清華大學 Tsinghua University

Dynamic Hypergraph Structure Learning

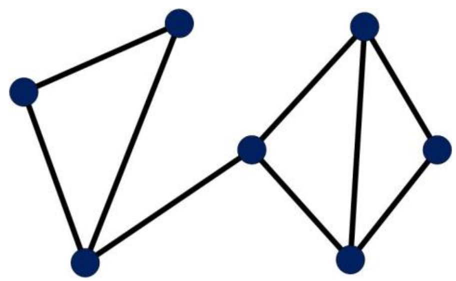
Zizhao Zhang , Yubo Zhang, Xibin Zhao, and Yue Gao

School of Software, Tsinghua University, China.

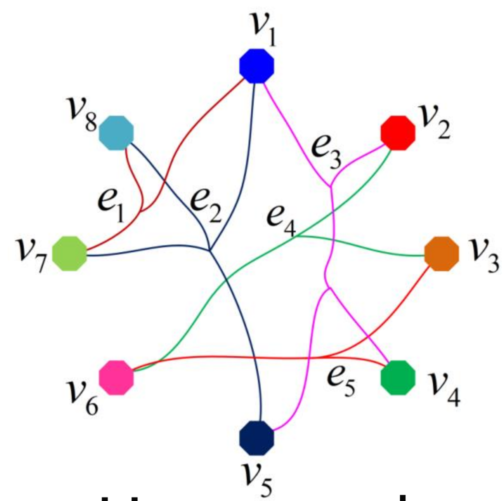
Contents

Research Background

Correlation Modelling



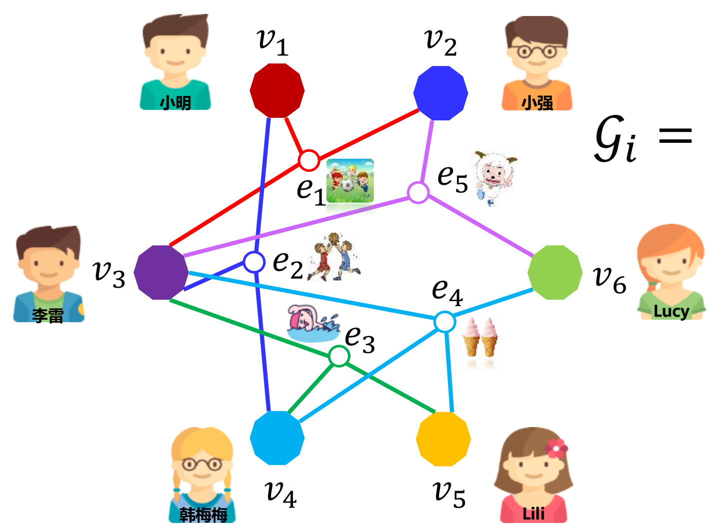
Graph



Hypergraph

Example

Name	foot-ball	baske-tball	swim-ming	sweet	carton
Ming	✓	✓	✗	✗	✗
Qiang	✓	✗	✗	✗	✓
Lei	✓	✓	✓	✓	✓
Mei	✗	✓	✓	✓	✗
Lili	✗	✗	✓	✓	✗
Lucy	✗	✗	✗	✓	✓

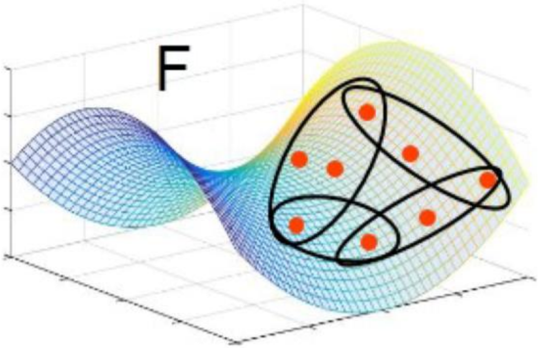


$$\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i, W_i)$$

each vertex represents an object

Research Background

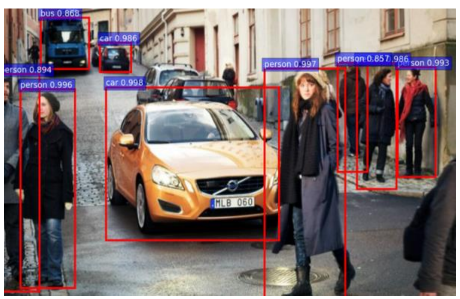
Hypergraph Learning



Hypergraph laplacian Δ

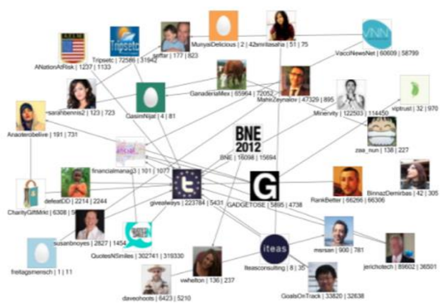
$$\arg\min_{\mathbf{F}} \left\{ \mathrm{tr}\left(\mathbf{F}^T \mathbf{\Delta} \mathbf{F}\right) + \lambda \|\mathbf{F} - \mathbf{Y}\|_F^2 \right\}$$

recognition



[An et al. , 2017]

sorting

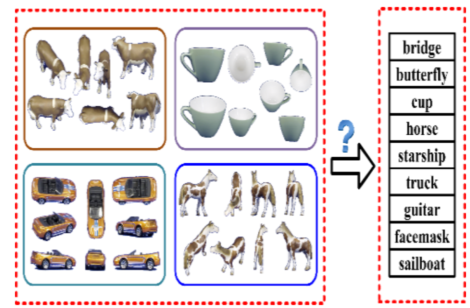


[Zhu et al. , 2015]

segmentation

[Huang et al. , 2009]

classification[4]

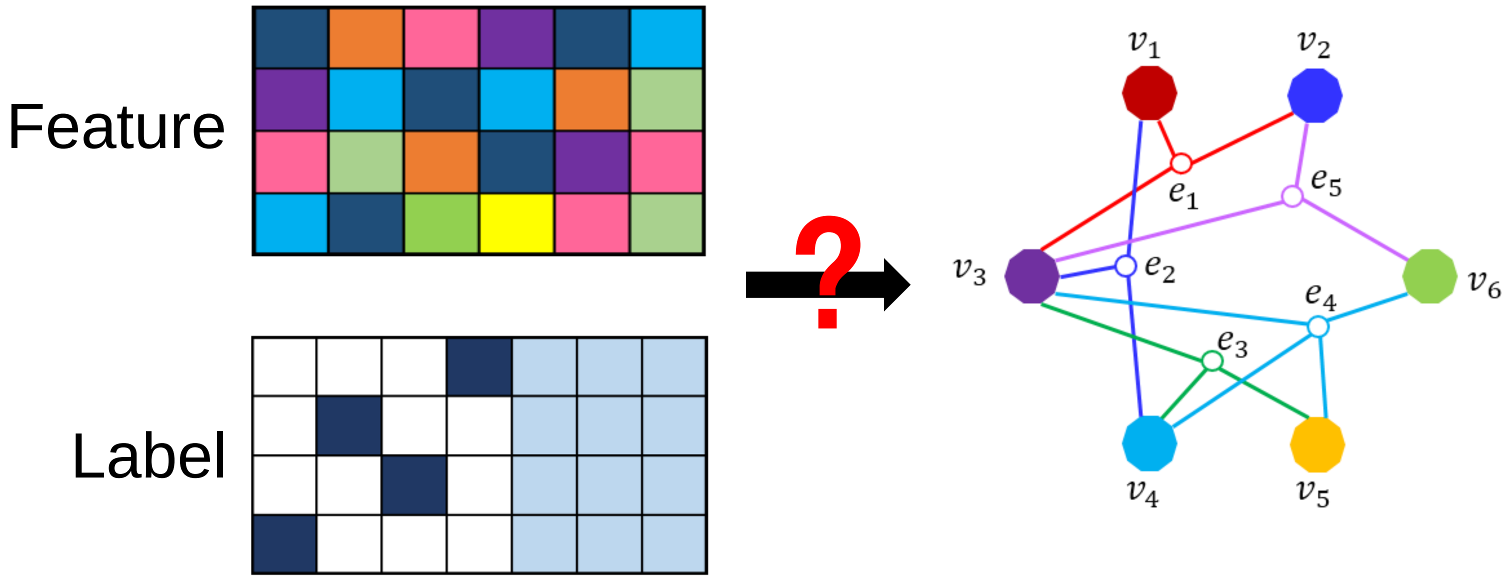


[Gao et al. , 2012]

A well constructed hypergraph structure can represent the data correlation accurately, yet leading to better performance.

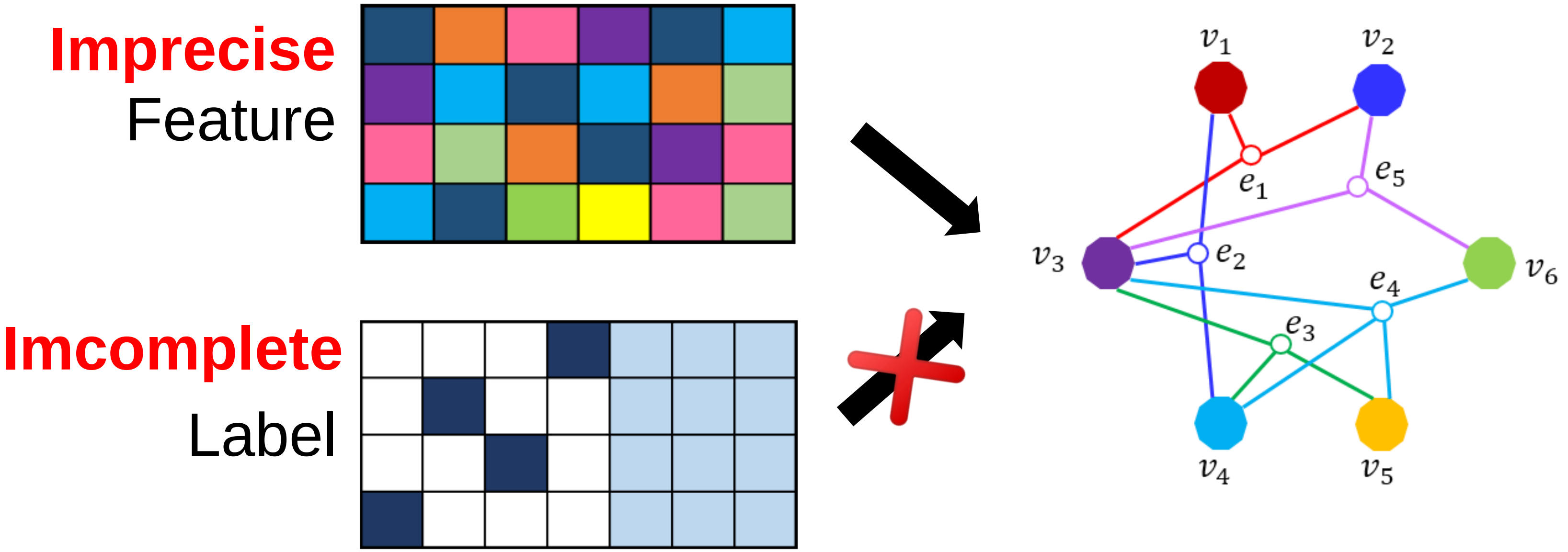
Research Background

Hypergraph Construction



Research Background

Hypergraph Construction



- ☐ k-nn method *(Huang et al., 2009)*
- ☐ clustering-based method *(Gao et al., 2012)*
- ☐ spare representation method *(Wang et al., 2015)*

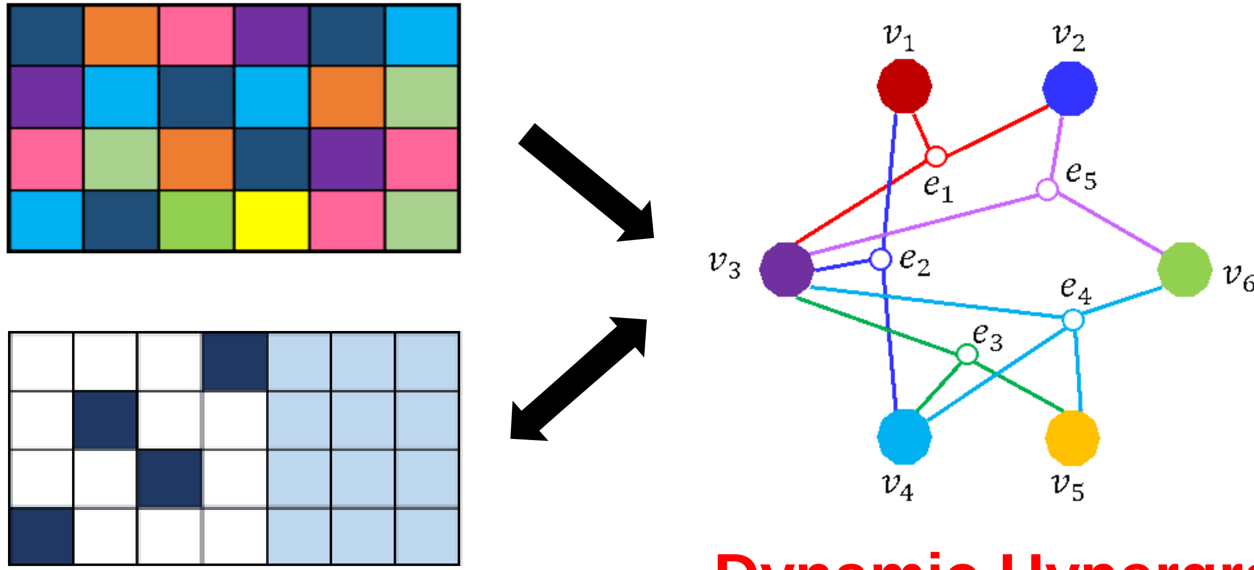
A static hypergraph structure cannot represent the data correlation accurately.

Motivation

Hypergraph Construction

Feature

Label

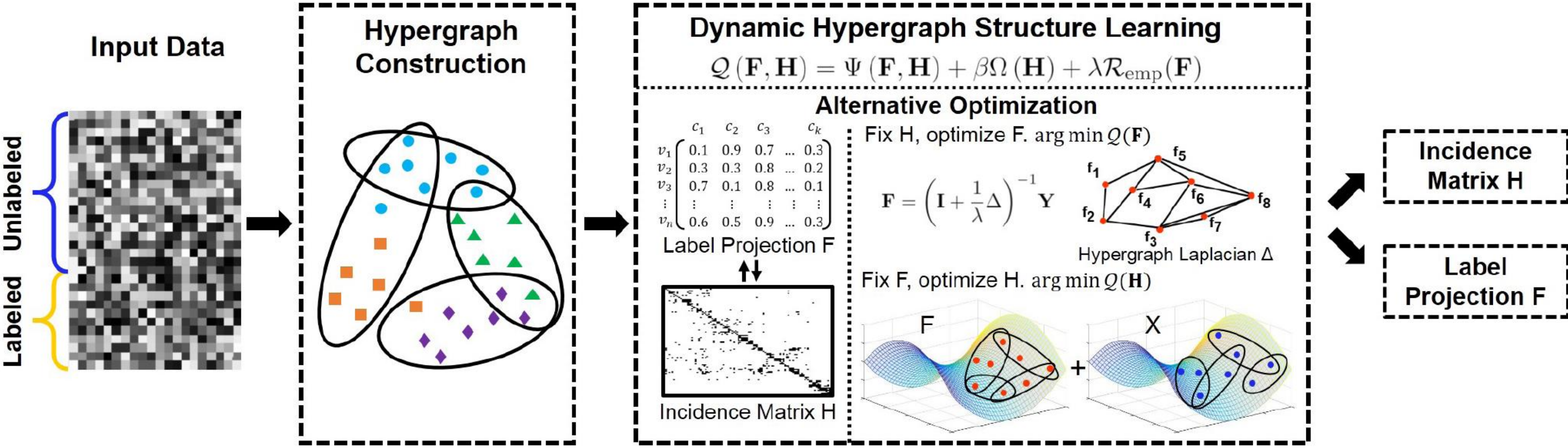


Gradually complete

Dynamic Hypergraph Structure

- construct the hypergraph structure using both the feature information and the label information
- gradually complete the label of all data and dynamically adapt hypergraph structure simultaneouly

Dynamic Hypergraph Structure Learning



- jointly learn the hypergraph structure and the label projection matrix
- optimize the hypergraph structure from both the data label space and the data feature space

The Formulation of DHSL

Cost function:

$$\arg\min_{\mathbf{F},0\preceq\mathbf{H}\preceq 1} \mathcal{Q}(\mathbf{F},\mathbf{H}) = \Psi(\mathbf{F},\mathbf{H}) + \beta\Omega(\mathbf{H}) + \lambda\mathcal{R}_{\mathrm{emp}}(\mathbf{F})$$

The objective should satisfy three conditions.

1. **The label projection matrix F should be smooth on H.**
2. H should be smooth on the data from both label space and feature space.
3. The empirical loss

$$\begin{aligned}
&\Psi(\mathbf{F},\mathbf{H}) \\
&= \frac{1}{2}\sum_{c=1}^{n_c}\sum_{e\in\mathcal{E}}\sum_{u,v\in\mathcal{V}} \frac{\mathbf{W}(e)\mathbf{H}(u,e)\mathbf{H}(v,e)}{\delta_i(e)}\left(\frac{\mathbf{F}(u,c)}{\sqrt{d(u)}} - \frac{\mathbf{F}(v,c)}{\sqrt{d(v)}}\right)^2 \\
&= \sum_{c=1}^{n_c}\sum_{e\in\mathcal{E}}\sum_{u,v\in\mathcal{V}} \frac{\mathbf{W}(e)\mathbf{H}(u,e)\mathbf{H}(v,e)}{\delta(e)}\left(\frac{\mathbf{F}(u,c)^2}{d(u)} - \frac{\mathbf{F}(u,c)\mathbf{F}(v,c)}{\sqrt{d(u)d(v)}}\right) \\
&= \sum_{c=1}^{n_c}\sum_{u\in\mathcal{V}}\mathbf{F}(u,c)^2\sum_{e\in\mathcal{E}}\frac{\mathbf{W}(e)\mathbf{H}(u,e)}{d(u)}\sum_{v\in\mathcal{V}}\frac{\mathbf{H}(v,e)}{\delta(e)} \\
&- \sum_{e\in\mathcal{E}}\sum_{u,v\in\mathcal{V}}\frac{\mathbf{F}(u,c)\mathbf{H}(u,e)\mathbf{W}(e)\mathbf{H}(v,e)\mathbf{F}(v,c)}{\sqrt{d(u)d(v)\delta(e)}} \\
&= \mathrm{tr}\left(\left(\mathbf{I} - \mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{D}_v^{-\frac{1}{2}}\right)\mathbf{F}\mathbf{F}^{\mathrm{T}}\right),
\end{aligned}$$

The Formulation of DHSL

Cost function:

$$\arg\min_{\mathbf{F},0\preceq\mathbf{H}\preceq 1} \mathcal{Q}(\mathbf{F},\mathbf{H}) = \Psi(\mathbf{F},\mathbf{H}) + \beta\Omega(\mathbf{H}) + \lambda\mathcal{R}_{\mathrm{emp}}(\mathbf{F})$$

The objective should satisfy three conditions.

1. The label projection matrix F should be smooth on H.
2. **H should be smooth on the data from both label space and feature space.**
3. The empirical loss

$$\boxed{\Omega(\mathbf{H}) = \mathrm{tr}\left(\left(\mathbf{I} - \mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{D}_v^{-\frac{1}{2}}\right)\mathbf{X}\mathbf{X}^{\mathrm{T}}\right)}$$

The Formulation of DHSL

Cost function:

$$\arg\min_{\mathbf{F},0\preceq\mathbf{H}\preceq 1} \mathcal{Q}(\mathbf{F},\mathbf{H}) = \Psi(\mathbf{F},\mathbf{H}) + \beta\Omega(\mathbf{H}) + \lambda\mathcal{R}_{\mathrm{emp}}(\mathbf{F})$$

The objective should satisfy three conditions.

1. The label projection matrix F should be smooth on H.
2. H should be smooth on the data from both label space and feature space.
3. **The empirical loss**

$$\boxed{\mathcal{R}_{\mathrm{emp}}(\mathbf{F}) = \|\mathbf{F}-\mathbf{Y}\|_F^2}$$

Optimization

Cost function:

$$\arg\min_{\mathbf{F},0\preceq\mathbf{H}\preceq 1} \mathcal{Q}(\mathbf{F},\mathbf{H}) = \Psi(\mathbf{F},\mathbf{H}) + \beta\Omega(\mathbf{H}) + \lambda\mathcal{R}_{\mathrm{emp}}(\mathbf{F})$$

$$= \mathrm{tr}((I - \mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{D}_v^{-\frac{1}{2}})(\mathbf{F}\mathbf{F}^{\mathrm{T}} + \beta\mathbf{X}\mathbf{X}^{\mathrm{T}})) + \lambda \parallel \mathbf{F} - \mathbf{Y} \parallel_F^2$$

Fix H, optimize $\mathcal{F}$

$$\arg\min_{\mathbf{F}} \mathcal{Q}(\mathbf{F}) = \Psi(\mathbf{F}) + \lambda\mathcal{R}_{\mathrm{emp}}(\mathbf{F}) = \mathrm{tr}\left(\boldsymbol{\Delta}\mathbf{F}\mathbf{F}^{\mathrm{T}}\right) + \lambda \parallel \mathbf{F} - \mathbf{Y} \parallel^2$$

$$\mathbf{F} = \left(\mathbf{I} + \frac{1}{\lambda}\Delta\right)^{-1}\mathbf{Y}$$

Optimization

Cost function:

$$\arg\min_{\mathbf{F},0\preceq\mathbf{H}\preceq 1} \mathcal{Q}(\mathbf{F},\mathbf{H}) = \Psi(\mathbf{F},\mathbf{H}) + \beta\Omega(\mathbf{H}) + \lambda\mathcal{R}_{\text{emp}}(\mathbf{F})$$

$$= \mathrm{tr}((I - \mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{D}_v^{-\frac{1}{2}})(\mathbf{F}\mathbf{F}^{\mathrm{T}} + \beta\mathbf{X}\mathbf{X}^{\mathrm{T}})) + \lambda \parallel \mathbf{F} - \mathbf{Y} \parallel_F^2$$

Fix $\mathcal{F}$, Optimize H

$$\arg\min_{0\preceq\mathbf{H}\preceq 1} \mathcal{Q}(\mathbf{H}) = \Psi(\mathbf{H}) + \beta\Omega(\mathbf{H}) = \mathrm{tr}\left(\left(I - \mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{D}_v^{-\frac{1}{2}}\right)\right)(\mathbf{F}^{\mathrm{T}} + \beta\mathbf{X}\mathbf{X}^{\mathrm{T}})$$

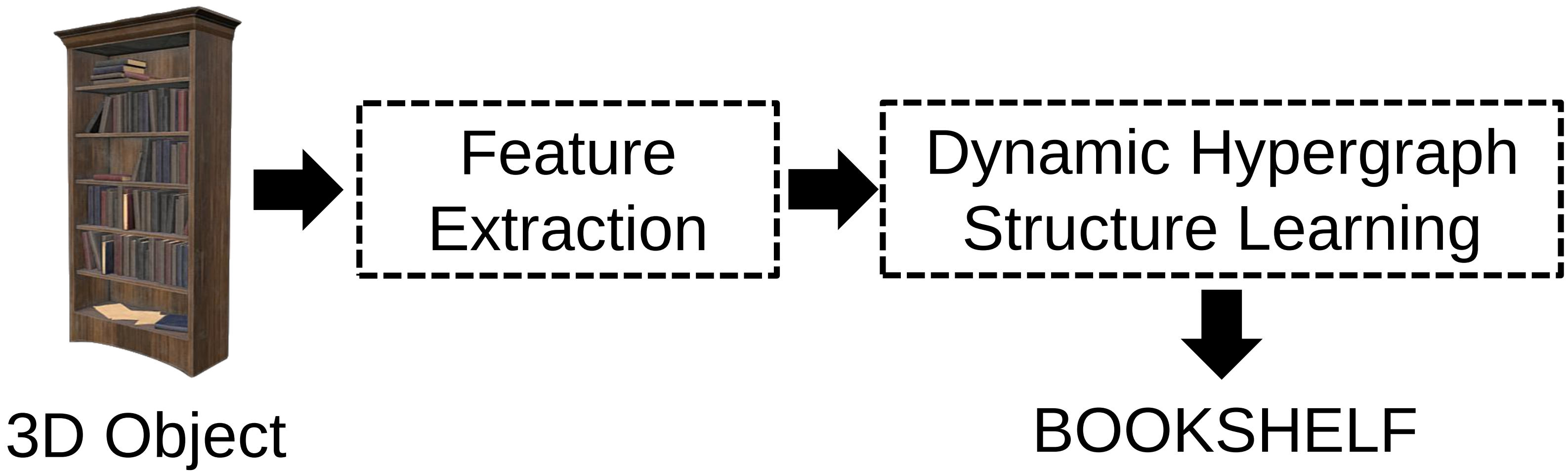
$$\nabla\mathcal{Q}(\mathbf{H}) = \mathbf{J}(\mathbf{I}\otimes \mathbf{H}^{\mathrm{T}}\mathbf{D}_v^{-\frac{1}{2}}\mathbf{K}\mathbf{D}_v^{-\frac{1}{2}}\mathbf{H})\mathbf{W}\mathbf{D}_e^{-2} + \mathbf{D}_v^{-\frac{3}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{D}_v^{-\frac{1}{2}}\mathbf{K}\mathbf{J}\mathbf{W} - 2\mathbf{D}_v^{-\frac{1}{2}}\mathbf{K}\mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}$$

projected gradient method

$$\mathbf{H}_{k+1} = \boldsymbol{P}[\mathbf{H}_k - \alpha\nabla\mathcal{Q}(\mathbf{H}_k)]$$

$$\boldsymbol{P}[h_{ij}] = \begin{cases} h_{ij} & \mathit{if}\; 0 \le h_{ij} \le 1 \\ 0 & \mathit{if}\; h_{ij} < 0 \\ 1 & \mathit{if}\; h_{ij} > 1 \end{cases}$$

3D Shape Recognition



Datasets

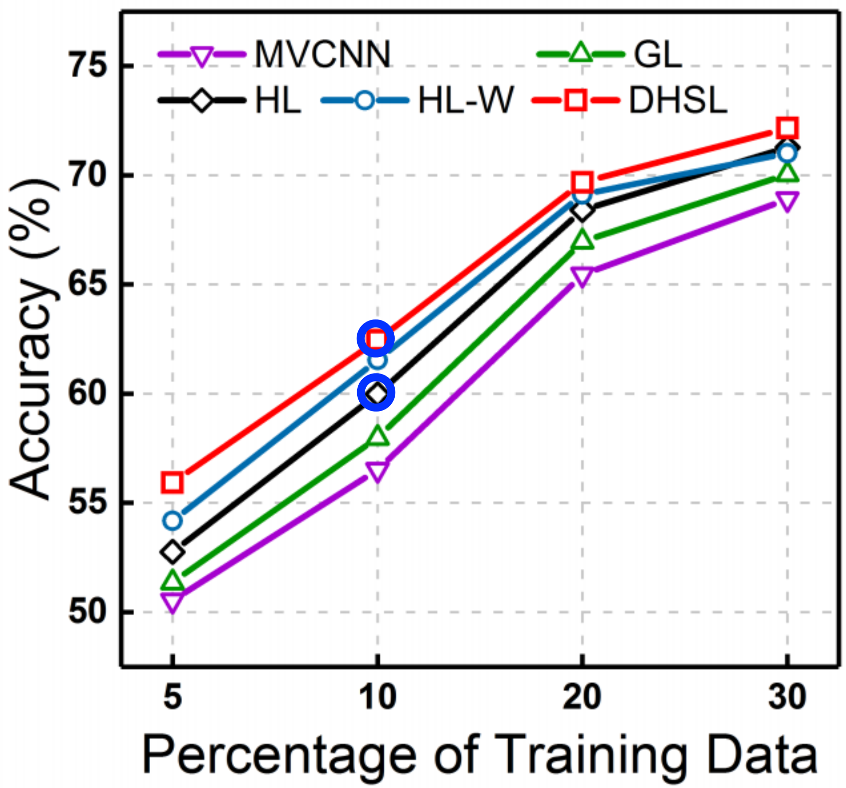
- The NTU dataset contains 2020 objects from 67 classes.
- The ESB dataset contains 866 objects from 43 classes.

State-of-the-art methods

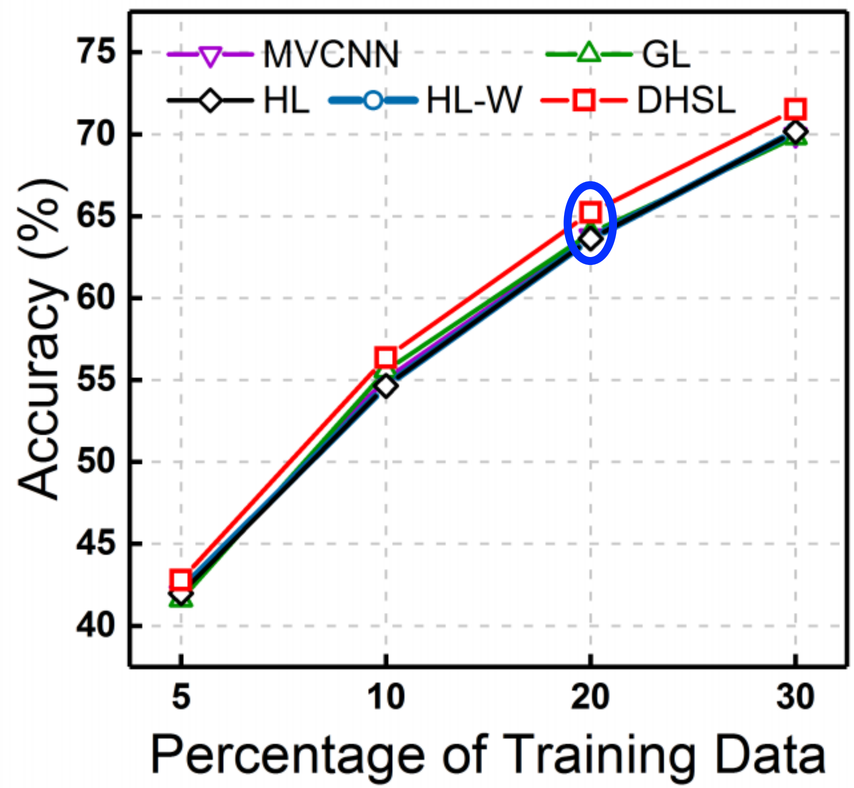
Multi-view Convolutional Neural Networks *[Su et al., 2015]*
Graph-based Learning *[Zhou et al., 2003]*
Traditional Hypergraph Learning *[Zhou et al., 2007]*
Hypergraph Learning with Hyperedge Weight Learning *[Gao et al., 2013]*
Dynamic Hypergraph Structure Learning (DHSL)

3D Shape Recognition

Experimental Results



(a) The NTU Dataset



(b) The ESB Dataset

Gesture Recognition

Feature Extraction → Dynamic Hypergraph Structure Learning → DRAW TICK

Depth Sequence

Datasets

- The MSRGesture3D dataset: 333 depth sequences from 12 classes.
- The Gesture3DMotion dataset: 384 depth sequences from 12 classes.

State-of-the-art methods

HON4D *[Oreifej and Liu, 2013]*

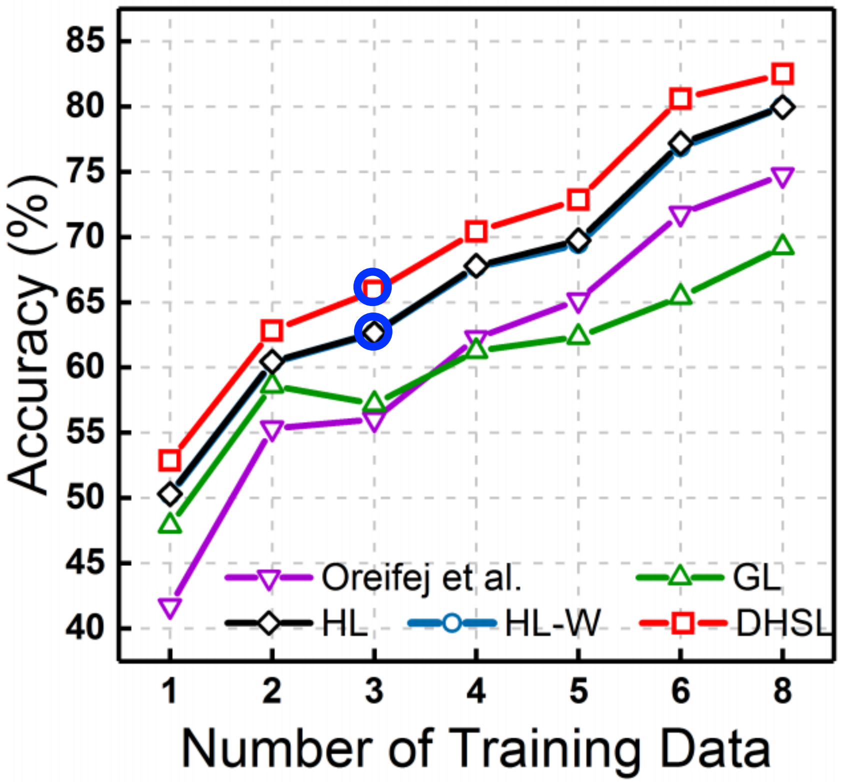
Graph-based Learning *[Zhou et al., 2003]*

Traditional Hypergraph Learning *[Zhou et al., 2007]*

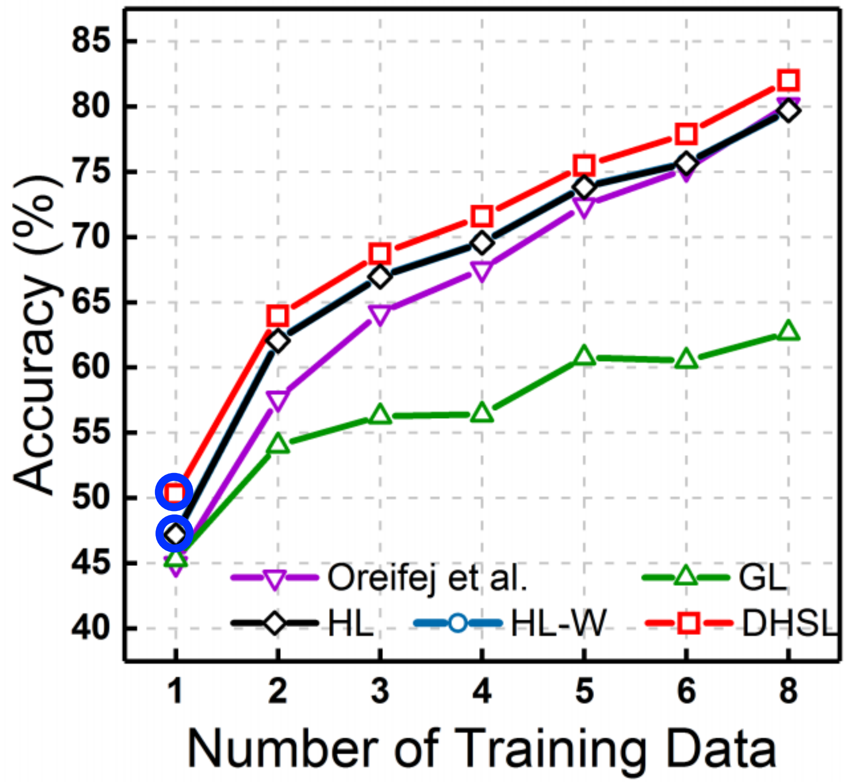
Hypergraph Learning with Hyperedge Weight Learning *[Gao et al., 2013]*

Dynamic Hypergraph Structure Learning (DHSL)

Gesture Recognition



(a) The MSRGesture3D Dataset



(b) The Gesture3DMotion Dataset

Conclusion

- We propose a **dynamic hypergraph structure learning algorithm** to jointly optimize the hypergraph structure and learn the label projection matrix.
- We apply the proposed method on the tasks of **3D shape recognition** and **gesture recognition**. The proposed method can achieve about **4%~6% improvement** compared with the traditional hypergraph learning method.

Reference

[1] Goldberger, J.; Gordon, S.; and Greenspan, H. 2003. An efficient image similarity measure based on approximations of kl-divergence between two gaussian mixtures. *In Proceedings of IEEE International Conference on Computer Vision*, 487. IEEE.

[2] Kusner, M.; Sun, Y.; Kolkin, N.; and Weinberger, K. 2015. From word embeddings to document distances. *In Proceedings of International Conference on Machine Learning*, 957–966.

[3] Li, P.;Wang, Q.; and Zhang, L. 2013. A novel earth mover's distance methodology for image matching with gaussian mixture models. *In Proceedings of IEEE International Conference on Computer Vision*, 1689–1696.

[4] Rubner, Yossi, Carlo Tomasi, and Leonidas J. Guibas. 2000. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision* 40(2): 99-121.

[5] Denny Zhou, Jiayuan Huang, and Bernhard Scholkopf. Learning with Hypergraphs: Clustering, Classification, and Embedding. *Advances in Neural Information Processing Systems 19*, pages 1601–1608, 2007.

[6] Dengyong Zhou, Olivier Bousquet, Thomas Navin Lal, Jason Weston, and Bernhard Scholkopf. Learning with Local and Global Consistency. *Proceedings of the 16th International Conference on Neural Information Processing Systems*, 2003.

[7] Y. Gao, M. Wang, Z. J. Zha, J. Shen, X. Li, and X. Wu. Visual-textual Joint Relevance Learning for Tagbased Social Image Search. *IEEE Transactions on Image Processing*, 22(1):363–376, 2013.

[8] Wang, R.; Guo, H.; Davis, L. S.; and Dai, Q. 2012. Covariance discriminative learning: A natural and efficient approach to image set classification. *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2496–2503. IEEE.

清華大學
Tsinghua University

Thanks!

zzz_14@126.com