

On Dataset Caltech (HOG, Gabor, and CENTRIST views), the ACC of single-view method DEC on the best view is 0.8820, and that on the worst view is 0.3775.

Fig. 1 shows the t -SEN [1] visualization of different views' representations learned by DSIMVC. One can observe that the CENTRIST view's performance is improved, but at the cost of HOG view's performance being reduced. The result is that the final ACC of DSIMVC (0.8629) is not better than the ACC of DEC on the best view (0.8820).

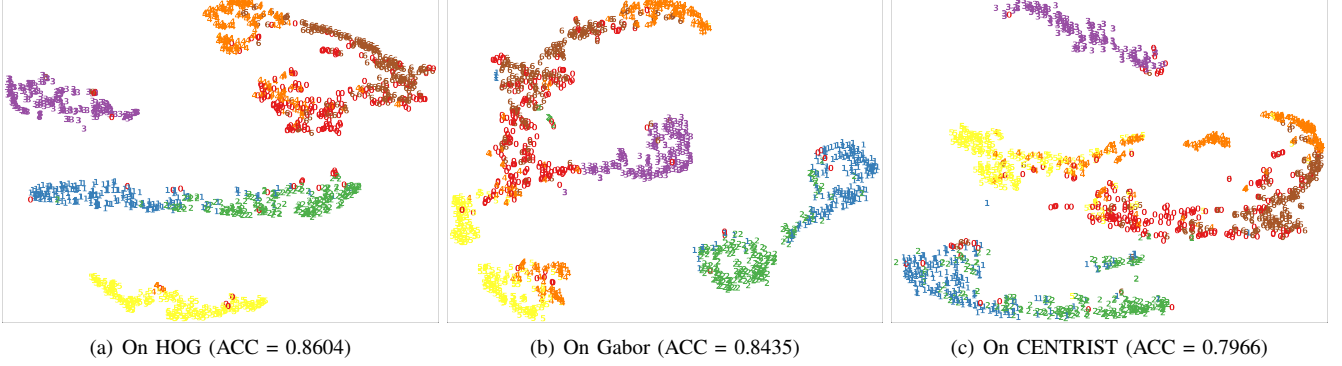


Fig. 1. t -SEN visualization of different views' representations learned by DSIMVC, whose final ACC is 0.8629.

Fig. 2 shows the t -SEN visualization of different views' representations learned by MvCAN. We can observe that the unclear cluster structures and noise in CENTRIST view have almost no negative impact on HOG view. This is because MvCAN has a decoupled framework and its clustering objective leverages un-shared parameters and inconsistent clustering predictions to avoid the side effects of noisy views. MvCAN further generates robust learning targets for mining multiple views' useful information, thus its final ACC (0.9160) is better than the ACC of DEC on the best view (0.8820).

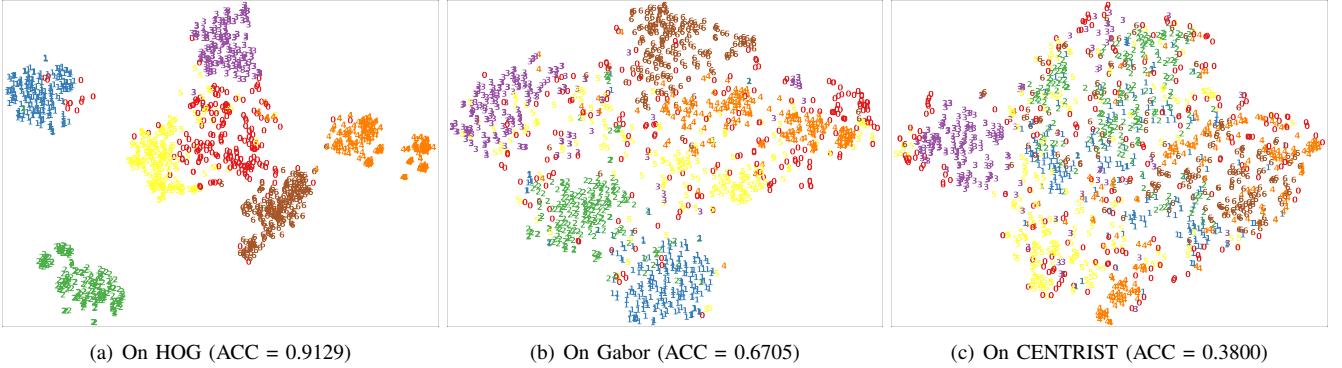


Fig. 2. t -SEN visualization of different views' representations learned by MvCAN, whose final ACC is 0.9160.

REFERENCES

- [1] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t -SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.