

Supplementary Materials

A. Model Architecture

We provide the detailed model architecture (Table 5 and Table 7) for each component in our model: Generator, Disentangler, Domain Classifier, Classifier and MINE.

Table 5. Model Architecture for ‘Digit-Five’. For each convolution layer, we list the input dimension, output dimension, kernel size, stride, and padding. For the fully-connected layer, we provide the input and output dimensions. For drop-out layers, we provide the probability of an element to be zeroed.

layer	configuration
Feature Generator	
1	Conv2D (3, 64, 5, 1, 2), BN, ReLU, MaxPool
2	Conv2D (64, 64, 5, 1, 2), BN, ReLU, MaxPool
3	Conv2D (64, 128, 5, 1, 2), BN, ReLU
Disentangler	
1	FC (8192, 3072), BN, ReLU
2	DropOut (0.5), FC (3072, 2048), BN, ReLU
Domain Identifier	
1	FC (2048, 256), LeakyReLU
2	FC (256, 2), LeakyReLU
Class Identifier	
1	FC (2048, 10), BN, Softmax
Reconstructor	
1	FC (4096, 8192)
Mutual Information Estimator	
fc1_x	FC (2048, 512)
fc1_y	FC (2048, 512), LeakyReLU
2	FC (512,1)

B. Details of datasets

We provide the detailed information of datasets (Table 6). For Digit-Five and the DomainNet dataset, we provide the train/test split for each domain and for Office-Caltech10, we provide the number of images in each domain.

Table 6. Detailed information for datasets

Digit-Five							
Splits	<i>mnist</i>	<i>mnist'm</i>	<i>svhn</i>	<i>syn</i>	<i>usps</i>	Total	
Train	55,000	55,000	25,000	25,000	7,348	167,348	
Test	10,000	10,000	14,549	9,000	1,860	37,309	
Office-Caltech10							
Splits	<i>amazon</i>	<i>caltech</i>	<i>dslr</i>	<i>webcam</i>	Total		
Total	958	1,123	157	295	2,533		
DomainNet							
Splits	<i>clp</i>	<i>inf</i>	<i>pnt</i>	<i>qdr</i>	<i>rel</i>	<i>skt</i>	Total
Train	34,019	37,087	52,867	120,750	122,563	49,115	416,401
Test	14,818	16,114	22,892	51,750	52,764	21,271	179,609

Table 7. Model Architecture for ‘Office-Caltech10’ and ‘DomainNet’. For each convolution layer, we list the input dimension, output dimension, kernel size, stride, and padding. For the fully-connected layer, we provide the input and output dimensions. For drop-out layers, we provide the probability of an element to be zeroed.

layer	configuration
Feature Generator: ResNet101 or AlexNet	
Disentangler	
1	Dropout(0.5), FC (2048, 2048), BN, ReLU
2	Dropout(0.5), FC (2048, 2048), BN, ReLU
Domain Identifier	
1	FC (2048, 256), LeakyReLU
2	FC (256, 2), LeakyReLU
Class Identifier	
1	FC (2048, 10), BN, Softmax
Reconstructor	
1	FC (4096, 2048)
Mutual Information Estimator	
fc1_x	FC (2048, 512)
fc1_y	FC (2048, 512), LeakyReLU
2	FC (512,1)