//□□□□□□□□□□□□□□□□□□□□□□□□ //*□□□□□□□□□□□□□□□*//

Table 1: **Performance of Token Reduction methods with DeiT-S backbone.** Model performance is measured across varying keep rates, r, denoted in percentage of tokens kept at each reduction stage. Scores exceeding the DeiT baseline are noted in **bold**, measured in Top-1 accuracy for ImageNet & NABirds and mean Average Precision for COCO & NUS-WIDE. The three best performing methods per keep rate are denoted in descending order with red, orange, and yellow, respectively. Similarly, the three worst performing methods are denoted in descending order with light blue, blue, and dark blue. Results with the DeiT-B and DeiT-T backbones are available in the supplementary material.

	ImageNet				NABirds				COCO				NUS-WIDE			
DeiT-S	79.85				80.57				78.11				63.23			
r (%)	25	50	70	90	25	50	70	90	25	50	70	90	25	50	70	90
ℓ_1	70.05	74.47	77.25	79.17	62.90	70.52	77.10	80.08	61.28	69.49	74.31	77.09	54.44	59.60	61.98	62.91
ℓ_2	70.54	74.86	77.41	79.27	64.28	72.09	77.53	80.11	62.23	70.30	74.66	77.19	55.31	60.22	62.07	62.71
ℓ_∞	70.58	74.03	77.48	79.23	63.36	70.19	77.23	79.96	61.50	69.11	74.73	77.27	55.10	59.34	62.11	62.77
Top-K	72.91	77.82	79.22	**79.87**	76.28	80.38	**80.70**	**80.60**	70.14	75.84	77.50	78.09	59.32	61.98	62.69	**63.26**
EViT	74.17	78.08	79.30	**79.87**	76.74	80.28	**80.73**	**80.64**	71.28	75.78	77.50	78.07	59.69	61.89	62.67	**63.25**
DynamicViT	60.32	77.84	79.17	79.79	70.60	**80.62**	**80.77**	**80.84**	39.18	69.02	75.43	77.69	39.20	57.83	61.96	63.16
ATS	72.95	77.86	79.09	79.63	73.46	78.89	80.36	80.55	70.13	75.66	77.23	77.83	60.20	62.35	62.93	63.18
ToMe	-	78.29	79.63	**79.92**	-	74.99	80.05	**80.68**	-	74.99	77.36	77.88	-	61.51	62.50	62.89
K-Medoids	68.94	76.44	78.74	79.73	65.28	76.95	79.75	80.46	66.26	74.15	76.76	77.94	57.78	61.48	62.47	63.12
DPC-KNN	75.01	77.95	78.85	79.54	68.77	74.14	76.70	78.88	72.15	75.70	77.06	77.74	60.78	62.11	62.67	62.93
SiT	74.65	77.16	77.52	77.71	62.82	62.02	60.72	58.50	57.65	57.33	57.11	57.13	57.95	58.84	59.29	59.59
PatchMerger	69.44	74.17	75.80	76.75	47.26	61.34	65.45	68.24	62.24	68.09	70.75	72.12	55.82	59.27	60.46	61.20
Sinkhorn	64.26	64.07	64.02	64.09	48.89	50.19	51.46	51.22	56.93	56.68	56.85	56.65	50.59	50.67	50.63	50.21

//* □□□□□□□□a□b□□□□□□□□□□□□□□□*//
//z□□□□□□□□□□□□□□□□□

Table 2: **Performance of Token Reduction methods.** Measured across varying keep rates, r, and backbone capacities. Scores exceeding the DeiT baseline are noted in **bold**, measured as Top-1 accuracy for ImageNet & NABirds and mean Average Precision for COCO and NUS-WIDE. The three best performing methods per keep rate are denoted in descending order with red, orange, and yellow, respectively. Similarly, the three worst performing methods are denoted in descending order with light blue, blue, and dark blue

(a) Performance comparison of token reduction methods trained with a DeiT-Base backbone.

	ImageNet				NABirds				COCO				NUS-WIDE			
DeiT-B	81.85				83.32				80.93				64.37			
r (%)	25	50	70	90	25	50	70	90	25	50	70	90	25	50	70	90
ℓ_1	71.23	74.96	78.94	81.04	59.79	71.57	78.92	82.42	58.28	69.27	76.23	79.65	53.01	60.10	63.25	64.14
ℓ_2	71.41	75.40	79.07	81.18	61.55	73.24	79.52	82.55	59.69	70.33	76.56	79.75	54.00	60.37	63.29	64.28
ℓ_∞	71.67	74.40	78.95	81.20	59.96	70.51	79.73	82.59	58.48	68.50	76.54	79.89	53.00	59.59	63.12	64.25
Top-K	73.63	78.97	80.91	**82.03**	74.71	82.22	83.20	**83.40**	67.63	76.91	79.95	**80.97**	58.51	62.78	63.92	**64.40**
EViT	75.26	79.22	80.99	**82.00**	74.73	82.00	83.19	**83.33**	68.93	76.92	79.87	80.92	59.00	62.88	63.90	**64.43**
DynamicViT	27.94	74.58	80.68	81.76	49.23	82.30	83.16	83.23	24.88	62.79	76.54	80.64	28.56	55.51	60.73	63.83
ATS	73.89	78.94	80.78	81.57	71.00	80.10	82.58	83.26	68.17	76.38	79.35	80.50	59.49	63.17	64.21	**64.48**
ToMe	-	78.89	81.05	**82.00**	-	73.67	81.59	82.98	-	74.11	78.82	80.48	-	62.38	64.06	64.35
K-Medoids	69.12	76.86	79.98	81.76	57.54	75.29	80.62	82.57	61.79	73.60	77.58	80.32	56.67	62.18	63.53	64.35
DPC-KNN	69.40	75.87	79.06	81.05	58.16	67.36	72.83	78.29	65.99	73.32	77.03	79.76	58.58	61.39	62.96	63.87
SiT	68.39	75.53	76.63	77.26	65.09	70.75	70.36	68.96	54.86	53.27	53.16	52.73	56.12	59.76	60.64	61.08
PatchMerger	58.78	70.63	74.52	76.76	40.38	57.21	62.20	67.06	54.25	66.22	70.97	73.72	51.80	58.83	60.79	62.09
Sinkhorn	63.37	63.33	63.36	63.50	42.89	42.33	41.72	42.86	52.57	52.33	52.21	52.12	47.55	47.41	47.26	47.48

(b) Performance comparison of token reduction methods trained with a DeiT-Tiny backbone.

	ImageNet				NABirds				COCO				NUS-WIDE			
DeiT-T	72.20				74.16				71.09				59.27			
r (%)	25	50	70	90	25	50	70	90	25	50	70	90	25	50	70	90
ℓ_1	58.58	62.27	67.91	71.06	51.82	59.25	68.36	73.47	49.09	58.27	67.03	70.24	44.81	52.64	57.30	58.73
ℓ_2	58.85	62.91	67.91	71.13	53.10	60.87	69.20	73.42	50.46	60.00	67.33	69.98	45.73	53.45	57.34	58.54
ℓ_∞	59.08	61.92	67.79	71.38	52.60	57.70	69.25	73.34	50.08	57.30	67.22	69.89	45.32	51.91	57.41	58.30
Top-K	62.19	68.55	70.96	71.85	62.14	73.19	**74.57**	**74.64**	60.31	67.47	70.20	**71.65**	52.20	57.02	58.60	**59.50**
EViT	64.11	68.69	71.06	71.83	64.13	73.24	**74.49**	**74.53**	61.44	67.62	70.25	**71.63**	53.09	57.26	58.64	**59.49**
DynamicViT	36.93	67.40	70.94	72.14	57.38	72.54	73.97	**74.30**	24.67	61.70	68.83	**71.30**	28.09	49.36	56.79	58.95
ATS	62.63	68.61	70.77	71.71	64.53	71.07	73.71	**74.43**	60.97	67.37	69.88	**71.10**	52.85	57.30	58.55	59.20
ToMe	-	69.72	71.74	72.16	-	66.61	73.65	**74.50**	-	65.66	69.70	**71.16**	-	55.32	57.78	58.98
K-Medoids	57.50	65.82	69.90	71.50	44.62	66.52	72.09	74.04	54.08	64.83	69.09	71.05	49.13	55.92	58.38	59.07
DPC-KNN	64.56	69.68	71.10	71.88	64.23	71.05	73.02	74.05	63.32	68.03	69.55	70.84	55.37	57.33	58.08	58.88
SiT	63.43	67.98	68.99	68.90	36.35	36.65	34.00	35.07	48.01	47.50	46.98	46.48	36.67	38.15	36.98	37.70
PatchMerger	60.38	64.80	66.81	68.09	38.83	54.20	59.94	62.60	52.49	59.69	62.63	64.30	47.56	52.69	54.33	55.37
Sinkhorn	53.61	53.49	53.19	53.51	36.94	35.98	37.29	36.19	50.47	49.52	49.12	49.01	45.77	44.81	44.52	44.20

Table 3: **Selected token reduction method hyperparameters - ImageNet.** We present the selected hyperparameters when searching on ImageNet for each token reduction method.

r (%)	25			50			70			90		
	W-E	B-LR	B-FE	W-E	B-LR	B-FE	W-E	B-LR	B-FE	W-E	B-LR	B-FE
ℓ_1	5	1	0	5	0.01	5	5	0.01	5	5	0.01	5
ℓ_2	5	1	0	5	0.01	5	5	0.01	5	5	0.01	5
ℓ_∞	5	1	0	5	0.01	5	5	0.01	5	5	0.01	5
Top-K	5	1	0	5	0.01	5	5	0.01	5	20	1	0
EViT	5	1	0	5	0.01	5	5	0.01	5	20	1	0
DynamicViT	20	0.01	5	5	0.01	5	20	0.01	5	20	0.01	5
ATS	5	1	0	5	0.01	5	20	0.01	5	5	0.01	5
ToMe	-	-	-	5	0.01	5	5	0.01	5	5	0.01	5
K-Medoids	5	1	0	5	0.01	5	20	0.01	5	20	1	0
DPC-KNN	5	0.01	5	20	0.01	5	5	0.01	5	5	0.01	5
SiT	5	0.01	5	5	0.01	5	5	0.01	5	20	0.01	5
PatchMerger	5	0.01	5	5	0.01	5	5	0.01	5	5	0.01	5
Sinkhorn	5	1	0	5	1	0	5	1	0	5	1	0