Associative Domain Adaptation
Supplementary Material

We provide additional information that is necessary to reproduce our results, as well as plots complementing the evaluation section of the main paper. To this end, we begin by stating implementation details for our neural network learning algorithm. Furthermore, we show additional t-SNE embeddings of source target domain for the different domain adaptation tasks analyzed in the paper.

1. Hyperparameters

We report the hyperparameters that we used for our experiments for the sake of reproducibility as detailed in Table 1.

2. t-SNE embeddings

We complement our analysis in Section 3.4.1 of the main document, *Qualitative evaluation: t-SNE embeddings*. In Figure 1 we show the t-SNE embeddings for all domain adaptation tasks that we have analyzed (cf. Table 3 of the main paper). The qualitative interpretation that we provide for the task Synthetic Digits to SVHN in the main paper is consistent across all tasks: when trained on source only, the target domain distribution is diffuse, the respective target classes can be visibly separated after domain adaptation and the separation is less clear when training with an MMD loss instead of our associative loss. Note that for the task Synthetic Signs to GTSRB, the target domain test error for the network trained on source only is already rather low. Subsequent domain adaptation improves the numerical result, which is, however, difficult to observe qualitatively due to the relatively small *coverage* compared to the previous settings.

Hyperparameter	Domains (source → target)			
	MNIST → MNIST-M	Syn. Digits → SVHN	SVHN → MNIST	Syn. Signs → GTSRB
New width/height	32	-	32	-
Source domain batch size	1000	1000	1000	1032
Target domain batch size	1000	1000	1000	1032
Learning rate decay steps	9000	9000	9000	9000
Visit loss weight	0.6	0.2	0.2	0.1
Delay (steps) for $\mathcal{L}_{assoc}$	500	2000	500	0

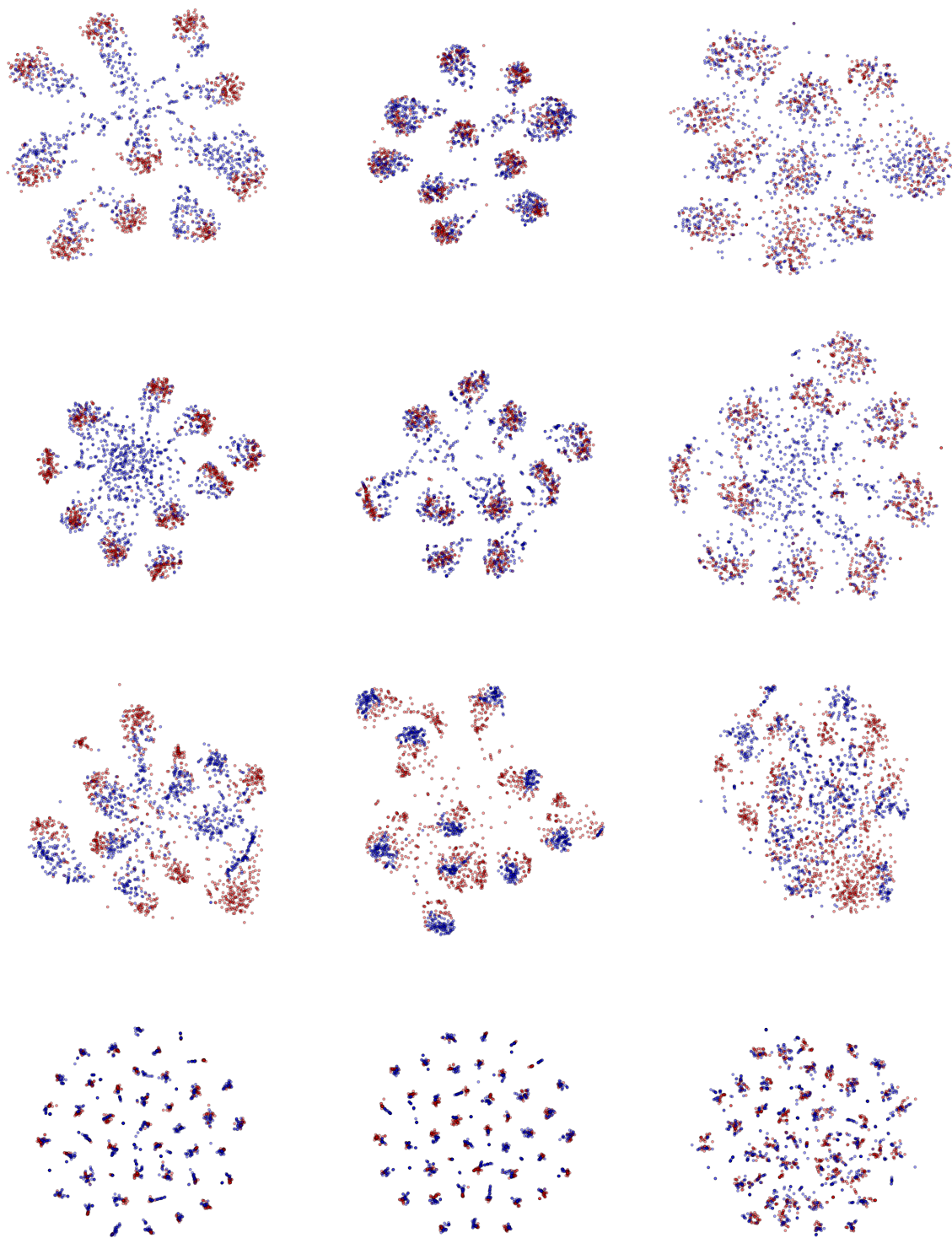Table 1: Hyperparameters for our domain adaptation experiments.

Figure 1: t-SNE embeddings of test samples for source (red) and target (blue). **First row:** MNIST to MNIST-M, perplexity 35. **Second row:** SVHN to MNIST, perplexity 35. **Third row:** Synthetic Signs to GTSRB, perplexity 25. 1,000 samples per domain, except for Synthetic Signs to GTSRB, where we took 60 samples for each of the 43 classes due to class imbalance in GTSRB. **Left**: After training on *source only*. **Middle**: after training with *associative domain adaptation* (DA_{assoc}). **Right**: after training with *MMD loss* (DA_{MMD}).