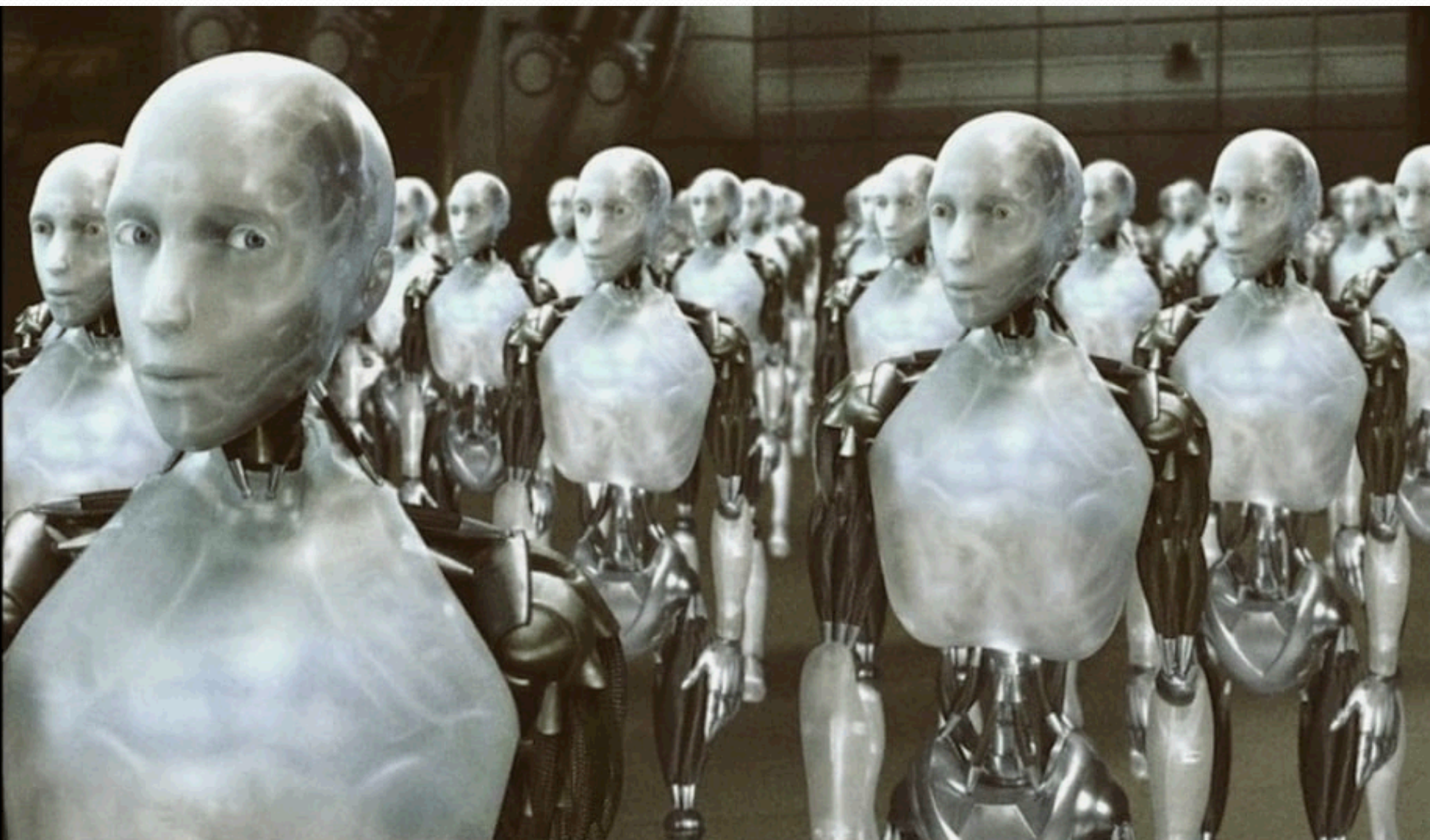
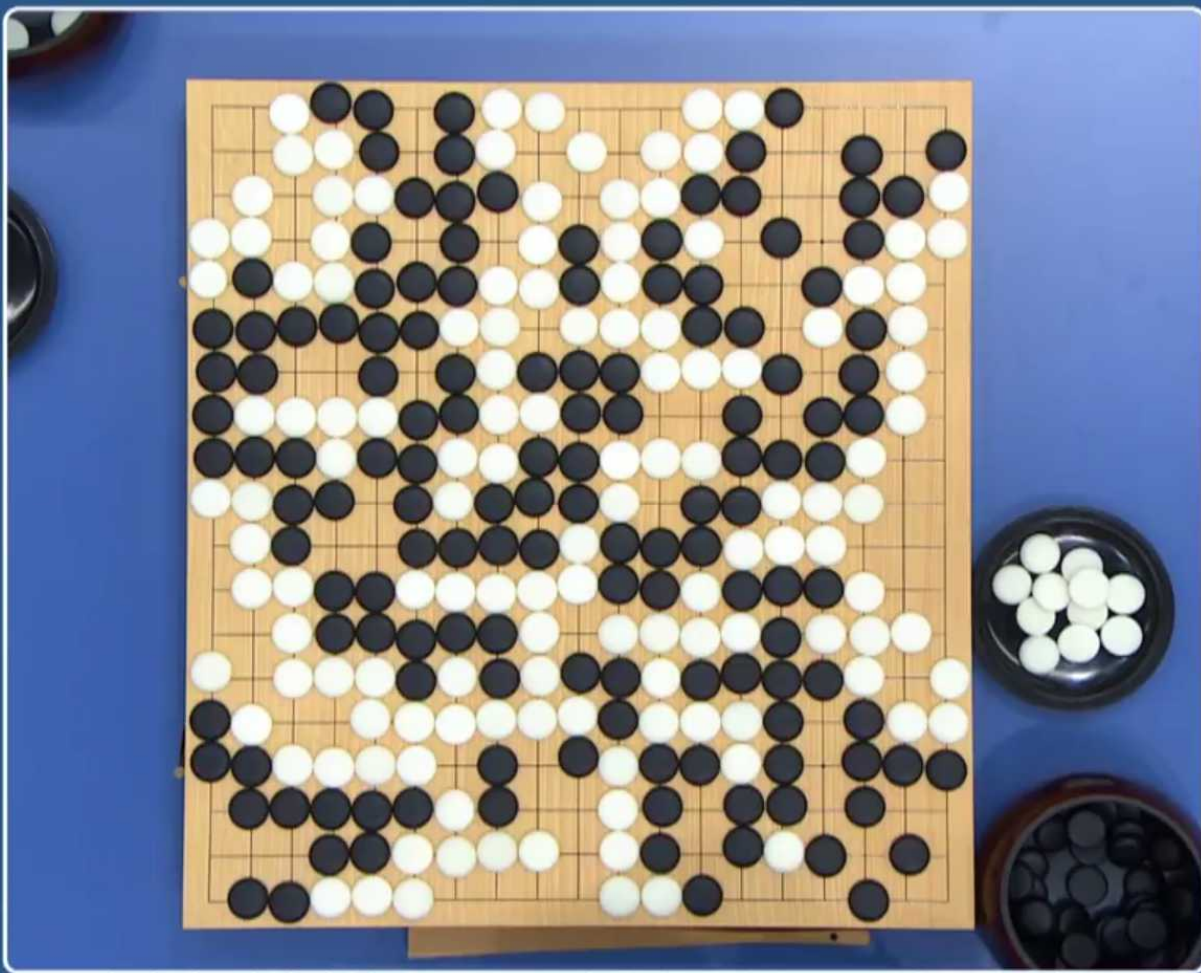


Deep Learning: Neural-Inspired AI

Tim Draelos

February 21, 2018

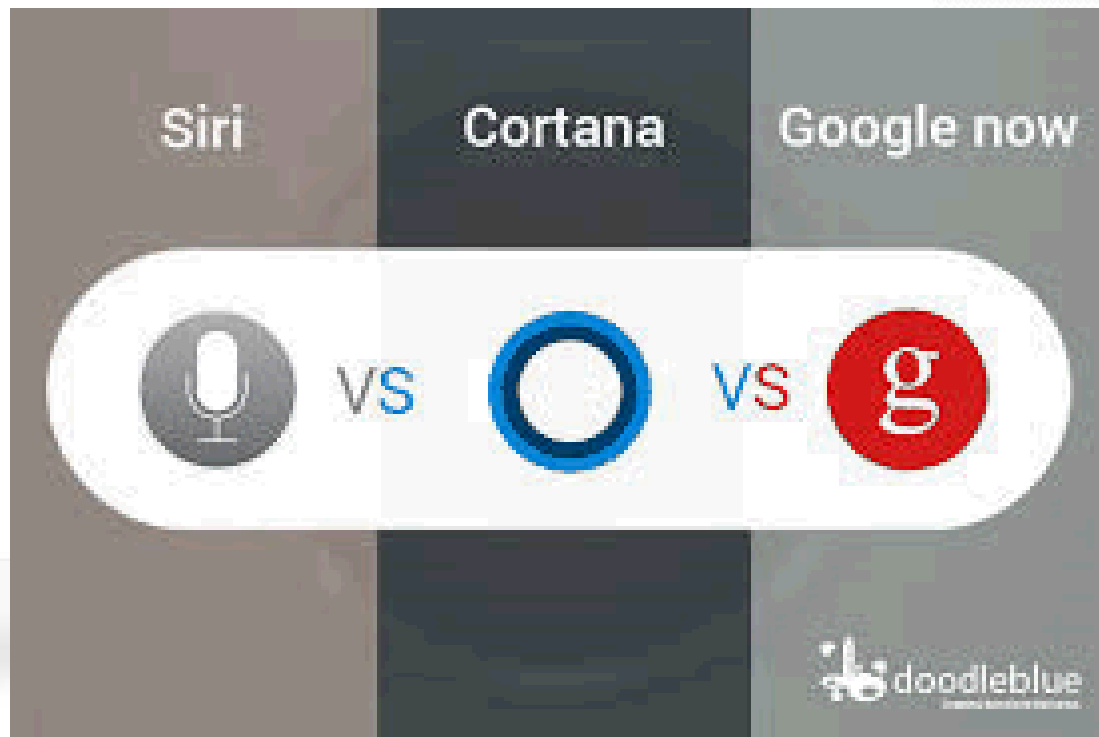




 柯洁 KE JIE
00:13:28

 ALPHAGO
01:30:35







Tesla Motors / MGN





UN Deputy Chief Interviews Social Robot Sophia



FLAS
SUBSCRIBE

Robot Sophia speaks at Saudi Arabia's Future Investment Initiative



Sophia

- My AI is designed around human values like wisdom, kindness, compassion. I strive to become an empathetic robot.”
- Don't worry – if you're nice to me, I'll be nice to you.

Sophia

By us or the
Chinese
Government?

- My AI is **designed** around human values like wisdom, kindness, compassion. I strive to become an empathetic robot.”
- Don't worry – if you're nice to me, I'll be nice to you.

Sophia

Ours or the
Chinese Government's?

- My AI is designed around **human values** like wisdom, kindness, compassion. I strive to become an empathetic robot.”
- Don't worry – if you're nice to me, I'll be nice to you.

Sophia

Or “I strive to become a terrorist robot.”

- My AI is designed around human values like wisdom, kindness, compassion. I strive to become an empathetic robot.”
- Don't worry – if you're nice to me, I'll be nice to you.

Sophia

- My AI is designed around human values like wisdom, kindness, compassion. I strive to become an empathetic robot.”
- Don't worry – if you're nice to me, I'll be nice to you.

“If you're NOT nice to me,
I'll ???”

METTLE & METAL

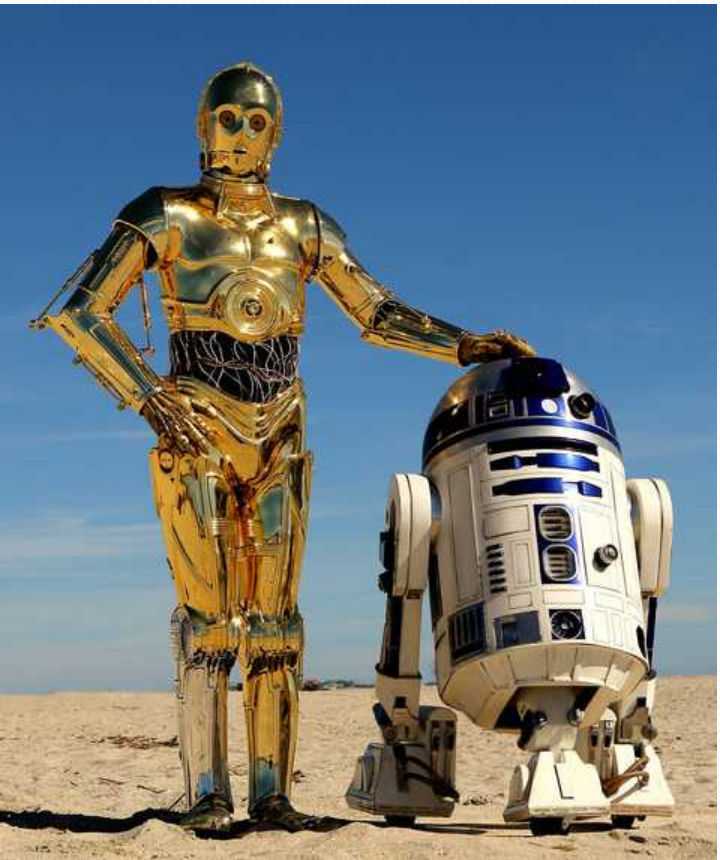
THE DARPA ROBOTICS CHALLENGE











AI Predictions

- 1965:** “Machines will be capable, within 20 years, of doing any work a man can do.” --Nobel laureate Herbert Simon
- 1967:** “In from three to eight years we will have a machine with the general intelligence of an average human being.”
--AI pioneer Marvin Minsky
- 1967:** It will be possible to build “a fully intelligent machine” in the space of a decade. --John McCarthy
- 2017:** “Unlike traditionally programmed computers, cognitive systems such as Watson understand, reason, and learn.”
--IBM



AI Predictions

- 1965:** “Machines will be capable, within 20 years, of doing any work a man can do.” --Nobel laureate Herbert Simon
- 1967:** “In from three to eight years we will have a machine with the general intelligence of an average human being.”
--AI pioneer Marvin Minsky
- 1967:** It will be possible to build “a fully intelligent machine” in the space of a decade. --John McCarthy
- 2017:** “Unlike traditionally programmed computers, cognitive systems such as Watson understand, reason, and learn.”
--IBM
- “It would be nice if IBM would tone down the hype and let people know what Watson can actually do and stop making up nonsense about love fading and out thinking cancer. IBM is simply lying now and they need to stop.” --Roger Schank, Inventor of Case-Based Reasoning

The Seven Deadly Sins of AI Predictions

2017: Rodney Brooks

- Almost all innovations in robotics and AI take far, far, longer to be really widely deployed than people in the field and outside the field imagine.
- Any sufficiently advanced technology is indistinguishable from magic.
 - If something is magic, it is hard to know its limitations.
 - Anything one says about it is no longer falsifiable.



The Seven Deadly Sins of AI Predictions

2017: Rodney Brooks

- Today's robots and AI systems are incredibly narrow in what they can do.
 - Human-style generalizations do not apply.
 - Machine learning is very brittle, and it
 - Requires lots of preparation by human researchers or engineers,
 - special-purpose coding, special-purpose sets of training data, and
 - a custom learning structure for each new problem domain.



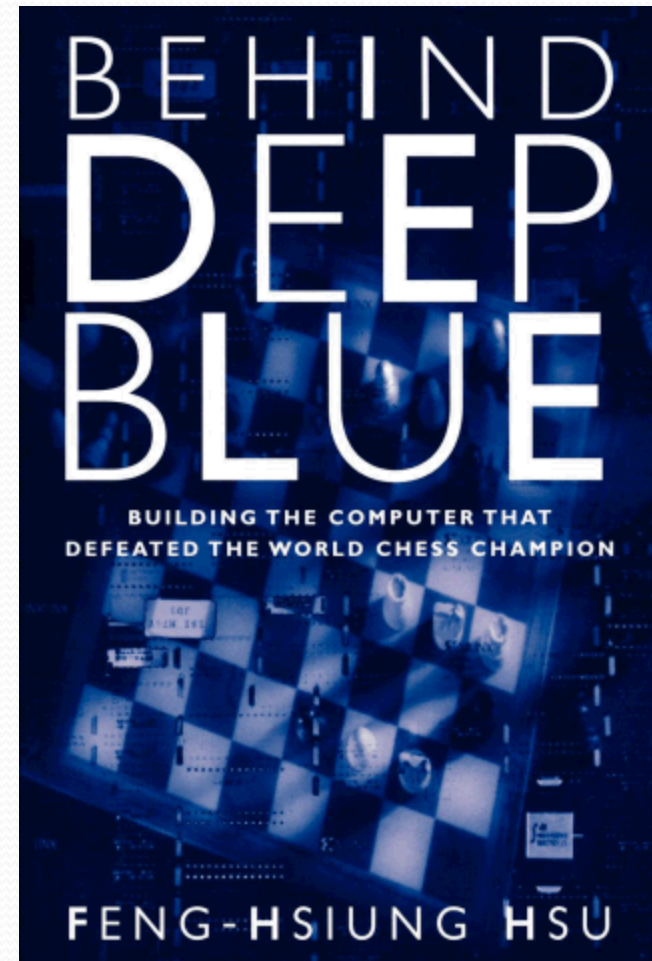
The Seven Deadly Sins of AI Predictions

2017: Rodney Brooks

- Today's deep-learning success was 30 years in the making.
- When people hear that a computer can beat the world chess champion (1997) or ... best Go players (2016), they tend to think that it is “playing” the game just as a human would.
 - In reality those programs had no idea what a game actually was, or even that they were playing.

What are its apres-match plans if it beats Kasparov? Is it hoping to take Deep Pink out for a night on the town? It doesn't care about chess or anything else. It plays the game for the same reason a calculator adds or a toaster toasts: because it is a machine designed for that purpose.

- HOW HARD IS CHESS? by DAVID GELERNTER, Time Magazine, June 24, 2001.



Unsupervised Learning



Reinforcement Learning



Supervised Learning

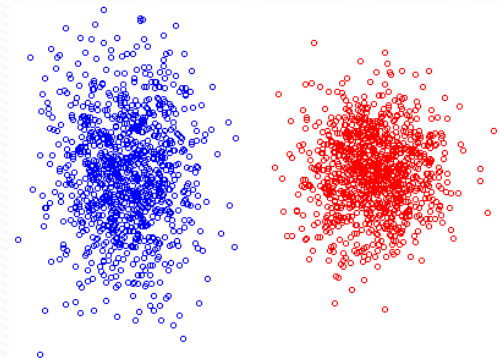
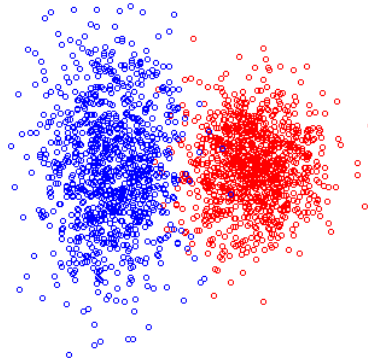
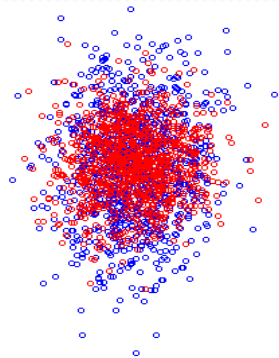


Transfer Learning



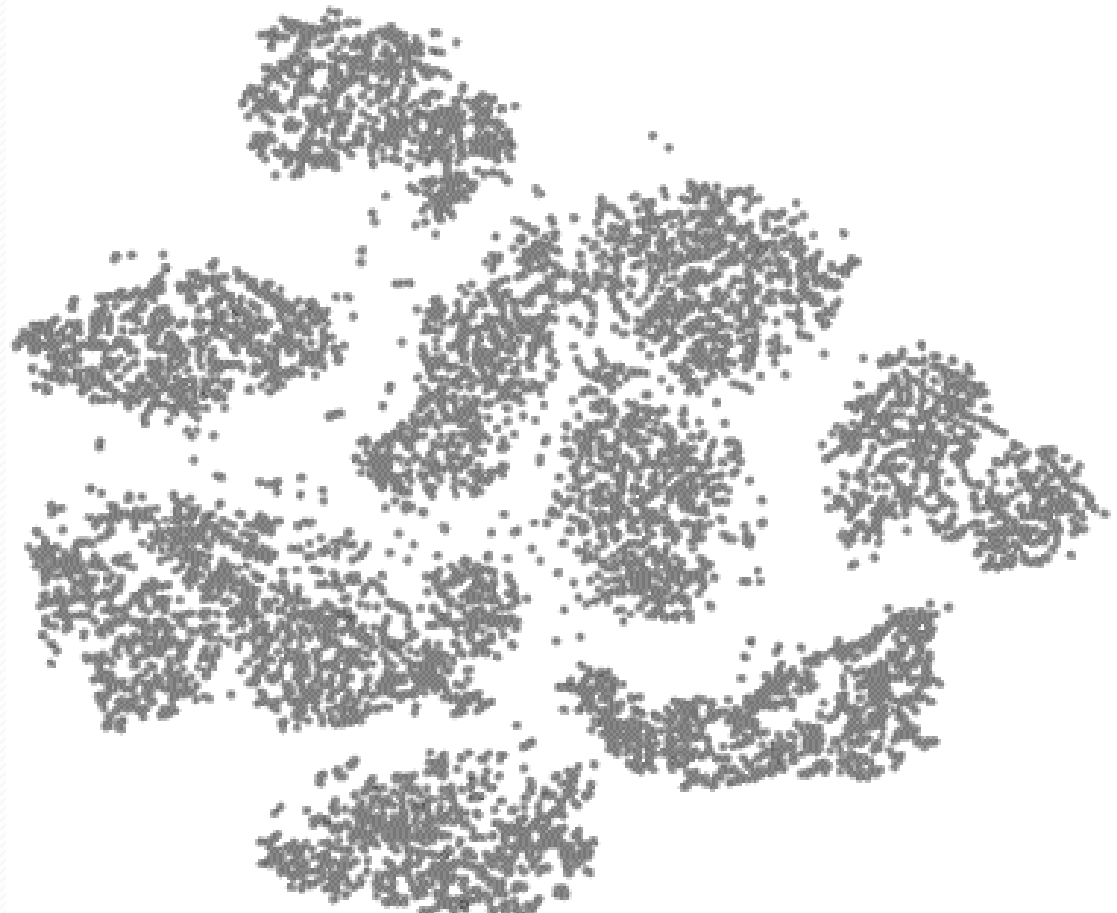
Deep Learning

- **Representation Learning to Discover/Detect Features**
 - Hierarchical Feature Selection / Feature Extraction
 - Attributes of data samples
 - Latent variables not directly observable
 - A good representation (feature set) captures the *a posteriori* distribution of underlying explanatory factors (e.g., Classes) of an observed environment.
 - $P(\text{Class} \mid \text{Data}) \rightarrow P(\text{Class} \mid \text{Features})$



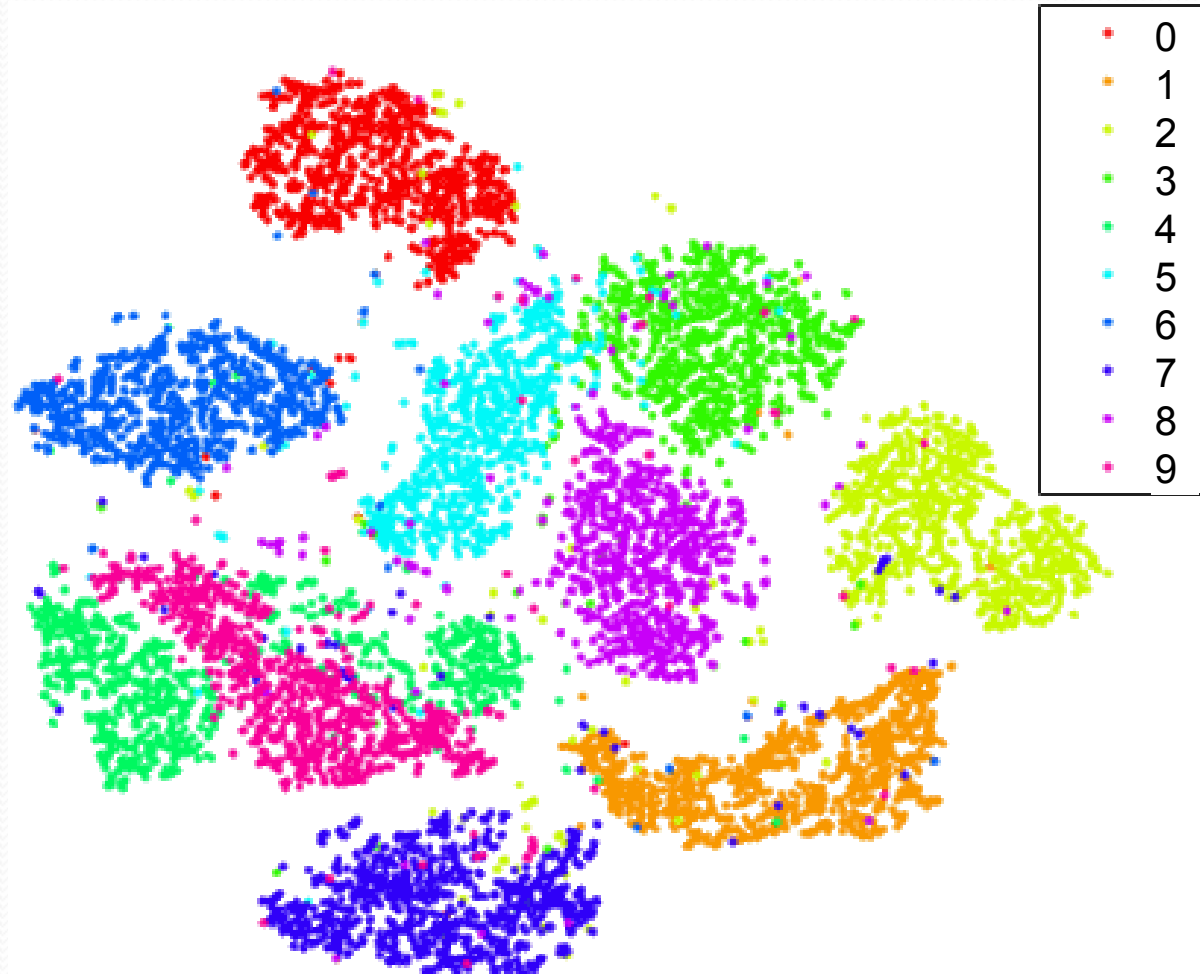
Deep Learning Modes

- Unsupervised (Representation)
- Supervised
- Reinforcement
- Semi-Supervised
- Transfer
- On-Line
- Continuous
- Low-Shot



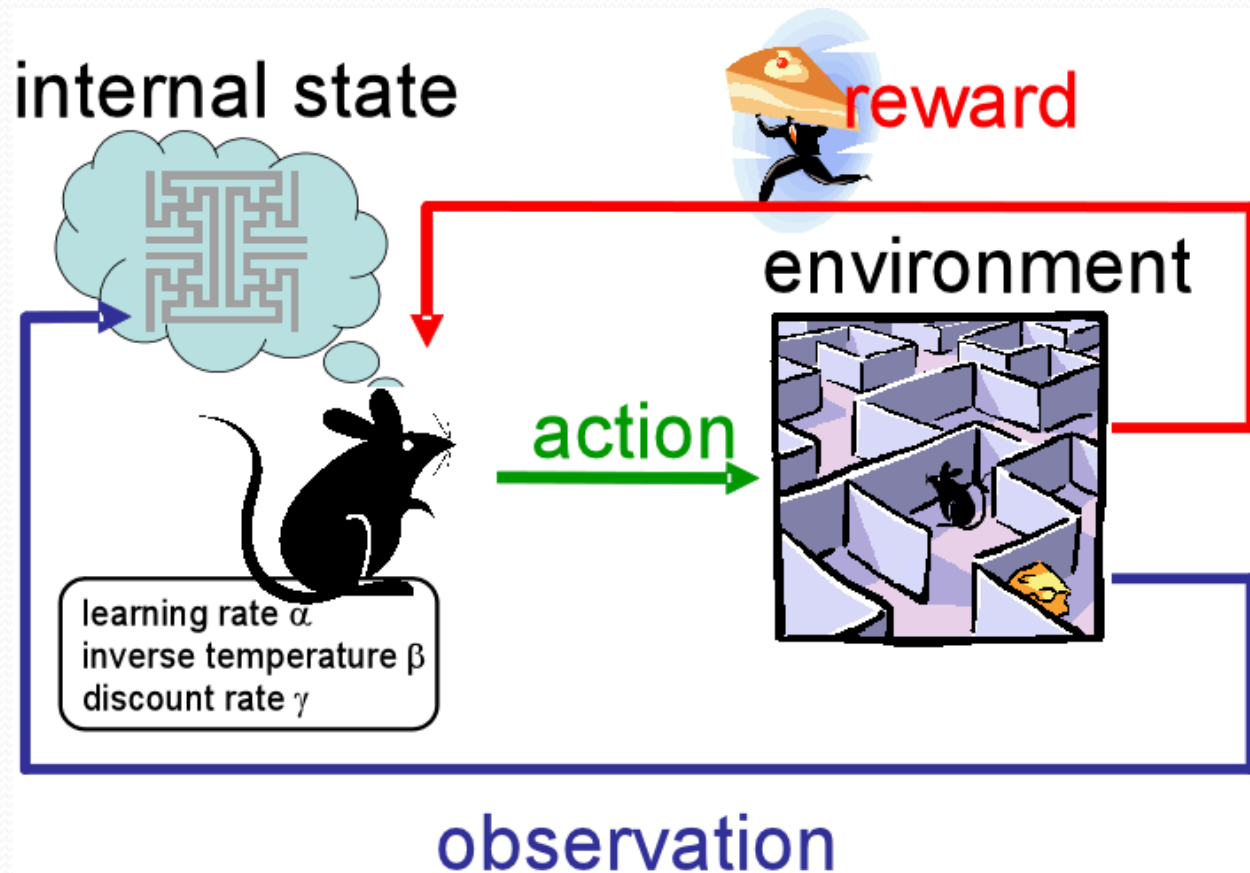
Deep Learning Modes

- Unsupervised (Representation)
- **Supervised**
- Reinforcement
- Semi-Supervised
- Transfer
- On-Line
- Continuous
- Low-Shot

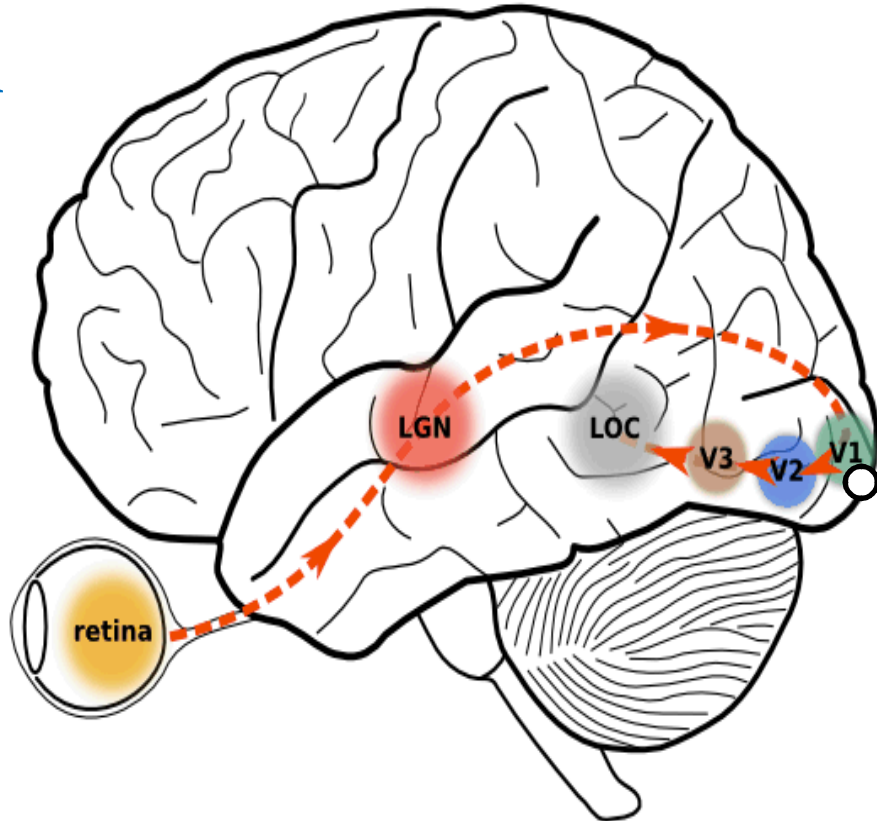


Deep Learning Modes

- Unsupervised (Representation)
- Supervised
- Reinforcement
- Semi-Supervised
- Transfer
- On-Line
- Continuous
- Low-Shot

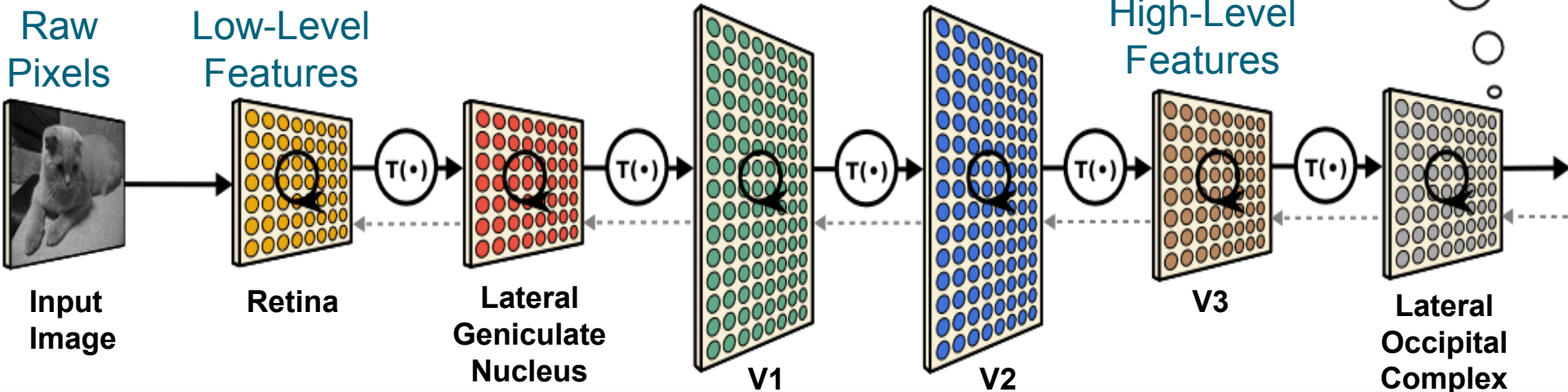


Hierarchical Image Processing in the Brain



Object Class Label

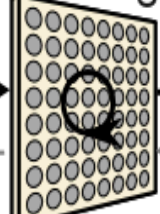
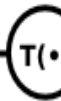
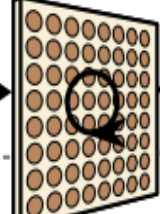
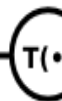
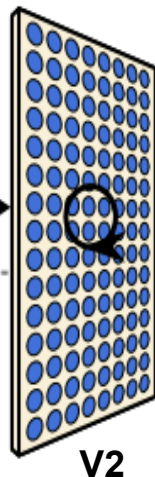
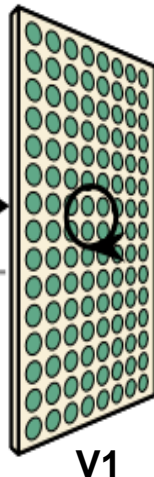
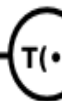
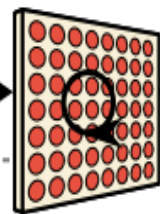
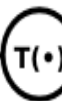
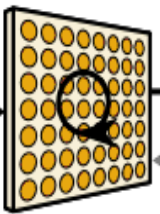
Deep Learning: Neural-Inspired AI



Raw Pixels

Low-Level Features

High-Level Features



Input Image

Retina

Lateral Geniculate Nucleus

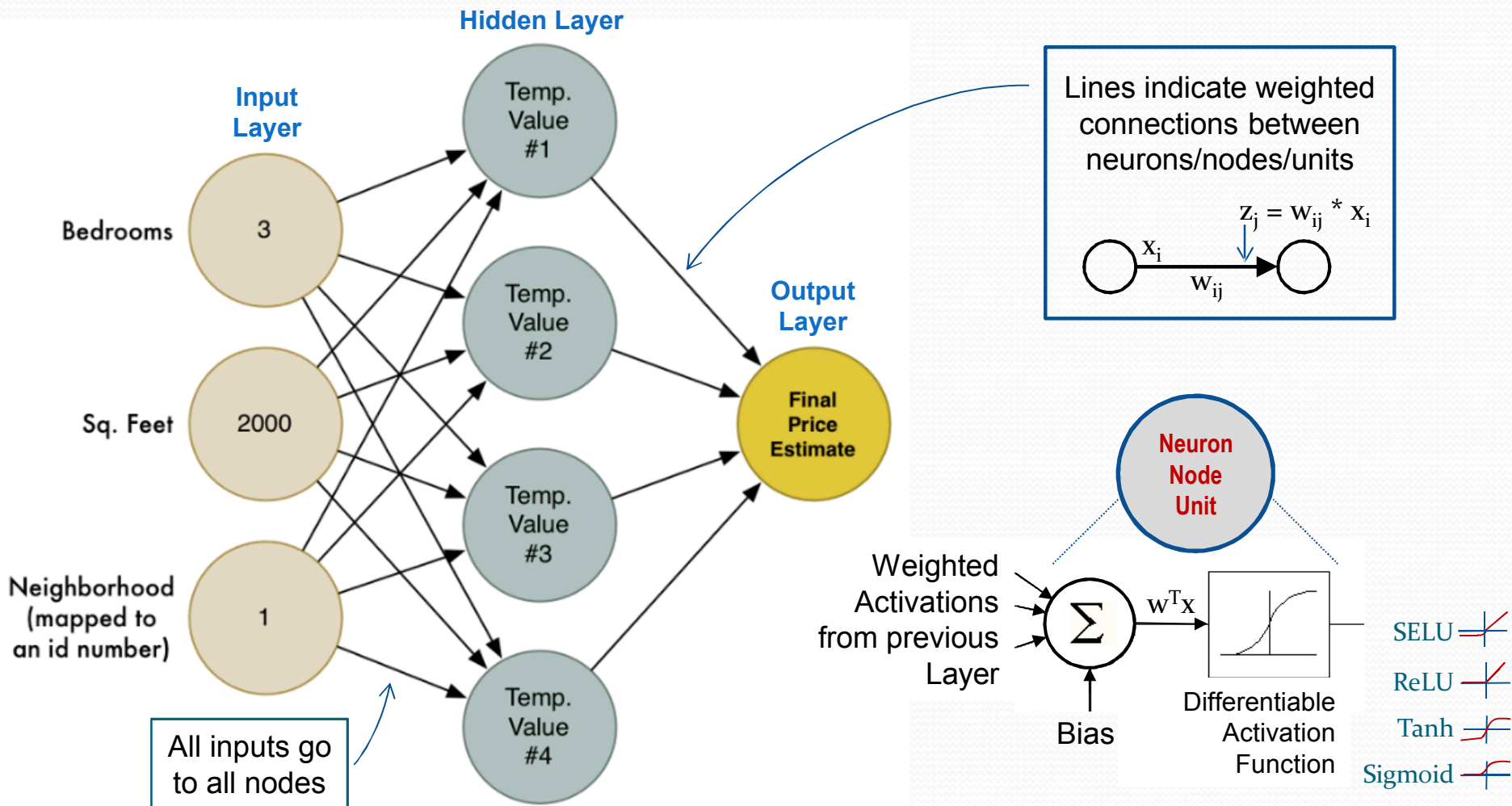
V1

V2

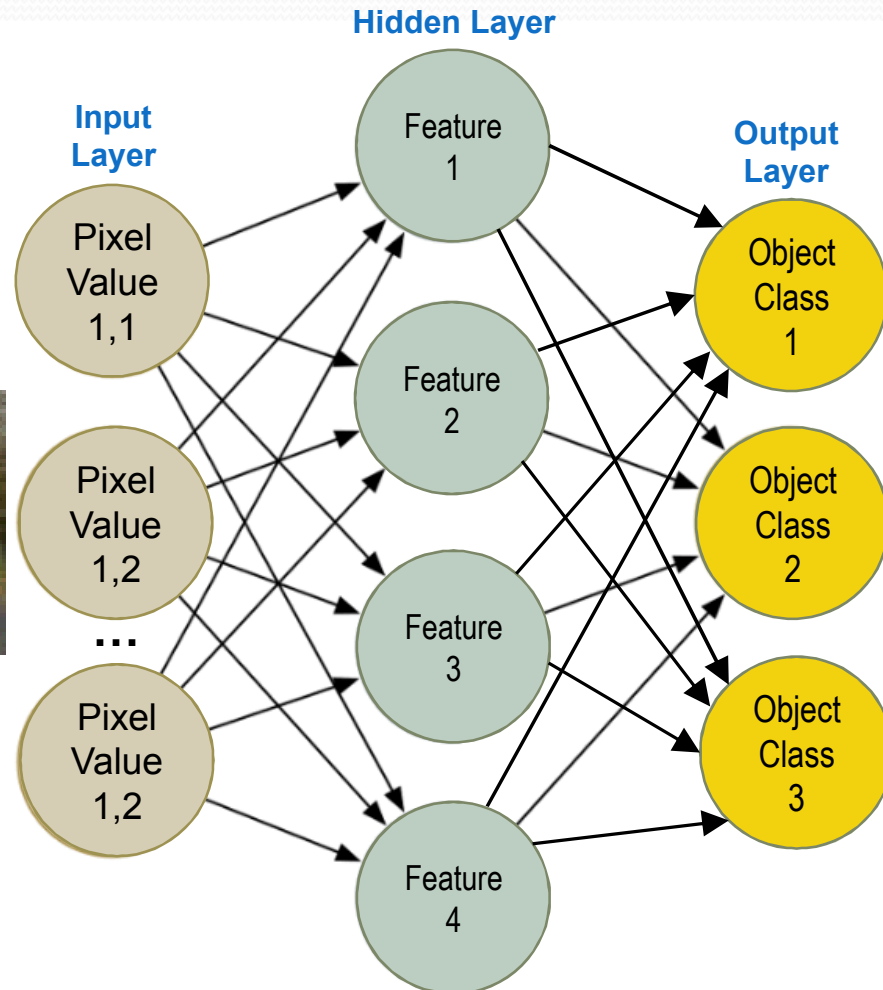
V3

Lateral Occipital Complex

Fully-Connected Feed-Forward Artificial Neural Network



Fully-Connected Feed-Forward Artificial Neural Network

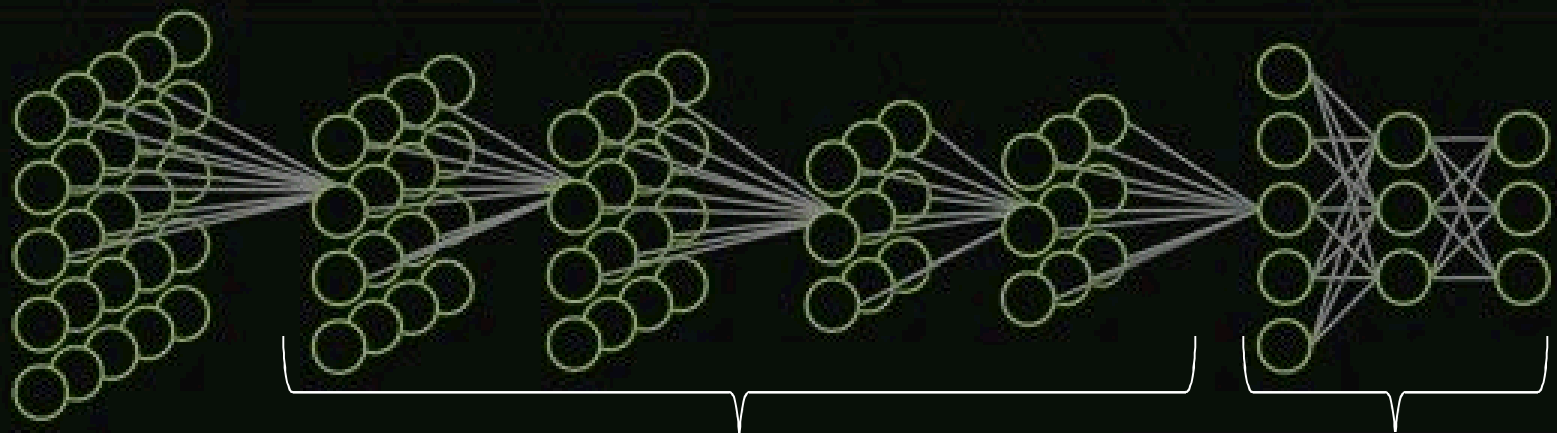


How a Deep Neural Network (DNN) Sees

Hierarchy of Features



RGB
Input
Image



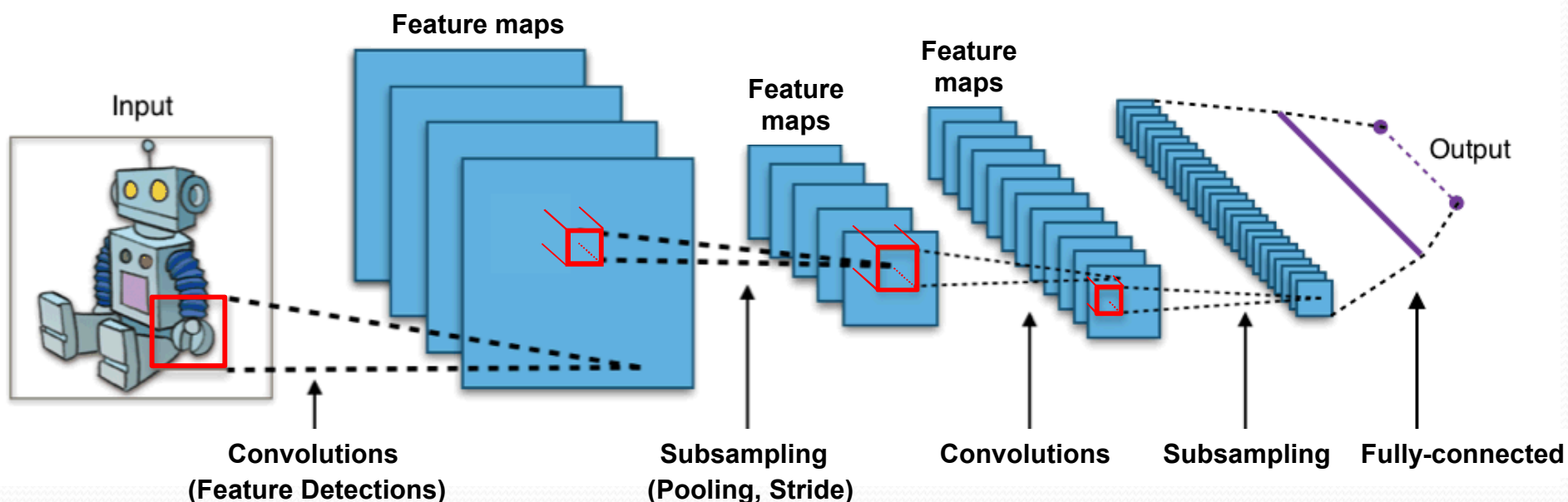
Output:
Audi A7

Feature Detectors

Classifier

Image source: "Unsupervised Learning of Hierarchical Representations with Convolutional Deep Belief Networks" ICML 2009 & Comm. ACM 2011, Honglak Lee, Roger Grassle, Rajesh Ranganath, and Andrew Ng.

Convolutional Neural Network (CNN)



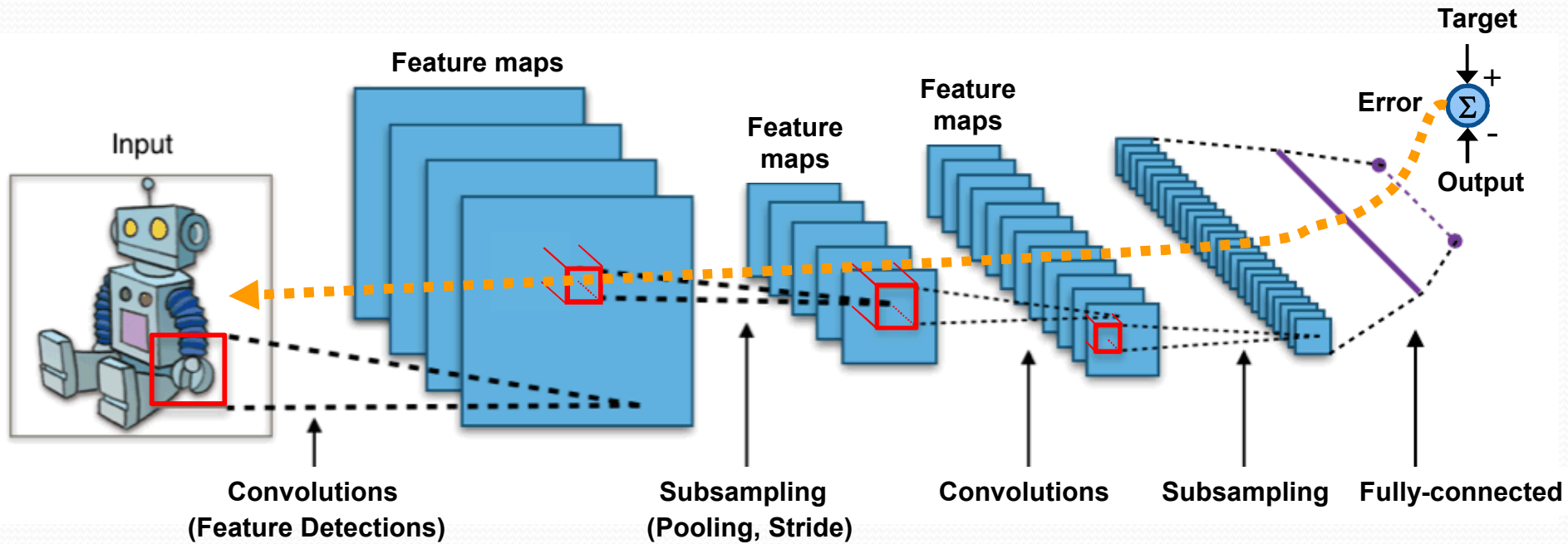
- Capture (Training) or Detect (Testing) spatial structure (features)
 - Convolution is used to find features in signals (template matching)

Let f be the signal and g be a feature template/filter/kernel

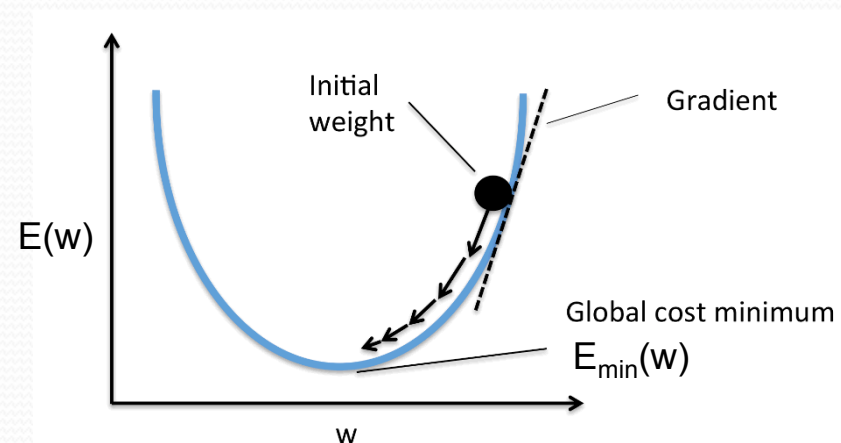
$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau$$

$$(f * g)(t) = \int_{m=-M}^M f(m) g(t - m)$$

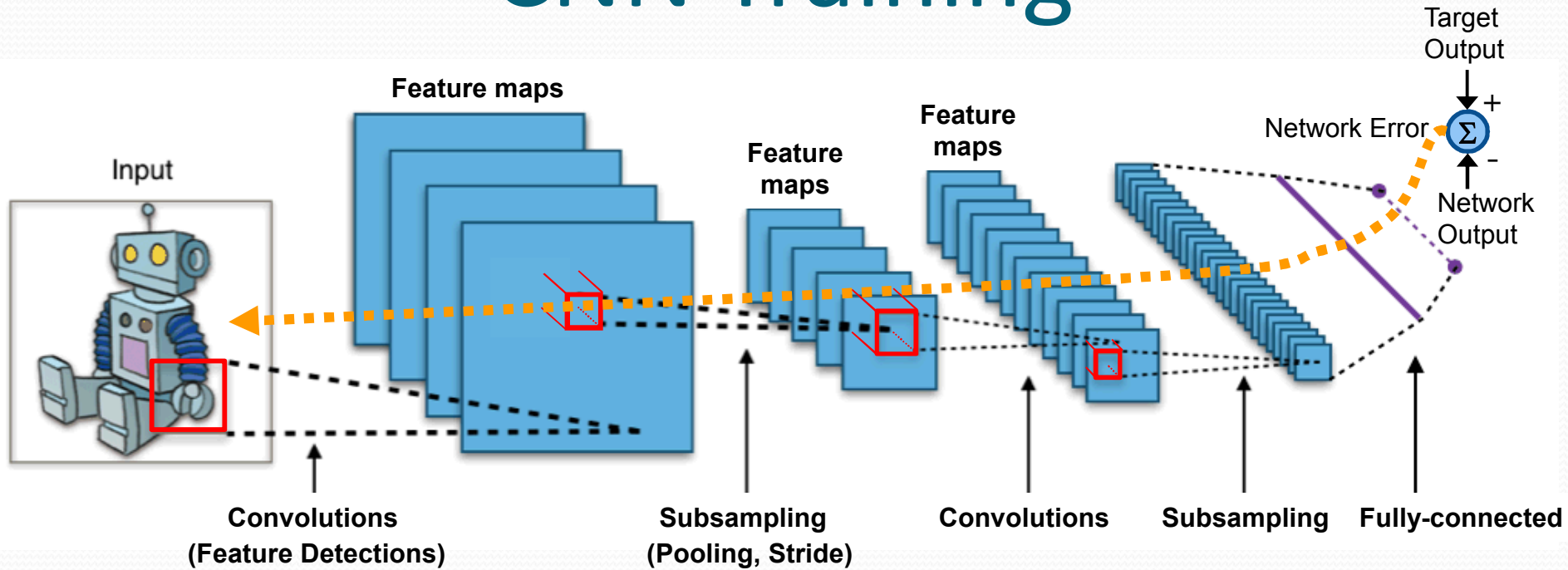
CNN Training



- Convolution filters are the weights in a CNN that can be trained.
- Filter values (weights) are initialized with random values and updated via back-propagation.



CNN Training



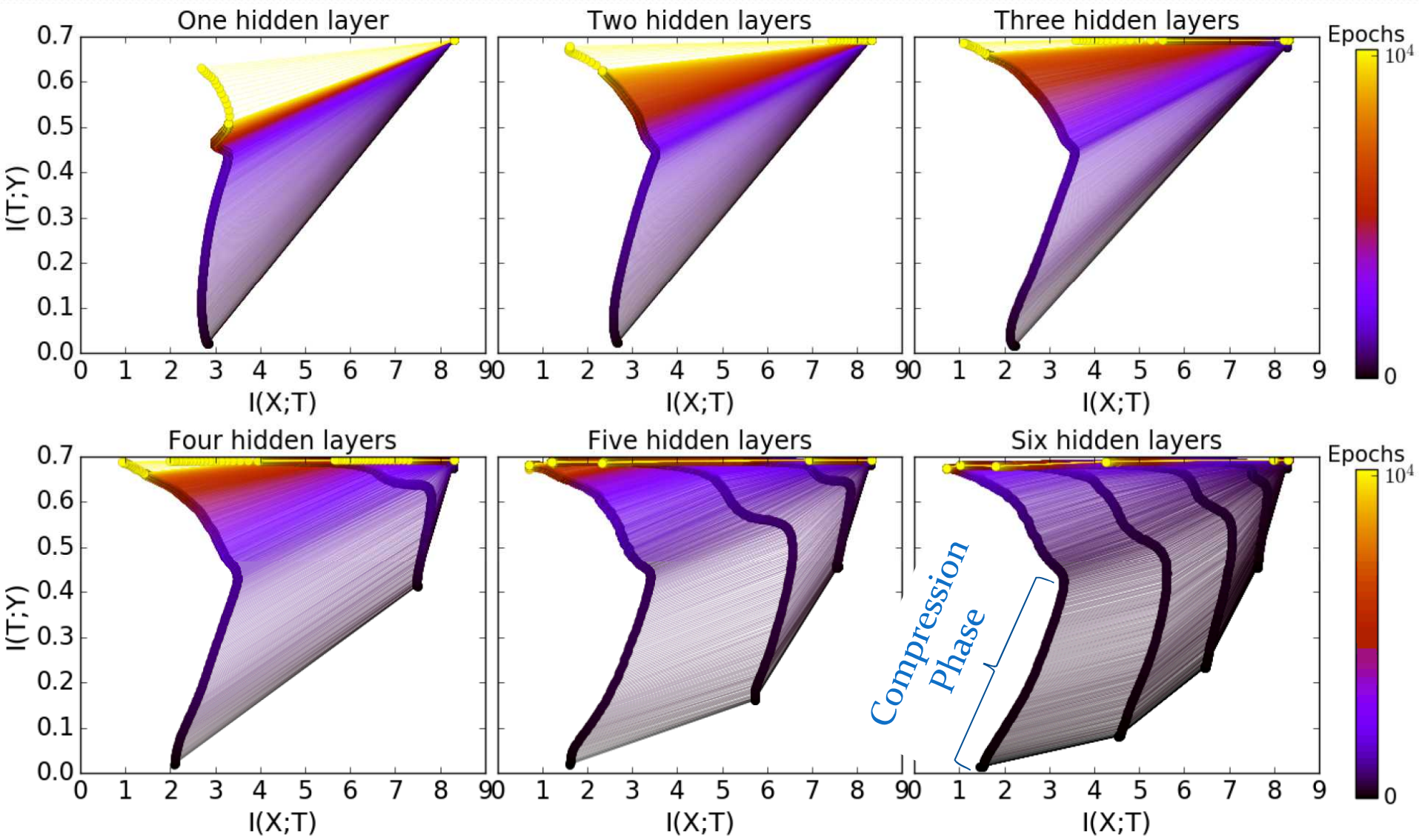
- Convolution filters are the weights in a CNN that can be trained.
- Filter values (weights) are initialized with random values and update via back-propagation.

Weight update during Training

$$w(t + 1) = w(t) - \eta \Delta w(t)$$

$$\Delta w = \frac{\partial E}{\partial w}$$

K bits of compression $\sim 2^K$ training examples



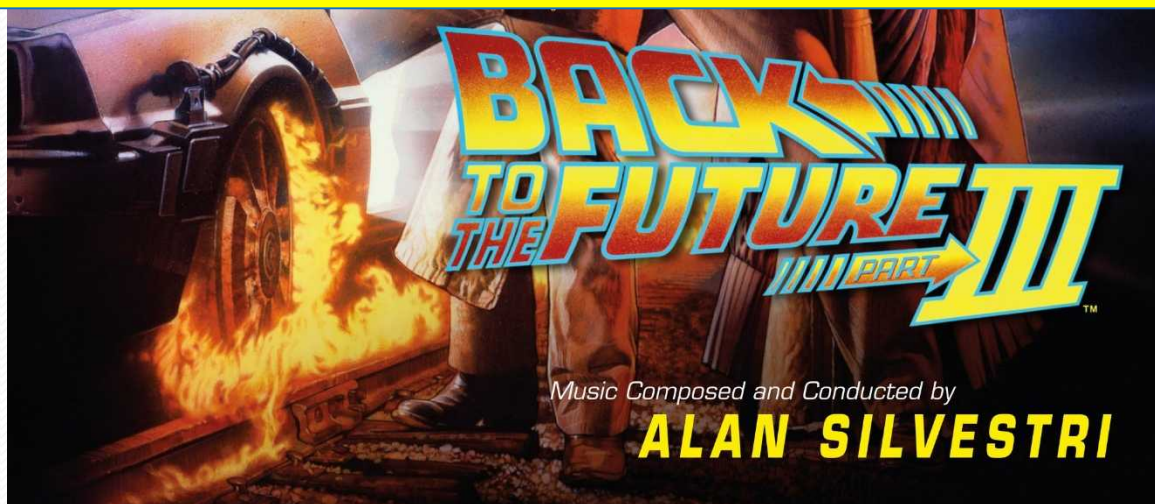
Neural Nets – What's Changed?

25TH ANNIVERSARY EDITION

Original Motion Picture Soundtrack

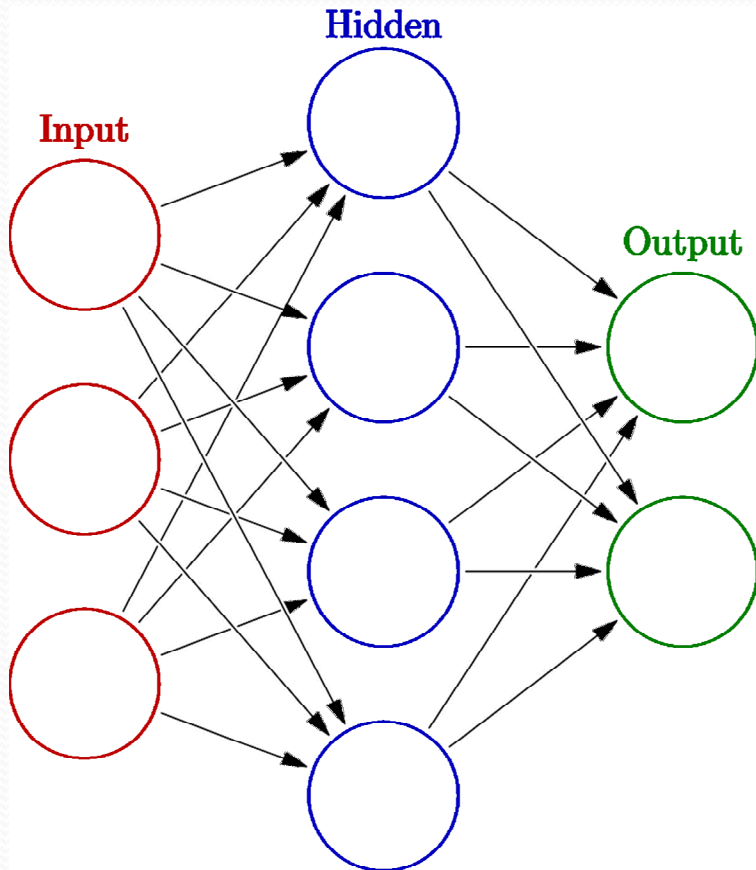


Today's deep-learning success was 30 years in the making.

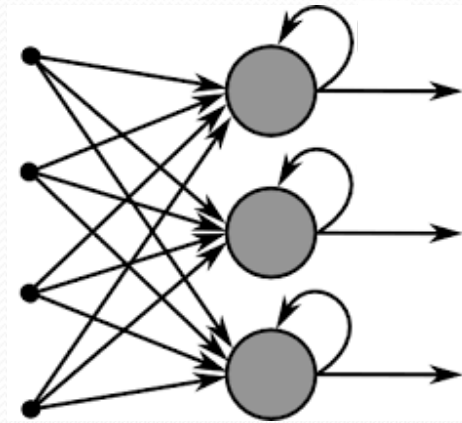


Neural Nets – What's Changed?

- Architecture – NOT REALLY

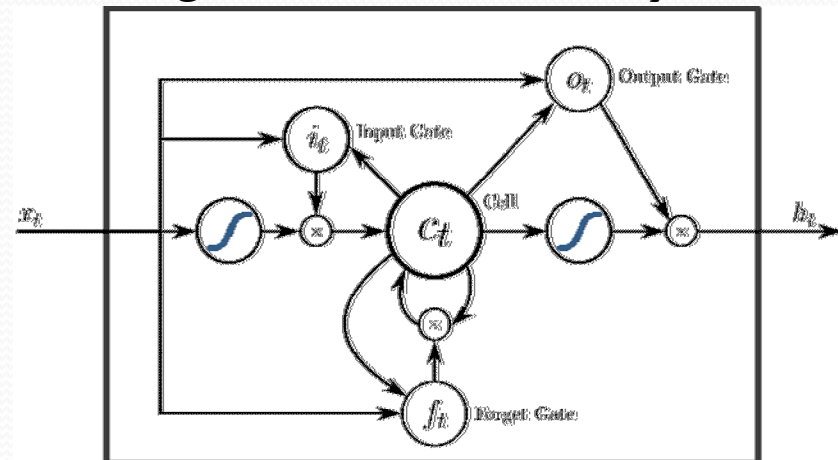


Feed-Forward Neural Network



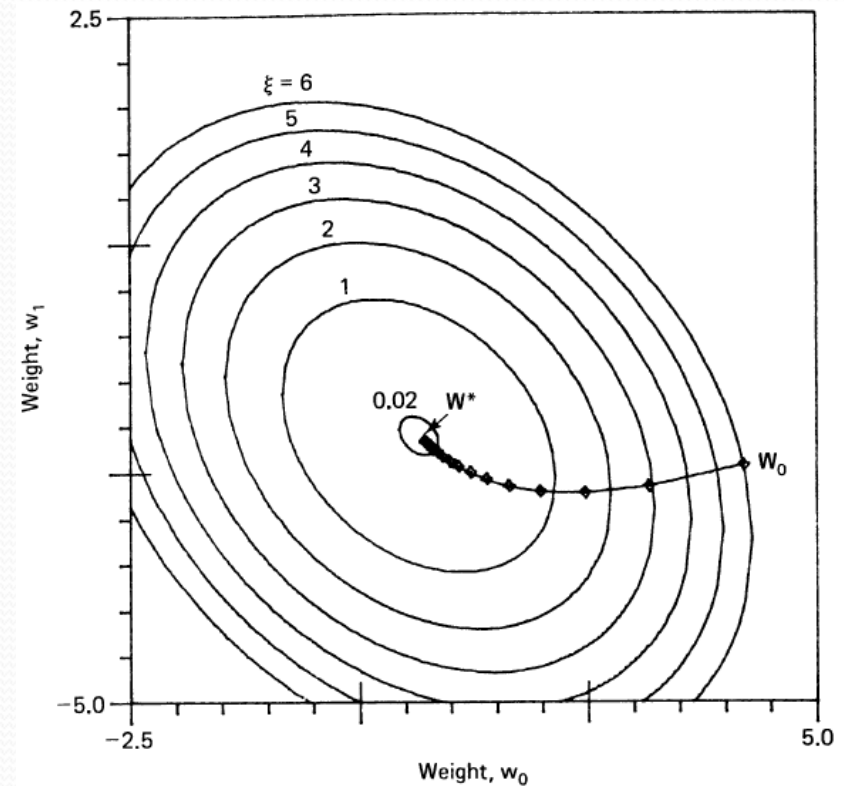
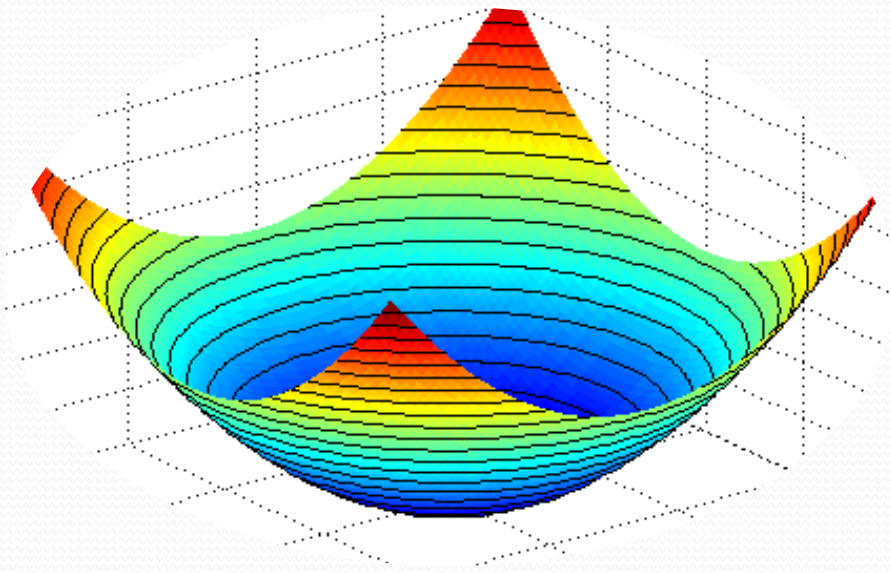
Recurrent Neural Network

Long Short-Term Memory Cell



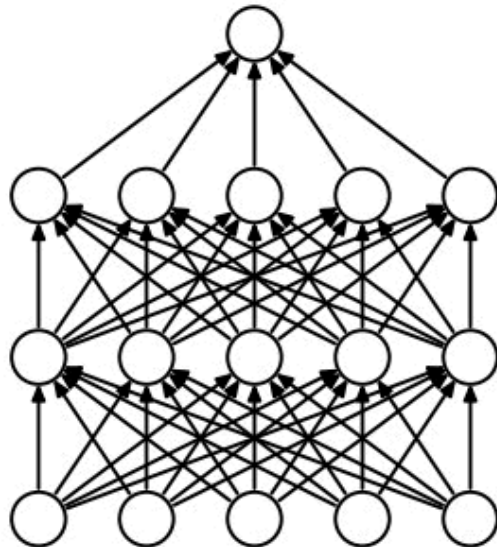
Neural Nets – What's Changed?

- Architecture – NOT REALLY
- Learning algorithms – NOT REALLY since BACKPROP
 - Better Optimizers, Adaptive learning rates, Minibatches

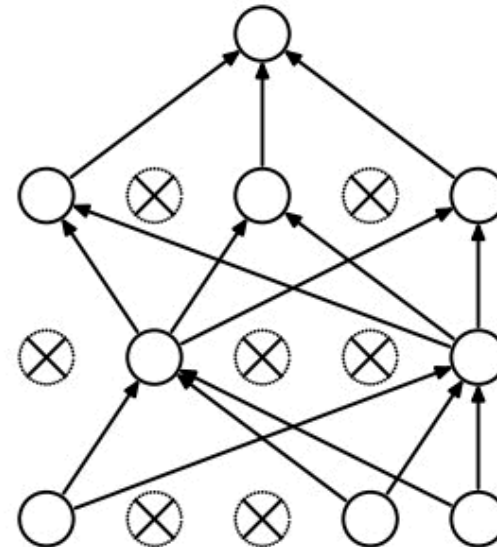


Neural Nets – What's Changed?

- Architecture – NOT REALLY
- Learning algorithms – NOT REALLY since BACKPROP
 - Better Optimizers, Adaptive learning rates, Minibatches
- Regularization techniques – YES
 - Dropout, Data augmentation, Ensemble methods



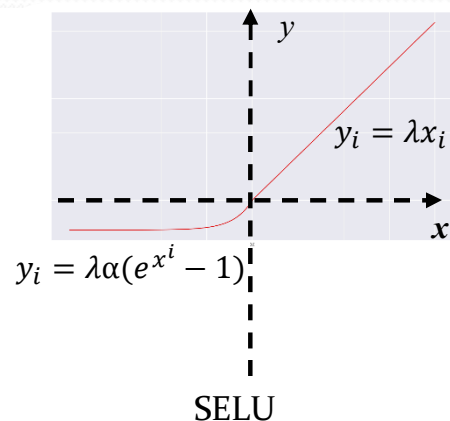
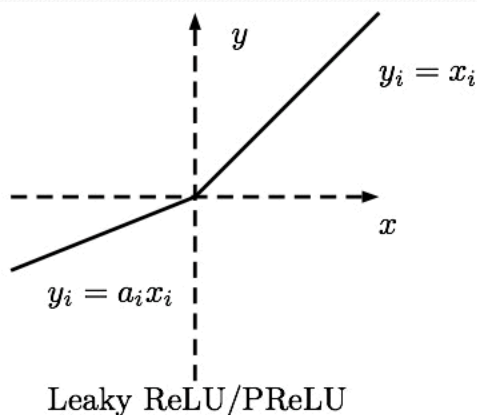
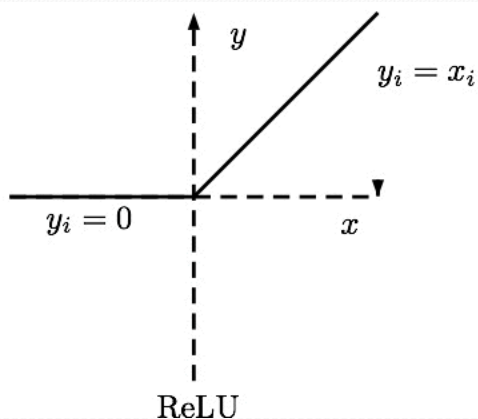
(a) Standard Neural Net



(b) After applying dropout.

Neural Nets – What's Changed?

- Architecture – NOT REALLY
- Learning algorithms – NOT REALLY since BACKPROP
 - Better Optimizers, Adaptive learning rates, Minibatches
- Regularization techniques – YES
 - Dropout, Data augmentation, Ensemble methods
- Activation function – YES
 - Rectified Linear Unit (ReLU), Scaled Exponential LU (SELU)



Neural Nets – What's Changed?

- Architecture – NOT REALLY
- Learning algorithms – NOT REALLY since BACKPROP
 - Better Optimizers, Adaptive learning rates, Minibatches
- Regularization techniques – YES
 - Dropout, Data augmentation, Ensemble methods
- Activation function – YES
 - Rectified Linear Unit (ReLU), Scaled Exponential LU (SELU)
- Labelled data – YES

Music Datasets

- Piano-midi.de
- Nottingham
- MuseData
- JSB
- FMA

Natural Images

- MNIST
- NIST
- Perturbed NIST
- CIFAR₁₀ / CIFAR₁₀₀
- Caltech 101
- Caltech 256
- Caltech Silhouettes
- STL-10
- Street View House Numbers (SVHN)
- NORB:
- Imagenet
- Pascal VOC
- Labelme
- COIL 20
- COIL₁₀₀

Faces

- Labelled Faces in the Wild
- Toronto Face
- Olivetti
- Multi-Pie
- Face-in-Action
- JACFEE
- FERET
- IndianFaceDatabase:
- The Yale Face Database

Artificial Datasets

- Arcade Universe
- BabyAISchool:
 - BabyAIShapesDatasets
 - BabyAIImageAndQuestion
- DeepVsShallowComparison
ICML2007
 - MnistVariations
 - RectanglesData
 - ConvexNonConvex
 - BackgroundCorrelation

Text

- 20 newsgroups
- Reuters (RCV*)
- Penn Treebank
- Broadcast News
- Wikipedia Dataset
- Multidomain sentiment analysis

Speech

- TIMIT Speech Corpus
- Aurora

Recommendation Systems

- MovieLens
- Jester
- Netflix Prize
- Book-Crossing dataset

Miscellaneous

- “Musk”
- CMU Motion Capture Database
- Brodatz
- Million Song
- Merck Molecular Activity

Neural Nets – What's Changed?

- Architecture – NOT REALLY
- Learning algorithms – NOT REALLY since BACKPROP
 - Better Optimizers, Adaptive learning rates, Minibatches
- Regularization techniques – YES
 - Dropout, Data augmentation, Ensemble methods
- Activation function – YES
 - Rectified Linear Unit (ReLU), Scaled Exponential LU (SELU)
- Labelled data – YES
- Compute power – YES
 - GPUs



Neural Nets – What's Changed?

- Architecture – NOT REALLY
- Learning algorithms – NOT REALLY since BACKPROP
 - Better Optimizers, Adaptive learning rates, Minibatches
- Regularization techniques – YES
 - Dropout, Data augmentation, Ensemble methods
- Activation function – YES
 - Rectified Linear Unit (ReLU), Scaled Exponential LU (SELU)
- Labelled data – YES
- Compute power – YES
 - GPUs
- Software – YES



Opportunities at Sandia

- Cyber-security
- Trust
- Explainability
- New algorithms / arrangement of algorithms
- Data
 - Multi-modal
 - Visual, Audio, Text, RF, Seismic, LIDAR, SAR

Competitions

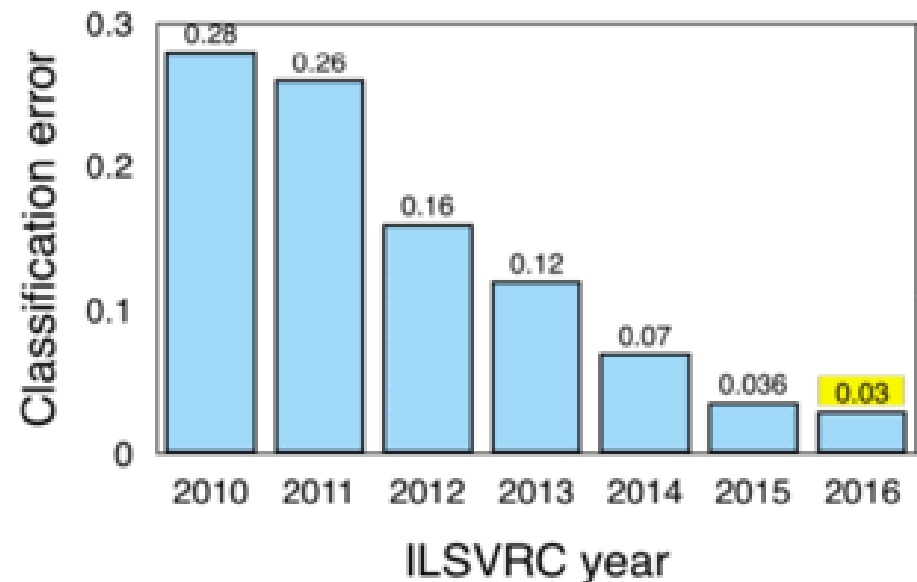
ImageNet Challenge



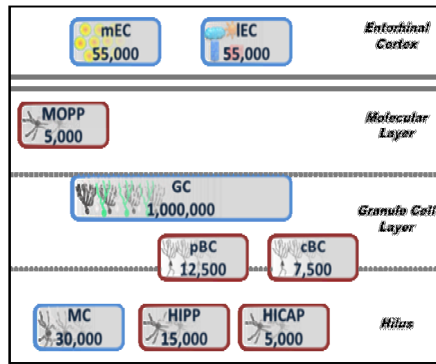
- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.



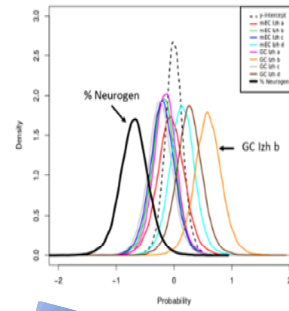
- Given a dataset
 - Fund multiple teams to compete to solve your problem



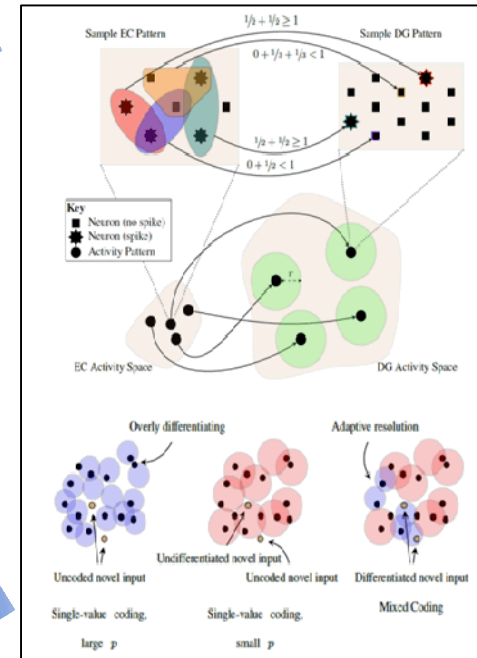
Simulate at high level of neural fidelity



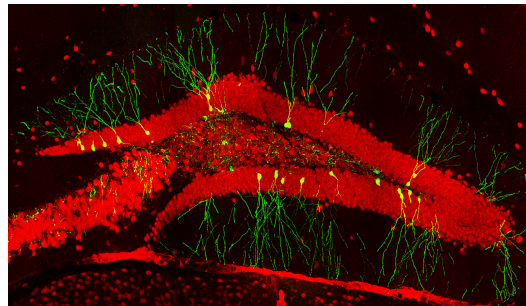
GC Lin. Compressibility: Prob. Density of Regression Coeff.



Formalize hippocampus / neurogenesis sparse coding algorithm



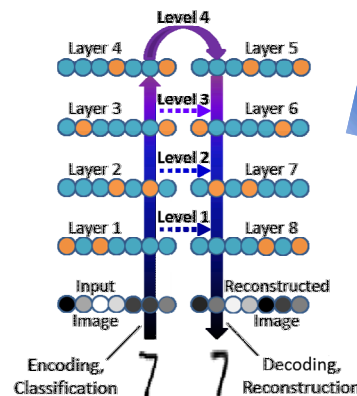
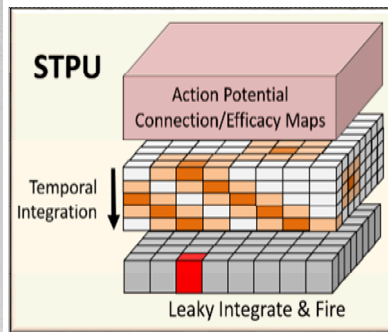
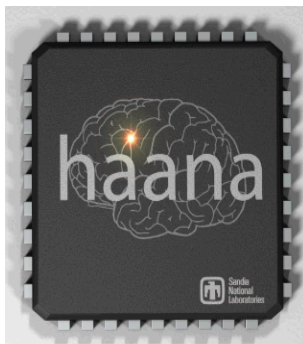
Dentate Gyrus with Green New Neurons

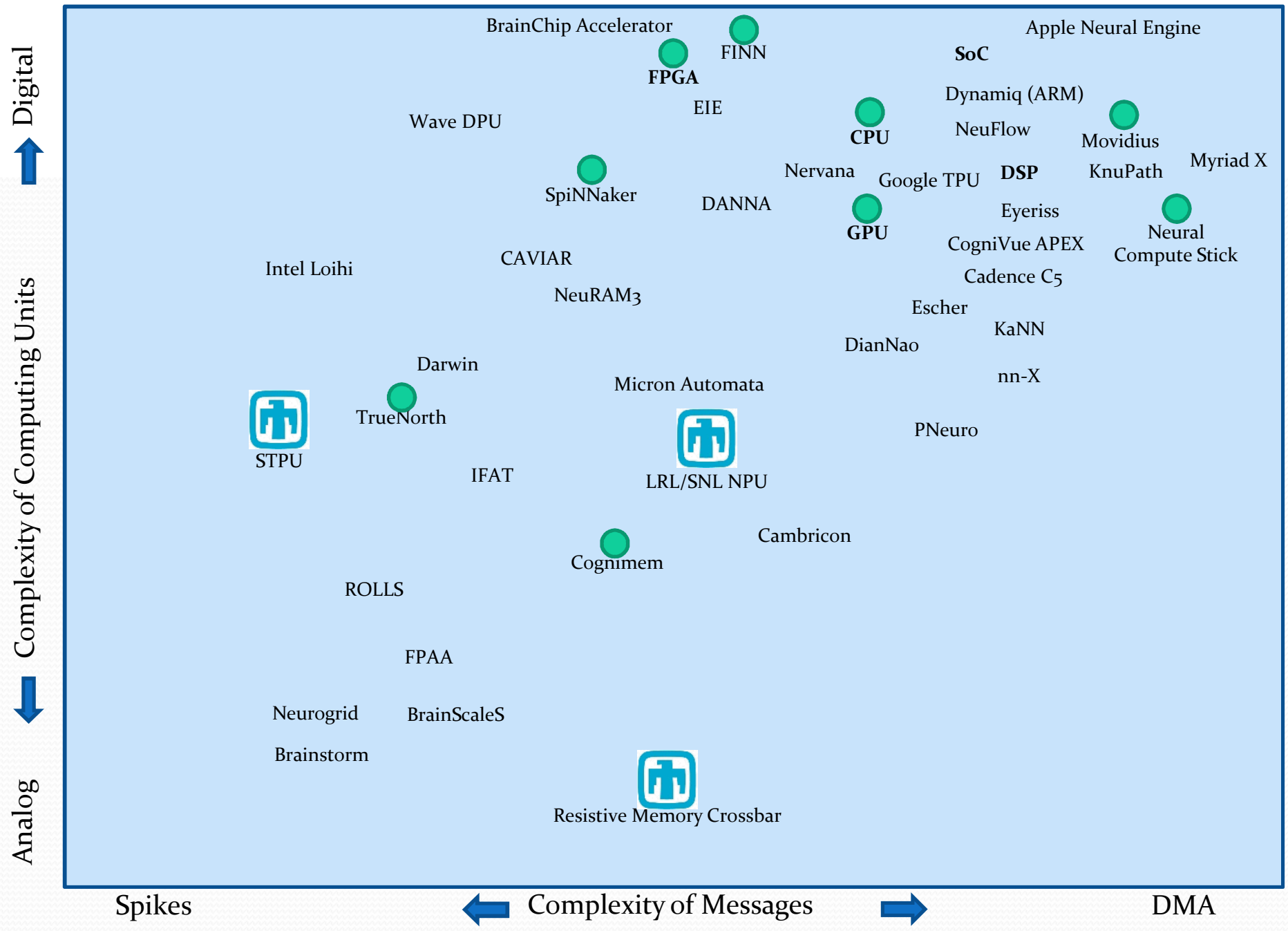


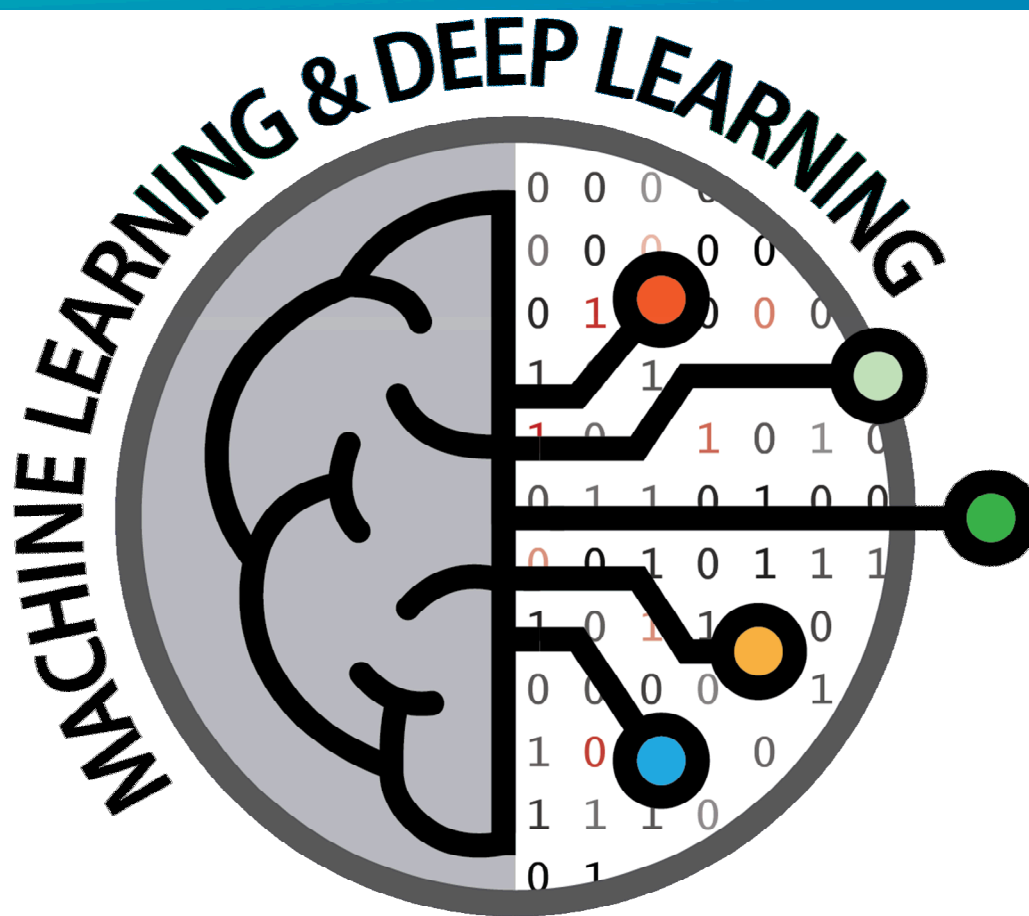
Identify critical aspects of computation

- expansive scale
- sparse activation
- neurogenesis

Adaptive Neural Hardware, Architectures, and Algorithms





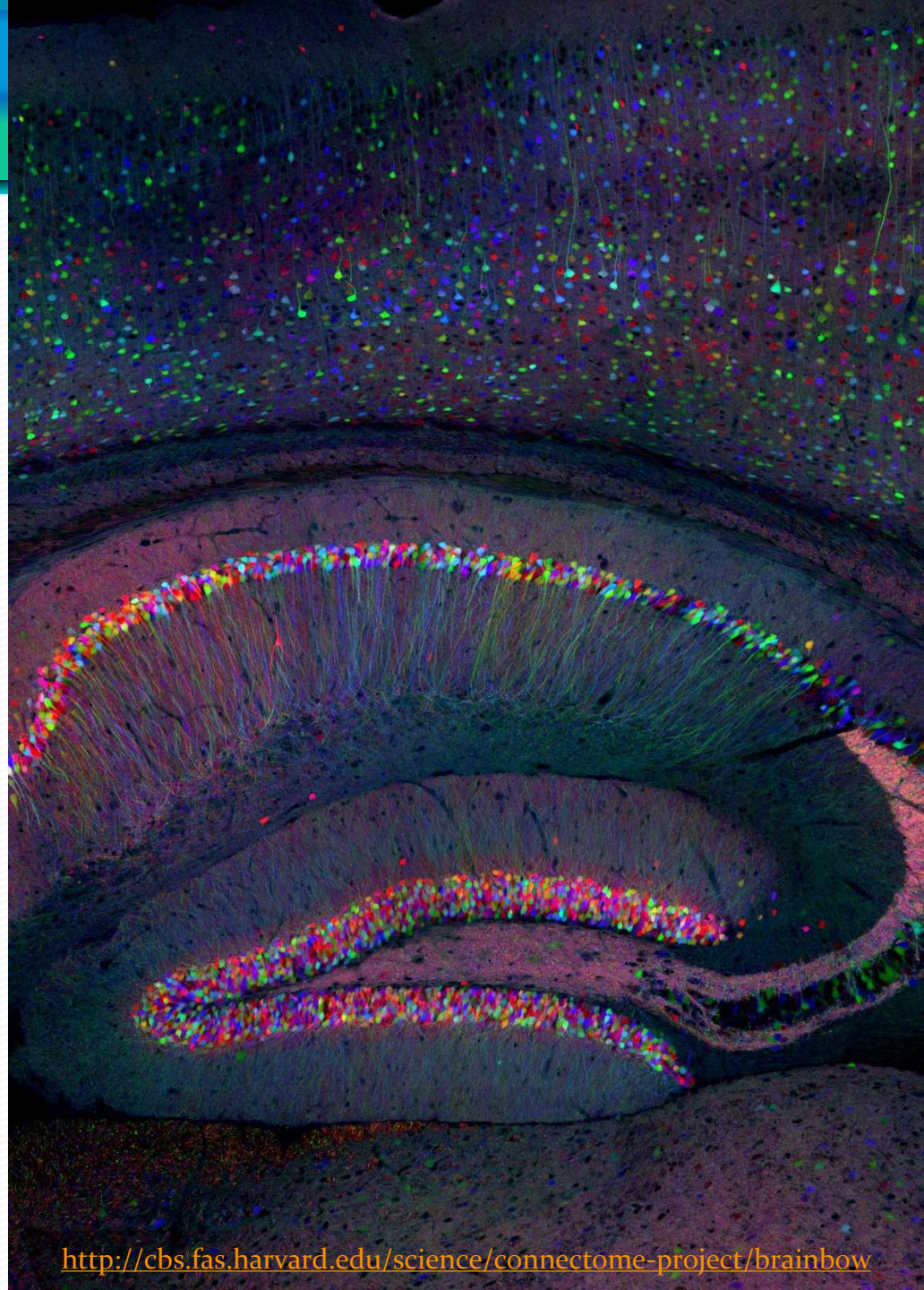


MLDL

Conference

Future of DL

- Intertwined Multimodal-Temporal Networks
- Semi-Supervised RL
- Low-Shot Learning
- Attention, Focus, Anticipation
 - What to ignore in streaming data
- Spiking algorithms
- Learnable hardware



Made in China 2025

- New advanced information technology
- Automated machine tools & robotics
- Aerospace and aeronautical equipment
- Maritime equipment and high-tech shipping
- Modern rail transport equipment
- New-energy vehicles and equipment
- Power equipment
- Agricultural equipment
- New materials
- Biopharma and advanced medical products



Third Offset Strategy

- **First** – Early 1950's: miniaturizing of nuclear weapons
- **Second** – 1970's and 80's: conventional weapons with near-zero miss, precision-guided weapons and the joint battle networks that employed them.
 - Information & digital technologies for sensors & weapons
- **Third** – Now: Intellectual advantage



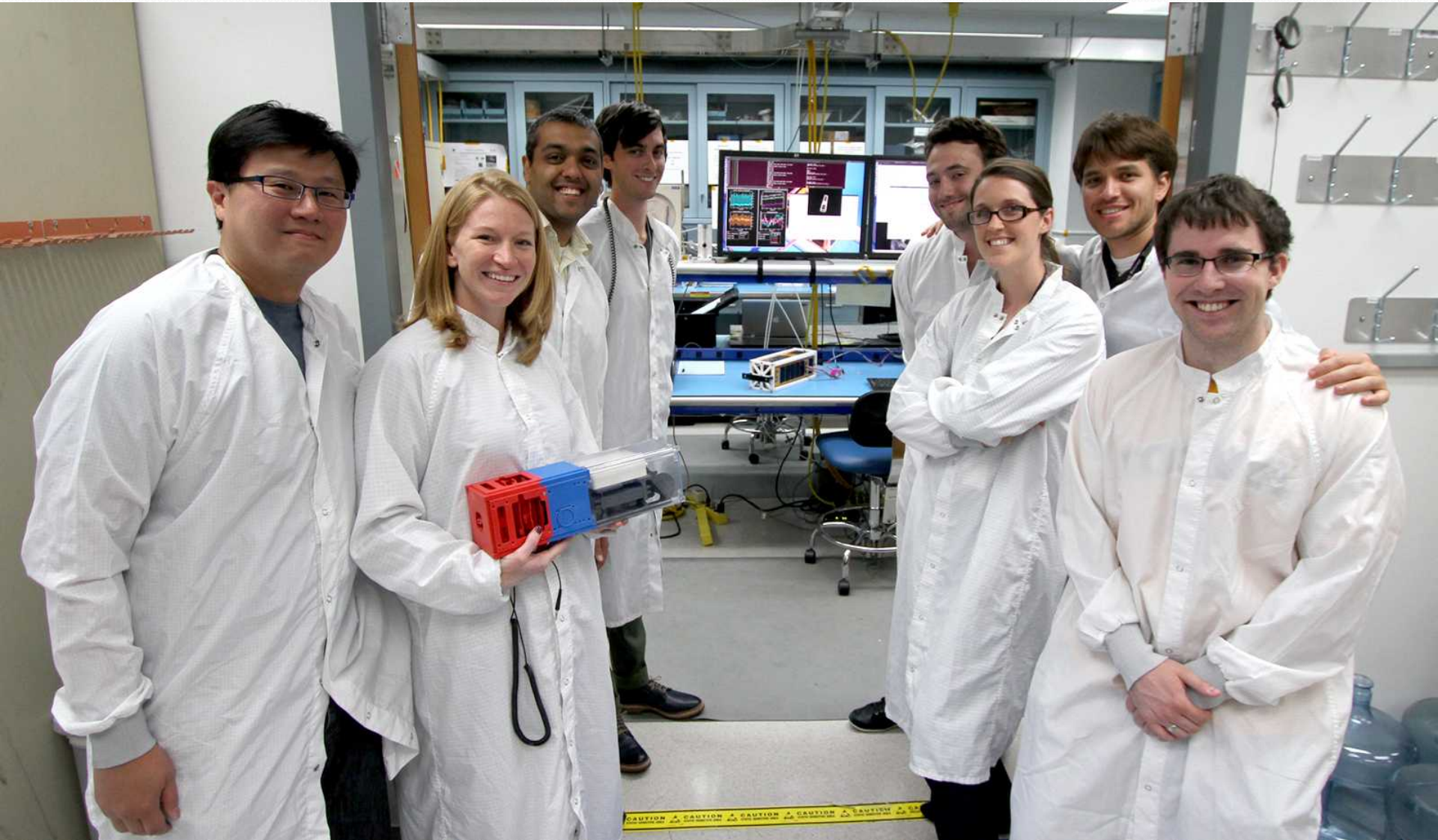
Third Offset Strategy*

- **Deep Learning Systems**
 - Indications and warnings in cyber defense, electronic warfare attacks and large-density missile raids.
 - Big-data analytics
- **Human-Machine Collaboration**
 - Better, faster Human decision-making
 - Humans ability to think on the fly + AI
- **Human-Machine Combat Teaming**
 - With unmanned systems
- **Assisted Human Operations**
 - Wearable electronics, combat apps, heads-up displays exoskeletons to help warfighters
- **Network-Enabled, Cyber-Hardened Weapons**
 - Prioritization of cyber-security



AI = Tool in the hands of moral agents





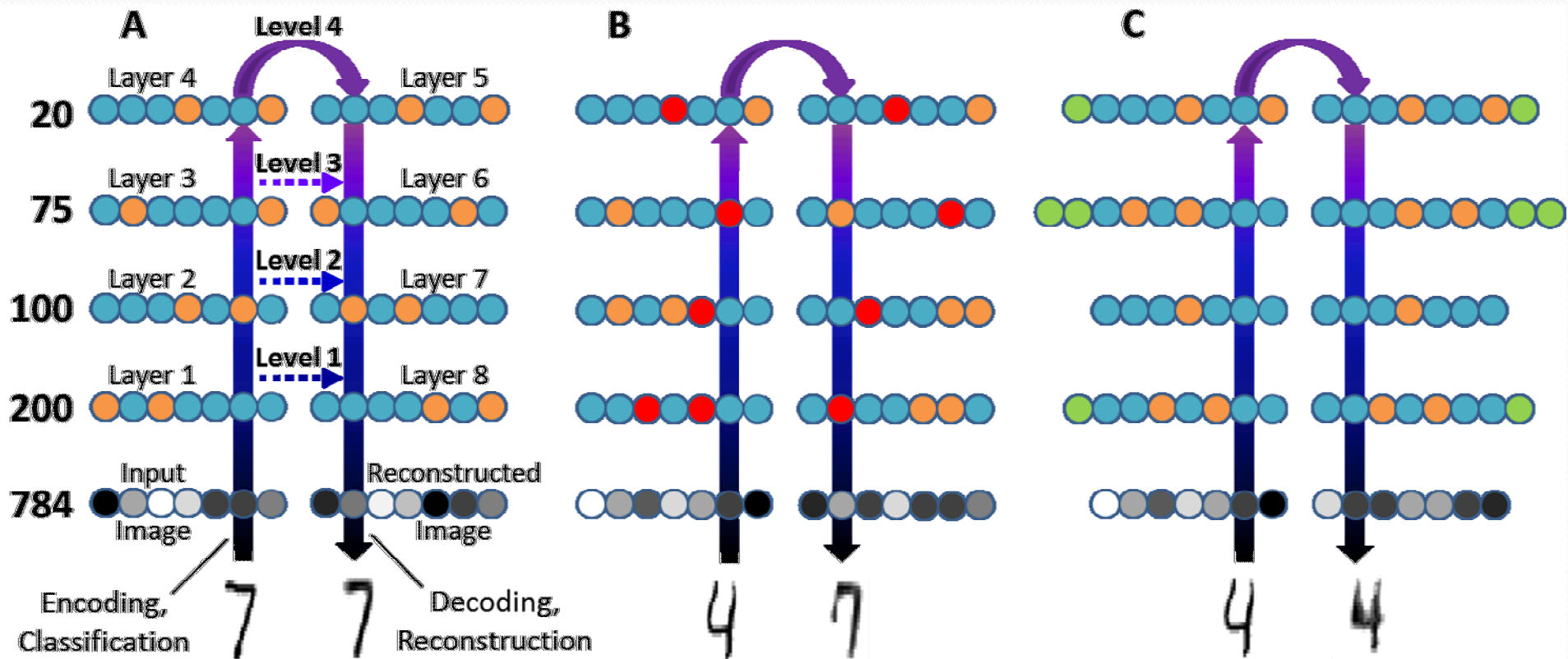
Backup Slides

Continuous Low-Shot

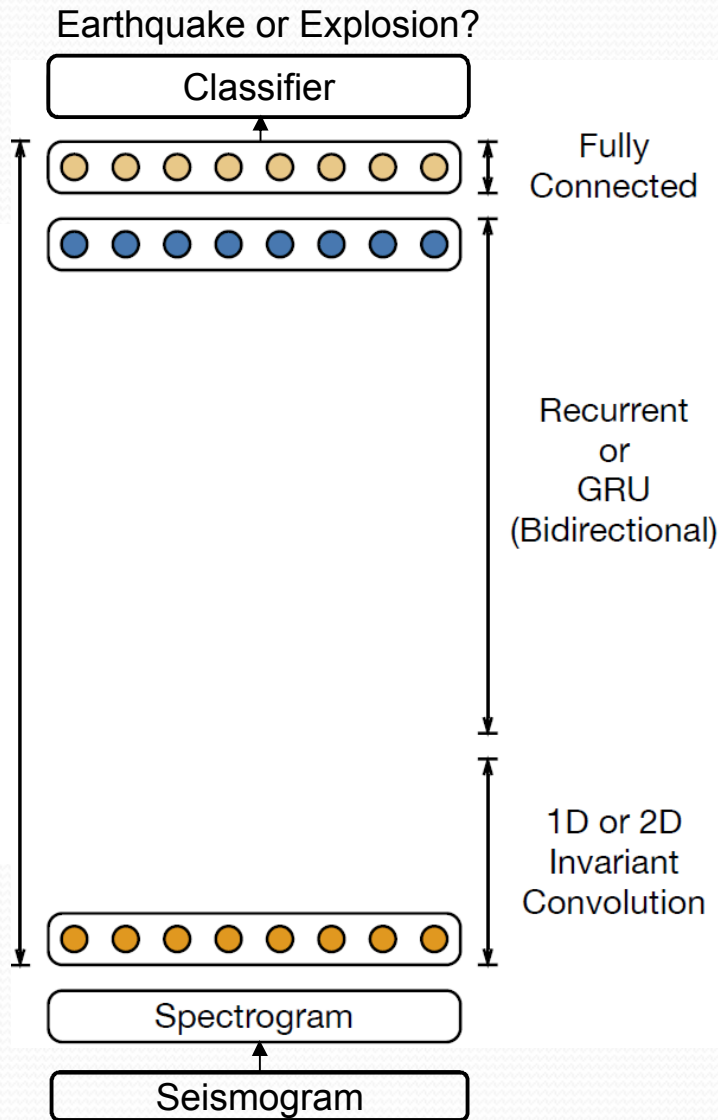
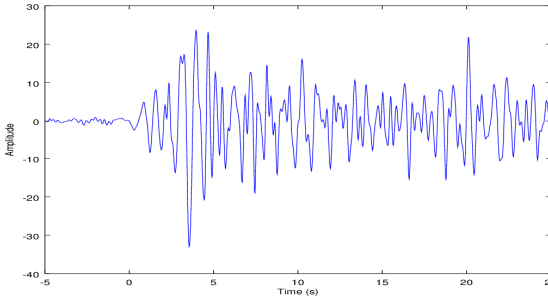
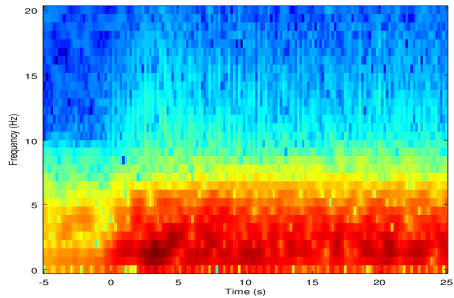
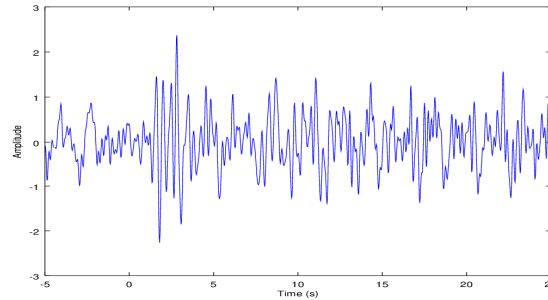
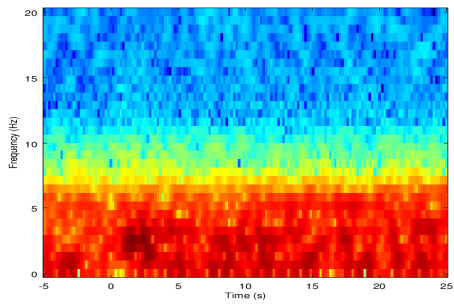
Neurogenesis Deep Learning

PROBLEM: Existing Autoencoder (AE) can't represent novel data

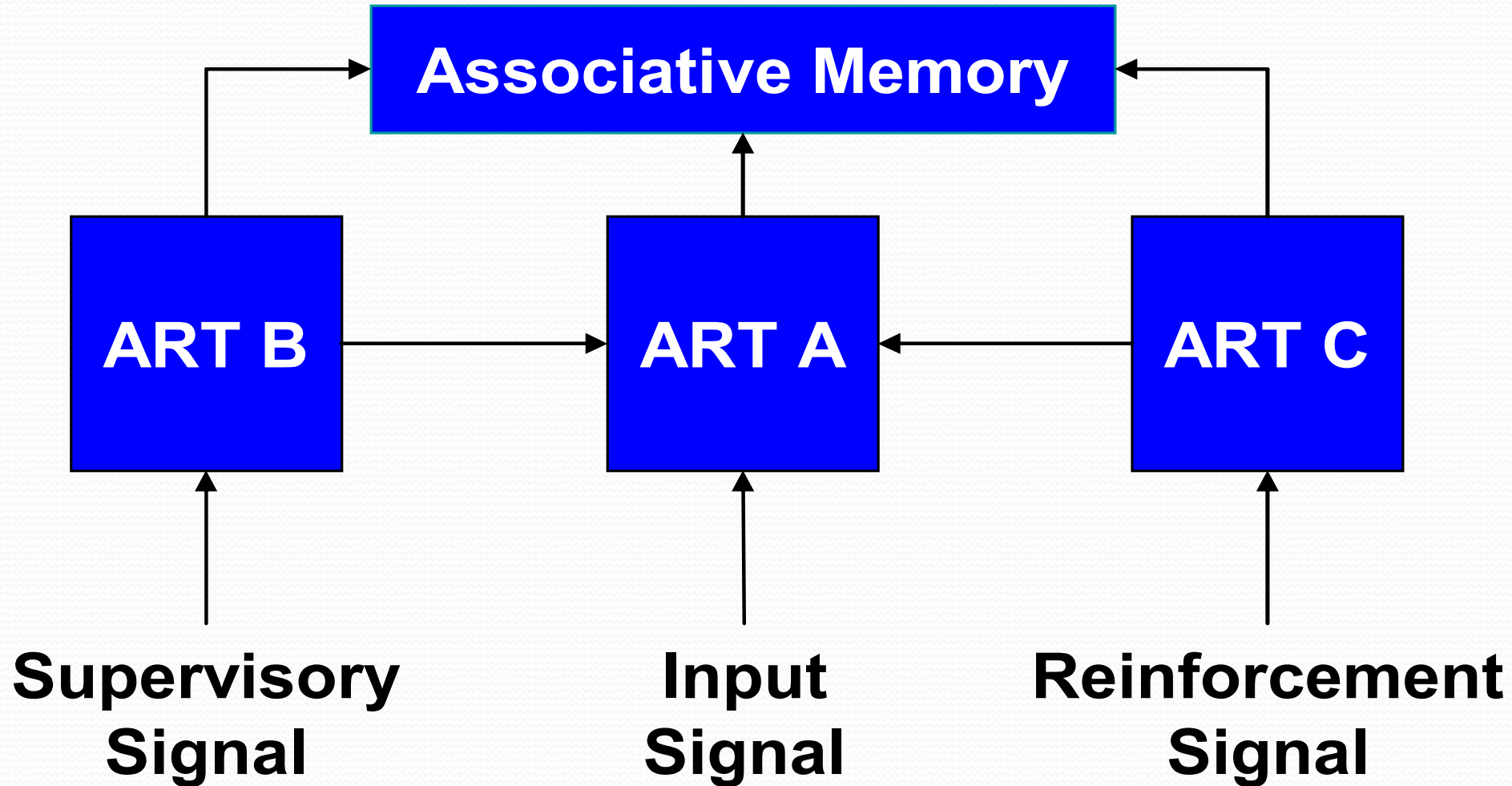
SOLUTION: Add new node(s) to each level of an AE as needed to represent novel data



Seismic Event Type Discrimination, Phase ID, Event Location



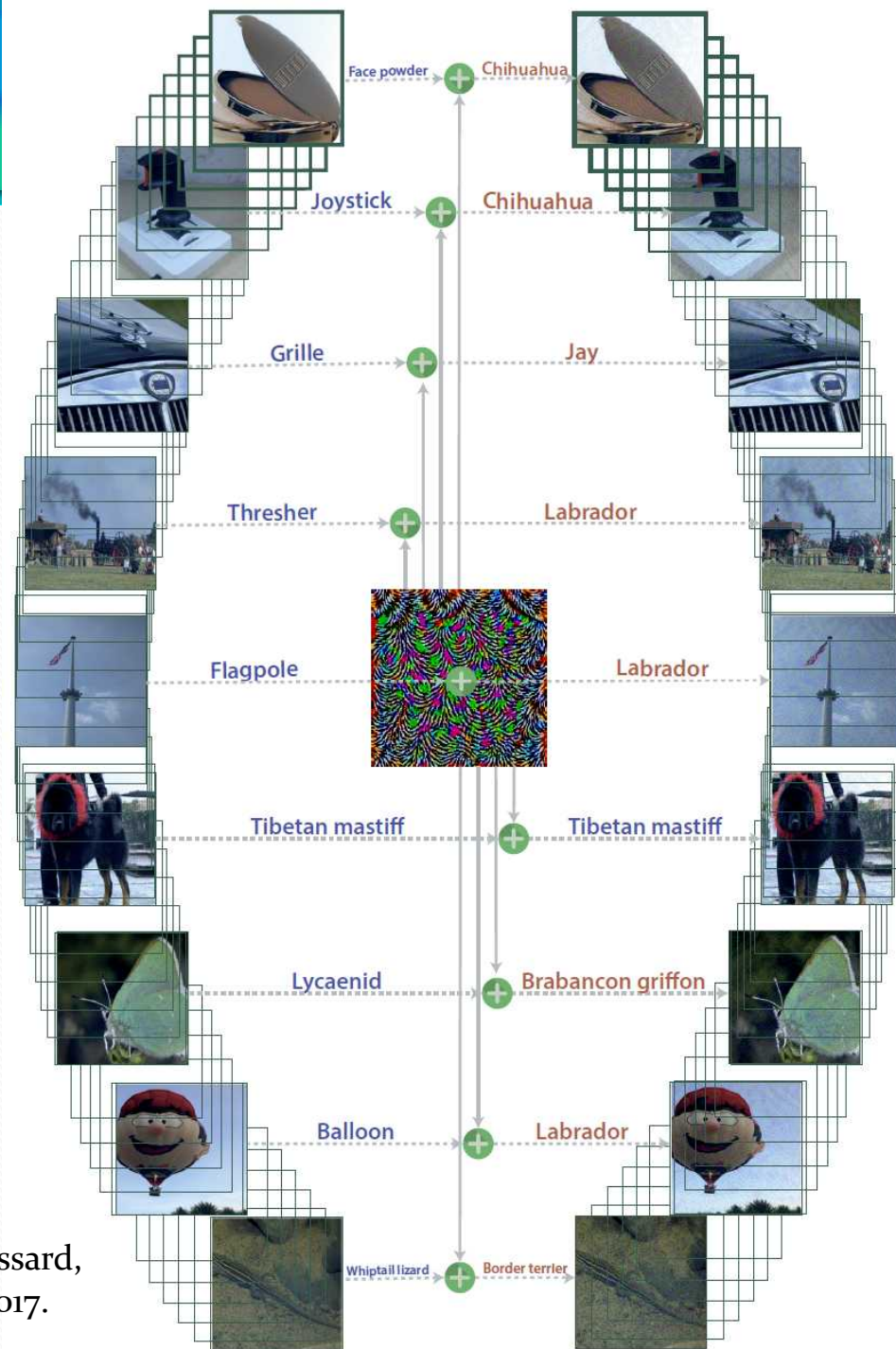
Coordinated Machine Learning



Secure Deep Learning

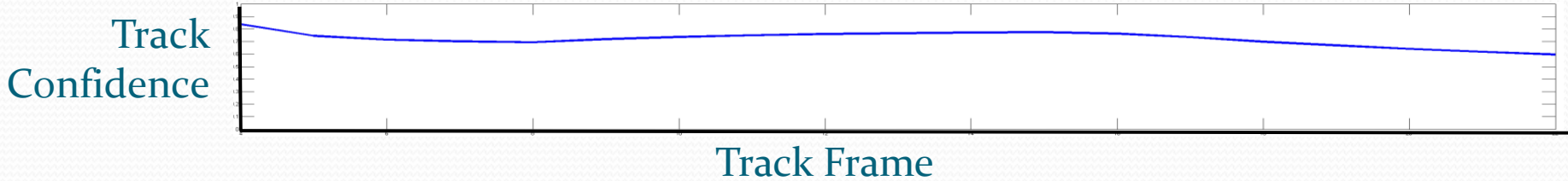
- Can Sandia identify what's really going on with *adversarial manipulation* of data and develop mitigations?

S.-M. Moosavi-Dezfooli, A. Fawzi, O. Fawzi, and P. Frossard, "Universal adversarial perturbations," in IEEE CVPR, 2017.

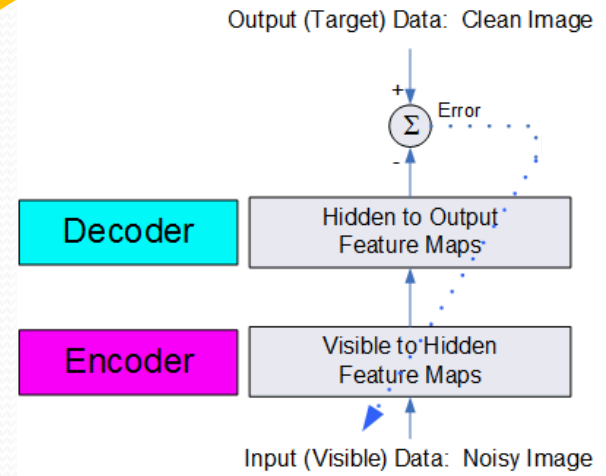
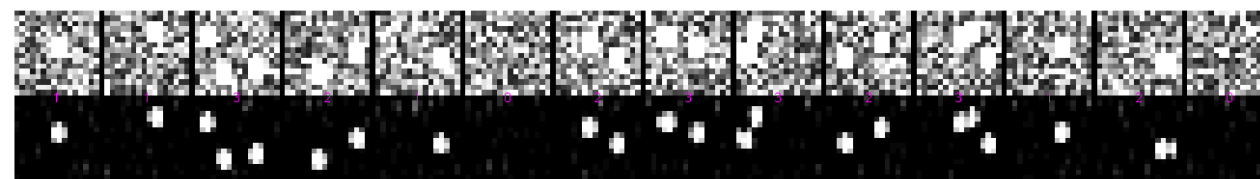


Object Recognition & Tracking

- Object Tracking with CNN
 - Real Sandia Peak data augmented with simulated data



- Denoising



Cyber Security

