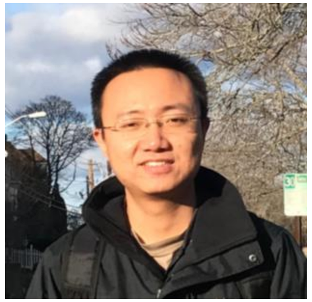
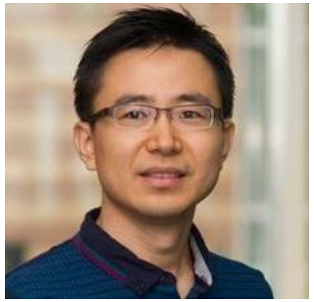
Adaptive Graph Guided Embedding for Multi-label Annotation





Lichen Wang
wanglichenxj@gmail.com

Zhengming Ding
allanding@ece.neu.edu

Yun Fu
yunfu@ece.neu.edu

SMILE Lab
Electrical & Computer Engineering
Northeastern University

Motivation

Multi-label annotation:

- Multi-label learning assigns multiple labels for each sample.

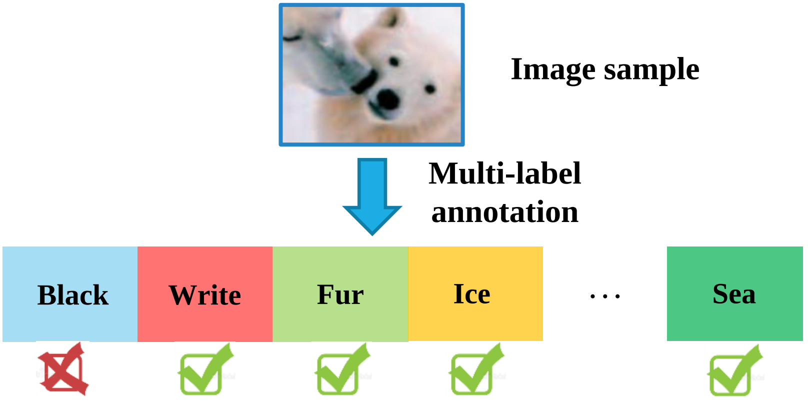
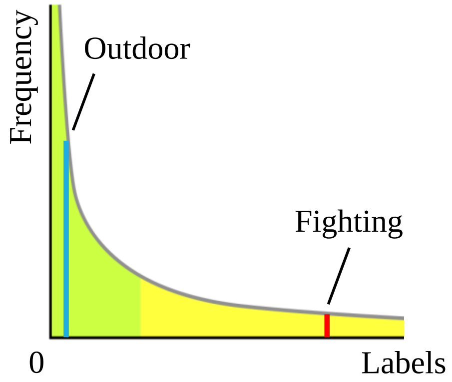


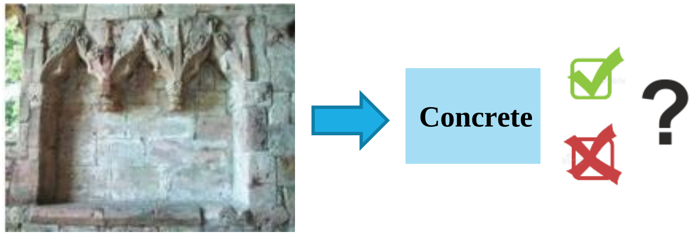
Fig 1: Goal of Multi-label annotation

General problems:

- Limited training data due to high labeling cost.
- Long-tailed label distribution.
- Inevitable labeling noise and outliers.



Long-tail distribution



Inevitable labeling noise

Fig 2: Challenges of multi-label learning

Our Algorithm

Adaptive Graph Guided Embedding (AG2E) for semi-supervised Multi-label Annotation:

- Utilizes limited labeled data associating with large-scale unlabeled data to facilitate learning performance.

Current model limitations:

- Sensitive to noise and outliers.
- Performance is highly depended on the pre-constructed graph.
- Not designed for multiple label scenarios.

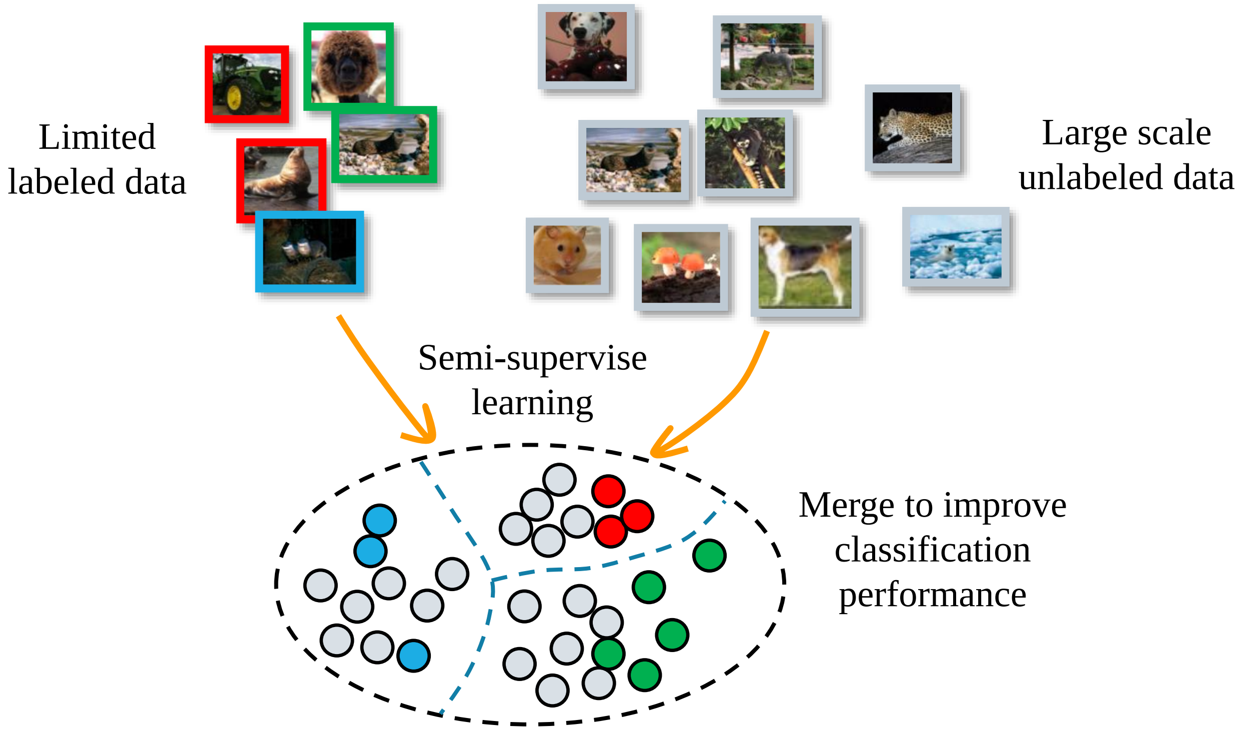


Fig 3: Goal of semi-supervised multi-label annotation

Our Algorithm

Adaptive Graph Guided Embedding (AG2E) for Multi-label Annotation:

- Propose to utilize adaptive graph to automatically capture data structure.
- Project original to latent space for obtaining distinctive representations.
- Efficient optimization strategy for real-world applications.

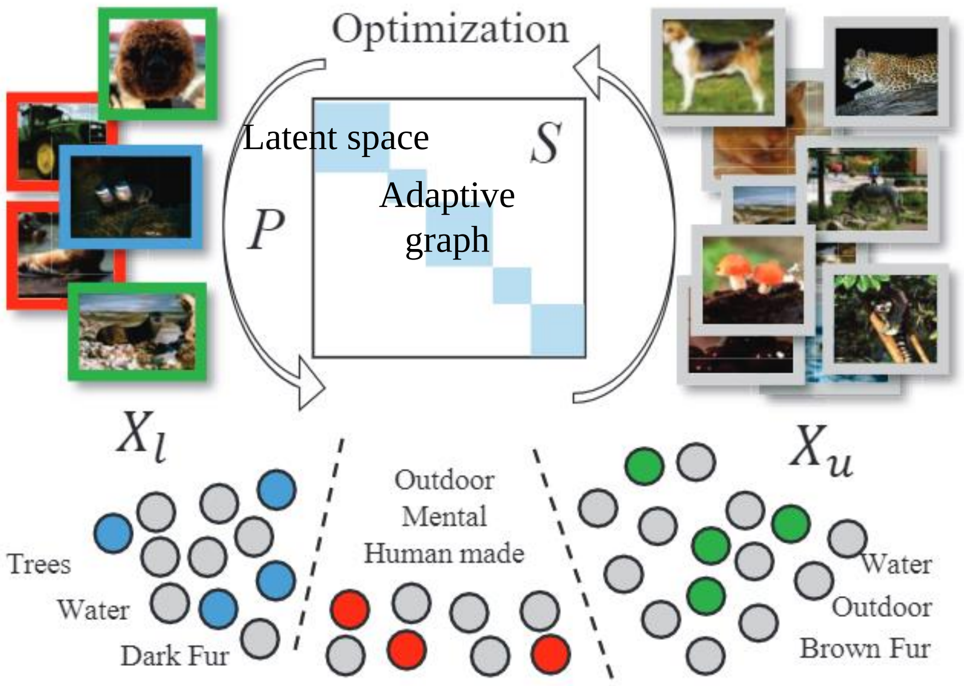


Fig 4: Adaptive Graph Guided Embedding model

Our Algorithm

Adaptive Graph Guided Embedding (AG2E) for Multi-label Annotation:

- Semi-supervise label regression.

$$\min_{F} \sum_{i,j} \|f_i - f_j\|_2^2 \boxed{s_{ij}}, \text{ s.t. } F_l = Y_l$$

$F = [F_l, F_u]$ – Label matrix
F_l, F_u – Labels of the labeled samples and unlabeled samples
S – Similarity matrix
Y_l – Ground truth of the labeled data

- Fixed model to obtain S

$$\boxed{s_{ij}} = \begin{cases} e^{-\|x_i - x_j\|_2^2 / 2\delta^2}, & \text{if } x_i \in \mathcal{N}_K(x_j) \text{ or } x_j \in \mathcal{N}_K(x_i), \\ 0, & \text{otherwise,} \end{cases}$$

x_i – Sample features

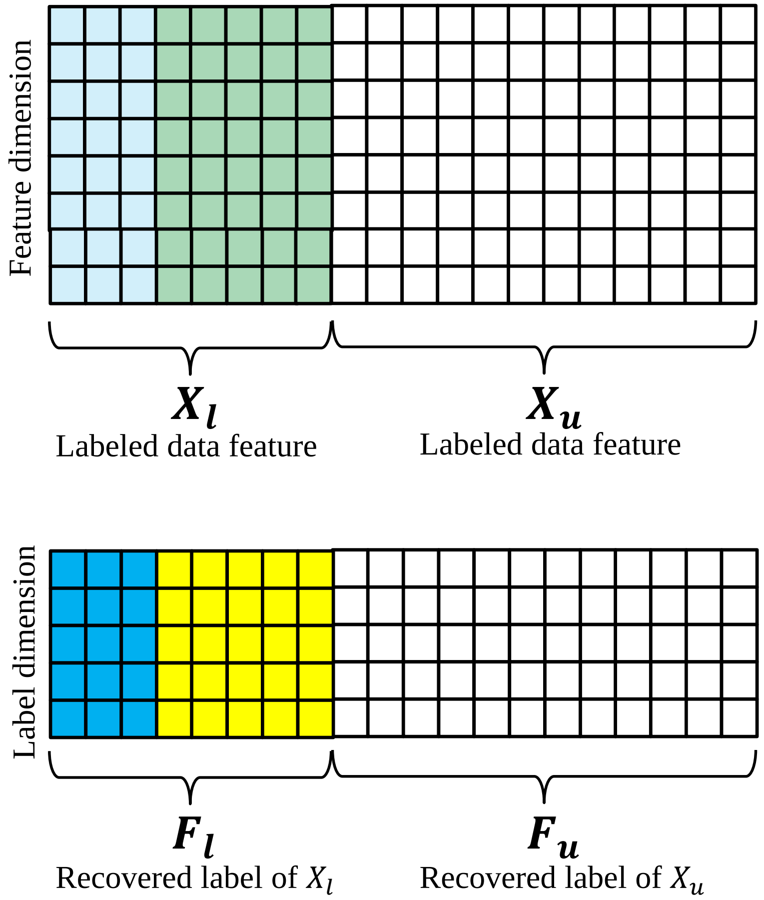


Fig 5: Definition and dimension of feature/label matrix

Our Algorithm

Adaptive Graph Guided Embedding (AG2E) for Multi-label Annotation:

- Adaptive graph embedding which optimize S in the training process:

$$\min_{F,S} \sum_{i,j} \|f_i - f_j\|_2^2 s_{ij} + \boxed{\mu \sum_{i,j} \|x_i - x_j\|_2^2 s_{ij}}$$
$$\text{s.t. } F_l = Y_l, S \geq 0.$$

- A pre-defined *S* guides and improves training process

$$\min_{F,S} \sum_{i,j} \|f_i - f_j\|_2^2 s_{ij} + \mu \sum_{i,j} \|x_i - x_j\|_2^2 s_{ij} + \boxed{\lambda \|S - \bar{S}\|_{\mathrm{F}}^2}$$
$$\text{s.t. } F_l = Y_l, S \geq 0.$$

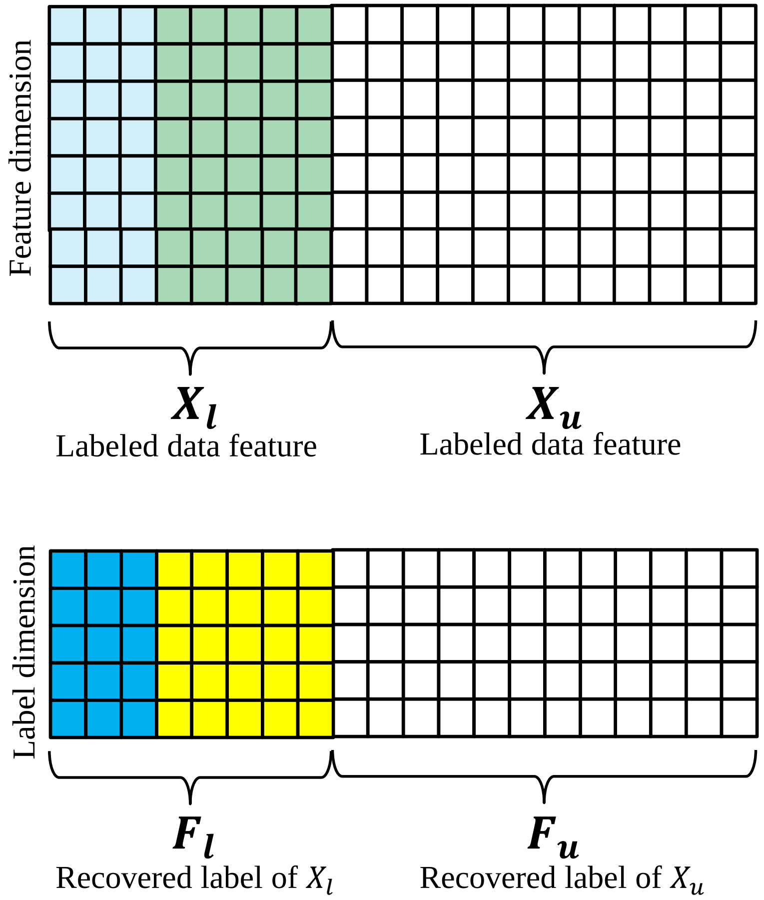


Fig 5: Definition and dimension of feature/label matrix

Our Algorithm

Adaptive Graph Guided Embedding for Multi-label Annotation:

- Final objective function with constraints:

$$\min_{F,P,S} \mathrm{tr}(FL_SF^\top) + \mu\mathrm{tr}(PXL_SX^\top P^\top) + \lambda\|S - \bar{S}\|_\mathrm{F}^2$$
$$\text{s.t.} \quad F_l = Y_l, \quad PXHX^\top P^\top = \mathrm{I}, \quad S\mathbf{1} = \mathbf{1}, \quad S \geq 0.$$

Advantages:

- Robust to noise & outliers
- Fully utilize unlabeled data structure (in latent space)
- Higher multi-label recovery accuracy

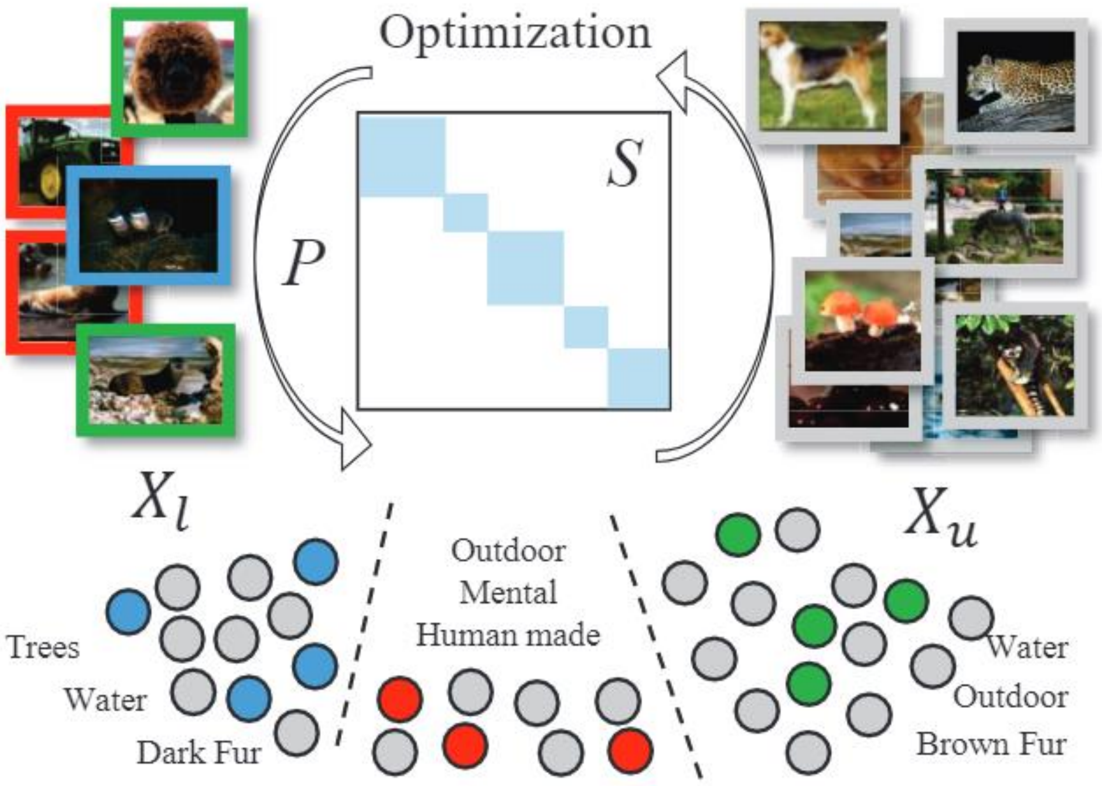


Fig 6: Challenges of multi-label learning

Experiments

Datasets:

- 3 visual + 1 acoustic + 1 music multi-label datasets:
  - Animal with attributes dataset [1]
  - SUN dataset [2]
  - Caltech-Ucsd Birds dataset [3]
  - BIRD dataset [4]
  - Emotion dataset [5]

Evaluations:

- Precision (Pre)
- Recall (Rec)
- F1-score (F1)
- Non-zero Recall (N-R)

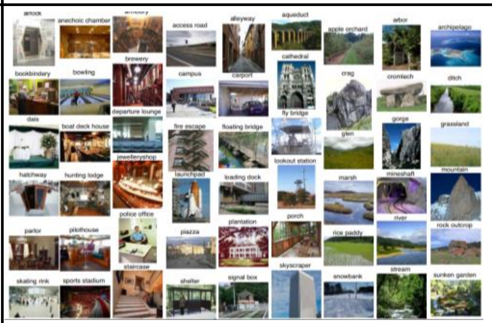
AWA Dataset [1]	SUN Dataset [2]	CUB Dataset [3]
Att:85 Cls:40/10 Sap: 24295/6180	Att: 102 Cls:707/10 Sap: Sap: 14140/200	Att: 312 Cls: 100/50 Sap: 8855/2933
		

Fig 7: Image samples of the 3 visual datasets

[1] Christoph H., Hannes Nickisch, and Stefan Harmeling. "Attribute-based classification for zero-shot visual object categorization." *TPAMI* 36.3 (2014): 453-465.
[2] Xiao, Jianxiong, et al. "Sun database: Large-scale scene recognition from abbey to zoo." Computer vision and pattern recognition (CVPR), 2010 IEEE conference on. IEEE, 2010.
[3] Wah, Catherine, et al. "The caltech-ucsd birds-200-2011 dataset." (2011).

Experiments

Performance comparison:

- Achieves highest performance in most of the metrics.
- Achieves highest performance if the data scale beyond a threshold.

Dataset	Method	Prec	Recall	F1	N-R
SUN	Regression	0.6318	0.1504	0.2429	101
	SSMLDR	0.5625	0.1239	0.2031	68
	FastTag	0.6187	0.1473	0.2379	101
	ML-PGD	0.7218	0.1521	0.2513	100
	AG^{2}E (Ours)	**0.7460**	**0.1625**	**0.2669**	**102**
CUB	Regression	0.2183	0.0247	0.0443	162
	SSMLDR	0.2162	0.0399	0.0674	164
	FastTag	0.3231	0.0496	0.0860	163
	ML-PGD	0.3029	0.0448	0.0781	132
	AG^{2}E	**0.3351**	**0.0525**	**0.0908**	**194**
AWA	Regression	**0.8198**	0.0819	0.1489	**75**
	SSMLDR	0.8085	0.0948	0.1698	74
	FastTag	0.7848	0.0857	0.1545	67
	ML-PGD	0.5283	0.0631	0.1127	45
	AG^{2}E	0.7745	**0.1285**	**0.2204**	72
EMO	Regression	0.3793	0.9114	0.5357	6
	SSMLDR	0.3556	0.8965	0.5093	6
	FastTag	0.3833	0.9459	0.5456	6
	ML-PGD	0.3784	0.9265	0.5373	6
	AG^{2}E	**0.3995**	**0.9714**	**0.5762**	**6**
BIRD	Regression	0.0764	0.3726	0.1268	13
	SSMLDR	0.0709	0.3465	0.1178	12
	FastTag	0.1005	0.3783	0.1601	16
	ML-PGD	0.0809	0.3883	0.1338	15
	AG^{2}E	**0.1021**	**0.4529**	**0.1653**	**17**

Table 1: Multi-label annotation of 5 datasets

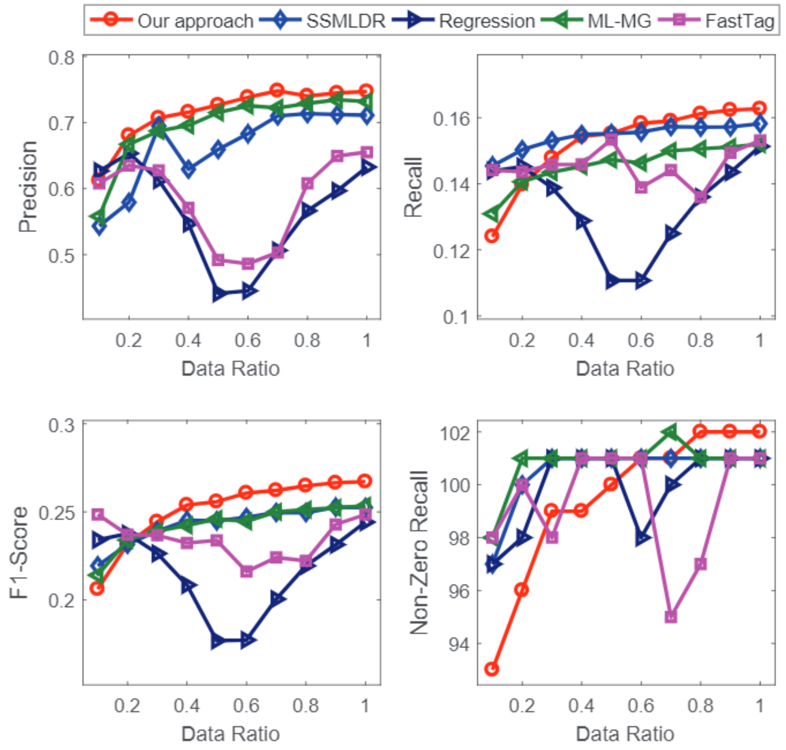


Fig 8: Annotation performance based on partial of the training set

Experiments

Performance comparison:

- Robust in zero-shot learning scenario.
- It is able to find the incorrect labels from test set.
- Parameter insensitive.

Fig 9: Multi-label annotation visualization samples

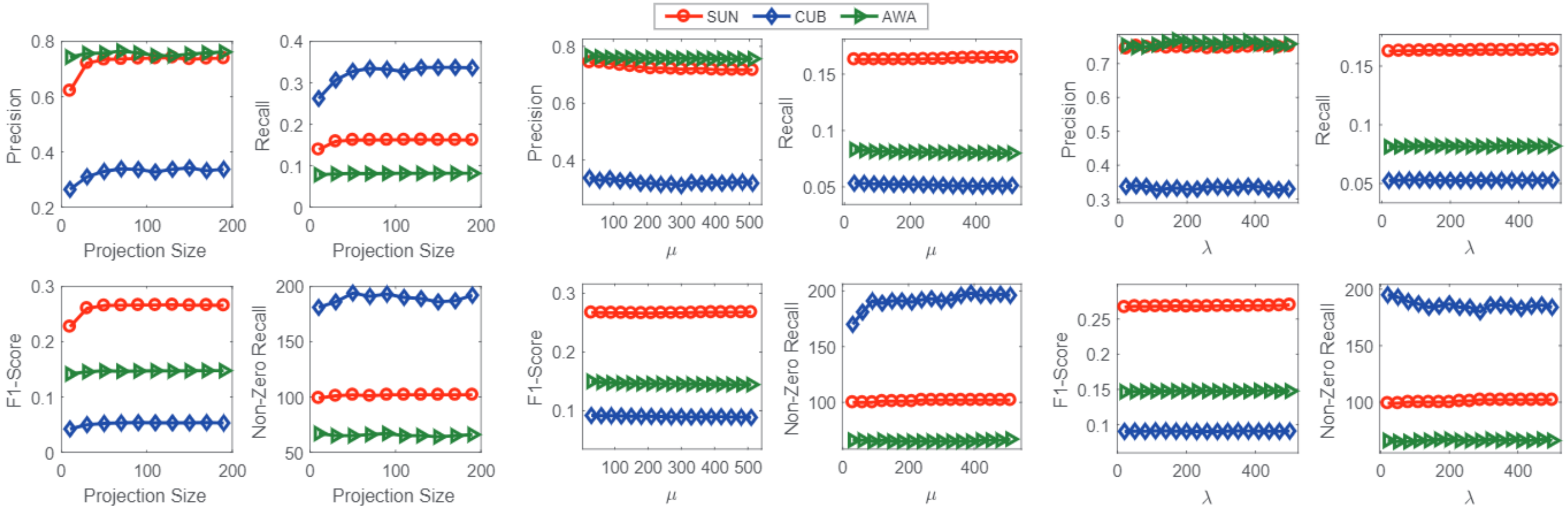


Fig 10: Parameter sensitivity analysis

Approaches	SUN	CUB	AWA
Labeled data	65.20	27.24	52.31
Regression	65.00	27.21	52.33
SSMLDR	66.00	32.19	53.64
FastTag	64.00	27.18	54.32
ML-PGD	65.40	28.48	54.93
AG^{2}E (Ours)	**67.40**	**32.53**	**55.71**

Table 2: Zero-shot multi-label annotation

Thank you!

Please contact: **wanglichenxj@gmail.com** for questions.

SMILE Lab
Electrical & Computer Engineering
Northeastern University