

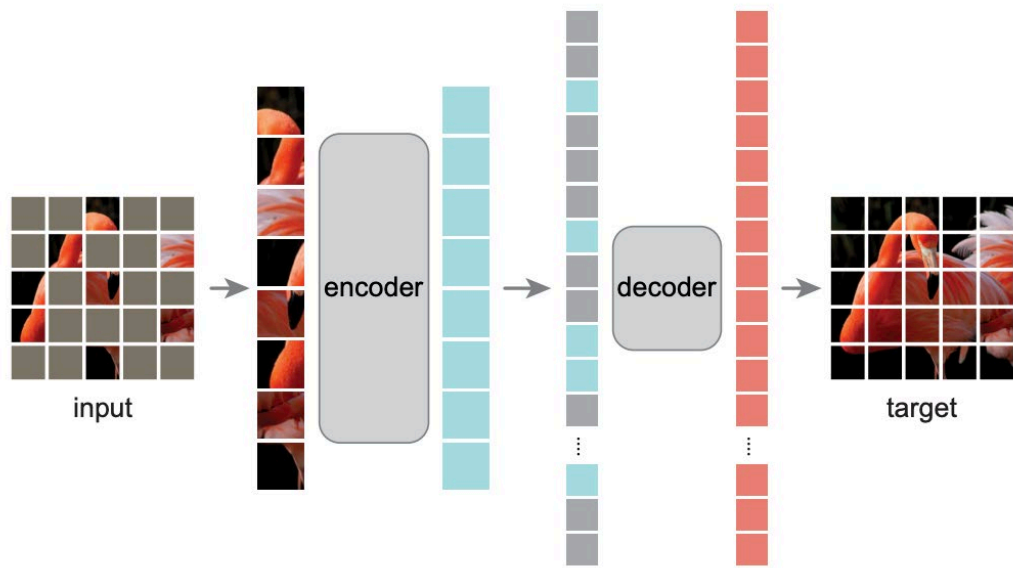


**ICLR**

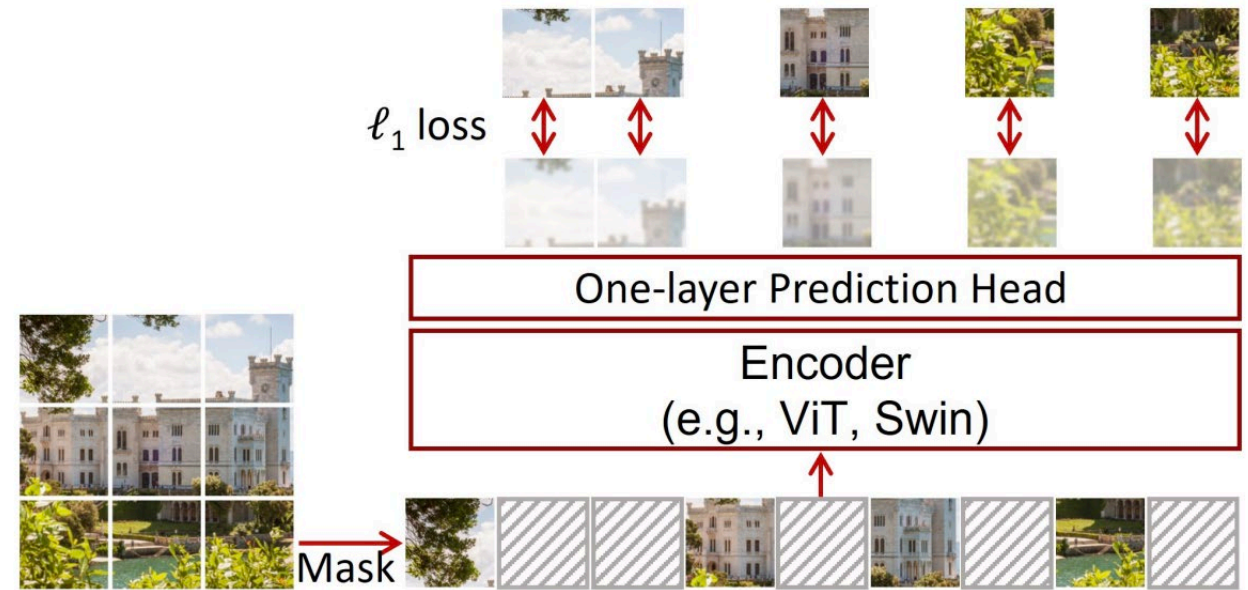


# Progressively Compressed Auto-encoder for Self-supervised Learning

# Masked Image Modeling:



MAE [1]

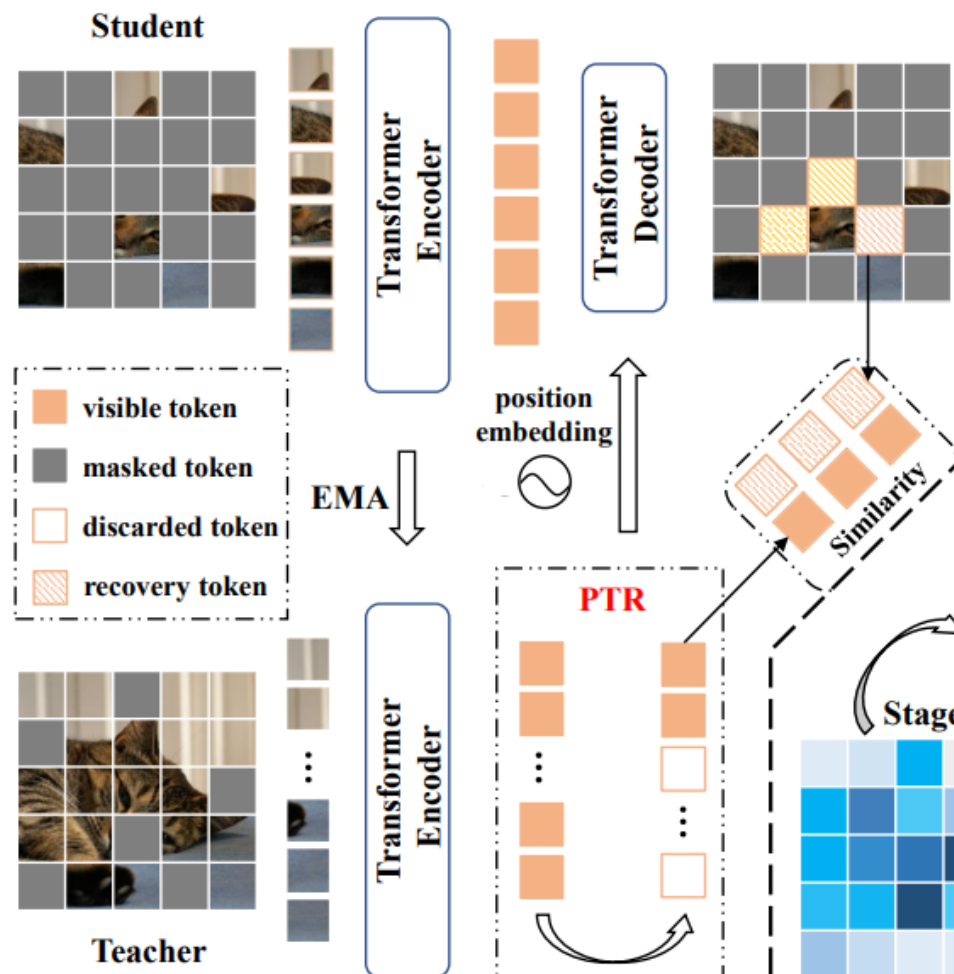


SimMIM [2]

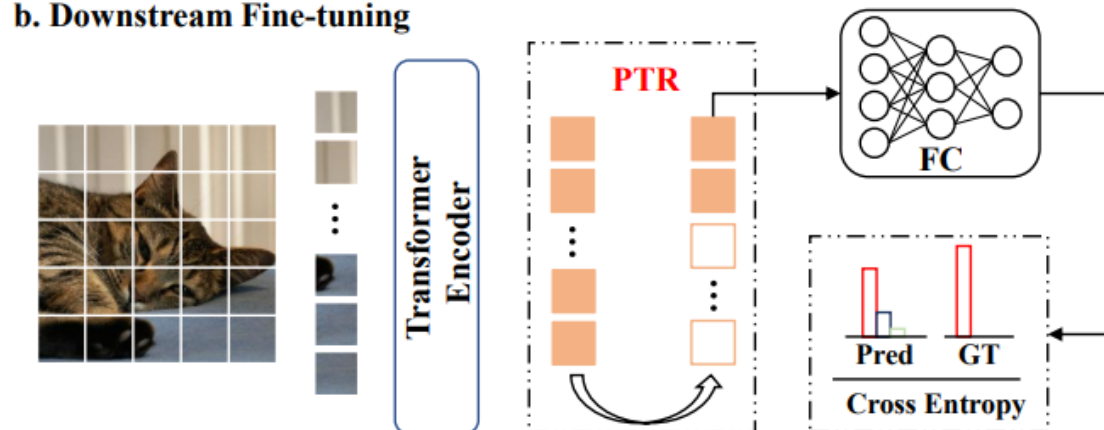
- Recover masked regions from visible ones

# PCAE

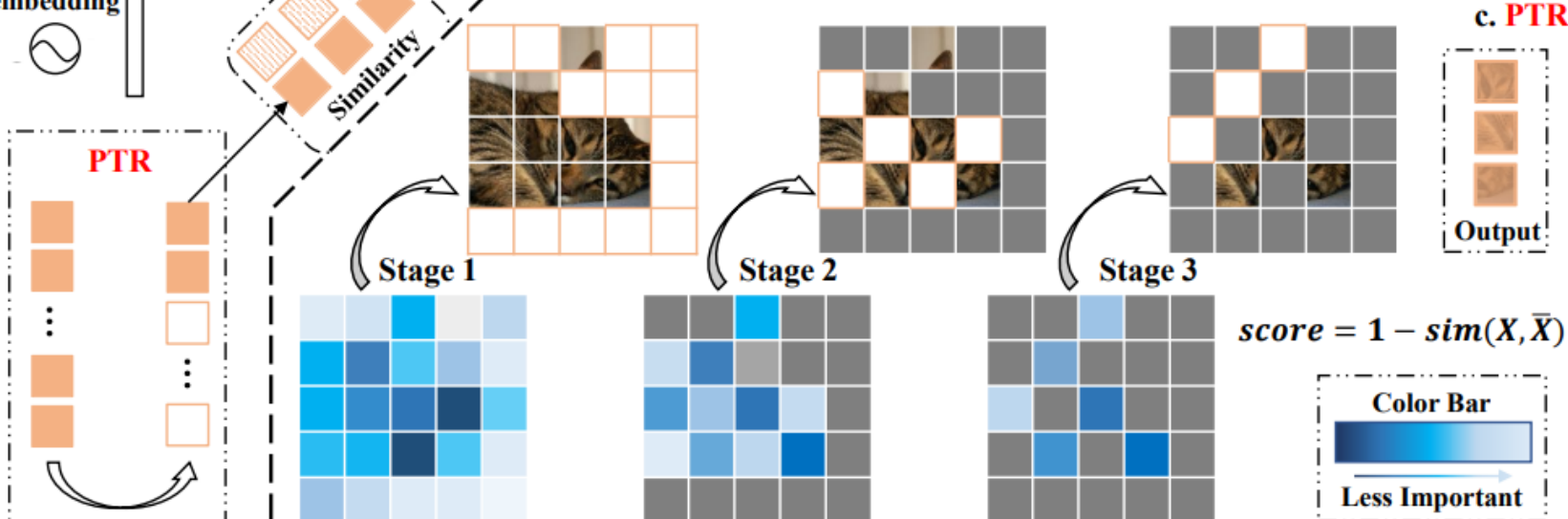
## a. Self-supervised Pretraining



## b. Downstream Fine-tuning



## c. PTR



# Evaluation to PCAE

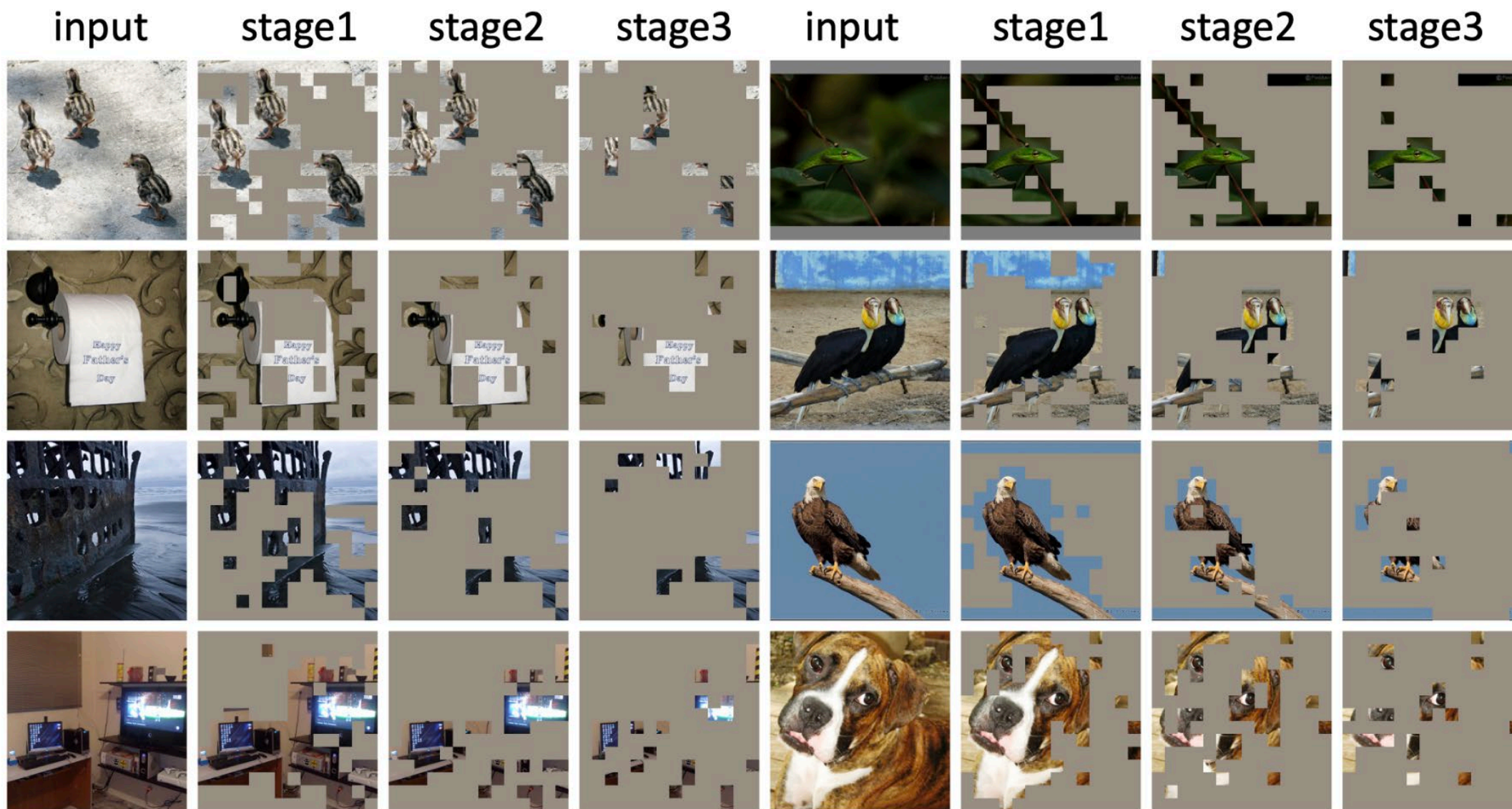
ImageNet 1k fine-tuning

Method	Epochs	Forwards	GPU Days	Accuracy
Train from Scratch	300	1	-	81.8
MoCo v3	300	2	-	83.2
DINO	400	12	122.5	83.3
BEiT <sup>†</sup>	300	1	-	83.0
CAE <sup>†</sup>	300	2	-	83.6
SimMIM	800	1	54.1	<b>83.8</b>
MAE	300	1	15.4	82.9
MAE	1600	1	82.1	83.6
PCAE	300	2	10.1	83.6
PCAE	800	2	26.9	<b>83.9</b>

COCO detection

Methods	Epochs	Object Detection			Instance Segmentation		
		AP <sup>b</sup>	AP <sub>50</sub> <sup>b</sup>	AP <sub>75</sub> <sup>b</sup>	AP <sup>m</sup>	AP <sub>50</sub> <sup>m</sup>	AP <sub>75</sub> <sup>m</sup>
<i>Supervised methods:</i>							
DeiT	300	46.9	68.9	51.0	41.5	65.5	44.4
<i>Contrastive learning methods:</i>							
MoCo v3	300	45.5	67.1	49.4	40.5	63.7	43.4
DINO	400	<b>46.8</b>	68.6	50.9	<b>41.5</b>	65.3	44.5
<i>Masked Image Modeling methods:</i>							
BEiT	300	39.5	60.6	43.0	35.9	57.7	38.5
BEiT	800	42.1	63.3	46.0	37.8	60.1	40.6
MAE	300	45.4	66.4	49.6	40.6	63.4	43.7
MAE	1600	48.4	69.4	53.1	42.6	66.1	45.9
PCAE	300	47.6	68.3	52.4	42.0	65.3	45.6
PCAE	800	<b>48.8</b>	69.6	53.6	<b>43.1</b>	66.9	46.4

# Visualization to PCAE



# Properties of PCAE

Table 4: The performance on ImageNet and COCO as well as the throughput of PCAE with different mask ratios.

Mask ratio	75%	85%	90%
FT Acc.	83.6	83.6	83.5
COCO Det.	47.6	47.7	47.4
Throughput	550.5	739.7	876.1

Table 6: Ablation on drop stage with fixed total drop ratio of 87.5%.

Exp id	drop stage	drop ratio	GPUd.	Acc.
Exp id = 1	1	87.5	3.0	81.5
Exp id = 2	3	50	3.3	<b>82.3</b>
Exp id = 3	6	29	3.5	82.0
Exp id = 4	12	16	3.6	81.5

Table 5: ImageNet classification under different drop ratios.

Ratio	Throughput	MACs	Acc.
0	286	17.57	83.8
0.1	330 ( $\uparrow 15\%$ )	15.34 ( $\downarrow 13\%$ )	83.7 ( $\downarrow 0.1$ )
0.2	385 ( $\uparrow 35\%$ )	13.19 ( $\downarrow 25\%$ )	83.5 ( $\downarrow 0.3$ )
0.3	449 ( $\uparrow 57\%$ )	11.29 ( $\downarrow 36\%$ )	83.2 ( $\downarrow 0.6$ )
0.4	538 ( $\uparrow 88\%$ )	9.53 ( $\downarrow 46\%$ )	82.6 ( $\downarrow 1.2$ )

Table 7: Ablation on the totally drop ratio.

Exp id	drop ratio	total	GPUd.	Acc.
Exp id = 1	37	75	3.6	81.8
Exp id = 2	41.5	80	3.5	<b>82.4</b>
Exp id = 3	47	85	3.4	81.9
Exp id = 4	50	87.5	3.3	82.3
Exp id = 5	53.6	90	3.3	81.7

**Thanks**