A Model Structure

The model structures for the generator, the discriminator, and the regressor are detailed in Table 2, Table 3, and Table 4 respectively.

Table 2: Model structure of the generator.

i	Oper	skip	Ker.	Stri.	Dim.	BN	Drop	activ.
0	conv	to 11	4x4	2x2	32	N	-	leaky
1	conv	to 10	4x4	2x2	64	Y	-	leaky
2	conv	to 9	4x4	2x2	128	Y	-	leaky
3	conv	to 8	4x4	2x2	256	Y	-	leaky
4	conv	to 7	4x4	2x2	256	Y	-	leaky
5	conv	-	4x4	2x2	256	Y	-	leaky
6	trans_conv	-	4x4	2x2	256	Y	0.5	relu
7	trans_conv	-	4x4	2x2	256	Y	0.5	relu
8	trans_conv	-	4x4	2x2	256	Y	-	relu
9	trans_conv	-	4x4	2x2	128	Y	-	relu
10	trans_conv	-	4x4	2x2	64	Y	-	relu
11	trans_conv	-	4x4	2x2	32	Y	-	relu
12	conv	-	4x4	1x1	1	N	-	relu

Table 3: Model structure of the discriminator; the input of discriminator layers is directly passed from shared layers.

Usage	Operation	Ker.	Stri.	Dim.	BN	activ.
Shared layers	conv	4x4	2x2	32	N	leaky
	conv	4x4	2x2	64	Y	leaky
	conv	4x4	2x2	128	Y	leaky
Predict m2n	conv	4x4	2x2	128	Y	leaky
	conv	4x4	2x2	256	Y	leaky
	linear	-	-	128	Y	relu
	linear	-	-	1	N	-
Discriminator layers	conv	4x4	1x1	256	Y	relu
	conv	4x4	1x1	1	N	-

Table 4: Model structure of the regressor. The first batch normalization layer right serves as z-score normalization. After the convolution layers, 10 dimension features are passed into linear layers along with the convolution layers' output.

Operation	Kernel	Strides	Dim.	BN	activ.
BN	-	-	-	Y	-
conv	4x4	2x2	16	Y	relu
conv	3x3	2x2	32	Y	relu
conv	3x3	2x2	64	Y	relu
conv	3x3	2x2	128	Y	relu
concatenate 10 additional features					
linear	-	-	256	Y	relu
linear	-	-	64	Y	relu
linear	-	-	1	N	-

B Hyperparameters

The hyperparameters used in the proposed model are shown in Table 5 for reproducibility.

Table 5: Hyper-parameters used in the 5-stage training.

Stage	α(L2)	β(regr)	γ(m2n)	Max epochs
Pre-training Regressor	-	-	-	70
VIS Generator	1000	0.0001	0.002	500
Pre-training Regressor	-	-	-	100
PMW Generator	10	0.001	-	200
Fine-tune regressor	-	-	-	300