EMBO Practical Course
Population genomics: Background, tools and programming

30 March – 06 April 2019 | Procida, Italy

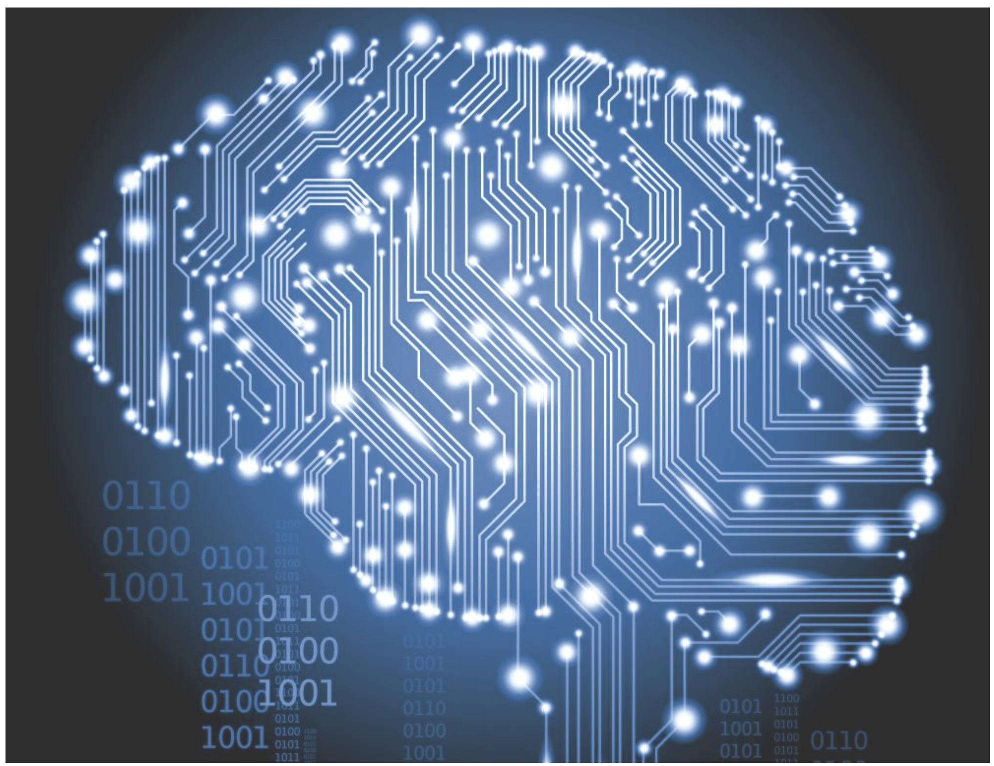
an introduction to deep learning

Marco Chierici and Margherita Francescatto
Fondazione Bruno Kessler
chierici@fbk.eu francescatto@fbk.eu

machine learning

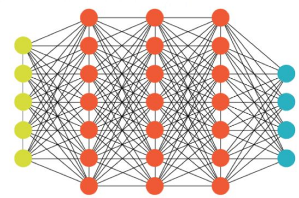
> *The ability to learn making **predictions from data** without being explicitly programmed*

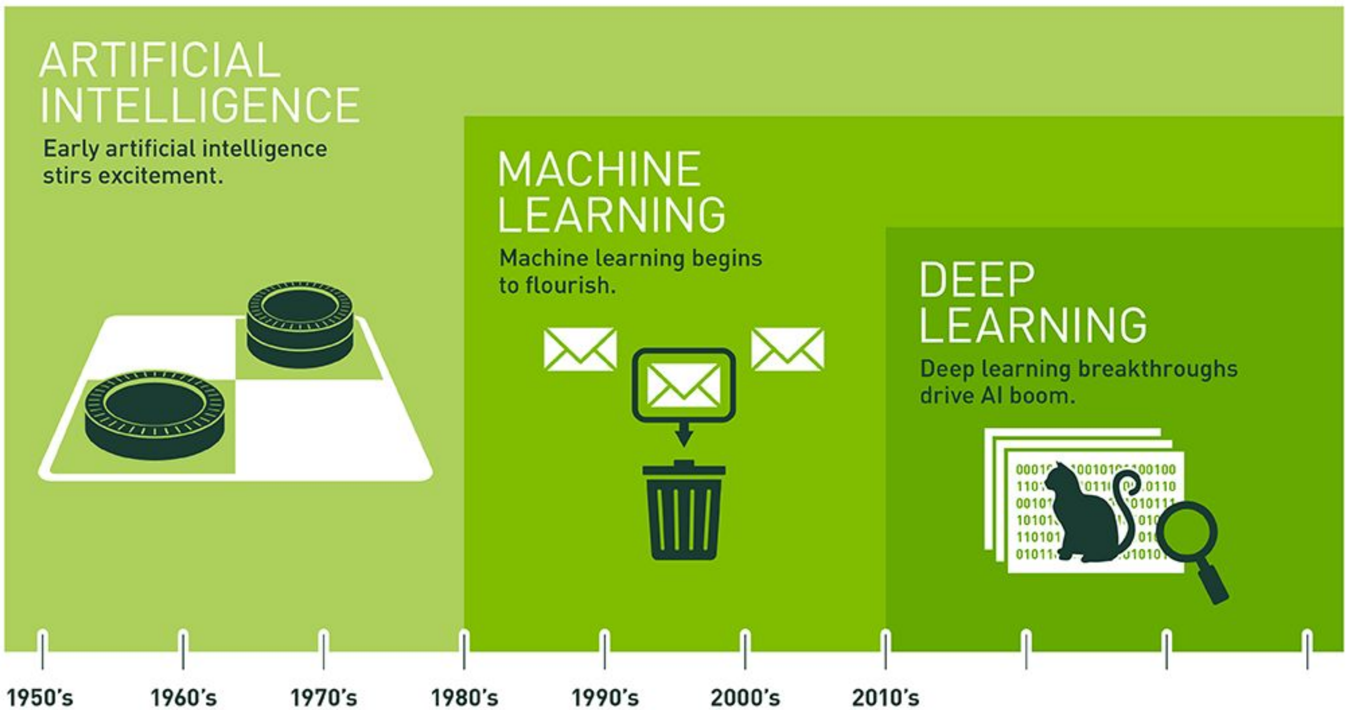
arthur samuel, 1959



deep learning

*The ability to **learn representations** of data with multiple levels of **abstraction** through multiple processing **layers***





FONDAZIONE BRUNO KESSLER

hitting the headlines

BBC | Sign in | News | Sport | Weather | Shop | Earth | Travel | Mo

NEWS

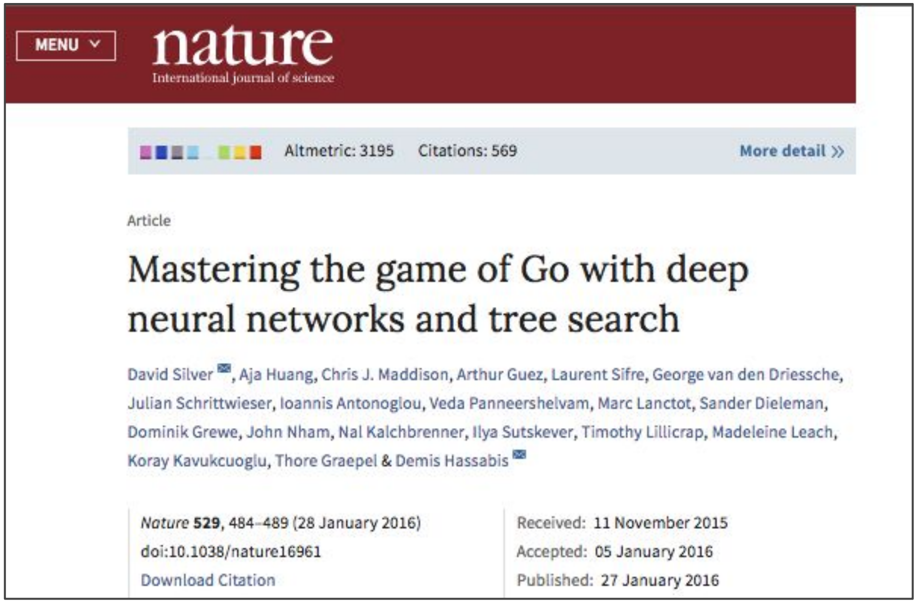
Home | Video | World | UK | Business | Tech | Science | Stories | Entertainment & Arts

Technology

Artificial intelligence: Google's AlphaGo beats Go master Lee Se-dol

12 March 2016

Share



MENU

nature

International journal of science

Altmetric: 3195 Citations: 569 More detail »

Article

Mastering the game of Go with deep neural networks and tree search

David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel & Demis Hassabis

Nature **529**, 484–489 (28 January 2016)
doi:10.1038/nature16961
Download Citation

Received: 11 November 2015
Accepted: 05 January 2016
Published: 27 January 2016

DeepMind

hitting the headlines

The Telegraph HOME | NEWS |

News | Science

News › Science

AlphaGo Zero: Google DeepMind supercomputer learns 3,000 years of human knowledge in 40 days

The Telegraph HOME | NEWS | SP

News | Science

News › Science

Entire human chess knowledge learned and surpassed by DeepMind's AlphaZero in four hours

MENU

nature

International journal of science

Altmetric: 2164 Citations: 1 More detail »

Article

Mastering the game of Go without human knowledge

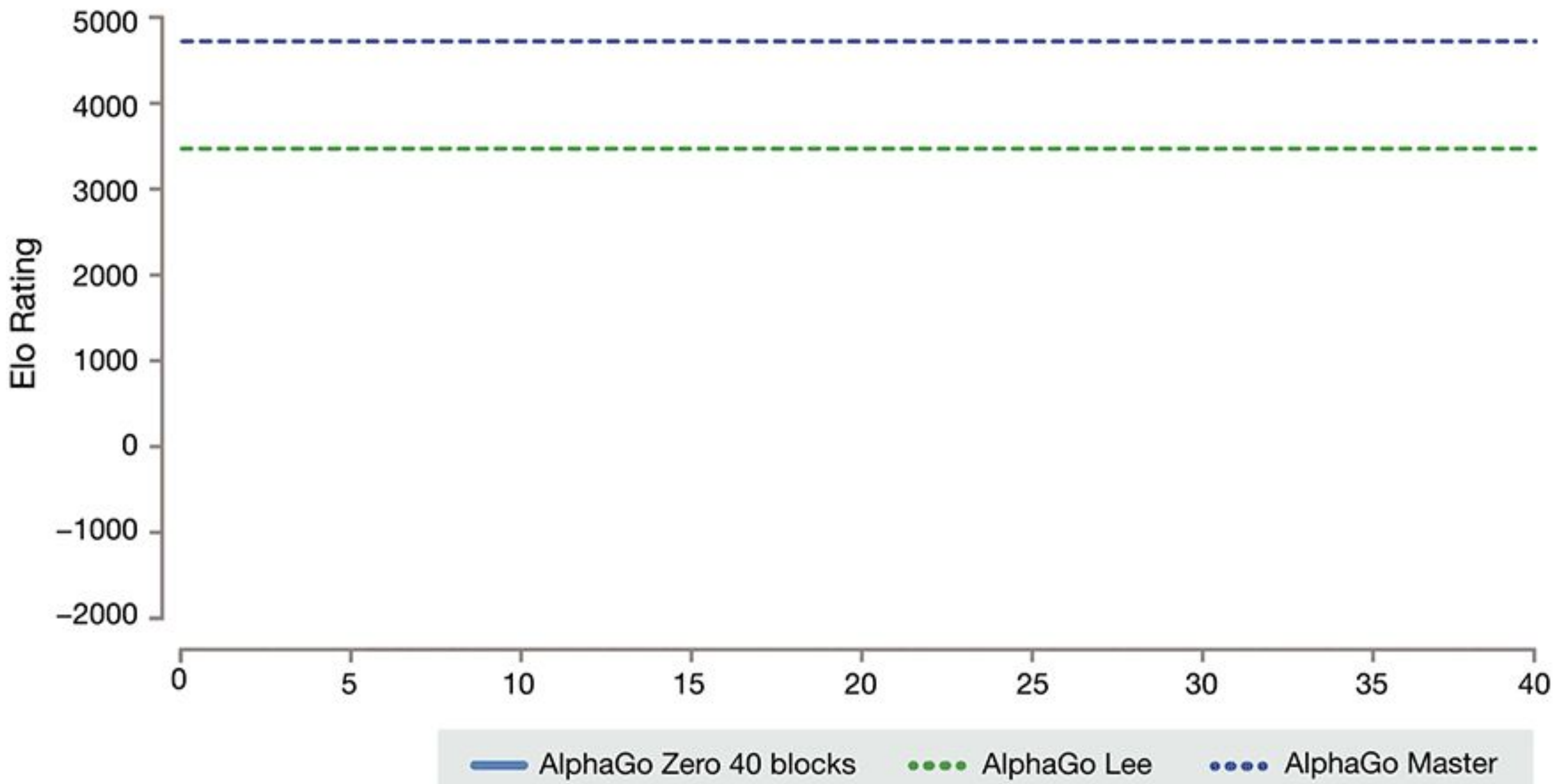
David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel & Demis Hassabis

Nature **550**, 354–359 (19 October 2017)
doi:10.1038/nature24270
Download Citation

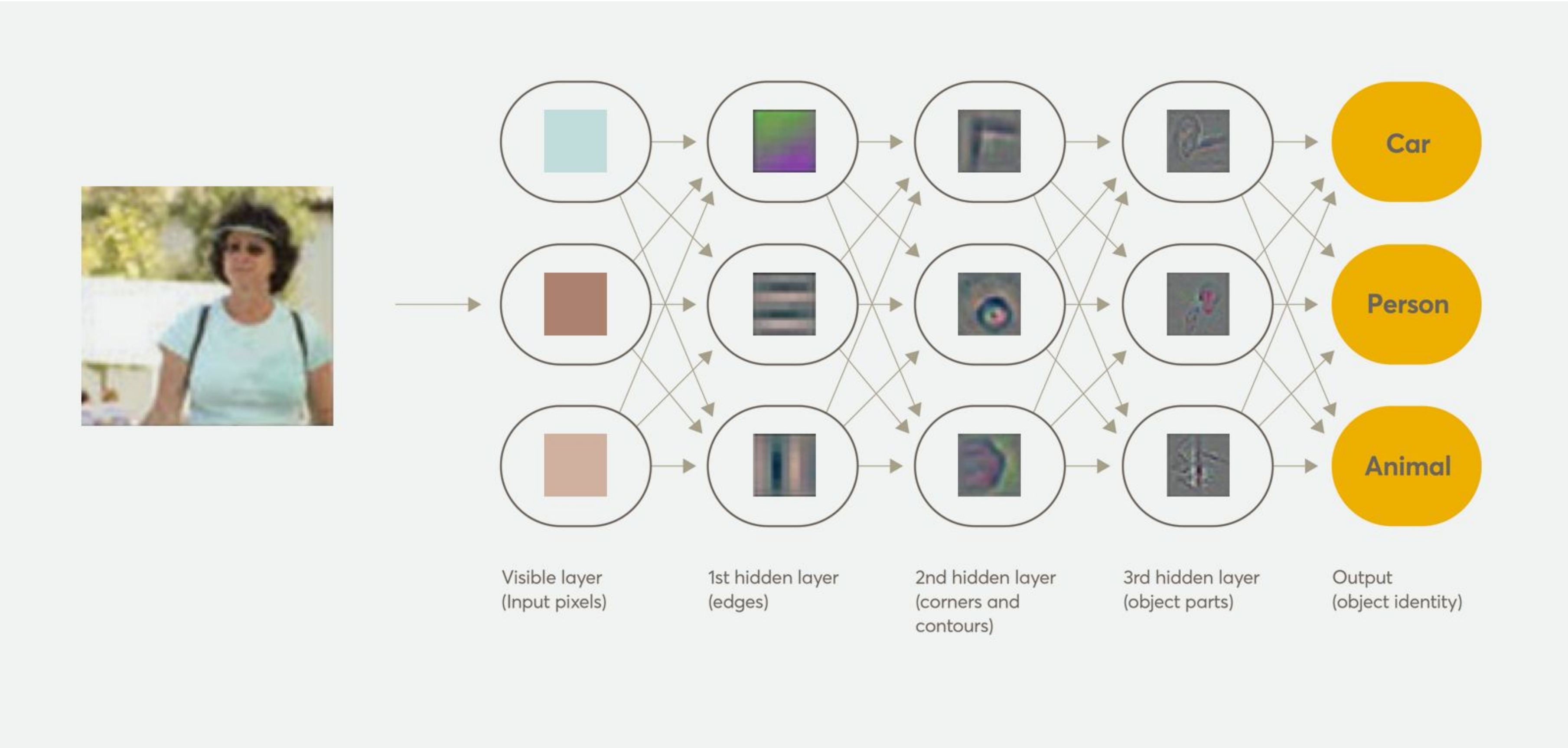
Received: 07 April 2017
Accepted: 13 September 2017
Published: 18 October 2017

FONDAZIONE BRUNO KESSLER | ICT CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGY

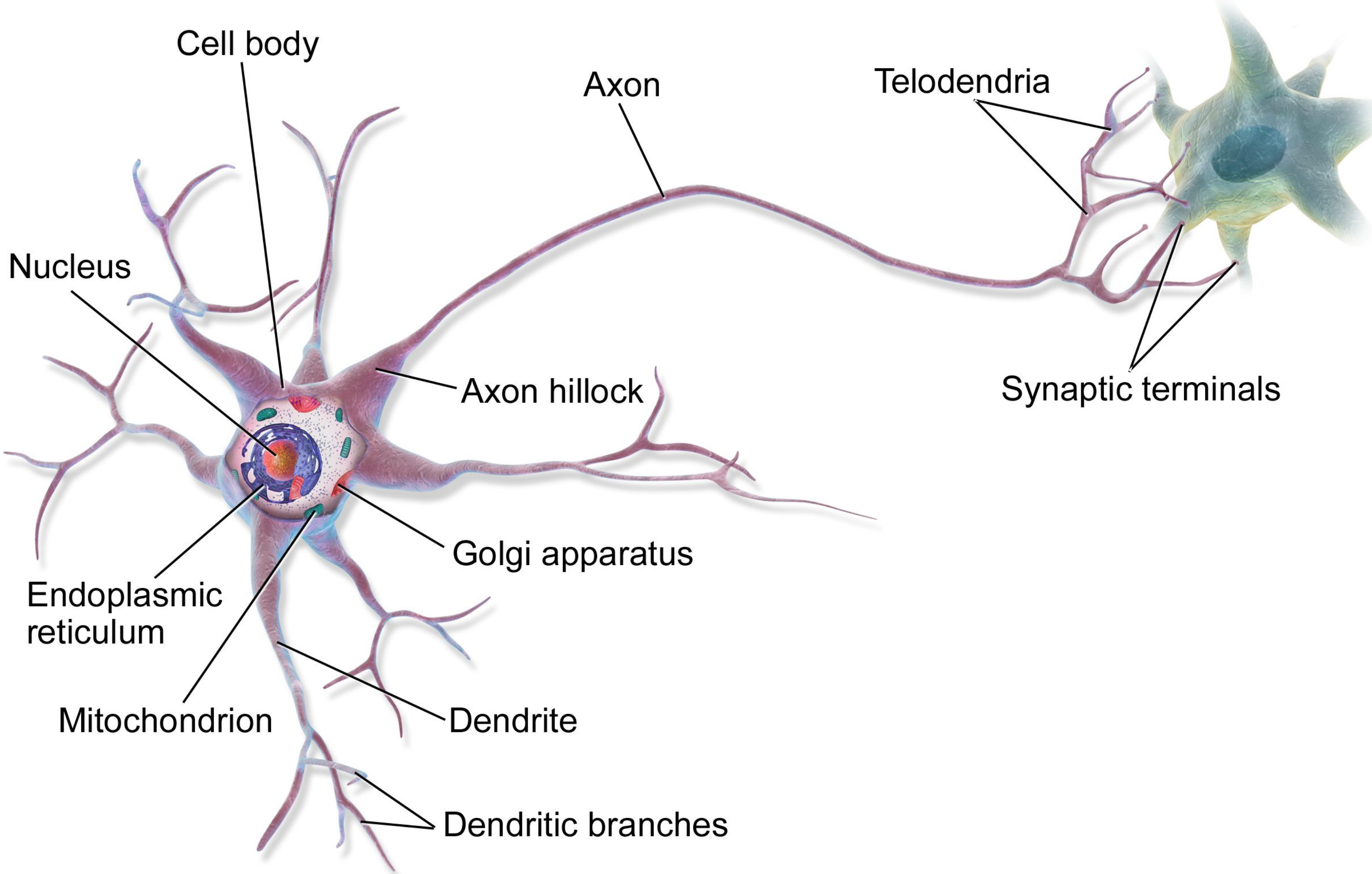
AlphaGo
training



deep learning
view

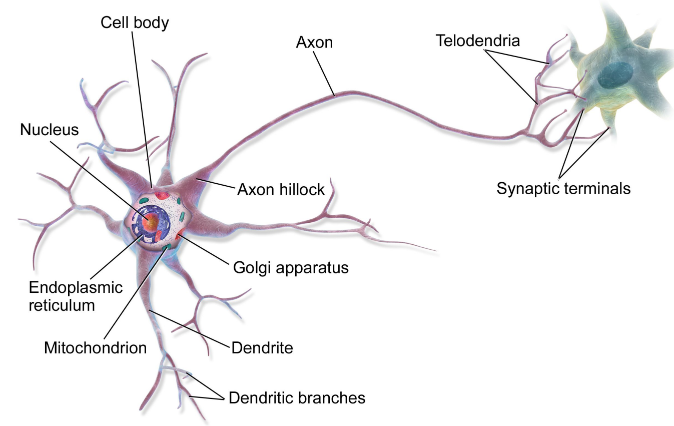


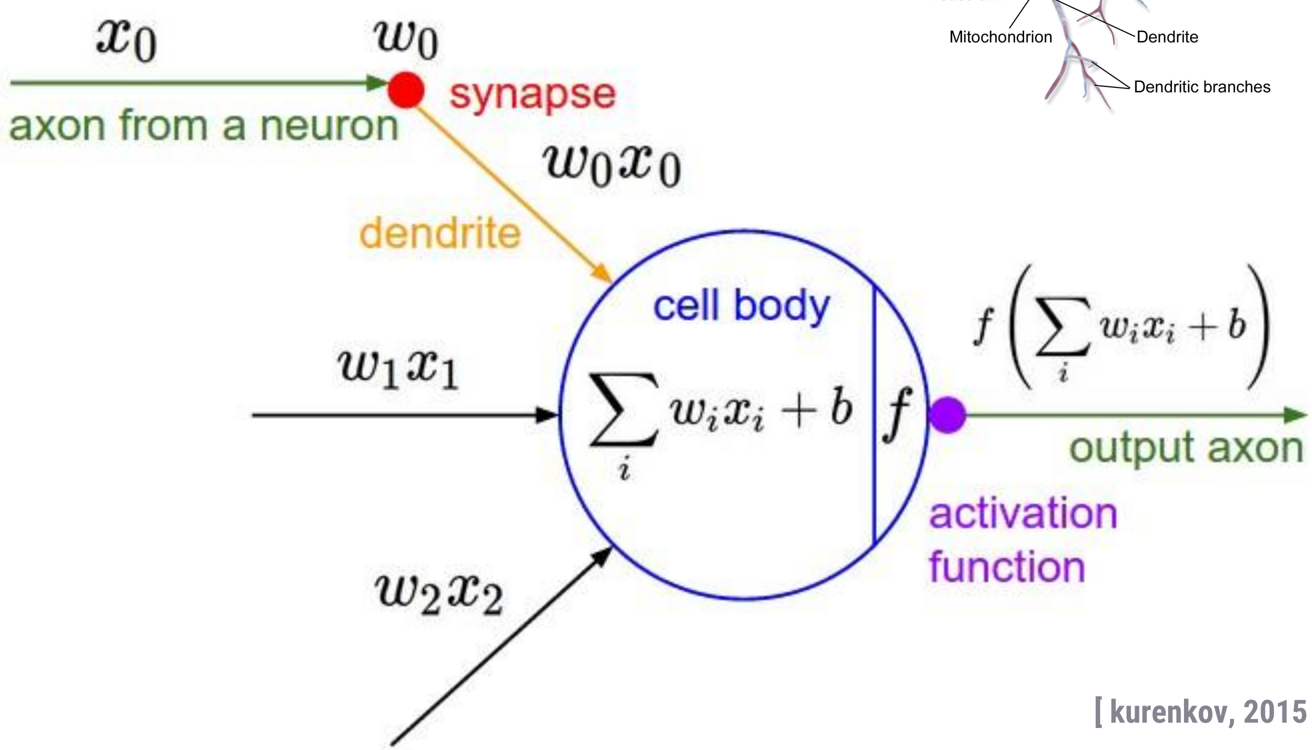
neurons



Cell body
Axon
Telodendria
Nucleus
Axon hillock
Synaptic terminals
Golgi apparatus
Endoplasmic reticulum
Mitochondrion
Dendrite
Dendritic branches

neuron model



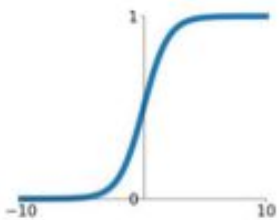


[kurenkov, 2015]

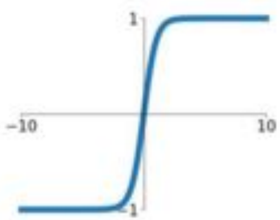
FONDAZIONE BRUNO KESSLER
ICT CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGY

activation functions

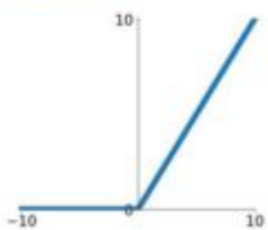
Sigmoid
$\sigma(x) = \frac{1}{1+e^{-x}}$



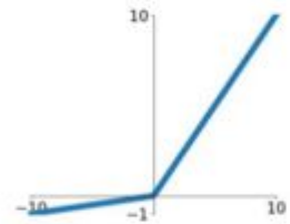
tanh
$\tanh(x)$



ReLU
$\max(0, x)$

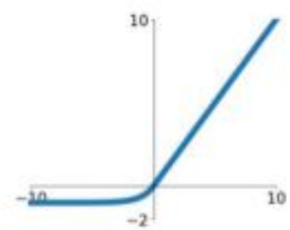


Leaky ReLU
$\max(0.1x, x)$



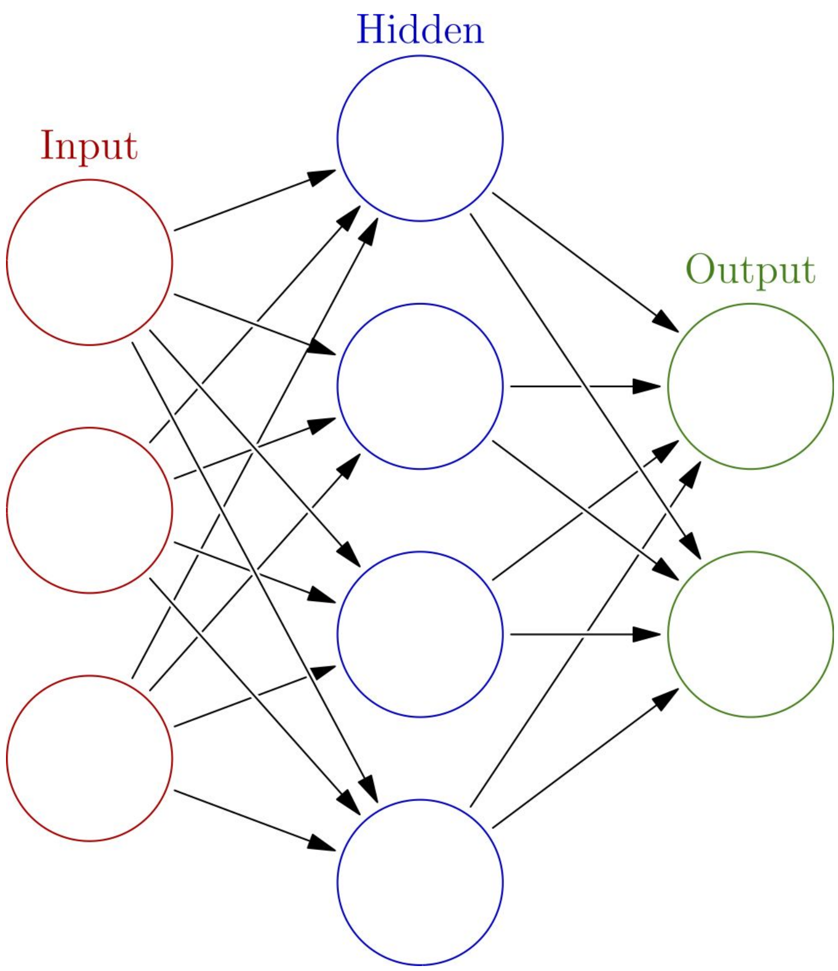
Maxout
$\max(w_1^T x + b_1, w_2^T x + b_2)$

ELU
$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$



[medium.com, 2017]

artificial
neural networks



layer

all neurons at the same depth

hidden layer

any layer but the input or output layer

deep learning

many hidden layers (> 3)

timeline

2016
The winning solution (AUC 0.99) of the **Camelyon challenge on detecting metastatic cancer** beats the human pathologist benchmark (AUC 0.96)

2015
A CNN by team from Microsoft beats the human benchmark (5% error rate) by bringing down the error rate to 3% in the ImageNet Challenge

2012
A CNN designed by a team at the University of Toronto wins the ImageNet Challenge bringing dow the error rate to 16% (compared to 25% 2011)

2009
A group around Andrew Ng introduce **Graphics Processing Units (GPUs)** for Deep Learning making them applicable on a large scale

2007
Fei Fei Li and colleagues at Princeton University start to collect a large database of annotated images, the **ImageNet**

2006
Hinton summarizes ideas of **multilayer neural networks** and training them to generate sensory data rather than to classify it

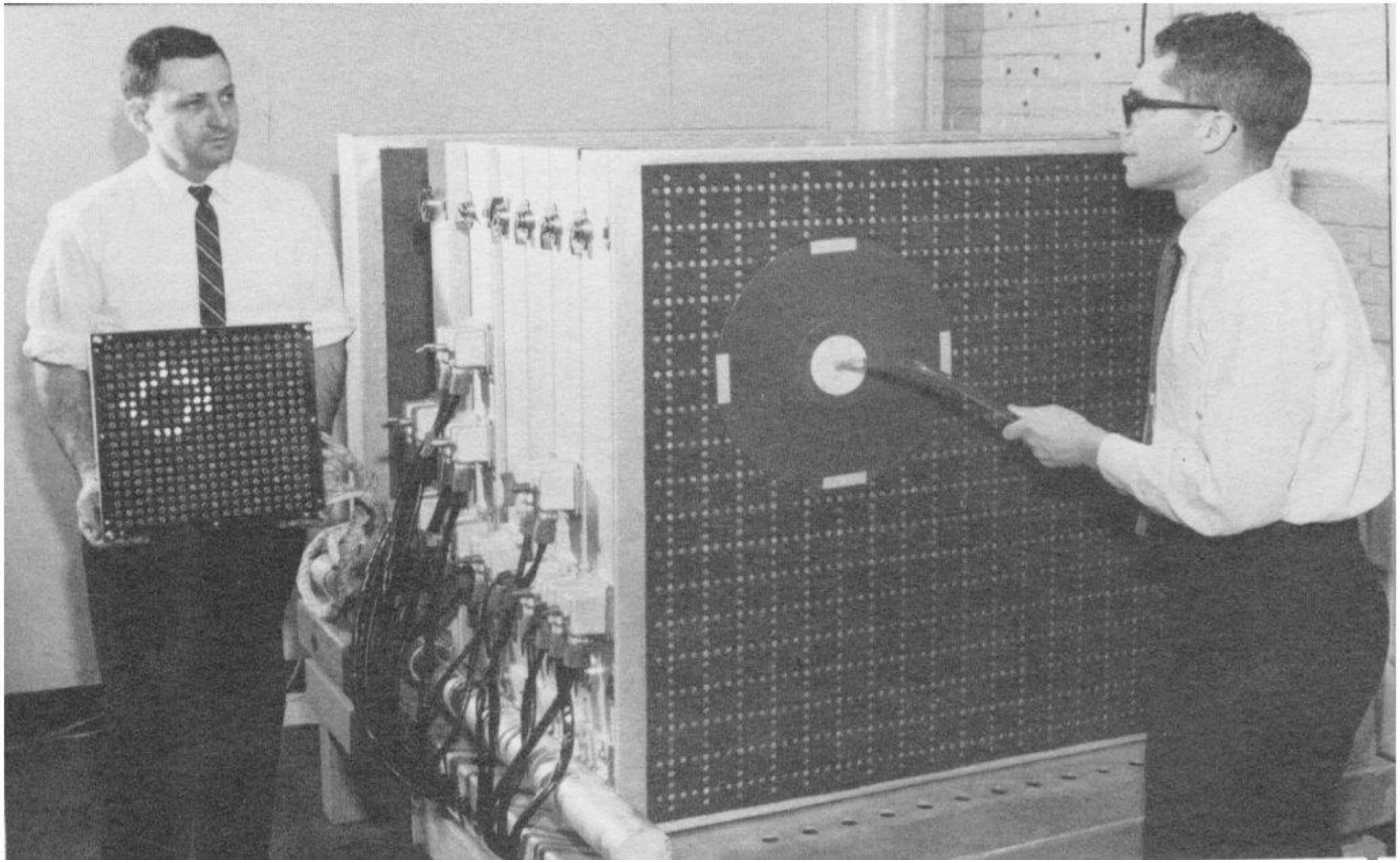
1989
A group around Yann LeCun successfully applies a back-propagation algorithm to a multi-layer neural network, **recognizing handwritten ZIP codes**

1986
Rumelhart, Hinton, and Williams introduce **backpropagation** as a learning procedure for "networks of neuron-like units"

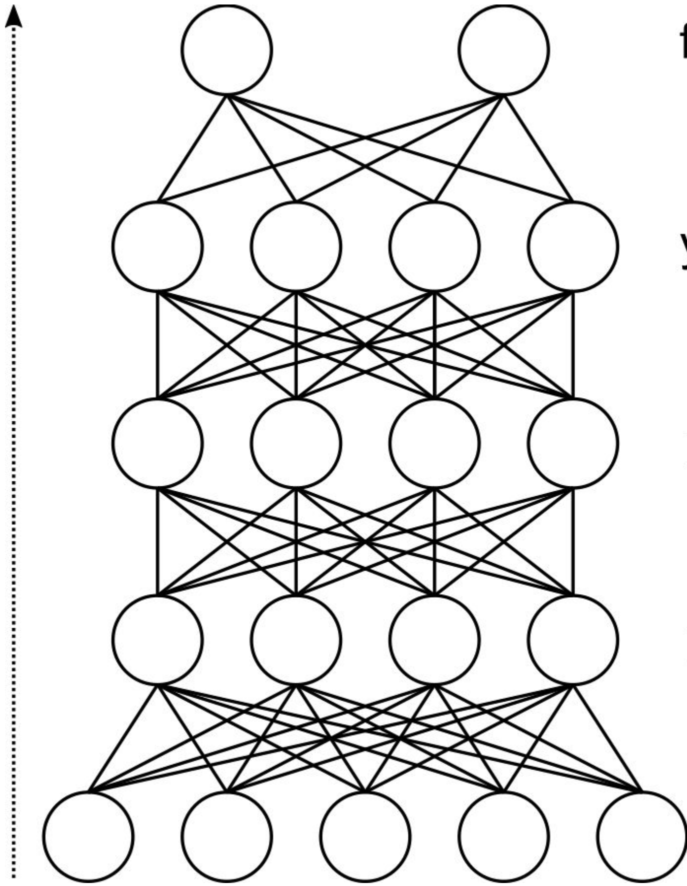
1957
Frank Rosenblatt develops the **Perceptron**, an early neural network enabling pattern recognition based on a two-layer learning network

[Wahrmann, 2017]

1957 Rosenblatt
perceptron



multilayer perceptron



$$f(x) = \varphi_4(W_4 y_3(x))$$

$$y_3(x) = \varphi_3(W_3 y_2(x)) = \varphi_3(W_3 \varphi_2(W_2 \varphi_1(W_1 x)))$$

$$y_2(x) = \varphi_2(W_2 y_1(x)) = \varphi_2(W_2 \varphi(W_1 x))$$

$$y_1(x) = \varphi_1(W_1(x))$$

x

universal approximation theorem

A feed-forward network with a single hidden layer containing a finite number of neurons (that is, a multilayer perceptron) approximates continuous functions.

Math. Control Signals Systems (1989) 2: 303–314

Mathematics of Control, Signals, and Systems

Approximation by Superpositions of a Sigmoidal Function*

G. Cybenko†

Abstract. In this paper we demonstrate that finite linear combinations of compositions of a fixed, univariate function and a set of affine functionals can uniformly approximate any continuous function of *n* real variables with support in the unit hypercube; only mild conditions are imposed on the univariate function. Our results settle an open question about representability in the class of single hidden layer neural networks. In particular, we show that arbitrary decision regions can be arbitrarily well approximated by continuous feedforward neural networks with only a single internal, hidden layer and any continuous sigmoidal nonlinearity. The paper discusses approximation properties of other possible types of nonlinearities that might be implemented by artificial neural networks.

Key words. Neural networks, Approximation, Completeness.

This works exponentially better with increasing depth

training /1

01 initialization

Randomly initialize network weights

02 forward propagation

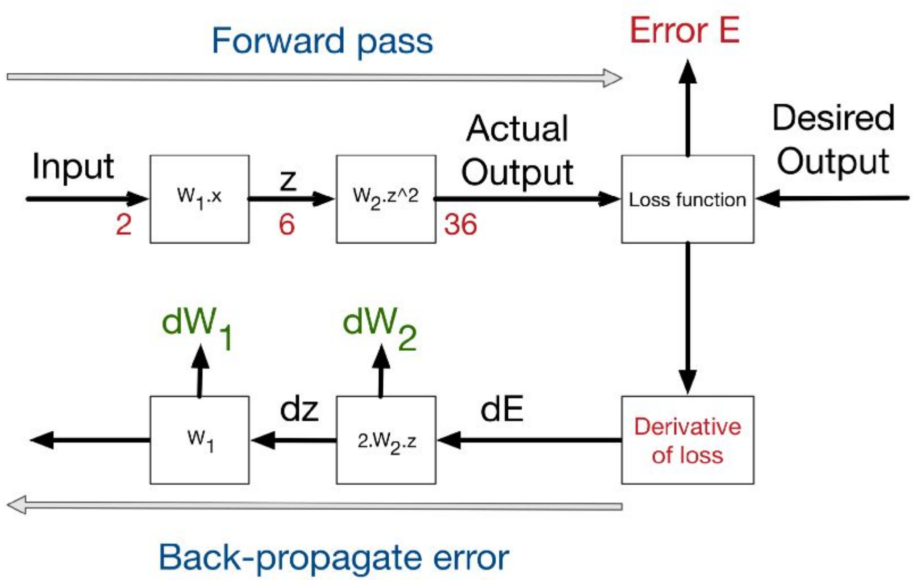
Propagate forward, generate output

03 loss function

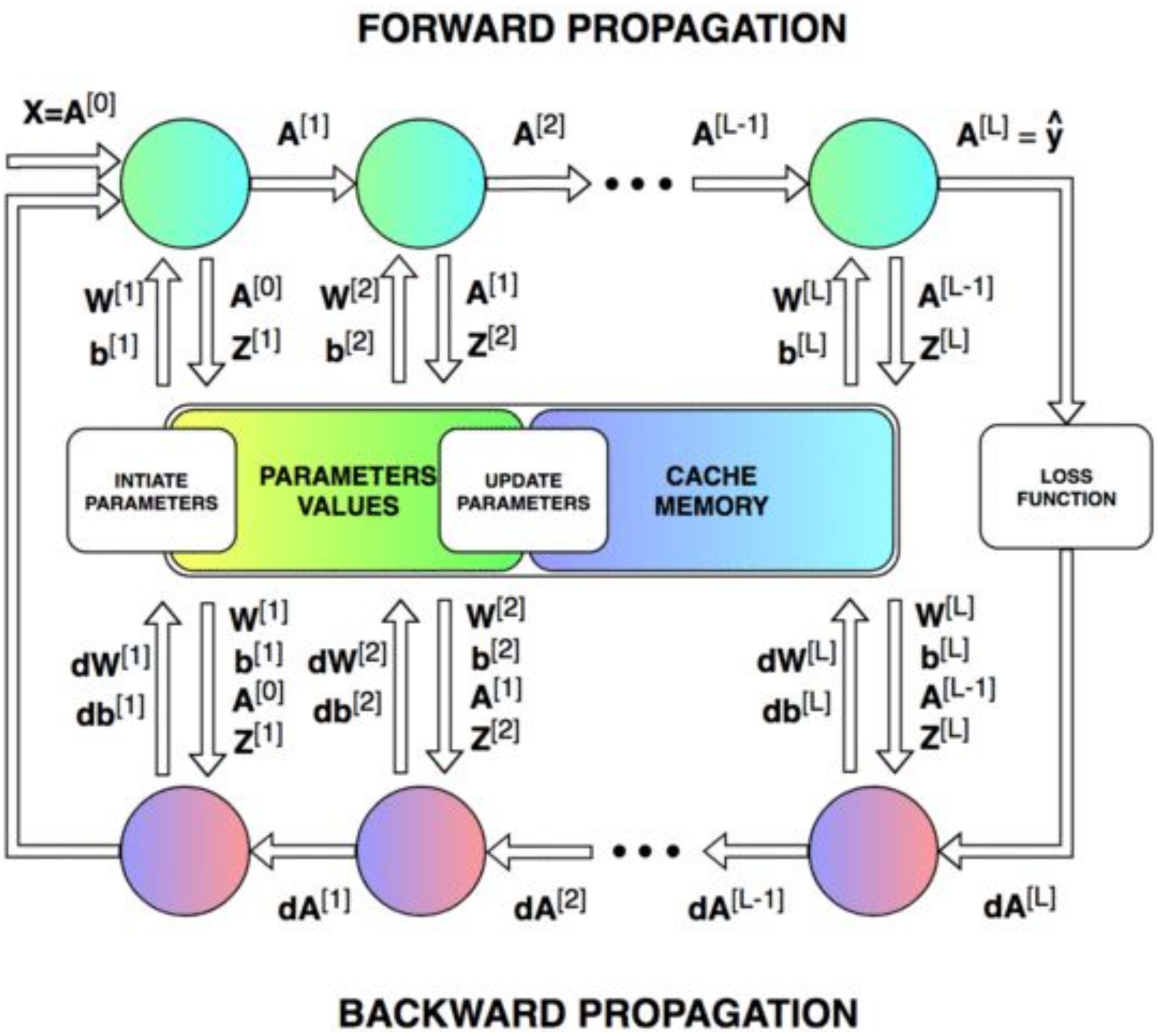
- error metric
- loss ~= (desired output) - (actual output)

04 back-propagation

- generate loss gradient
- propagate back the error from output to input

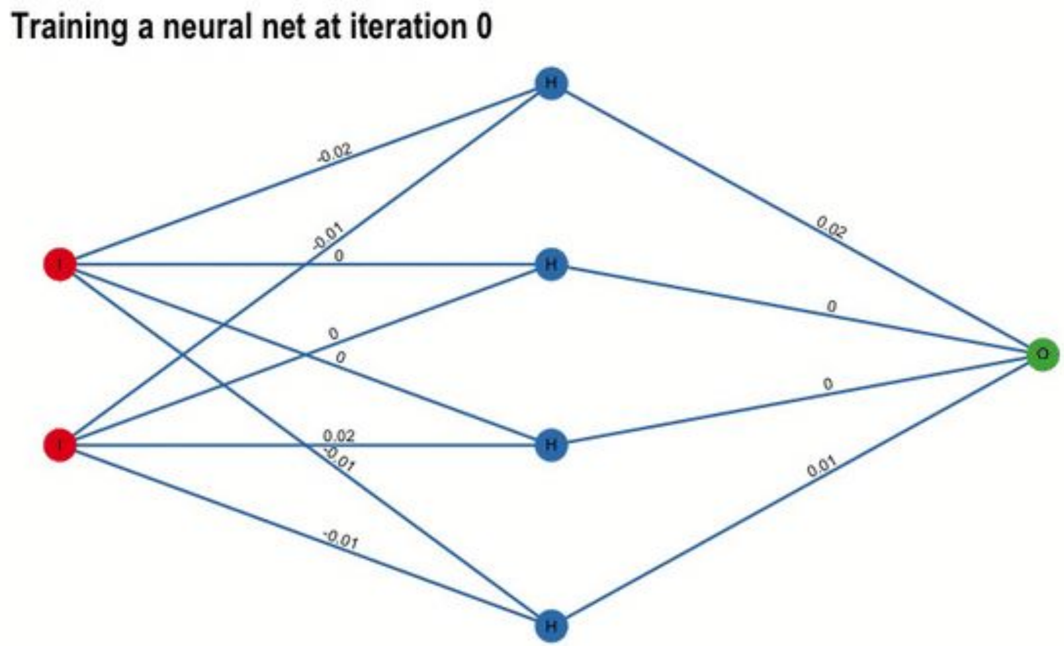


training /2

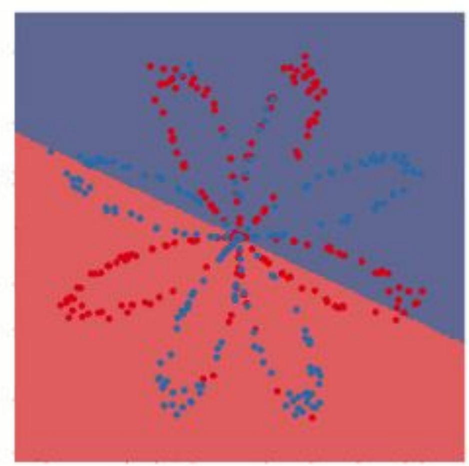


FORWARD PROPAGATION
X=A[0]
A[1]
A[2]
A[L-1]
A[L] = ŷ
W[1] b[1]
A[0] Z[1]
W[2] b[2]
A[1] Z[2]
W[L] b[L]
A[L-1] Z[L]
INTIATE PARAMETERS
PARAMETERS VALUES
UPDATE PARAMETERS
CACHE MEMORY
LOSS FUNCTION
dW[1] db[1]
W[1] b[1] A[0] Z[1]
dW[2] db[2]
W[2] b[2] A[1] Z[2]
dW[L] db[L]
W[L] b[L] A[L-1] Z[L]
dA[1]
dA[2]
dA[L-1]
dA[L]
BACKWARD PROPAGATION

training /3

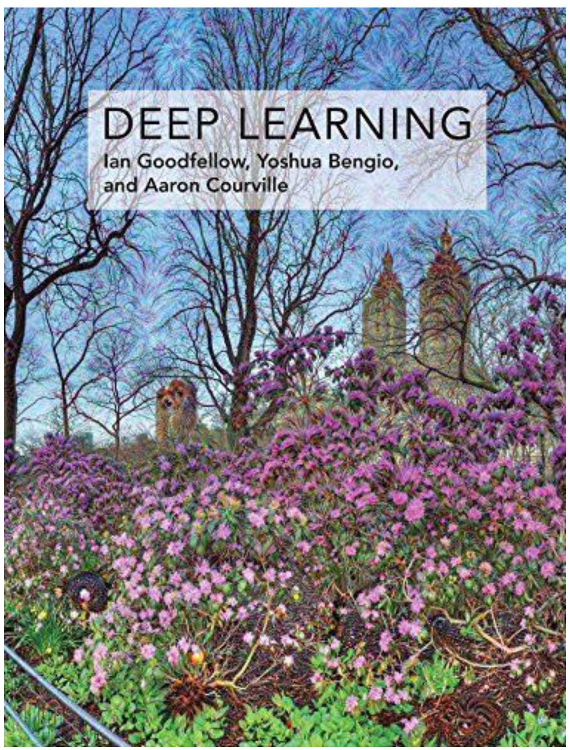


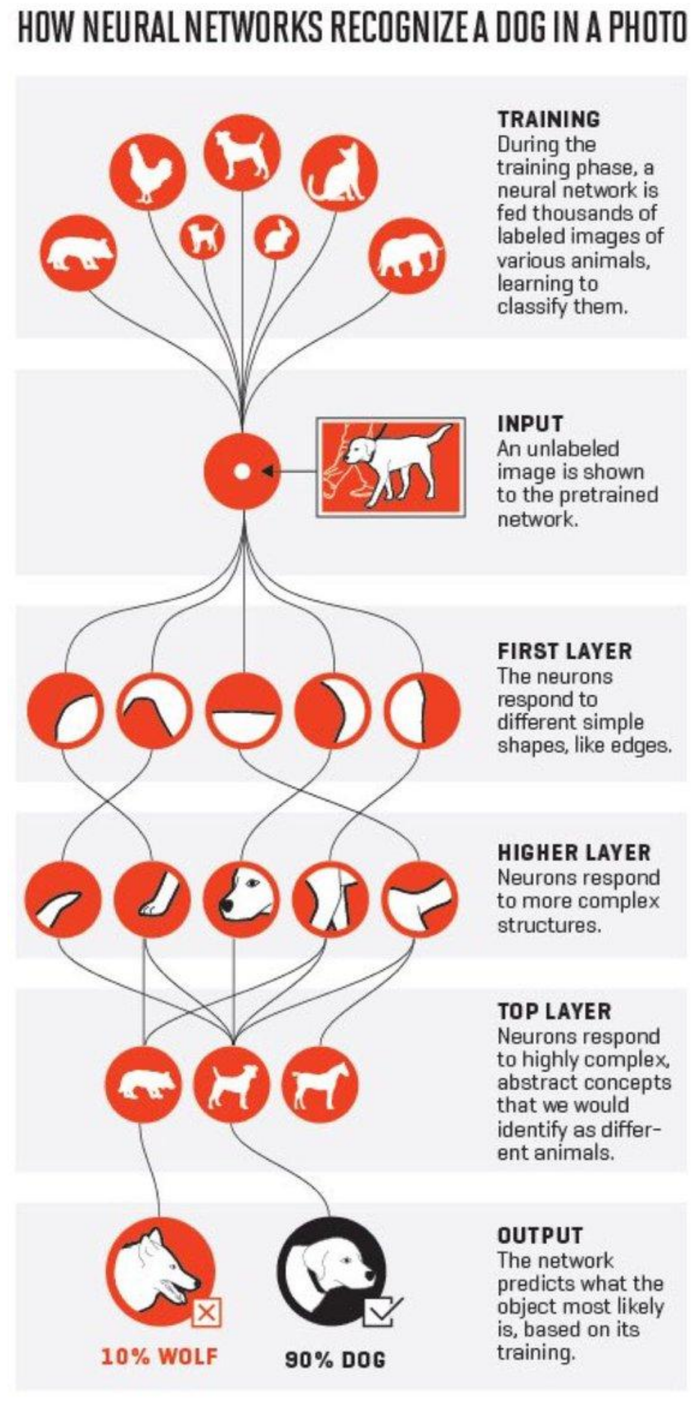
loss (error)



decision boundary

take-home message



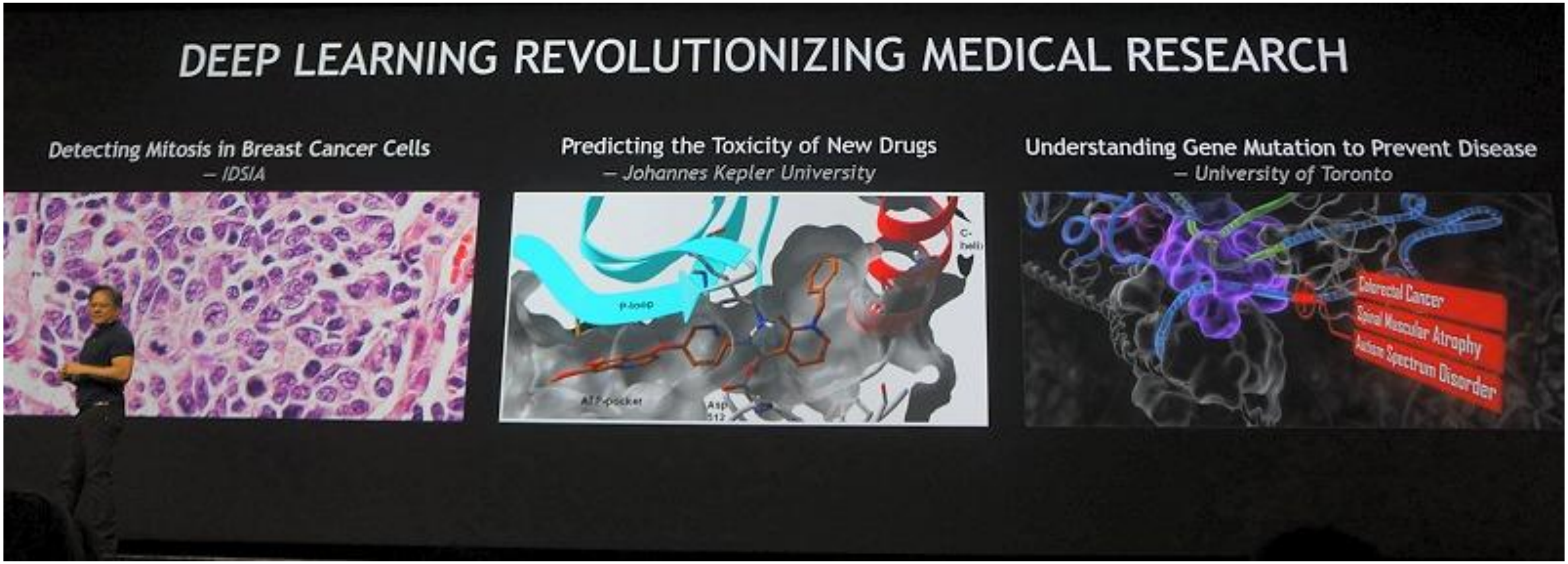


challenges

@teenybiscuit (Karen Zack)

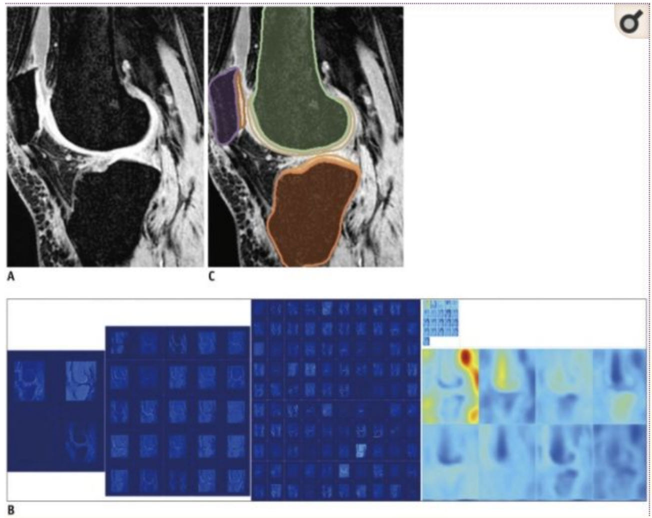
deep learning in biology.

impacting biodomain

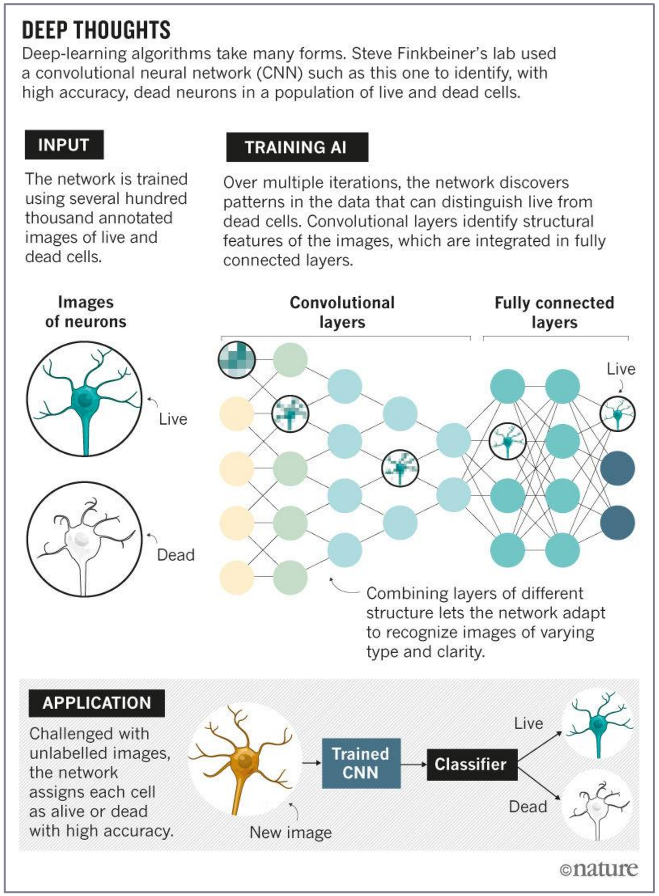


[nvidia, 2018]

imaging classification

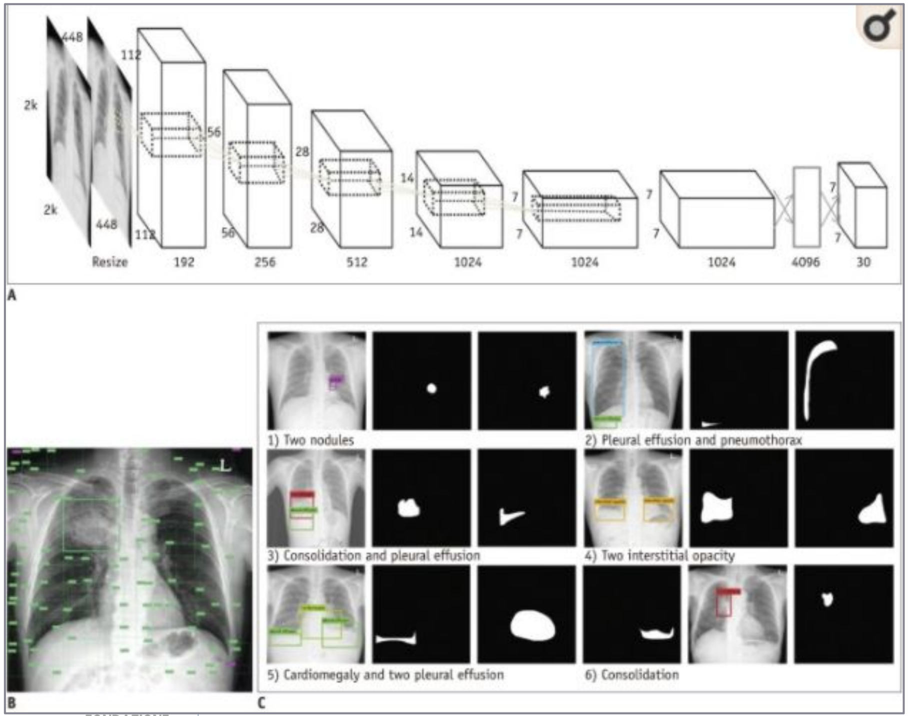


[lee et al, 2017]

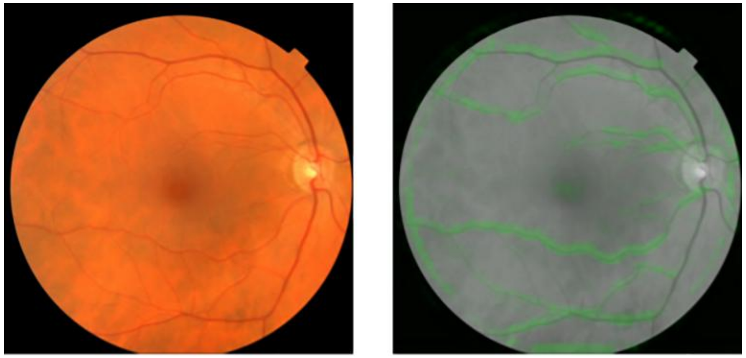


[finkbeiner, 2018]

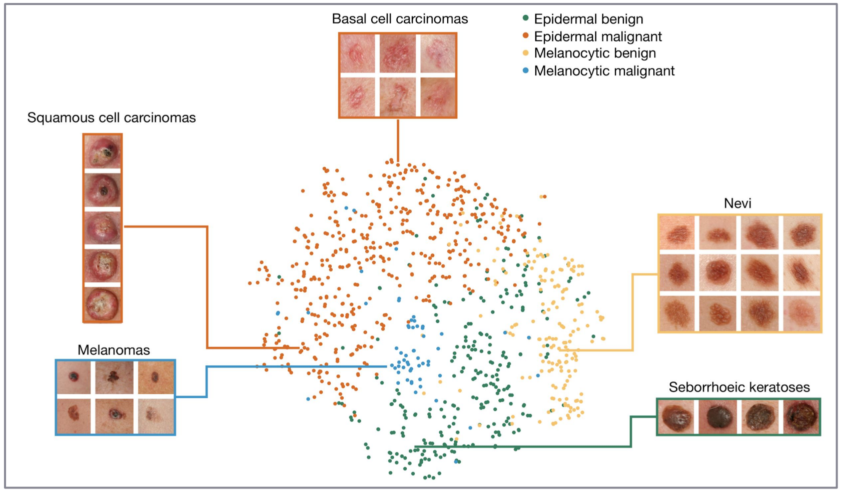
[ren et al, 2015]



[google & stanford, 2018]

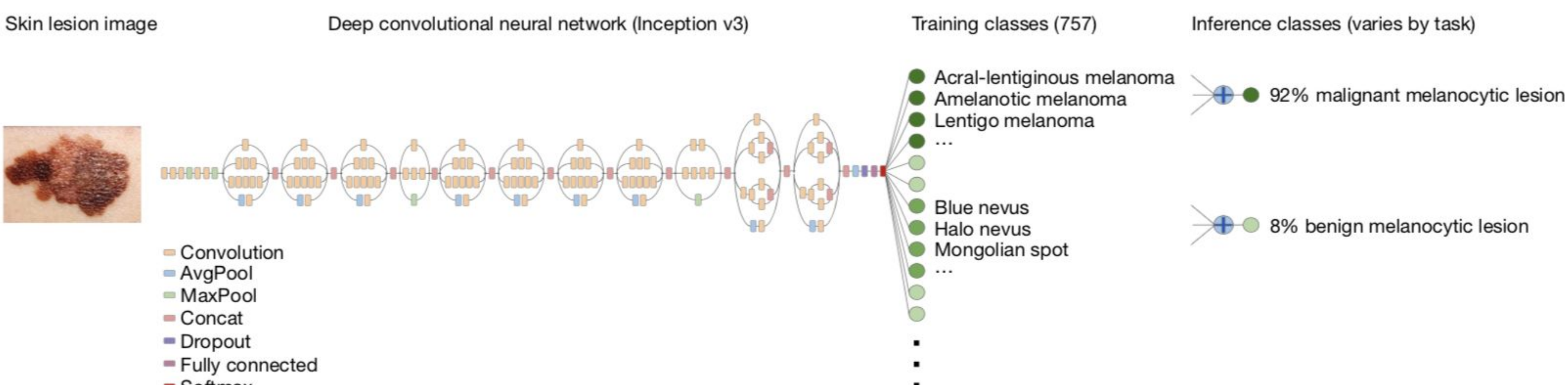


breakthrough




MENU
nature
International journal of science
Altmetric: 2716 Citations: 85 More detail
Letter
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
Nature **542**, 115–118 (02 February 2017)
doi:10.1038/nature21056
Download Citation
Diagnosis Machine learning Skin cancer
Received: 28 June 2016
Accepted: 14 December 2016
Published: 25 January 2017
Corrigendum: 28 June 2017

129,450 images, 2,032 diseases: DL ~ 21 board-certified dermatologists



other clinical apps

PNAS

Predicting cancer outcomes from histology and genomics using convolutional networks

Pooya Mobadersany[a], Safoora Yousefi[a], Mohamed Amgad[a], David A. Gutman[b], Jill S. Barnholtz-Sloan[c], José E. Velázquez Vega[d], Daniel J. Brat[e], and Lee A. D. Cooper[a,f,g,1]

[a]Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA 30322; [b]Department of Neurology, Emory University School of Medicine, Atlanta, GA 30322; [c]Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH 44106; [d]Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, GA 30322; [e]Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611; [f]Winship Cancer Institute, Emory University, Atlanta, GA 30322; and [g]Department of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA 30322

EXPERT REVIEW OF PRECISION MEDICINE AND DRUG DEVELOPMENT, 2017
VOL. 2, NO. 5, 239–241
https://doi.org/10.1080/23808993.2017.1380516

EDITORIAL

The role of artificial intelligence in precision medicine

Bertalan Mesko

The Medical Futurist Institute; Department of Behavioral Sciences, Semmelweis University, Budapest, Hungary

THE LANCET

Online First | Current Issue | All Issues | Special Issues | Multimedia | Information for Authors

All Content | Search | Advanced Search

< Previous Article | Volume 391, No. 10127, p1260, 31 March 2018 | Next Article >

Correspondence

Artificial intelligence in health care: enabling informed care

Lionel Tarassenko, Peter Watkinson

Published: 31 March 2018

SCIENTIFIC REPORTS

Article | OPEN

Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models

Safoora Yousefi, Fatemeh Amrollahi, Mohamed Amgad, Chengliang Dong, Joshua E. Lewis, Congzheng Song, David A. Gutman, Sameer H. Halani, Jose Enrique Velazquez Vega, Daniel J. Brat & Lee A. D. Cooper

Scientific Reports **7**, Article number: 11707 (2017)
doi:10.1038/s41598-017-11817-6

Received: 23 May 2017
Accepted: 30 August 2017
Published online: 15 September 2017

SCIENTIFIC REPORTS

Article | OPEN

Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records

Riccardo Miotto, Li Li, Brian A. Kidd & Joel T. Dudley

Scientific Reports **6**, Article number: 26094 (2016)
doi:10.1038/srep26094
Download Citation

Received: 28 January 2016
Accepted: 27 April 2016
Published: 17 May 2016

other
clinical apps

New Results

Opportunities And Obstacles For Deep Learning In Biology And Medicine

Travers Ching, Daniel S. Himmelstein, Brett K. Beaulieu-Jones, Alexandr A. Kalinin, Brian T. Do, Gregory P. Way, Enrico Ferrero, Paul-Michael Agapow, Wei Xie, Gail L. Rosen, Benjamin J. Lengerich, Johnny Israeli, Jack Lanchantin, Stephen Woloszynek, Anne E. Carpenter, Avanti Shrikumar, Jinbo Xu, Evan M. Cofer, David J. Harris, Dave DeCaprio, Yanjun Qi, Anshul Kundaje, Yifan Peng, Laura K. Wiley, Marwin H. S. Segler, Anthony Gitter, Casey S. Greene

doi: https://doi.org/10.1101/142760

Conclusions

Deep learning-based methods now match or surpass the previous state of the art in a diverse array of tasks in patient and disease categorization, fundamental biological study, genomics, and treatment development. Returning to our central question: given this rapid progress, has deep learning transformed the study of human disease? Though the answer is highly

deep genomics

CHANGING THE COURSE OF GENOMIC MEDICINE

Today, machine learning and experimental biology are advancing at an exponential pace. Deep Genomics is where these disciplines meet. Our systems predict the molecular effects of genetic variation, opening a new and exciting path to discovery for disease diagnostics and therapies.

LEARN MORE

DEEP LEARNING GENOMICS PRECISION MEDICINE

Briefings in Bioinformatics

Deep learning in bioinformatics

Seonwoo Min, Byunghan Lee, Sungroh Yoon

Brief Bioinform bbw068. **DOI:** https://doi.org/10.1093/bib/bbw068
Published: 25 July 2016 **Article history ▾**

Although deep learning holds promise, it is not a silver bullet and cannot provide great results in ad hoc bioinformatics applications. There remain many potential challenges, including limited or imbalanced data, interpretation of deep learning results, and selection of an appropriate architecture and hyperparameters. Furthermore, to fully exploit the capabilities of

608 IEEE TRANSACTIONS ON NANOBIOSCIENCE, VOL. 14, NO. 6, SEPTEMBER 2015

Multi-Layer and Recursive Neural Networks for Metagenomic Classification

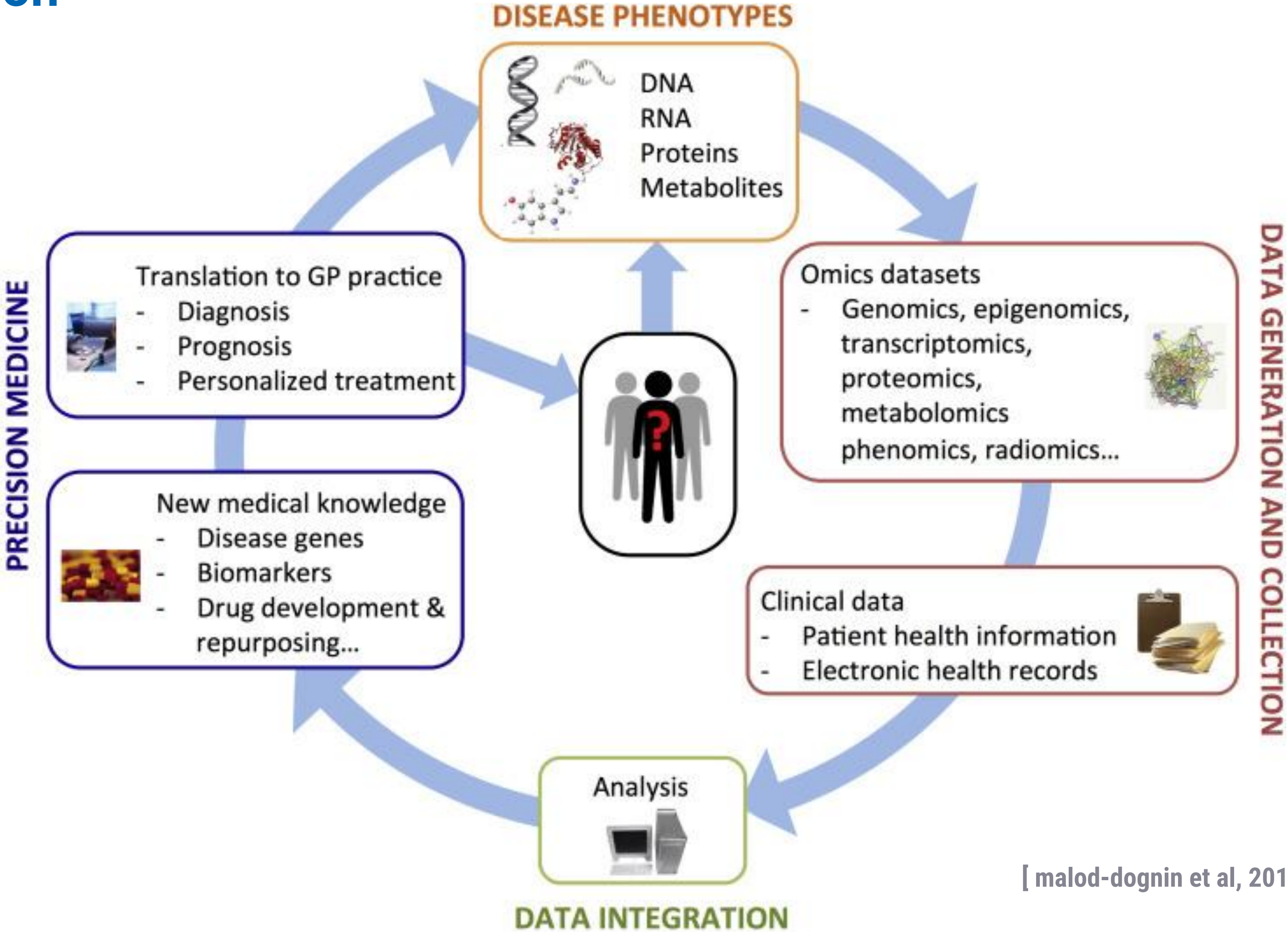
Gregory Ditzler*, *Member, IEEE*, Robi Polikar, *Senior Member, IEEE*, and Gail Rosen, *Senior Member, IEEE*

V. DISCUSSION

The experiments discussed in the previous section demonstrated that: i) the deep learning approaches are not superior, at least on the data sets we evaluated, and ii) traditional MLPNNs are quite competitive with the RFCs, and in general perform better. However, none of the classifiers—deep or shallow—uniformly performs better than the RFCs across different experiments. The performance of the deep learning approaches may be improved upon with data sets that are much larger. It appears that—at least based on accuracy alone—the deep learning approaches may not be suitable for metagenomic applications. Accuracy, however, is not the only figure of merit.

FONDAZIONE BRUNO KESSLER ICT CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGY

vision



[malod-dognin et al, 2018]

thank you

Acks

Giuseppe Jurman
Cesare Furlanello