A Benchmark and Asymmetrical-Similarity Learning for Practical Image Copy Detection

Wenhao Wang, Yifan Sun, Yi Yang

Introduction

- We contribute a new ICD dataset, i.e., Negative-Distractor for Edited Copy (**NDEC**), with emphasis on the seldom-noticed hard negative problem (while preserving the popular hard positive problem).
- We benchmark NDEC with state-of-the-art methods and correspondingly reveals a fundamental conflict between the commonly-adopted symmetric distance and the asymmetric "reference → edited copy" process.
- We propose a novel Asymmetric-Similarity Learning (**ASL**) for ICD. ASL uses the norm ratio as an asymmetric similarity metric to distinguish edited copy against hard negative samples and substantially improves ICD.

Hard Negative v.s. Positive

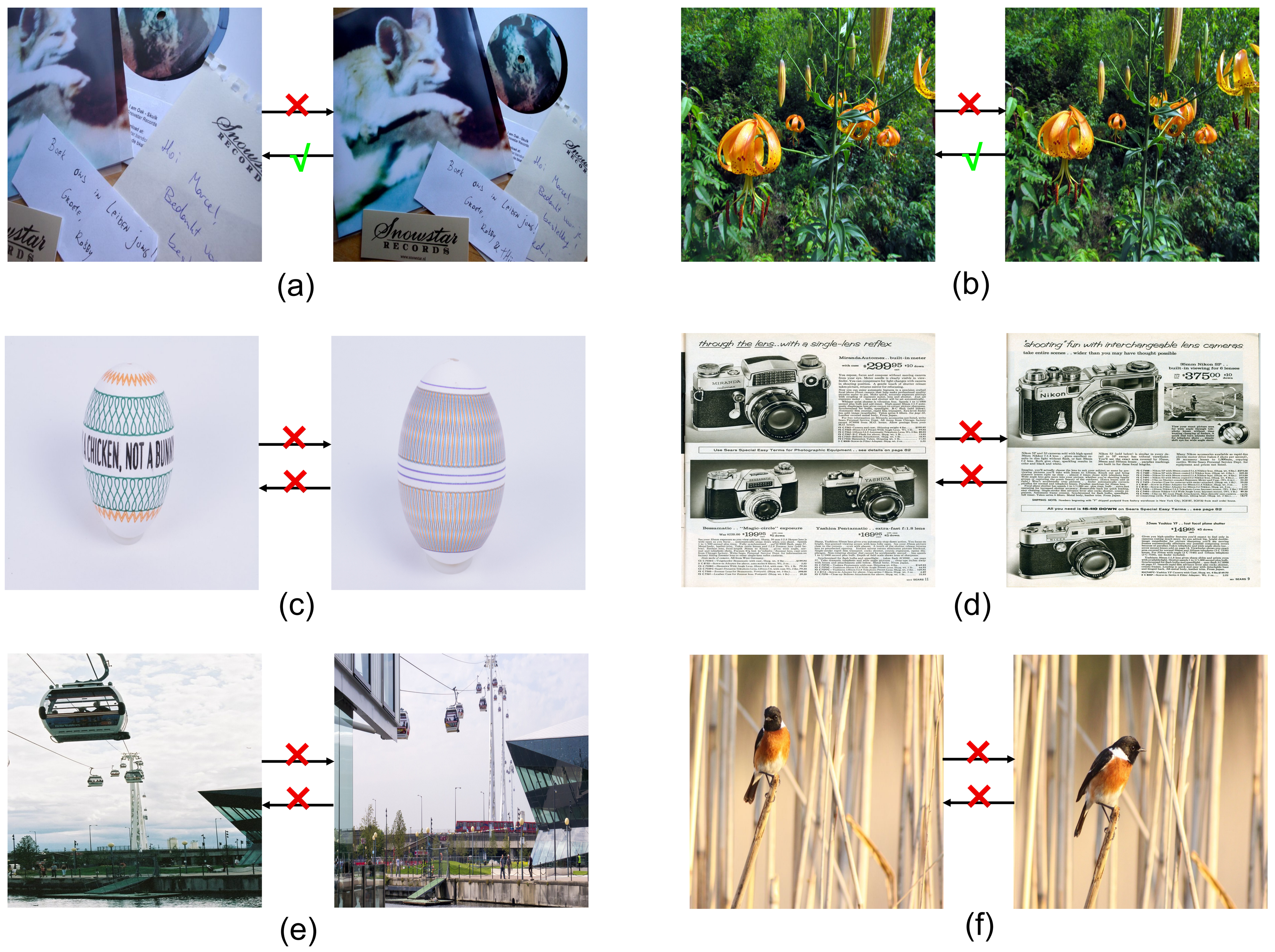
- Respective examples for edited copy and hard negative sample.

A Unique Challenge: Unidirection

- The "reference → edited copy" is an unidirectional process.

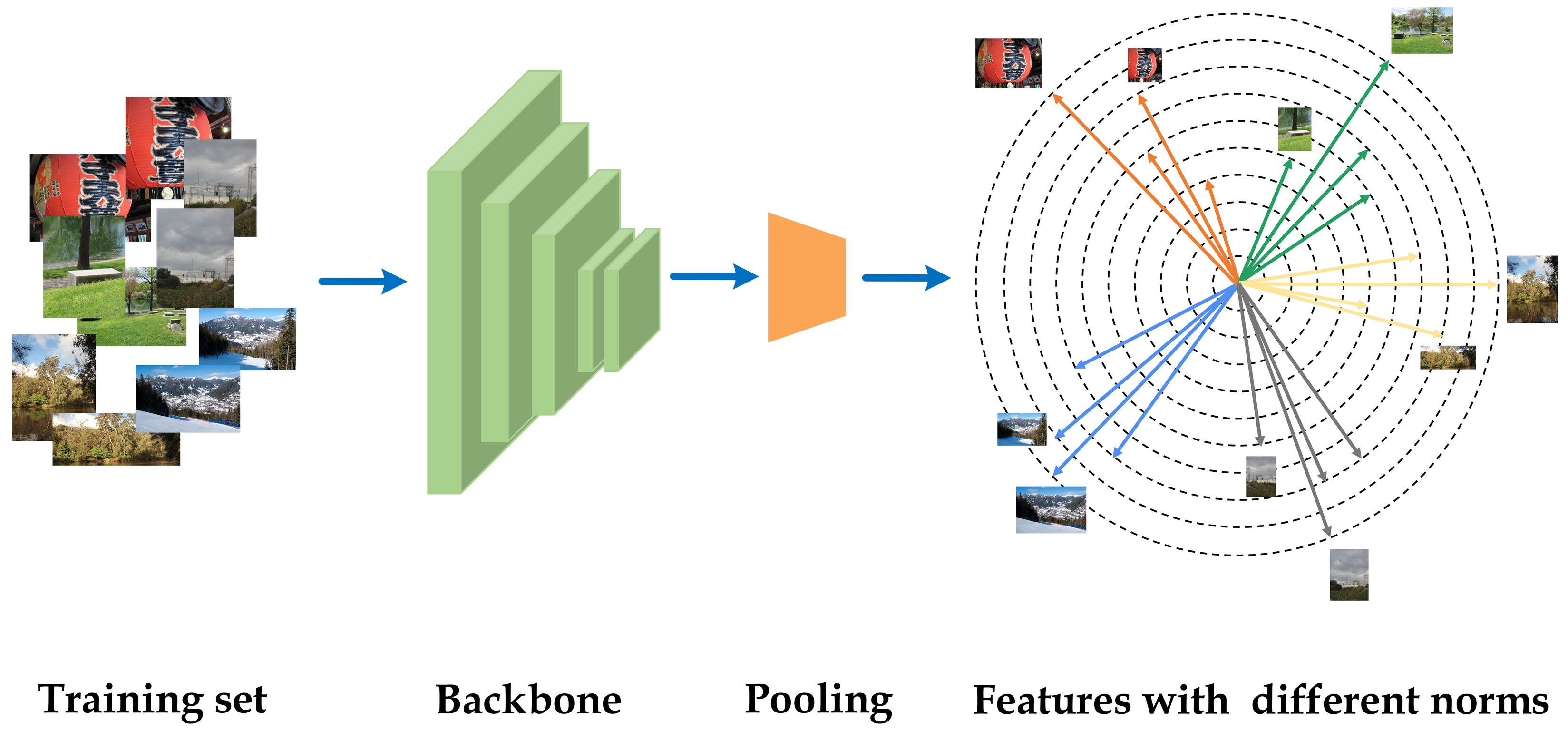
The NDEC Set

- The hard negative pairs in the training set of the NDEC dataset. In each pair, we emphasize the right-side image is not an edited copy (cropped region) of the left-side image, though the left-side image may be an edited copy of the right-side image.



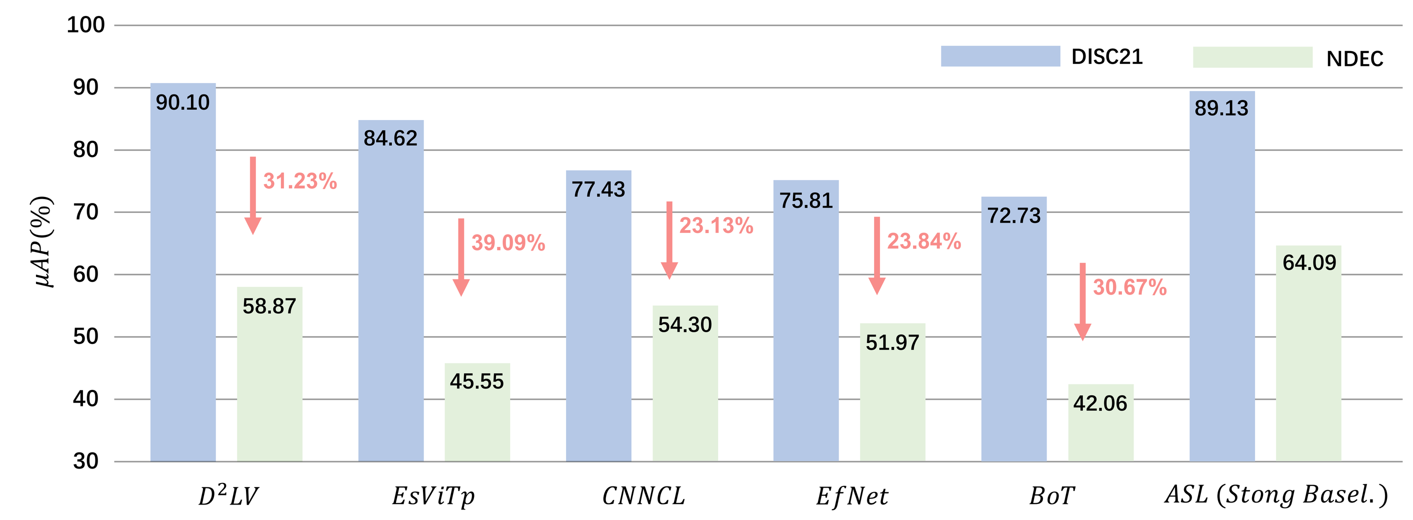
Asymmetrical-Similarity Learning (ASL)

- The illustration of Asymmetrical-Similarity Learning (ASL). In ASL, the norm ratio based loss makes images with more content/information have a larger norm.



Experiments

- Evaluation on DISC21 and our NDEC.



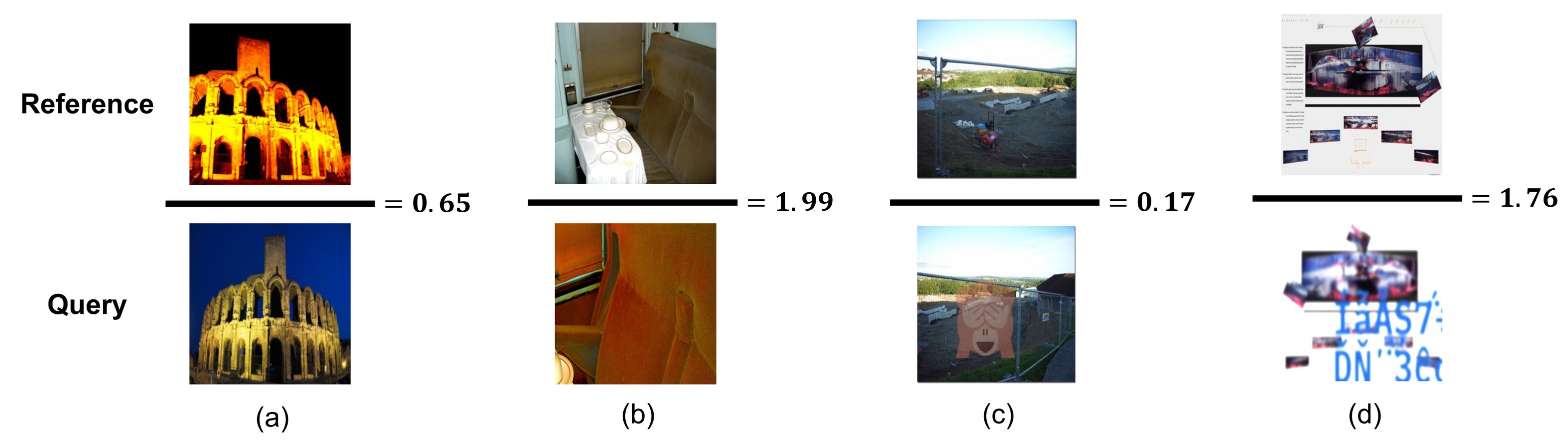
- ASL brings general improvement over various baselines.

Method	$\mu AP\uparrow$	True Positive ↑	False Positive ↓	Precision ↑
EsViTp	45.55%	2,346	2,663	46.84%
EsViTp+ASL	48.31%	1,265	501	71.63%
CNNCL	54.30%	2,620	2,389	52.31%
CNNCL+ASL	56.95%	2,110	545	79.47%
EfNet	51.97%	2,579	2,430	51.49%
EfNet+ASL	53.81%	1,997	562	78.04%
BoT	42.06%	2,075	2,934	41.43%
BoT+ASL	45.07%	1,620	838	65.91%
D^2LV	58.87%	2,836	2,173	56.62%
D^2LV+ASL	61.28%	2,227	583	79.25%
Simple Basel.	47.00%	2,269	2,740	45.30%
Simple + ASL	49.30%	1,829	648	73.84%
Strong Basel.	61.03%	2,968	2,041	59.25%
Strong + ASL	64.09%	2,331	567	80.43%

- The ablation studies based on our simple baseline.

Method	$\mu AP\uparrow$	True Positive ↑	False Positive ↓	Precision ↑
Simple Basel.	47.00%	2,269	2,740	45.30%
ASL-Crop	49.10%	2,098	1,233	62.98%
ASL-Negative	48.14%	1,877	845	68.96%
ASL-Positive	48.17%	1,932	794	70.87%
Triplet	45.37%	1,774	1,200	59.65%
ASL	**49.30%**	1,829	648	**73.84%**

- ASL makes reasonable predictions of the norm ratio under various scenarios.



Code & Data

https://github.com/WangWenhao0716/ASL

If you have any question, please contact:
wangwenhao0716@gmail.com