

Atari ゲームに対する Transformer ベース強化学習のロバスト性検証 補足資料 (Supplementary Material)

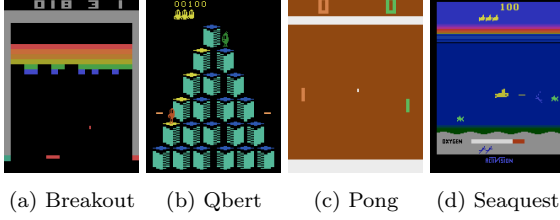


図 1: 使用する 4 種類のゲームタスク

1. ゲーム環境

図 1 に使用した 4 種類のゲーム環境を示す.

2. ノイズで訓練した実験結果

訓練に使用するデータセットは 100% で学習する.

- 表 2～表 26 に各ゲームにおける各 seed 値の平均と標準偏差を示す.
- 表 27～表 30 に各ゲームにおける 3seed 値の平均と標準偏差を示す.
- 図 2～図 5 に各ゲームにおける 3seed 値のスコア分布を示す.
- 図 6～図 9 に各ゲームにおける 3seed 値の平均とヒートマップを示す.

3. ノイズ比率を変化させて訓練した実験結果

訓練に使用するデータセットは 0, 20, 40, 60, 80, 100% の比率で学習する.

- 表 31～表 46 に各ゲームにおける平均と標準偏差を示す. (seed 値は 123 で固定)
- 図 10～図 25 に各ゲームにおけるスコア分布を示す. (seed 値は 123 で固定)
- 図 26～図 41 に平均とヒートマップを示す. (seed 値は 123 で固定)
- 図 42～図 57 にノイズ比率の変化に対する平均報酬を示す. (seed 値は 123 で固定)

4. 訓練とテストの観測画像

- 図 58～図 61 に訓練に使用した観測画像を示す.
- 図 62～図 65 にテスト時の観測画像を示す.

5. 付録

- 表 1 に Decision Transformer のハイパーパラメータを示す.

表 1: Atari ゲームに対する DT のハイパーパラメータ

Hyperparameter	Value
Batch size	128
Context length K	50 Pong 30 Breakout 30 Qbert 30 Seaquest
Return-to-go conditioning	90 Breakout 14000 Qbert 20 Pong 1150 Seaquest
Max epochs	5
Replay buffer capacity	500000

表 2: Clean 訓練, Clean テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	57.28 \pm 45.52	36.7 \pm 22.8	50.04 \pm 48.91	48.00 \pm 41.65
Qbert	1806.0 \pm 2751.6	2655.5 \pm 2734.2	2155.0 \pm 2760.9	2205.5 \pm 2770.9
Pong	12.6 \pm 8.49	13.36 \pm 8.337	-20.48 \pm 0.6997	1.826 \pm 17.21
Seaquest	868.8 \pm 366.0	805.2 \pm 401.9	863.2 \pm 491.3	845.7 \pm 424.0

表 3: Clean 訓練, Gaussian テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	2.28 \pm 3.56	2.60 \pm 3.66	2.58 \pm 2.13	2.486 \pm 3.199
Qbert	358.5 \pm 219.7	320.0 \pm 231.4	360.0 \pm 181.2	346.16 \pm 212.71
Pong	-20.38 \pm 0.7717	-19.54 \pm 1.22	-20.52 \pm 0.7547	-20.14 \pm 1.035
Seaquest	553.2 \pm 261.8	535.2 \pm 233.1	418.0 \pm 222.0	502.1 \pm 246.9

表 4: Clean 訓練, Shot テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	39.74 \pm 40.91	44.02 \pm 37.71	33.2 \pm 17.8	38.98 \pm 34.03
Qbert	932.5 \pm 1862.1	1541.0 \pm 2052.6	826.5 \pm 1228.4	1100.0 \pm 1778.3
Pong	-20.12 \pm 1.012	-19.36 \pm 1.546	-20.3 \pm 0.806	-19.92 \pm 1.233
Seaquest	754.8 \pm 380.7	624.0 \pm 275.3	656.4 \pm 329.1	678.4 \pm 335.8

表 5: Clean 訓練, Impulse テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	1.9 \pm 2.7	1.52 \pm 2.03	1.74 \pm 2.55	1.72 \pm 2.47
Qbert	298.0 \pm 142.0	104.5 \pm 79.79	252.0 \pm 154.5	218.16 \pm 153.70
Pong	-20.74 \pm 0.5589	-20.54 \pm 0.7539	-20.42 \pm 0.7769	-20.56 \pm 0.7156
Seaquest	406.8 \pm 142.1	408.0 \pm 183.1	369.2 \pm 188.8	394.6 \pm 173.5

表 6: Clean 訓練, Speckle テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	38.76 \pm 20.60	31.66 \pm 25.27	44.32 \pm 34.78	38.24 \pm 28.01
Qbert	601.5 \pm 773.9	5043.0 \pm 3193.1	845.0 \pm 1247.5	2163.1 \pm 2876.4
Pong	-20.08 \pm 0.9558	-19.72 \pm 1.096	-20.18 \pm 0.8411	-19.99 \pm 0.9899
Seaquest	782.8 \pm 348.9	802.0 \pm 322.3	859.2 \pm 523.3	814.6 \pm 409.3

表 7: Gaussian 訓練, Clean テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	18.06 \pm 11.15	6.84 \pm 6.06	11.42 \pm 8.466	12.10 \pm 9.942
Qbert	7014.5 \pm 4135.3	5970.0 \pm 4787.5	5927.0 \pm 4859.5	6303.8 \pm 4633.0
Pong	-19.66 \pm 1.893	16.88 \pm 3.314	-20.3 \pm 0.728	-7.693 \pm 17.52
Seaquest	741.6 \pm 356.1	668.0 \pm 325.8	696.0 \pm 371.8	701.8 \pm 353.1

表 8: Gaussian 訓練, Gaussian テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	19.48 \pm 10.34	22.7 \pm 21.3	15.38 \pm 8.854	19.18 \pm 14.94
Qbert	359.5 \pm 590.7	992.0 \pm 1446.3	1957.5 \pm 2326.3	1103.0 \pm 1746.2
Pong	9.14 \pm 9.16	12.78 \pm 4.883	-20.5 \pm 0.728	0.4733 \pm 16.07
Seaquest	772.4 \pm 332.4	672.0 \pm 294.3	713.6 \pm 313.5	719.3 \pm 316.5

表 9: Gaussian 訓練, Shot テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	18.48 \pm 9.77	7.6 \pm 5.26	13.76 \pm 10.91	13.28 \pm 10.03
Qbert	2916.5 \pm 3527.0	3636.0 \pm 3547.3	2899.5 \pm 3388.2	3150.6 \pm 3505.0
Pong	-19.4 \pm 2.06	12.88 \pm 5.28	-20.58 \pm 0.5688	-9.033 \pm 15.84
Seaquest	827.2 \pm 344.8	690.0 \pm 292.2	628.4 \pm 286.2	715.2 \pm 319.8

表 10: Gaussian 訓練, Impulse テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	1.38 \pm 2.05	2.06 \pm 2.52	2.56 \pm 4.42	2.0 \pm 3.2
Qbert	315.5 \pm 206.5	259.5 \pm 180.8	432.5 \pm 248.8	335.83 \pm 225.76
Pong	-9.04 \pm 6.95	0.12 \pm 6.89	-20.52 \pm 0.7277	-9.813 \pm 10.17
Seaquest	790.8 \pm 257.5	592.4 \pm 259.6	611.2 \pm 270.6	664.8 \pm 277.4

表 11: Gaussian 訓練, Speckle テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	14.0 \pm 7.26	6.76 \pm 5.53	20.2 \pm 34.2	13.65 \pm 21.21
Qbert	3160.0 \pm 3366.0	6488.0 \pm 4029.2	1847.5 \pm 2703.0	3831.8 \pm 3929.2
Pong	10.52 \pm 11.23	14.9 \pm 4.83	-20.38 \pm 0.7717	1.68 \pm 17.22
Seaquest	792.4 \pm 294.3	757.6 \pm 285.4	510.4 \pm 268.4	686.8 \pm 309.5

表 12: Shot 訓練, Clean テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	59.36 \pm 47.59	38.22 \pm 34.07	44.22 \pm 39.16	47.26 \pm 41.62
Qbert	5312.5 \pm 4775.1	3347.5 \pm 3598.1	3143.0 \pm 3880.8	3934.3 \pm 4230.0
Pong	9.6 \pm 13.7	15.82 \pm 5.058	-20.64 \pm 0.6560	1.593 \pm 18.02
Seaquest	914.8 \pm 376.7	780.4 \pm 376.3	706.8 \pm 330.6	800.6 \pm 372.0

表 13: Shot 訓練, Gaussian テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	2.04 \pm 2.74	1.14 \pm 1.14	1.9 \pm 2.4	1.693 \pm 2.256
Qbert	577.5 \pm 916.2	574.5 \pm 753.6	1460.0 \pm 1920.8	870.66 \pm 1368.4
Pong	7.2 \pm 11.8	8.96 \pm 6.82	-20.62 \pm 0.66	-1.486 \pm 15.68
Seaquest	752.0 \pm 350.7	756.8 \pm 351.8	612.0 \pm 310.1	706.9 \pm 344.7

表 14: Shot 訓練, Shot テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	40.88 \pm 22.89	45.92 \pm 36.29	35.24 \pm 27.94	40.68 \pm 29.88
Qbert	1424.5 \pm 1850.6	890.5 \pm 1052.3	2736.5 \pm 2892.8	1683.8 \pm 2214.0
Pong	7.24 \pm 13.01	11.64 \pm 5.175	-20.4 \pm 0.8	-0.5066 \pm 16.33
Seaquest	677.2 \pm 406.9	638.0 \pm 308.5	762.0 \pm 386.7	692.4 \pm 373.4

表 15: Shot 訓練, Impulse テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	1.46 \pm 1.47	2.74 \pm 3.43	1.36 \pm 2.07	1.853 \pm 2.546
Qbert	214.0 \pm 111.0	223.5 \pm 170.9	300.5 \pm 200.8	246.0 \pm 169.6
Pong	-6.82 \pm 9.23	-9.96 \pm 4.75	-20.44 \pm 0.6974	-12.40 \pm 8.369
Seaquest	611.2 \pm 235.9	520.8 \pm 199.2	474.0 \pm 192.0	535.3 \pm 217.5

表 16: Shot 訓練, Speckle テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	50.34 \pm 36.10	43.72 \pm 51.55	40.1 \pm 23.2	44.72 \pm 38.97
Qbert	2379.0 \pm 2368.0	971.0 \pm 1209.5	3555.5 \pm 4113.1	2301.8 \pm 3018.6
Pong	7.1 \pm 12.9	13.84 \pm 5.742	-20.48 \pm 0.64	0.1533 \pm 16.96
Seaquest	738.0 \pm 402.8	613.6 \pm 357.4	675.2 \pm 366.7	675.6 \pm 379.6

表 17: Impulse 訓練, Clean テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	2.14 \pm 2.32	1.56 \pm 1.43	1.3 \pm 1.66	1.666 \pm 1.878
Qbert	7703.0 \pm 4955.9	4307.0 \pm 2691.2	7970.5 \pm 4872.2	6660.1 \pm 4614.6
Pong	10.7 \pm 11.2	12.68 \pm 5.832	-20.38 \pm 0.8919	1.0 \pm 16.82
Seaquest	616.0 \pm 286.7	482.4 \pm 219.2	632.8 \pm 369.5	577.0 \pm 305.7

表 18: Impulse 訓練, Gaussian テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	2.34 \pm 3.94	2.04 \pm 1.88	2.3 \pm 3.32	2.226 \pm 3.175
Qbert	1967.5 \pm 2379.1	1164.0 \pm 1705.1	1529.0 \pm 2168.6	1553.5 \pm 2128.7
Pong	8.28 \pm 9.52	11.2 \pm 7.07	-20.54 \pm 0.6696	-0.3533 \pm 15.88
Seaquest	564.4 \pm 302.6	418.0 \pm 229.8	532.8 \pm 260.1	505.0 \pm 273.2

表 19: Impulse 訓練, Shot テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	2.84 \pm 2.73	2.92 \pm 3.19	1.84 \pm 1.59	2.533 \pm 2.644
Qbert	3336.0 \pm 3178.6	4334.0 \pm 3766.8	4569.0 \pm 3286.4	4079.6 \pm 3461.7
Pong	5.5 \pm 8.97	11.4 \pm 7.21	-20.52 \pm 0.6705	-1.206 \pm 15.38
Seaquest	551.6 \pm 281.7	570.0 \pm 278.2	622.4 \pm 381.9	581.3 \pm 319.0

表 20: Impulse 訓練, Impulse テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	1.78 \pm 1.61	2.1 \pm 1.95	1.48 \pm 1.73	1.786 \pm 1.791
Qbert	802.0 \pm 1058.9	790.5 \pm 1121.5	1054.0 \pm 1347.6	882.16 \pm 1188.8
Pong	5.54 \pm 9.17	9.62 \pm 6.56	-20.48 \pm 0.7277	-1.773 \pm 14.84
Seaquest	542.4 \pm 298.5	546.8 \pm 304.3	560.4 \pm 256.3	549.8 \pm 287.3

表 21: Impulse 訓練, Speckle テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	1.8 \pm 1.66	2.84 \pm 3.30	2.0 \pm 2.13	2.213 \pm 2.504
Qbert	4137.0 \pm 3494.9	5386.5 \pm 4612.4	4970.0 \pm 4212.6	4831.1 \pm 4165.1
Pong	9.44 \pm 11.7	14.02 \pm 5.540	-20.52 \pm 0.5741	0.98 \pm 17.04
Seaquest	423.2 \pm 241.9	617.6 \pm 268.1	648.4 \pm 332.5	563.0 \pm 300.4

表 22: Speckle 訓練, Clean テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	45.16 \pm 33.74	38.24 \pm 27.45	36.46 \pm 20.99	39.95 \pm 28.13
Qbert	3406.0 \pm 3602.8	1416.5 \pm 1882.4	2985.0 \pm 2755.9	2602.5 \pm 2961.8
Pong	9.4 \pm 14.8	16.3 \pm 3.73	-20.46 \pm 0.6069	1.746 \pm 18.24
Seaquest	903.6 \pm 402.5	786.4 \pm 434.0	961.2 \pm 435.9	883.7 \pm 430.6

表 23: Speckle 訓練, Gaussian テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	2.58 \pm 3.19	1.62 \pm 1.49	2.58 \pm 3.71	2.26 \pm 2.99
Qbert	550.5 \pm 547.1	519.0 \pm 692.6	346.5 \pm 786.1	472.0 \pm 688.3
Pong	-17.36 \pm 2.833	-12.92 \pm 3.474	-20.3 \pm 0.8306	-16.86 \pm 4.016
Seaquest	614.4 \pm 289.2	795.6 \pm 335.1	630.8 \pm 261.9	680.2 \pm 308.0

表 24: Speckle 訓練, Shot テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	27.1 \pm 19.8	30.6 \pm 25.3	41.26 \pm 22.60	32.98 \pm 23.48
Qbert	2466.0 \pm 3006.8	946.5 \pm 1140.5	1823.5 \pm 3071.0	1745.3 \pm 2641.7
Pong	-9.76 \pm 5.82	-6.3 \pm 8.01	-20.44 \pm 0.6974	-12.16 \pm 8.311
Seaquest	830.8 \pm 367.0	875.6 \pm 402.6	798.8 \pm 389.6	835.0 \pm 388.0

表 25: Speckle 訓練, Impulse テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	1.74 \pm 1.64	2.1 \pm 2.83	1.52 \pm 1.72	1.786 \pm 2.149
Qbert	258.5 \pm 176.9	272.0 \pm 201.7	270.0 \pm 194.4	266.83 \pm 191.43
Pong	-16.5 \pm 2.74	-18.22 \pm 1.526	-20.26 \pm 0.7158	-18.32 \pm 2.412
Seaquest	520.8 \pm 220.2	484.4 \pm 171.4	300.8 \pm 160.6	435.3 \pm 209.3

表 26: Speckle 訓練, Speckle テストにおけるスコア平均と標準偏差

Game Name	Three types of seed values			
	123	231	312	3seed
Breakout	42.02 \pm 18.07	34.8 \pm 38.2	26.22 \pm 16.66	34.34 \pm 27.01
Qbert	1582.0 \pm 2262.1	525.0 \pm 573.0	2512.0 \pm 3077.8	1539.6 \pm 2373.1
Pong	13.2 \pm 6.06	14.26 \pm 4.412	-20.44 \pm 0.6374	2.34 \pm 16.68
Seaquest	983.2 \pm 423.8	686.0 \pm 323.9	775.6 \pm 357.6	814.9 \pm 391.1

表 27: Breakout における 3seed 値の平均と標準偏差

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Clean	48.00 \pm 41.65	2.486 \pm 3.199	38.98 \pm 34.03	1.72 \pm 2.47	38.24 \pm 28.01
Gaussian	12.10 \pm 9.942	19.18 \pm 14.94	13.28 \pm 10.03	2.0 \pm 3.2	13.65 \pm 21.21
Shot	47.26 \pm 41.62	1.693 \pm 2.256	40.68 \pm 29.88	1.853 \pm 2.546	44.72 \pm 38.97
Impulse	1.666 \pm 1.878	2.226 \pm 3.175	2.533 \pm 2.644	1.786 \pm 1.791	2.213 \pm 2.504
Speckle	39.95 \pm 28.13	2.26 \pm 2.99	32.98 \pm 23.48	1.786 \pm 2.149	34.34 \pm 27.01

表 28: Qbert における 3seed 値の平均と標準偏差

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Clean	2205.5 \pm 2770.9	346.16 \pm 212.71	1100.0 \pm 1778.3	218.16 \pm 153.70	2163.1 \pm 2876.4
Gaussian	6303.8 \pm 4633.0	1103.0 \pm 1746.2	3150.6 \pm 3505.0	335.83 \pm 225.76	3831.8 \pm 3929.2
Shot	3934.3 \pm 4230.0	870.66 \pm 1368.4	1683.8 \pm 2214.0	246.0 \pm 169.6	2301.8 \pm 3018.6
Impulse	6660.1 \pm 4614.6	1553.5 \pm 2128.7	4079.6 \pm 3461.7	882.16 \pm 1188.8	4831.1 \pm 4165.1
Speckle	2602.5 \pm 2961.8	472.0 \pm 688.3	1745.3 \pm 2641.7	266.83 \pm 191.43	1539.6 \pm 2373.1

表 29: Pong における 3seed 値の平均と標準偏差

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Clean	1.826 \pm 17.21	-20.14 \pm 1.035	-19.92 \pm 1.233	-20.56 \pm 0.7156	-19.99 \pm 0.9899
Gaussian	-7.693 \pm 17.52	0.4733 \pm 16.07	-9.033 \pm 15.84	-9.813 \pm 10.17	1.68 \pm 17.22
Shot	1.593 \pm 18.02	-1.486 \pm 15.68	-0.5066 \pm 16.33	-12.40 \pm 8.369	0.1533 \pm 16.96
Impulse	1.0 \pm 16.82	-0.3533 \pm 15.88	-1.206 \pm 15.38	-1.773 \pm 14.84	0.98 \pm 17.04
Speckle	1.746 \pm 18.24	-16.86 \pm 4.016	-12.16 \pm 8.311	-18.32 \pm 2.412	2.34 \pm 16.68

表 30: Seaquest における 3seed 値の平均と標準偏差

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Clean	845.7 \pm 424.0	502.1 \pm 246.9	678.4 \pm 335.8	394.6 \pm 173.5	814.6 \pm 409.3
Gaussian	701.8 \pm 353.1	719.3 \pm 316.5	715.2 \pm 319.8	664.8 \pm 277.4	686.8 \pm 309.5
Shot	800.6 \pm 372.0	706.9 \pm 344.7	692.4 \pm 373.4	535.3 \pm 217.5	675.6 \pm 379.6
Impulse	577.0 \pm 305.7	505.0 \pm 273.2	581.3 \pm 319.0	549.8 \pm 287.3	563.0 \pm 300.4
Speckle	883.7 \pm 430.6	680.2 \pm 308.0	835.0 \pm 388.0	435.3 \pm 209.3	814.9 \pm 391.1

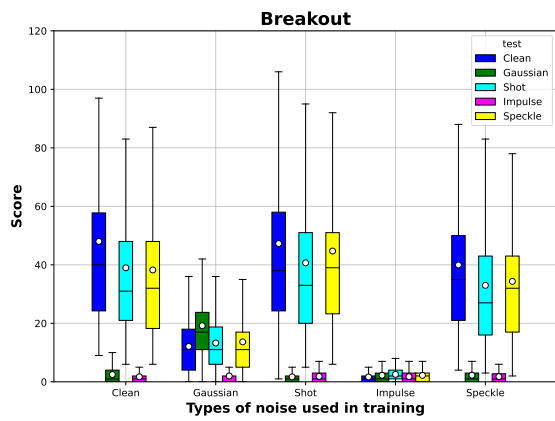


図 2: Breakout での各ノイズ訓練に対するスコア

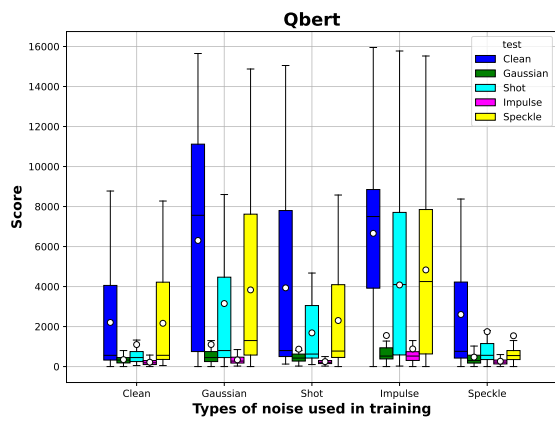


図 3: Qbert での各ノイズ訓練に対するスコア

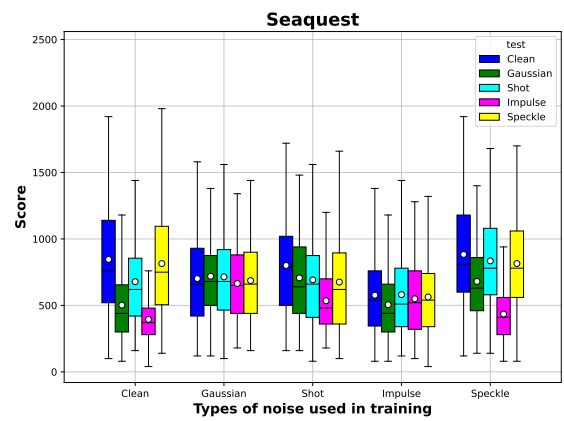


図 5: Seaquest での各ノイズ訓練に対するスコア

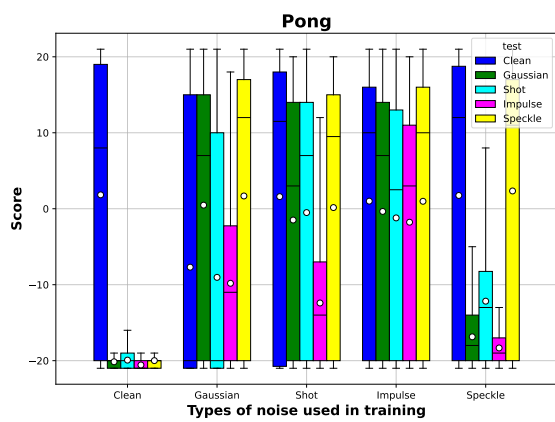


図 4: Pong での各ノイズ訓練に対するスコア

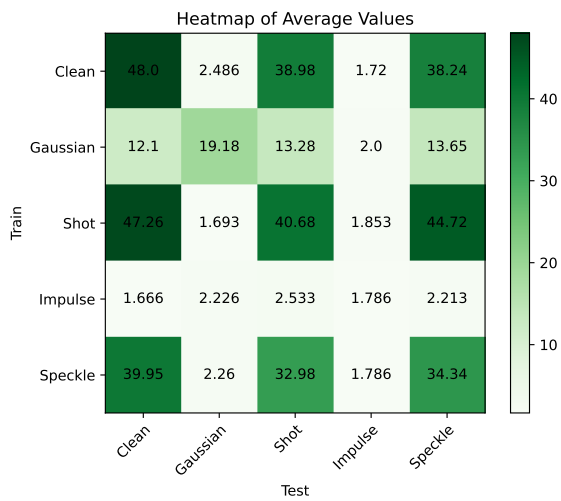


図 6: Breakout における 3seed 値の平均とヒートマップ

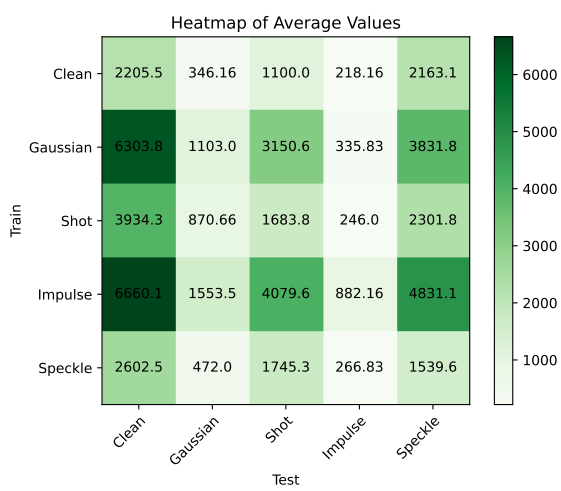


図 7: Qbert における 3seed 値の平均とヒートマップ

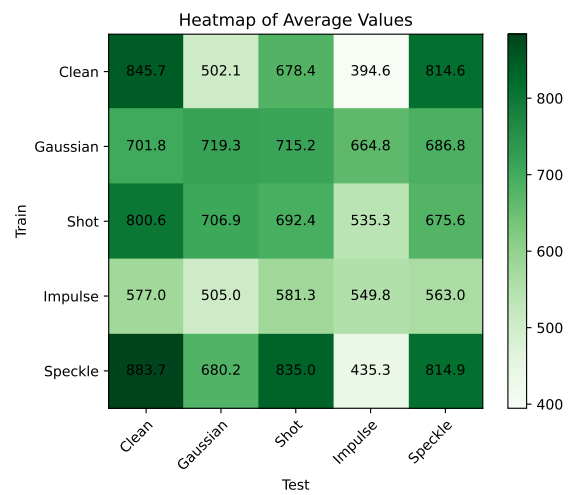


図 9: Seaquest における 3seed 値の平均とヒートマップ

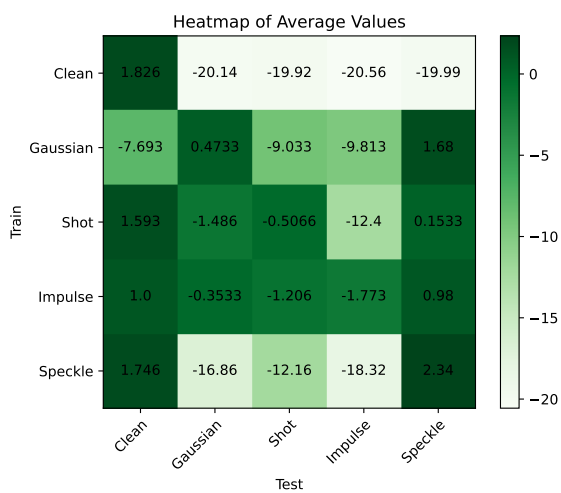


図 8: Pong における 3seed 値の平均とヒートマップ

表 31: Gaussian 比率データセットで訓練した Breakout における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Gaussian 0%	55.36 \pm 49.47	3.96 \pm 3.37	34.8 \pm 23.88	2.68 \pm 2.411	54.16 \pm 49.25
Gaussian 20%	67.32 \pm 74.26	9.8 \pm 4.50	48.24 \pm 43.99	1.16 \pm 2.344	39.76 \pm 19.87
Gaussian 40%	34.68 \pm 20.99	19.56 \pm 11.01	29.28 \pm 12.63	1.68 \pm 3.355	43.92 \pm 24.94
Gaussian 60%	33.64 \pm 21.55	22.28 \pm 19.47	31.56 \pm 25.89	1.84 \pm 3.318	40.44 \pm 44.17
Gaussian 80%	54.56 \pm 53.34	23.0 \pm 13.53	44.16 \pm 70.63	1.48 \pm 2.515	28.68 \pm 12.60
Gaussian 100%	5.6 \pm 3.65	16.56 \pm 7.62	8.68 \pm 4.78	1.08 \pm 1.622	5.92 \pm 4.298

表 32: Gaussian 比率データセットで訓練した Qbert における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Gaussian 0%	2813.0 \pm 3166.7	335.0 \pm 174.7	1365.0 \pm 1871.9	282.0 \pm 217.7	1053.0 \pm 1764.8
Gaussian 20%	2510.0 \pm 2606.6	493.0 \pm 781.7	3829.0 \pm 3699.4	534.0 \pm 781.2	2936.0 \pm 2941.6
Gaussian 40%	2144.0 \pm 2761.7	1363.0 \pm 1829.4	1496.0 \pm 1941.1	458.0 \pm 185.0	2074.0 \pm 2761.2
Gaussian 60%	2747.0 \pm 2848.8	3283.0 \pm 3005.9	1701.0 \pm 1714.4	413.0 \pm 187.8	2006.0 \pm 2289.3
Gaussian 80%	4695.0 \pm 4294.9	2057.0 \pm 2601.3	2129.0 \pm 2351.6	385.0 \pm 268.4	1474.0 \pm 2204.5
Gaussian 100%	6254.0 \pm 3768.3	2466.0 \pm 2645.8	5487.0 \pm 3840.8	339.0 \pm 219.5	3602.0 \pm 3980.2

表 33: Gaussian 比率データセットで訓練した Pong における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Gaussian 0%	10.16 \pm 11.25	-20.68 \pm 0.4664	-20.44 \pm 0.5713	-20.2 \pm 0.8	-19.64 \pm 1.162
Gaussian 20%	14.84 \pm 6.631	-16.16 \pm 4.076	-14.72 \pm 5.510	-18.8 \pm 1.811	-0.52 \pm 9.944
Gaussian 40%	15.52 \pm 6.080	3.56 \pm 9.329	-1.52 \pm 10.31	-14.56 \pm 5.123	12.16 \pm 5.842
Gaussian 60%	11.4 \pm 10.05	1.0 \pm 14.74	4.76 \pm 12.74	-10.24 \pm 9.948	11.08 \pm 8.206
Gaussian 80%	11.8 \pm 11.8	3.68 \pm 12.39	5.36 \pm 10.07	-6.8 \pm 8.499	6.96 \pm 12.62
Gaussian 100%	7.16 \pm 13.52	5.2 \pm 11.79	7.44 \pm 11.49	-6.48 \pm 9.470	9.96 \pm 10.92

表 34: Gaussian 比率データセットで訓練した Seaquest における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Gaussian 0%	1030.4 \pm 340.1	584.0 \pm 231.7	786.4 \pm 328.0	388.0 \pm 141.1	1098.4 \pm 418.1
Gaussian 20%	910.4 \pm 259.1	792.8 \pm 377.8	881.6 \pm 332.0	640.0 \pm 231.9	737.6 \pm 357.4
Gaussian 40%	1004.8 \pm 454.9	766.4 \pm 318.3	713.6 \pm 283.9	788.0 \pm 260.7	847.2 \pm 273.9
Gaussian 60%	954.4 \pm 421.4	682.4 \pm 310.4	932.8 \pm 470.1	588.8 \pm 219.9	684.0 \pm 353.3
Gaussian 80%	792.0 \pm 275.6	721.6 \pm 279.3	723.2 \pm 345.3	810.4 \pm 285.1	764.8 \pm 344.9
Gaussian 100%	787.2 \pm 369.1	686.4 \pm 378.4	736.0 \pm 373.4	585.6 \pm 286.5	792.8 \pm 331.6

表 35: Shot 比率データセットで訓練した Breakout における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Shot 0%	69.8 \pm 72.13	2.32 \pm 3.56	35.28 \pm 21.93	1.96 \pm 2.749	57.68 \pm 40.96
Shot 20%	32.8 \pm 25.86	2.52 \pm 3.689	53.84 \pm 25.81	1.96 \pm 2.521	42.76 \pm 22.92
Shot 40%	53.76 \pm 37.10	2.2 \pm 2.545	59.28 \pm 38.37	1.24 \pm 1.817	47.36 \pm 31.38
Shot 60%	47.68 \pm 26.48	2.24 \pm 2.846	52.52 \pm 45.18	1.64 \pm 1.382	37.24 \pm 19.52
Shot 80%	44.32 \pm 23.75	1.88 \pm 2.804	60.24 \pm 45.04	0.92 \pm 1.916	38.64 \pm 25.38
Shot 100%	49.44 \pm 45.55	1.84 \pm 2.033	40.4 \pm 30.32	1.8 \pm 1.496	37.92 \pm 18.73

表 36: Shot 比率データセットで訓練した Qbert における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Shot 0%	2219.0 \pm 3191.5	298.0 \pm 149.4	608.0 \pm 686.6	347.0 \pm 217.5	1344.0 \pm 2289.1
Shot 20%	2802.0 \pm 3273.7	900.0 \pm 1232.8	1887.0 \pm 2059.4	265.0 \pm 214.1	1262.0 \pm 1875.6
Shot 40%	1079.0 \pm 1465.4	328.0 \pm 185.9	1028.0 \pm 1388.3	230.0 \pm 132.8	2306.0 \pm 2746.2
Shot 60%	1981.0 \pm 2298.0	321.0 \pm 815.6	4209.0 \pm 4479.9	252.0 \pm 198.3	2682.0 \pm 3119.0
Shot 80%	4466.0 \pm 4697.6	596.0 \pm 864.4	1436.0 \pm 2144.4	326.0 \pm 278.6	2601.0 \pm 2748.1
Shot 100%	530.0 \pm 715.6	581.0 \pm 1036.0	1094.0 \pm 1371.7	356.0 \pm 251.1	497.0 \pm 240.9

表 37: Shot 比率データセットで訓練した Pong における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Shot 0%	7.64 \pm 15.43	-20.2 \pm 0.8	-19.88 \pm 1.032	-20.32 \pm 1.377	-19.0 \pm 1.766
Shot 20%	13.04 \pm 8.710	-15.2 \pm 3.334	-12.68 \pm 6.721	-16.88 \pm 3.421	-3.92 \pm 11.48
Shot 40%	7.6 \pm 14.80	-6.04 \pm 9.459	-10.28 \pm 8.660	-15.84 \pm 4.636	-0.2 \pm 14.55
Shot 60%	15.08 \pm 7.386	4.32 \pm 9.02	3.36 \pm 9.61	-3.96 \pm 10.20	6.24 \pm 13.86
Shot 80%	9.84 \pm 13.74	5.36 \pm 13.19	8.16 \pm 8.058	-9.56 \pm 8.202	8.88 \pm 12.50
Shot 100%	11.0 \pm 9.674	4.76 \pm 9.065	7.6 \pm 9.528	-4.64 \pm 8.88	10.36 \pm 9.418

表 38: Shot 比率データセットで訓練した Seaquest における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Shot 0%	995.2 \pm 356.0	455.2 \pm 237.2	825.6 \pm 282.5	418.4 \pm 203.3	864.8 \pm 413.1
Shot 20%	809.6 \pm 372.9	677.6 \pm 272.3	866.4 \pm 392.8	636.0 \pm 275.6	1016.8 \pm 428.6
Shot 40%	768.0 \pm 392.8	776.8 \pm 353.0	948.8 \pm 391.1	608.8 \pm 339.6	773.6 \pm 294.0
Shot 60%	1080.8 \pm 285.6	856.8 \pm 351.5	744.0 \pm 310.5	683.2 \pm 271.5	937.6 \pm 410.0
Shot 80%	762.4 \pm 367.6	937.6 \pm 419.6	724.8 \pm 360.2	534.4 \pm 235.7	1013.6 \pm 470.2
Shot 100%	994.4 \pm 332.3	812.8 \pm 298.2	712.8 \pm 367.7	577.6 \pm 276.9	779.2 \pm 306.0

表 39: Impulse 比率データセットで訓練した Breakout における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Impulse 0%	37.72 \pm 19.72	2.72 \pm 4.132	42.2 \pm 26.97	1.08 \pm 1.597	37.56 \pm 33.51
Impulse 20%	38.76 \pm 19.40	2.52 \pm 1.941	46.88 \pm 42.72	1.28 \pm 1.249	35.04 \pm 33.41
Impulse 40%	27.48 \pm 44.48	1.28 \pm 1.970	18.4 \pm 9.967	1.56 \pm 1.358	14.6 \pm 10.47
Impulse 60%	8.92 \pm 5.606	2.2 \pm 3.059	4.44 \pm 4.391	3.88 \pm 3.076	11.96 \pm 5.758
Impulse 80%	5.0 \pm 5.011	2.12 \pm 1.750	4.36 \pm 4.279	2.08 \pm 1.895	4.84 \pm 3.042
Impulse 100%	3.8 \pm 2.842	0.84 \pm 1.189	1.64 \pm 2.133	1.6 \pm 1.649	2.48 \pm 3.645

表 40: Impulse 比率データセットで訓練した Qbert における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Impulse 0%	4121.0 \pm 2823.9	344.0 \pm 123.7	1647.0 \pm 2205.4	164.0 \pm 124.1	543.0 \pm 709.4
Impulse 20%	6283.0 \pm 3792.3	1722.0 \pm 2267.8	1661.0 \pm 2000.0	417.0 \pm 234.9	3741.0 \pm 3377.
Impulse 40%	2618.0 \pm 3309.1	1539.0 \pm 1740.9	3009.0 \pm 3082.7	1263.0 \pm 1537.5	3014.0 \pm 2694.1
Impulse 60%	4652.0 \pm 3349.7	2638.0 \pm 2401.6	1572.0 \pm 1670.5	821.0 \pm 1399.7	1898.0 \pm 2041.5
Impulse 80%	1860.0 \pm 2108.3	1541.0 \pm 1871.2	2781.0 \pm 2950.7	1482.0 \pm 2271.4	1708.0 \pm 2248.5
Impulse 100%	7373.0 \pm 4604.4	2352.0 \pm 3086.1	5181.0 \pm 4725.7	1434.0 \pm 1682.4	5016.0 \pm 3941.7

表 41: Impulse 比率データセットで訓練した Pong における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Impulse 0%	7.92 \pm 15.00	-20.44 \pm 0.8039	-20.12 \pm 1.142	-20.04 \pm 0.9156	-20.0 \pm 1.166
Impulse 20%	12.12 \pm 10.08	-12.36 \pm 5.305	-14.88 \pm 3.581	-17.24 \pm 2.249	-1.0 \pm 8.818
Impulse 40%	10.16 \pm 13.99	0.12 \pm 11.68	-0.4 \pm 14.57	-10.68 \pm 6.839	8.88 \pm 10.37
Impulse 60%	9.52 \pm 13.55	3.56 \pm 11.95	7.56 \pm 11.76	-0.8 \pm 9.541	7.88 \pm 12.32
Impulse 80%	8.92 \pm 14.96	4.52 \pm 12.07	5.52 \pm 10.17	-0.04 \pm 10.05	11.88 \pm 6.420
Impulse 100%	6.16 \pm 12.46	3.52 \pm 13.04	5.36 \pm 12.10	2.08 \pm 10.79	10.32 \pm 11.19

表 42: Impulse 比率データセットで訓練した Seaquest における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Impulse 0%	917.6 \pm 340.3	514.4 \pm 217.4	639.2 \pm 290.7	436.0 \pm 180.3	896.4 \pm 380.4
Impulse 20%	904.0 \pm 447.7	524.0 \pm 349.3	701.6 \pm 341.3	525.6 \pm 316.2	782.4 \pm 332.6
Impulse 40%	796.0 \pm 390.7	488.0 \pm 276.1	564.8 \pm 324.5	472.8 \pm 171.6	623.2 \pm 283.2
Impulse 60%	746.4 \pm 384.8	568.0 \pm 278.2	668.0 \pm 385.6	692.8 \pm 332.6	761.6 \pm 307.8
Impulse 80%	759.2 \pm 366.0	477.6 \pm 366.9	550.4 \pm 311.9	746.4 \pm 349.5	688.0 \pm 358.3
Impulse 100%	677.6 \pm 321.9	555.2 \pm 327.9	545.6 \pm 348.1	620.8 \pm 322.2	688.0 \pm 329.7

表 43: Speckle 比率データセットで訓練した Breakout における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Speckle 0%	56.44 \pm 53.83	2.2 \pm 2.433	54.52 \pm 58.13	1.6 \pm 1.6	62.0 \pm 68.80
Speckle 20%	45.52 \pm 30.37	2.6 \pm 4.166	59.0 \pm 54.25	2.76 \pm 4.120	62.68 \pm 47.23
Speckle 40%	48.64 \pm 52.92	2.48 \pm 3.226	51.8 \pm 52.91	2.0 \pm 1.766	31.04 \pm 18.63
Speckle 60%	51.92 \pm 29.68	2.64 \pm 4.078	32.8 \pm 18.35	2.08 \pm 1.597	42.36 \pm 15.60
Speckle 80%	57.88 \pm 48.06	2.68 \pm 2.344	39.04 \pm 16.38	1.28 \pm 2.441	46.24 \pm 21.69
Speckle 100%	54.44 \pm 42.50	2.44 \pm 4.030	32.08 \pm 18.91	1.04 \pm 1.708	39.68 \pm 15.76

表 44: Speckle 比率データセットで訓練した Qbert における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Speckle 0%	3308.0 \pm 3203.2	520.0 \pm 748.4	997.0 \pm 1239.1	226.0 \pm 147.7	888.0 \pm 1112.1
Speckle 20%	1384.0 \pm 1609.0	326.0 \pm 168.0	1174.0 \pm 2134.7	290.0 \pm 243.7	1264.0 \pm 1524.2
Speckle 40%	3635.0 \pm 3568.8	732.0 \pm 1033.5	1560.0 \pm 2000.6	322.0 \pm 200.4	871.0 \pm 1437.8
Speckle 60%	2376.0 \pm 3220.2	541.0 \pm 822.5	2048.0 \pm 2047.1	327.0 \pm 167.9	2330.0 \pm 2380.7
Speckle 80%	4089.0 \pm 3830.8	773.0 \pm 1053.5	1924.0 \pm 2303.6	223.0 \pm 140.3	2898.0 \pm 2039.8
Speckle 100%	4841.0 \pm 3824.5	546.0 \pm 741.5	474.0 \pm 239.2	227.0 \pm 176.6	2578.0 \pm 2902.9

表 45: Speckle 比率データセットで訓練した Pong における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Speckle 0%	8.84 \pm 13.68	-19.88 \pm 1.142	-20.0 \pm 0.8944	-20.04 \pm 1.28	-17.88 \pm 2.250
Speckle 20%	1.72 \pm 16.29	-16.44 \pm 3.150	-16.72 \pm 1.949	-18.64 \pm 1.873	-3.84 \pm 8.278
Speckle 40%	11.64 \pm 10.87	-15.76 \pm 2.371	-12.88 \pm 3.658	-16.28 \pm 2.341	4.64 \pm 8.740
Speckle 60%	14.84 \pm 5.647	-13.16 \pm 5.604	-10.32 \pm 7.497	-18.4 \pm 1.876	11.04 \pm 7.443
Speckle 80%	8.76 \pm 14.91	-12.48 \pm 4.535	-8.76 \pm 6.924	-19.16 \pm 1.460	10.92 \pm 9.967
Speckle 100%	12.32 \pm 9.727	-8.56 \pm 4.613	-10.48 \pm 6.268	-15.68 \pm 3.221	14.16 \pm 6.239

表 46: Speckle 比率データセットで訓練した Seaquest における平均と標準偏差のスコア (修正後)

Train	Test				
	Clean	Gaussian	Shot	Impulse	Speckle
Speckle 0%	923.2 \pm 421.8	688.0 \pm 300.8	814.4 \pm 293.6	504.8 \pm 222.0	800.8 \pm 320.4
Speckle 20%	936.0 \pm 355.3	520.8 \pm 255.3	648.0 \pm 381.8	533.6 \pm 219.3	1026.4 \pm 383.7
Speckle 40%	1011.2 \pm 404.6	800.8 \pm 316.2	945.6 \pm 386.3	421.6 \pm 175.8	981.6 \pm 527.6
Speckle 60%	968.8 \pm 345.1	965.6 \pm 347.8	1020.8 \pm 420.7	640.8 \pm 241.9	951.2 \pm 469.2
Speckle 80%	782.4 \pm 416.2	930.4 \pm 323.6	954.4 \pm 391.0	537.6 \pm 232.0	873.6 \pm 341.7
Speckle 100%	891.2 \pm 395.6	711.2 \pm 316.8	1021.6 \pm 376.0	478.4 \pm 149.3	964.0 \pm 442.5

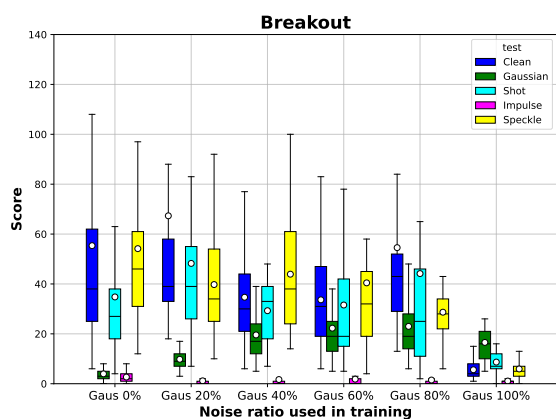


図 10: Breakout での Gaussian 比率訓練に対するスコア

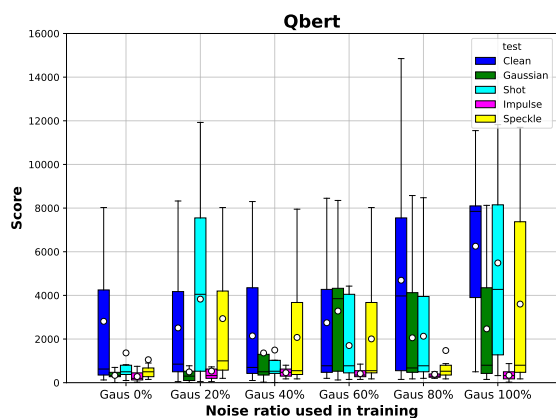


図 11: Qbert での Gaussian 比率訓練に対するスコア

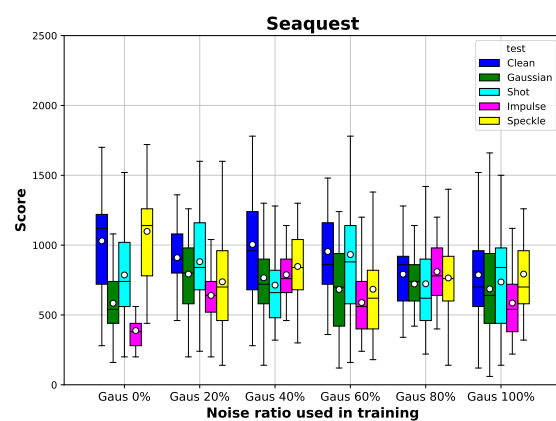


図 13: Seaquest での Gaussian 比率各訓練に対するスコア

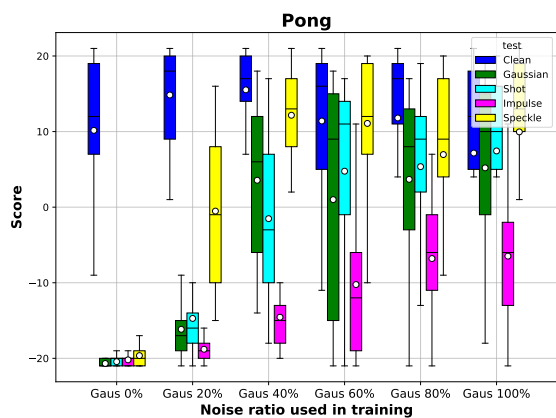


図 12: Pong での Gaussian 比率訓練に対するスコア

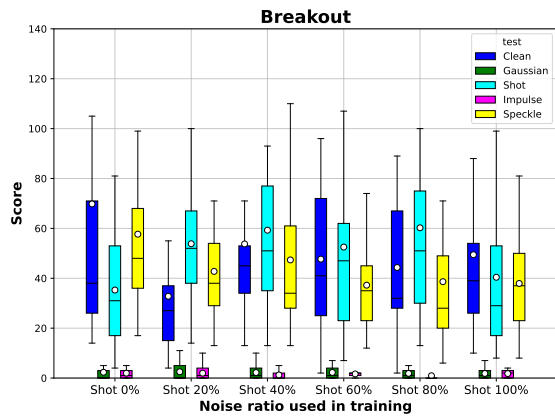


図 14: Breakout での Shot 比率訓練に対するスコア

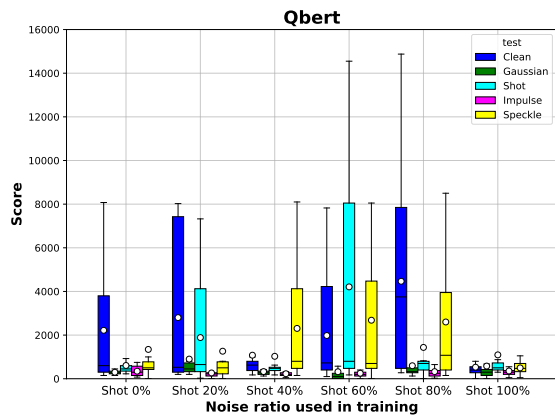


図 15: Qbert での Shot 比率訓練に対するスコア

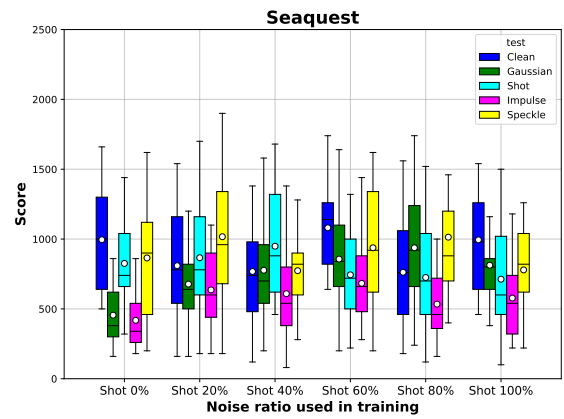


図 17: Seaquest での Shot 比率各訓練に対するスコア

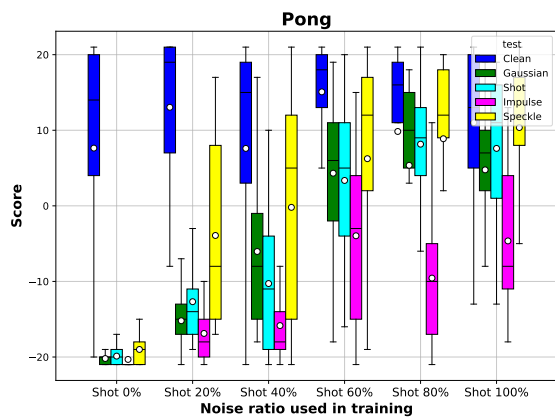


図 16: Pong での Shot 比率訓練に対するスコア

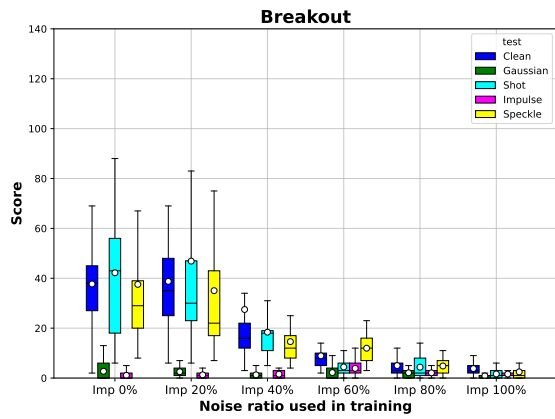


図 18: Breakout での Impulse 比率訓練に対するスコア

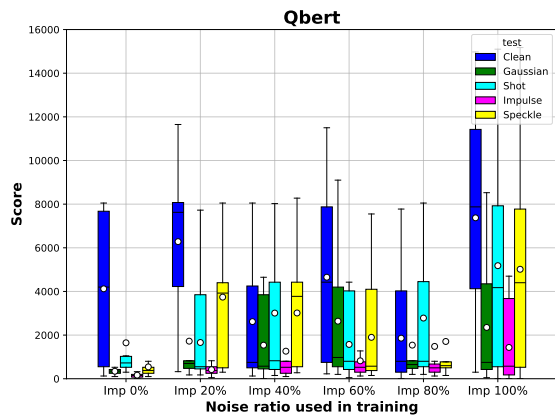


図 19: Qbert での Impulse 比率訓練に対するスコア

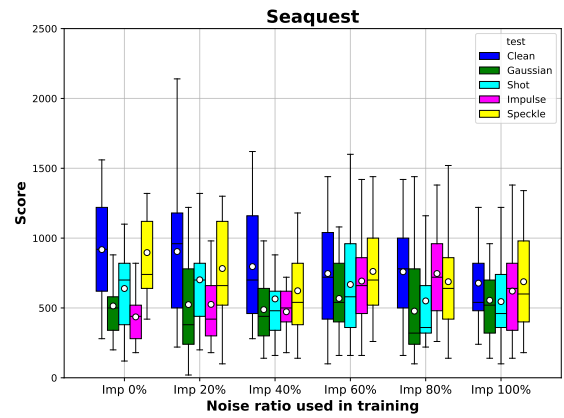


図 21: Seaquest での Impulse 比率各訓練に対するスコア

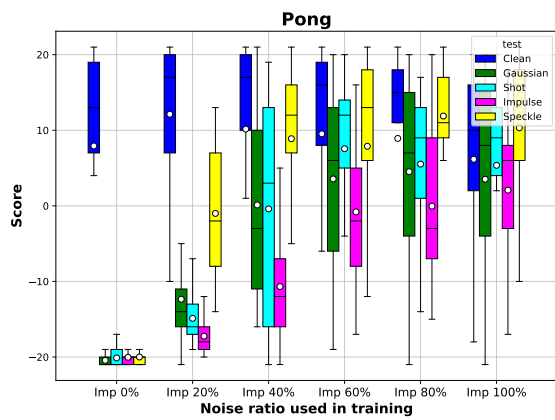


図 20: Pong での Impulse 比率訓練に対するスコア

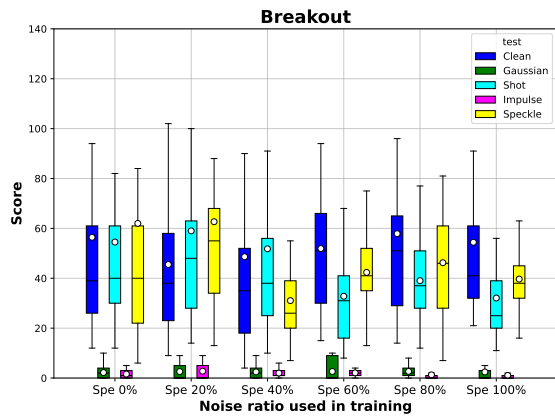


図 22: Breakout での Speckle 比率訓練に対するスコア

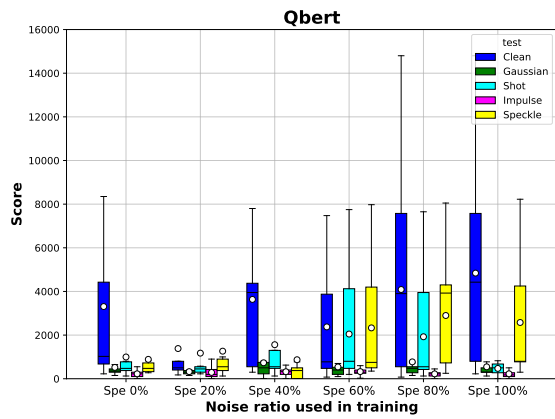


図 23: Qbert での Speckle 比率訓練に対するスコア

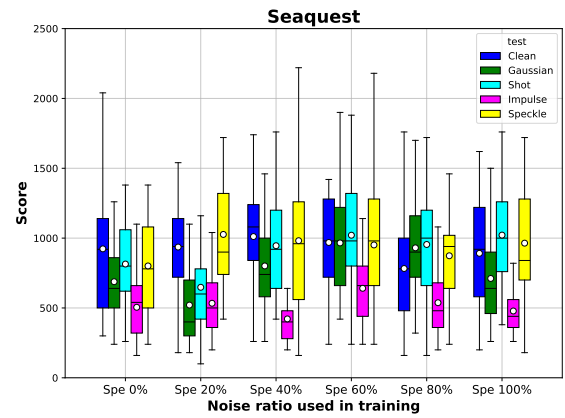


図 25: Seaquest での Speckle 比率各訓練に対するスコア

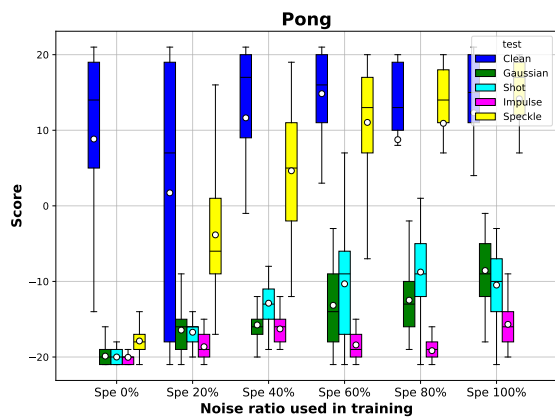


図 24: Pong での Speckle 比率訓練に対するスコア

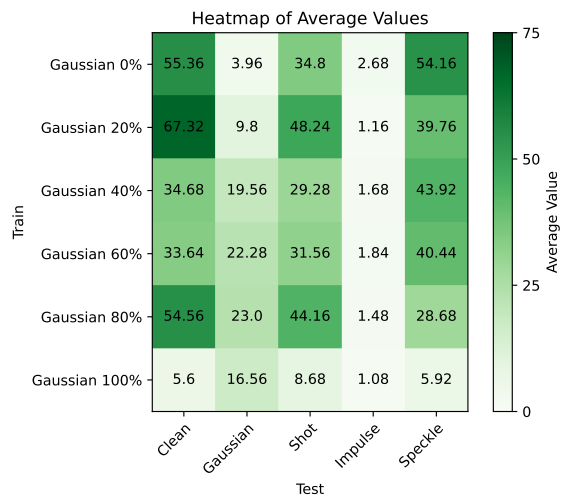


図 26: Gaussian 比率データセットにおける Breakout の平均とヒートマップ

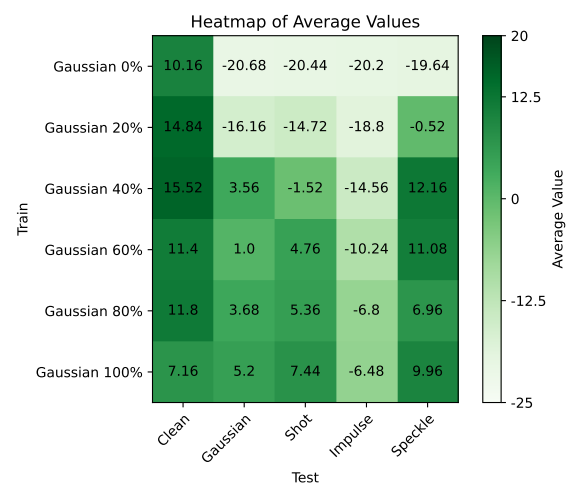


図 28: Gaussian 比率データセットにおける Pong の平均とヒートマップ

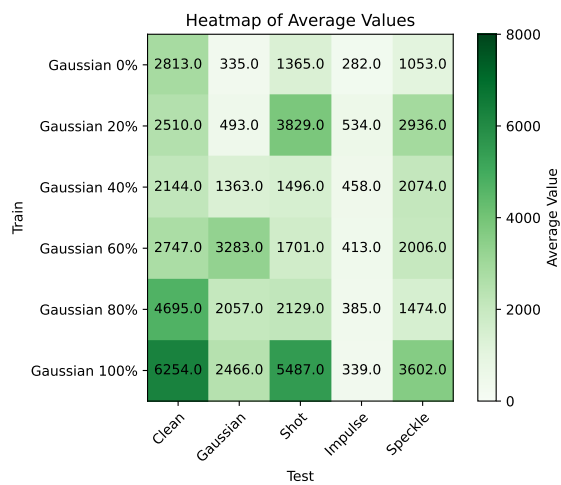


図 27: Gaussian 比率データセットにおける Qbert の平均とヒートマップ

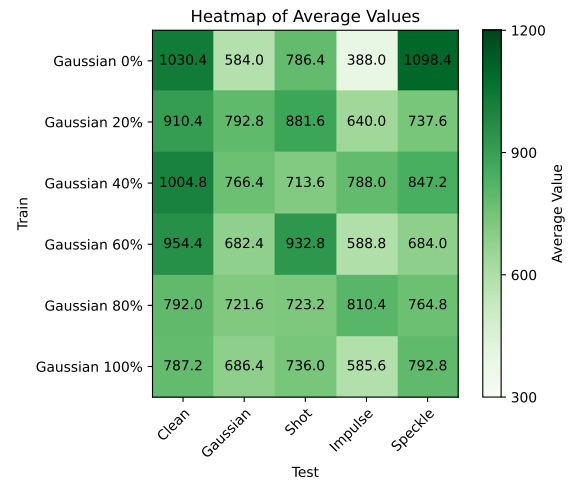


図 29: Gaussian 比率データセットにおける Seaquest の平均とヒートマップ

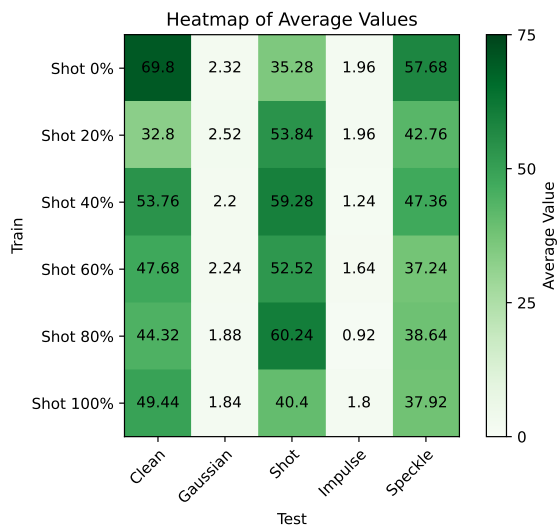


図 30: Shot 比率データセットにおける Breakout の平均とヒートマップ

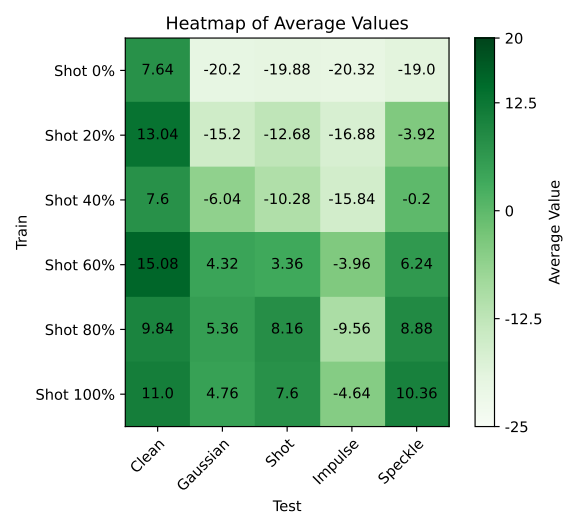


図 32: Shot 比率データセットにおける Pong の平均とヒートマップ

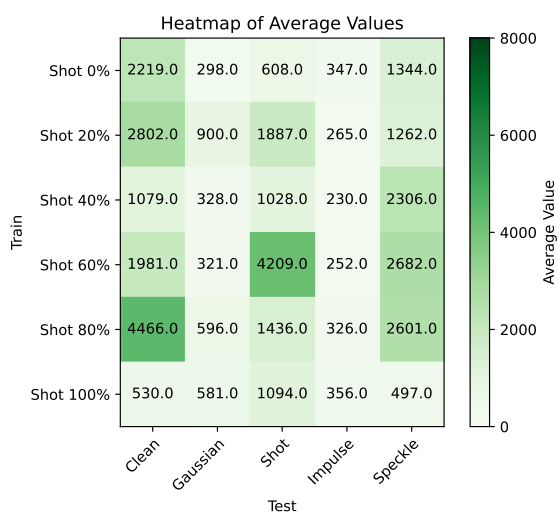


図 31: Shot 比率データセットにおける Qbert の平均とヒートマップ

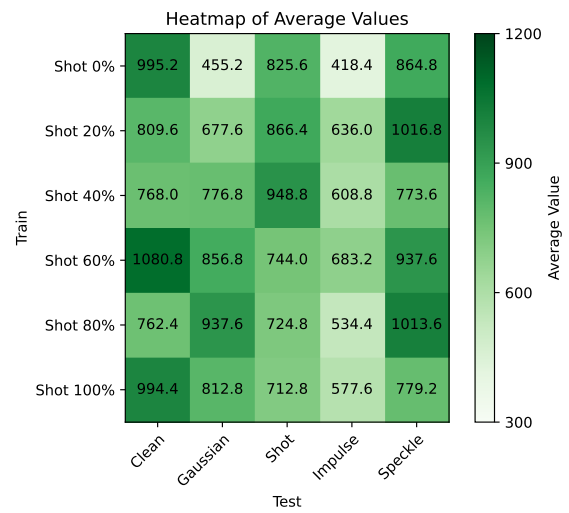


図 33: Shot 比率データセットにおける Seaquest の平均とヒートマップ

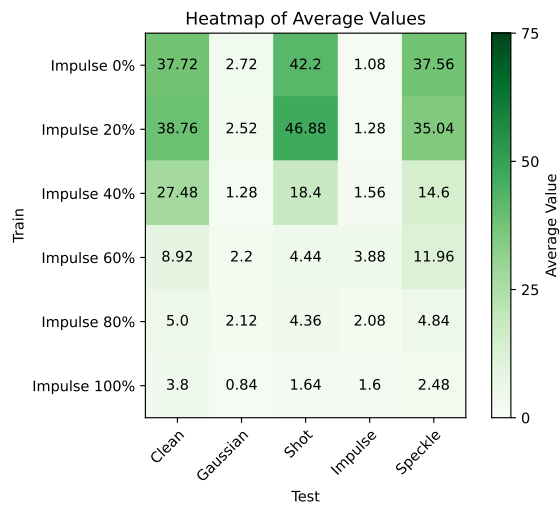


図 34: Impulse 比率データセットにおける Breakout の平均とヒートマップ

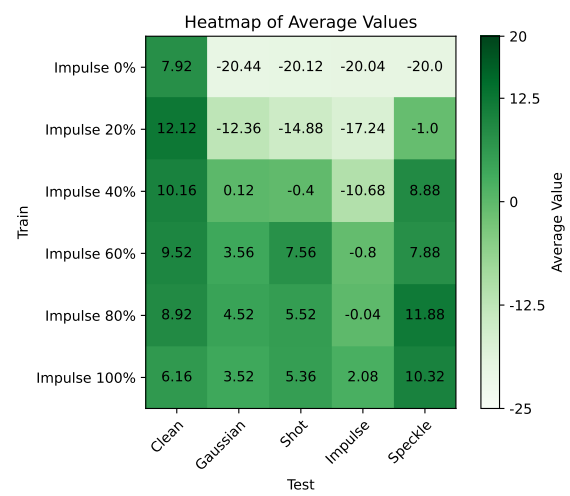


図 36: Impulse 比率データセットにおける Pong の平均とヒートマップ

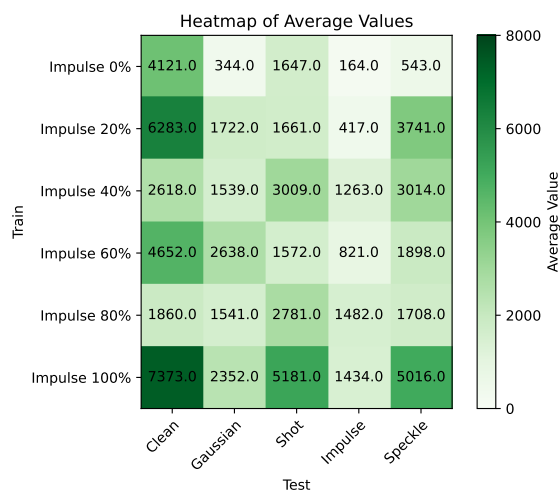


図 35: Impulse 比率データセットにおける Qbert の平均とヒートマップ

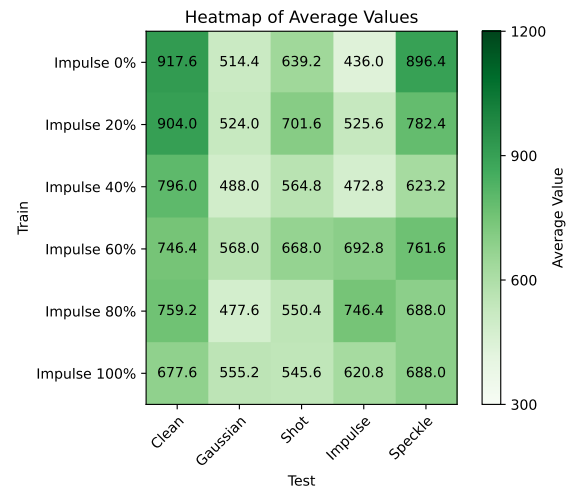


図 37: Impulse 比率データセットにおける Seaquest の平均とヒートマップ

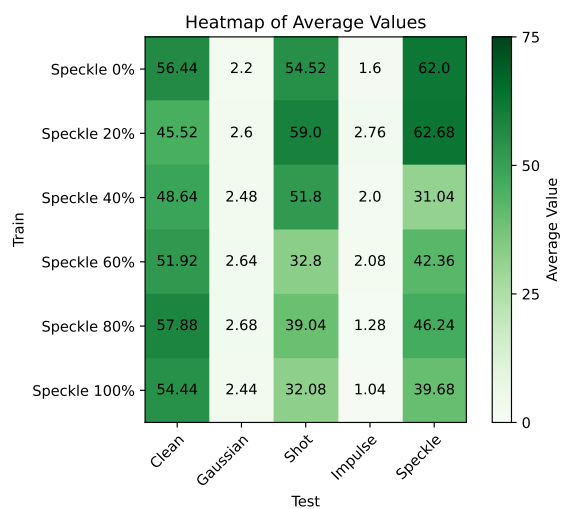


図 38: Speckle 比率データセットにおける Breakout の平均とヒートマップ

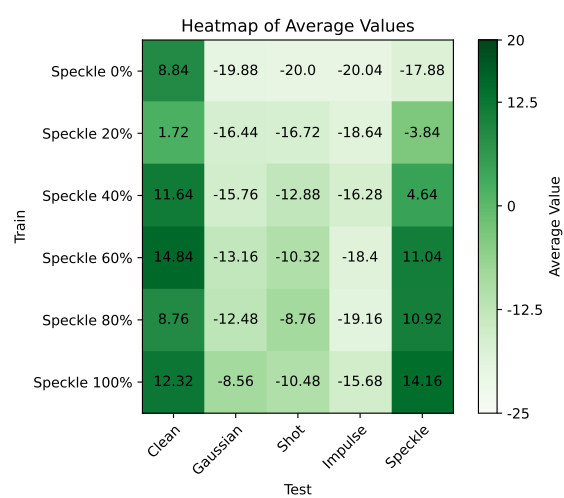


図 40: Speckle 比率データセットにおける Pong の平均とヒートマップ

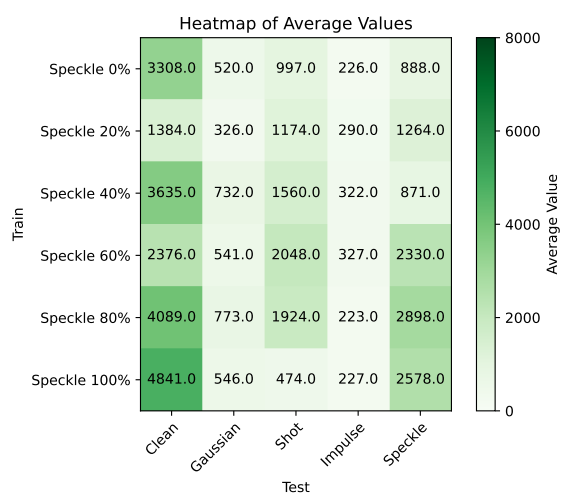


図 39: Speckle 比率データセットにおける Qbert の平均とヒートマップ

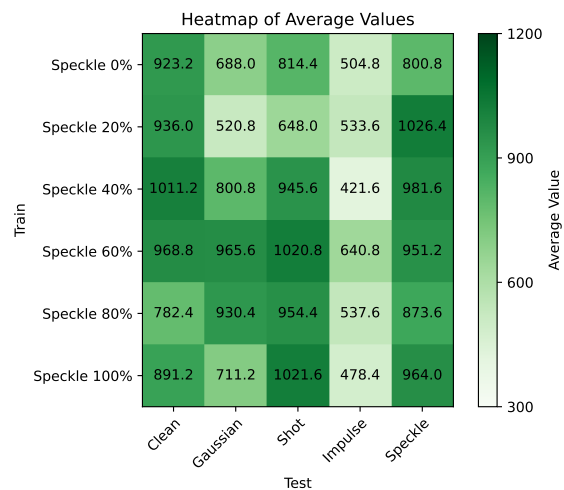


図 41: Speckle 比率データセットにおける Seaquest の平均とヒートマップ

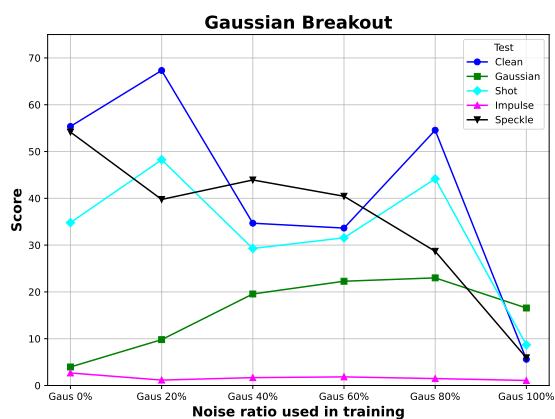


図 42: Gaussian 比率訓練における Breakout の平均報酬

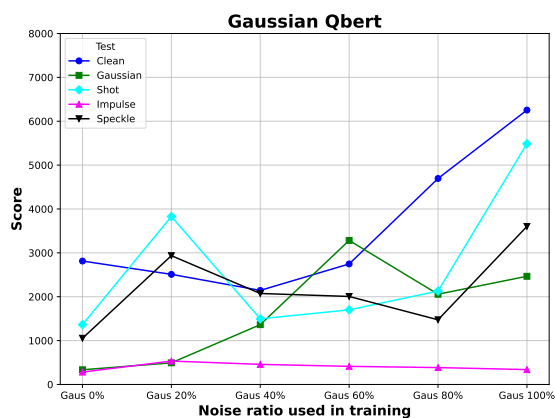


図 43: Gaussian 比率訓練における Qbert の平均報酬

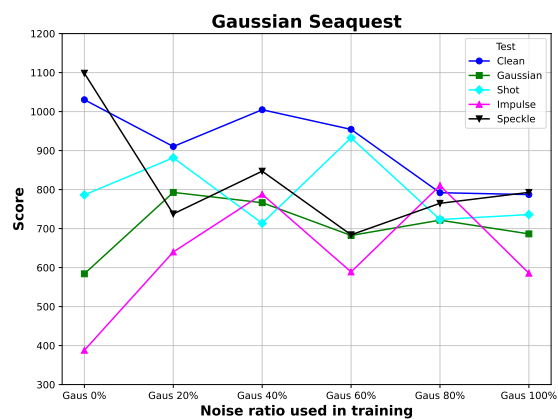


図 45: Gaussian 比率訓練における Seaquest の平均報酬

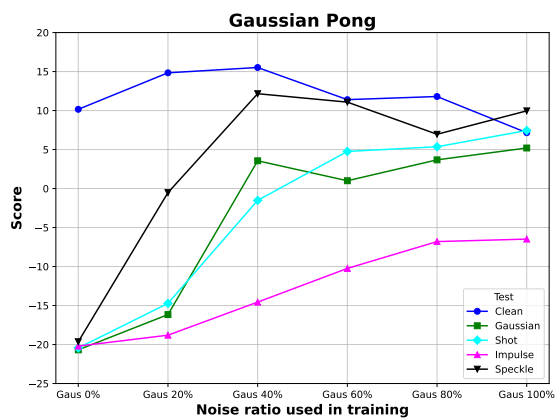


図 44: Gaussian 比率訓練における Pong の平均報酬

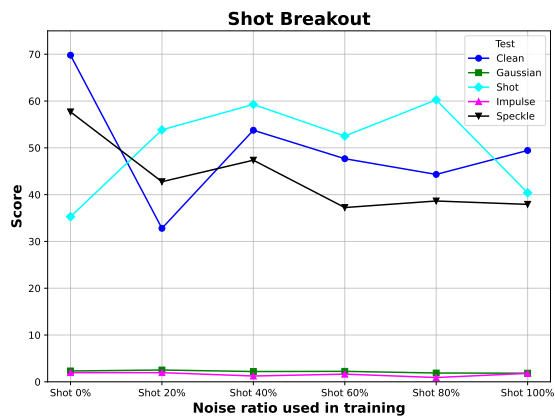


図 46: Shot 比率訓練における Breakout の平均報酬

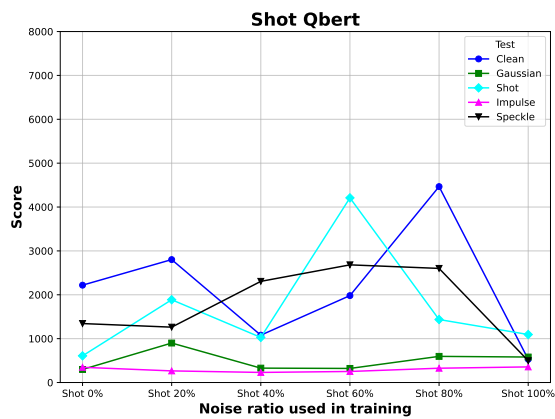


図 47: Shot 比率訓練における Qbert の平均報酬

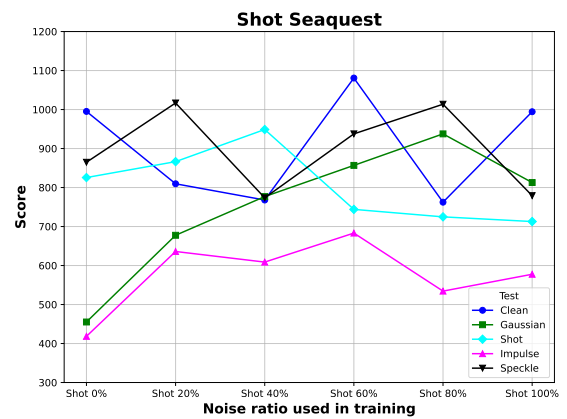


図 49: Shot 比率訓練における Seaquest の平均報酬

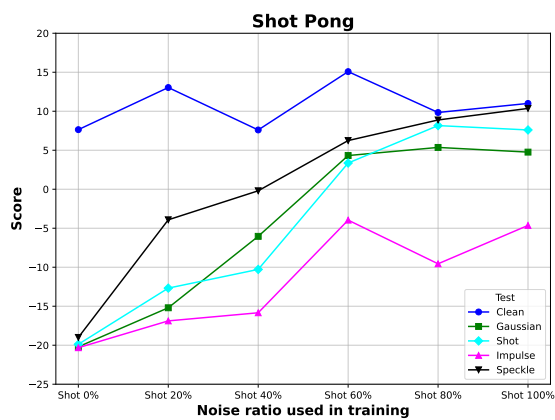


図 48: Shot 比率訓練における Pong の平均報酬

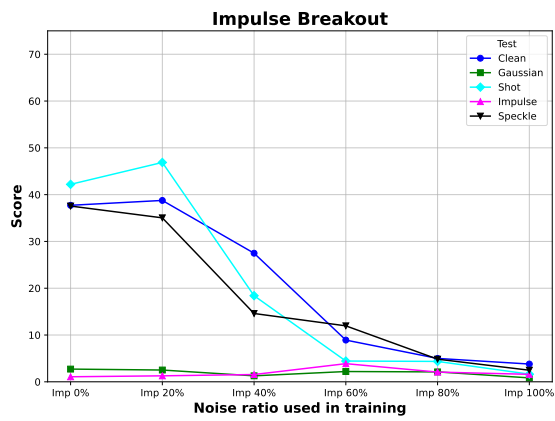


図 50: Impulse 比率訓練における Breakout の平均報酬

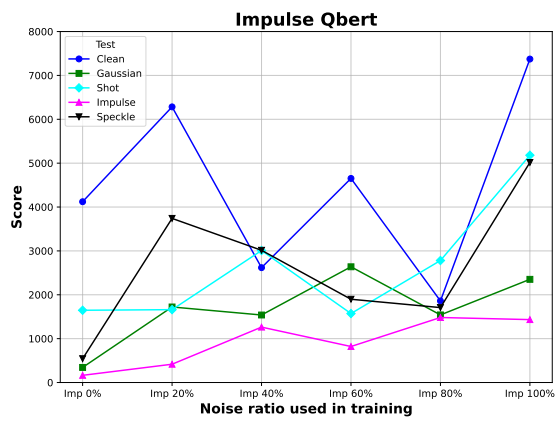


図 51: Impulse 比率訓練における Qbert の平均報酬

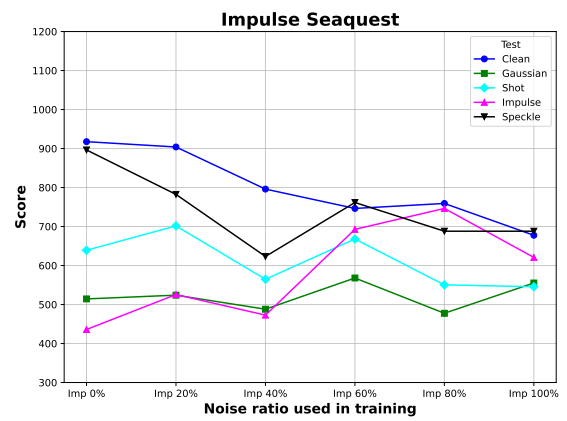


図 53: Impulse 比率訓練における Seaquest の平均報酬

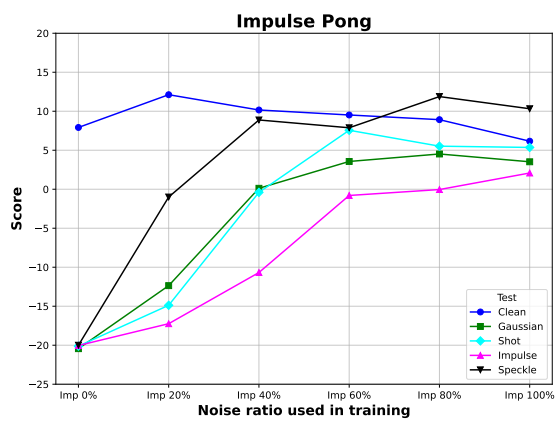


図 52: Impulse 比率訓練における Pong の平均報酬

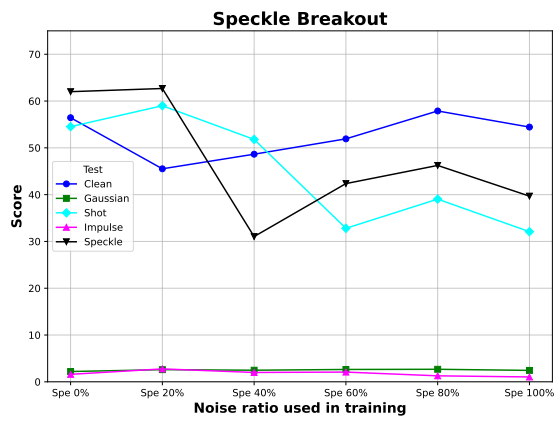


図 54: Speckle 比率訓練における Breakout の平均報酬

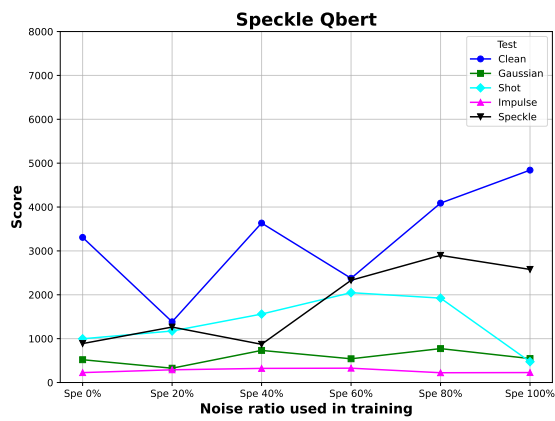


図 55: Speckle 比率訓練における Qbert の平均報酬

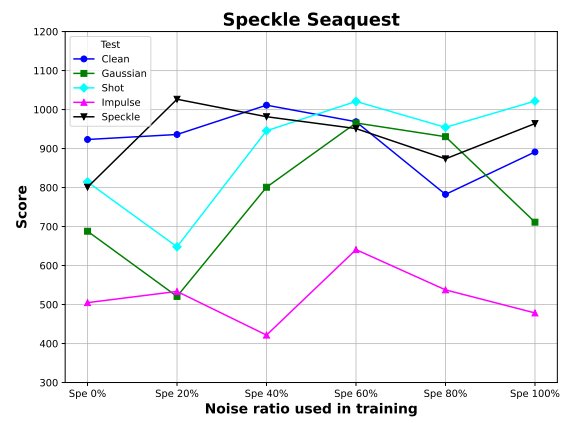


図 57: Speckle 比率訓練における Seaquest の平均報酬

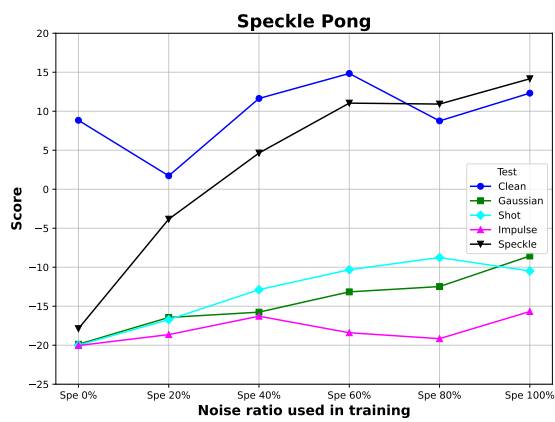


図 56: Speckle 比率訓練における Pong の平均報酬

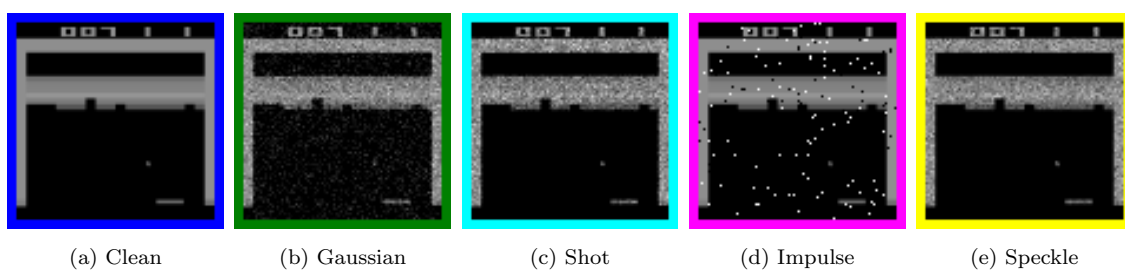


図 58: Breakout タスクの訓練で使用する観測画像の例

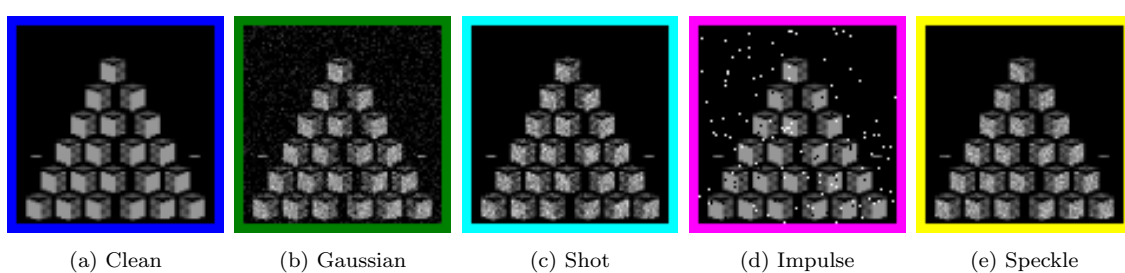


図 59: Qbert タスクの訓練で使用する観測画像の例

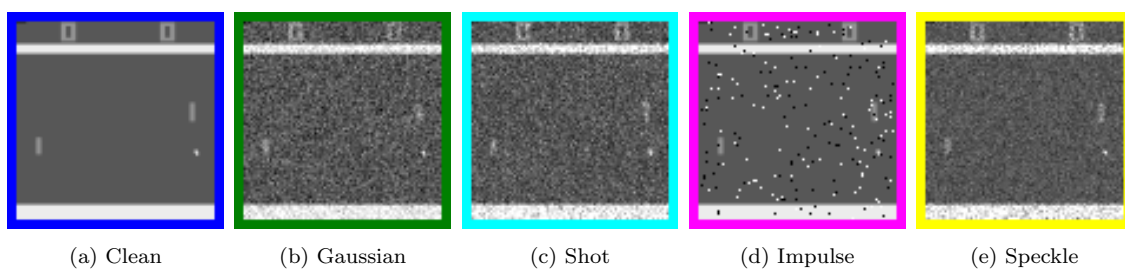


図 60: Pong タスクの訓練で使用する観測画像の例

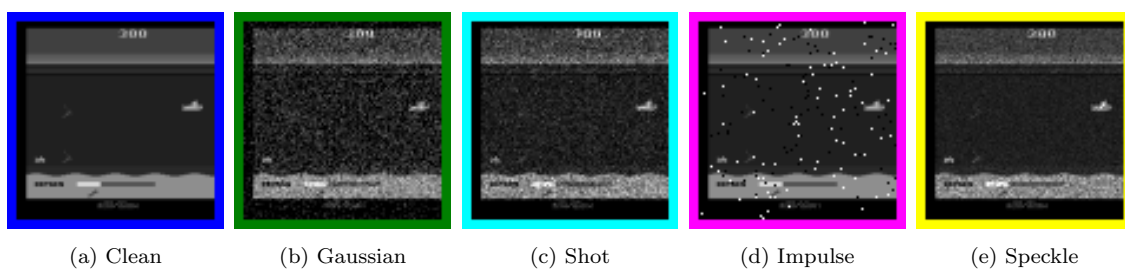


図 61: Seaquest タスクの訓練で使用する観測画像の例

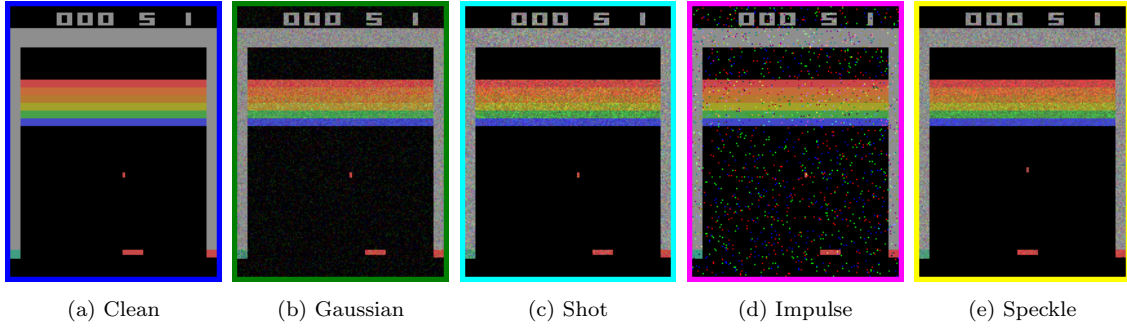


図 62: Breakout タスクにおける Clean 訓練に対する各テストの例

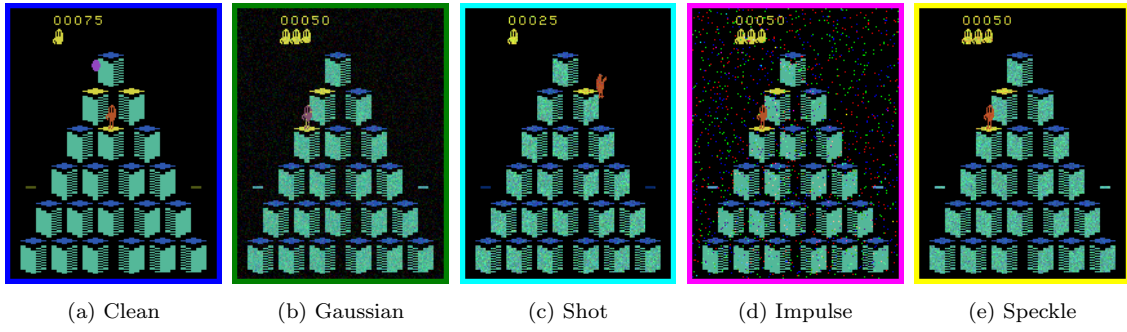


図 63: Qbert タスクにおける Clean 訓練に対する各テストの例

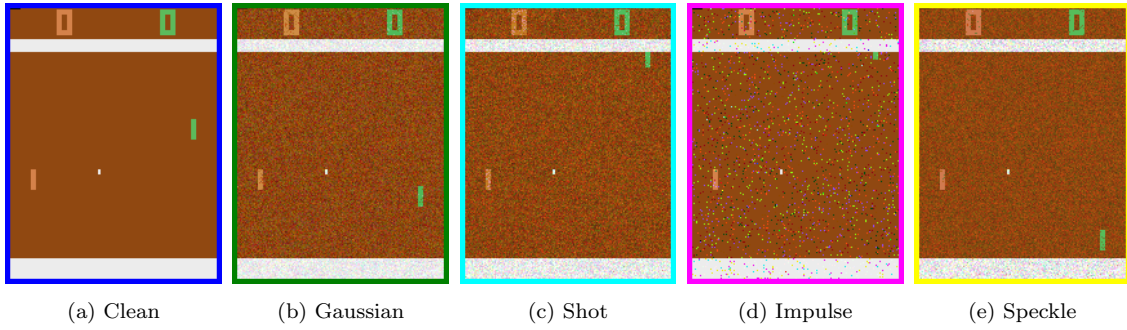


図 64: Pong タスクにおける Clean 訓練に対する各テストの例

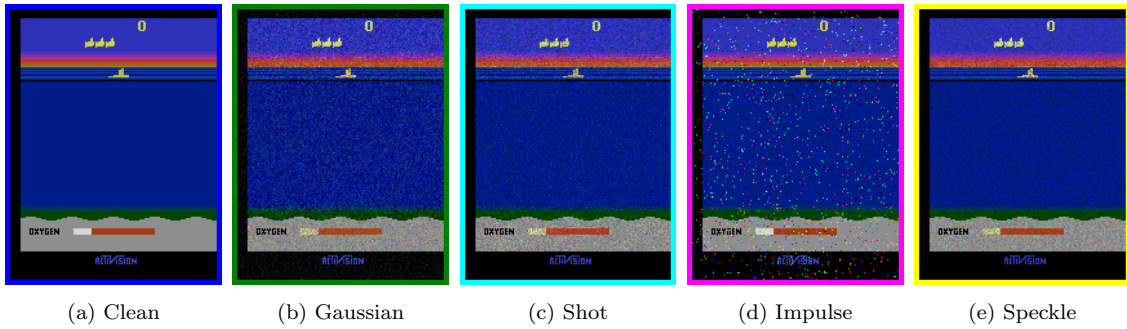


図 65: Seaquest タスクにおける Clean 訓練に対する各テストの例