Carl Doersch

Research Scientist
Google DeepMind
London, UK

web: http://www.carldoersch.com/

Research Interests

- **Understanding the Visual World:** Learning representations, recognizing objects, and modeling geometry.
- **Self-Supervised Learning:** Developing tasks that drive visual learning without requiring manual annotation.
- **Visual Data Mining & Visualization:** Discovering the objects, object parts, and other patterns from visual data without human intervention.

Education

- **Ph.D. in Machine Learning, Carnegie Mellon University (2016)**
  Advisor: Prof. Alexei A. Efros & Prof. Abhinav Gupta
  Thesis: Supervision Beyond Manual Annotations for Learning Visual Representations
- **M.S. in Machine Learning, Carnegie Mellon University (2013)**
  4.0 GPA
  Advisor: Prof. Alexei A. Efros
  Thesis: What Makes Paris Look Like Paris?
- **B.S. in Computer Science, Carnegie Mellon University (2010)**
  Double Major in Computer Science and Cognitive Science with minor in Computational Neuroscience
  Dean's List all eight semesters, 4.0 GPA
  Elected to Phi Beta Kappa as a Junior
  Advisor: Prof. Tai Sing Lee
  Senior Thesis: Temporal Continuity Learning for Convolutional Deep Belief Nets

Research Experience

2016 - Ongoing | Google Deepmind, London UK, with Prof. Andrew Zisserman

- Computer vision via self-supervised learning.

2010 - 2016 | Carnegie Mellon University, Pittsburgh PA, with Prof. Alexei A. Efros and Prof. Abhinav Gupta

- Unsupervised object discovery and representation learning by context prediction. [3],[5]
- Visual Datamining via discriminative patches, including mining Google Street View for distinctive elements of city architecture. [6], [7]

2014 & 2015 | Visiting Scholar at UC Berkeley, Berkeley CA, with Prof. Alexei A. Efros

- Spring and Summer in both 2014 and 2015.

2013 | Summer Internship, Google, Venice CA, with Dr. Florian Schroff and Dr. Taehee Lee

- Implemented the discriminative patch discovery algorithm of [6] to run on Google-scale data.

2009 | Summer REU, University of Massachusetts, Amherst, MA, with Prof. Erik Learned-Miller

- Implemented a convolutional deep belief network, an algorithm for unsupervised learning of image features, with the goal of applying these networks to face recognition.
- Helped develop an OCR program that can learn document fonts on the fly. [9]

2008-2010 | Independent Research, Carnegie Mellon University, Pittsburgh, PA

- Designed and programmed a deep belief network for unsupervised visual representation learning based on temporal coherence in videos.

Honors and Awards

- **Google Fellowship**: 2 years, $75,000/year, approx 30 fellows internationally each year across all Computer Science, 2014-2016
- **NDSEG Fellowship**: 3 years, $70,000/year, funding by US Department of Defense, approx 200 fellows nationally each year across all STEM disciplines, 2011-2014
- Honorable mention for the NSF Graduate Research Fellowship Program, 2010 & 2011
- CMU/Pitt Undergraduate Fellowship in Computational Neuroscience, 2009
- NSF Summer Research Experience for Undergraduates Fellowship, 2009
- Elected to Phi Beta Kappa as Junior, 2009
- Elected to Psi-Chi National Honors Society in Psychology, 2008

Refereed Publications

[1] Jacob Walker, Carl Doersch, Abhinav Gupta, and Martial Hebert: An Uncertain Future: Forecasting from Static Images using Variational Autoencoders. ECCV, 2016

[2] Philipp Krähenbühl, Carl Doersch, Jeff Donahue, and Trevor Darrell: Data-dependent Initializations of Convolutional Neural Networks. ICLR, 2016

[3] Carl Doersch, Abhinav Gupta, Alexei Efros: Unsupervised visual representation learning by context prediction. ICCV, 2015 **(oral)**

[4] Carl Doersch, Saurabh Singh, Josef Sivic, Abhinav Gupta, Alexei Efros: What makes Paris look like Paris? Communications of the ACM, 2015

[5] Carl Doersch, Abhinav Gupta, Alexei Efros: Context as supervisory signal: discovering objects with predictable context. ECCV, 2014

[6] Carl Doersch, Abhinav Gupta, Alexei Efros: Mid-level visual element discovery as discriminative mode seeking. NIPS, 2013

[7] Carl Doersch, Saurabh Singh, Josef Sivic, Abhinav Gupta, Alexei Efros: What makes Paris look like Paris? SIGGRAPH, 2012 **(oral)**

[8] Gary B. Huang, Andrew Kae, Carl Doersch, and Erik Learned-Miller: Bounding the probability of error for high precision optical character recognition. JMLR, 2012

[9] Andrew Kae, Gary B. Huang, Carl Doersch, and Erik Learned-Miller: Improving state-of-the-art OCR through high-precision document-specific modeling. CVPR, 2010

Tutorials

[10] Carl Doersch, Tutorial on Variational Autoencoders. arXiv:1606.05908, 2016

Teaching Experience

Fall 2013	16-824 Learning Based Methods in Vision Teaching assistant for Prof. Abhinav Gupta Responsibilities included lecturing, designing and grading assignments.
Spring 2011	10-701 Machine Learning Teaching assistant for Prof. Tom Mitchell 10-701 is intended for first and second year PhD students, but I TA'd as a first-year. Responsibilities included teaching recitations, writing and grading homeworks and exams.

Industry & Other Experience

Summer 2013	Summer Intern, Google, Venice, CA Re-implemented my own research on visual element discovery on Google's distributed architecture, for the purpose of identifying and visualizing the most discriminative objects in user photos.
Summer 2010	Summer Intern, Basis Technologies, Cambridge, MA Developed a system for detecting events from a stream of twitter tweets. The system included a novel distance metric for documents which captures both word salience and word order.
Summer 2008	Summer Intern, Oracle, Nashua, NH Used Java/Struts/Javascript/AJAX/Toplink skills to add features to an Oracle Review, which enables the collaborative review of documents.
Summer 2007	Information Systems Intern, Hamilton Sundstrand (UTC), Windsor Locks, CT
2010-2011 2012-2013	President, Student Pugwash, CMU Student Pugwash promotes discussion of ethical issues in the application of science and technology. It was nearly defunct as of fall 2010, but under my leadership has grown to approx. 20 members regularly attending.

2003 | Eagle Scout, Avon, CT

Advanced Courses

- Statistical Machine Learning
- Geometry-based Vision
- Graduate Statistics
- Databases and Datamining
- Learning-based Vision
- Computer Vision
- Artificial Intelligence
- Applied Regression
- Real Analysis I-II
- Parallel Computer Architecture
- Optimization
- Algorithms

Popular Press

August 2012	What Makes Paris Look Like Paris? Featured in The Atlantic, The Wall Street Journal, Popular Science, Wired, TechCrunch, CNet, Slate, and many more.