

dataset	our results	our stddev	vpt paper
cifar-100	71.13	6.23	<b>78.8</b>
caltech101	87.65	2.94	<b>90.8</b>
dtd	<b>66.73</b>	0.96	65.8
flowers102	94.97	1.67	<b>98.0</b>
pets	86.94	0.40	<b>88.3</b>
svhn	70.72	<u>13.02</u>	<b>78.1</b>
sun397	43.87	2.01	<b>49.6</b>
patch camelyon	78.04	2.40	<b>81.8</b>
eurosat	95.96	0.28	<b>96.1</b>
resisc45	81.68	1.75	<b>83.4</b>
retinopathy	<b>69.19</b>	3.16	68.4
clevr/count	<u>55.03</u>	8.83	<b>68.5</b>
clevr/distance	58.24	1.32	<b>60.0</b>
dmlab	41.01	1.59	<b>46.5</b>
kitti/distance	65.86	7.47	<b>72.8</b>
dsprite/location (*)	<b>76.87</b>	t.b.d.	73.6
dsprites/orientation (**)			47.9
smallnorb/azimuth	<u>17.07</u>	2.67	<b>32.9</b>
smallnorb/elevation	<u>30.86</u>	4.00	<b>37.8</b>

(\*) did not complete 5 runs yet

(\*\*) did not complete any runs

**Bold** indicates which result was higher

Underlined indicates surprising results