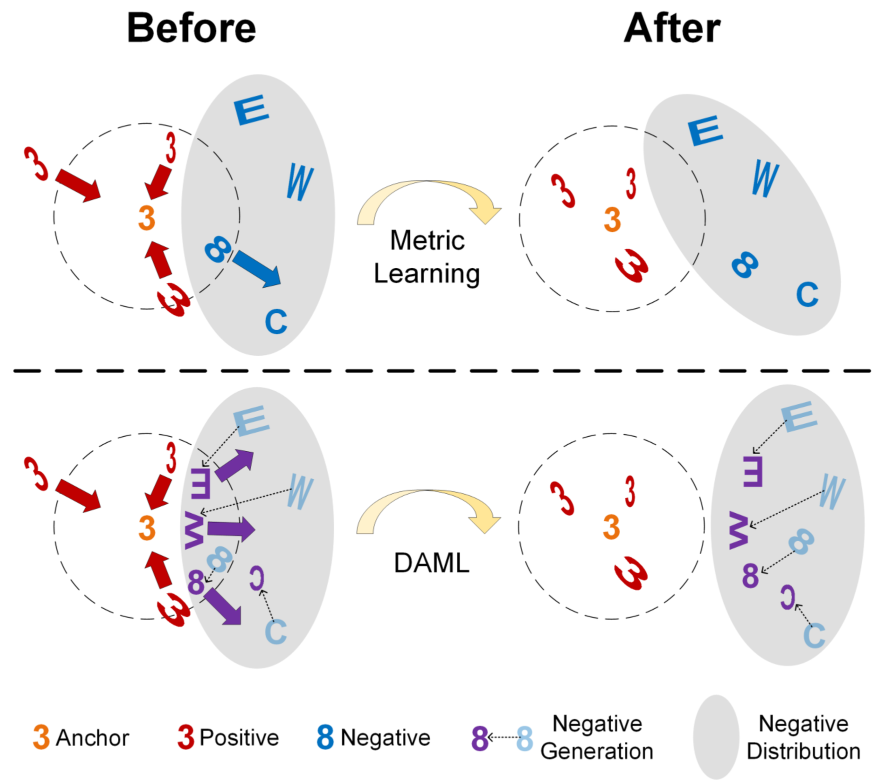
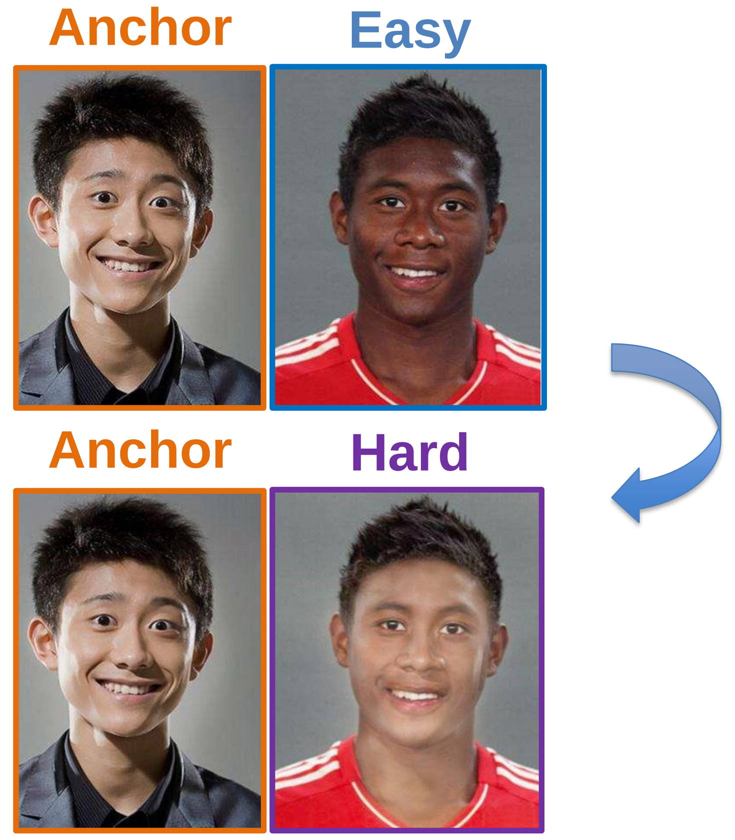
Deep Adversarial Metric Learning

Yueqi Duan, Wenzhao Zheng, Xudong Lin, Jiwen Lu, Jie Zhou

Motivation

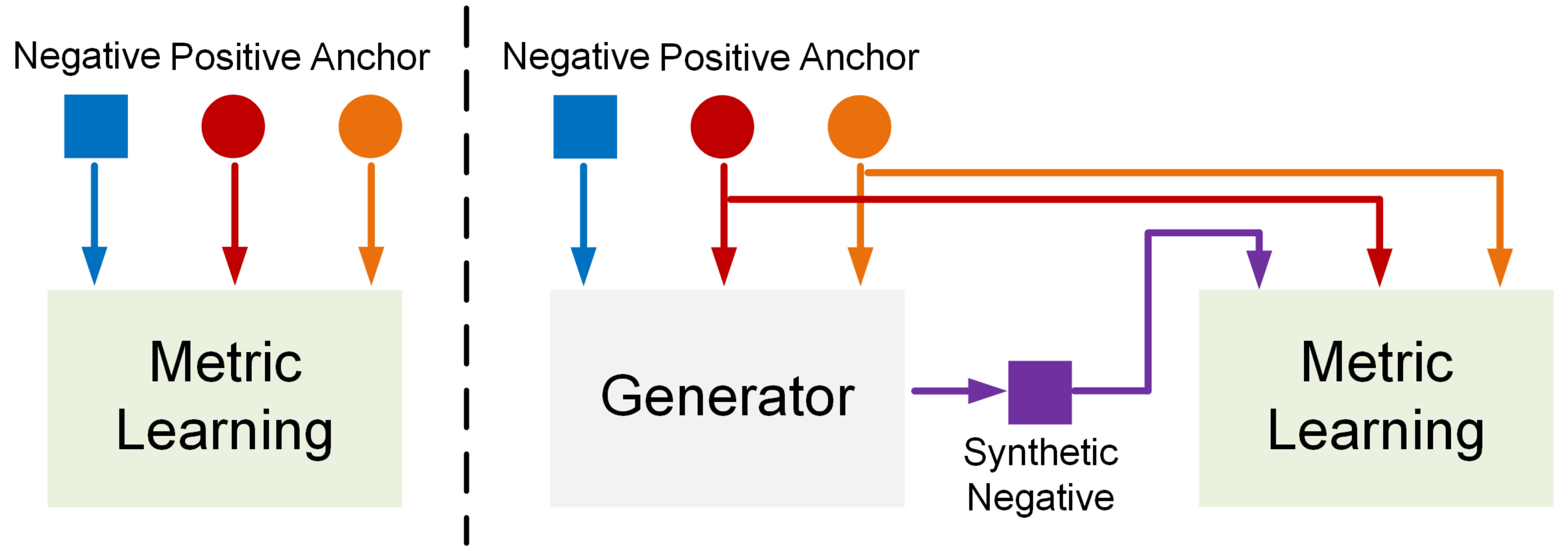
- Deep metric learning methods usually utilize **hard negative mining** strategy to better exploit large-scale negative data[1]





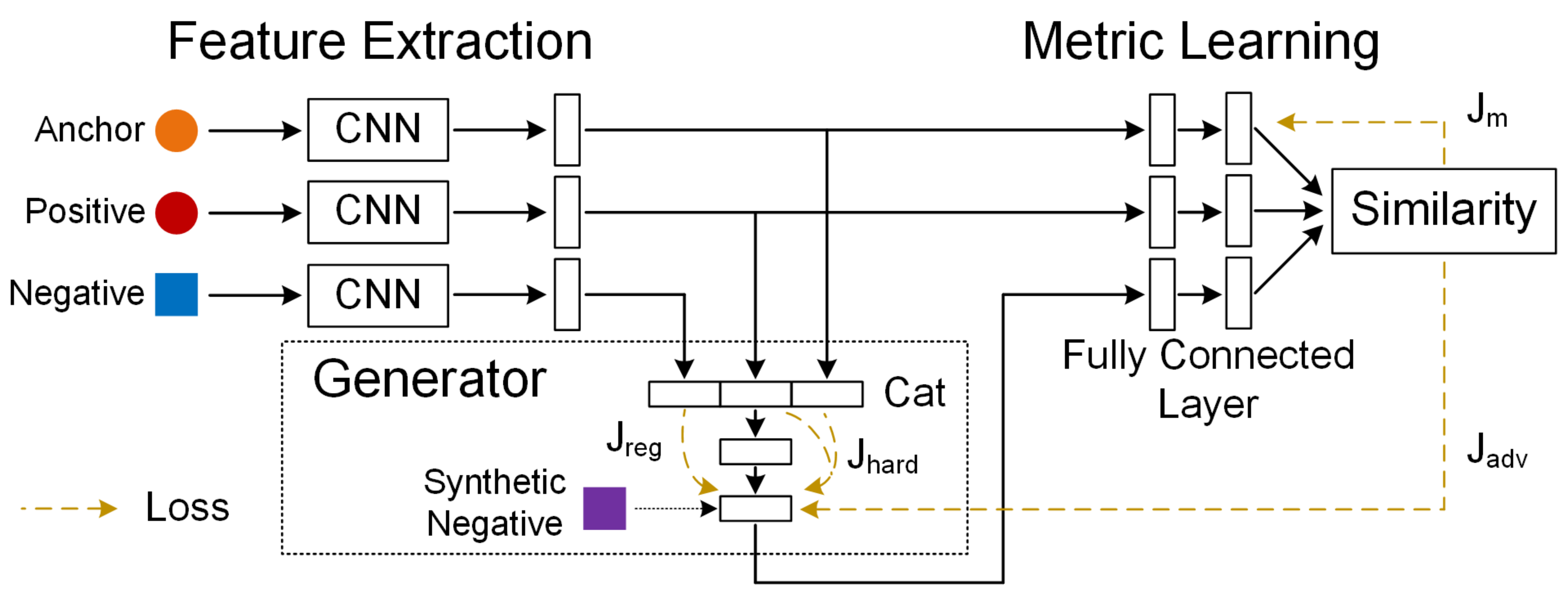
- There are two key limitations of only mining observed data:
  - The observed hard negatives may not be enough to fully describe the distributions of negative samples near the margin
  - A large number of easy negative samples are wasted which produce gradients close to zero

Flowchart



- Utilize the generated hard negatives to train the metric
- Jointly train the generator and the metric adversarially

Network Architecture



- Take the input with features extracted from CNNs and generate synthetic hard negatives for deep metric learning
- Employ existing metric learning losses generally

Objective Function

- Overall

$$\min_{\theta_g,\theta_f} J = J_{\text{gen}} + \lambda J_{\text{m}}$$

- Hard negative generator

$$\begin{aligned}\min_{\theta_g} J_{\text{gen}} &= J_{\text{hard}} + \lambda_1 J_{\text{reg}} + \lambda_2 J_{\text{adv}} \\ &= \sum_{i=1}^{N}(||\tilde{\mathbf{x}}_i^- - \mathbf{x}_i||_2^2 + \lambda_1 ||\tilde{\mathbf{x}}_i^- - \mathbf{x}_i^-||_2^2 \\ &+ \lambda_2 [D(\tilde{\mathbf{x}}_i^-, \mathbf{x}_i)^2 - D(\mathbf{x}_i^+, \mathbf{x}_i)^2 - \alpha]_+)\end{aligned}$$

  - J_{hard}: Encourage hard negative sample in the original space
  - J_{reg}: Regularize the generated negative sample not too far away
  - J_{adv}: Make the generator adversarial to the learned metric

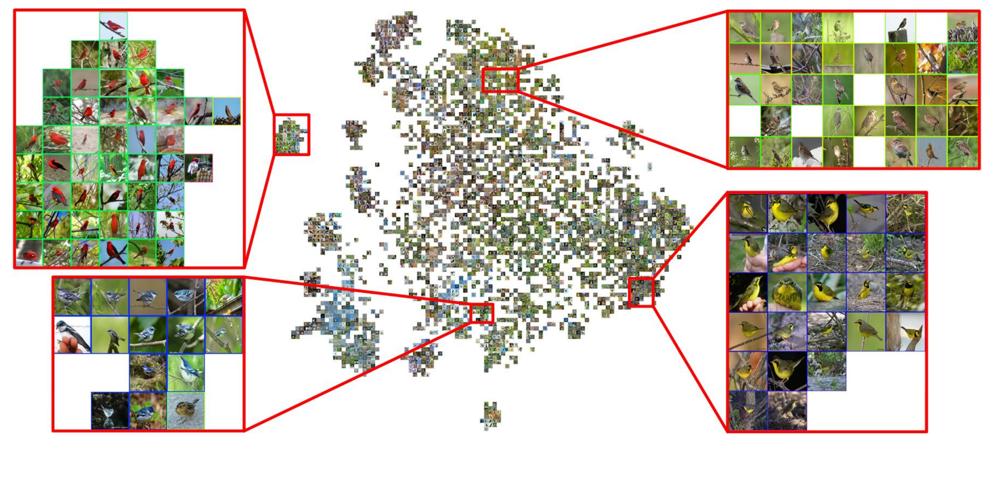
References

[1] F. Schroff, D. Kalenichenko, and J. Philbin, Facenet: A unified embedding for face recognition and clustering, CVPR, 2015.

Experiments

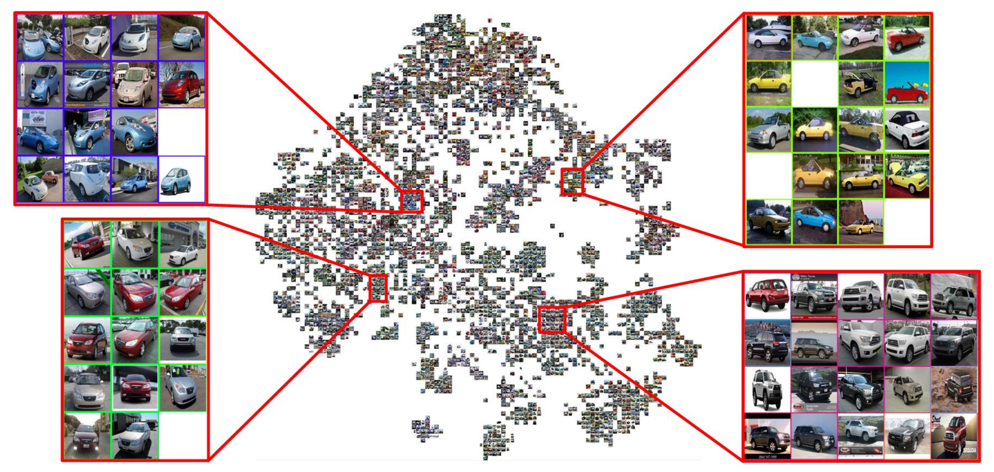
- The CUB-200-2011 dataset

Method	NMI	F_1	R@1	R@2	R@4	R@8
DDML	47.3	13.1	31.2	41.6	54.7	67.1
Triplet+N-pair	54.1	20.0	42.8	54.9	66.2	77.6
Angular	61.0	30.2	53.6	65.0	75.3	83.7
Contrastive	47.2	12.5	27.2	36.3	49.8	62.1
DAML (cont)	**49.1**	**16.2**	**35.7**	**48.4**	**60.8**	**73.6**
Triplet	49.8	15.0	35.9	47.7	59.1	70.0
DAML (tri)	**51.3**	**17.6**	**37.6**	**49.3**	**61.3**	**74.4**
Lifted	56.4	22.6	46.9	59.8	71.2	81.5
DAML (lifted)	**59.5**	**26.6**	**49.0**	**62.2**	**73.7**	**83.3**
N-pair	60.2	28.2	51.9	64.3	74.9	83.2
DAML (N-pair)	**61.3**	**29.5**	**52.7**	**65.4**	**75.5**	**84.3**



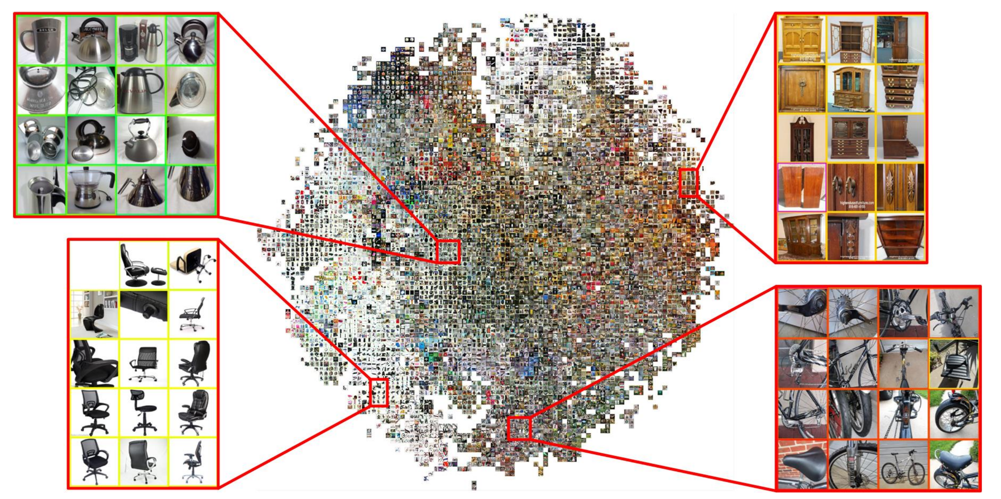
- The Cars196 dataset

Method	NMI	F_1	R@1	R@2	R@4	R@8
DDML	41.7	10.9	32.7	43.9	56.5	68.8
Triplet+N-pair	54.3	19.6	46.3	59.9	71.4	81.3
Angular	62.4	31.8	71.3	80.7	87.0	91.8
Contrastive	42.3	10.5	27.6	38.3	51.0	63.9
DAML (cont)	**42.6**	**11.4**	**37.2**	**49.6**	**61.8**	**73.3**
Triplet	52.9	17.9	45.1	57.4	69.7	79.2
DAML (tri)	**56.5**	**22.9**	**60.6**	**72.5**	**82.5**	**89.9**
Lifted	57.8	25.1	59.9	70.4	79.6	87.0
DAML (lifted)	**63.1**	**31.9**	**72.5**	**82.1**	**88.5**	**92.9**
N-pair	62.7	31.8	68.9	78.9	85.8	90.9
DAML (N-pair)	**66.0**	**36.4**	**75.1**	**83.8**	**89.7**	**93.5**



- The Stanford Online Products dataset

Method	NMI	F_1	R@1	R@10	R@100
DDML	83.4	10.7	42.1	57.8	73.7
Triplet+N-pair	86.4	21.0	58.1	76.0	89.1
Angular	87.8	26.5	67.9	83.2	92.2
Contrastive	82.4	10.1	37.5	53.9	71.0
DAML (cont)	**83.5**	**10.9**	**41.7**	**57.5**	**73.5**
Triplet	86.3	20.2	53.9	72.1	85.7
DAML (tri)	**87.1**	**22.3**	**58.1**	**75.0**	**88.0**
Lifted	87.2	25.3	62.6	80.9	91.2
DAML (lifted)	**89.1**	**31.7**	**66.3**	**82.8**	**92.5**
N-pair	87.9	27.1	66.4	82.9	92.1
DAML (N-pair)	**89.4**	**32.4**	**68.4**	**83.5**	**92.3**



Email: duanyq14@mails.tsinghua.edu.cn