DeepIndex for Accurate and Efficient Image Retrieval

Yu Liu, Yanming Guo, Song Wu, Michael S. Lew

LIACS Media Lab, Leiden University {y.iu, y.guo, s.wu, m.s.lew}@liacs.leidenuniv.nl

Abstract

The success of deep features extracted from ConvNets has shown promising results toward bridging the semantic gap. Inspired by this, we attempt to introduce deep features into the inverted index based image retrieval and propose the DeepIndex framework. We further incorporate multiple deep features from different fully connected layers, resulting in the multiple DeepIndex. Extensive experiments demonstrate DeepIndex method is competitive with the state-of-the-art. In addition, our method is efficient in terms of both memory and time cost.

Contributions

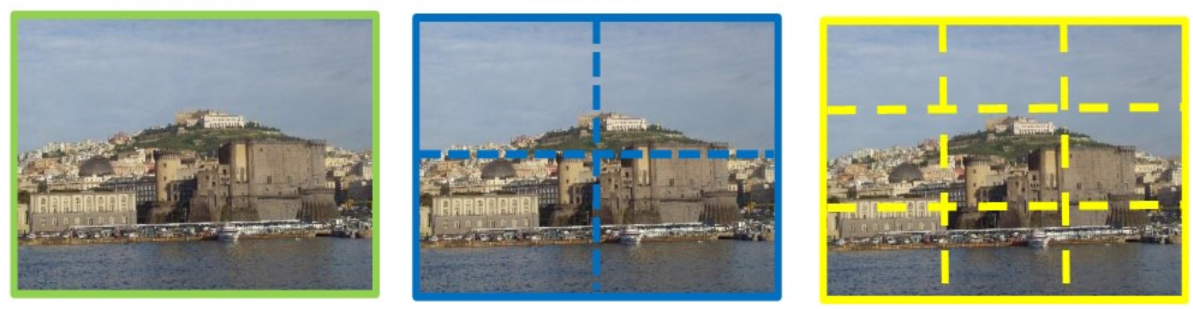
1. We introduce the DeepIndex framework that uses the inverted index scheme for deep features.
2. We propose to use multiple deep features with multiple DeepIndex which improves retrieval accuracy. It consists of two variants: intra-CNN and inter-CNN.
3. We further employ the deep feature of the whole image as a global signature for improving matching accuracy, which is called global image signature (GIS).

Acknowledgments

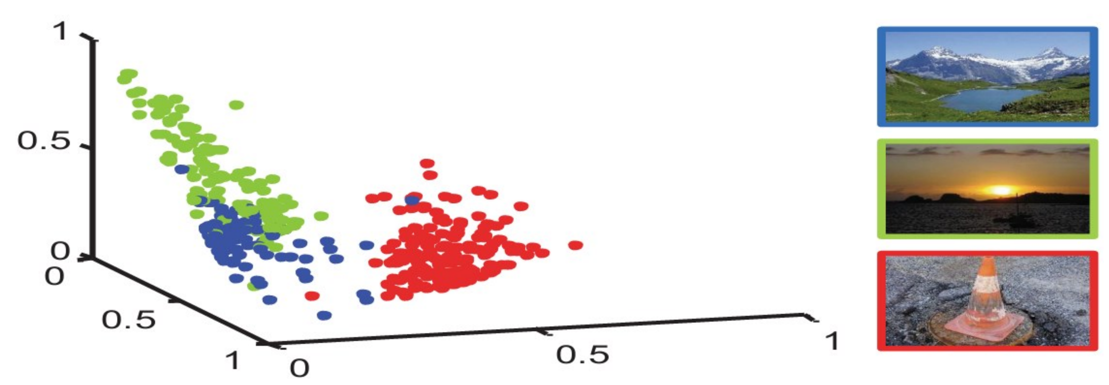
This work was supported by the LIACS Media Lab at Leiden University.

The proposed approach

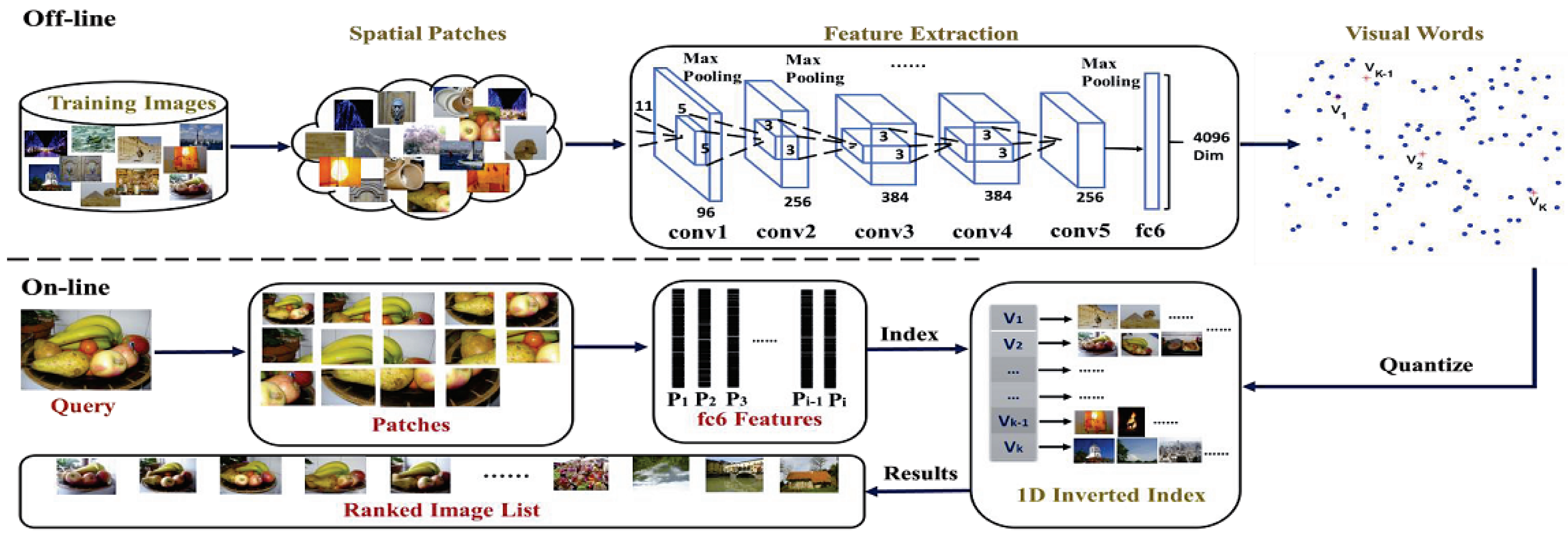
Spatial patches



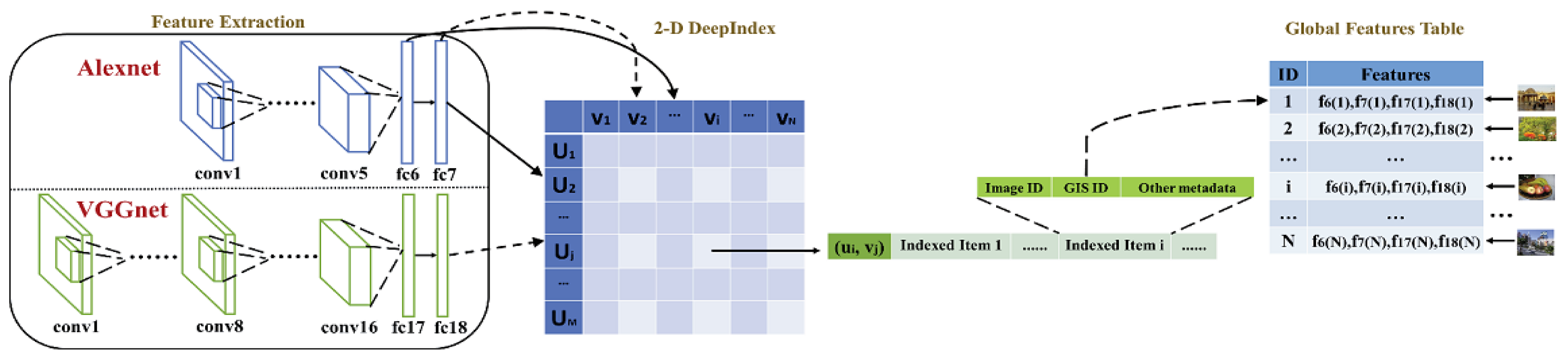
Visualizing patches deep features



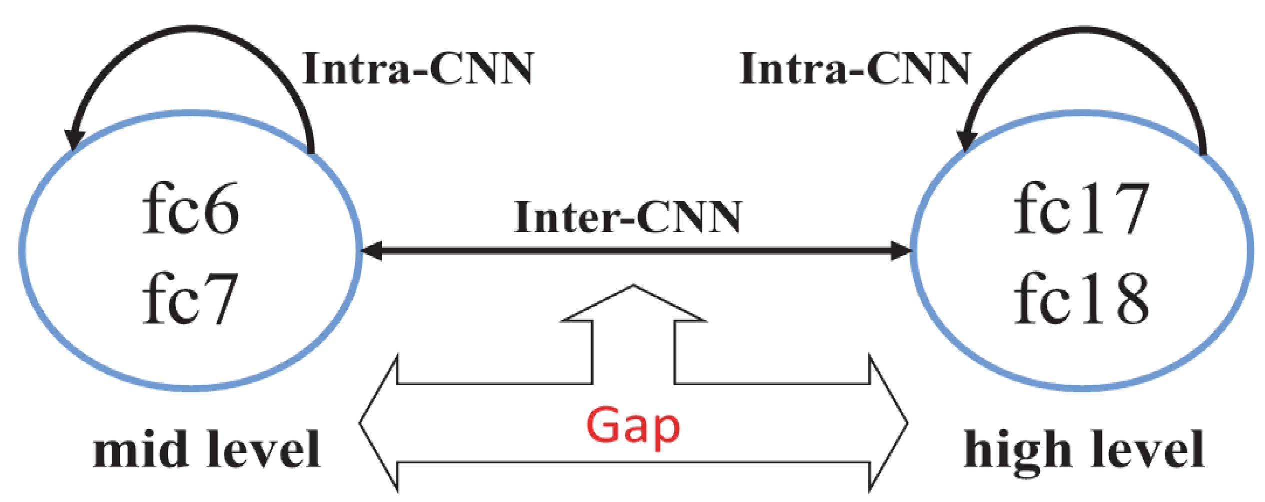
DeepIndex framework



Multiple DeepIndex with GIS

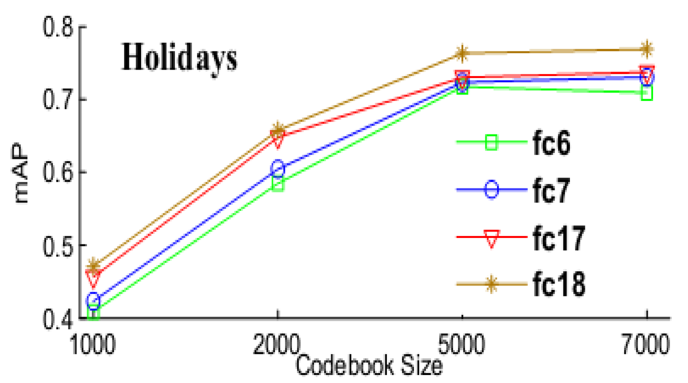


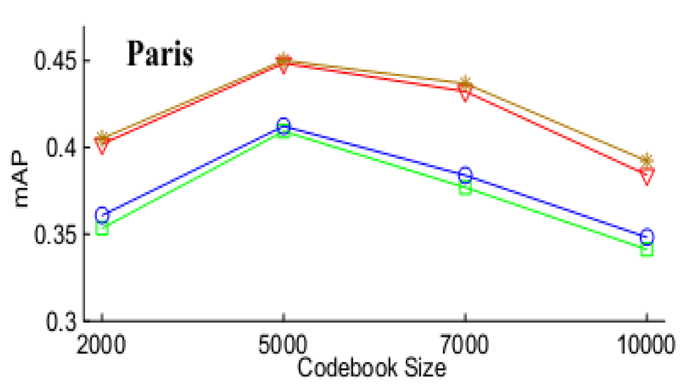
Intra-CNN and Inter-CNN

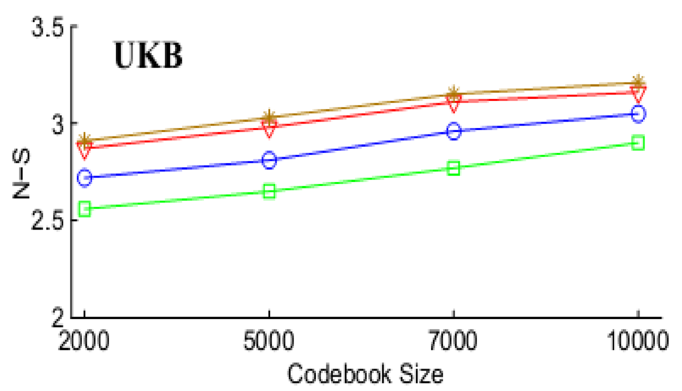


Experimental results and discussions

(1) Codebook size







(2) Overall evaluation

Dataset	Holidays Dataset			Paris Dataset			UKB Dataset		
Methods	MA=1	MA=3	MA=5	MA=1	MA=5	MA=10	MA=1	MA=5	MA=10
DPI_6	0.7173	0.7354	0.7201	0.4094	0.5689	0.6521	2.90	3.03	3.02
DPI_7	0.7234	0.7490	0.7358	0.4124	0.5745	0.6578	3.05	3.12	3.04
DPI_{17}	0.7302	0.7322	0.7262	0.4487	0.6101	0.7024	3.16	3.19	3.15
DPI_{18}	0.7631	0.7672	0.7563	0.4503	0.6123	0.7133	3.21	3.25	3.19
DPI_{6+7}	0.7200	0.7888	0.7717	0.2935	0.6289	0.7120	3.02	3.13	3.05
DPI_{17+18}	0.7575	0.7996	0.7934	0.3228	0.6329	0.7169	3.16	3.25	3.26
DPI_{7+17}	0.7401	0.8053	0.8020	0.3345	0.6412	0.7324	3.21	3.25	3.19
DPI_{6+17}	0.7332	0.8162	0.8115	0.3395	0.6508	0.7435	3.22	3.26	3.22
DPI_{7+18}	0.7466	0.8123	0.8174	0.3656	0.6618	**0.7535**	3.26	**3.37**	3.32
DPI_{6+18}	0.7382	0.8164	**0.8238**	0.3412	0.6540	0.7452	3.19	3.23	3.29

(3) GIS improvement

Method	Holidays	Paris	UKB
Inter-CNN without GIS	82.38	75.35	3.37
Inter-CNN with GIS	83.30	78.24	3.68

(4) PCA reduction

Datasets	Holidays	Paris	UKB
Dim=4096	0.833	0.7824	3.68
Dim=2048	0.8411	0.7945	3.72
Dim=1024	0.8463	0.8065	3.74
Dim=512	**0.8565**	**0.8124**	**3.76**
Dim=256	0.8367	0.7875	3.71
Dim=128	0.8272	0.7724	3.65

Comparison

Method	Group	Holidays	Paris	UKB
ASMK-small (G. Tolias, 2013)	Non-CNN	82.20	78.20	—
c-Multi-Index(L. Zheng, 2014)	Non-CNN	84.02	—	3.71
ASMK-large(G. Tolias, 2013)	Non-CNN	88.00	80.50	—
CNNaug-ss (A. S. Razavian, 2014)	CNN	84.30	79.50	91.1mAP
DF.FC1+SL(J. Wan, 2014)	CNN	—	**86.83**	—
Ours	CNN	85.65	81.24	3.76
Binary(L. Zheng, 2014)	SIFT-CNN	85.30	—	3.79
Float(L. Zheng, 2014)	SIFT-CNN	**88.08**	—	**3.85**

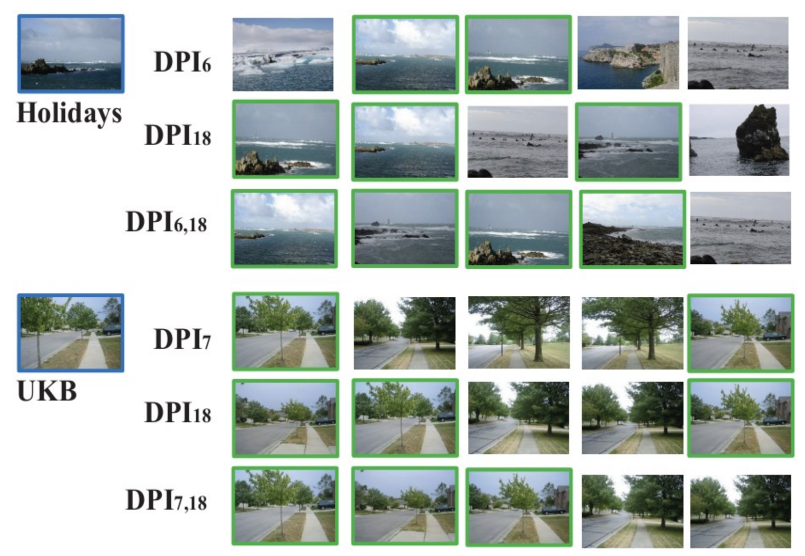
Complexity (bytes and seconds)

Memory (Bytes)	Binary(L. Zheng, 2014)	1-D DPI	2-D DPI
ImageID	4 × 500	4 × 14	4 × 14
Signature	10.18KB	4 × 512	4 × 2 × 512
Total Memory	12.13KB	2.06KB	4.06KB
Query Time	**2.32**	**0.25**	**0.45**

* For a fair comparison, we only report results that exclude post-processing like spatial re-ranking and query expansion. Also, we do not consider fine-tuning.

* Our query time does not include the feature extraction.

Example



Discussions:

1. Multiple assignment (MA) is useful to increase recall.
2. Mostly, 2-D DPI > 1-D DPI.
3. Mostly, Intra-CNN> Inter-CNN.
4. GIS is useful to develop matching strength as a kind of global constraint.
5. PCA is useful to reduce memory complexity.

The source code can be downloaded from: http://press.liacs.nl/lml/deepindex.

References

[1]G. Tolias, Y. Avrithis, and H. Jegou. To aggreagate or not aggregate: selective match kernels for image search. In ICCV, 2013.

[2] L. Zheng, S. Wang, Z. Liu, and Q. Tian. Packing and padding: Coupled multi-index for accurate image retrieval. In CVPR, 2014.

[3] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson. CNN features off-the-shelf: an astounding baseline for recognition. In CVPR, DeepVision workshop, 2014.

[4] J. Wan, D. Wang, S. C. H. Hoi, P. Wu, J. Zhu, Y. Zhang, J. Li. Deep learning for context-based image retrieval: A comprehensive study. In ACM Multimedia, 2014.

[5] L. Zheng, S. Wang, F. He, and Q. Tian. Seeing the big picture: Deep embedding with contextual evidences. CoRR, abs/1406.0132, 2014.

Conclusions

We proposed the DeepIndex framework for accurate and efficient image retrieval that introduced the inverted index into deep features. Moreover, we integrated multiple deep features with the multiple DeepIndex which attempted to bridge the gap between mid-level and high-level representations. Experimental results showed that our method achieved competitive performance with low complexity. In the future, it is promising to investigate the possibilities of using posting-processing or finetuning stage .