Ranked List Loss for Deep Metric Learning

Xinshao Wang, Yang Hua, Elyor Kodirov, Guosheng Hu, Romain Garnier, Neil M. Robertson

QUEEN'S UNIVERSITY BELFAST
anyvision A BETTER TOMORROW
CVPR LONG BEACH CALIFORNIA June 16-20, 2019

What is deep metric learning (DML)?

- DML learns a deep embedding space where the relative locations of data points are based on their semantic labels to achieve intraclass compactness and interclass separability.

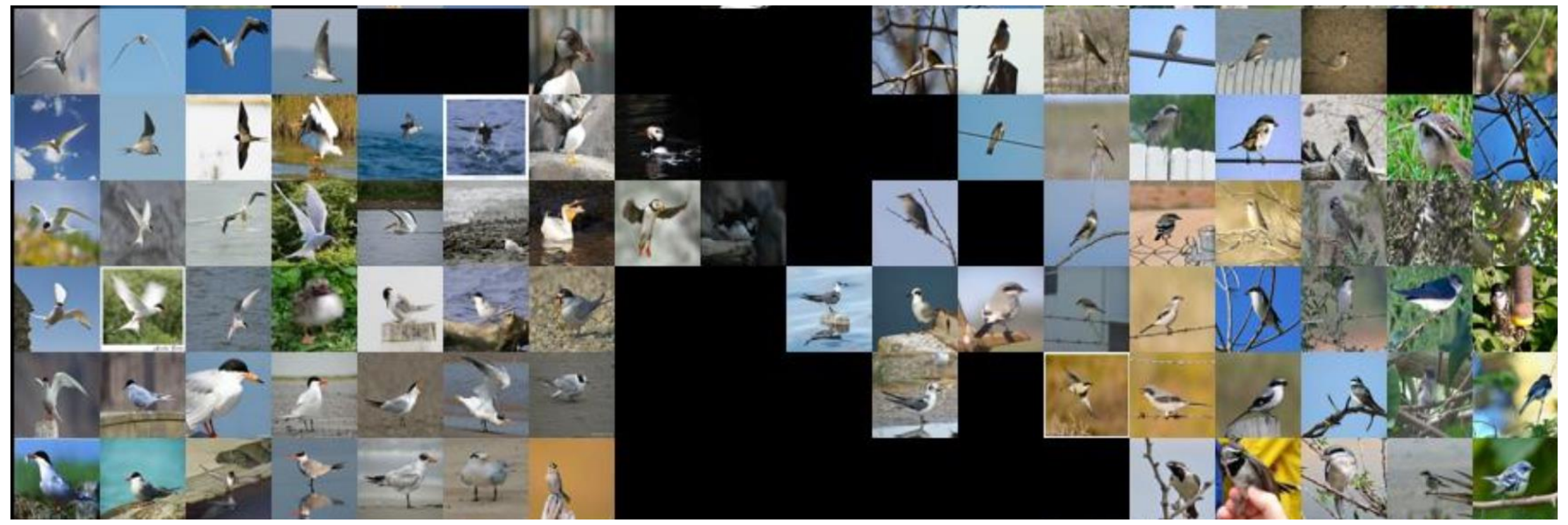

Figure 1: An excerpt of t-SNE plot on CUB-200-2011 test set.

What are the limitations of methods in the literature?

- Exploiting only a proportion of informative examples. We find that utilising all of them jointly is better.
- Tending to shrink class-level distribution into one point, i.e., kill all intraclass variances in the metric space. We argue that regularising the metric space by pulling all intraclass examples into a hypersphere is better.
- Imbalanced number of positive examples and negative ones.

How do we address the key limitations?

- Leveraging all non-trivial training examples.
- Hypersphere regulator: Non-excessively reduce intraclass variances.
- Treating the positive and negative sets of a query equally.

Illustration of our ranked list loss (RLL)

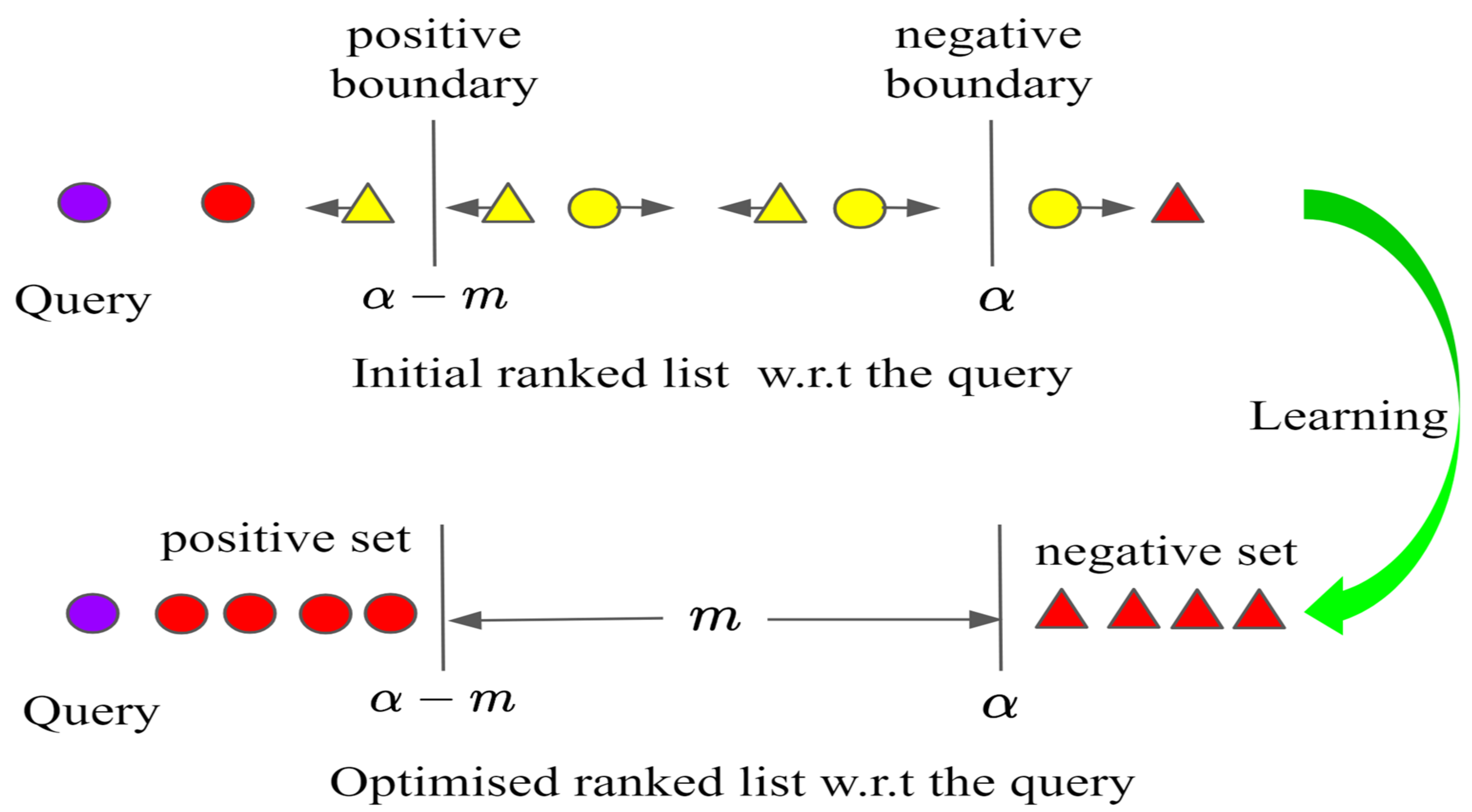

Figure 2: The optimisation objective of RLL. Different shapes represent different classes. The red and yellow denote trivial and informative examples, respectively. Arrows indicate pull/push effects to the query by non-trivial examples in the gallery. During training, the query's representation is updated based on a weighted combination of their effects. *We only calculate the gradient w.r.t. the query per loss.* In other words, we iteratively update a query's embedding while fixing the gallery.

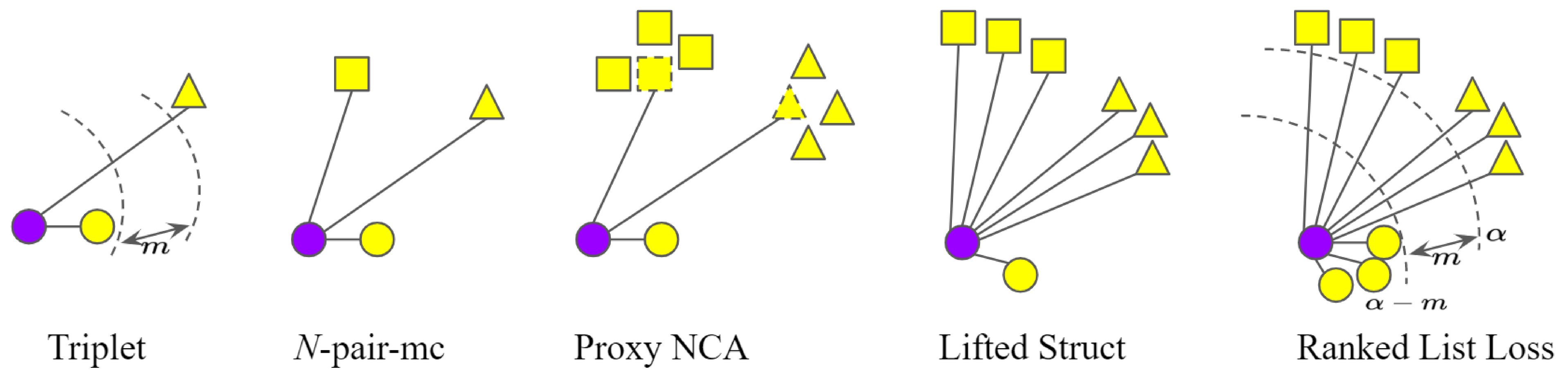

Figure 3: Comparing RLL with other ranking-motivated structured losses. The blue circle is an anchor (query). The intraclass hypersphere diameter: $\alpha - m$

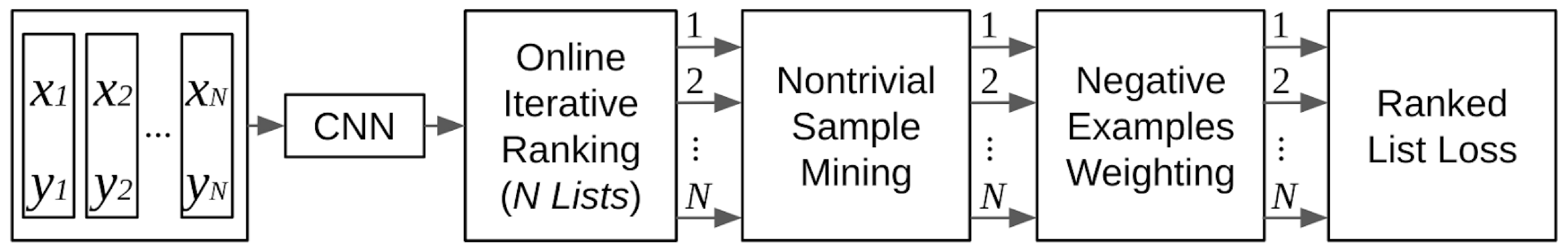

Figure 4: Pipeline of RLL. Computational complexity: O(N^2)

SOTA results and ablation study on SOP

	SOP			
	R@1	R@10	R@100	NMI
Triplet Semihard [25]	66.7	82.4	91.9	89.5
Lifted Struct [21]	62.5	80.8	91.9	88.7
N-pair-mc [29]	66.4	83.2	93.0	89.4
Struct Clust [30]	67.0	83.7	93.2	89.5
Spectral Clust [17]	67.6	83.7	93.3	89.4
Proxy NCA [20]	73.7	–	–	**90.6**
RLL-H	**76.1**	**89.1**	**95.4**	89.7
RLL-(L,M,H)	**79.8**	**91.3**	**96.3**	90.4

Embedding	R@1	R@10	R@100
L	76.1	88.8	94.9
M	76.9	89.6	95.5
H	76.1	89.1	95.4
(L,M,H)	**79.8**	**91.3**	**96.3**

Batch size	R@1	R@10	R@100
$120 = 40 \times 3$	79.2	90.9	96.2
$150 = 50 \times 3$	79.5	91.1	96.2
$165 = 55 \times 3$	79.7	91.2	**96.3**
$180 = 60 \times 3$	**79.8**	**91.3**	**96.3**
$195 = 65 \times 3$	**79.8**	**91.3**	**96.3**

General hyper-factors study

$m = 0.4, T = 10$	R@1	R@10	R@100
$\alpha = 1.4$	76.2	89.4	95.6
$\alpha = 1.2$	**79.8**	**91.3**	**96.3**
$\alpha = 1.0$	78.7	90.5	95.9

RLL's hyper-parameters study

$\alpha = 1.2, T = 10$	R@1	R@10	R@100
$m = 0$	76.1	89.8	95.7
$m = 0.2$	79.0	91.2	**96.3**
$m = 0.4$	**79.8**	**91.3**	**96.3**
$m = 0.6$	79.2	90.6	96.0
$m = 1.2$	79.1	90.5	95.8

$m = 0.4, \alpha = 1.2$	R@1	R@10	R@100
$T = 0$	78.8	90.7	96.1
$T = 5$	79.1	91.0	96.2
$T = 10$	**79.8**	**91.3**	**96.3**
$T = 15$	79.3	90.9	96.0
$T = 20$	78.6	90.5	95.7

Summary

- RLL is simple, effective, intuitive (easy to interpret).
- Extensive ablation studies are presented and valuable for practice.