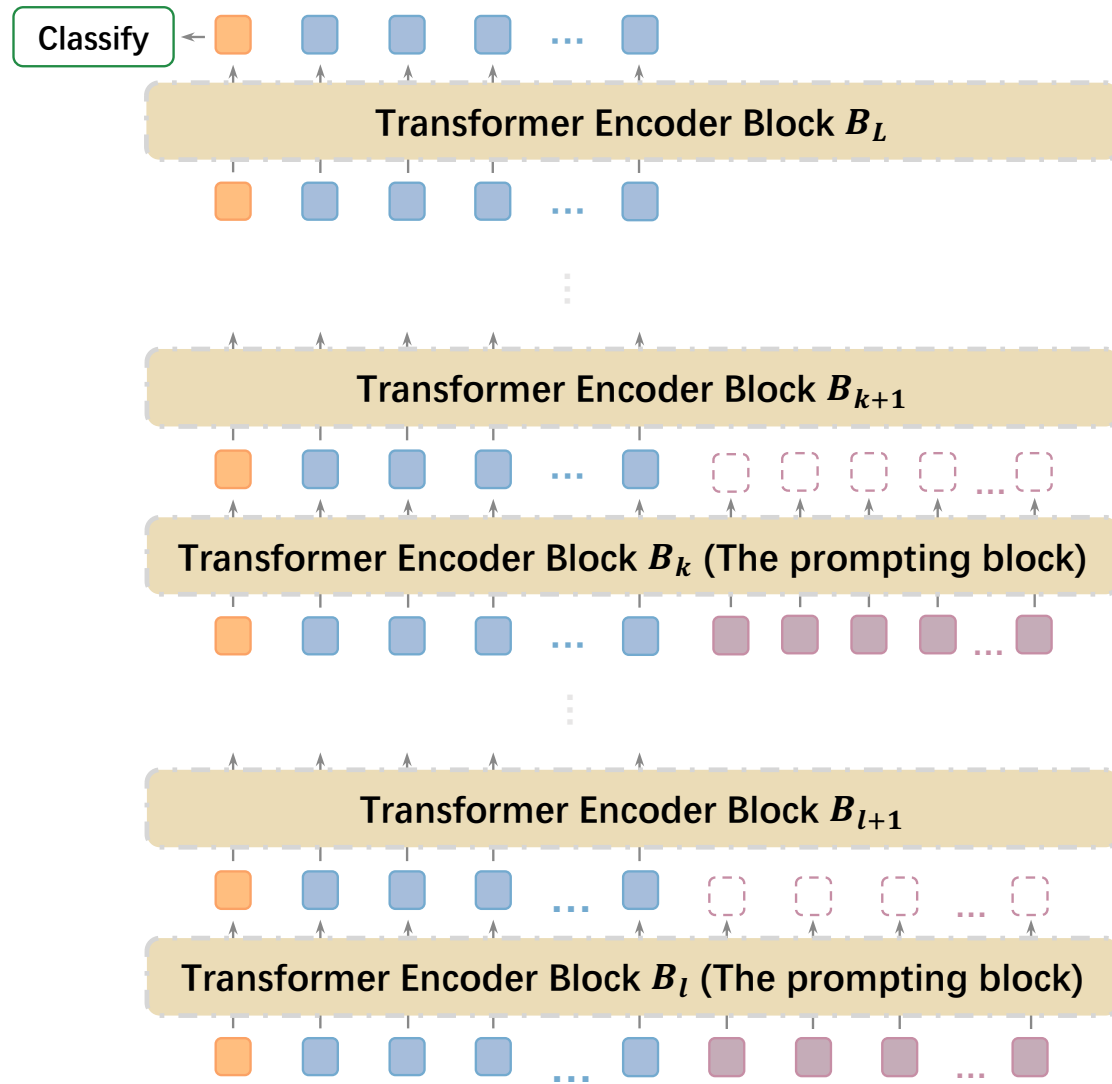
Accuracy (%)	100%	50%	20%	10%
Baseline	76.21	67.87	44.60	25.24
Guided	76.05	67.74	45.02	25.67
HiMulConE	77.52	69.23	48.50	30.76
TransHP	78.65	70.74	53.71	37.93

TransHP Improves Model Explainability

Visualization of the attention map for analyzing the receptive field.



Multiple layers of hierarchy



THANKS FOR LISTENING