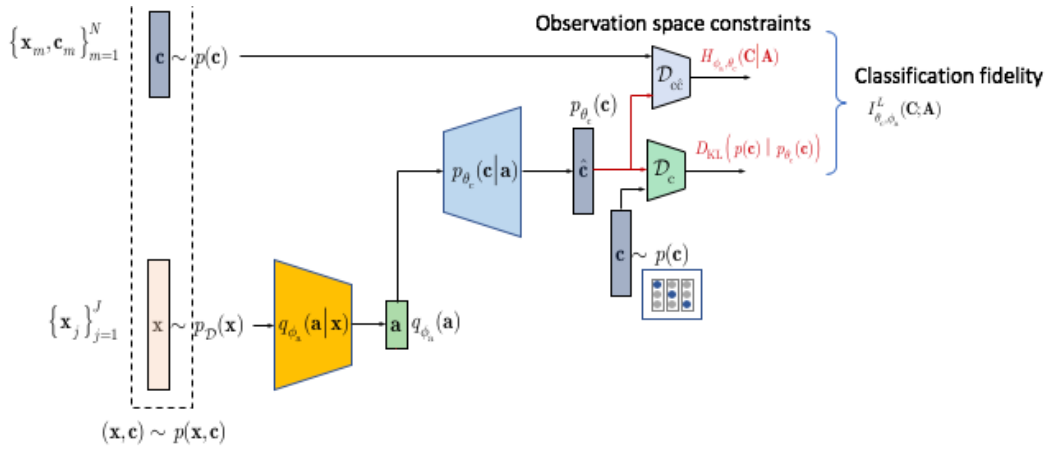
1 BCE+Dc



Encoder	
Size	Layer
$28 \times 28 \times 1$	Input
$14 \times 14 \times 32$	Conv2D, LeakyReLU
$7 \times 7 \times 64$	Conv2D, LeakyReLU
$4 \times 4 \times 128$	Conv2D, LeakyReLU
2048	Flatten
1024	FC, ReLU
500	FC, ReLU
10	FC, Softmax

Encoder BN	
Size	Layer
$28 \times 28 \times 1$	Input
$14 \times 14 \times 32$	Conv2D, BN, LeakyReLU
$7 \times 7 \times 64$	Conv2D, BN, LeakyReLU
$4 \times 4 \times 128$	Conv2D, BN, LeakyReLU
2048	Flatten
1024	FC, ReLU
500	FC, ReLU
10	FC, Softmax

Dc	
Size	Layer
10	Input
500	FC, ReLU
500	FC, ReLU
1	FC, Sigmoid

Training:

$$\mathrm{BCE} + \alpha \cdot \mathrm{Dc} \tag{1}$$

Results:

# supervised samples – 100							# supervised samples – 1000							# supervised samples – All						
		runs							runs						runs					
alpha	model	1	2	3	mean	std	alpha	model	1	2	3	mean	std	alpha	model	1	2	3	mean	std
0	Encoder + Dc	26.562	26.241	28.044	26.95	0.96	0	Encoder + Dc	7.742	6.991	6.971	7.23	0.44	0	Encoder + Dc	0.831	0.831	0.741	0.80	0.05
0.005		20.442	21.935	18.98	20.45	1.48	0.005		5.618	6.059	5.598	5.76	0.26	0.005		0.831	0.821	0.881	0.84	0.03
0.0005		18.549	20.432	20.592	19.86	1.14	0.0005		6.3	6.119	6.019	6.15	0.14	0.0005		0.861	0.921	0.821	0.87	0.05
1		19.23	22.415	20.572	20.74	1.60	1		5.999	6.27	6.28	6.18	0.16	1		0.721	0.851	0.871	0.81	0.08
0	Encoder BN + Dc	29.366	29.266	30.618	29.75	0.75	0	Encoder BN + Dc	7.451	6.951	7.522	7.31	0.31	0	Encoder BN + Dc	0.731	0.671	0.791	0.73	0.06
0.005		27.974	28.024	26.272	27.42	1.00	0.005		5.568	5.078	5.218	5.29	0.25	0.005		0.721	0.731	0.701	0.72	0.02
0.0005		25.991	23.697	24.469	24.72	1.17	0.0005		5.608	6.049	6.219	5.96	0.32	0.0005		0.751	0.771	0.721	0.75	0.03
1		27.784	31.981	35.877	31.88	4.05	1		6.049	6.41	5.819	6.09	0.30	1		0.67	0.681	0.731	0.69	0.03