A Graph-Theoretic Framework for Understanding Open-World Semi-Supervised Learning

NeurIPS 2023 (Spotlight)

Yiyou Sun
UW-Madison

Zhenmei Shi
UW-Madison

Yixuan Li
UW-Madison
A Paradigm Shift from Closed-world to the Open-world

**Closed-world ML:**
Handle data with the **known** classes

**Open-world ML:**
Handle data with both **novel and known** classes

(Figures are powered by GPT-4V)
Goal: correctly classify known and cluster novel classes.
This research area starts to gain attention!

A Graph-Theoretic Framework for Understanding Open-world Semi-Supervised Learning [SSL, NeurIPS 23]
An Open Research Question

“what is the role of the label information in shaping representations for both known and novel classes?”
An Intuitive Example

Starting Point: All *Unlabeled* Samples
An Intuitive Example

We label the first two images as "traffic lights"…
An Intuitive Example

**Question:** Will other “traffic light” samples get closer to each other?

**Known Class**

---

A Graph-Theoretic Framework for Understanding Open-world Semi-Supervised Learning [SSL, NeurIPS 23]
An Intuitive Example

Question: Will other “green” samples get closer to “red” samples?

Novel Class
(Strong relationship)
An Intuitive Example

Question: Will unrelated novel class be affected?
An Intuitive Example

A formal understanding is needed!
Methodology
Augmentation Graph

**Node**: Augmented Images.

**Edge Weight**: Probability of two images are considered as positive pair.
Label Perturbation
Adding labels changes the graph structure.

Unlabeled Augmentation Graph
Label Perturbation

Adding labels perturbs the graph structure.

Unlabeled Augmentation Graph

Augmentation Graph with labels
Label Perturbation
Adding labels changes the graph structure.

How do representations change?
How do cluster results change?

Unlabeled Augmentation Graph  Augmentation Graph with labels

Add labels
Contrastive Learning learns the augmentation graph.

Learning Goal
Features
Augmented Image
Source Image

Make Close
Keep away
Spectral Open-world Representation Learning (SORL)
Contrastive loss derived from Matrix Factorization

\[ \mathcal{L}_{mf}(F, A) = \left\| \text{normalize}(A) - FF^T \right\|^2_F \]

\[ \mathcal{L}_{sort}(f) \triangleq -2\alpha \mathcal{L}_1(f) - 2\beta \mathcal{L}_2(f) + \alpha^2 \mathcal{L}_3(f) + 2\alpha\beta \mathcal{L}_4(f) + \beta^2 \mathcal{L}_5(f) \]

Make Close Positive Pairs
Keep away Negative Pairs

See more details in paper!
SORL has the closed-form solution.

\[
\mathcal{L}_{mf}(F, A) = \left\| \text{normalize}(A) - FF^\top \right\|_F^2
\]

Optimal Solution (Eckart–Young–Mirsky Theorem)

A Graph-Theoretic Framework for Understanding Open-world Semi-Supervised Learning [SSL, NeurIPS 23]
The closed-form solution is known!

$$\mathcal{L}_{mf}(F, A) = \left\| \text{normalize}(A) - FF^\top \right\|_F^2$$

Good! We can analyze the feature space with **spectral analysis** of the adjacency matrix!
Theory
Main Intuition of the Theorem

Cluster Performance Gain by adding labels for Class $c$.

\[
\Delta_{\pi_c}(\delta) = \left( I_{\pi_c} - \frac{1}{N} \right) - 2 \left( 1 - \frac{\left| \pi_c \right|}{N} \right) \left( \mathbb{E}_{i \in \pi_c} \mathbb{E}_{j \in \pi_c} z_i^T z_j - \mathbb{E}_{i \in \pi_c} \mathbb{E}_{j \notin \pi_c} z_i^T z_j \right).
\]

Connection from class $c$ to the labeled data.

Intra-class similarity

Inter-class similarity
Main Theorem (Case Study)

Case Study 1 (unlabeled data from known class):

\[
\Delta_{\pi_c}(\delta) = (I_{\pi_c} - \frac{1}{N}) - 2(1 - \frac{|\pi_c|}{N})(E_{i \in \pi_c} E_{j \in \pi_c} z_i^T z_j - E_{i \in \pi_c} E_{j \notin \pi_c} z_i^T z_j).
\]

connection >> (intra-sim - inter-sim)

(very large) \quad (...) \quad (...) 

Conclusion: Unlabeled traffic lights will be better clustered!
Main Theorem (Case Study)

$$\Delta_{\pi_c}(\delta) = \left(I_{\pi_c} - \frac{1}{N}\right) - 2\left(1 - \frac{|\pi_c|}{N}\right) \left(\mathbb{E}_{i \in \pi_c, j \in \pi_c} z_i^T z_j - \mathbb{E}_{i \in \pi_c, j \notin \pi_c} z_i^T z_j\right).$$

Connection from class $c$ to the labeled data.

**Intra-class similarity**

**Inter-class similarity**

Case Study 2 (novel class with **strong** connection):

connection \(>\) (intra-sim - inter-sim)

(large) \(>\) (Low) \(>\) (…)

**Conclusion:** Green and red apple will be close to each other!
Main Theorem (Case Study)

Case Study 3 (novel class with \textit{weak} connection):

\[
\Delta_{\pi_c}(\delta) = \left( \mathbf{I}_{\pi_c} - \frac{1}{N} \right) - 2(1 - \frac{|\pi_c|}{N})(\mathbb{E}_{i \in \pi_c} \mathbb{E}_{j \in \pi_c} z_i^T z_j - \mathbb{E}_{i \in \pi_c} \mathbb{E}_{j \notin \pi_c} z_i^T z_j).
\]

\text{Connection from class } c \text{ to the labeled data.}

Connection < (intra-sim - inter-sim)

(Low) \hspace{2cm} \text{High} \hspace{2cm} (\ldots)

Conclusion: Add labels may not be beneficial to flower class.
A Toy Example

\[ A = \eta_u A^{(u)} + \]

\[ \eta_u A^{(u)} = \begin{bmatrix}
\tau_1^2 + \tau_s^2 + \tau_c^2 & 2\tau_1\tau_s & 2\tau_1\tau_c & 2\tau_c\tau_s & 0 & 0 \\
2\tau_1\tau_s & \tau_1^2 + \tau_s^2 + \tau_c^2 & 2\tau_1\tau_s & 2\tau_c\tau_s & 0 & 0 \\
2\tau_1\tau_c & 2\tau_c\tau_s & \tau_1^2 + \tau_s^2 + \tau_c^2 & 2\tau_1\tau_s & 0 & 0 \\
2\tau_c\tau_s & 2\tau_1\tau_s & 2\tau_1\tau_s & \tau_1^2 + \tau_s^2 + \tau_c^2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \]

By Eq. (1,2,3)

See more details in paper!
Experiment
Set Up

**Model**

**Dataset (CIFAR-10/100)**

1. Separate all classes into 50% known and 50% novel.
2. Divide known-class samples into 50% labeled and 50% unlabeled.
SORL is also appealing for practical usage!

### CIFAR-10

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCD</td>
<td>86</td>
</tr>
<tr>
<td>ORCA</td>
<td>88</td>
</tr>
<tr>
<td>OpenCon</td>
<td>90</td>
</tr>
<tr>
<td>SORL</td>
<td>94</td>
</tr>
</tbody>
</table>

### CIFAR-100

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCD</td>
<td>30</td>
</tr>
<tr>
<td>ORCA</td>
<td>37.5</td>
</tr>
<tr>
<td>OpenCon</td>
<td>45</td>
</tr>
<tr>
<td>SORL</td>
<td>52.5</td>
</tr>
</tbody>
</table>

A Graph-Theoretic Framework for Understanding Open-world Semi-Supervised Learning [SSL, NeurIPS 23]
Thank you!

Our code is available at https://github.com/deeplearning-wisc/SORL.