Supplemental Material for "Deep Multi-View Self-Paced Clustering"

method	MNIST	USPS	STL-10	CIFAR-10
SP-J-view-1	0.000406	0.000891	0.001628	0.000174
SP-J-view-2	0.000900	0.001694	0.000037	0.001009
SP-J-view-3	0.001203	0.000098	0.004481	0.000038
DMSC(ours)	**0.000172**	**0.000025**	**0.000009**	**0.000025**

Table 1: The variances of ACC on MNIST, USPS, STL-10 and CIFAR-10.

In the supplemental material, we present additional two ablation studies for the proposed DMSC, including sensitivity for initialization and the robustness for integrating different views.

A Ablation study

Sensitivity for initialization: Since the proposed DMSC contains the initialization stage, we investigate the sensitivity for initialization to validate whether or not the proposed model can achieve stable performance in most cases. To this end, we report the variances of the ACC for both DMSC and the single-view baseline **SP-J-View-i** on some grayscale and three-channel datasets, including MNIST, USPS, STL-10 and CIFAR-10. We run all the methods for 10 times to obtain the variance, and the results are given in Table 1. It shows that relatively low variances are kept by DMSC, demonstrating that the proposed model has good stability.

Robustness for integrating different views: In the above experiments, the number of views for the proposed DMSC is fixed to 3. To investigate the impact of the number of views on our model, we build our model with two, three and four network branches, where the configurations are (Conv-VAE + SAE), (CAE + Conv-VAE + SAE), and (CAE + Conv-VAE + SAE + VAE), respectively. Due to the limitation of the space, we only report the results of **SP-J-view-i** and our DMSC on USPS, which are shown in Table 2.

When compared with **SP-J-view-1** and **SP-J-view-3**, we can find that the two-view DMSC can beat these two single-view baselines (75.86% vs 69.35% and 72.80% for ACC), which demonstrates that in DMSC, the complementary information between **SP-J-view-1** and **SP-J-view-3** can be well exploited, therefore the resulting performance will be better. Moreover, comparing with three-view DMSC, which is added with a well-performed single-view baseline **SP-J-view-2**, the

method	ACC	NMI	ARI
SP-J-view-1 (CAE)	0.6935	0.7267	0.6255
SP-J-view-2 (Conv-VAE)	0.7334	0.7562	0.6746
SP-J-view-3 (SAE)	0.7280	0.7118	0.6276
SP-J-view-4 (VAE)	0.7159	0.7047	0.6238
DMSC (CAE + SAE)	0.7586	0.7601	0.6840
DMSC (CAE + Conv-VAE + SAE)	**0.7729**	**0.7811**	**0.7104**
DMSC (CAE + Conv-VAE + SAE + VAE)	0.7696	0.7683	0.7000

Table 2: The clustering results with incorporating different types of views on USPS.

two-view DMSC also hold highly competetive performance. The four-view DMSC, which is added by a degenerative view **SP-J-view-4**, shows little performance decrease than three-view DMSC. These experimental results fully demonstrate the robustness of the proposed DMSC for incorporating different views.