Computer Vision

Approximate Nearest Neighbors Algorithms
KGraph

Vladislav Belov

FNSPE CTU

May 3, 2020

Product Quantization

In this report, we cover results of our experiments with the KGraph[1] index [1] for the nearest neighbors search. The report includes the following:

- Showcase of KGraph performance on Oxford105K data (image descriptors with total number of samples N $=$ 104933 of dimension d $=$ 128);
- Comparison of KGraph with FLANN[2][3] and Product Quantization[3][2].

Technical Details

The respective notebook is available on our GitHub page.

[1] KGraph library with default parameters is utilized.

[2] PyFLANN library with pre-set target precision is used.

[3] We utilize Fair AI Similarity Search (faiss) library with fixed parameters of the index: `nlist`$=$ 256, m $=$ 16, `nbits`$=$ 8, and `nprobe`$=$ 32 of IVFPQ index.

NN Search with IVFPQ on CNN-based Image Descriptors

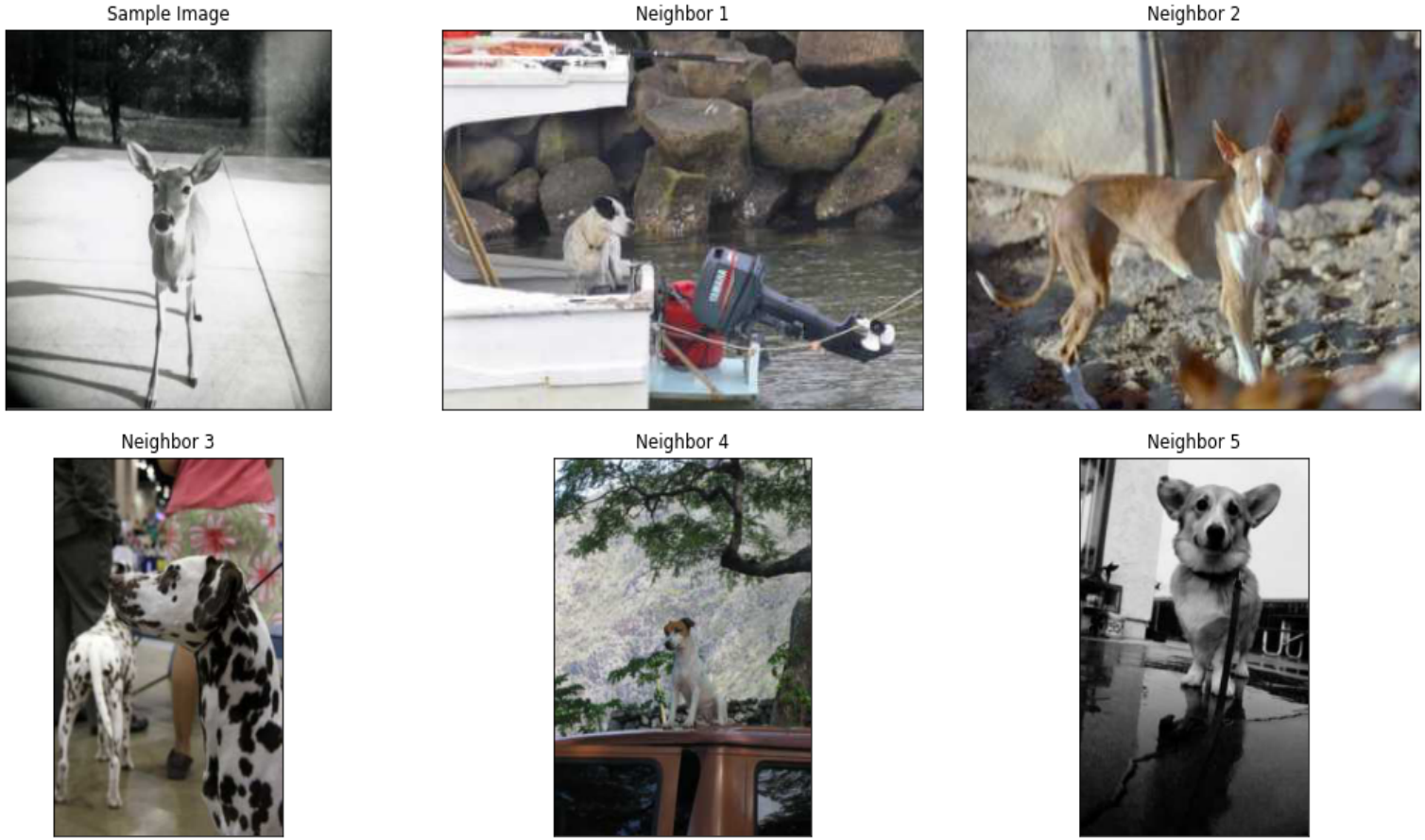

Figure 1: 5 Nearest neighbors for the sample image on CNN-based descriptors.

Framework of the Experiment

- We compare performance of KGraph with the auto-tuned FLANN with target precision set to 0.99 (FLANN-TARGET) and with the IVFPQ index by means of estimation of recall $R(k, K)$.
- Recall is measured as the average fraction over queries (all images from the data set) of first k true nearest neighbors found during the search of K approximate nearest neighbors, $k \leq K$. Results for $K = 100$ can be viewed in Fig. 2.
- For each approach, we also measure the time to build the index and to perform the query. Results are presented in Tab. 1.

Method	Time to build the index, [s]	Time to perform the query, [s]
IVFPQ[4]	5.75 + 0.729	6.11
KGraph	105	-
FLANN-TARGET	0.136	4.35
Exact NN	-	2980

Table 1: Timings of ANN methods to perform 105K queries for 100 nearest neighbors.

[3] Timings for both training on all data points and the point insertion are given.

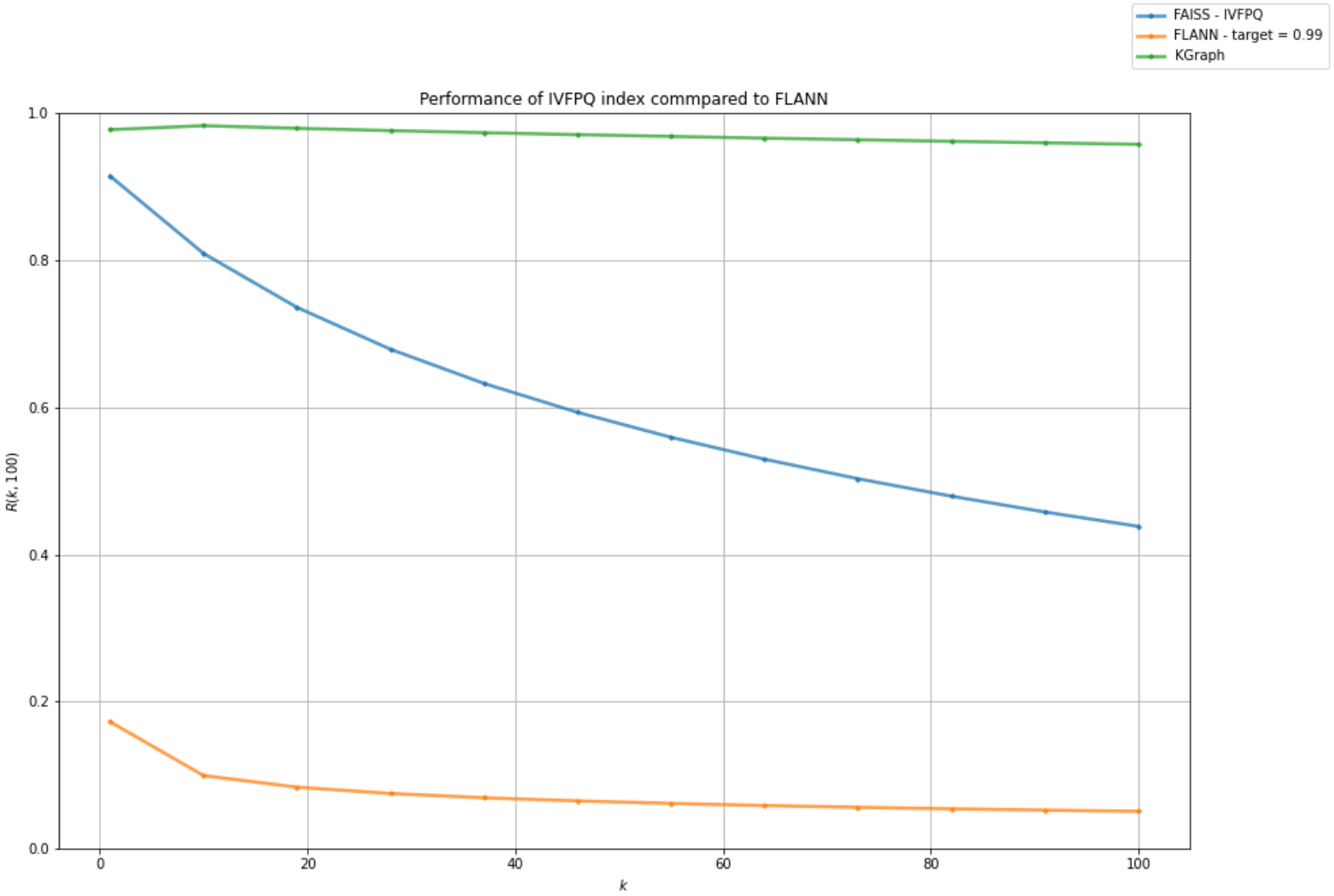

Figure 2: Values of R(k, 100). In terms of recall, KGraph demonstrates great superiority to FLANN and IVFPQ.

Appendix

W. Dong, C. Moses, and K. Li.

Efficient k-nearest neighbor graph construction for generic similarity measures.
In Proceedings of the 20th International Conference on World Wide Web, WWW '11, page 577–586, New York, NY, USA, 2011. Association for Computing Machinery.

H. Jégou, M. Douze, and C. Schmid.

Product quantization for nearest neighbor search.
IEEE Trans. Pattern Anal. Mach. Intell., 33(1):117–128, 2011.

M. Muja and D. Lowe.

Fast approximate nearest neighbors with automatic algorithm configuration.
VISAPP 2009 - Proceedings of the 4th International Conference on Computer Vision Theory and Applications, 1:331–340, Jan 2009.