Görüntü Tanımada Çizge Tabanlı Yöntemler

Ahmet İşcen

Google Research
Outline

01 Label Propagation with Graphs
02 Ranking on Manifolds with Graphs
03 Metric Learning with Graphs
04 Deep Semi-supervised Learning with Graphs
Label Propagation with Graphs
Label propagation

- Semi supervised classification
- Transductive learning
- $N$ examples, $C$ classes

- **Smoothness assumption:**

  “if two points are close, then, so should be the corresponding output”

[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]
Label propagation

- Semi supervised classification
- Transductive learning
- \(N\) examples, \(C\) classes

\[ F^t = \alpha S F^{t-1} + (1 - \alpha) Y \]

\(Y\): \(N \times C\) 1-hot-coded label matrix
\(S\): \(N \times N\) normalized affinity matrix (k-nearest neighbor graph)
\(\alpha\): propagation parameter in (0,1)
\(F\): \(N \times C\) class confidence matrix

[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]
Label propagation iterations

iteration 0

[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]
Label propagation iterations

iteration 1

[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]
Label propagation iterations

iteration 2
Label propagation iterations

iteration 5

[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]
Label propagation iterations

iteration 10

[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]
Label propagation iterations

[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]
Label propagation iterations

[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]
Ranking on Manifolds with Graphs
Ranking on manifolds

- Similarity search (unsupervised)
- \( N \) examples, 1 query

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Euclidean similarity

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Euclidean similarity

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Ranking on manifolds

- Similarity search (unsupervised)
- \( N \) examples, 1 query

\[
f^t = \alpha S f^{t-1} + (1 - \alpha) y
\]

- \( y \): \( N \times 1 \) 1-hot-coded vector which defines the query
- \( S \): \( N \times N \) normalized affinity matrix (k-nearest neighbor graph)
- \( \alpha \): propagation parameter in (0,1)
- \( f \): \( N \times 1 \) similarity vector w.r.t the query

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Similarity propagation iterations

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Simularity propagation iterations

iteration 1

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Similarity propagation iterations

iteration 5

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Similarity propagation iterations

iteration 10

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Similarity propagation iterations

iteration 50

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Similarity propagation iterations

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Euclidean vs similarity propagation

Euclidean similarity

Similarity propagation

[Zhou et al, NeurIPS 2003, Ranking on data manifolds]
Euclidean vs similarity propagation
Euclidean vs similarity propagation

manifold similarity for image retrieval:
[Zhang et al., ECCV’12, Query specific fusion for image retrieval]
[Donoser & Bischof, CVPR’13, Diffusion processes for retrieval revisited]
[Bai et al. ICCV’17, Ensemble diffusion for retrieval]
[Iscen, Tolias, Avrithis, Furon, Chum. CVPR’17. Efficient Diffusion on Region Manifolds: Recovering Small Objects with Compact CNN Representations]

......
Efficient solution

- **Iterative solution (inefficient)**
  \[ f^t = \alpha S f^{t-1} + (1 - \alpha) y \]

- **Closed-form solution**
  \[ f^* = (I - \alpha S)^{-1} y \]
  Not sparse, inefficient

- **Conjugate gradient for linear system**
  \[ (I - \alpha S) f^* = y \]
  Sparse, efficient

[Iscen et al, CVPR 2017, Efficient Diffusion on Region Manifolds: Recovering Small Objects with Compact CNN Representations]
Metric Learning with Graphs
Metric learning

- Learn a metric which assigns small distances to images that are visually similar
Metric learning

- Learn a metric which assigns small distances to images that are visually similar

\[
l_c(z^r, z^+, z^-) = \|z^r - z^+\|^2 + [m - \|z^r - z^-\|]_+^2
\]
Mining on manifolds (MoM)

- automatically mine pairs of matching and non-matching images

[Iscen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]
Mining on manifolds (MoM)

Anchor selection

- Nearest neighbor graph $A$
- Stationary probability distribution $\pi$ of random walk on $A$
  
  $\pi = \pi P,$
  $P = D^{-1}A$
  $D = \text{degree matrix of } A$

[İsčen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]
Mining on manifolds (MoM)

- Anchor: black dot
- Euclidean NN (orange): $E(y)$
- Manifold NN (purple): $M(y)$
- Hard positives (green): $S^+ = M(y) \setminus E(y)$
- Hard negatives (red): $S^- = E(y) \setminus M(y)$

[Iscen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]
Mining on manifolds (MoM)

[Iscen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]
Experiments on instance search

initialize: pre-training on ImageNet -- fine-tune: MoM on $10^6$ images

[Iscen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]
Experiments on fine-grained recognition

**Cyclic Match:** [Li et al. ECCV’16. Unsupervised visual representation learning by graph-based consistent constraints]

**Triplets (semi-hard):** [Schroff et al. CVPR’15. Facenet: A unified embedding for face recognition and clustering]

**Lifted Structure:** [Song et al. CVPR’16. Deep metric learning via lifted structured feature embedding]

**Smart Mining:** [Harwood et al. ICCV’17. Smart mining for deep metric learning. In ICCV, 2017]

[İscen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]
Ablation experiments

<table>
<thead>
<tr>
<th>Anchors</th>
<th>CUB</th>
<th>Oxford5k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>35.0</td>
<td>52.6</td>
</tr>
<tr>
<td>$\text{NN}^e_5$</td>
<td>$X \setminus \text{NN}^e_5$</td>
<td>38.5</td>
</tr>
<tr>
<td>$P^+$</td>
<td>$X \setminus \text{NN}^e_5$</td>
<td>43.0</td>
</tr>
<tr>
<td>$\text{NN}^e_5$</td>
<td>$P^-$</td>
<td>42.1</td>
</tr>
<tr>
<td>$P^+$</td>
<td>$P^-$</td>
<td>43.5</td>
</tr>
<tr>
<td>$P^+ + W$</td>
<td>$P^- + W$</td>
<td>45.3</td>
</tr>
</tbody>
</table>

random anchors vs proposed anchors
## Ablation experiments

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
<th>CUB</th>
<th>Oxford5k</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anchors</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Initial</td>
<td>All</td>
<td>35.0</td>
<td>52.6</td>
</tr>
<tr>
<td>$\text{NN}_5^e$</td>
<td>$X \setminus \text{NN}_5^e$</td>
<td>38.5</td>
<td>37.4</td>
</tr>
<tr>
<td>$P^+$</td>
<td>$X \setminus \text{NN}_5^e$</td>
<td>43.0</td>
<td>48.2</td>
</tr>
<tr>
<td>$\text{NN}_5^e$</td>
<td>$P^-$</td>
<td>42.1</td>
<td>57.8</td>
</tr>
<tr>
<td>$P^+$</td>
<td>$P^-$</td>
<td>43.5</td>
<td>64.4</td>
</tr>
<tr>
<td>$P^+ + W$</td>
<td>$P^- + W$</td>
<td>45.3</td>
<td>67.0</td>
</tr>
</tbody>
</table>

- **hard positives vs easy positives**

[Iscen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]
Ablation experiments

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
<th>CUB</th>
<th>Oxford5k</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Random</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anchors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial</td>
<td>35.0</td>
<td>52.6</td>
<td></td>
</tr>
<tr>
<td>$\text{NN}_5^e$</td>
<td>$X \setminus \text{NN}_5^e$</td>
<td>38.5</td>
<td>37.4</td>
</tr>
<tr>
<td>$P^+$</td>
<td>$X \setminus \text{NN}_5^e$</td>
<td>43.0</td>
<td>48.2</td>
</tr>
<tr>
<td>$\text{NN}_5^e$</td>
<td>$P^-$</td>
<td>42.1</td>
<td>57.8</td>
</tr>
<tr>
<td>$P^+$</td>
<td>$P^-$</td>
<td>43.5</td>
<td>64.4</td>
</tr>
<tr>
<td>$P^+ + W$</td>
<td>$P^- + W$</td>
<td>45.3</td>
<td>67.0</td>
</tr>
</tbody>
</table>

[Iselen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]
Deep
Semi-supervised
Learning with
Graphs
Label Propagation for Deep Semi-supervised Learning

Given labeled examples $X_L$, and unlabeled examples $X_U$

We want to learn:

- A feature map $\phi_\theta : \mathcal{X} \rightarrow \mathbb{R}^d$
- A classifier (FC layer after the feature map)

By combining:

- Transductive learning (all test points are seen) with inductive learning (none of the test points are seen)

Label Propagation for Deep Semi-supervised Learning
Label Propagation for Deep Semi-supervised Learning

Label Propagation for Deep Semi-supervised Learning

Label Propagation for Deep Semi-supervised Learning

- Train with $L_s(X_L, Y_L; \theta)$ for $T$ epochs
- Use $\phi_\theta$ for classification
- Features $\phi_\theta(X)$
- Affinity $A$
- $W \leftarrow A + A^T$
- $W \leftarrow D^{-1/2}WD^{-1/2}$
- Label propagation

Label Propagation for Deep Semi-supervised Learning

Label Propagation for Deep Semi-supervised Learning

Label Propagation for Deep Semi-supervised Learning

- Supervised loss
  \[ L_s(X_L, Y_L; \theta) := \sum_{i \in L} \ell_s(f_\theta(x_i), y_i) \]  Cross entropy loss

- Weighted pseudo-label loss
  \[ L_w(X_U, \hat{Y}_U; \theta) := \sum_{i \in U} \omega_i \zeta_{\hat{y}_i} \ell_s(f_\theta(x_i), \hat{y}_i) \]

- Certainty of pseudo-label prediction
  \[ \omega_i := 1 - \frac{H(\hat{z}_i)}{\log c} \]

- Class weight for balancing class distribution:
  \[ \zeta_j := (|L_j| + |U_j|)^{-1} \]
Label Propagation for Deep Semi-supervised Learning

Correctly and incorrectly pseudo-labeled examples

Ablation experiment on CIFAR10 (Error rate, less is better)

<table>
<thead>
<tr>
<th>Pseudo-labeling</th>
<th>$\omega_i$</th>
<th>$\zeta_j$</th>
<th>CIFAR-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion (7)</td>
<td>✓</td>
<td>✓</td>
<td>$36.53 \pm 1.42$</td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td>✓</td>
<td>$36.17 \pm 1.98$</td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td>✓</td>
<td>$33.32 \pm 1.53$</td>
</tr>
<tr>
<td>Network (1)</td>
<td>✓</td>
<td>✓</td>
<td>$32.40 \pm 1.80$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$35.17 \pm 2.46$</td>
</tr>
</tbody>
</table>

### Classification error on CIFAR10

<table>
<thead>
<tr>
<th>Dataset</th>
<th>CIFAR-10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>500</td>
</tr>
<tr>
<td>Nb. labeled images</td>
<td></td>
</tr>
<tr>
<td>Fully supervised</td>
<td>49.08 ± 0.83</td>
</tr>
<tr>
<td>TDCNN [33]†</td>
<td>-</td>
</tr>
<tr>
<td>Ours—(1)</td>
<td>35.17 ± 2.46</td>
</tr>
<tr>
<td>Ours</td>
<td>32.40 ± 1.80</td>
</tr>
<tr>
<td>VAT [23]†</td>
<td>-</td>
</tr>
<tr>
<td>Π model [20]†</td>
<td>-</td>
</tr>
<tr>
<td>Temporal Ensemble [20]†</td>
<td>-</td>
</tr>
<tr>
<td>MT [35]†</td>
<td>-</td>
</tr>
<tr>
<td>MT [35]</td>
<td>27.45 ± 2.64</td>
</tr>
<tr>
<td>MT + Ours</td>
<td>24.02 ± 2.44</td>
</tr>
</tbody>
</table>

## Classification error on CIFAR100 and mini-ImageNet

<table>
<thead>
<tr>
<th>Dataset</th>
<th>CIFAR-100</th>
<th>Mini-ImageNet-top1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb. labeled images</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td>55.43 ± 0.11</td>
<td>74.78 ± 0.33</td>
</tr>
<tr>
<td>10000</td>
<td>40.67 ± 0.49</td>
<td>60.25 ± 0.29</td>
</tr>
<tr>
<td>Fully supervised</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td>46.20 ± 0.76</td>
<td>70.29 ± 0.81</td>
</tr>
<tr>
<td>MT [35]</td>
<td>45.36 ± 0.49</td>
<td>72.51 ± 0.22</td>
</tr>
<tr>
<td>MT + Ours</td>
<td>43.73 ± 0.20</td>
<td>57.35 ± 1.66</td>
</tr>
</tbody>
</table>

Other graph-based methods
Additional graph-based methods

- **Fast spectral ranking for similarity search** (Iscen et al, CVPR 2018)
  - Embedding for manifold search based on low-rank decomposition

- **Hybrid diffusion: Spectral-temporal graph filtering for manifold ranking** (Iscen et al, ACCV 2018)
  - Even more efficient label propagation for similarity search

- **Graph convolutional networks for learning with few clean and many noisy labels** (Iscen et al, ECCV 2020)
  - Graph convolutional networks for cleaning label noise
co-authors

Slide credits: Giorgos Tolias and Yannis Avrithis
Questions