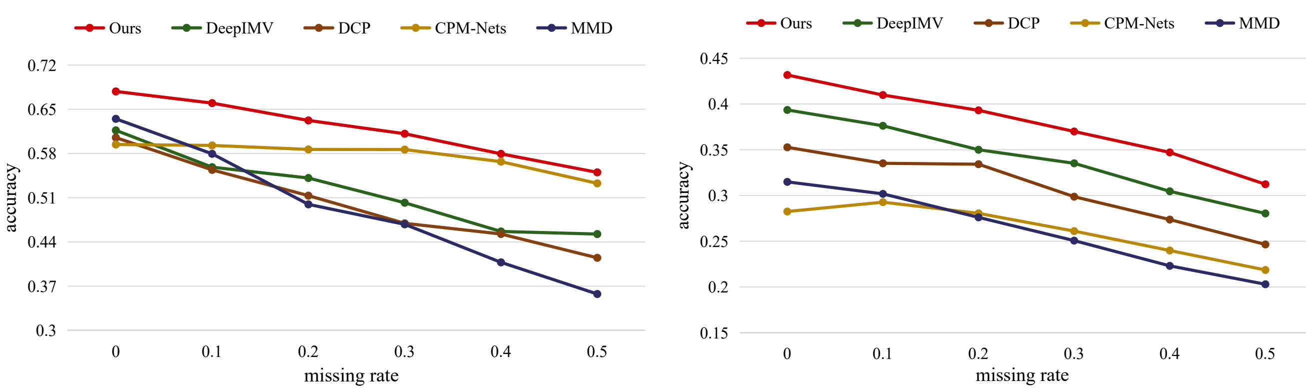
Supplementary Material for
Exploring and Exploiting Uncertainty for Incomplete Multi-View Classification

Mengyao Xie[‡], Zongbo Han[‡], Changqing Zhang[*], Yichen Bai, Qinghua Hu
College of Intelligence and Computing, Tianjin University
zhangchangqing@tju.edu.cn

1. Large-Scale Multi-view Datasets Experiments

To validate the effectiveness of the proposed method on large-scale datasets, we conduct experiments on two datasets. **Caltech-101** [1] is a 6-view dataset consists of pictures of objects belonging to 101 classes, plus one background clutter class. Each class contains roughly 40 to 800 images, totalling 9144 images. **NUS-WIDE-OBJECT** [2] is a 5-view dataset consists of pictures belonging to 31 classes, totally 30,000 images.

We compare the proposed algorithm with 4 state-of-the-art multi-view classification methods. **CPM-Nets** [3] directly learns the joint latent representations for all views with available data, and maps the latent representation to classification predictions. **DeepIMV** [4] applies the information bottleneck (IB) framework to obtain marginal and joint representations with the available data, and constructs the view-specific and multi-view predictors to obtain the classification predictions. **MMD** [5] models both the feature-level and modality-level dynamicities and introduces a sparse gating strategy to trustworthily fuse the complete multi-view data. **DCP** [6] provides an information theoretical framework under which the consistency learning and data recovery are treated as a whole. The missing views are recovered by minimizing the conditional entropy through dual prediction. For the MMD method which can not deal with incomplete multi-view data, we first impute the missing views with the corresponding means, and then train MMD to obtain the classification predictions. As shown in Fig. S1, our method still achieves the superior performance on large-scale datasets.



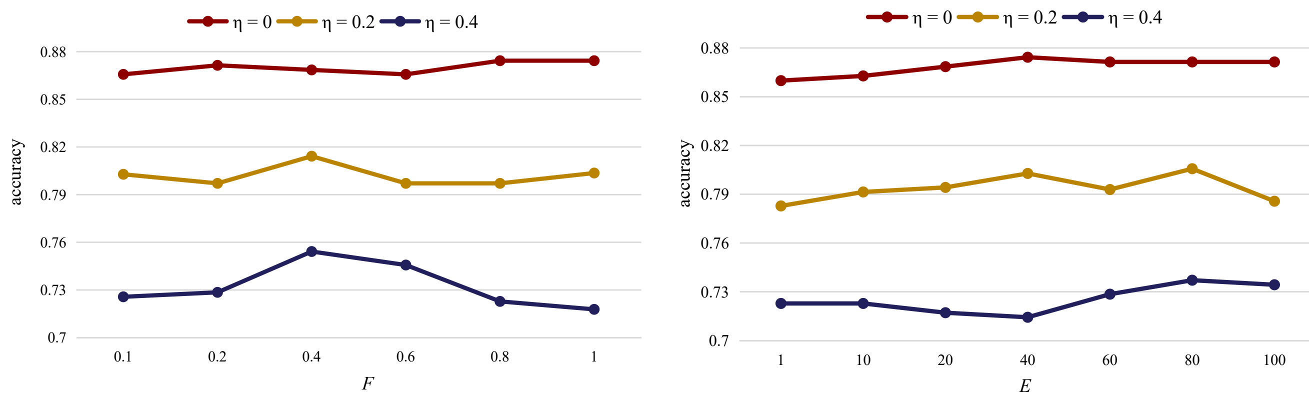
(a) Caltech-101 of 9144 samples

(b) NUS-WIDE-OBJECT of 30,000 samples

Figure S1. Classification performance with $\eta = [0, 0.1, 0.2, 0.3, 0.4, 0.5]$.

2. Details and Impact of Hypeparameter λ

Details. Same as previous methods [7, 8], we set $\lambda = min(F, e/E)$ where F, e and E denote the final value of λ, iteration epoch and the parameter controlling the decay rate of λ. ***Impact.*** we conduct experiments on ROSMAP with different missing rates η to investigate the influence of λ in terms of decay rate and the final value. Specifically, we change the final value F or parameter E to control the decay rate. As shown in Fig. S2, our method is robust to the final value and decay rate of λ.



(a) Final value

(b) Decay rate

Figure S2. Classification performance with different final values and decay rates of λ on ROSMAP under $\eta = [0, 0.2, 0.4]$ where we set $E = 50$ in Fig. S2a and $F = 1$ in Fig. S2b.

3. Details of Constructing the Incomplete Multi-view Data.

Same as previous methods [3, 4], we construct the incomplete multi-view data to satisfy the following requirements. 1. Missing at random with the desired missing rate η. 2. Each sample has at least one view that is not missing. Therefore, the following steps are conducted. 1. Calculate how many views are missing with $N^m = \eta NV$, where N and V are number of samples and views respectively. 2. Randomly assign N^m to each sample while ensure that at least one view is available for each sample (i.e., $N^m = \sum_{i=1}^{N} N_i^m$ and $N_i^m < V$, where N_i^m is the number assigned to each sample). 3. Randomly select $V - N_i^m$ available views for each sample.

References

[1] Li Fei-Fei, Rob Fergus, and Pietro Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. In *Conference on Computer Vision and Pattern Recognition Workshops*, page 178, 2004. 1

[2] Tat-Seng Chua, Jinhui Tang, Richang Hong, Haojie Li, Zhiping Luo, and Yan-Tao Zheng. Nus-wide: A real-world web image database from national university of singapore. In *Proceedings of the ACM International Conference on Image and Video Retrieva*, 2009. 1

[3] Changqing Zhang, Zongbo Han, Yajie Cui, Huazhu Fu, Joey Tianyi Zhou, and Qinghua Hu. Cpm-nets: Cross partial multi-view networks. In *Advances in Neural Information Processing Systems*, pages 557–567, 2019. 1, 2

[4] Changhee Lee and Mihaela van der Schaar. A variational information bottleneck approach to multi-omics data integration. In *The 24th International Conference on Artificial Intelligence and Statistics*, 2021. 1, 2

[5] Zongbo Han, Fan Yang, Junzhou Huang, Changqing Zhang, and Jianhua Yao. Multimodal dynamics: Dynamical fusion for trustworthy multimodal classification. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20675–20685, 2022. 1

[6] Yijie Lin, Yuanbiao Gou, Xiaotian Liu, Jinfeng Bai, Jiancheng Lv, and Xi Peng. Dual contrastive prediction for incomplete multi-view representation learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):4447–4461, 2023. 1

[7] Murat Sensoy, Lance M. Kaplan, and Melih Kandemir. Evidential deep learning to quantify classification uncertainty. *Advances in neural information processing systems*, 31:3183–3193, 2018. 1

[8] Zongbo Han, Changqing Zhang, Huazhu Fu, and Joey Tianyi Zhou. Trusted multi-view classification with dynamic evidential fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 1