

Index	Method	OxfordPets	Flowers102	FGVCAircraft	DTD	EuroSAT	StanfordCars	Food101	SUN397	Caltech101	UCF101	ImageNet	Average
0	Handcrafted prompt	85.77	66.14	17.28	42.32	37.56	55.61	77.31	58.52	86.29	61.46	58.18	
1													
2	[V]x16, pos=end, shots=1	85.89	68.12	9.64	44.39	50.63	55.59	74.32	60.29	87.53	61.92	57.15	
3	[V]x16, pos=end, shots=2	82.64	77.51	18.68	45.15	61.50	58.28	72.49	59.48	87.93	64.09	57.81	
4	[V]x16, pos=end, shots=4	86.70	86.20	21.87	53.49	70.18	62.62	73.33	63.47	89.55	67.03	59.99	
5	[V]x16, pos=end, shots=8	85.32	91.18	26.13	59.97	76.73	68.43	71.82	65.52	90.21	71.94	61.56	
6	[V]x16, pos=end, shots=16	87.01	94.51	31.26	63.58	83.53	73.36	74.67	69.26	91.83	75.71	62.95	
7	[V]x16, pos=mid, shots=1	85.55	69.48	6.58	41.86	53.26	56.02	72.43	59.28	86.52	61.47	57.33	
8	[V]x16, pos=mid, shots=2	80.57	79.40	14.18	45.92	60.68	57.09	71.06	59.55	87.30	65.05	58.42	
9	[V]x16, pos=mid, shots=4	86.06	86.52	22.02	52.72	70.93	61.62	72.64	63.67	88.53	67.06	59.96	
10	[V]x16, pos=mid, shots=8	83.58	91.81	28.18	59.14	77.65	67.32	72.04	65.64	90.33	72.74	62.04	
11	[V]x16, pos=mid, shots=16	86.16	94.80	32.29	63.16	83.55	73.27	74.46	69.12	91.62	75.29	63.08	
12	[V]x16, pos=end, CSC, shots=1	64.65	62.89	14.61	31.54	49.67	34.14	50.63	36.95	71.40	45.04	32.40	
13	[V]x16, pos=end, CSC, shots=2	66.89	78.82	17.70	39.64	58.45	42.39	54.86	45.19	76.56	52.96	38.80	
14	[V]x16, pos=end, CSC, shots=4	74.91	87.10	22.90	49.96	70.45	53.33	60.26	53.31	83.42	60.36	44.93	
15	[V]x16, pos=end, CSC, shots=8	76.12	92.61	29.52	58.33	78.00	63.06	64.52	58.41	86.71	67.42	50.11	
16	[V]x16, pos=end, CSC, shots=16	81.10	95.59	36.93	62.87	84.55	72.20	70.02	63.83	90.58	73.23	54.57	
17	[V]x16, pos=mid, CSC, shots=1	65.07	62.58	14.04	28.92	48.79	34.26	46.92	35.66	69.20	42.72	31.01	
18	[V]x16, pos=mid, CSC, shots=2	66.69	76.79	18.48	39.66	58.99	42.61	53.06	44.77	76.65	52.28	38.19	
19	[V]x16, pos=mid, CSC, shots=4	74.35	87.04	22.92	50.37	70.55	53.68	59.71	53.04	83.31	59.23	44.67	
20	[V]x16, pos=mid, CSC, shots=8	75.96	92.53	29.47	57.84	78.37	63.22	64.55	58.31	86.33	67.97	50.29	
21	[V]x16, pos=mid, CSC, shots=16	80.87	95.52	37.48	62.57	84.87	72.32	70.04	63.76	90.57	73.35	54.65	
22	linear probe CLIP, shots=1	30.14	58.07	12.89	29.59	51.00	24.64	30.13	32.80	70.62	41.43	22.07	
23	linear probe CLIP, shots=2	43.47	73.35	17.85	39.48	61.58	36.53	42.79	44.44	78.72	53.55	31.95	
24	linear probe CLIP, shots=4	56.35	84.80	23.57	50.06	68.27	48.42	55.15	54.59	84.34	62.23	41.29	
25	linear probe CLIP, shots=8	65.94	92.00	29.55	56.56	76.93	60.82	63.82	62.17	87.78	69.64	49.55	
26	linear probe CLIP, shots=16	76.42	94.95	36.39	63.97	82.76	70.08	70.17	67.15	90.63	73.72	55.87	
27													
28													
29													
30		oxford_pets	oxford_flowers	fgvc_aircraft	scribable_textur	eurosat	stanford_cars	food101	sun397	caltech101	ucf101	imagenet	
31	How many context words to use?	OxfordPets	Flowers102	FGVCAircraft	DTD	EuroSAT	StanfordCars	Food101	SUN397	Caltech101	UCF101	ImageNet	
32	[V]x4, pos=end, shots=16	87.12	93.18	28.47	62.92	83.08	69.36	76.56	69.19	91.44	74.51	63.33	72.65
33	[V]x8, pos=end, shots=16	86.75	94.29	30.02	62.98	83.41	71.33	75.58	69.57	92.02	75.99	63.27	73.20
34	[V]x16, pos=end, shots=16	87.01	94.51	31.26	63.58	83.53	73.36	74.67	69.26	91.83	75.71	62.95	73.42
35	[V]x4, pos=mid, shots=16	86.73	92.83	28.45	63.02	81.72	68.93	76.62	68.91	91.71	75.66	63.24	72.53
36	[V]x8, pos=mid, shots=16	86.49	94.29	31.11	63.67	84.00	71.41	75.67	69.10	91.79	74.91	63.22	73.24
37	[V]x16, pos=mid, shots=16	86.16	94.80	32.29	63.16	83.55	73.27	74.46	69.12	91.62	75.29	63.08	73.35
38													
39	Context initialization												
40	a photo of a (CLASS)	86.37	92.53	28.16	63.24	83.32	69.36	77.12	69.05	91.48	75.36	63.11	72.65

	Source	Target			
	ImageNet	ImageNetV2	ImageNet-Sketch	ImageNet-A	ImageNet-R
RN50					
Zero-shot CLIP, a photo of a {class}	58.18	51.34	33.32	21.65	56.00
Ours, [V]x16, pos=end, shots=16	62.95	55.11	32.74	22.12	54.96
Ours, [V]x8, pos=end, shots=16	63.27	55.76	33.44	22.38	55.62
Ours, [V]x4, pos=end, shots=16	63.33	55.40	34.67	23.06	56.60
linear probe CLIP, 16 shots	55.87	45.97	19.07	12.74	34.86
RN101					
Zero-Shot CLIP	61.62	54.81	38.71	28.05	64.38
Ours, [V]x16, pos=end, shots=16	66.60	58.66	39.08	28.89	63.00
Ours, [V]x4, pos=end, shots=16	65.98	58.60	40.40	29.60	64.98
linear probe CLIP, 16 shots	59.75	50.05	26.80	19.44	47.19
ViT-B/32					
Zero-Shot CLIP	62.05	54.79	40.82	29.57	65.99
Ours, [V]x16, pos=end, shots=16	66.85	58.08	40.44	30.62	64.45
Ours, [V]x4, pos=end, shots=16	66.34	58.24	41.48	31.34	65.78
linear probe CLIP, 16 shots	59.58	49.73	28.06	19.67	47.20
ViT-B/16					
Zero-Shot CLIP	66.73	60.83	46.15	47.77	73.96
Ours, [V]x16, pos=end, shots=16	71.92	64.18	46.71	48.41	74.32
Ours, [V]x4, pos=end, shots=16	71.73	64.56	47.89	49.93	75.14
linear probe CLIP, 16 shots	65.85	56.26	34.77	35.68	58.43

	RN50	RN101	ViT-B/32	ViT-B/16
Prompt engineering	58.18	61.26	62.05	66.73
Prompt ensembling	60.41	62.54	63.71	68.74
CoOp	62.95	66.60	66.85	71.92

	classes	train	val	test	task
oxford_pets	37	2,944	736	3,669	fine-grained
oxford_flowers	102	4,093	1,633	2,463	fine-grained
fgvc_aircraft	100	3,334	3,333	3,333	fine-grained
describable textures	47	2,820	1,128	1,692	textures
eurosat	10	13,500	5,400	8,100	satellite images
stanford_cars	196	6,509	1,635	8,041	fine-grained
food101	101	50,500	20,200	30,300	food
sun397	397	15,880	3,970	19,850	scene
caltech101	100	4,128	1,649	2,465	object
ucf101	101	7,639	1,898	3,783	action
imagenet	1,000	1.28M	N/A	50,000	object
imagenetv2	1,000	N/A	N/A	10,000	
imagenet_sketch	1,000	N/A	N/A	50,889	
imagenet_a	200	N/A	N/A	7,500	
imagenet_r	200	N/A	N/A	30,000	
objectnet	113	N/A	N/A	18574	