Görüntü Tanımada Çizge Tabanlı Yöntemler

Ahmet İşcen

Google Research

Outline

Google Research

Google

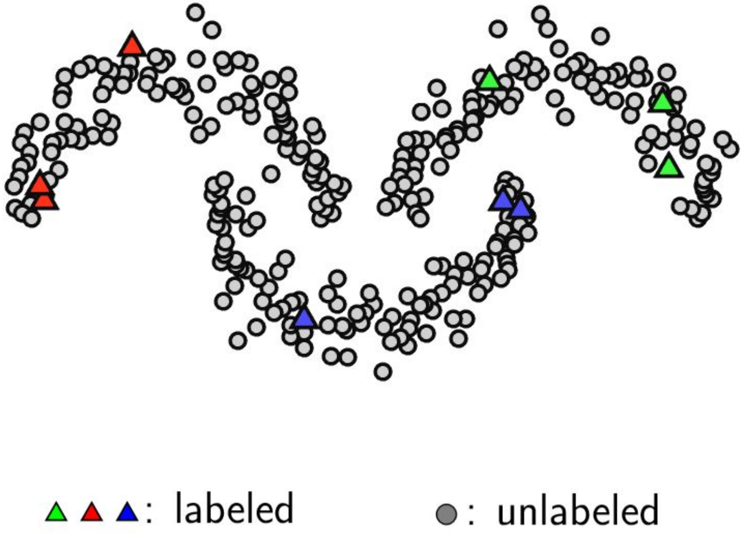
Label Propagation with Graphs

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”



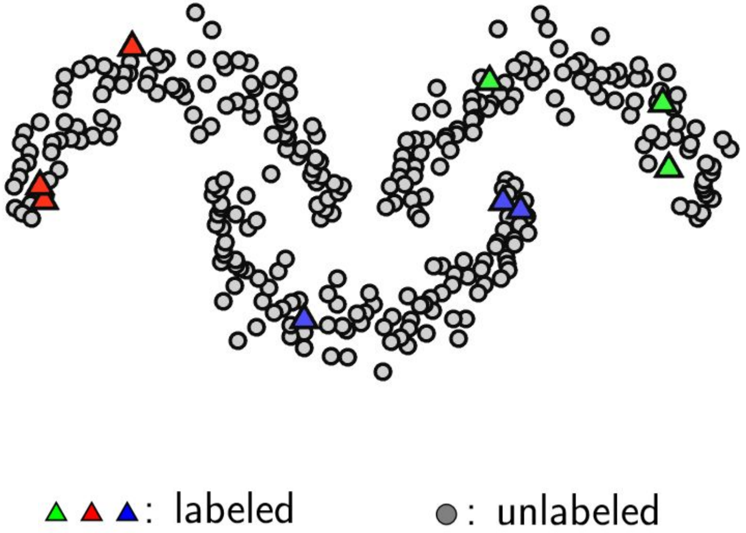
[Zhou et al, NeurIPS 2003, Learning with Local and Global Consistency]

Label propagation

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

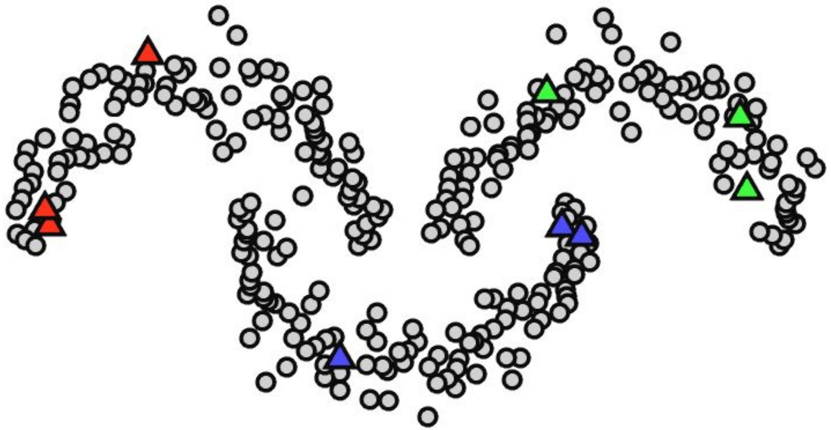
$$F^t = \alpha S F^{t-1} + (1 - \alpha) Y$$

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



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

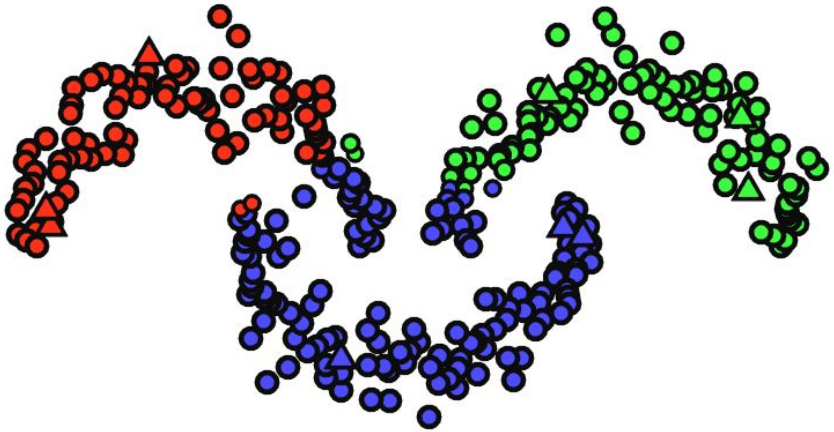
Label propagation iterations



iteration 0

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

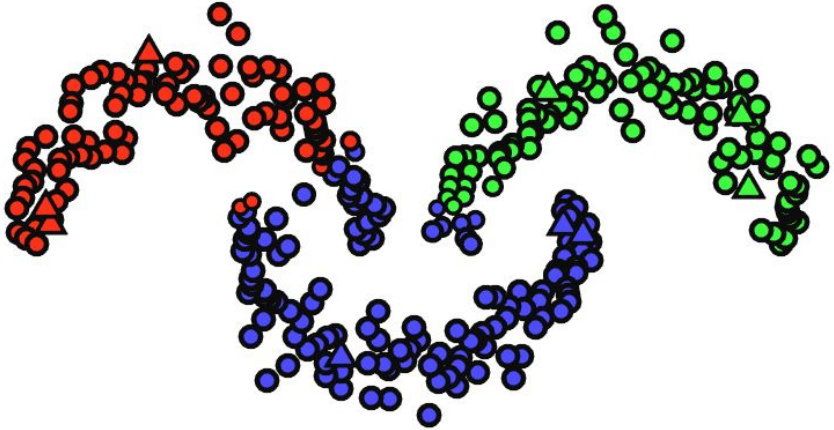
Label propagation iterations



iteration 1

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

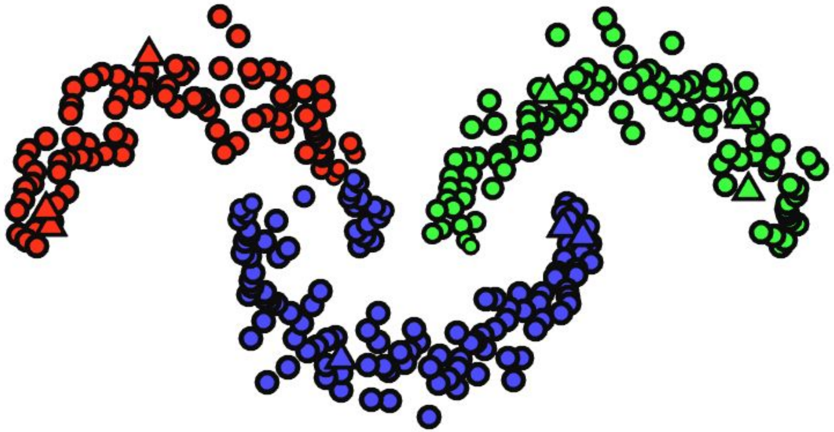
Label propagation iterations



iteration 2

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

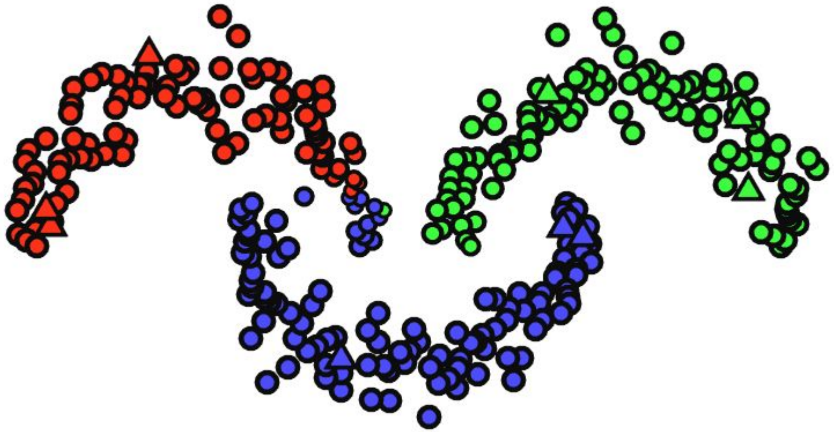
Label propagation iterations



iteration 5

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

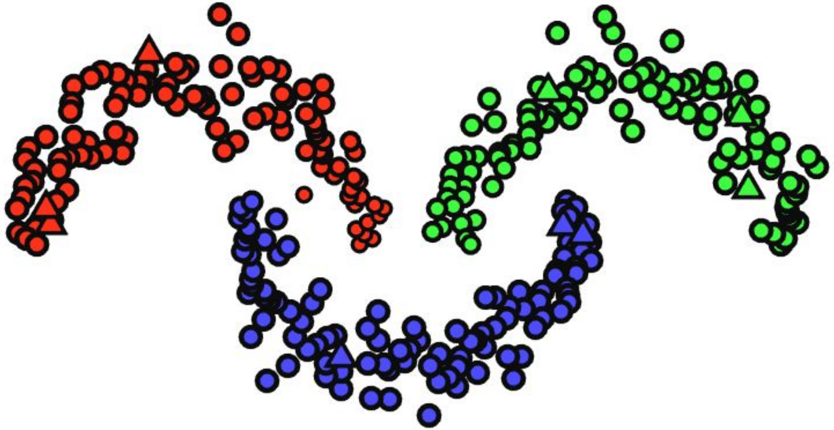
Label propagation iterations



iteration 10

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

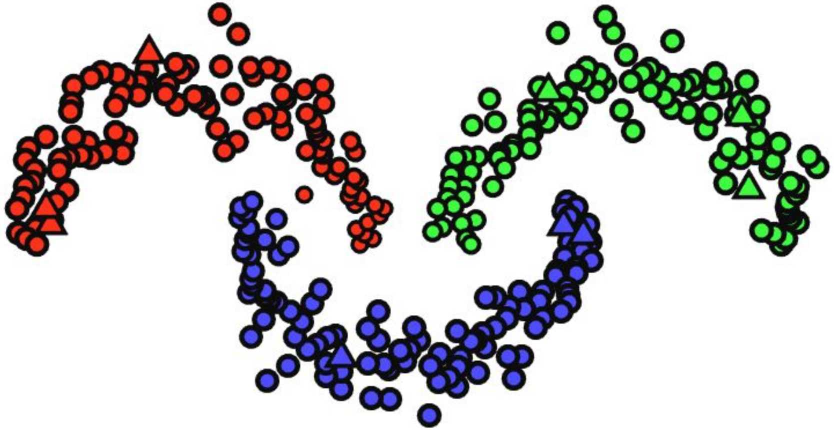
Label propagation iterations



iteration 50

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

Label propagation iterations



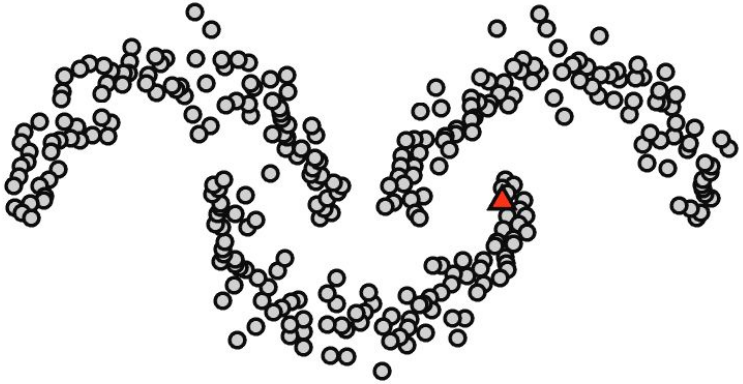
iteration 100

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

Ranking on Manifolds with Graphs

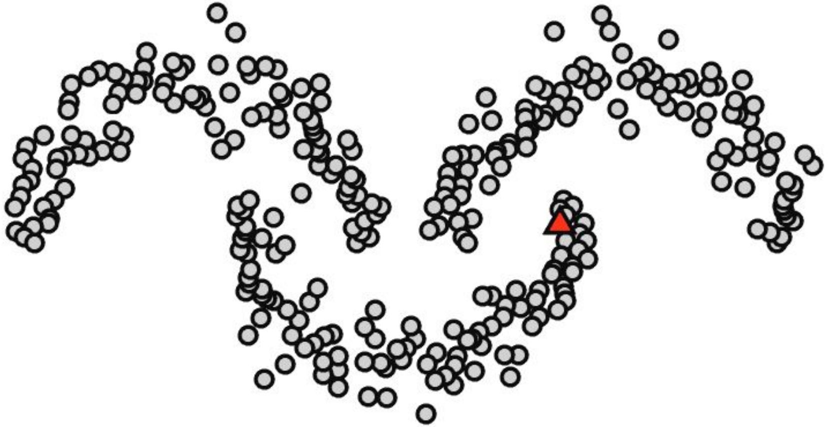
Ranking on manifolds

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



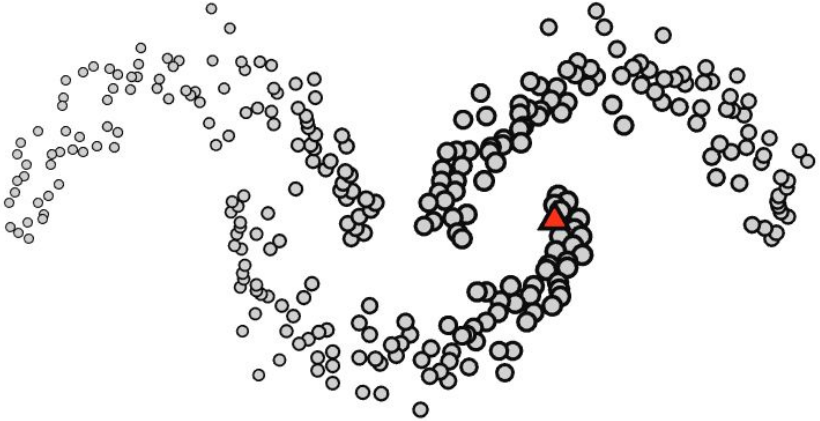
[Zhou et al, NeurIPS 2003, Ranking on data manifolds]

Euclidean similarity



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

Euclidean similarity

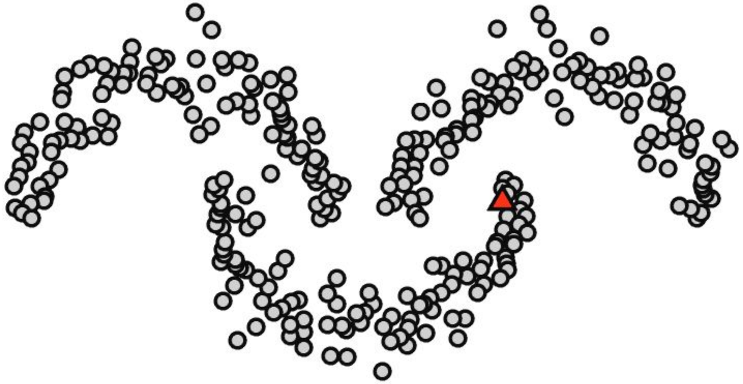


[Zhou et al, NeurIPS 2003, Ranking on data manifolds]

Ranking on manifolds

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

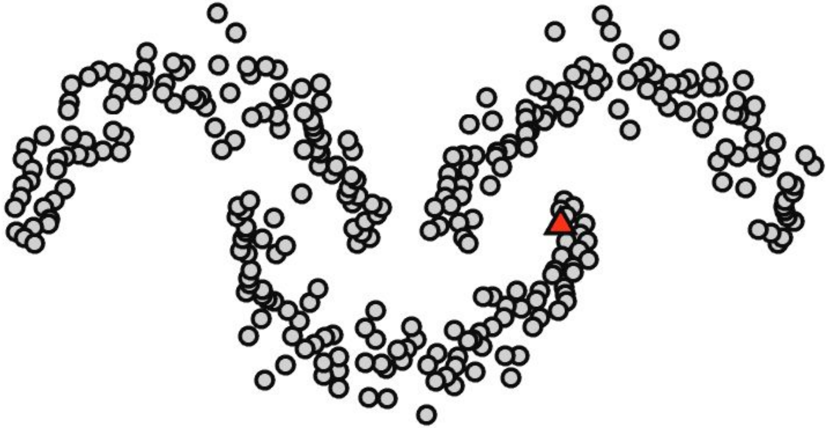
$$\mathbf{f}^t = \alpha S \mathbf{f}^{t-1} + (1-\alpha)\mathbf{y}$$



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

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

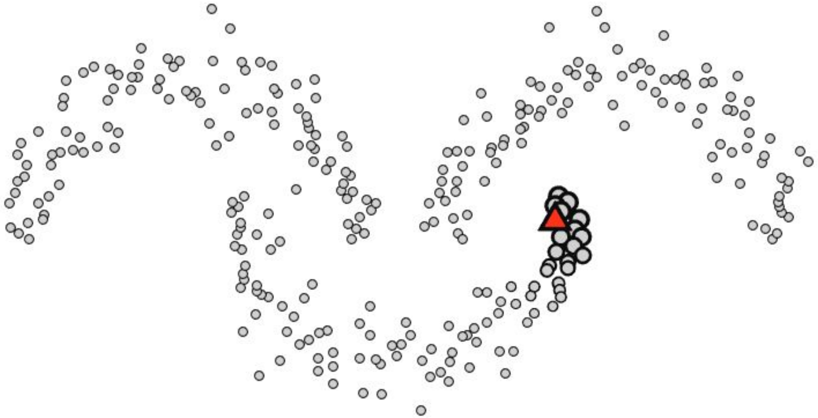
Similarity propagation iterations



iteration 0

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

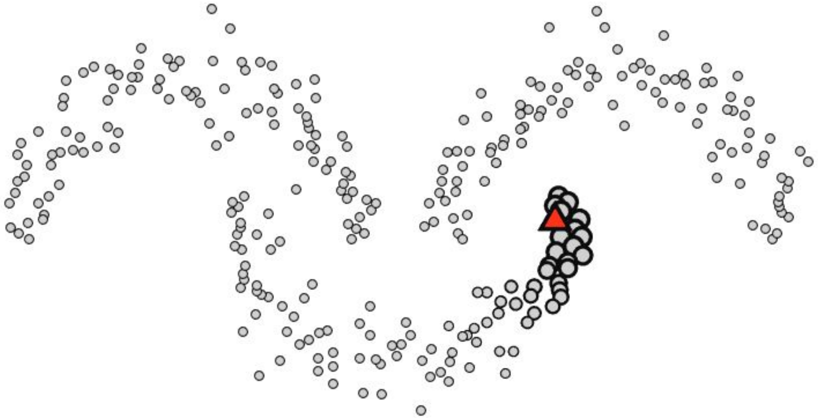
Similarity propagation iterations



iteration 1

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

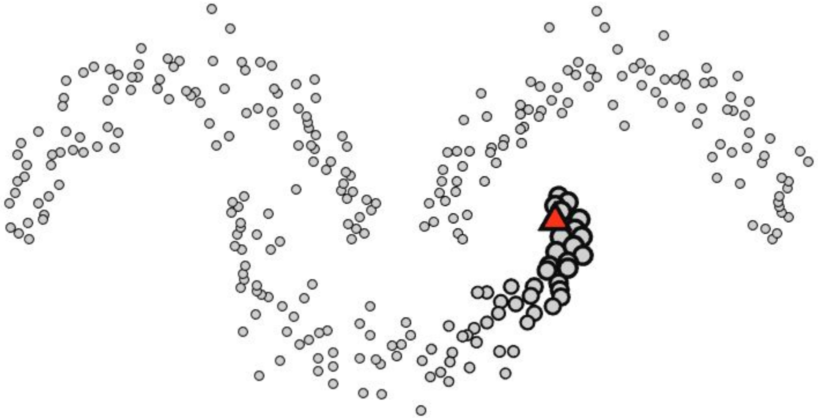
Similarity propagation iterations



iteration 5

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

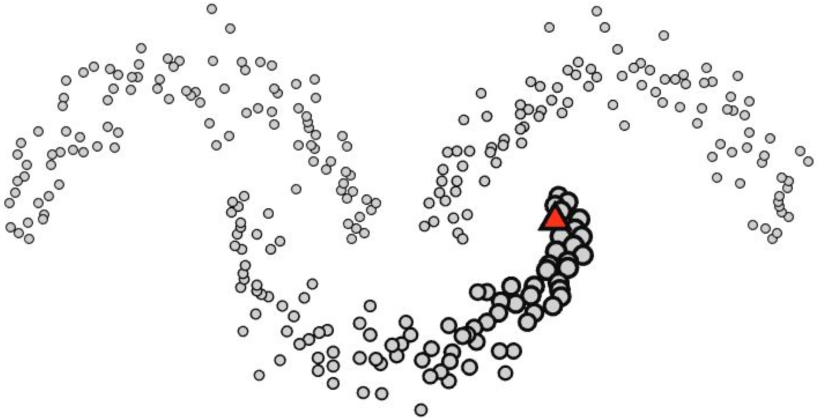
Similarity propagation iterations



iteration 10

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

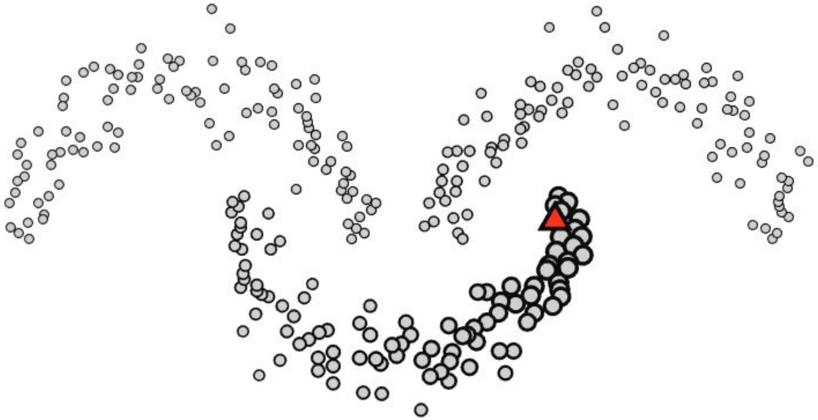
Similarity propagation iterations



iteration 50

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

Similarity propagation iterations

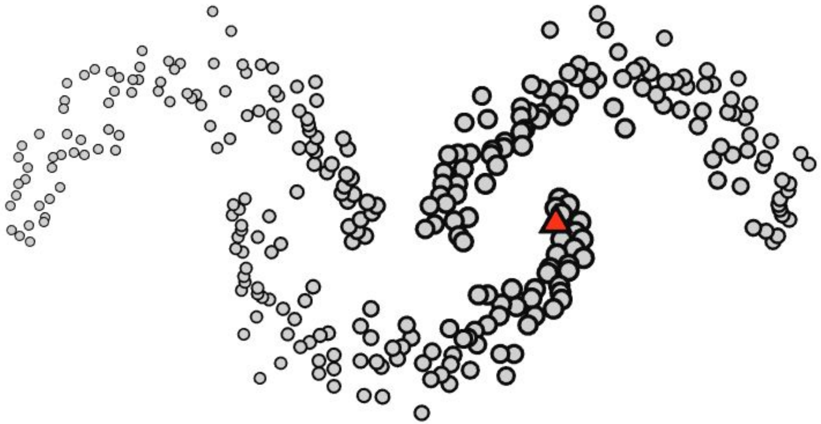


iteration 100

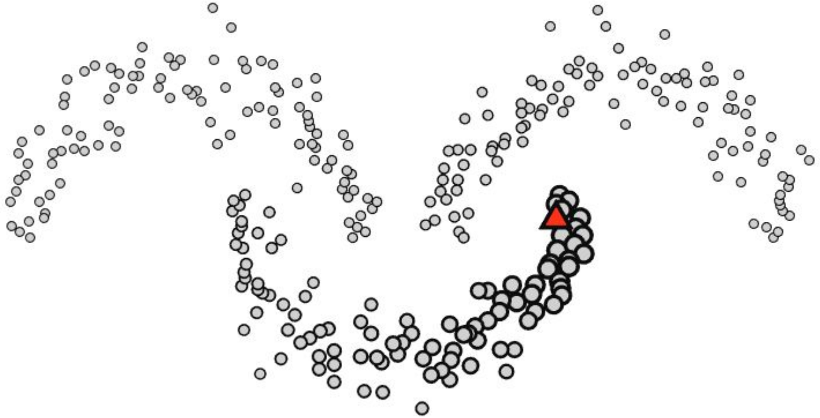
Google

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

Euclidean vs similarity propagation



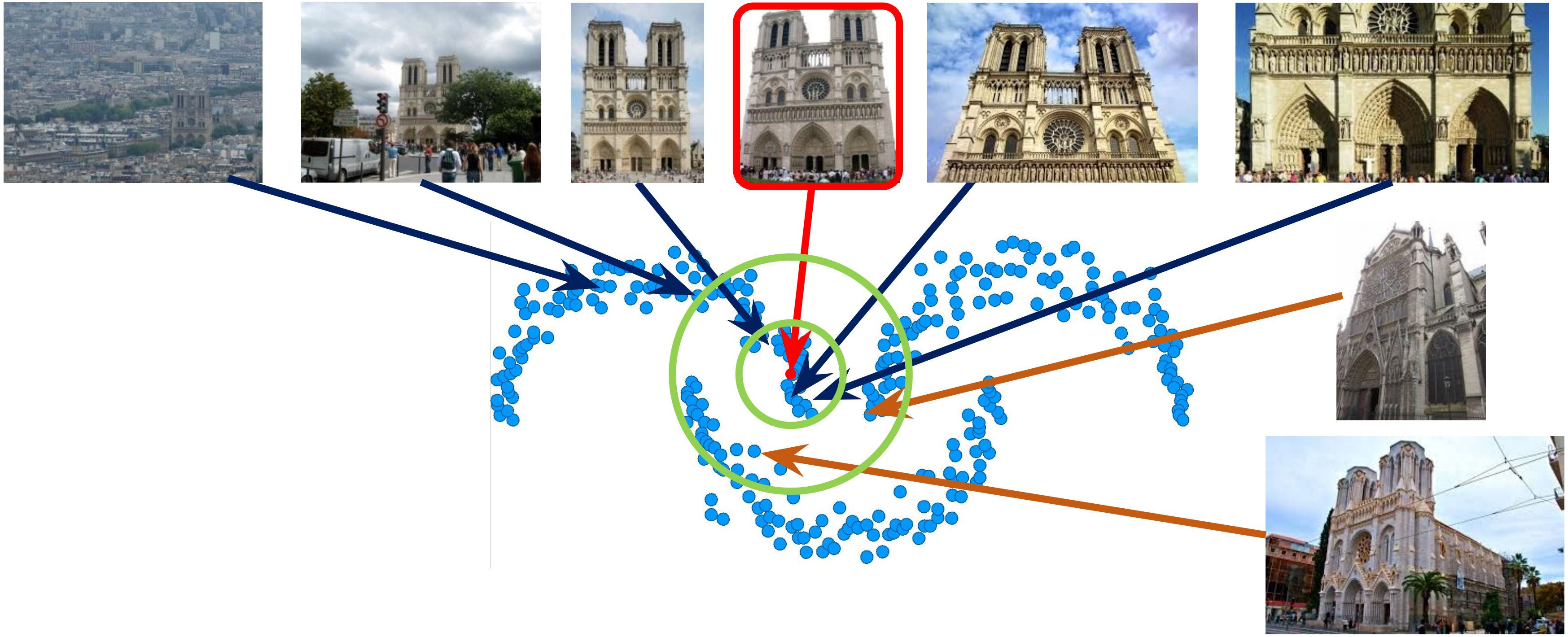
Euclidean similarity



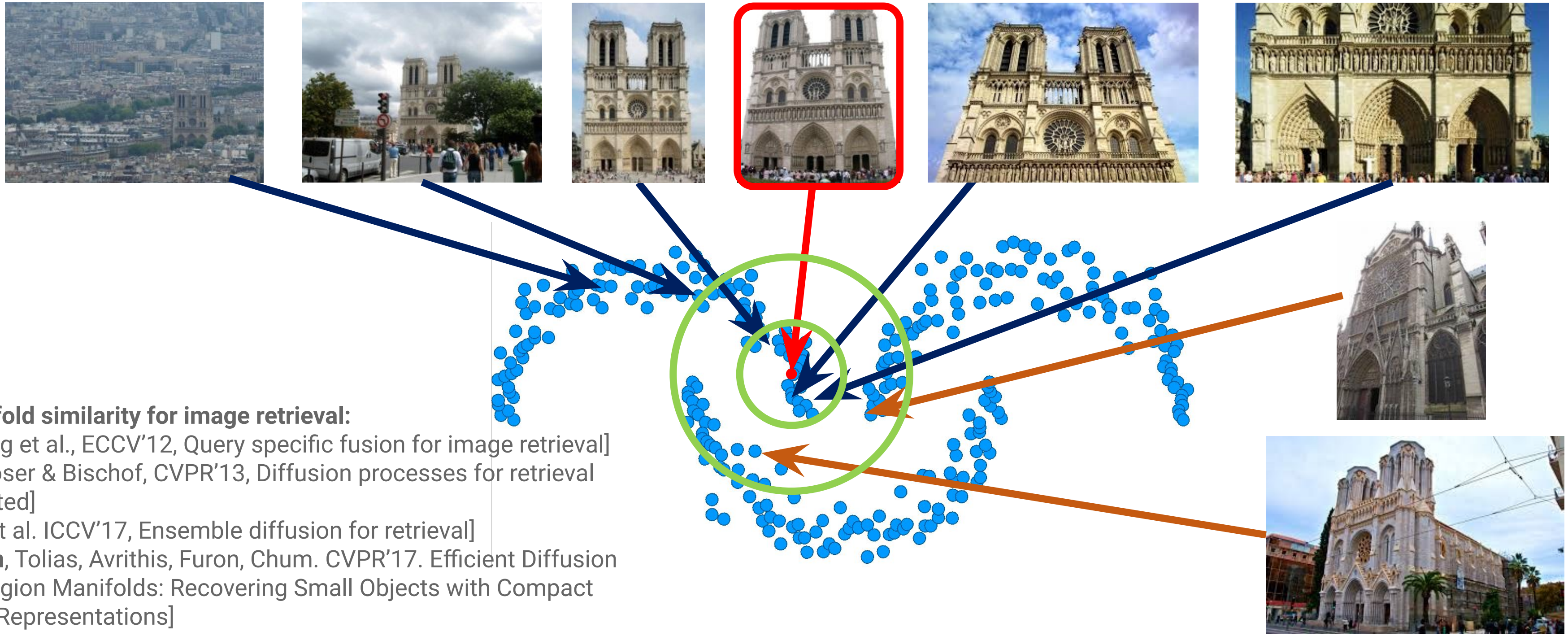
Similarity propagation

Google [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]
……

Google

Efficient solution

- Iterative solution (inefficient)

$$\mathbf{f}^t = \alpha S \mathbf{f}^{t-1} + (1-\alpha)\mathbf{y}$$

- Closed-form solution

$$\mathbf{f}^\star = \boxed{(I - \alpha S)^{-1}}\mathbf{y}$$

Not sparse,
inefficient

- Conjugate gradient for linear system

$$\boxed{(I - \alpha S)}\mathbf{f}^\star = \mathbf{y}$$

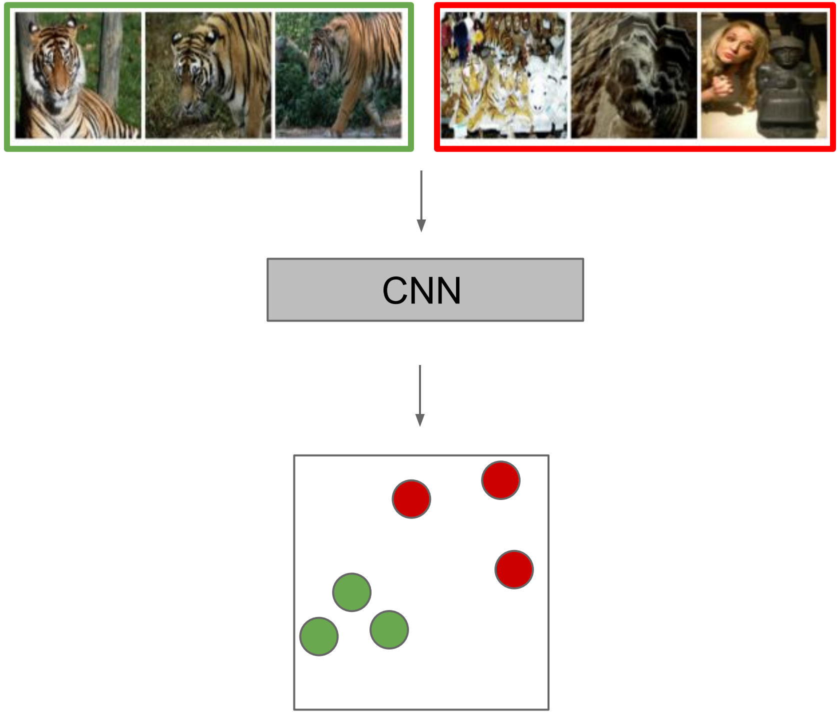
Sparse,
efficient

[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



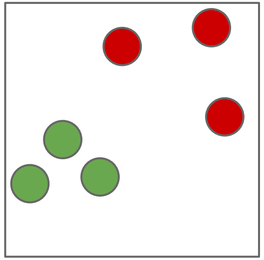
Metric learning

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

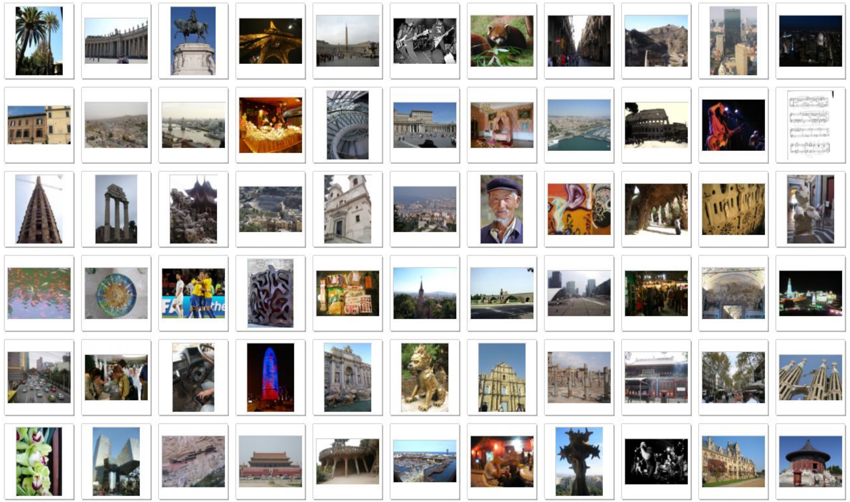
CNN

Contrastive loss

$$l_c(\mathbf{z}^r, \mathbf{z}^+, \mathbf{z}^-) = \|\mathbf{z}^r - \mathbf{z}^+\|^2 + [m - \|\mathbf{z}^r - \mathbf{z}^-\|]_+^2$$



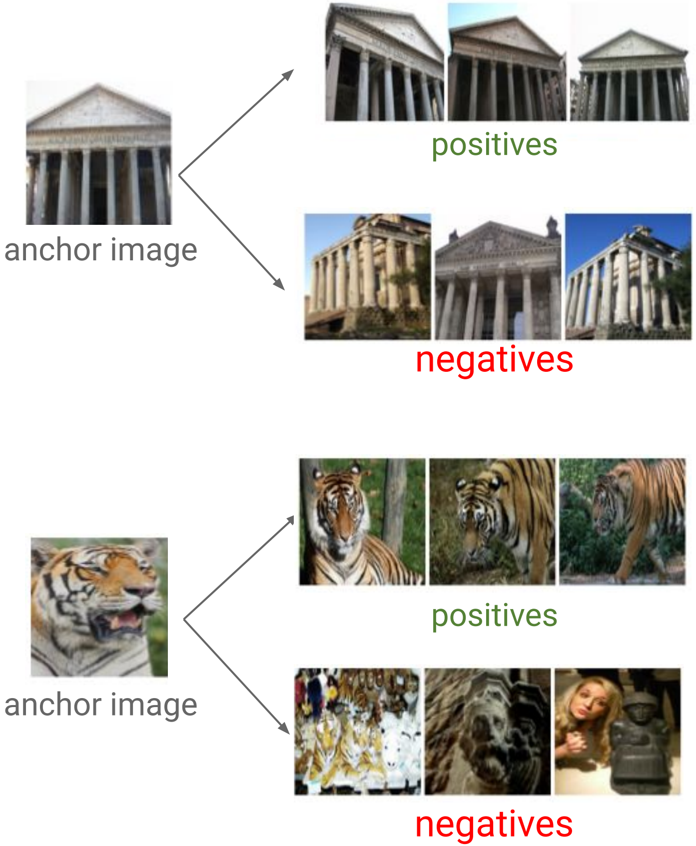
Mining on manifolds (MoM)



raw image collection

pre-trained features

MoM



- automatically mine pairs of matching and non-matching images

Google

[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 π of random walk on *A*

  $\pi = \pi P$,

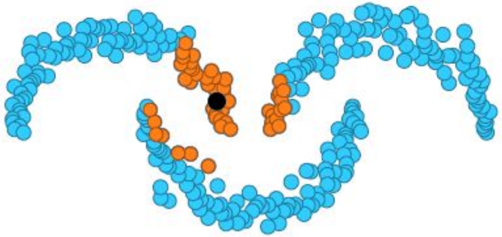
  $P = D^{-1}A$

  D = degree matrix of A

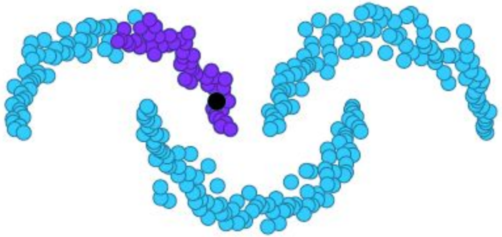
[Iscen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]

Mining on manifolds (MoM)

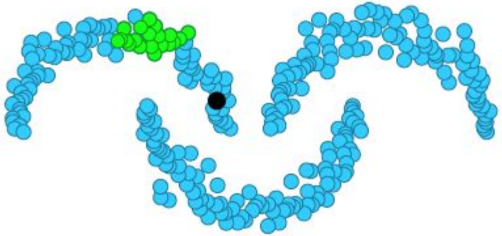
● : anchor



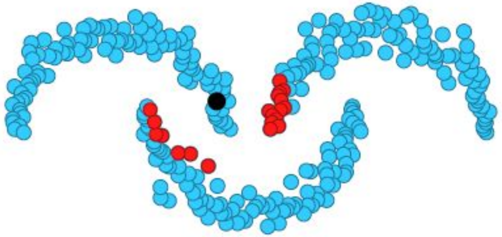
Euclidean NN (orange): $E(\mathbf{y})$



Manifold NN (purple): $M(\mathbf{y})$



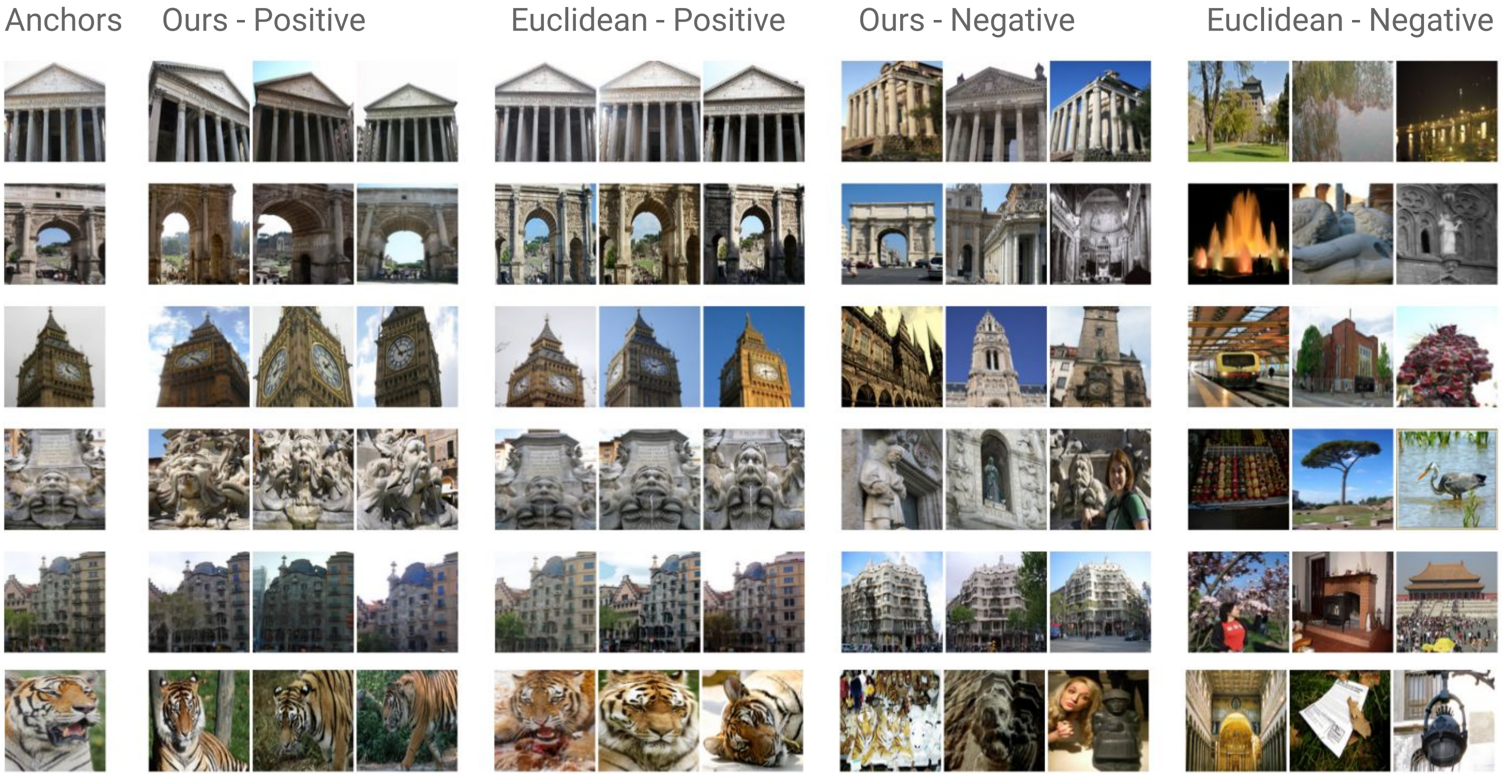
Hard positives (green): $S^+ = M(\mathbf{y}) \setminus E(\mathbf{y})$



Hard negatives (red): $S^- = E(\mathbf{y}) \setminus M(\mathbf{y})$

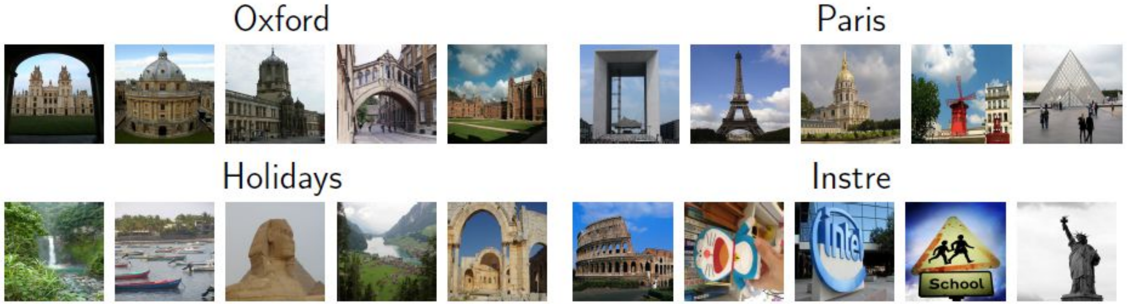
[Iscen et al, CVPR 2018, Mining on Manifolds: Metric Learning without Labels]

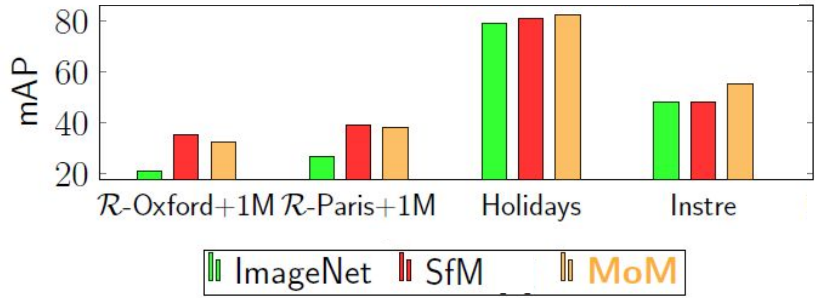
Mining on manifolds (MoM)



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

Experiments on instance search



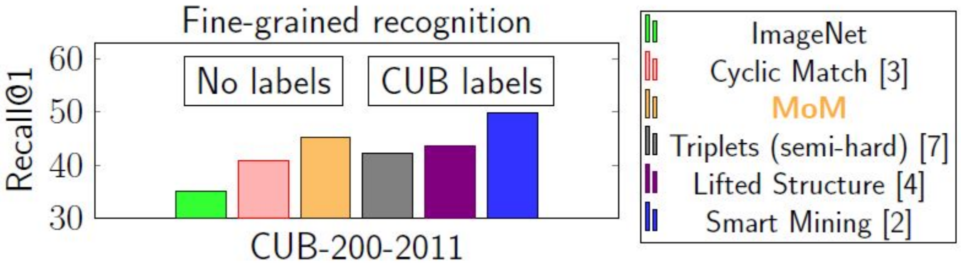


initialize: pre-training on ImageNet -- fine-tune: MoM on 10^6 images

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

Experiments on fine-grained recognition

CUB-200-2011



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]

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

Ablation experiments

Positive	Negative	CUB	Oxford5k	
Anchors		All	Random	$\mathcal{A}$
Initial		35.0	52.6	
NN_5^e	$X \setminus \mathrm{NN}_5^e$	38.5	37.4	41.9
P^+	$X \setminus \mathrm{NN}_5^e$	43.0	48.2	38.1
NN_5^e	P^-	42.1	57.8	71.3
P^+	P^-	43.5	64.4	73.7
$P^+ + W$	$P^- + W$	**45.3**	67.0	**76.7**

random anchors vs proposed anchors

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

Ablation experiments

Positive	Negative	CUB	Oxford5k	
Anchors		All	Random	$\mathcal{A}$
Initial		35.0	52.6	
NN_5^e	$X \setminus \text{NN}_5^e$	38.5	37.4	41.9
P^+	$X \setminus \text{NN}_5^e$	43.0	48.2	38.1
NN_5^e	P^-	42.1	57.8	$\boxed{71.3}$
P^+	P^-	43.5	64.4	$\boxed{73.7}$
$P^+ + W$	$P^- + W$	**45.3**	67.0	**76.7**

hard positives vs easy positives

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

Ablation experiments

Positive	Negative	CUB	Oxford5k	
Anchors		All	Random	$\mathcal{A}$
Initial		35.0	52.6	
NN_5^e	$X \setminus \mathrm{NN}_5^e$	38.5	37.4	41.9
P^+	$X \setminus \mathrm{NN}_5^e$	43.0	48.2	$\boxed{38.1}$
NN_5^e	P^-	42.1	57.8	71.3
P^+	P^-	43.5	64.4	$\boxed{73.7}$
$P^+ + W$	$P^- + W$	**45.3**	67.0	**76.7**

hard negatives vs easy negatives

[Iscen 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 learn:

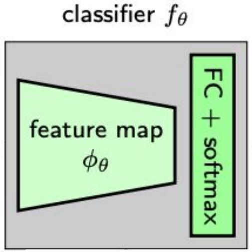
- A feature map $\phi_\theta : \mathcal{X} \rightarrow \mathbb{R}^d$
- A classifier (FC layer after the feature map)

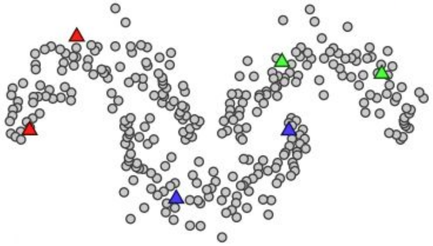
By combining:

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

[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

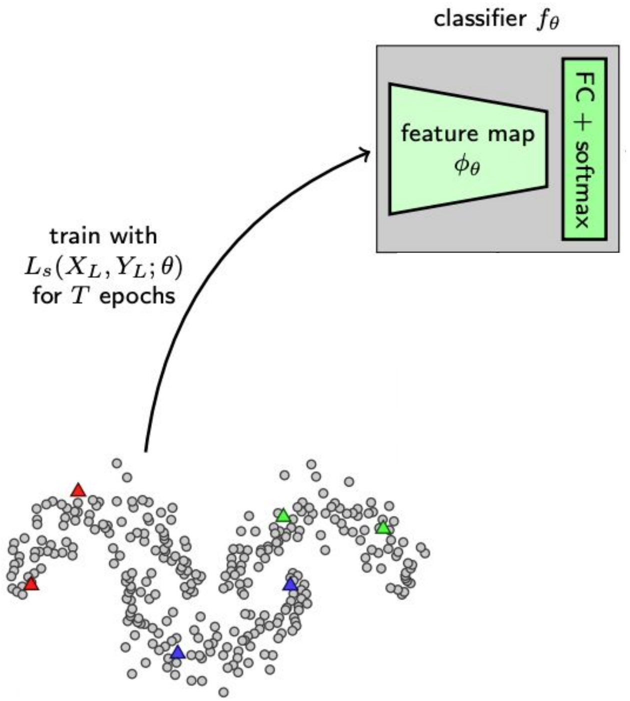
Label Propagation for Deep Semi-supervised Learning





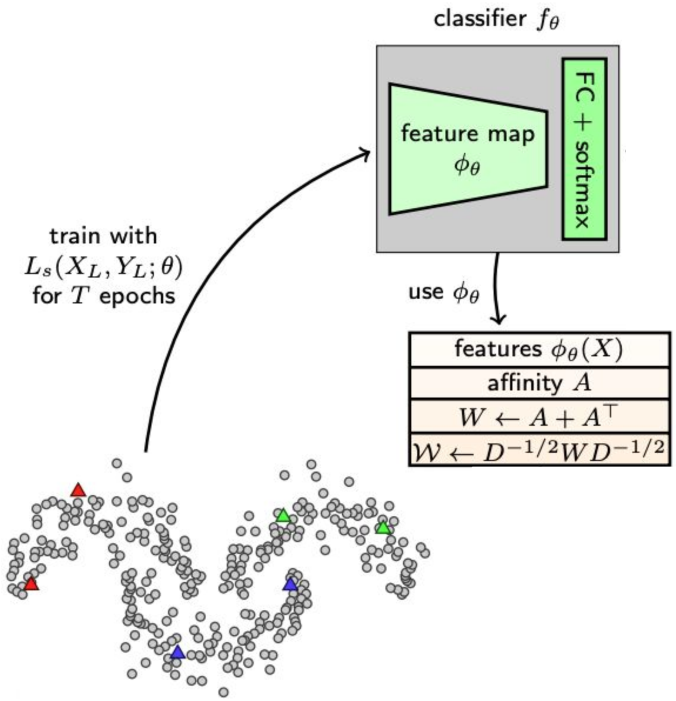
[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

Label Propagation for Deep Semi-supervised Learning



[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

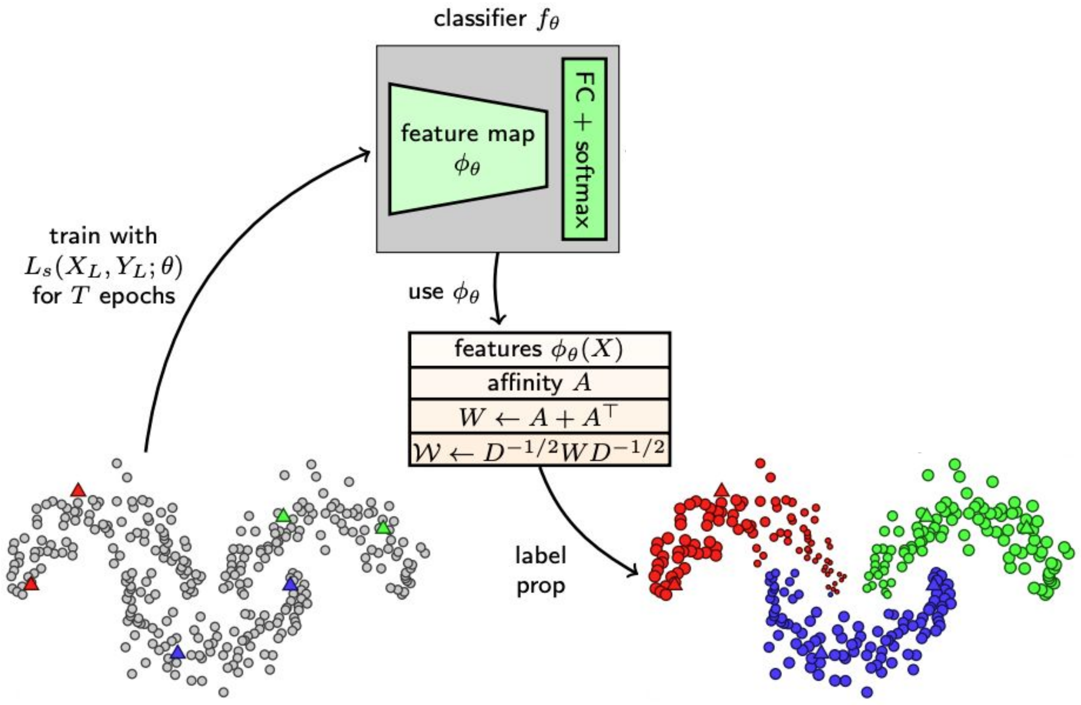
Label Propagation for Deep Semi-supervised Learning



Google

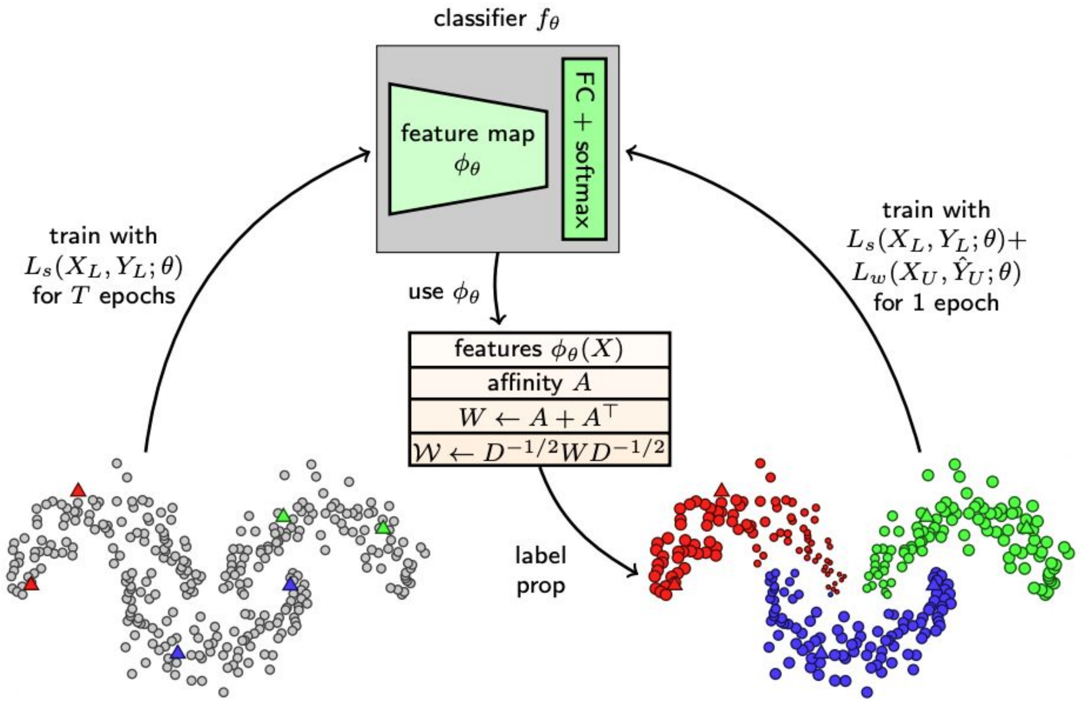
[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

Label Propagation for Deep Semi-supervised Learning



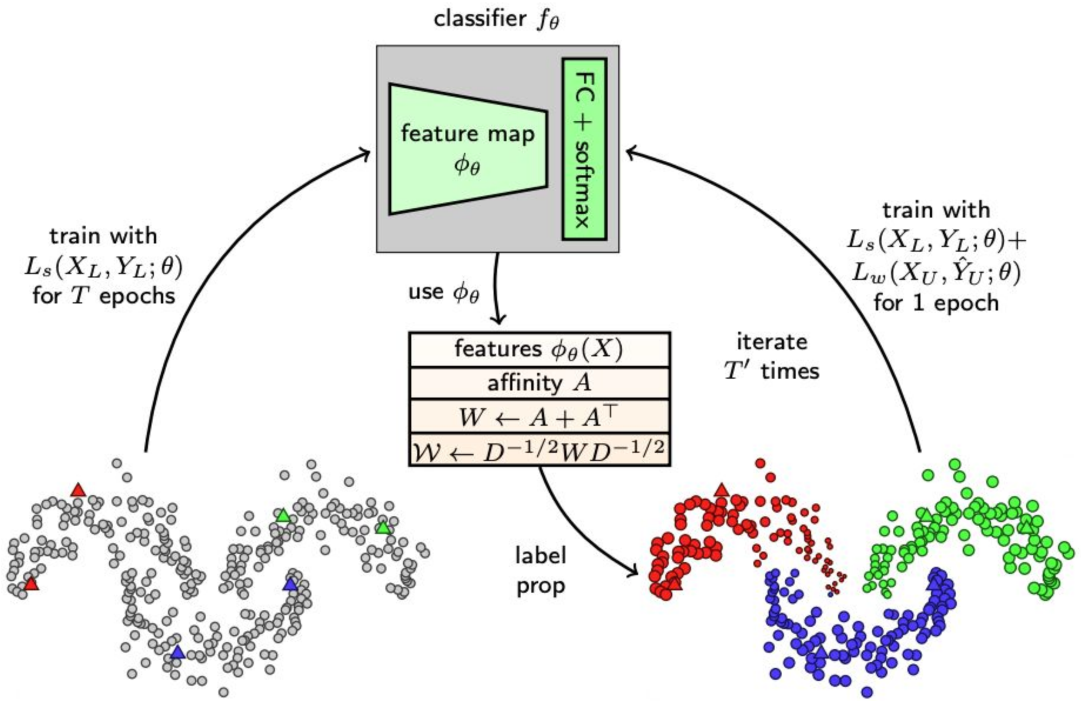
[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

Label Propagation for Deep Semi-supervised Learning



[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

Label Propagation for Deep Semi-supervised Learning



[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

Label Propagation for Deep Semi-supervised Learning

- Supervised loss

$$L_s(X_L, Y_L; \theta) := \sum_{i \in L} \boxed{\ell_s}(f_\theta(x_i), y_j)$$

Cross entropy loss

- Weighted pseudo-label loss

$$L_w(X_U, \hat{Y}_U; \theta) := \sum_{i \in U} \omega_i \zeta_{\hat{y}_i} \ell_s(f_\theta(x_i), \hat{y}_i)$$

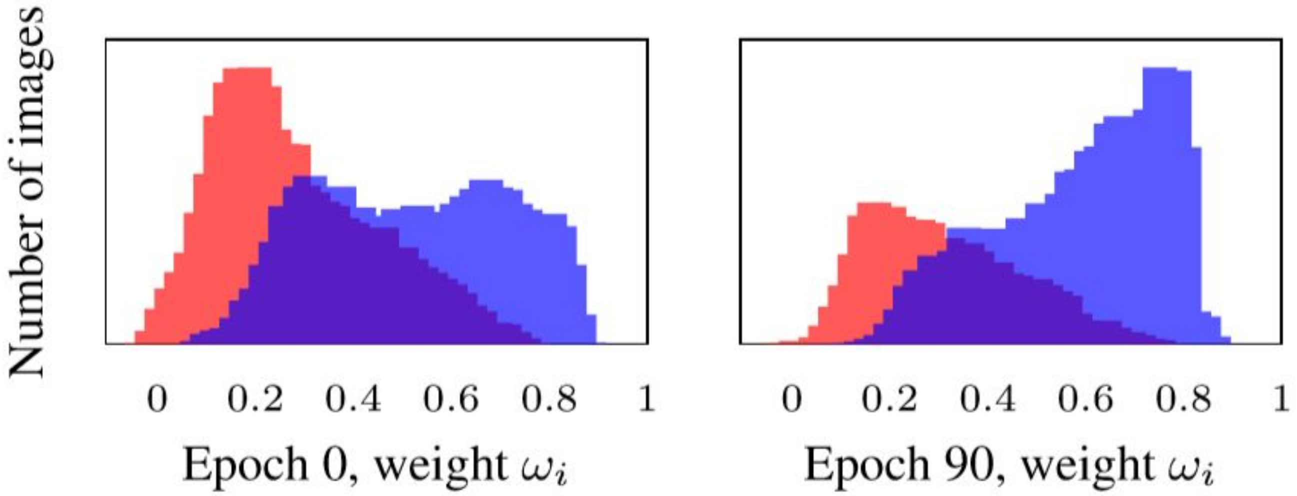
- Certainty of pseudo-label prediction

$$\omega_i := 1 - \frac{H(\hat{\mathbf{z}}_i)}{\log c}$$

- Class weight for balancing class distribution: $\zeta_j := (|L_j| + |U_j|)^{-1}$

[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

Label Propagation for Deep Semi-supervised Learning



Correctly and incorrectly pseudo-labeled examples

[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

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

example weight

class balancing

Pseudo-labeling	ω_i	ζ_j	CIFAR-10
Diffusion (7)			36.53 ± 1.42
		✓	36.17 ± 1.98
	✓		33.32 ± 1.53
	✓	✓	$\mathbf{32.40 \pm 1.80}$
Network (1)	✓	✓	35.17 ± 2.46

[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

Classification error on CIFAR10

Dataset	CIFAR-10			
Nb. labeled images	500	1000	2000	4000
Fully supervised	49.08 ± 0.83	40.03 ± 1.11	29.58 ± 0.93	21.63 ± 0.38
TDCNN [33]†	-	32.67 ± 1.93	22.99 ± 0.79	16.17 ± 0.37
Ours–(1)	35.17 ± 2.46	23.79 ± 1.31	16.64 ± 0.48	13.21 ± 0.61
Ours	32.40 ± 1.80	22.02 ± 0.88	15.66 ± 0.35	12.69 ± 0.29
VAT [23]†	-	-	-	11.36
Π model [20]†	-	-	-	12.36 ± 0.31
Temporal Ensemble [20]†	-	-	-	12.16 ± 0.24
MT [35]†	-	27.36 ± 1.30	15.73 ± 0.31	12.31 ± 0.28
MT [35]	27.45 ± 2.64	19.04 ± 0.51	14.35 ± 0.31	11.41 ± 0.25
MT + Ours	$\mathbf{24.02 \pm 2.44}$	$\mathbf{16.93 \pm 0.70}$	$\mathbf{13.22 \pm 0.29}$	$\mathbf{10.61 \pm 0.28}$

[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

Classification error on CIFAR100 and mini-ImageNet

Dataset	CIFAR-100		Mini-ImageNet-*top1*	
Nb. labeled images	4000	10000	4000	10000
Fully supervised	55.43 ± 0.11	40.67 ± 0.49	74.78 ± 0.33	60.25 ± 0.29
Ours	46.20 ± 0.76	38.43 ± 1.88	$\mathbf{70.29 \pm 0.81}$	57.58 ± 1.47
MT [35]	45.36 ± 0.49	36.08 ± 0.51	72.51 ± 0.22	57.55 ± 1.11
MT + Ours	$\mathbf{43.73 \pm 0.20}$	$\mathbf{35.92 \pm 0.47}$	72.78 ± 0.15	$\mathbf{57.35 \pm 1.66}$

Google

[Iscen et al, CVPR 2019, Label Propagation for Deep Semi-supervised Learning]

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

Giorgos Tolias

Yannis Avrithis

Teddy Furon

Ondřej Chum

Cordelia Schmid

co-authors

Slide credits: Giorgos Tolias and Yannis Avrithis

Questions