



## NEURAL NETWORKS SHOW NEW PROMISE FOR MACHINE VISION

By Pam Frost Gorder

**T**WENTY YEARS AGO, GEOFFREY HINTON HAD AN IDEA THAT WAS AHEAD OF ITS TIME: TO HELP COMPUTERS LEARN FROM THEIR MISTAKES. HE WANTED TO CREATE artificial neural networks that could learn to “see” images and recognize patterns through a kind of trial and error in a way that mimicked some models of human brain function. He and his colleagues developed the first practical method for training neural networks, but there was just one problem.

“We could never make it work properly,” Hinton says. At least, not the way they wanted to. Obstacles such as the lack of computing power stood in their way, and controversy erupted when studies in neuroscience suggested that the human brain probably didn’t function like their model network.

Flash ahead to 2006, and computing power is no longer a problem. Neuroscience has discovered much about how the brain works, but much of how we see is still a mystery. Hinton, a computer scientist at the University of Toronto and the Canadian Institute for Advanced Research, has discovered some creative strategies to help neural networks fulfill their potential in pattern recognition and artificial intelligence, which he reported in a recent issue of *Science* (vol. 313, no. 5786, 2006, pp. 504–507). Machine vision is his near-term goal, but the real prize could be insight into the human brain.

### Basic Training

Despite the fact that Hinton couldn’t meet his original goal in 1986, neural networks have since become computational workhorses and perform a wide variety of statistical analyses in research and commercial software. Their model “neurons” are arranged in successive layers: a bottom layer receives the input data, and each layer processes the data before passing it to the layer above. The network can encode data—take data sets of many variables or dimensions and fold them down into smaller numbers that are computa-

tionally more manageable. It can also decode data—unfold encoded data so that higher-dimensional data can emerge. High-dimensional data requires many steps to fold or unfold, and several neuron layers to do the job.

Face images are a prime example of high-dimensional data that researchers would like to be able to analyze with machine vision systems. A computer that could identify people by their faces would be an invaluable security tool, but it hasn’t proved easy to design. Because each pixel in a digital image is a variable, the average megapixel camera could take a portrait that would require many layers of neurons to process.

The training method Hinton and his colleagues developed 20 years ago—backward propagation of error, or *backpropagation* for short—is now a standard procedure for training neural networks. When a network tries to identify a known object but comes up with the wrong answer, users feed the right answer back into the network. The process repeats, and if all goes well, the network eventually learns to provide the right answer. However, backpropagation isn’t an effective way to train networks with many layers because the correction (for example, “this is the right answer”) only penetrates the outermost neuron layers. The deeper layers are unreachable, so they function as a kind of black box. To keep the deeper layers from getting stuck on the wrong answer, the user must design the network algorithmically to be close to the right answer from the start. That’s an especially tricky job when the network has several layers or the right answers aren’t well known.

What makes Hinton’s latest work special is that it shows how to reach the deepest layers of a neural network: he and doctoral student Ruslan Salakhutdinov found an algorithm that can pretrain each layer as the network is built. Once a network is pretrained, backpropagation becomes an effective method for fine-tuning it.

Hinton and Salakhutdinov used their method to construct networks for three visual tasks: finding compact representations of curved lines, handwritten numbers, and faces in photographs. They joined an encoder network for each task to a decoder and tested whether an image fed into the encoder would successfully re-emerge from the decoder. The networks learned their tasks on thousands of training im-

ages, but were asked to encode–decode new images that they had never “seen” before. Even so, this method outperformed two other common methods, and did so while more closely emulating what Hinton suspects is happening in the human brain. It’s the realization of the system that he wanted to build back in ’86.

### So Many Dimensions, So Little Time

Hinton studied psychology before computer science, so even as he began to work on backpropagation, he knew that it alone didn’t make a good model for human learning. Neurons send chemical signals forward through adaptive connections in milliseconds, but they can’t send rapid signals backward through the same connections.

“When we’re learning to see, nobody’s telling us what the right answers are—we just look,” Hinton says. “Every so often, your mother says ‘that’s a dog,’ but that’s very little information. You’d be lucky if you got a few bits of information—even one bit per second—that way. The brain’s visual system requires  $10^{14}$  [neural] connections. And you only live for  $10^9$  seconds. So it’s no use learning one bit per second. You need more like  $10^5$  bits per second. And there’s only one place you can get that much information—from the input itself.”

That idea, known in psychology as *generative learning*, led him to his encoder–decoder. In this scheme, every pixel in an inputted image becomes an opportunity to learn, so having a megapixel image isn’t a bad thing. “If you make mistakes, you make mistakes in millions of pixels, so you get lots of error information,” he says.

Twenty years ago, computers weren’t powerful enough to run such a system, nor were there data sets large enough to train it. Compounding this was the problem of training deep layers in the neural network. Today, powerful CPUs are plentiful, as are large data sets. You could have predicted that they would both catch up to Hinton’s plan, but the last obstacle—finding a way to train the network—wasn’t a given. As to how he and Salakhutdinov came up with their pre-training algorithm, he explains it this way: “If you keep thinking hard about anything for 20 years, you’ll go a long way.” The entire code for the pretraining method is available on Hinton’s Web site ([www.cs.toronto.edu/~hinton/](http://www.cs.toronto.edu/~hinton/)).

Robert P.W. Duin, associate professor of electrical engineering, mathematics, and computer science at the Delft University of Technology in the Netherlands, develops learning strategies for both neural networks and a competing methodology called the support vector machine (SVM). He’s impressed that Hinton demonstrated his technique’s

effectiveness for a series of applications, but suspects that mapping the right model for a given problem isn’t as straightforward and could only be done by an expert. For his part, Hinton says that adapting his algorithm for different applications doesn’t involve a lot of tweaking so much as changing the network’s size.

Yiannis Aloimonos, a computer scientist at the University of Maryland, thinks that Hinton’s work will be “quite influential,” and not just because it opens the door to a variety of new, powerful learning techniques. He points to the April 2006 Columbia Theory Day conference ([www.columbia.edu/theory/sp06.html](http://www.columbia.edu/theory/sp06.html)), where Princeton University scientist Bernard Chazelle gave a presentation on data-driven algorithm design. “In that presentation, Chazelle argued that algorithmic design as we know it has reached its limitations and now moves into a new phase, where as we design new algorithms, we also have access to gargantuan amounts of data. We do statistical analysis on this data, and



## Seeking Exceptional Computational Scientists

for graduate fellowship program

### Department of Energy Computational Science Graduate Fellowship

The DOE CSGF program supports students pursuing doctoral studies at U.S. universities in scientific and engineering disciplines that rely upon high-performance computation in their thesis research.

#### Benefits include:

- Payment of tuition and required fees
- A competitive yearly stipend
- Research practicum at a DOE laboratory
- Yearly fellows conference
- Renewable for up to four years

Application Deadline: January 10, 2007.

Apply Now: [www.krellinst.org/csgf/application](http://www.krellinst.org/csgf/application)



Sponsored by the U.S. Department of Energy Office of Science and NNSA Programs.

Administered by USDOE by the Krell Institute under contract DE-FG02-97ER25308. This is an equal opportunity program that is open to all qualified persons without regard to race, sex, creed, age, physical disability or national origin.

Learn more about the Fellowship in DEIXIS, the DOE CSGF Annual Magazine. Request a free copy at <http://www.krellinst.org/csgf/reqmag/>



Contact: The Krell Institute

1609 Golden Aspen Drive, Suite 101, Ames, IA 50010  
515.956.3696 • [csgf@krellinst.org](mailto:csgf@krellinst.org) • [www.krellinst.org/csgf](http://www.krellinst.org/csgf)

# computing

in SCIENCE & ENGINEERING

## EDITOR IN CHIEF

**Norman Chonacky**  
cise-editor@aip.org

## ASSOCIATE EDITORS IN CHIEF

**Jim X. Chen, George Mason Univ.**  
jchen@cs.gmu.edu

**Denis Donnelly, Siena College**  
donnelly@siena.edu

**Douglass E. Post, Carnegie Mellon Univ.**  
post@ieee.org

**John Rundle, Univ. of California, Davis**  
rundle@physics.ucdavis.edu

## EDITORIAL BOARD MEMBERS

Klaus-Jürgen Bathe, Mass. Inst. of Technology, kjb@mit.edu

Antony Beris, Univ. of Delaware, beris@che.udel.edu

Michael W. Berry, Univ. of Tennessee, berry@cs.utk.edu

Bruce Boghosian, Tufts Univ., bruce.boghosian@tufts.edu

George Cybenko, Dartmouth College, gvc@dartmouth.edu

Jack Dongarra, Univ. of Tennessee, dongarra@cs.utk.edu

Rudolf Eigenmann, Purdue Univ., eigenman@ecn.purdue.edu

David Eisenbud, Mathematical Sciences Research Inst.,  
de@msri.org

William J. Feiereisen, Los Alamos Nat'l Lab, ill@feiereisen.net

Geoffrey Fox, Indiana State Univ., gcf@grids.ucs.indiana.edu

Sharon Glotzer, Univ. of Michigan, sglotzer@umich.edu

Steven Gottlieb, Indiana Univ., sg@indiana.edu

Anthony C. Hearn, RAND, hearn@rand.org

Charles J. Holland, Darpa, charles.holland@darpa.mil

M.Y. Hussaini, Florida State Univ., myh@cse.fsu.edu

Rachel Kuske, Univ. of British Columbia, rachel@math.ubc.ca

David P. Landau, Univ. of Georgia, dlandau@hal.physast.uga.edu

B. Vincent McKoy, California Inst. of Technology,  
mckoy@its.caltech.edu

Jill P. Mesirov, Whitehead/MIT Ctr. for Genome Research,  
mesirov@genome.wi.mit.edu

Charles Peskin, Courant Inst. of Mathematical Sciences,  
peskin@cims.nyu.edu

Constantine Polychronopoulos, Univ. of Illinois,  
cdp@csr.d.uiuc.edu

William H. Press, Los Alamos Nat'l Lab., wpress@lanl.gov

John Rice, Purdue Univ., jrr@cs.purdue.edu

Ahmed Sameh, Purdue Univ., sameh@cs.purdue.edu

Henrik Schmidt, MIT, henrik@keel.mit.edu

Greg Wilson, Univ. of Toronto, gvwilson@third-bit.com

## CONTRIBUTING EDITORS

Francis Sullivan, fran@super.org

Paul F. Dubois, paul@pfdubois.com

## OBSERVATOIRE LANDAU

By Rubin Landau, News Editor



This is the first issue of *CISE* in which I'm officially the news editor, so it's probably a good time to introduce myself and my function. I believe my editorial job is to work with the *CISE* staff in deciding what topics might be good news items, what points we should emphasize in features, and what news items

might be better suited as full articles (suggestions always welcome). I also get to write sidebars such as this in which I pretend to be a columnist.

### My Background

As a research scientist, my specialty has been computational few-body systems in particle and nuclear physics. As an educator, I now direct the undergraduate program in computational physics at Oregon State University, where I've developed five computation classes and written five books on these and related subjects. I'm active in national groups with interests in better integrating computation into education and widening the groups of people that use computation. (In fact, I just received an announcement from the US National Science Foundation entitled *CISE Pathways to Revitalized Undergraduate Computing Education*, whose focus is to establish groups that will "transform undergraduate computing education on a national scale"; [www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=nsf06608](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf06608).)

Although I've been working in multidisciplinary computational and physics education for two decades while also conducting basic research, it's only in the past few years that I've stopped feeling like a lone man out in the physics community. I've found kindred spirits in the computational science community, which appears to embrace multidisciplinary education more so than traditional educators do. I

we use the results to gain intuition for better modeling of our original problem," Aloimonos remembers.

To him, Hinton's ideas couple well with Chazelle's to form a new methodology, one in which large data sets let researchers map relationships among relevant variables. "We let the data itself tell us how things are related," he says. "Of course, the data cannot tell us everything—we still need to do some modeling, but now our modeling will be guided by the statistical analysis. In this new era, Hinton's deep auto-encoders and dimensionality reducers will become a basic tool for anyone developing new algorithms."

In his own work with Kwabena Boahen of Stanford Uni-

hope that by having my feet in several camps I'll be able to provide a reasoned perspective in this space. I believe the times they are now a-changing, and the four conferences I participated in this summer appear to prove that.

### My Observations

The 51st Annual Conference of the South African Institute of Physics was held in Capetown from 3–7 July 2006. Its theme was the broad relationship between physics and computers, both in showcasing computing's role in physics and in highlighting how physics research leads to improvements in computing. What I found particularly interesting and moving about this conference was the large number of students present and the mix of the various groups of people that constitute present-day South Africa. In fact, one of the stated reasons for focusing on how computing is changing physics education was that the students were interested in acquiring an education that equips them with the skills and confidence to use computers in problem solving, and thus find gainful employment in a developing economy in which only a few will end up as traditional physicists. In addition, I couldn't help but be awed by how the latest communication technologies have placed as distant a country as South Africa right in the heart of research being conducted with the CERN Large Hadron Collider Computing Grid and with the International Virtual Observatory Alliance's AstroGrid (with links to the South African Large Telescope).

Two conferences in New York also had sessions, or a major focus, on how computation affects education in the sciences and engineering. The first, the Summer Meeting of the American Association of Physics Teachers (AAPT) in Syracuse, 22–26 July, included special sessions organized by Norman Chonacky and David Winch on computation in undergraduate physics courses. Detailed reports on the sessions appear in the September/October 2006 issue of *CiSE*. I'll only add the comment that it's nice to see the AAPT beginning to recognize computation's importance in education after trying for years to keep computer programs out of its journals.

The International Conference on Computational Science

and Education, held in Rochester, NY, from 7–10 August, blended computational education and research together, possibly in recognition of how important proper education is for the future growth of computational science. In contrast to the AAPT meeting, which focused on undergraduate education, here we saw examples of how computation has been introduced into K–20 classes throughout several school districts within metropolitan Rochester. This ambitious project claims that introducing of computation within the scientific problem-solving paradigm leads to better science and math education—even for inner-city students.

The Conference on Computational Physics 2006 was held 29 August–1 September in Gyeongju, Republic of Korea. Here, too, I was struck by how important computation was in the reported research advances (in simulations involving multimillions of atoms and relativistic star systems, for example) and by the continuing reluctance of many researchers to discuss their algorithms and code verifications. But I was also struck by the large number of students in attendance (a larger fraction than at the US conferences), by their serious involvement in research, and by the researchers' interest in the education sessions. As in South Africa, we have vibrant countries with growing economies, and many appear to view computation as one way to support that growth and to acquire flexibility for employment.

**M**y conclusion? Computational science is gaining acceptance among some faculty and many students in different parts of the world. They view it as a modern way to learn science and to do research, and as a career opportunity that might provide them with employment, even if not in their specialty. Institutions and governments, in turn, appear to view computational science as a way to assist economic and intellectual development. As was pointed out at the conferences, three times as many degree or focused programs now exist in the computational sciences than five years ago. This is the type of good “news” that is a pleasure to report.

versity, Aloimonos designs chips that he hopes will one day integrate computer vision, hearing, and language capabilities into one cognitive system. It's a difficult enterprise that is tied to our understanding of how those capabilities are entwined in the human brain.

### Open Box

Hinton's strategy harkens back to the very early days of neural networks—the 1950s—when people wanted to train one neural layer at a time. His *Science* paper marks the first time anyone has penetrated the black box to show that this can indeed be done, even for very deep networks. Encoding and de-

coding a complex image such as a human face is an important first step toward developing machines that can see in useful ways beyond handwriting analysis or simple shape recognition.

Still, he'd like to eventually develop machines that do even more—ones that see the way we do and therefore act as tools to help us better understand ourselves. “I'm interested in two things. One is how to solve tough problems in artificial intelligence like shape recognition and speech recognition,” he says. “But the second is answering the question, how does the brain actually do it? And the new algorithm we've developed for pretraining these neural nets is probably much more like what the brain is using.”

## DEPARTMENT EDITORS

**Book Reviews:** Mario Belloni, Davidson College, mabelloni@davidson.edu

**Computing Prescriptions:** Isabel Beichl, Nat'l Inst. of Standards and Tech., isabel.beichl@nist.gov, and Julian Noble, Univ. of Virginia, jvn@virginia.edu

**Computer Simulations:** Muhammad Sahimi, University of Southern California, moe@iran.usc.edu, and Dietrich Stauffer, Univ. of Köln, stauffer@thp.uni-koeln.de

**Education:** David Winch, Kalamazoo College, winch@TaosNet.com  
**News:** Rubin Landau, Oregon State Univ., rubin@physics.oregonstate.edu

**Scientific Programming:** Konstantin Läufer, Loyola University, Chicago, klaufer@cs.luc.edu, and George K. Thiruvathukal, Loyola University, Chicago, gkt@cs.luc.edu

**Technologies:** James D. Myers, jimmyers@ncsa.uiuc.edu

**Visualization Corner:** Jim X. Chen, George Mason Univ., jchen@cs.gmu.edu, and R. Bowen Loftin, Texas A&M University, Galveston, loftin@tamug.edu

**Your Homework Assignment:** Dianne P. O'Leary, Univ. of Maryland, oleary@cs.umd.edu

## STAFF

**Senior Editor:** Jenny Stout, jstout@computer.org

**Group Managing Editor:** Steve Woods

**Staff Editors:** Kathy Clark-Fisher, Rebecca L. Deuel, and Brandi Ortega

**Contributing Editor:** Cheryl Baltes and Joan Taylor

**Production Editor:** Monette Velasco

**Magazine Assistant:** Hazel Kosky, cise@computer.org

**Technical Illustrator:** Alex Torres

**Publisher:** Angela Burgess, aburgess@computer.org

**Associate Publisher:** Dick Price

**Advertising Coordinator:** Marian Anderson

**Marketing Manager:** Georgann Carter

**Business Development Manager:** Sandra Brown

## AIP STAFF

**Circulation Director:** Jeff Bebee, jbebee@aip.org

**Editorial Liaison:** Charles Day, cday@aip.org

IEEE ANTENNAS AND  
PROPAGATION SOCIETY LIAISON

Don Wilton, Univ. of Houston, wilton@uh.edu

## IEEE SIGNAL PROCESSING SOCIETY LIAISON

Elias S. Manolakos, Northeastern Univ., elias@neu.edu

## CS PUBLICATIONS BOARD

Jon Rokne (chair), Michael R. Blaha, Mark Christensen, Frank Ferrante, Roger U. Fujii, Phillip Laplante, Bill N. Schilit, Linda Shafer, Steven L. Tanimoto, Wenping Wang

## CS MAGAZINE OPERATIONS COMMITTEE

Bill N. Schilit (chair), Jean Bacon, Pradip Bose, Arnold (Jay) Bragg, Doris L. Carver, Kwang-Ting (Tim) Cheng, Norman Chonacky, George Cybenko, John C. Dill, Robert E. Filman, David A. Grier, Warren Harrison, James Hendler, Sethuraman (Panch) Panchanathan, Roy Want

## EDITORIAL OFFICE

COMPUTING in SCIENCE & ENGINEERING

10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720

phone +1 714 821 8380; fax +1 714 821 4010; www.computer.org/cise/



IEEE Antennas &  
Propagation Society

IEEE

Signal Processing Society



Within the machine vision community, there is much discussion about whether neural networks are the best platform for this kind of research. The SVM is a notable competitor because it uses statistical algorithms to identify visual features in a single processing layer. SVMs offer simpler computation and—until now—improved performance over neural networks, as well as transparency. Hinton says with some pride that for the handwritten digit-recognition task, the best reported error rate for SVMs was 1.4 percent, but his pre-trained neural network with back-propagation achieved 1.2 percent. Other researchers have suggested that both SVMs and neural networks have a place in machine vision because some tasks seem better suited to one or the other.

Aloimonos agrees that, historically, there has been tension among researchers who use computer models such as neural networks and those who use purely statistical methods such as SVMs to simulate learning. “In my view, the future will bring the two groups together—the modelers will use the tools of the statistical learners in order to do better modeling.” He predicts that the joining of the two concepts will be the hottest topic in the discipline for years to come, especially as data sets grow even larger. “Somehow it is the spirit of the times,” he says. “Something felt on a daily basis by the use of Google.”

CS  
SE

**Pam Frost Gorder** is a freelance science writer based in Columbus, Ohio.



## Stay on Track

IEEE Internet Computing reports emerging tools, technologies, and applications implemented through the Internet to support a worldwide computing environment.

IEEE  
**Internet Computing**

www.computer.org/internet/