SoDeep: a Sorting Deep net to learn ranking loss surrogates

Martin Engilberge[1,2], Louis Chevallier[2], Patrick Pérez[3], Matthieu Cord[1,3]

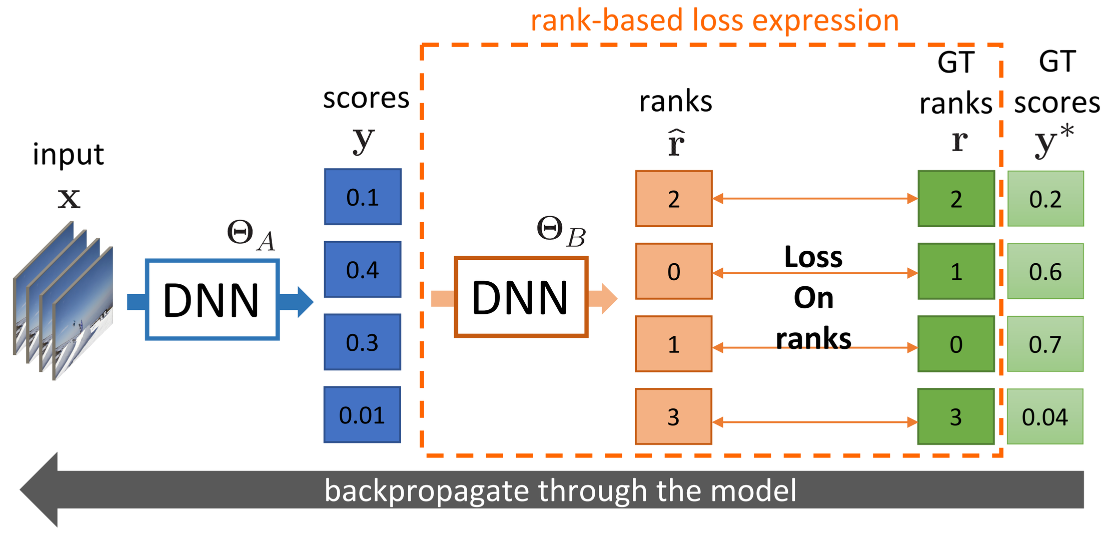
[1]Sorbonne Université, Paris, France, [2]Technicolor, Cesson Sévigné, France, [3]Valeo.ai, Paris, France

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

Overview

Non-differentiable ranking metrics:

The projection from continuous score to discrete rank used in ranking metrics makes them non-differentiable



Approach to make ranking metrics differentiable:

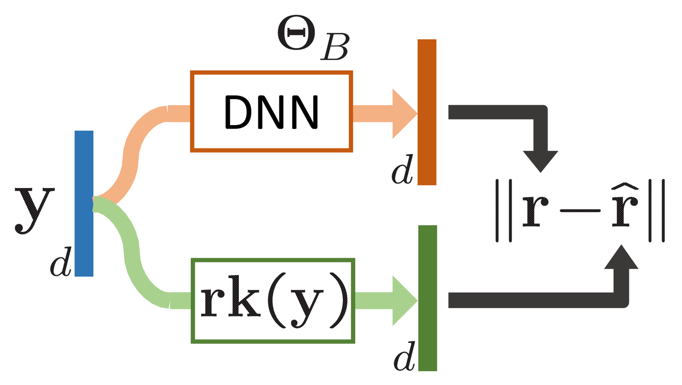
- DNN sorter trained to approximate the rank function **rk**
- Ranking metrics expressed as a function of the sorter
- Ranking metrics used as loss function

Training a differentiable sorter

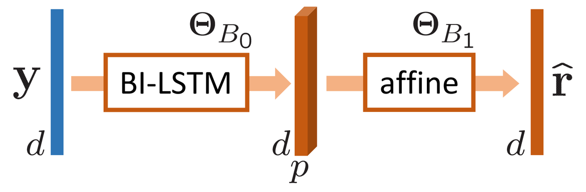
Using only synthetic data:

- Uniform distribution over [-1,1]
- Normal distribution with $\mu = 0$ and $\sigma = 1$
- Evenly spaced numbers in random sub-range of [-1,1]

Training:

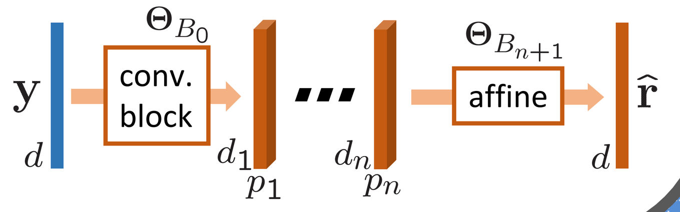


Sorter architecture:



- LSTM based

- Convolution based



Loss functions

Spearman correlation as loss function:

- Spearman correlation $\quad r_s = 1 - \frac{6\|\mathbf{rk}(\mathbf{y}) - \mathbf{rk}(\mathbf{y}')\|_2^2}{d(d^2-1)}$
- Maximizing Spearman correlation $\quad \min_{\Theta_A} \sum_{n=1}^{N} \left\|\mathbf{rk}(\mathbf{y}^{(n)}) - \mathbf{rk}(\mathbf{y}^{*(n)})\right\|_2^2$
- Replacing **rk** with our trained sorter

$$\mathcal{L}_{SPR}(\mathbf{\Theta}_A, \mathcal{B}) = \sum_{n=1}^{N} \left\| f_{\mathbf{\Theta}_B}(\mathbf{y}(\mathbf{\Theta}_A)^{(n)}) - \mathbf{rk}(\mathbf{y}^{*(n)}) \right\|_2^2$$

Mean average precision as loss function:

$$\mathcal{L}_{mAP}(\mathbf{\Theta}_A, \mathcal{B}) = \sum_{c=1}^{C} \langle f_{\mathbf{\Theta}_B}(\mathbf{y}_c), \mathbf{y}_c^* \rangle$$

Recall as loss function:

$$\mathcal{L}_{REC}(\mathbf{\Theta}_A, \mathcal{B}) = \frac{1}{d} \sum_{i \in \mathcal{B}} \max_{c \neq p, c \neq i} \operatorname{loss}(\mathbf{Y}[i], p, c)$$

$$\operatorname{loss}(\mathbf{Y}[i], p, c) = \max\left\{0, \alpha + f_{\mathbf{\Theta}_B}(\mathbf{Y}[i])_p - f_{\mathbf{\Theta}_B}(\mathbf{Y}[i])_c\right\}$$

Handcrafted sorter

Baseline: a ranking algorithm

- A differentiable comparator: the sigmoid function

$$\sigma_{comp}(a, b) = \frac{1}{1 + e^{-\lambda(b-a)}}$$

- Comparing an element with the rest of the sequence yields its rank

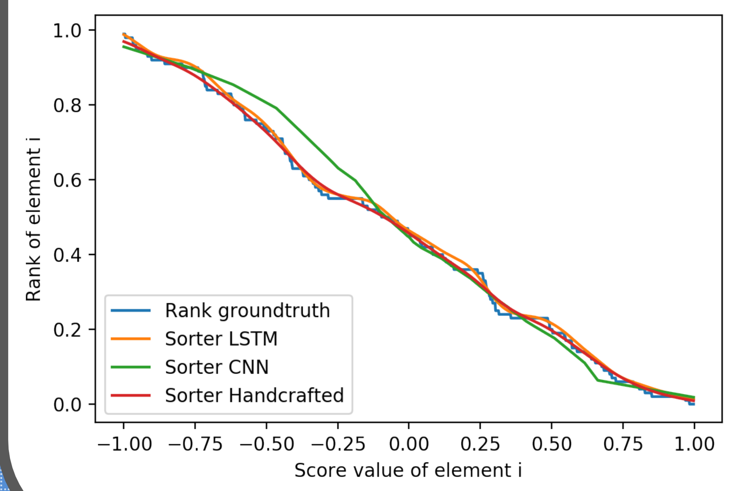
$$f_h(\mathbf{y}, i) = \sum_{j: j \neq i} \sigma_{comp}(\mathbf{y}_i, \mathbf{y}_j)$$

- Value of λ, a trade off between accuracy and gradient flow

Sorter comparison

Sorter behavior analysis:

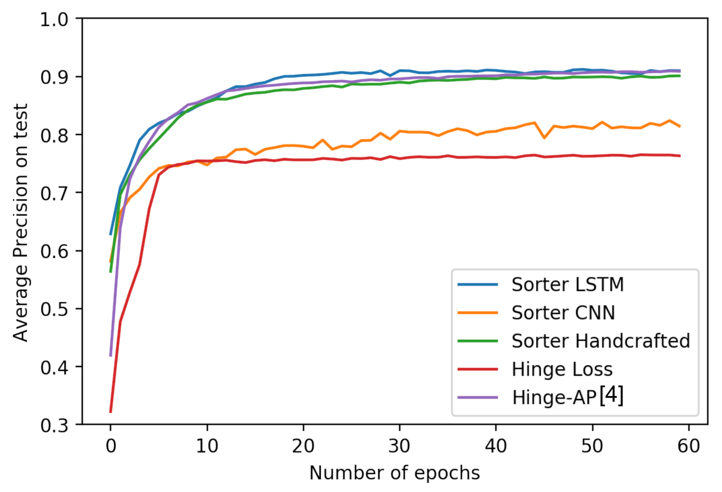
- The learned sorters are able to estimate the groundtruth rank
- RNN sorters perform better than CNN ones



Memorability Loss	Spear. corr. val.
Sorter baseline	45.7
Sorter CNN	46.6
Sorter LSTM	**50.4**

Toy experiments:

- Binary classification on synthetic data
- Competitive with less complexity



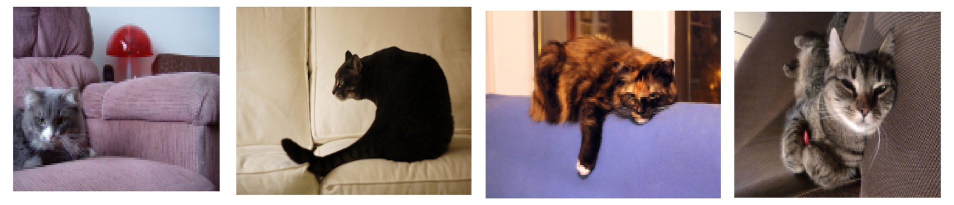
Evaluation

Object recognition: evaluated on the VOC 2007 challenge

Model	mAP
VGG 16	89.3
WILDCAT [1]	95.0
WILDCAT*	93.2
WILDCAT* + SoDeep	94.0

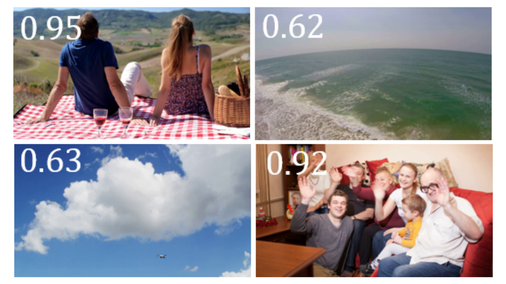
Cross modal retrieval: evaluated on MS-CoCo image/caption pairs

Query: A cat on a sofa



	Caption retrieval				Image retrieval			
Model	R@1	R@5	R@10	Med. r	R@1	R@5	R@10	Med. r
DSVE-Loc[2]	69.8	91.9	96.6	1	55.9	86.9	94.0	1
GXN[3]	68.5	-	**97.9**	1	**56.6**	-	**94.5**	1
SoDeep	**71.5**	**92.8**	97.1	1	56.2	**87.0**	94.3	1

Memorability prediction: memorability score reflects the probability of a video to be remembered



Model	Spear. corr.
Baseline	46.0
Sem-Emb + MSE loss	48.6
Sem-Emb + SoDeep	**49.4**

References

[1] T. Durand et al. Wildcat: Weakly supervised learning of deep convnets for image classification, pointwise localization and segmentation. CVPR, 2017.

[2] M. Engilberge et al. Finding beans in burgers: Deep semantic-visual embedding with localization. CVPR, 2018.

[3] J. Gu et al. Look, imagine and match: Improving textual-visual cross-modal retrieval with generative models. CVPR, 2018.

[4] Y. Yue et al. A support vector method for optimizing average precision. ACM SIGIR, 2007.

Code available on Github

https://github.com/technicolor-research/sodeep