

Neuron Id	Label(s)	Images	Coverage	Target%	Non-Target%
0	Soil and Island	80	0.928	<b>86.250</b>	87.211
7	Mountain and Sea_water	80	0.863	78.750	33.413
11	Garbage_can and Wastebin	80	1.000	33.750	23.221
12	Mount and Grass	80	1.000	67.500	54.975
15	Building and Building	80	1.000	6.250	8.3894
18	Sea	80	0.812	<b>98.750</b>	93.125
20	Sash and Mattress	80	1.000	<b>92.500</b>	42.836
21	Table	80	1.000	57.500	67.956
22	Mountain and Slope	80	0.970	<b>82.500</b>	49.326
23	Tureen and Tureen	80	1.000	77.500	58.221
25	Dress and Apparel	80	1.000	20.000	38.100
26	Trees and Grass	80	1.000	<b>93.750</b>	49.134
35	Spot and Picture	80	0.982	<b>82.500</b>	74.086
39	Showcase and Card	80	0.903	97.500	83.028
41	Chimney and Bike	80	0.978	56.250	52.067
47	Sea	80	0.961	<b>87.500</b>	63.245
54	Bed and Drawer	80	0.978	<b>100.000</b>	83.725
58	Door and Door	80	0.791	38.750	36.250
60	Sea	80	0.857	72.500	29.639
63	Armchair and Coffee_table	80	0.950	<b>92.500</b>	53.076

Table 10: Representative data from dataset 3, solution 1, showing evaluation results. The table lists neurons, their assigned labels, and the percentage of target and non-target images that activated each neuron. For example, Neuron 26, associated with the label “Trees and Grass”, was activated by 93.75% of target images and 49.134% of non-target images. Neurons with a target activation percentage greater than or equal to 80% (highlighted in bold) were and selected for further hypothesis validation. A total of 19 confirmed labels were identified for this purpose.

Neuron Id	Label(s)	Images	Activations(%)		Mean		Median		z-score	p-value
			targ	non-t	targ	non-t	targ	non-t		
0	Soil and Island	20	95	86.15	2.14	1.33	1.97	1.54	1.57	0.11460241
13	Sash and Mattress	20	90	54.03	0.81	0.07	1.22	0.50	3.64	<b>0.000132412</b>
16	Windowpane	20	80	85.48	0.87	1.62	0.99	1.76	-2.51	<b>0.011627039</b>
18	Sea	20	95	92.50	2.93	2.65	3.09	2.68	0.79	0.428070095
19	Dress and Apparel	20	90	68.75	1.03	0.90	1.07	1.20	0.39	0.692150265
20	Sash and Mattress	20	85	42.69	2.35	0.00	2.42	0.54	4.97	<b>3.85531E-08</b>
22	Mountain and Slope	20	85	47.98	1.72	0.00	1.81	0.52	4.53	<b>1.11043E-06</b>
24	Mountain	20	100	26.05	2.14	0.00	2.19	0.22	7.23	<b>3.37924E-20</b>
26	Trees and Grass	20	75	45.86	1.72	0.00	1.51	0.65	3.17	<b>0.000549313</b>
35	Spot and Picture	20	90	73.65	2.29	0.79	2.31	1.08	4.12	<b>3.21312E-05</b>
37	Brand_name and Signboard	20	90	80.86	1.61	0.92	1.47	1.11	1.93	0.05268508
39	Showcase and Card	20	100	82.01	2.25	1.34	2.43	1.51	3.21	<b>0.001275969</b>
45	Dress and Apparel	20	95	57.21	1.20	0.21	1.16	0.61	3.63	<b>0.000155636</b>
47	Sea	20	90	62.50	1.39	0.41	1.44	0.71	3.64	<b>0.000183727</b>
48	Mount and Mountain	20	95	70.57	1.48	0.63	1.59	0.92	3.33	<b>0.000749184</b>
49	Sea	20	95	53.65	1.97	0.15	1.95	0.74	4.17	<b>1.16734E-05</b>
52	Drapery and Mantle	20	100	65.09	1.67	0.61	1.83	1.02	3.39	<b>0.000521596</b>
54	Bed and Drawer	20	100	83.75	3.99	1.82	3.95	1.99	4.81	<b>1.41834E-06</b>
63	Armchair and Coffee_table	20	85	53.55	1.23	0.14	1.47	0.69	3.26	<b>0.000594428</b>

Table 11: Verification results for dataset 3, solution 1. For example, Neuron 54, associated with the label “Bed and Drawer”, was activated by 100% of target images and 83.75% of non-target images, with a mean activation of 3.99 for target images and 1.82 for non-target images. The Mann-Whitney U test yielded a z-score of 4.81 and a p-value < 0.05, confirming the label hypothesis for this neuron. In total, the null hypothesis was rejected for 15 out of 19 neurons, indicating that their activations were significantly different for target and non-target images, thus confirming the corresponding label hypotheses.