清華大學
Tsinghua University

Deep Adversarial Metric Learning

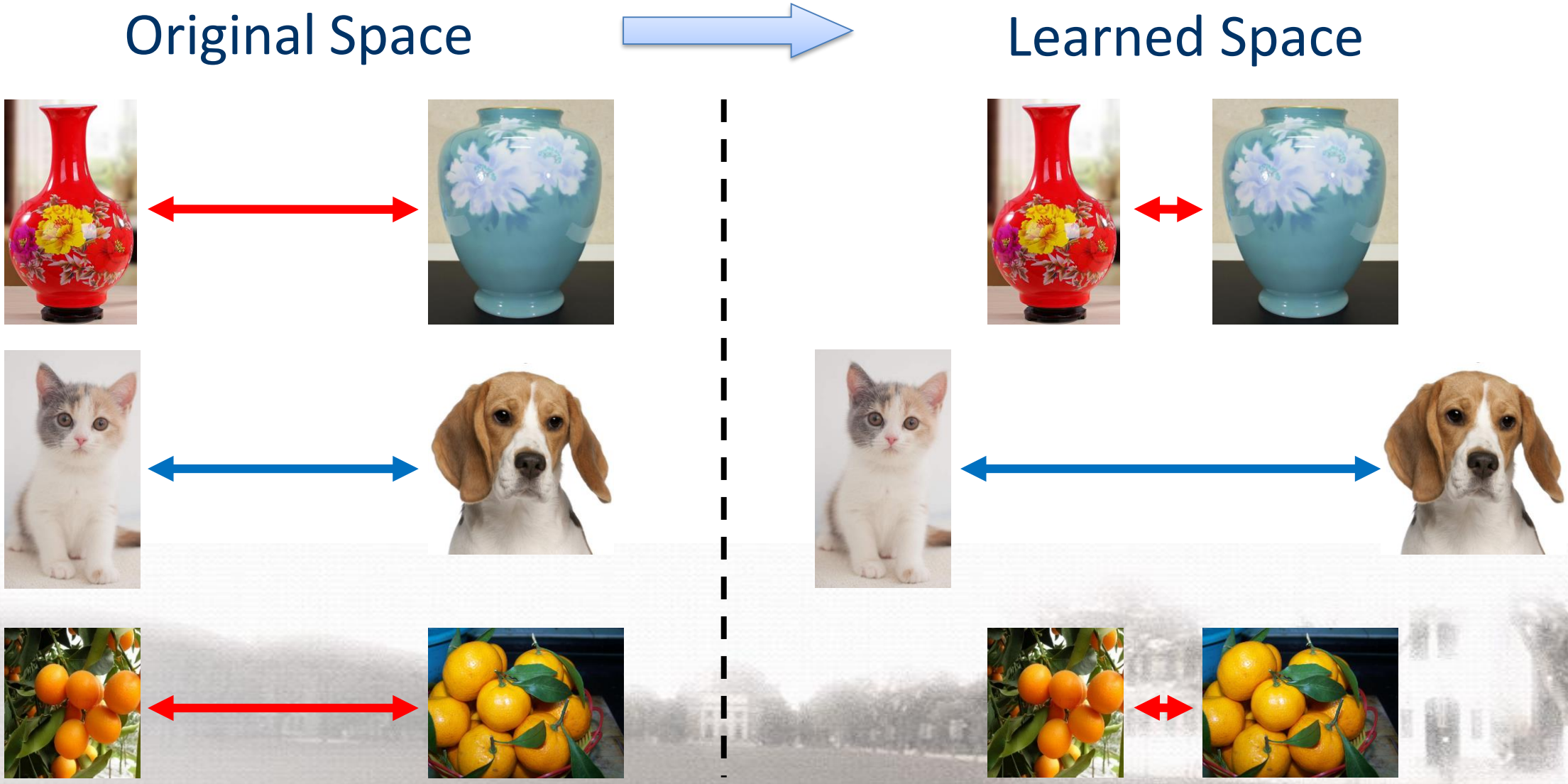
Yueqi Duan, *Wenzhao Zheng, Xudong Lin,*

Jiwen Lu, and Jie Zhou

Department of Automation, Tsinghua University, China

http://ivg.au.tsinghua.edu.cn/people/Yueqi_Duan/

Metric Learning

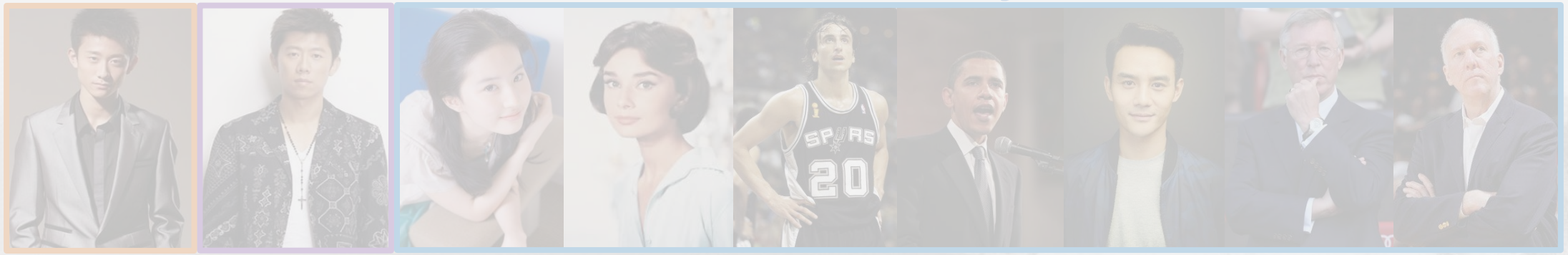


Sampling Matters

- Objective functions
  - Contrastive loss: $L_{\text{cont}}(x_i, x_j; \theta) = \mathbf{1}\{y_i = y_j\} D_{ij}^2 + \mathbf{1}\{y_i \neq y_j\} [\alpha - D_{ij}]_+^2$
  - Triplet loss: $L_{\text{tri}}(x_a, x_p, x_n; \theta) = [D_{ap}^2 - D_{an}^2 + \alpha]_+$

Easy negatives account for the vast majority.

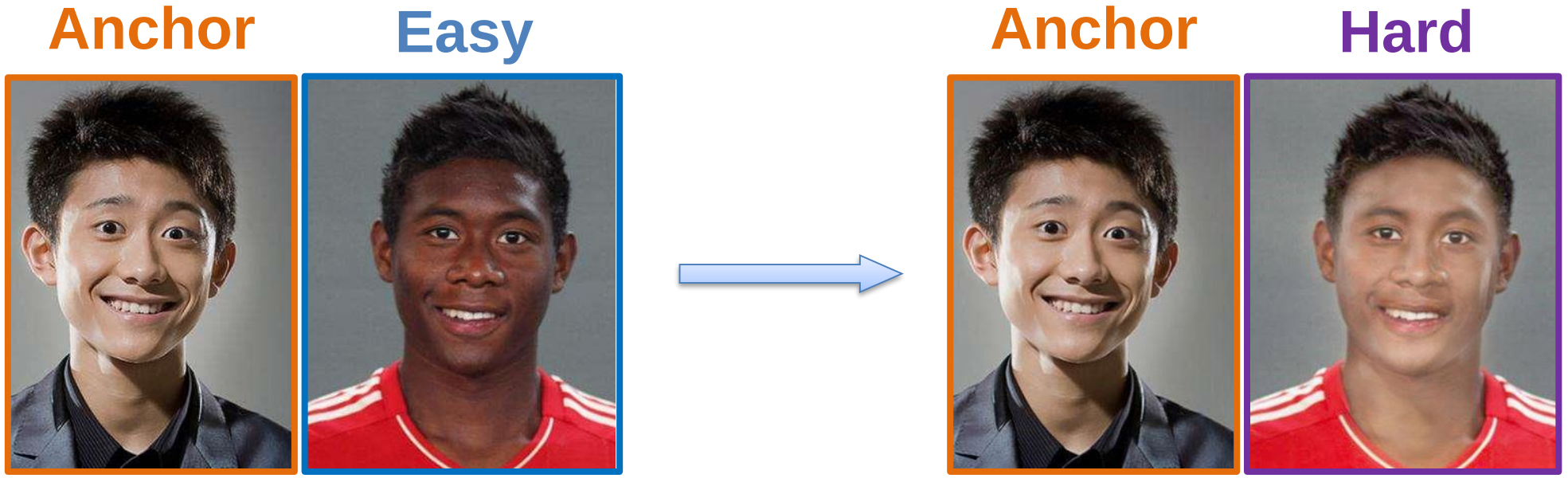
Are easy negatives really useless?



[Wu et al., ICCV'17; Harwood et al., ICCV'17; Yuan et al., ICCV'17]

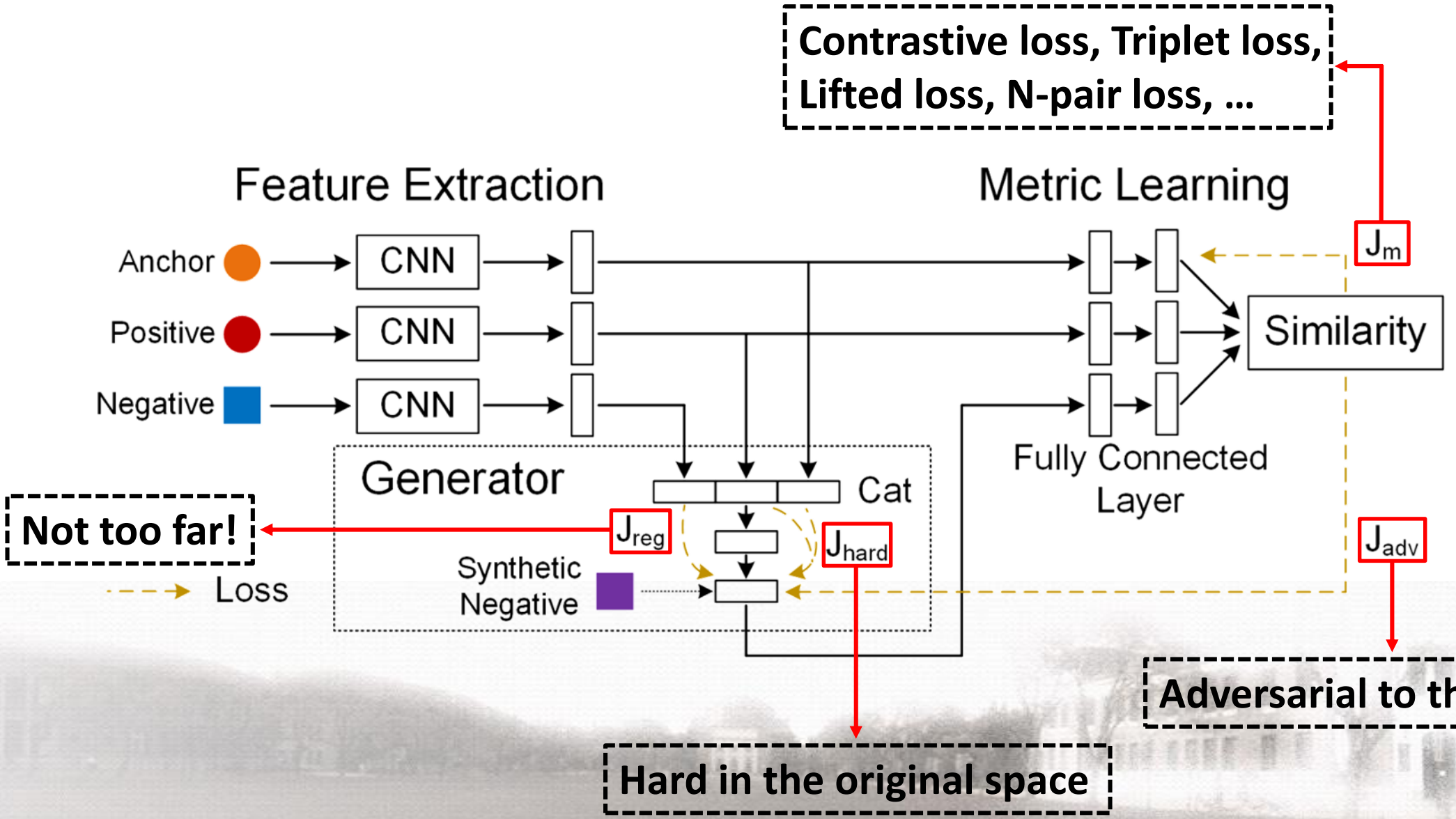
Hard Negative Mining!

- Easy negatives may have potential to become hard negatives



- Ignore easy negatives?
- DAML: Exploit easy negatives through adversarial hard negative generation

Deep Adversarial Metric Learning



Objective Function

- Overall:

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

- Hard negative generator:

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

Datasets

- CUB-200-2011 includes 11,788 images of 200 bird species.
- Cars196 contains 16,185 images of 196 cars models.
- Stanford Online Products has 120,053 images of 22,634 products from eBay.com.

Experiments on CUB-200-2011

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**

Experiments on Cars196

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**

Experiments on Stanford Online Products

清華大學 Tsinghua University

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**

Summary

- Hard negative mining ignores large numbers of easy negatives
- DAML taps their potential through adversarial hard negative generation
- The synthetic hard negatives provide essential complements for effective metric learning