
# How do Recent Machine Learning Advances Impact the Data Visualization Research Agenda?

Timo Ropinski, Ulm University
Daniel Archambault, Swansea University
Min Chen, Oxford University
Ross Maciejewski, Arizona State University
Klaus Mueller, Stony Brook University
Alexandru Telea, University of Groningen
Martin Wattenberg, Google


Source: World Economic Forum (CC BY-NC-SA 2.0)

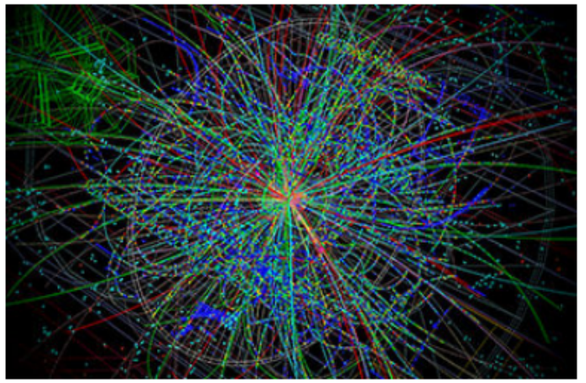
# AI is changing how we do science. Get a glimpse

By Science News Staff | Jul. 5, 2017 , 11:00 AM

## AI's early proving ground: the hunt for new particles

Particle physicists began fiddling with artificial intelligence (AI) in the late 1980s, just as the term "neural network" captured the public's imagination. Their field lends itself to AI and machine-learning algorithms because nearly every experiment centers on finding subtle spatial patterns in the countless, similar readouts of complex particle detectors—just the sort of thing at which AI excels. "It took us several years to convince people that this is not just some magic, hocus-pocus, black box stuff," says Boaz Klima, of Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois, one of the first physicists to embrace the techniques. Now, AI techniques number among physicists' standard tools.

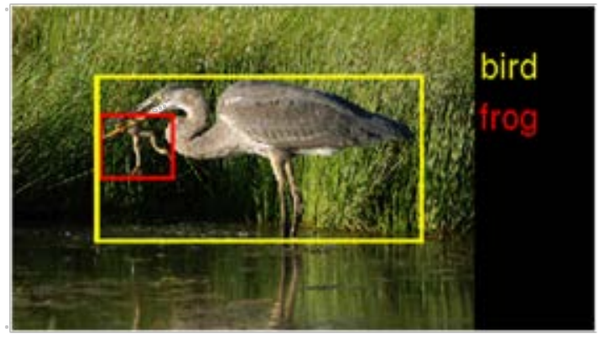
Particle physicists strive to understand the inner workings of the universe by smashing subatomic particles together with enormous energies to blast out exotic new bits of matter. In 2012, for example, teams working with the world's largest proton collider, the Large Hadron Collider (LHC) in Switzerland, discovered the long-predicted Higgs boson, the fleeting particle that is the linchpin to physicists' explanation of how all other fundamental particles get their mass.

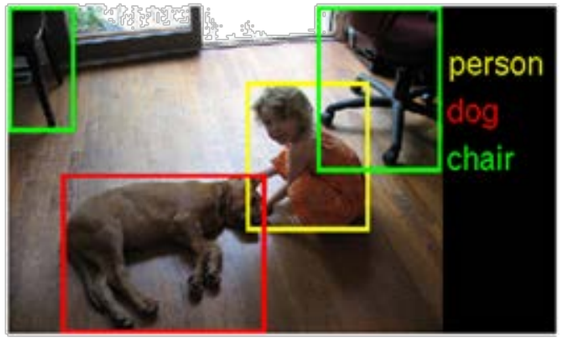


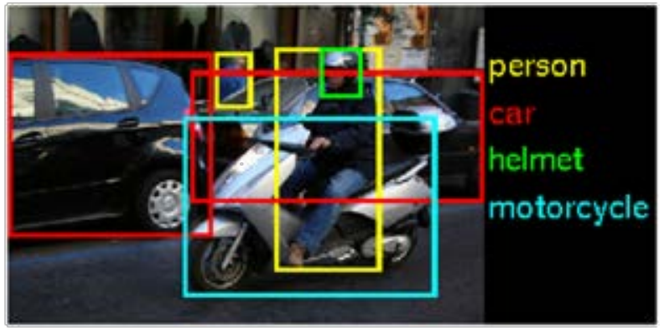
Neural networks search for fingerprints of new particles in the debris of collisions at the LHC.

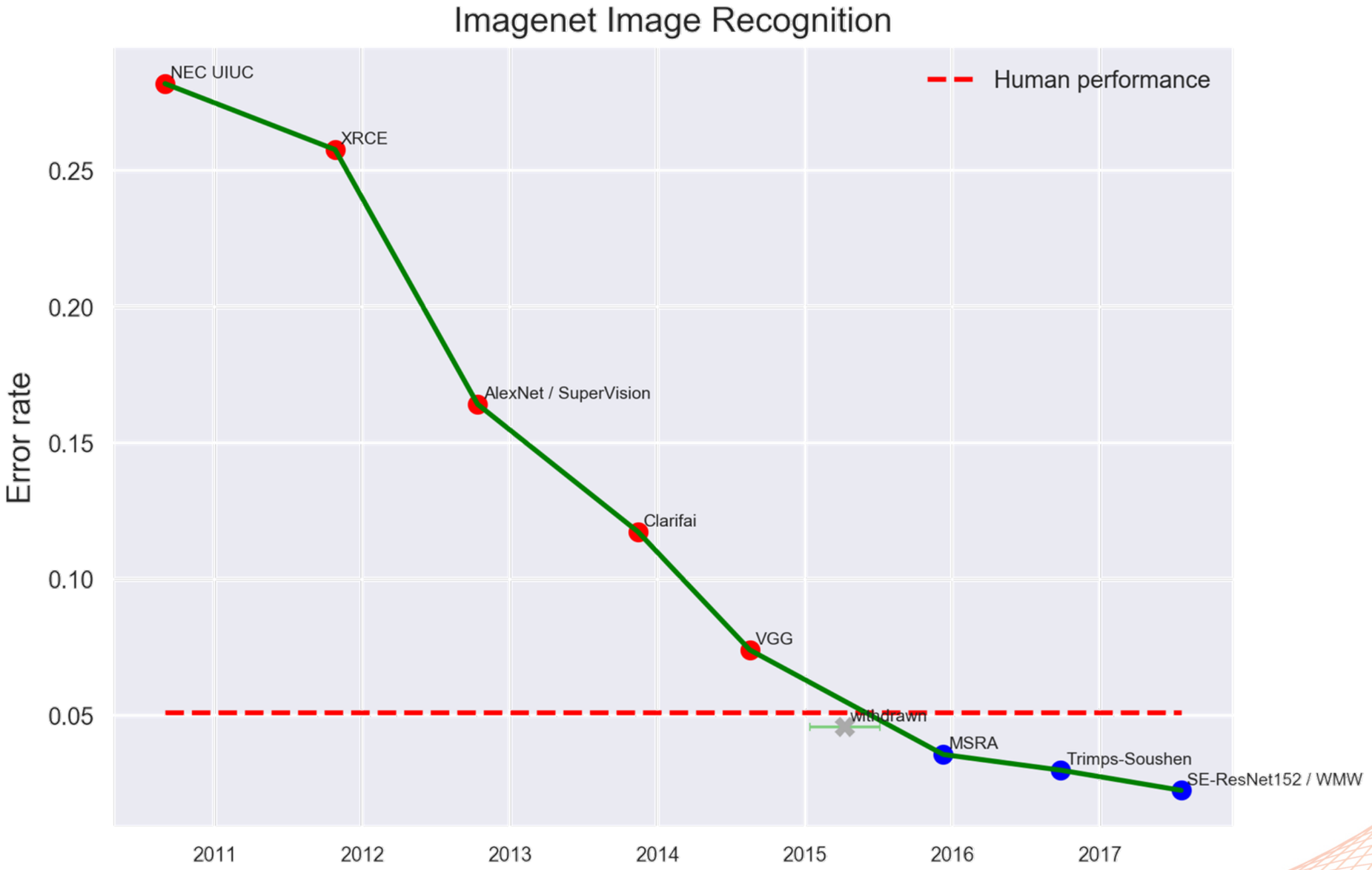


Such exotic particles don't come with labels, however. At the LHC, a Higgs boson emerges from roughly one out of every 1 billion proton collisions, and within a billionth of a picosecond it decays into other particles, such as a pair of photons or a




VIS 2017

Source: Electronic Frontier Foundation AI Progress Measurement

ARTICLE PREVIEW

view full access options ▸

*NATURE* | LETTER

日本語要約

# Dermatologist-level classification of skin cancer with deep neural networks

**Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun**

**Affiliations** | **Contributions** | **Corresponding authors**

*Nature* **542**, 115–118 (02 February 2017) | doi:10.1038/nature21056

Received 28 June 2016 | Accepted 14 December 2016 | Published online 25 January 2017

**Corrigendum (June, 2017)**



Skin cancer, the most common human malignancy[1, 2, 3], is primarily diagnosed visually, beginning with an initial clinical screening and followed potentially by dermoscopic analysis, a biopsy and histopathological examination. Automated classification of skin lesions using images is a challenging task owing to the fine-grained variability in the appearance of skin lesions. Deep

**Editor's summary** العربية

Andre Esteva *et al.* used 129,450 clinical images of skin disease to train a deep convolutional neural network to classify skin lesions. The result is an algorithm that can classify lesions from photog…

**Associated links**

News & Views

Medicine: The final frontier in cancer diagnosis

by Leachman and Merlino

**Related video**

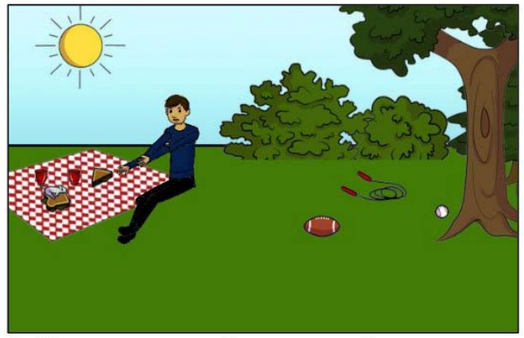
Digital doctor: AI singles out skin cancer from photos



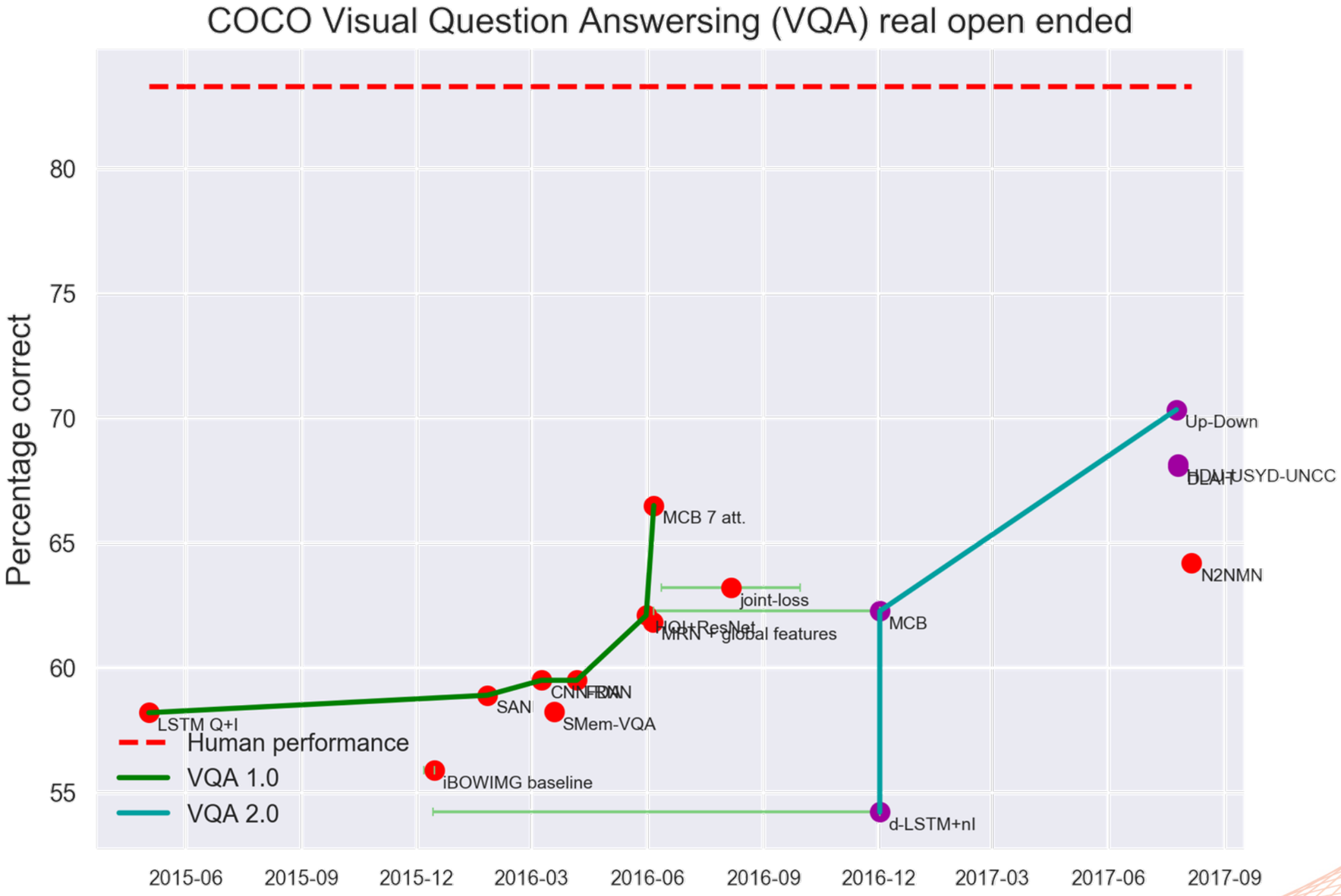
**Source: Nature 542, 115-118 (February 2017)**

How many slices of pizza are there?
Is this a vegetarian pizza?

What color are her eyes?
What is the mustache made of?



Is this person expecting company?
What is just under the tree?




VIS 2017

Source: Electronic Frontier Foundation AI Progress Measurement

# Endangering the Human-in-the-Loop

- Is the human-in-the-loop paradigm threatened by the recent machine learning advances?
- Are application areas of visualization shrinking due to recent advances in machine learning?

- Is the user's main task shifting from making decisions towards analyzing machine made decisions?

# Visualizing and Understanding Convolutional Networks

Matthew D. Zeiler and Rob Fergus

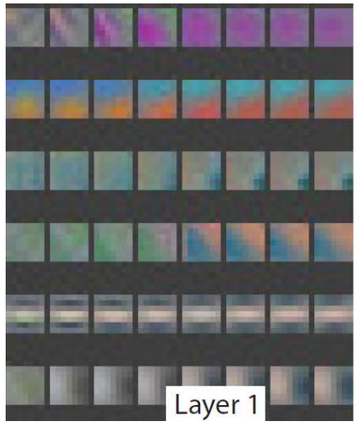
Dept. of Computer Science,
New York University, USA
{zeiler,fergus}@cs.nyu.edu

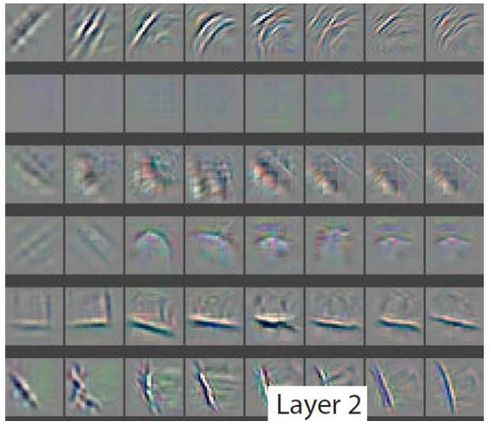
**Abstract.** Large Convolutional Network models have recently demonstrated impressive classification performance on the ImageNet benchmark Krizhevsky *et al.* [18]. However there is no clear understanding of why they perform so well, or how they might be improved. In this paper we explore both issues. We introduce a novel visualization technique that gives insight into the function of intermediate feature layers and the operation of the classifier. Used in a diagnostic role, these visualizations allow

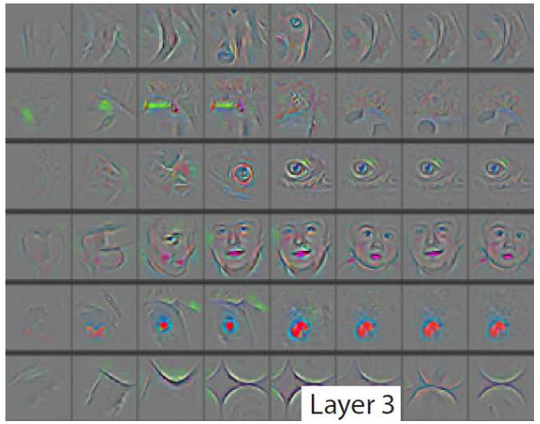
eral papers have shown that they can also deliver outstanding performance on more challenging visual classification tasks. Ciresan *et al.* [4] demonstrate state-of-the-art performance on NORB and CIFAR-10 datasets. Most notably, Krizhevsky *et al.* [18] show record beating performance on the ImageNet 2012 classification benchmark, with their convnet model achieving an error rate of 16.4%, compared to the 2nd place result of 26.1%. Following on from this work, Girshick *et al.* [10] have shown leading detection performance on the PASCAL VOC dataset. Several factors are responsible for this dramatic improvement in performance: (i) the availability of much larger training sets, with millions of labeled examples; (ii) powerful GPU implementations, making the training of very large models practical and (iii) better model regularization strategies, such as Dropout [14].

Despite this encouraging progress, there is still little insight into the internal operation and behavior of these complex models, or how they achieve such good performance. From a scientific standpoint, this is deeply unsatisfactory. Without clear understanding of how and why they work, the development of better models is reduced to trial-and-error. In this paper we introduce a visualization

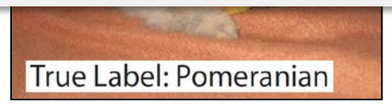
D. Fleet et al. (Eds.): ECCV 2014, Part I, LNCS 8689, pp. 818–833, 2014.

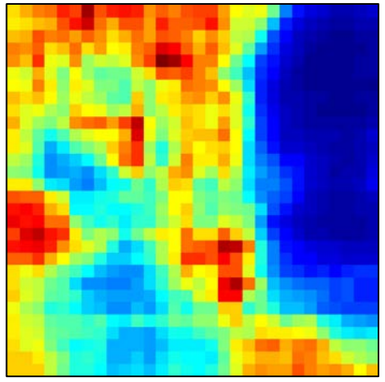




Used in a diagnostic role,
these visualizations allow us to find model architectures
that outperform Krizhevsky *et al.* on the ImageNet classification benchmark.

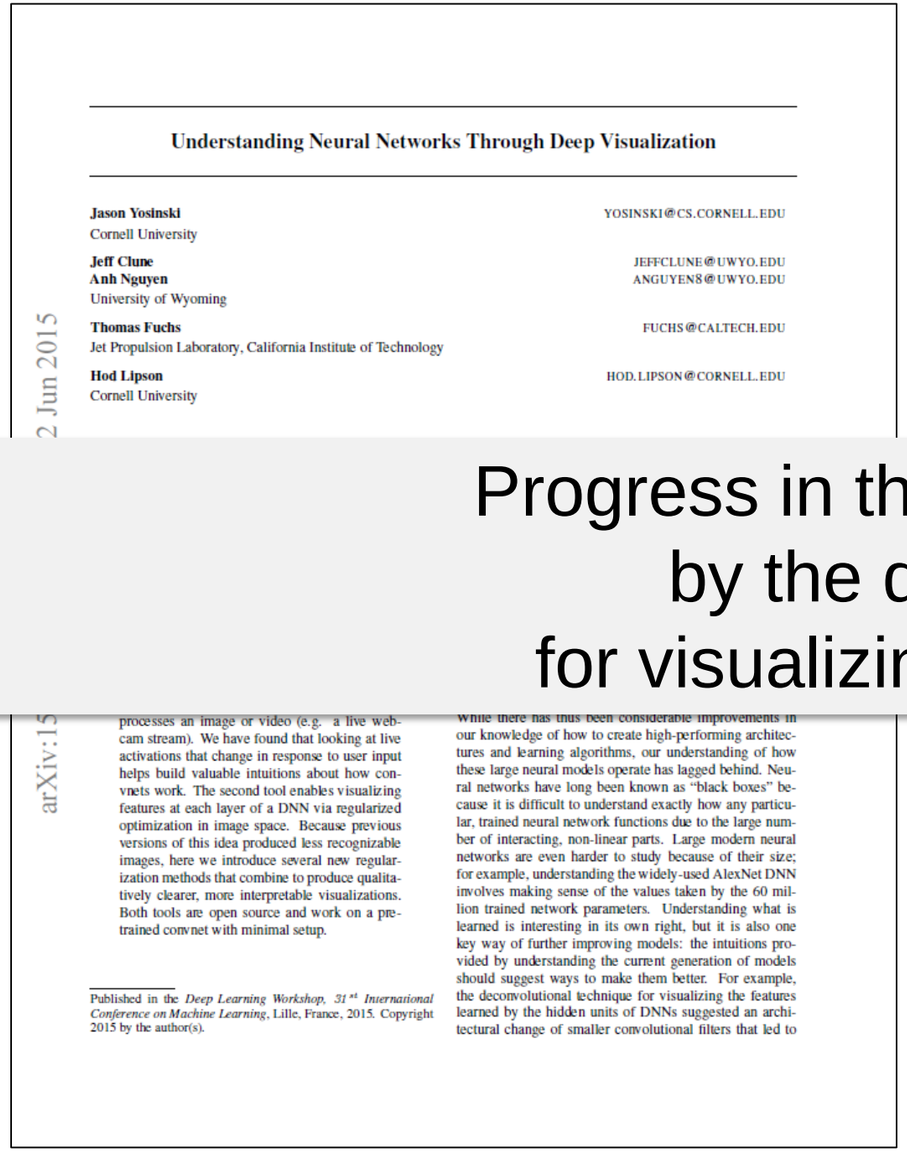






VIS 2017

[Zeiler & Ferguson, ECCV 2014]



Understanding Neural Networks Through Deep Visualization


**Jason Yosinski** YOSINSKI@CS.CORNELL.EDU
Cornell University

**Jeff Clune** JEFFCLUNE@UWYO.EDU
**Anh Nguyen** ANGUYEN8@UWYO.EDU
University of Wyoming

**Thomas Fuchs** FUCHS@CALTECH.EDU
Jet Propulsion Laboratory, California Institute of Technology

**Hod Lipson** HOD.LIPSON@CORNELL.EDU
Cornell University





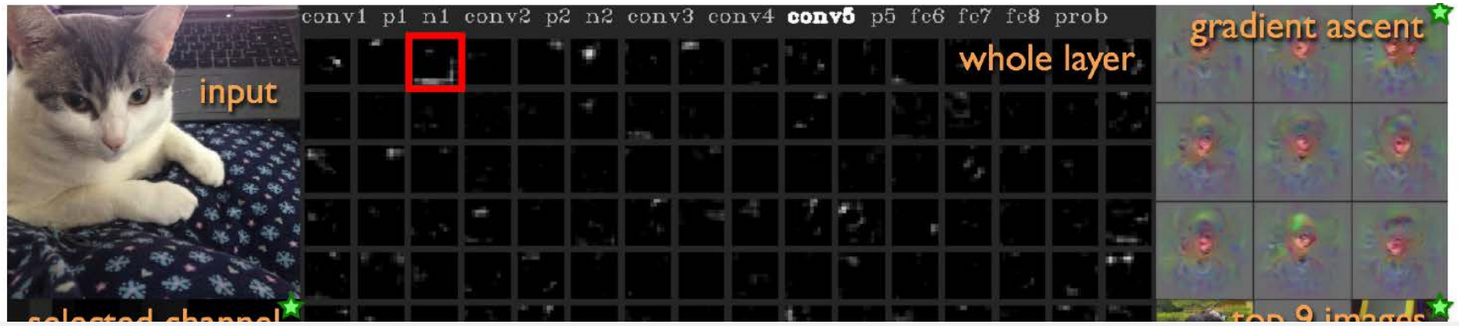


processes an image or video (e.g. a live webcam stream). We have found that looking at live activations that change in response to user input helps build valuable intuitions about how convnets work. The second tool enables visualizing features at each layer of a DNN via regularized optimization in image space. Because previous versions of this idea produced less recognizable images, here we introduce several new regularization methods that combine to produce qualitatively clearer, more interpretable visualizations. Both tools are open source and work on a pre-trained convnet with minimal setup.

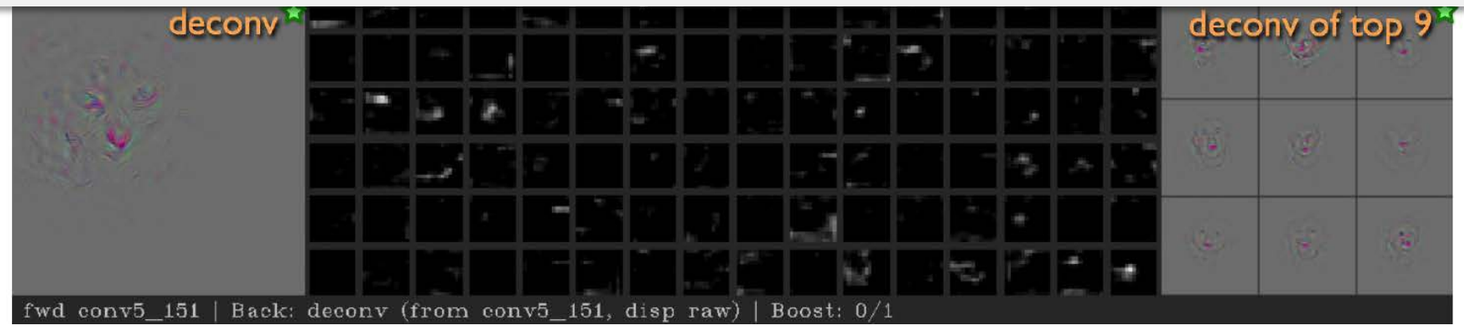


Published in the *Deep Learning Workshop, 31st International Conference on Machine Learning*, Lille, France, 2015.
 

While there has thus been considerable improvements in our knowledge of how to create high-performing architectures and learning algorithms, our understanding of how these large neural models operate has lagged behind. Neural networks have long been known as "black boxes" because it is difficult to understand exactly how any particular, trained neural network functions due to the large number of interacting, non-linear parts. Large modern neural networks are even harder to study because of their size; for example, understanding the widely-used AlexNet DNN involves making sense of the values taken by the 60 million trained network parameters. Understanding what is learned is interesting in its own right, but it is also one key way of further improving models: the intuitions provided by understanding the current generation of models should suggest ways to make them better. For example, the deconvolutional technique for visualizing the features learned by the hidden units of DNNs suggested an architectural change of smaller convolutional filters that led to




Progress in the field will be further accelerated
by the development of better tools
for visualizing and interpreting neural nets.




VIS 2017

[Yosinsky et al., Deep Learning Workshop 2015]

# ML Sessions @ VIS 2017

## 4:40–5:40 PM

**Keynote**

**Visualization: The Secret Weapon For Machine Learning**

Speakers: Fernanda Viegas and Martin Wattenberg, *Google*

Machine learning is playing an increasingly influential role in the world, due to dramatic technical leaps in recent years. But these new developments bring their own questions. What is the best way to train models and to debug them? How can we understand what is going on under the hood of deep neural networks? It turns out that visualization can play a central role in answering these questions. We'll discuss recent work that shows how interactive exploration can help people use, interpret, and learn about machine intelligence.

*Room 106-ABC*

**Tutorial (8:30 AM–12:10 PM)**

**Vis+ML: Symbiosis of Visualization and Machine Learning**

Contributors: Abon Chaudhuri, Yifan Hu, Xiaotong Liu, Yang Wang

Visualization and machine learning (ML) have come close to each other in recent years more than ever before. Visualization has emerged as a popular technique to understand the inner working and performance of machine learning and of late, deep learning algorithms. At the same time, machine learning techniques such as dimensionality reduction, clustering, and classification have been

*Room 101-ABC*

**Workshop (2:00–5:55 PM)**

**VADL 2017: Workshop on Visual Analytics for Deep Learning**

Contributors: Jaegul Choo, Shixia Liu, Jason Yosinski, Deokgun Park

VADL 2017, the workshop on visual analytics for deep learning, is a half-day workshop held in conjunction with the IEEE VIS 2017 Conference. The primary goal of the workshop is to bridge the gap by bringing together researchers from both the machine learning and visual analytics fields, which allows us to push the boundary of deep learning. The workshop should provide an opportunity to discuss and explore ways to harmonize the power of automated techniques and exploratory nature of interactive visualization.

*Room 301-C*

**VAST Papers**

**ML1: Deep Learning**

Chair: Ross Maciejewski

**[J] Analyzing the Training Processes of Deep Generative Models,** Mengchen Liu, Jiaxin Shi, Kelei Cao, Jun Zhu, Shixia Liu

**[C] Understanding Hidden Memories of Recurrent Neural Networks,** Yao Ming, Shaozu CAO, Ruixiang Zhang, Zhen Li, Yuanzhe Chen, Yangqiu Song, Huamin Qu

**[J] ActiVis: Visual Exploration of Industry-Scale Deep Neural Network Models,** Minsuk Kahng, Pierre Y. Andrews, Aditya Kalro, Duen Horng (Polo) Chau

**[J] DeepEyes: Progressive Visual Analytics for Designing Deep Neural Networks,** Nicola Pezzotti, Thomas Höllt, Jan van Gemert, Boudewijn P.F. Lelieveldt, Elmar Eisemann, Anna Vilanova

*207 Lecture Hall*

**VIS Panel**

**How do Recent Machine Learning Advances Impact the Data Visualization Research Agenda?**

Panelists: Timo Ropinski (Organizer), Daniel Archambault, Min Chen, Ross Maciejewski, Klaus Mueller, Alexandru Telea, Martin Wattenberg

Nowadays, machine learning approaches have revolutionized many domains. As this pushes the human out of the loop, the human-in-the-loop paradigm might be endangered. Thus, we would like to investigate, which old visualization challenges are rendered obsolete, and which new visualization challenges arise from the recent advances in machine learning.

*Room 301-C*

**VAST Papers**

**ML2: Cluster Analysis**

Chair: Tatiana von Landesberger

**[J] Visualizing Confidence in Cluster-based Ensemble Weather Forecast Analyses,** Alexander Kumpf, Bianca Tost, Marlene Baumgart, Michael Riemer, Rüdiger Westermann, Marc Rautenhaus

**[J] SOMFlow: Guided Exploratory Cluster Analysis with Self-Organizing Maps and Analytic Provenance,** Dominik Sacha, Matthias Kraus, Jürgen Bernard, Michael Behrisch, Tobias Schreck, Yuki Asano, Daniel A. Keim

**[J] Towards a Systematic Combination of Dimension Reduction and Clustering in Visual Analytics,** John Wenskovitch, Ian Crandell, Naren Ramakrishnan, Leanna House, Scotland Leman, Chris North

**[J] Clustervision: Visual Supervision of Unsupervised Clustering,** Bum Chul Kwon, Ben Eysenbach, Janu Verma, Kenney Ng, Christopher deFilippi, Walter F. Stewart, Adam Perer

*Room 102-ABC*

**VAST Papers**

**ML3: Classification**

Chair: Huamin Qu

**[J] Do Convolutional Neural Networks Learn Class Hierarchy?** Bilal Alsallakh, Amin Jourabloo, Mao Ye, Xiaoming Liu, Liu Ren

**[J] Visual Diagnosis of Tree Boosting Methods,** Shixia Liu, Jiannan Xiao, Junlin Liu, Xiting Wang, Jing Wu, Jun Zhu

**[C] A Workflow for Visual Diagnostics of Binary Classifiers using Instance-Level Explanations,** Josua Krause, Aritra Dasgupta, Jordan Swartz, Yindalon Aphinyanaphongs, Enrico Bertini

**[J] TreePOD: Sensitivity-Aware Selection of Pareto-Optimal Decision Trees,** Thomas Mühlbacher, Lorenz Linhardt, Torsten Möller, Harald Piringer

VIS 2017

# ML Sessions @ VIS 2016

*Holiday 4+5*

**VAST Papers**

**Machine Learning**
Chair: Torsten Möller

**[J] Squares: Supporting Interactive Performance Analysis for Multiclass Classifiers,** Donghao Ren, Saleema Amershi, Bongshin Lee, Jina Suh, Jason D. Williams

**[J]** [Best Paper Award] **An Analysis of Machine- and Human-Analytics in Classification,** Gary K. L. Tam, Vivek Kothari, Min Chen

**[J] Multi-Resolution Climate Ensemble Parameter Analysis with Nested Parallel Coordinates Plots,** Junpeng Wang, Xiaotong Liu, Han-Wei Shen, Guang Lin

**[J] Towards Better Analysis of Deep Convolutional Neural Networks,** Mengchen Liu, Jiaxin Shi, Zhen Li, Chongxuan Li, Jun Zhu, Shixia Liu

**[J] Visualizing the Hidden Activity of Artificial Neural Networks,** Paulo E. Rauber, Samuel G. Fadel, Alexandre X. Falcão, Alexandru C. Telea

VIS 2017

# Extending the Visualization Research Agenda

- How to make machine-made decisions more transparent to humans?
- How can visualization researchers capitalize from the shift from programmable systems to learning systems?
- When perceptual problems are solvable, what does this mean for more cognitive problems, where the user has to do some planning?

- What changes when enabling the human to judge machine-made decisions, rather than making man-made decisions?

Daniel Archambault
Swansea University

Min Chen
Oxford University

Ross Maciejewski
Arizona State University

Klaus Mueller
Stony Brook University

Alexandru Telea
University of Groningen

Martin Wattenberg
Google Inc.

VIS 2017