

TECHNISCHE UNIVERSITÄT DRESDEN

ZENTRUM FÜR INFORMATIONSDIENSTE  
UND HOCHLEISTUNGSRECHNEN  
PROF. DR. WOLFGANG E. NAGEL

## Master-Arbeit

zur Erlangung des akademischen Grades  
Master of Science

## Image Retrieval für Historische Bilder

Philipp Langen  
(Geboren am 26. Dezember 1994 in Münsterlingen)

Hochschullehrer: Prof. Dr. Wolfgang E. Nagel  
Betreuer: Dr. Christoph Lehmann & Dr. Taras Lazariv

Dresden, 1. Juni 2020

---

**Hier Aufgabenstellung einfügen!**

---

# Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die von mir am heutigen Tag dem Prüfungsausschuss der Fakultät Informatik eingereichte Master-Arbeit zum Thema:

*Image Retrieval für Historische Bilder*

vollkommen selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Dresden, den 1. Juni 2020

Philipp Langen

---

**Kurzfassung**

**Abstract**

# Inhaltsverzeichnis

|                             |          |
|-----------------------------|----------|
| 0.1 zitate . . . . .        | 2        |
| <b>Literaturverzeichnis</b> | <b>3</b> |

## Motivation

Mit der zunehmenden Digitalisierung unserer Datenwelt stehen jedem von uns heute mit wenigen Klicks mehr Informationen zur Verfügung, als wir je analog aufnehmen könnten. Auch wenn diese Flut an Informationen ein unglaubliches Potential für uns darstellt, so bürgt sie auch neue Herausforderungen. Um große Datenmengen sinnvoll nutzbar zu machen sind effiziente Suchwerkzeuge von zentraler Bedeutung. Dies gilt insbesondere auch für die Suche in großen Bilderdatenbanken. Bei der klassischen Bildersuche wird vom Nutzer eine Anfrage als String formuliert, auf welche das System eine Liste an Bildern mit groß möglichstem Bezug zur Anfrage liefert. Dabei werden häufig nützliche Zusatzinformationen, sogenannte Metadaten der Bilder, wie Titel, Beschreibung, Aufnahmeort und Datum genutzt, um bessere Ergebnisse liefern zu können. Eine alternativer Ansatz der Suche ist die inhaltsbasierte Bildersuche (engl. Content-Based Image Retrieval kurz CBIR). Hierbei werden statt formulierten Anfragen Bilder als Anfragen gestellt. Ziel ist es Bilder mit gleichem oder ähnlichem Bildinhalt wie in der Anfrage als Ergebnis zurückzugeben. Dabei arbeitet das System direkt mit den Pixelinformationen der Bilder. Zusätzliche Metadaten sind daher nicht erforderlich. Im Folgenden wird die inhaltsbasierte Suche auch als Image Retrieval bezeichnet.

Ein interessanter Anwendungsbereich für Image Retrieval Systeme ist die Suche auf historischen Bildern. Diese weit gefasste Domäne enthält sehr heterogene Daten. Dabei finden sich nicht nur sehr unterschiedliche Bildmotive, wie Gebäuden, Naturaufnahmen, Portraits oder Gegenständen, sondern auch unterschiedliche Aufnahmetechniken bedingt durch den technischen Fortschritt über Zeichnungen, Malerei und Druck bis hin zur Photographie. Da Metadaten zu historischen Bildern erst bei der Digitalisierung hinzugefügt werden können sind diese oft gar nicht oder nur lückenhaft vorhanden, was die inhaltsbasierte Suche zu einem besonders geeigneten Ansatz macht. Mit der Umsetzung einer unterstützenden Suche für das UrbanHistory4D Projekt [1] ergibt sich ein konkreter Anwendungsfall für Image Retrieval Systeme. Hierbei handelt es sich um einen aktiven Forschungsbereich, in dem momentan unterschiedliche Suchsysteme analysiert werden.

Bei dem Image Retrieval Verfahren DELF (attentive DEep Local Features) [2], welches in dieser Arbeit untersucht wird handelt es sich um einen Deep Learning Ansatz. Durch den raschen Fortschritt im Bereich tiefer neuronaler Netzwerkarchitekturen der letzten Jahre erfreuen sich gelernte Ansätze immer größerer Beliebtheit. DELF erzielt auf bekannten Benchmarkdatensätzen wie Oxford5k [3] und Paris6k [4] sehr gute Ergebnisse. Besonders gut schneidet DELF im Vergleich auf dem eigens erstellten Google Landmarks Datensatz [5] ab. Dieser enthält mit über 1 mio. Bilder und 13k unterschiedlichen Motiven eine deutlich heterogener Mischung an Objekten. Die gute Performanz auf diesem Datensatz lässt als hoffen, dass sich das Verfahren auch für die historische Domäne eignet.

## 0.1 zitate

[6][7][8][9][10][11][12][13][14][15][16][17][18][19][20][21][22][23][24][25][26][2][27]

## Literaturverzeichnis

- [1] Ferdinand Maiwald, Jonas Bruschke, Christoph Lehmann, and Florian Niebling. A 4d information system for the exploration of multitemporal images and maps using photogrammetry, web technologies and vr/ar. *Virtual Archaeology Review*, 10:1, 07 2019.
- [2] Hyeonwoo Noh, Andre Araujo, Jack Sim, Tobias Weyand, and Bohyung Han. Large-scale image retrieval with attentive deep local features. pages 3476–3485, 10 2017.
- [3] James Philbin, Ondrej Chum, Michael Isard, Josef Sivic, and Andrew Zisserman. Object retrieval with large vocabularies and fast spatial matching. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.
- [4] James Philbin, Ondrej Chum, Michael Isard, Josef Sivic, and Andrew Zisserman. Lost in quantization: Improving particular object retrieval in large scale image databases. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [5] Yan-Tao Zheng, Ming Zhao, Yang Song, Hartwig Adam, Ulrich Buddemeier, Alessandro Bissacco, Fernando Brucher, Tat-Seng Chua, Hartmut Neven, and Jay Yagnik. Tour the world: A technical demonstration of a web-scale landmark recognition engine. In *Proceedings of the 17th ACM International Conference on Multimedia*, MM '09, page 961–962, New York, NY, USA, 2009. Association for Computing Machinery.
- [6] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(12):1349–1380, December 2000.
- [7] David Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–, 11 2004.
- [8] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. volume 3951, pages 404–417, 07 2006.
- [9] Yan Ke and Rahul Sukthankar. Pca-sift: A more distinctive representation for local image descriptors. volume 2, pages II–506, 05 2004.
- [10] Miaojing Shi, Yannis Avrithis, and Hervé Jégou. Early burst detection for memory-efficient image retrieval. 06 2015.
- [11] Yin Zhang, Rong Jin, and Zhi-Hua Zhou. Understanding bag-of-words model: A statistical framework. *International Journal of Machine Learning and Cybernetics*, 1:43–52, 12 2010.
- [12] Hervé Jégou, Matthijs Douze, Cordelia Schmid, and Patrick Perez. Aggregating local descriptors into a compact image representation. pages 3304 – 3311, 07 2010.

- [13] Jerome Friedman, Jon Bentley, and Raphael Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Softw.*, 3:209–226, 09 1977.
- [14] Hervé Jégou, Matthijs Douze, and Cordelia Schmid. Product quantization for nearest neighbor search. *IEEE transactions on pattern analysis and machine intelligence*, 33:117–28, 01 2011.
- [15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. 7, 12 2015.
- [16] Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 25, 01 2012.
- [17] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv 1409.1556*, 09 2014.
- [18] Matthew Zeiler and Rob Fergus. Visualizing and understanding convolutional neural networks. volume 8689, 11 2013.
- [19] Ali Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. Cnn features off-the-shelf: an astounding baseline for recognition. *Arxiv*, 03 2014.
- [20] Artem Babenko, Anton Slesarev, Alexandr Chigorin, and Victor Lempitsky. Neural codes for image retrieval. volume 8689, 04 2014.
- [21] Joe Ng, Fan Yang, and Larry Davis. Exploiting local features from deep networks for image retrieval. pages 53–61, 06 2015.
- [22] Eva Mohedano, Kevin McGuinness, Noel O’Connor, Amaia Salvador, Ferran Marques, and Xavier Giró-i Nieto. Bags of local convolutional features for scalable instance search. pages 327–331, 04 2016.
- [23] Albert Gordo, Jon Almazan, Jerome Revaud, and Diane Larlus. Deep image retrieval: Learning global representations for image search. volume 9910, pages 241–257, 10 2016.
- [24] Filip Radenović, Giorgos Tolias, and Ondřej Chum. Cnn image retrieval learns from bow: Unsupervised fine-tuning with hard examples. volume 9905, pages 3–20, 10 2016.
- [25] Ali S. Razavian, Josephine Sullivan, Stefan Carlsson, and Atsuto Maki. [paper] visual instance retrieval with deep convolutional networks. *ITE Transactions on Media Technology and Applications*, 4(3):251–258, 2016.
- [26] Lingxi Xie, Richang Hong, Bo Zhang, and Qi Tian. Image classification and retrieval are one. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval, ICMR ’15*, page 3–10, New York, NY, USA, 2015. Association for Computing Machinery.
- [27] Amaia Salvador, Xavier Giró-i Nieto, Ferran Marques, and Shin’ichi Satoh. Faster r-cnn features for instance search. *Arxiv*, 04 2016.