EFFICIENT SEMI-SUPERVISED ANNOTATION WITH PROXY-BASED LOCAL CONSISTENCY PROPAGATION

Lei Huang, Yang Wang†, Xianglong Liu*, Bo Lang**

*State Key Lab of Software Development Environment, Beihang University, Beijing, China

†National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing, China

1. Introduction

Background

- Automatic image annotation is an effective solution to manage images which increases tremendously.
- Semi-Supervised Learning (SSL), are promising to build more accurate models.
- Many graph-based SSL methods , which involves graph construction and label propagation, have been applied to image or video annotation.

Main Issues

- Most of Graph-based SSL don't consider the difference between the labeled and unlabeled data when learning the manifold.
- Most of Graph-based SSL face the limitation that learning must be performed in a batch mode.

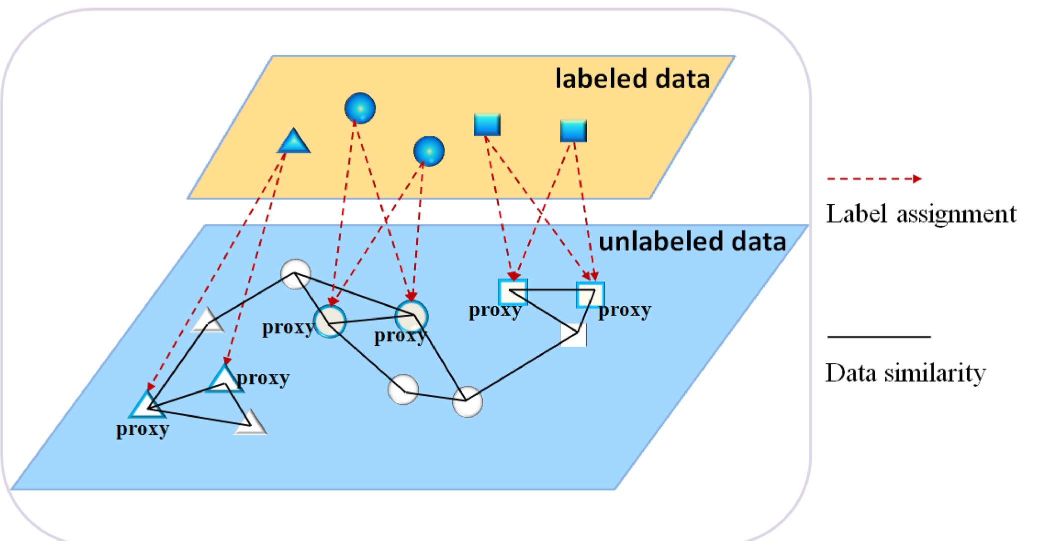
Our Work

- Propose a novel label propagation algorithm named PLCP, in which the label information is first propagated from labeled samples to their unlabeled neighbors, and then spreads only among unlabeled ones like a spreading activation network.
- Propose an online semi-supervised framework and develop an incremental learning method for PLCP.

2. Proxy-Based Local Consistency Propagation

Notation

- A point set $\mathcal{X} = (\mathcal{X}_L, \mathcal{X}_U) = \{\mathbf{x}_1, \dots, \mathbf{x}_l, \mathbf{x}_{l+1}, \dots, \mathbf{x}_n\}$
  - Points $\mathcal{X}_L = \{\mathbf{x}_1, \dots, \mathbf{x}_l\}$ are labeled $y_i \in \mathcal{L} = \{1, \dots, c\}$
  - Predict the label of unlabeled points $\mathcal{X}_U = \{\mathbf{x}_{l+1}, \dots, \mathbf{x}_n\}$



Algorithm

- Pairwise similarity
  - $w_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$
- Initial information:
  - $\mathbf{y}_i = \sum_{j=1}^{l} \delta\left(\mathbf{x}_i \in N(\mathbf{x}_j)\right) \cdot w_{ij}\mathbf{y}_j$
  - Matrix form: $\hat{\mathbf{Y}}_U = \mathbf{W}_{UL}\mathbf{Y}_L$
- Propagation iteratively
  - $\mathbf{F}_U(t+1) = \alpha\mathbf{S}_U\mathbf{F}_U(t) + (1-\alpha)\hat{\mathbf{Y}}_U$
- Converge to stable state
  - $\mathbf{F}^* = \lim_{t\to\infty} \mathbf{F}_U(t) = (1-\alpha)(\mathbf{I} - \alpha\mathbf{S}_U)^{-1}\hat{\mathbf{Y}}_U = (1-\alpha)(\mathbf{I} - \alpha\mathbf{S}_U)^{-1}\mathbf{W}_{UL}\mathbf{Y}_L$

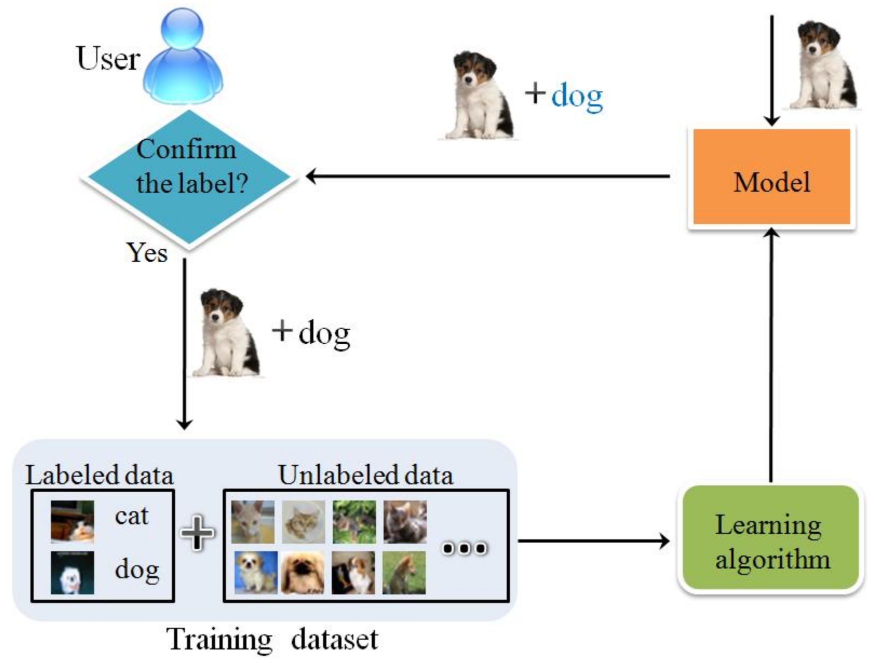
Object function

$$\mathcal{Q}(\mathbf{F}_U) = \frac{1}{2}\left(\sum_{i,j=1}^{u} w_{ij} \left\| \frac{\mathbf{F}_U^i}{\sqrt{\mathbf{D}_U^i}} - \frac{\mathbf{F}_U^j}{\sqrt{\mathbf{D}_U^j}} \right\|^2 + \mu \sum_{i=1}^{u} \| \mathbf{F}_U^i - [\mathbf{W}_{UL}\mathbf{Y}_L]_i \|^2\right)$$

3. Online Semi-Supervised Annotation

Framework

- Initially, all data are utilized to train the model in a semi-supervised manner.
- For a new image the system makes prediction using its current predictor and shows the prediction to the user.
- If the user confirms the label , the new data should be treated as training data to retrain the model, which can update the label information of the unlabeled data and the predictor.



Incremental learning

- Key idea: fix the transfer matrix
  - $\mathbf{T}_U = (1-\alpha)(\mathbf{I} - \alpha\mathbf{S}_U)^{-1}$
- For a new data, predict:
  - $\mathbf{F}_U' = \mathbf{T}_U[\mathbf{W}_{UL}, W_{U,new}] \cdot [\mathbf{Y}_L^T, \mathbf{y}_{new}^T]^T = \mathbf{F}_U + \mathbf{T}_U W_{U,new}\mathbf{y}_{new}$
- Trick: We denote $\Delta = W_{U,new}\mathbf{y}_{new}$, for multi-class, only update the *j-th* column of $\mathbf{F}_U$ with an increment $\Delta\mathbf{F}_U^j = \mathbf{T}_U \cdot \Delta_{.j}$
- Incrementally add *m* data:
  - $\mathbf{F}_U' = \mathbf{F}_U + \mathbf{T}_U\mathbf{W}_{UM}\mathbf{Y}_M$
  - Time complexity: $O(m \times u^2)$

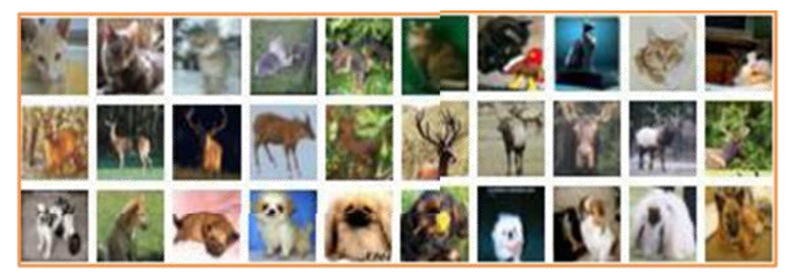
4. Experiments

Datasets

- MNIST: 70K, 784D Pixel
- CIFAR-10: 60K, 384D GIST

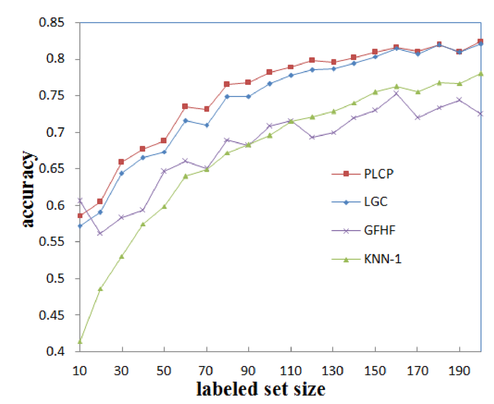
MNIST

CIFAR-10

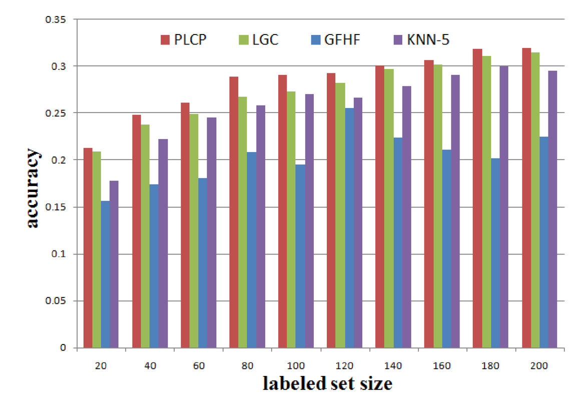


Results

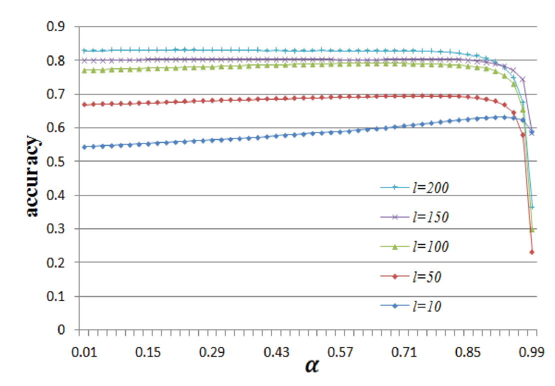
- Compared methods: KNN, GFHF, LGC
- Evaluation: transductive accuracy on unlabeled data



MNIST



CIFAR-10



Robustness of α

k-th round	PLCP time(s)	GFHF time(s)	LGC time(s)	PLCP-INC time(s)
k=1	17.30	8.48	21.83	0.0605
k=2	17.61	10.10	27.20	0.0640
k=3	17.79	11.50	34.14	0.0572
k=4	18.16	13.18	42.08	0.0595
k=5	18.28	15.45	51.02	0.0588
k=6	18.36	19.22	60.92	0.0651
k=7	18.65	21.40	72.49	0.0588
k=8	19.11	24.12	85.22	0.0609
k=9	19.78	28.11	99.53	0.0645
k=10	20.07	31.57	115.17	0.0589

Computational cost

Conclusion

- Efficient Semi-Supervised Annotation
  - Achieve better accuracy and has a promising performance
  - Satisfy the requirement of online real-time annotation

北京航空航天大學 BEIHANG UNIVERSITY

International Conference on Multimedia and Expo (ICME) 2013
San Jose, USA

国家重点实验室 STATE KEY LABORATORY