

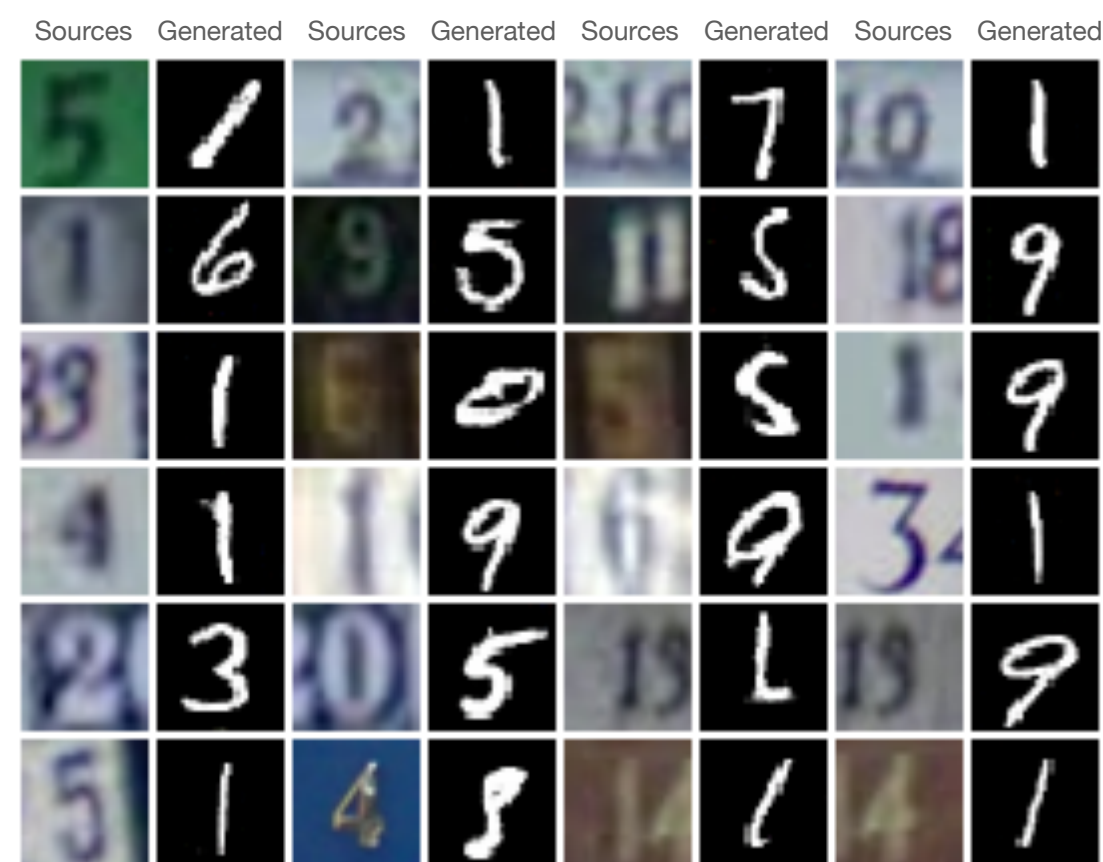
Integrating Categorical Semantics into Unsupervised Domain Translation

Samuel Lavoie, Faruk Ahmed and Aaron Courville

Problem formulation

Semantics preserving Unsupervised Domain Translation

SVHN→MNIST using MUNIT



MNIST→SVHN using MUNIT



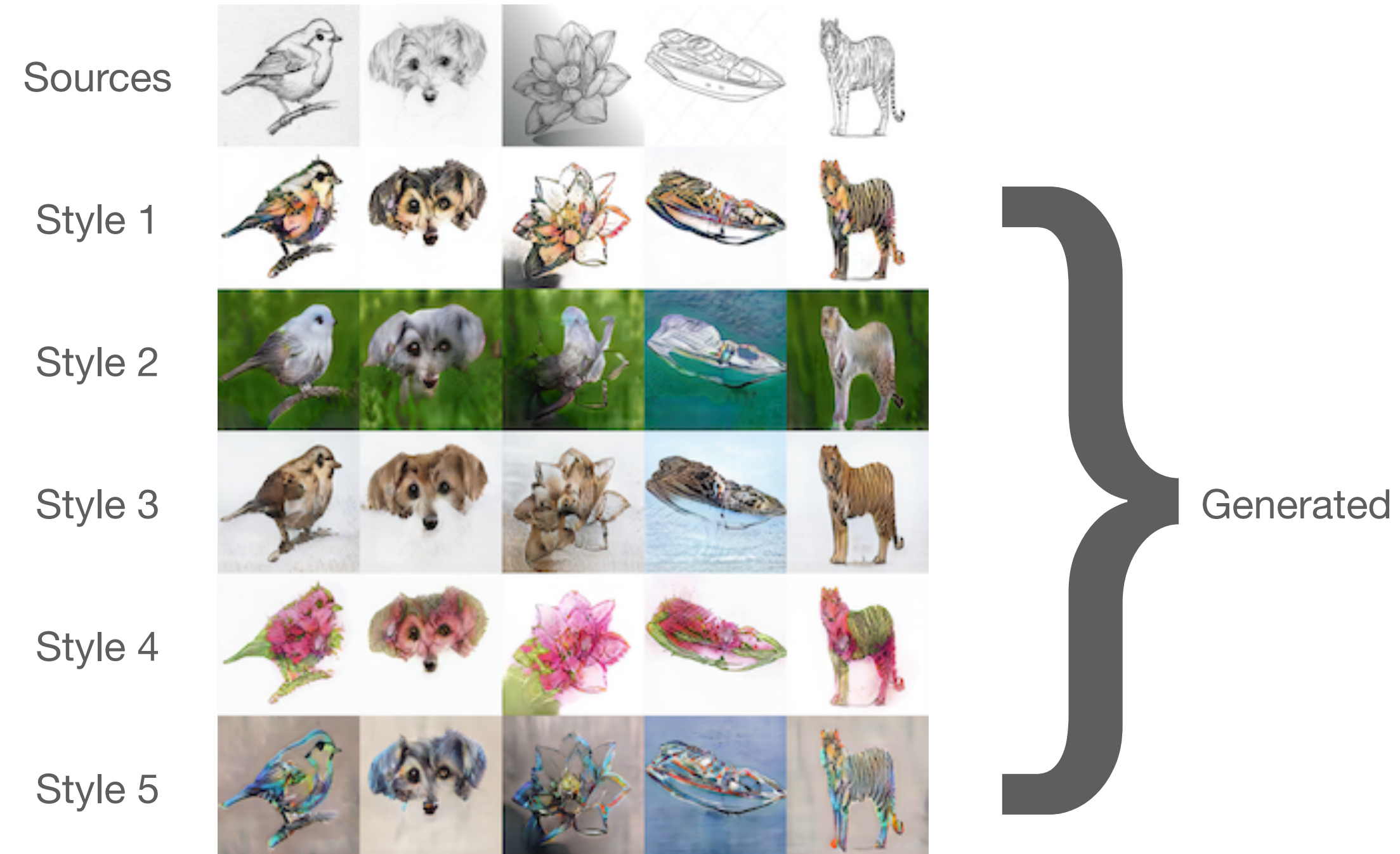
MNIST: handwritten digits. See: <http://yann.lecun.com/exdb/mnist/index.html>

SVHN: Street view house numbers. See: <http://ufldl.stanford.edu/housenumbers/>

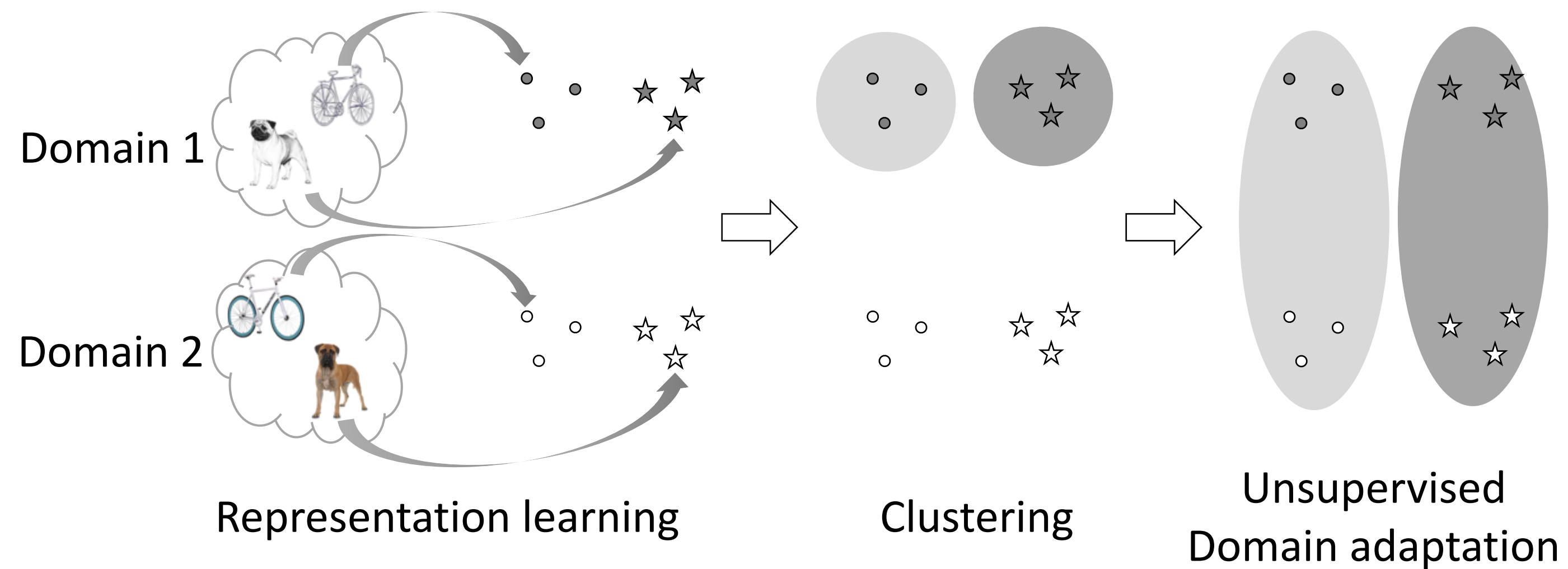
Problem formulation

Style-heterogeneous Domain Translation

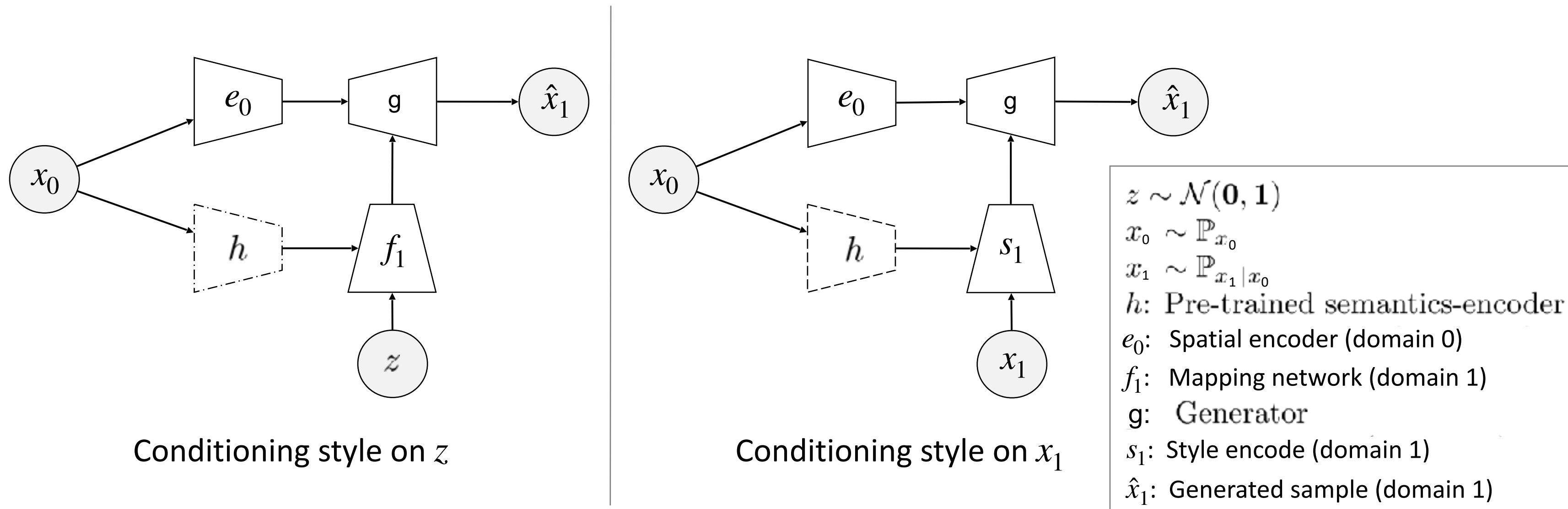
Sketches → Reals using StarGAN-V2



Domain invariant categorical semantics



Conditioning the style encoder



$$\mathcal{L}_{\text{sem}}^f := -\mathbb{E}_y \left[\mathbb{E}_{\mathbf{x}_y, \mathbf{z}} [h(\mathbf{x}_y) \log(h(g(e_y(\mathbf{x}_y), f_{\bar{y}}(h(\mathbf{x}_y), \mathbf{z})))))] \right],$$

$$\mathcal{L}_{\text{sem}}^s := -\mathbb{E}_y \left[\mathbb{E}_{\mathbf{x}_y} \mathbb{E}_{\mathbf{x}_{\bar{y}} \sim \mathbb{P}_{x_{\bar{y}}|h(\mathbf{x}_y)}} [h(\mathbf{x}_y) \log(h(g(e_y(\mathbf{x}_y), s_{\bar{y}}(h(\mathbf{x}_y), \mathbf{x}_{\bar{y}})))))] \right].$$

Results

MNIST↔SVHN

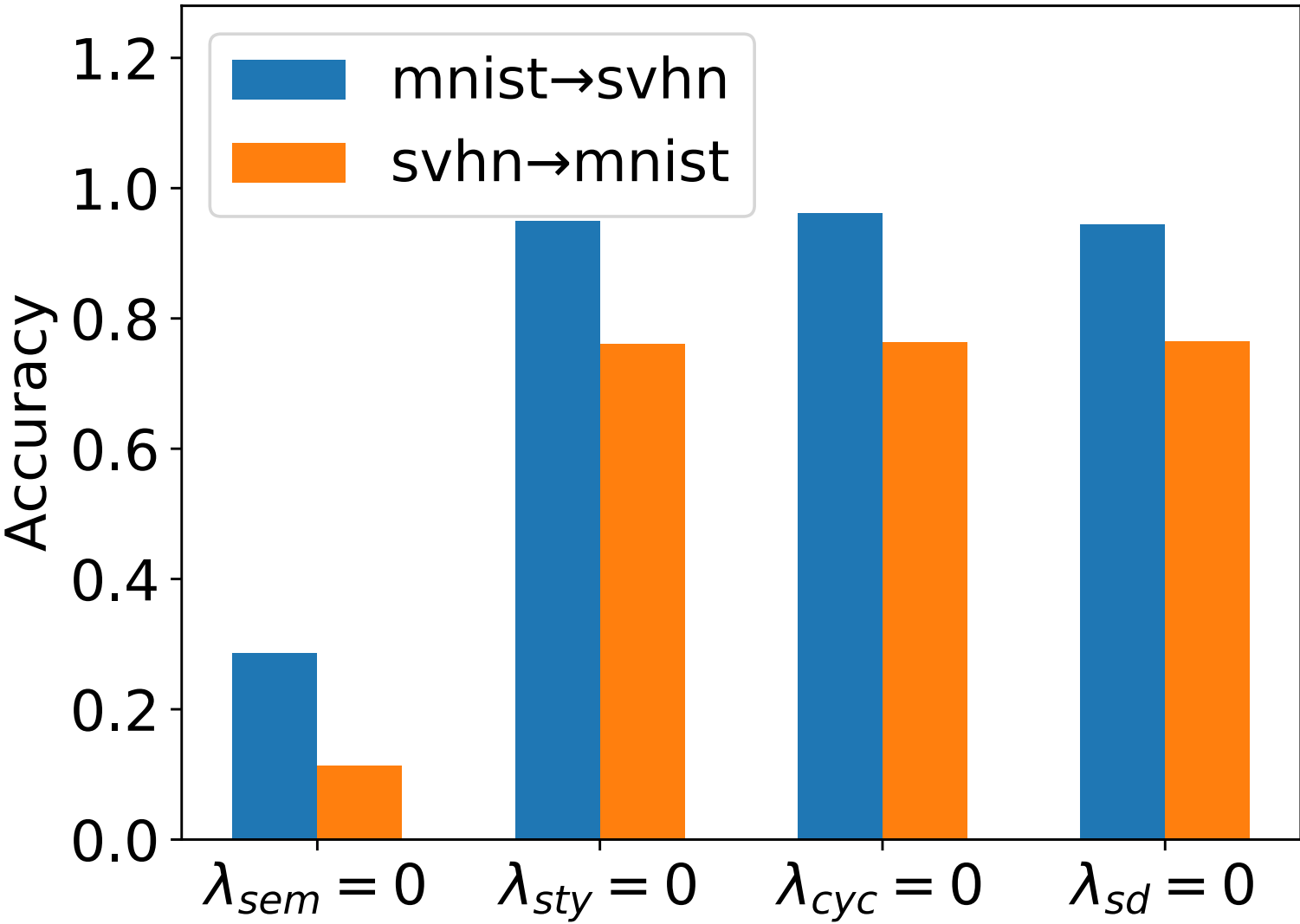
Table 1: **Comparison with the baselines.** Domain translation accuracy and FID obtained on MNIST (M) ↔SVHN (S) for the different methods considered. The last column is the test classification accuracy of the classifier used to compute the metric. *: Using weak supervision.

	Data	CycleGAN	MUNIT	DRIT	Stargan-V2	EGSC-IT*	CatS-UDT	Target
Acc	M→S	10.89	10.44	13.11	28.26	47.72	95.63	98.0
	S→M	11.27	10.12	9.54	11.58	16.92	76.49	99.6
FID	M→S	46.3	55.15	127.87	66.54	72.43	39.72	-
	S→M	24.8	30.34	20.98	26.27	19.45	6.60	-

MNIST→SVHN using our method



Ablating losses



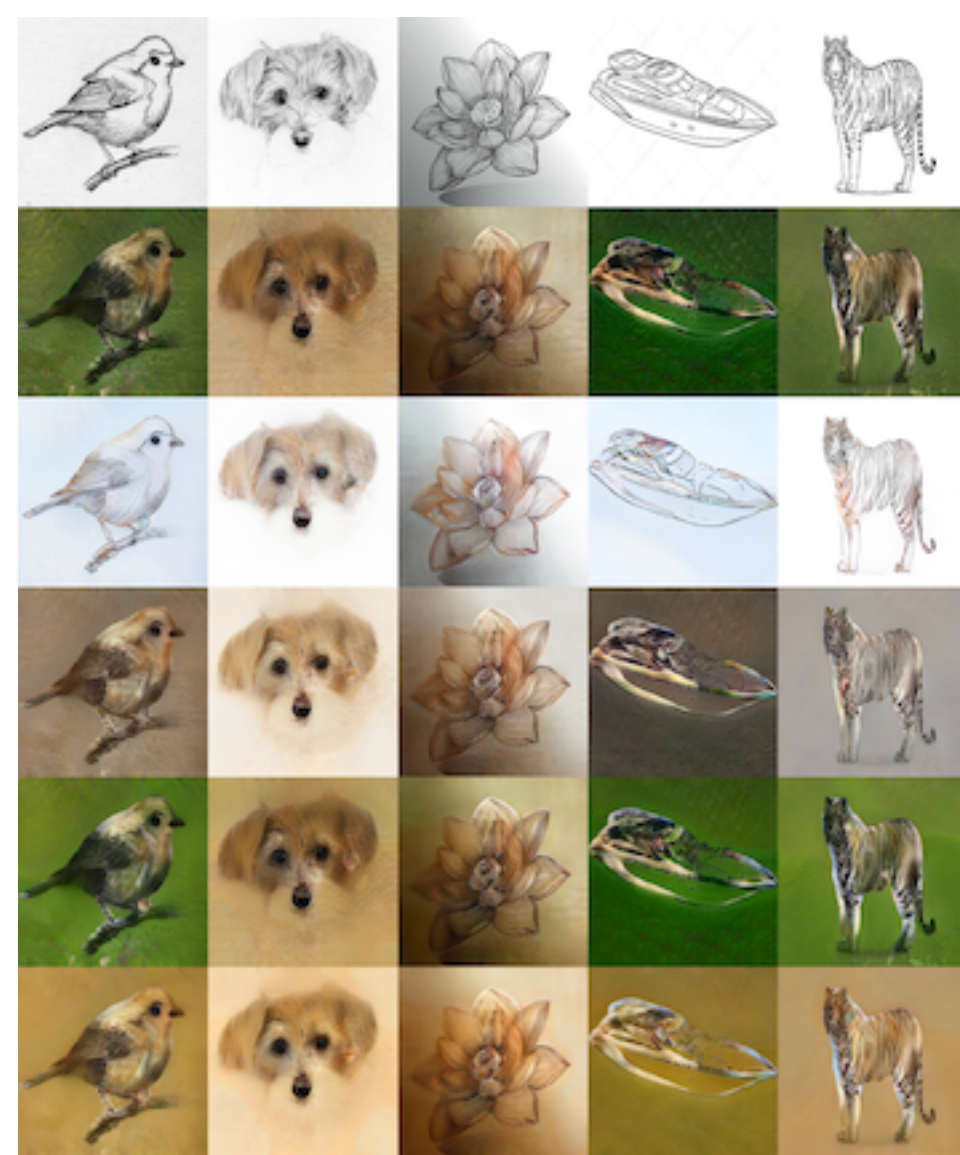
Results

Sketches→Reals

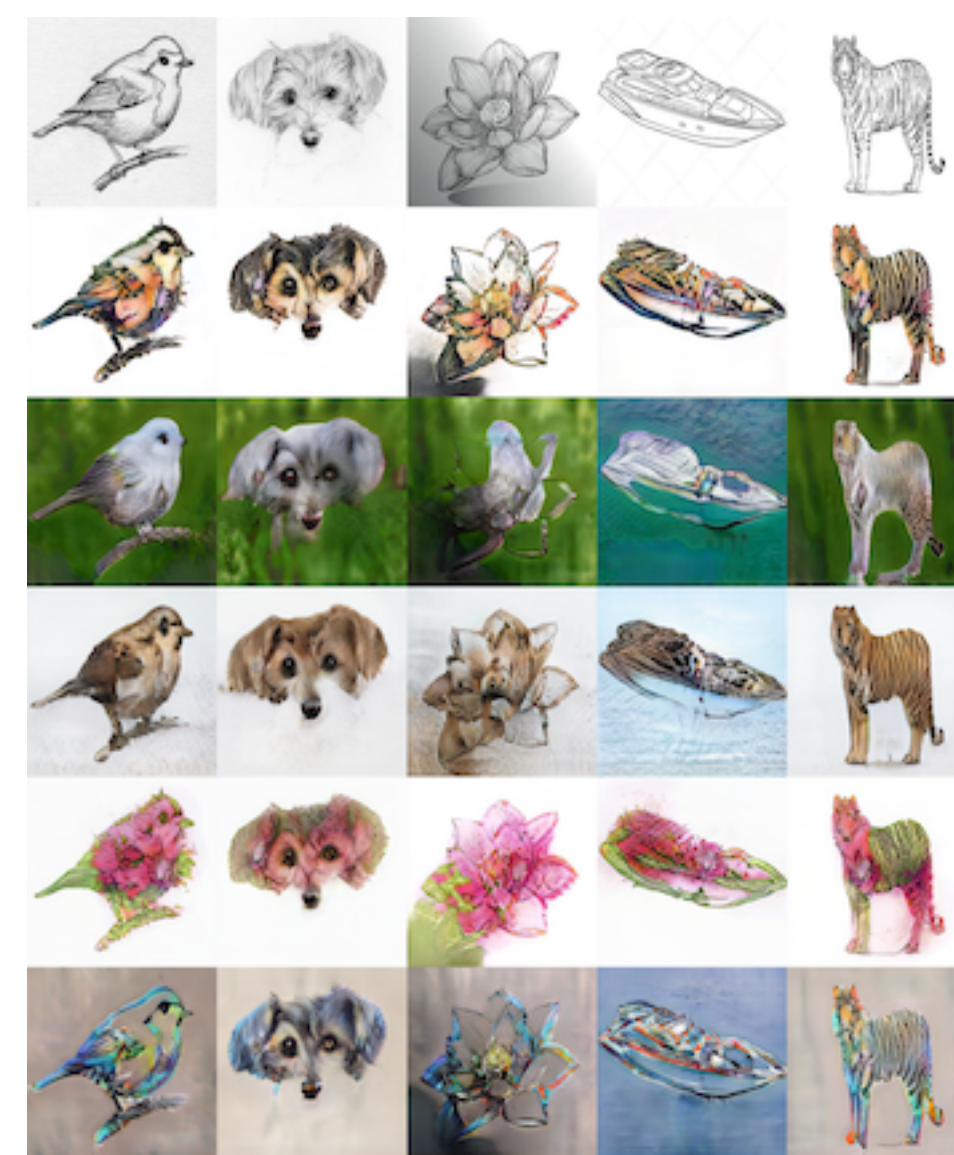
DRIT



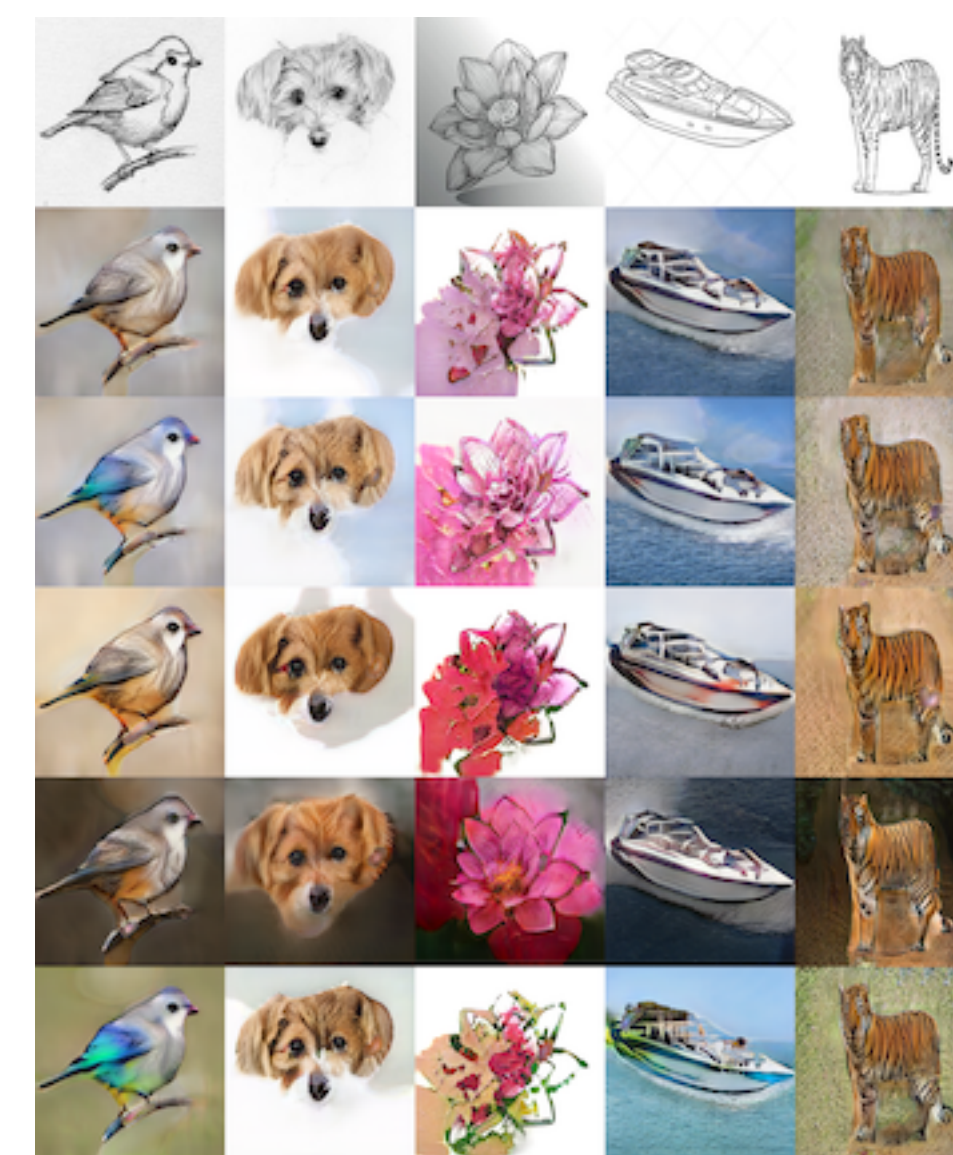
EGST-IT



StarGAN-V2



Cats-UDT (ours)



Thank you

- Paper: <https://openreview.net/forum?id=IMPA6MndSXU>
- Code: <https://github.com/lavoiems/Cats-UDT>