The following document contains supplementary tables for ICLR submission #828 detailing network architectures, training hyperparameters, and numeric results of experiments discussed in the main document.

Table 1. Model Accuracy of MI Source Task

Training Runs	Std_A	Ort_A	RR_A^1	Pre_A	lRR_A	RR_A	Std_{AA}	Ort_{AA}	RR_{AA}^1	Pre_{AA}	lRR_{AA}	RR_{AA}
1	0.976	0.969	0.836	0.59	0.8350	0.772	0.978	0.968	0.885	0.638	0.9770	0.992
2	0.975	0.921	0.68	0.548	0.8360	0.784	0.983	0.926	0.918	0.402	0.9770	0.994
3	0.964	0.957	0.8	0.566	0.8140	0.725	0.963	0.958	0.993	0.397	0.9740	0.993
4	0.987	0.991	0.657	0.676	0.8420	0.771	0.989	0.993	0.917	0.404	0.9980	0.989
5	0.957	0.979	0.825	0.533	0.8400	0.774	0.954	0.979	0.944	0.527	0.9710	0.993
Avg.	0.972	0.963	0.760	0.5826	0.8334	0.765	0.973	0.965	0.931	0.4736	0.9794	0.992
Std. dev.	0.012	0.027	0.085	0.056	0.0112	0.023	0.015	0.025	0.040	0.1069	0.0107	0.002

Table 2. Model Accuracy of MI Transfer Task

Training Runs	Std_{AB}	Ort_{AB}	Std_B	RR_{AB}^1	Pre_{AB}	lRR_{AB}	RR_{AB}
1	0.136	0.984	0.852	0.919	0.656	0.998	0.998
2	0.136	0.136	0.991	0.954	0.594	0.991	0.997
3	0.989	0.957	0.953	0.977	0.443	0.978	0.995
4	0.136	0.991	0.909	0.963	0.597	0.989	0.997
5	0.96	0.977	0.965	0.963	0.517	0.969	1.000
Avg.	0.471	0.809	0.934	0.955	0.5614	0.9850	0.997
Std. dev.	0.459	0.376	0.055	0.022	0.0826	0.0115	0.002

Table 3. Model Accuracy of CIFAR Tests (Resnet 110 network)

Training Runs	Std_{c1}	Ort_{c1}	RR_{c1}	Std_{c10}	Ort_{c10}	RR_{c10}	Std_{c10}	Ort_{c10}	RR_{c10}
	$CIFAR\ 10$						$CIFAR\ 10.1$		
1	0.9082	0.9042	0.9076	0.9194	0.9255	0.9277	0.8275	0.8375	0.8435
2	0.909	0.9081	0.9026	0.9228	0.9231	0.9267	0.8315	0.8315	0.846
3	0.908	0.9085	0.9028	0.92	0.9239	0.927	0.8305	0.8415	0.834
4	0.9031	0.9098	0.9052	0.9176	0.9257	0.928	0.8325	0.8395	0.8335
5	0.9104	0.9106	0.9108	0.9249	0.9259	0.9286	0.835	0.8435	0.8475
Avg.	0.9077	0.9082	0.9058	0.9209	0.9248	**0.9276**	0.8314	0.8387	**0.8409**
Std. dev.	0.0028	0.0025	0.0035	0.0029	0.0012	0.0008	0.0027	0.0046	0.0067

Table 4. Model L^2 Loss of Smoke Tests. Results for $B_2: 10^7$

Training Runs	Pre_{AB_1}	Std_{AB_1}	RR_{AB_1}	Pre_{AB_2}	Std_{AB_2}	RR_{AB_2}
1	1013	338.4	197.7	1.41	1.93	1.62
2	411.4	380.2	203.6	12.4	1.63	1.55
3	621.5	2096.2	206	3.74	2.40	1.56
4	526.5	214.3	203.1	11.5	1.52	1.54
5	358.5	219.9	200.8	9.19	1.49	1.51
Avg.	586.18	649.8	202.24	7.63	1.80	1.56
Std. dev.	259.50	811.82	3.14	4.83	0.382	0.0381

Table 5. Forward and Reverse Pass Neural Networks for MNIST and *peak* Tests

Forward pass:
$MNIST$ (Fig. 1):$I(784) \rightarrow FC(10) + b_1 \rightarrow O(10)$. $peak$ (Fig. 2 & Fig.4 (a)):$I(784) \rightarrow FC(2) + b_1 \rightarrow O(2)$.
$peak$ (Fig. 4 (b)):$I(784) \rightarrow relu(BN(FC(2) + b_1) \rightarrow O(2)$.
$peak$ (Fig. 4 (c)):$I(784) \rightarrow relu(BN(FC(128) + b_1) \rightarrow relu(BN(FC(2) + b_2) \rightarrow O(2)$.

Reverse pass:
$MNIST$ (Fig. 1):$I(10) - b_1 \rightarrow FC(784) \rightarrow O(784)$ $peak$ (Fig. 2 & Fig.4 (a)): $I(10) - b_1 \rightarrow FC(784) \rightarrow O(784)$
$peak$ (Fig. 4 (b)):$I(2) - b_1 \rightarrow FC(784) \rightarrow O(784)$.
$peak$ (Fig. 4 (c)):$I(2) - b_2 \rightarrow relu(BN(FC(128)) - b_1 \rightarrow FC(784) \rightarrow O(784)$.

Table 6. Forward and Reverse Pass Neural Network for MI Tests

Forward pass:
$I(12) \rightarrow tanh(FC(10) + b_1) = d_1 \rightarrow tanh(FC(7) + b_2) = d_2 \rightarrow tanh(FC(5) + b_3) = d_3$
$\rightarrow tanh(FC(4) + b_4) = d_4 \rightarrow tanh(FC(3) + b_5) = d_5 \rightarrow tanh(FC(2) + b_6) = d_6 \rightarrow O(2)$.

Reverse pass (RR_A):
$I(2) - b_6 \rightarrow tanh\big(FC(3)\big) - b_5 \rightarrow tanh\big(FC(4)\big) - b_4 \rightarrow tanh(FC(5)) - b_3 \rightarrow tanh(FC(7)) - b_2$
$\rightarrow tanh\big(FC(10)\big) - b_1 \rightarrow tanh\big(FC(12)\big) \rightarrow O(12)$.

Reverse pass (IRR_A):
$d_6 - b_6 \rightarrow tanh\big(FC(3)\big) = d_5', d_5 - b_5 \rightarrow tanh\big(FC(4)\big) = d_4', d_4 - b_4 \rightarrow tanh\big(FC(5)\big) = d_3'$,
$d_3 - b_3 \rightarrow tanh\big(FC(7)\big) = d_2', d_2 - b_2 \rightarrow tanh\big(FC(10)\big) = d_1', d_1 - b_1 \rightarrow tanh\big(FC(12)\big) = d_0'$

Table 7. Hyperparameters of MI Tests

Batch Size	512	Learning Rate	0.0004
λ	RR_A : $\lambda_{1\sim6} = 10^{-2}$ IRR_A: $\lambda_1 = 8 \times 10^{-2}, \lambda_2 = 6 \times 10^{-2}, \lambda_3 = 4 \times 10^{-2}, \lambda_4 = 2 \times 10^{-2}, \lambda_{1,2} = 10^{-2}$		
Training Epochs	20000 for $Pre_{A/AA/AB}$, $RR_{A/AA/AB}$, $IRR_{A/AA/AB}$, $Std_{A/AA/AB}$ and Std_B		

Table 8. Forward and Reverse Pass Neural Network for Digit Generation Tests

Forward pass:
$I(74) \rightarrow relu(BN(FC(1024))) \rightarrow relu(BN(FC(6272)) \rightarrow relu(BN(C(4,64,2))) \rightarrow relu(BN(C(4,1,2))) = I_r$.

Reverse pass ($RR_{A/AA/AB}$):
$I_r \rightarrow relu(BN(D(4,64,2))) \rightarrow relu(BN(D(4,128,2))) \rightarrow relu(BN(FC(1024))) \rightarrow relu(FC(74)) \rightarrow O(74)$.

Table 9. Hyperparameters of Digit Generation Tests

Batch Size	256	λ	$\lambda_{c_1} = 2 \times 10^{-3}; \lambda_{c_{2,3}} = 1.5 \times 10^{-3}$
Learning Rate	0.001 for RR^1 and Std	Training Steps	150000 steps for RR^1 and Std

Table 10. Forward and Reverse Pass Network for Texture-shape Tests

Forward pass:
$I(224,224,3) \rightarrow relu(C(4,8,2) + b_1) \rightarrow relu(C(4,8,2) + b_2) \rightarrow relu(C(4,8,2) + b_3) \rightarrow relu(C(4,8,2) + b_4)$
$\rightarrow relu(C(4,8,2) + b_5) \rightarrow relu(C(4,8,1) + b_6) = I_r \rightarrow FC(16) \rightarrow O(16)$

Reverse pass:
$I_r - b_6 \rightarrow relu\big(D(4,8,1)\big) - b_5 \rightarrow relu\big(D(4,8,2)\big) - b_4 \rightarrow relu\big(D(4,8,2)\big) - b_3 \rightarrow relu\big(D(4,8,2)\big) - b_2$
$\rightarrow relu(D(4,8,2)) - b_1 \rightarrow relu(D(4,3,2)) \rightarrow O(224,224,3)$

Table 11. Hyperparameters of Texture-shape Tests

Batch Size	64	$\lambda_{1\sim6}$	5×10^{-12}
Learning Rate	0.0001	Training Epochs	200

Table 12. Forward and Reverse Pass Neural Network for CIFAR Tests (Resnet 110 network)

Example residual block in forward pass: $I(32,32,3) \rightarrow \cdots \rightarrow BN, relu \rightarrow d_{m_{in}} \rightarrow C(3,64,1) \rightarrow d_{m_{out}} \rightarrow BN, relu \rightarrow d_{m+1_{in}} \rightarrow C(3,64,1) \rightarrow d_{m+1_{out}} \rightarrow \bigoplus \rightarrow \cdots$
Example block in reverse pass: $d_{m+1_{out}} \rightarrow D(3,64,1) \rightarrow d'_{m+1_{in}}, d_{m_{out}} \rightarrow D(3,64,1) \rightarrow d'_{m_{in}}, \dots \dots$

Table 13. Hyperparameters of CIFAR Tests (Resnet 110 network)

Batch Size	256	λ	10^{-8}	Training Epochs	200 epochs
Learning Rate	RR_{c1}, Ort_{c1} and Std_{c1}: 0.01 RR_{c10}, Ort_{c10} and Std_{c10}: 0.01 (0-80 epochs), 0.001 (80-120 epochs) 0.0001 (120-160 epochs), 0.00001 (160-200 epochs)				

Table 14. Forward and Reverse Pass Neural Network for Smoke Tests

Generator forward pass: $I(16,16,1) \rightarrow relu(C(5,64,1)+b_1) \rightarrow UP \rightarrow relu(C(5,128,1)+b_2) \rightarrow UP \rightarrow relu(C(5,128,1)+b_3) \rightarrow relu(C(5,64,1)+b_4) \rightarrow relu(C(5,32,1)+b_5) \rightarrow relu(C(5,1,1)+b_6) \rightarrow O(64,64,1) = I_r$.
Generator reverse pass: $I_r - b_6 \rightarrow relu(D(5,32,1)) - b_5 \rightarrow relu(D(5,64,1)) - b_4 \rightarrow relu(D(5,128,1)) - b_3 \rightarrow relu(D(5,128,1)) \rightarrow MP - b_2 \rightarrow relu(D(5,64,1)) \rightarrow MP - b_1 \rightarrow relu(D(5,1,1)) \rightarrow O(16,16,1)$.
Discriminator: $I(64,64,2) \rightarrow lrelu(BN(C(5,32,1)+b_1)) \rightarrow lrelu(BN(C(5,64,1)+b_2)) \rightarrow lrelu(BN(C(5,128,1)+b_3)) \rightarrow lrelu(BN(C(5,256,1)+b_4)) \rightarrow FC(1)+b_5 \rightarrow O(1)$.

Table 15. Autoencoder Architecture for Smoke Tests

Std_{AB_1}: $I(64,64,1) \rightarrow relu(C(5,32,1)+b_1) \rightarrow relu(C(5,64,1)+b_2) \rightarrow relu(C(5,128,1)+b_3) \rightarrow relu(C(5,128,1)+b_4) \rightarrow MP \rightarrow relu(C(5,64,1)+b_5) \rightarrow MP \rightarrow relu(C(5,1,1)+b_6) \rightarrow relu(C(5,64,1)+b_7) \rightarrow UP \rightarrow relu(C(5,128,1)+b_8) \rightarrow UP \rightarrow relu(C(5,128,1)+b_9) \rightarrow relu(C(5,64,1)+b_{10}) \rightarrow relu(C(5,32,1)+b_{11}) \rightarrow relu(C(5,1,1)+b_{12}) \rightarrow O(64,64,1)$.
Pre_{AB_1}: $I(64,64,1) \rightarrow relu(C(5,64,1)+b_1) \rightarrow MP \rightarrow relu(C(5,128,1)+b_2) \rightarrow MP \rightarrow relu(C(5,128,1)+b_3) \rightarrow relu(C(5,64,1)+b_4) \rightarrow relu(C(5,32,1)+b_5) \rightarrow relu(C(5,1,1)+b_6) \rightarrow relu(D(5,32,1)+b_7) \rightarrow relu(C(5,64,1)+b_8) \rightarrow relu(C(5,128,1)+b_9) \rightarrow relu(C(5,128,1)+b_{10}) \rightarrow UP \rightarrow relu(C(5,64,1)+b_{11}) \rightarrow UP \rightarrow relu(C(5,1,1)+b_{12}) \rightarrow O(64,64,1)$.
RR_{AB_1}: $I(64,64,1) - b_1 \rightarrow relu\big(D(5,32,1)\big) - b_2 \rightarrow relu\big(D(5,64,1)\big) - b_3 \rightarrow relu\big(D(5,128,1)\big) - b_4 \rightarrow relu(D(5,128,1)) \rightarrow MP - b_5 \rightarrow relu\big(D(5,64,1)\big) \rightarrow MP - b_6 \rightarrow relu(D(5,1,1)) \rightarrow relu(C(5,64,1)+b_7) \rightarrow UP \rightarrow relu(C(5,128,1)+b_8) \rightarrow UP \rightarrow relu(C(5,128,1)+b_9) \rightarrow relu(C(5,64,1)+b_{10}) \rightarrow relu(C(5,32,1)+b_{11}) \rightarrow relu(C(5,1,1)+b_{12}) \rightarrow O(64,64,1)$.
Pre_{AB_2}, Std_{AB_2} *and* RR_{AB_2}: $I(64,64,3) \rightarrow relu(C(5,32,1)+b_1) \rightarrow relu(C(5,64,1)+b_2) \rightarrow relu(C(5,128,1)+b_3) \rightarrow relu(C(5,128,1)+b_4) \rightarrow MP \rightarrow relu(C(5,64,1)+b_5) \rightarrow MP \rightarrow relu(C(5,1,1)+b_6) \rightarrow relu(C(5,64,1)+b_7) \rightarrow UP \rightarrow relu(C(5,128,1)+b_8) \rightarrow UP \rightarrow relu(C(5,128,1)+b_9) \rightarrow relu(C(5,64,1)+b_{10}) \rightarrow relu(C(5,32,1)+b_{11}) \rightarrow relu(C(5,3,1)+b_{12}) \rightarrow O(64,64,3)$.

Table 16. Hyperparameters of Smoke Tests

Batch Size	64	Learning Rate	0.0002	λ	$\lambda_1 = 10; \lambda_{2\sim6} = 0.1$
Training Epochs	40000 for Pre_A, RR_A and Std_A; 1000 for Pre_{AB_1}, Pre_{AB_2}, RR_{AB_1}, RR_{AB_2}, Std_{AB_1} and Std_{AB_2}.				

Table 17. $I(X;\mathcal{D}_m)$ and $I(\mathcal{D}_m;Y)$ Values of All MI Models

	Results of 5 runs	$I(X;\mathcal{D}_2)$	$I(X;\mathcal{D}_3)$	$I(X;\mathcal{D}_4)$	$I(X;\mathcal{D}_5)$	$I(X;\mathcal{D}_6)$	$I(X;\mathcal{D}_7)$	$I(\mathcal{D}_2;Y)$	$I(\mathcal{D}_3;Y)$	$I(\mathcal{D}_4;Y)$	$I(\mathcal{D}_5;Y)$	$I(\mathcal{D}_6;Y)$	$I(\mathcal{D}_7;Y)$
Std_A	Avg.	11.9992	11.9879	11.7279	7.9088	2.9612	1.2367	0.9992	0.9992	0.9973	0.9947	0.9786	0.9554
	Std. dev.	0.0011	0.0138	0.2007	1.5376	0.8695	0.0973	0.0000	0.0000	0.0012	0.0031	0.0112	0.0233
Pre_A	Avg.	11.6338	11.5846	11.5518	11.4495	10.8126	4.4459	0.9468	0.9371	0.9289	0.9037	0.6987	0.0188
	Std. dev.	0.2079	0.1342	0.1192	0.0732	0.2083	0.4858	0.0295	0.0149	0.0108	0.0140	0.0732	0.0122
RR_A^1	Avg.	11.5487	10.3822	8.8948	7.8285	6.1742	3.9528	0.9564	0.8418	0.7463	0.6921	0.5969	0.4937
	Std. dev.	0.9988	1.1225	0.5715	0.7409	0.7583	0.2322	0.0940	0.0785	0.0440	0.0131	0.0652	0.0736
lRR_A	Avg.	11.5521	6.7737	4.0047	4.0087	3.8456	3.7120	0.9388	0.6372	0.6013	0.6002	0.6008	0.5997
	Std. dev.	0.0558	2.3736	0.1872	0.2787	0.2033	0.0909	0.0070	0.0334	0.0244	0.0233	0.0231	0.0236
RR_A	Avg.	5.2845	4.8838	4.2668	3.9356	4.7603	4.6238	0.5654	0.5369	0.5358	0.5182	0.5261	0.5336
	Std. dev.	0.4545	0.1569	0.3115	0.1279	0.0783	0.0173	0.0337	0.0039	0.0111	0.0060	0.0115	0.0031
Ort_A	Avg.	11.9976	11.9702	11.4784	7.7023	2.6314	1.0755	0.9992	0.9992	0.9955	0.9849	0.9681	0.9493
	Std. dev.	0.0051	0.0505	0.5529	0.8500	0.4071	0.0629	0.0000	0.0000	0.0054	0.0128	0.0254	0.0325
Std_{AA}	Avg.	11.9984	11.9770	11.7029	7.3237	2.6905	1.2204	0.9992	0.9990	0.9973	0.9932	0.9752	0.9583
	Std. dev.	0.0021	0.0292	0.2305	1.2948	0.3707	0.1700	0.0000	0.0004	0.0009	0.0080	0.0169	0.0242
Pre_{AA}	Avg.	8.5178	6.7796	5.8199	4.9982	4.6117	3.8572	0.6125	0.4349	0.3312	0.2596	0.2187	0.1649
	Std. dev.	1.2999	0.8952	0.4626	0.3502	0.2979	0.3172	0.1350	0.1015	0.0423	0.0346	0.0272	0.0310
RR_{AA}^1	Avg.	11.9989	11.8699	11.0450	9.0914	3.9232	1.3239	0.9992	0.9934	0.9813	0.9778	0.9404	0.9054
	Std. dev.	0.0016	0.2075	0.9404	1.8047	1.3972	0.4193	0.0000	0.0118	0.0217	0.0374	0.0420	0.0528
lRR_{AA}	Avg.	11.9527	10.2473	8.1597	6.7212	2.4402	1.1113	0.9992	0.9743	0.9662	0.9917	0.9859	0.9706
	Std. dev.	0.0472	1.6567	2.3965	1.8774	0.4665	0.0512	0.0000	0.0227	0.0124	0.0047	0.0110	0.0197
RR_{AA}	Avg.	11.9200	11.5621	10.6095	7.7398	2.2372	1.0416	0.9992	0.9988	0.9985	0.9975	0.9906	0.9875
	Std. dev.	0.0540	0.1463	0.5465	0.8093	1.0104	0.0505	0.0000	0.0005	0.0007	0.0024	0.0073	0.0046
Ort_{AA}	Avg.	11.9970	11.9727	11.3343	6.1690	1.9642	1.0985	0.9992	0.9991	0.9971	0.9832	0.9625	0.9477
	Std. dev.	0.0061	0.0504	0.6796	1.2529	0.5865	0.0891	0.0000	0.0002	0.0032	0.0166	0.0269	0.0320
Std_{AB}	Avg.	11.9989	11.9794	10.2985	3.0652	1.2677	0.5283	0.9992	0.9988	0.9029	0.4041	0.3923	0.3860
	Std. dev.	0.0019	0.0116	1.5095	4.0345	1.8717	0.7445	0.0000	0.0006	0.0918	0.5359	0.5329	0.5288
Pre_{AB}	Avg.	8.1364	6.0174	5.6115	4.9866	4.5760	3.7471	0.5445	0.2819	0.2525	0.2119	0.1885	0.1429
	Std. dev.	1.2778	0.5196	0.3912	0.2972	0.2805	0.2942	0.1498	0.0481	0.0276	0.0165	0.0158	0.0133
RR_{AB}^1	Avg.	11.9895	11.7982	9.7800	7.4955	3.2139	1.3122	0.9992	0.9901	0.9644	0.9661	0.9472	0.9308
	Std. dev.	0.0075	0.1710	1.0207	1.4405	1.2750	0.1537	0.0000	0.0131	0.0310	0.0357	0.0349	0.0344
lRR_{AB}	Avg.	11.9047	10.8970	8.3860	5.4257	1.7110	1.0694	0.9989	0.9915	0.9813	0.9929	0.9812	0.9735
	Std. dev.	0.0890	0.8474	1.9089	1.1511	0.4913	0.0529	0.0006	0.0089	0.0145	0.0085	0.0209	0.0235
RR_{AB}	Avg.	11.9093	11.5761	9.4227	5.6190	2.2286	1.0100	0.9992	0.9992	0.9984	0.9989	0.9962	0.9935
	Std. dev.	0.0957	0.1450	1.7109	1.4609	0.4611	0.0091	0.0000	0.0000	0.0014	0.0004	0.0032	0.0035
Ort_{AB}	Avg.	11.9946	11.9546	8.9865	3.2802	1.4296	0.8571	0.9992	0.9985	0.8048	0.7851	0.7756	0.7675
	Std. dev.	0.0104	0.0561	3.3571	2.7102	0.9848	0.4858	0.0000	0.0007	0.4279	0.4391	0.4342	0.4296
Std_B	Avg.	12.0000	11.9913	11.7602	8.6184	3.2490	1.3171	0.9992	0.9990	0.9975	0.9880	0.9458	0.9160
	Std. dev.	0.0000	0.0113	0.3400	2.0542	1.2740	0.2025	0.0000	0.0003	0.0007	0.0114	0.0476	0.0565

Table 18. Per Category Accuracy of RR_{TS} for Texture-shape Tests

Data Sets		stylized data																	edge data	filled data
Model	Runs	airplane	bear	bicycle	bird	boat	bottle	car	cat	chair	clock	dog	elephant	keyboard	knife	oven	truck	Acc.	Acc.	Acc.
RR_{TS}	0	0.9	0	0.7	0.6	0	0.9	0.6	0.8	0.2	0.7	0	1	0.6	0.5	0	0.8	0.5214	0.3063	0.4500
	1	0.9	0.6	0	0.4	0.7	1	0	0.9	0.5	0	0.4	0	0.7	0.5	0.9	1	0.5145	0.2813	0.4375
	2	1	0.9	1	0.6	0.8	0	0.7	0	0.7	0.9	0.5	0.8	0	0	1	0.9	0.6043	0.1938	0.4500
	3	0	0.8	1	0.9	0.5	0	1	0	0	0	0.3	0	0.9	0.7	0.9	0	0.4402	0.2125	0.3313
	4	0.8	0.9	0	0.7	0.5	1	0	0.5	0	0.9	0	1	0.4	0.7	0.8	1	0.5746	0.2125	0.3625
	5	0.9	0.7	0.8	0	0	1	1	0	1	0.8	0.6	0	1	0.9	0.8	1	0.6471	0.2500	0.4563
	6	0.9	0.9	0.9	0.9	0.4	0.9	0.8	0.4	0.6	0.8	0.3	0	0	0	0	1	0.5465	0.2563	0.4375
	7	1	0.9	0.8	0.7	0.6	0	0.3	0.7	0	0.9	0.6	0.4	0	0.8	0.8	1	0.5897	0.2813	0.4313
	8	0.8	0.6	0	0.5	0.5	0	0.6	0.8	0	0.8	0.3	0	0.6	0.8	0.7	0.8	0.4850	0.1813	0.3563
	9	0.7	0.8	0	0.8	0	1	0	0.9	0	0.9	0.5	1	0.9	0.7	0.7	0	0.5440	0.1938	0.3625
	Avg.	0.79	0.71	0.52	0.61	0.4	0.58	0.5	0.5	0.3	0.67	0.35	0.42	0.51	0.56	0.66	0.75	0.5467	0.2369	0.4075
	Std. dev.	0.2923	0.2767	0.4566	0.2685	0.2981	0.5007	0.4000	0.3801	0.3712	0.3592	0.2173	0.4756	0.3929	0.3204	0.3596	0.4035	0.0604	0.0438	0.0481

Table 19. Per Category Accuracy of Ort_{TS} for Texture-shape Tests

Data Sets		stylized data																	edge data	filled data
Model	Runs	airplane	bear	bicycle	bird	boat	bottle	car	cat	chair	clock	dog	elephant	keyboard	knife	oven	truck	Acc.	Acc.	Acc.
Ort_{TS}	0	0.7	0	0.8	0.3	0	1	0.7	0.8	0.5	0	0.4	0.5	0.7	0.5	0.9	0	0.4987	0.1688	0.4188
	1	0.6	0.3	0.6	0.1	0.4	1	0.4	0.5	0.5	1	0.4	0.1	0.6	0	0.6	0.8	0.4877	0.2688	0.3313
	2	0.7	0.4	0.7	0.1	0.8	0	0.8	0	0	0.6	0	0.6	0	0.5	0.6	1	0.42	0.0875	0.3313
	3	1	0.9	0.4	0.2	0.7	1	0	0.3	0.8	0.7	0.2	0.4	0.7	0	0.6	1	0.543	0.2188	0.4
	4	0	0.2	0	0.7	0.9	1	0.9	0.6	1	0	0	0.3	0.8	0.4	0.7	1	0.521	0.1813	0.3688
	5	0.5	0.2	0.7	0.4	0.1	0.9	0.3	0.4	0	0.4	0.1	0.2	0.3	0.5	0.6	0.6	0.3864	0.3438	0.3375
	6	0.9	0	0.9	0.5	0	1	0.9	0.3	0	0.7	0.4	0	0	0.9	0	1	0.4683	0.1375	0.3938
	7	0	0	0.9	0	0.6	0	0.4	0.7	0.9	0	0.3	0	0.8	0	0.8	1	0.3952	0.2375	0.3313
	8	0	0.7	0.9	0.7	0.7	1	0	0.6	0	0.9	0.7	0	0.8	0.6	0.8	0	0.525	0.2438	0.3813
	9	0.6	0.8	0.9	0.4	0.2	1	0.6	0	0.7	0	0.6	0	0.7	0.7	0	0	0.4554	0.1438	0.4250
	Avg.	0.5	0.35	0.68	0.34	0.44	0.79	0.5	0.42	0.44	0.43	0.31	0.21	0.54	0.41	0.56	0.64	0.47007	0.2031	0.3719
	Std. dev.	0.3742	0.3408	0.2898	0.2459	0.3438	0.4175	0.3367	0.2741	0.4088	0.4029	0.2378	0.2283	0.3204	0.3143	0.3134	0.4600	0.0552	0.0746	0.0373

Table 20. Per Category Accuracy of Std_{TS} for Texture-shape Tests

Data Sets		stylized data																	edge data	filled data
Model	Runs	airplane	bear	bicycle	bird	boat	bottle	car	cat	chair	clock	dog	elephant	keyboard	knife	oven	truck	Acc.	Acc.	Acc.
Std_{TS}	0	0.8	0.4	0	0.8	0.6	1	0.9	0.3	0.3	0	0.8	0	0	0	0	0	0.3700	0.1250	0.325
	1	0.9	0.1	0.7	0.5	0.7	0.9	0.3	0.1	0.5	0.9	0.6	0.7	0	0	0.8	0.8	0.5228	0.1688	0.3875
	2	1	0	0	0.8	0.8	0	0.6	0	0.4	0.9	0	0.7	0.8	0	0	0	0.3679	0.1625	0.2625
	3	0	0.2	1	0.2	0.1	0	0	0.8	0	0.9	0.4	0	0.9	0.4	0.9	1	0.4216	0.1438	0.2688
	4	0.8	0	0	0	0.1	1	0.7	0	0	0	0.1	0.9	0.7	0.7	0.4	0	0.3460	0.1438	0.2813
	5	0.6	0.6	0	0.7	0	0	0	0.8	0.8	0.9	0.2	0	0.8	0.7	0.8	1	0.4937	0.1813	0.3438
	6	0.8	0.4	0.6	0.5	0.7	0	0.8	0.8	0.6	1	0.8	0	0.6	0	0.3	0.8	0.5409	0.1250	0.3750
	7	0	0	0.8	0.6	0	1	0	0	0.7	0.9	0.5	0	0.9	0	0.7	0	0.3812	0.1750	0.3375
	8	0.5	0.8	0.6	0.4	0.6	0.6	1	0.4	0.3	0.7	0.3	0.7	0.5	0.3	0.9	0	0.5313	0.2313	0.4438
	9	0.7	0	0	0.7	0.4	0	0.7	0.6	0.5	0.9	0.3	0.9	0	0	0.6	0.8	0.4442	0.1188	0.2813
	Avg.	0.61	0.25	0.37	0.52	0.4	0.45	0.5	0.38	0.41	0.71	0.4	0.39	0.52	0.21	0.54	0.44	0.4420	0.1575	0.3306
	Std. dev.	0.3510	0.2877	0.4057	0.2616	0.3197	0.4882	0.3916	0.3490	0.2685	0.3814	0.2749	0.4175	0.3795	0.2961	0.3471	0.4695	0.0753	0.0341	0.0593

Table 21. Per Category Accuracy of Pre_{TS} for Texture-shape Tests

Data Sets		stylized data																	edge data	filled data
Model	Runs	airplane	bear	bicycle	bird	boat	bottle	car	cat	chair	clock	dog	elephant	keyboard	knife	oven	truck	Acc.	Acc.	Acc.
Pre_{TS}	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0.0625	0.0625	0.0875
	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0.0625	0.0625	0.075
	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0.0625	0.0625	0.06875
	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0.0625	0.0625	0.03125
	4	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0	0.0875	0.0625	0.06875
	5	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0.0625	0.0625	0.0625
	6	1	0	0	0	0	0	0.1	0.2	0	0	0	0	0	0	0	0	0.08125	0.0625	0.1125
	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0625	0.0625	0.0875
	8	0	0	0	1	0	0	0.1	0	0	0	0	0	0	0	0	0	0.06875	0.0625	0.08125
	9	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0.0625	0.0625	0.0625
	Avg.	0.2	0.08	0	0.2	0.1	0.1	0.12	0.22	0	0	0	0	0	0	0.06	0	0.0675	0.0625	0.07375
	Std. dev.	0.4216	0.2530	0.0000	0.4216	0.3162	0.3162	0.3120	0.4158	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1897	0.0000	0.0092	0.0000	0.0212

Table 22. Performance Comparison of Texture-shape Models

Models	Parameters	Cost (s/epoch)	Acc.
Std_{TS}	11840	0.387	0.442
Ort_{TS}	11840	0.407	0.470
RR_{TS}	11840	0.627	0.547
$RR_{TS} - reduced$	8044	0.552	0.483