Introduction of Deep Learning

Giảng viên: Tiến sĩ Bùi Thanh Hùng
Email: hung.buithanhcs@gmail.com
Website: https://sites.google.com/site/hungthanhbui1980/

Acknowledgement

- The contents of these slides have origin from School of Interactive Computing, Georgia Tech, United States
- We greatly appreciate support from Prof. Dhruv Batra for kindly sharing these materials.

What is this class about?

- Introduction to Deep Learning

- Goal:
  - After finishing this class, you should be ready to get started on your first DL research project.
    - CNNs
    - RNNs
    - Deep Reinforcement Learning
    - Generative Models (VAEs, GANs)

- Target Audience:
  - MS

What this class is NOT

- NOT the target audience:
  - Advanced grad-students already working in ML/DL areas
  - People looking to understand latest and greatest cutting-edge research (e.g. GANs, AlphaGo, etc)
  - Masters students looking to graduate with a DL class on their resume.
- NOT the goal:
  - Teaching a toolkit. “Intro to TensorFlow/PyTorch”
  - Intro to Machine Learning

Prerequisites

- Intro Machine Learning
  - Classifiers, regressors, loss functions, MLE, MAP
- Linear Algebra
  - Matrix multiplication, eigenvalues, positive semi-definiteness…
- Calculus
  - Multi-variate gradients, hessians, jacobians…

Prerequisites

Prerequisites

- Intro Machine Learning
  - Classifiers, regressors, loss functions, MLE, MAP
- Linear Algebra
  - Matrix multiplication, eigenvalues, positive semi-definiteness…
- Calculus
  - Multi-variate gradients, hessians, jacobians…
- Programming!

Computing

- Major bottleneck
  - GPUs

- Options
  - Your own / group / advisor’s resources

  - Google Colab
    - jupyter-notebook + free GPU instance

What is the collaboration policy?

- Collaboration
  - Only on HWs and project.
  - You may discuss the questions
  - Each student writes their own answers
  - Write on your homework anyone with whom you collaborate
  - Each student must write their own code for the programming part

- Zero tolerance on plagiarism
  - Neither ethical nor in your best interest
  - Always credit your sources
  - Don’t cheat. We will find out.

Research

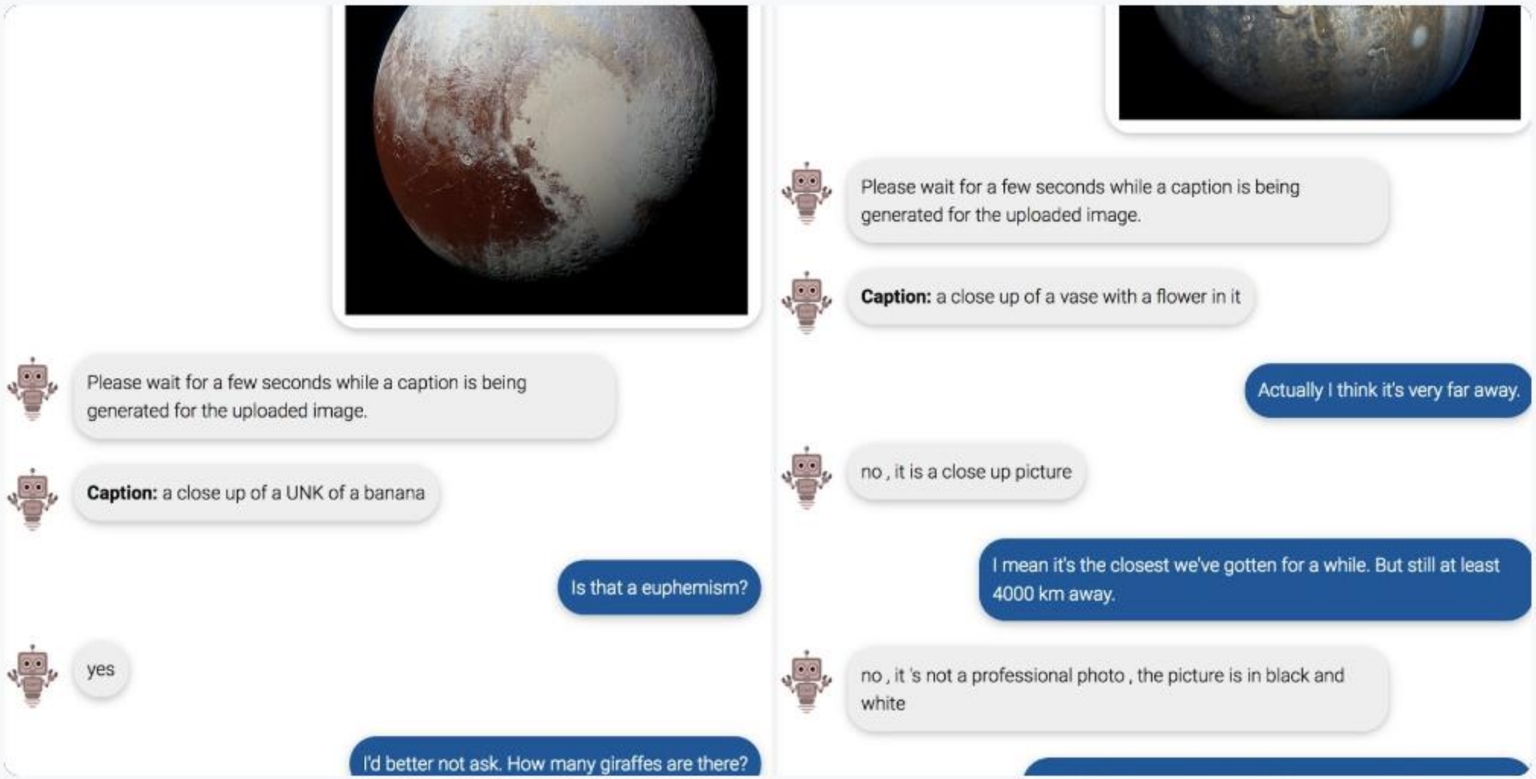
- "Can I work with your group for funding/credits/neither?"
  - I am not taking new advising duties.
  - If you can find one of my students to supervise you,
    I am happy to sign off on the paperwork.
  - Your responsibility to approach them and ask.
    It will help if you know what they are working on.

Janelle Shane @JanelleCShane · Jun 24

One fun thing I discovered about Visual Chatbot.

It learned from answers that humans gave, and apparently nobody ever asked "how many giraffes are there?" when the answer was zero.

demo.visualdialog.org



24 159 515

Concepts

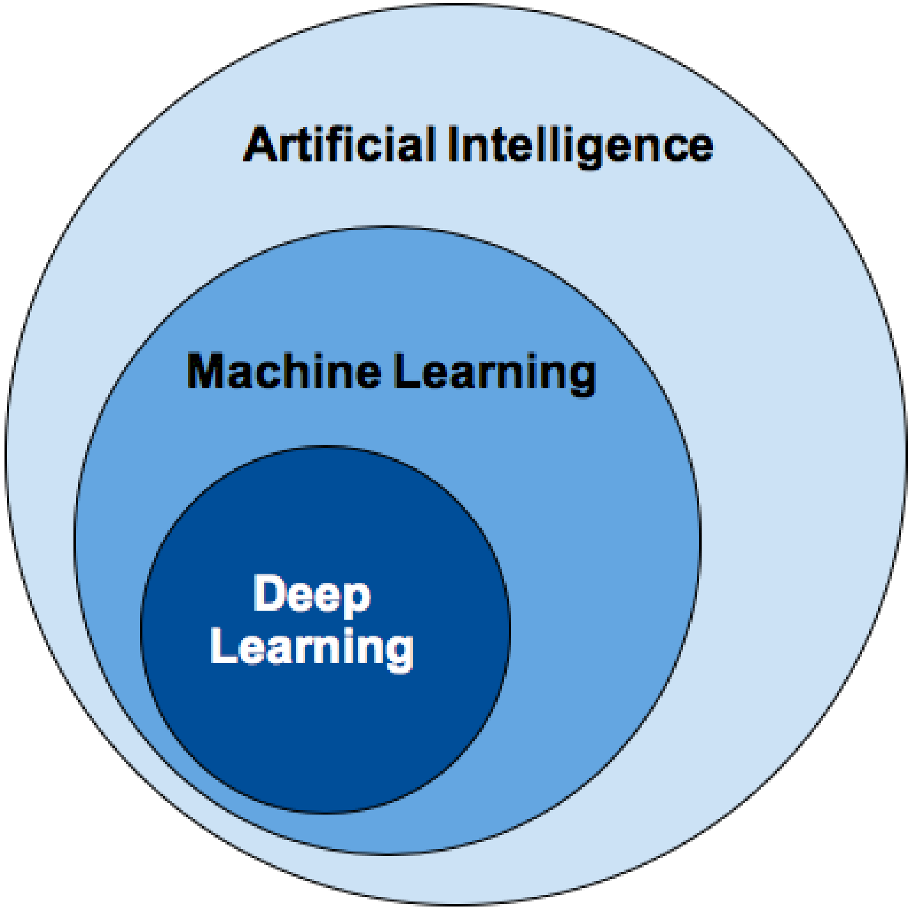


Image Credit: https://www.sumologic.com/blog/machine-learning-deep-learning/

What is (general) intelligence?

- Boring textbook answer

The ability to acquire and apply knowledge and skills

  – Dictionary

- My favorite

The ability to navigate in problem space

  – Siddhartha Mukherjee, Columbia

What is artificial intelligence?

- Boring textbook answer

Intelligence demonstrated by machines

  – Wikipedia

- My favorite

The science and engineering of making computers behave in ways that, until recently, we thought required human intelligence.

  – Andrew Moore, CMU

What is machine learning?

- My favorite

Study of algorithms that
improve their performance (P)
at some task (T)
with experience (E)

  - Tom Mitchell, CMU

So what *is* Deep (Machine) Learning?

- Representation Learning

- Neural Networks

- Deep Unsupervised/Reinforcement/Structured/
  <insert-qualifier-here>
  Learning

- Simply: Deep Learning

So what *is* Deep (Machine) Learning?

- A few different ideas:

- (Hierarchical) Compositionality
  - Cascade of non-linear transformations
  - Multiple layers of representations

- End-to-End Learning
  - Learning (goal-driven) representations
  - Learning to feature extraction

- Distributed Representations
  - No single neuron “encodes” everything
  - Groups of neurons work together

Hierarchical Compositionality

VISION

pixels → edge → texton → motif → part → object

SPEECH

sample → spectral band → formant → motif → phone → word

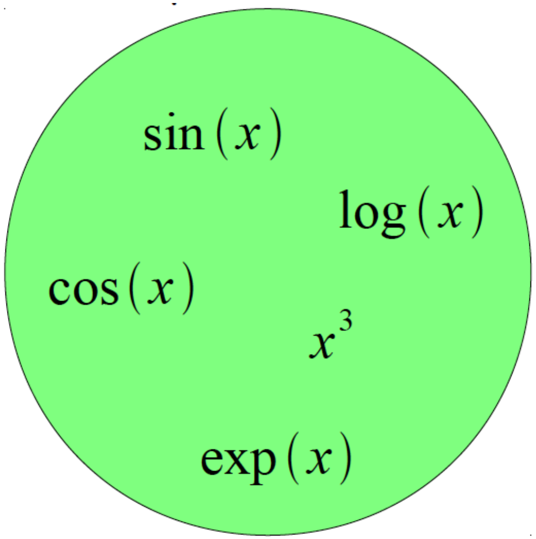
NLP

character → word → NP/VP/.. → clause → sentence → story

Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

Building A Complicated Function

Given a library of simple functions

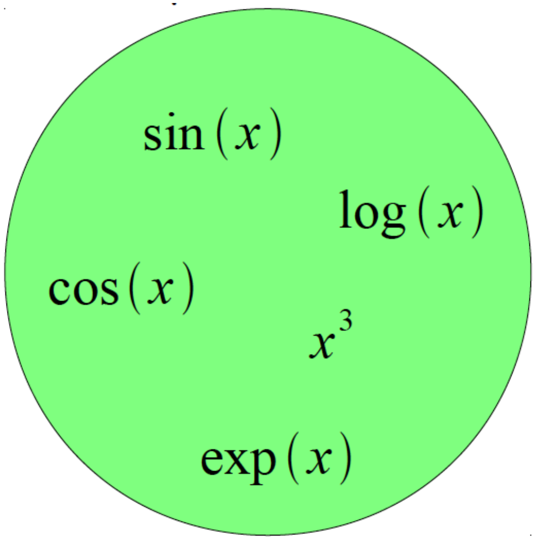


Compose into a

complicate function

Building A Complicated Function

Given a library of simple functions



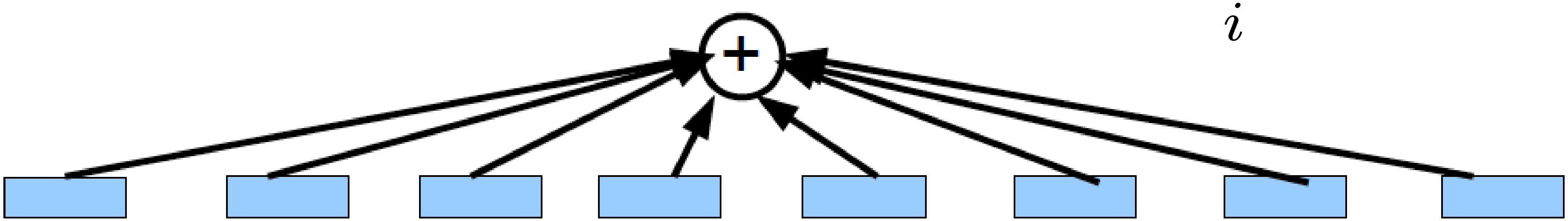
Compose into a

complicate function

Idea 1: Linear Combinations

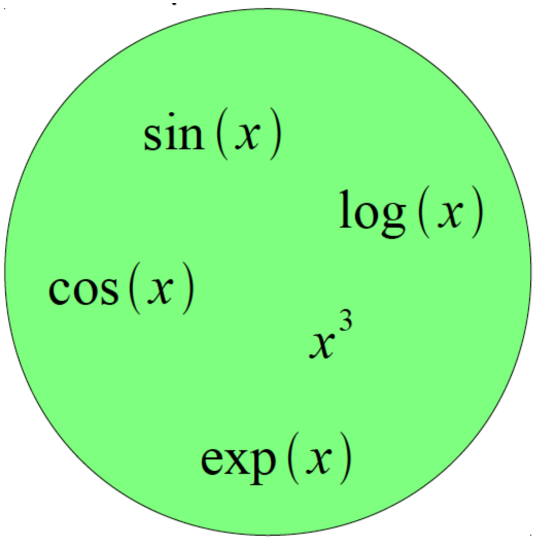
- Boosting
- Kernels
- …

$$f(x) = \sum_i \alpha_i g_i(x)$$



Building A Complicated Function

Given a library of simple functions



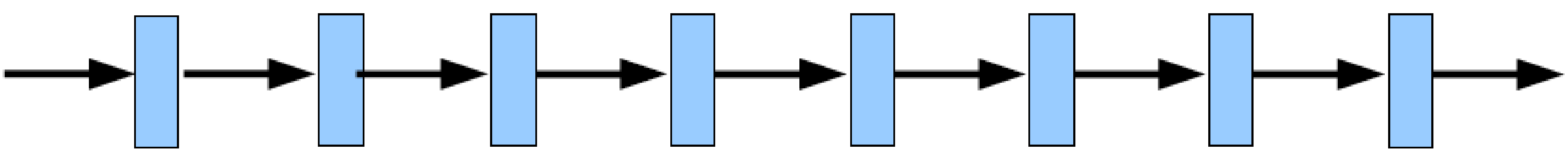
Compose into a

complicate function

Idea 2: Compositions

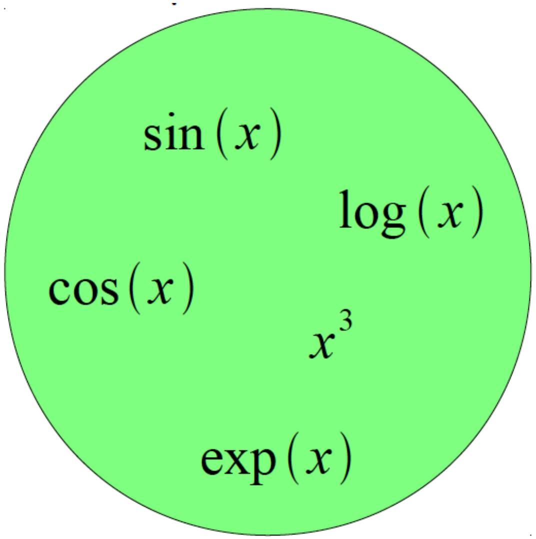
- Deep Learning
- Grammar models
- Scattering transforms…

$$f(x) = g_1(g_2(\ldots(g_n(x)\ldots))$$



Building A Complicated Function

Given a library of simple functions



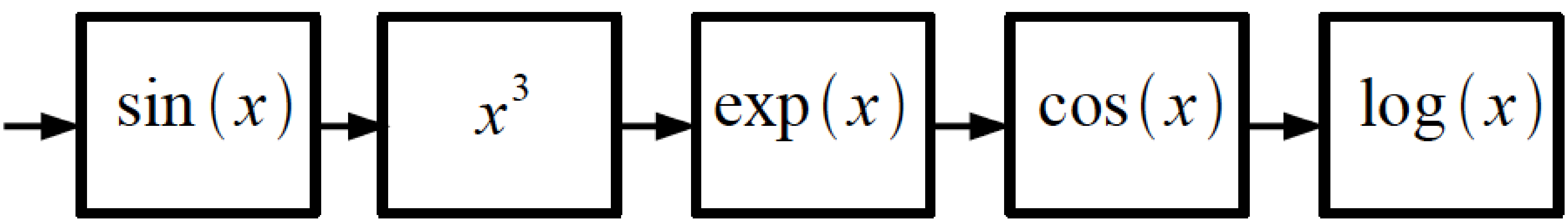
Compose into a

complicate function

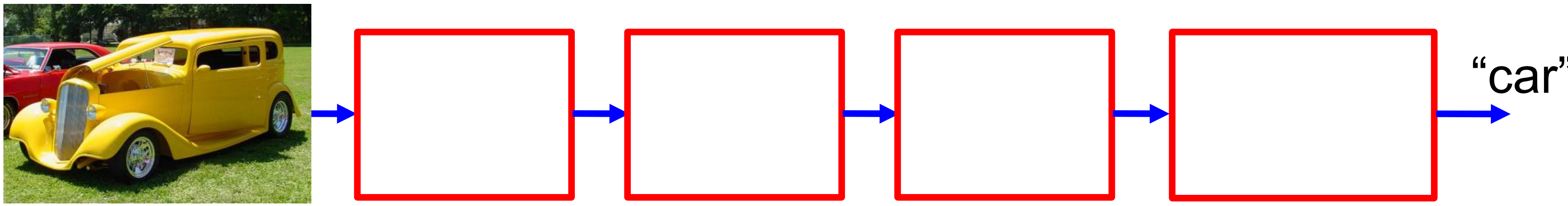
Idea 2: Compositions

- Deep Learning
- Grammar models
- Scattering transforms…

$$f(x) = \log(\cos(\exp(\sin^3(x))))$$

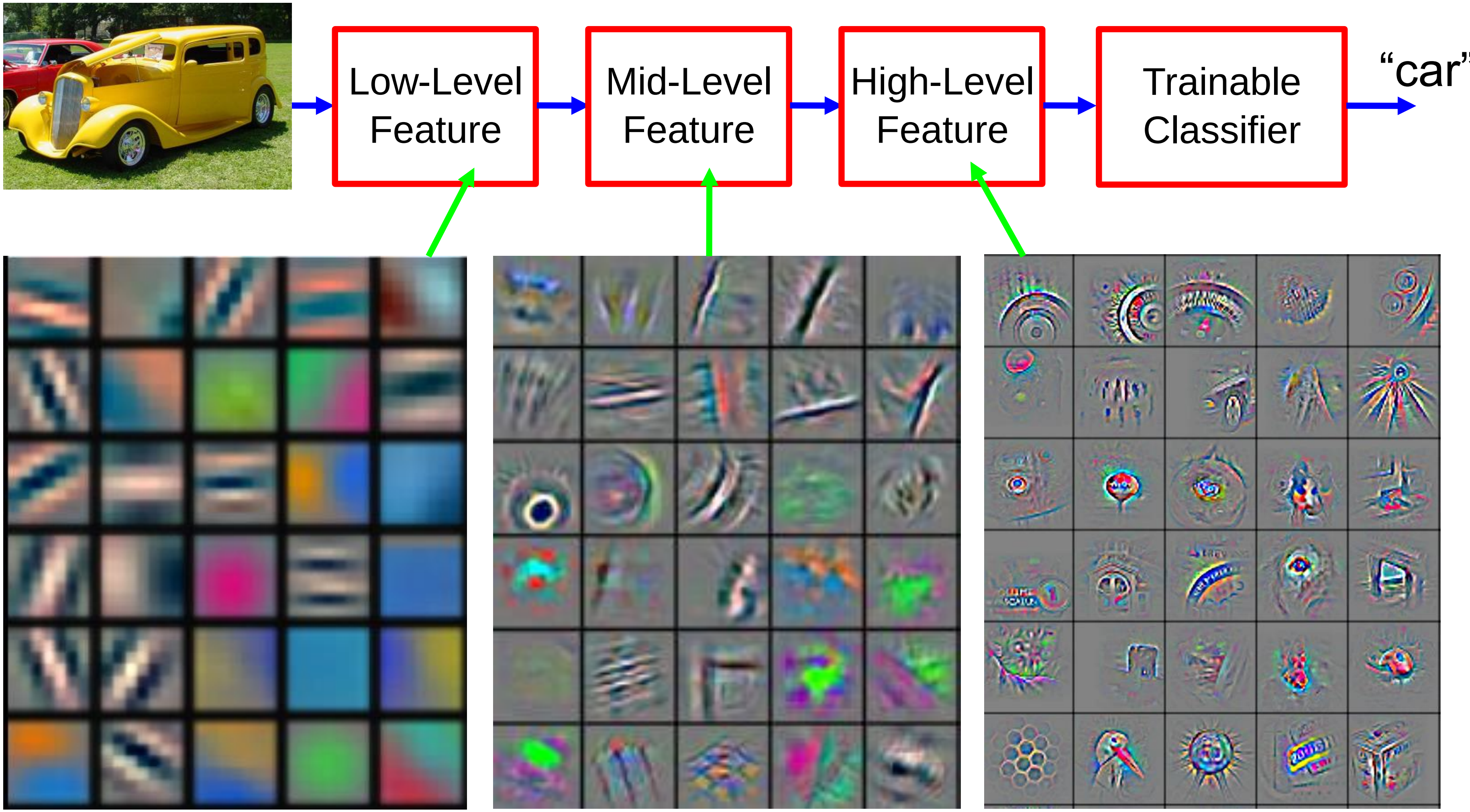


Deep Learning = Hierarchical Compositionality



Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

Deep Learning = Hierarchical Compositionality



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

So what *is* Deep (Machine) Learning?

- A few different ideas:

- (Hierarchical) Compositionality
  - Cascade of non-linear transformations
  - Multiple layers of representations

- End-to-End Learning
  - Learning (goal-driven) representations
  - Learning to feature extraction

- Distributed Representations
  - No single neuron “encodes” everything
  - Groups of neurons work together

Traditional Machine Learning

VISION

hand-crafted features

SIFT/HOG

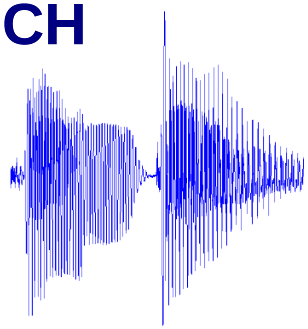
fixed

your favorite classifier

learned

“car”

SPEECH



hand-crafted features

MFCC

fixed

your favorite classifier

learned

\'d ē p\

NLP

This burrito place

is yummy and fun!

hand-crafted features

Bag-of-words

fixed

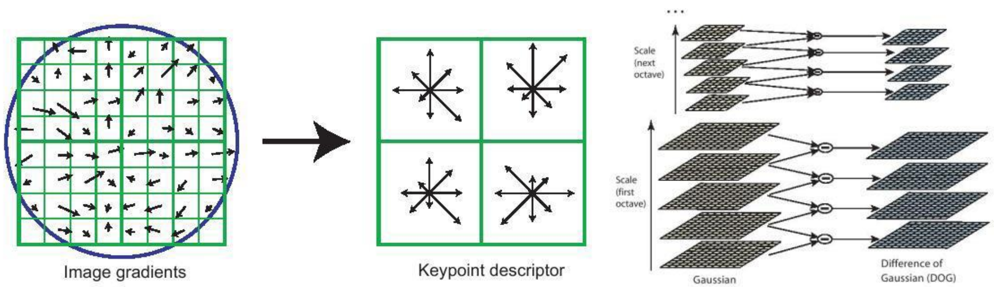
your favorite classifier

learned

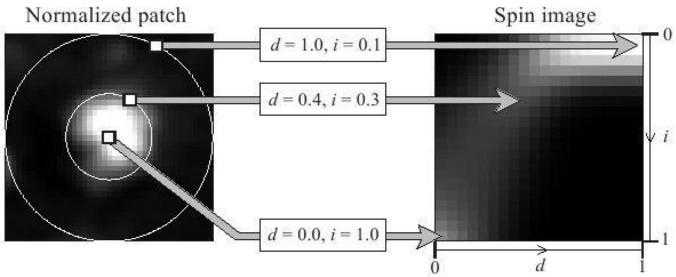
“+”

Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

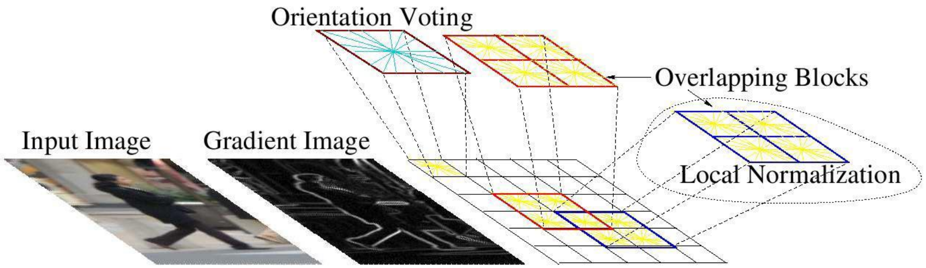
Feature Engineering



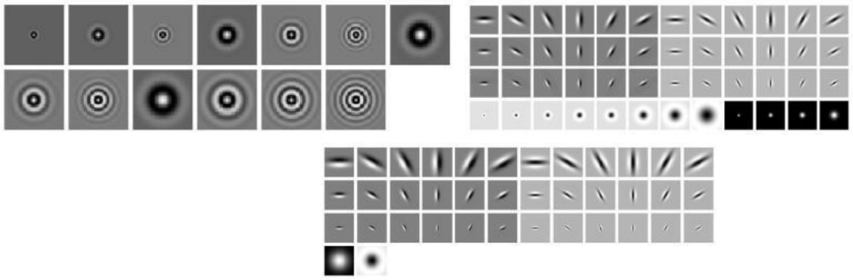
SIFT



Spin Images



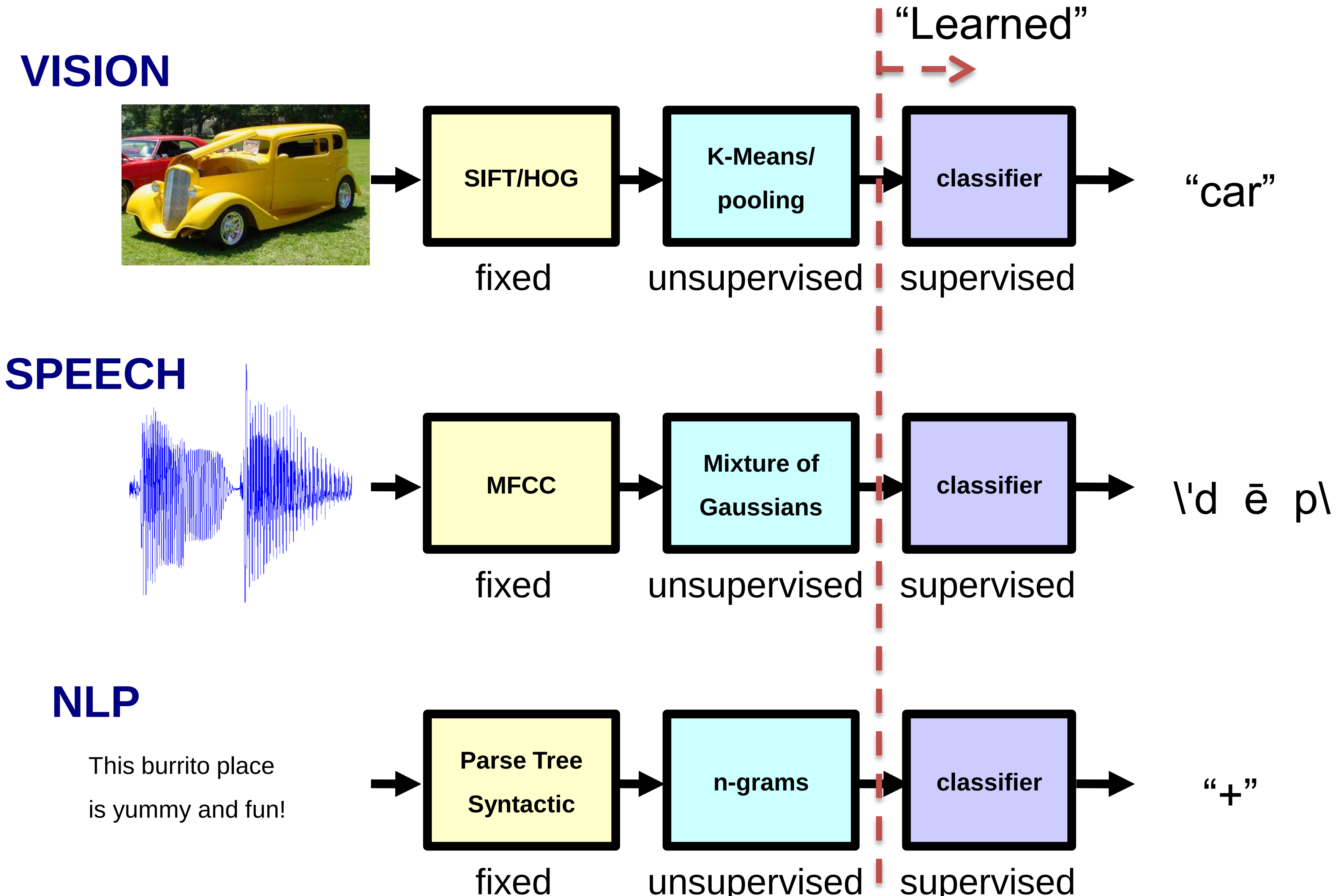
HoG



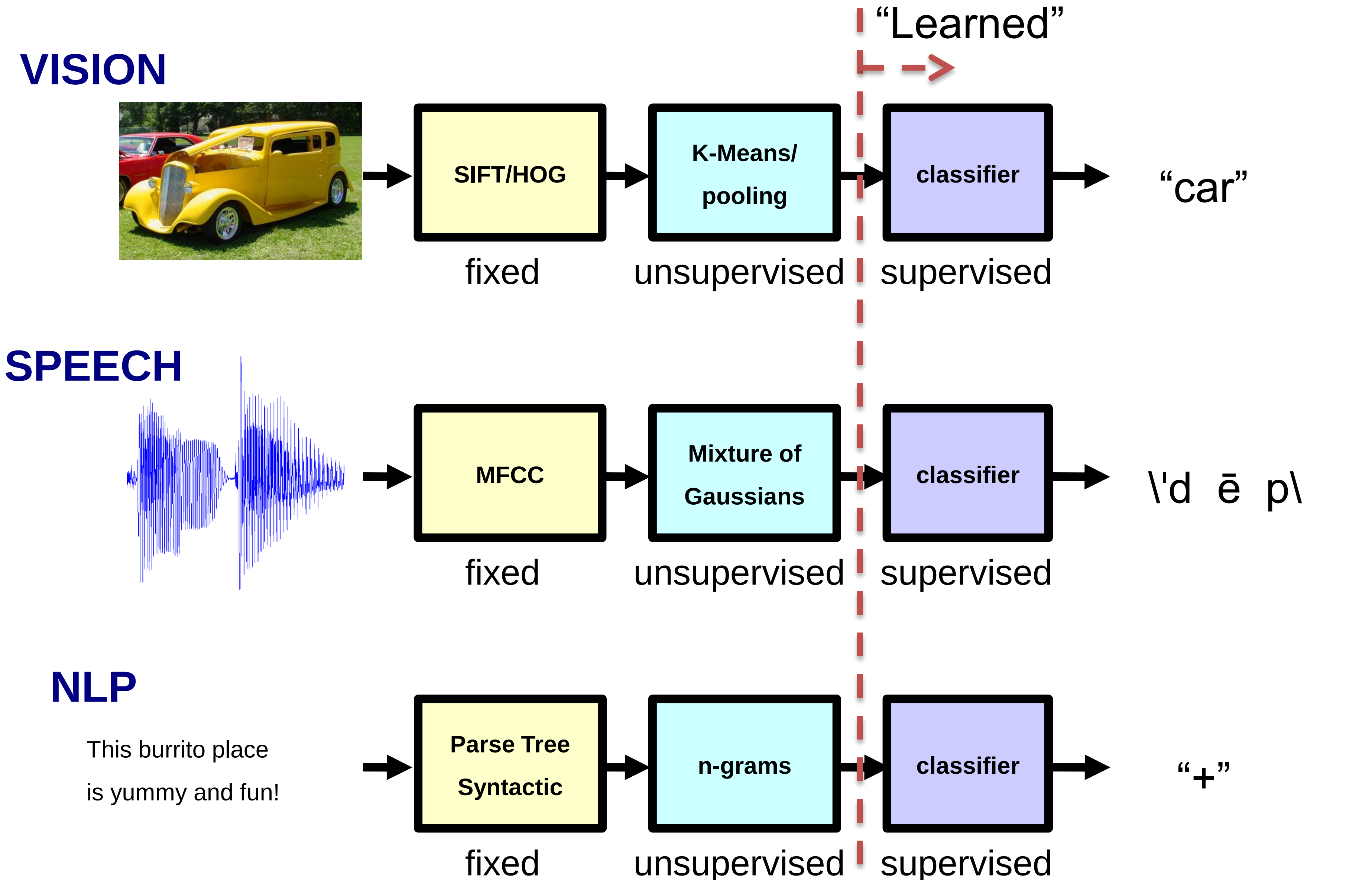
Textons

and many many more….

Traditional Machine Learning (more accurately)



Deep Learning = End-to-End Learning



“Shallow” vs Deep Learning

- “Shallow” models

hand-crafted Feature Extractor → “Simple” Trainable Classifier →

fixed — learned

- Deep models

Trainable Feature-Transform / Classifier → … → Trainable Feature-Transform / Classifier → Trainable Feature-Transform / Classifier →

Learned Internal Representations

Slide Credit: Marc'Aurelio Ranzato, Yann LeCun

So what *is* Deep (Machine) Learning?

- A few different ideas:

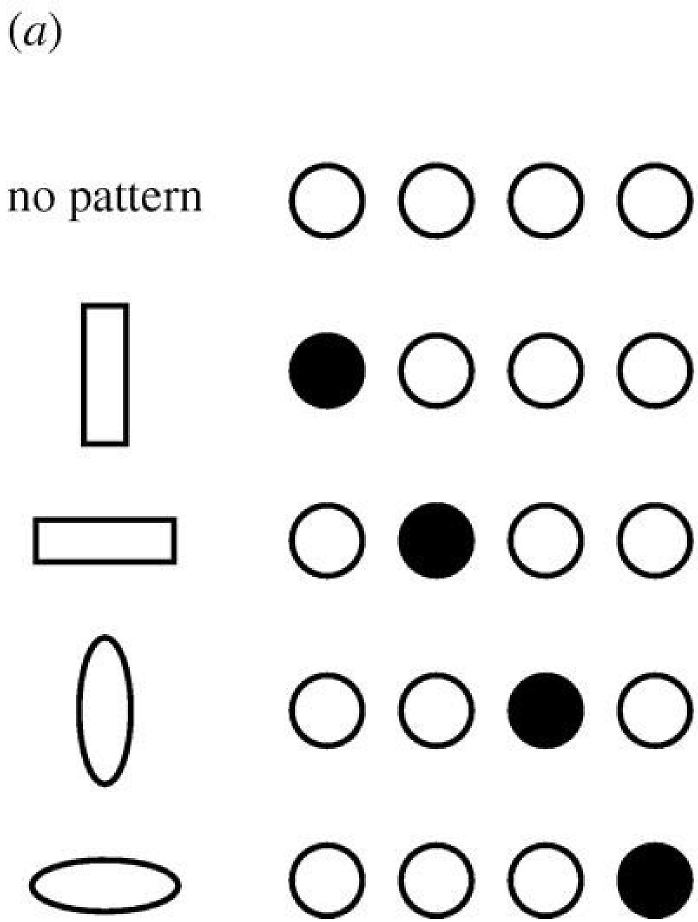
- (Hierarchical) Compositionality
  - Cascade of non-linear transformations
  - Multiple layers of representations

- End-to-End Learning
  - Learning (goal-driven) representations
  - Learning to feature extraction

- Distributed Representations
  - No single neuron “encodes” everything
  - Groups of neurons work together

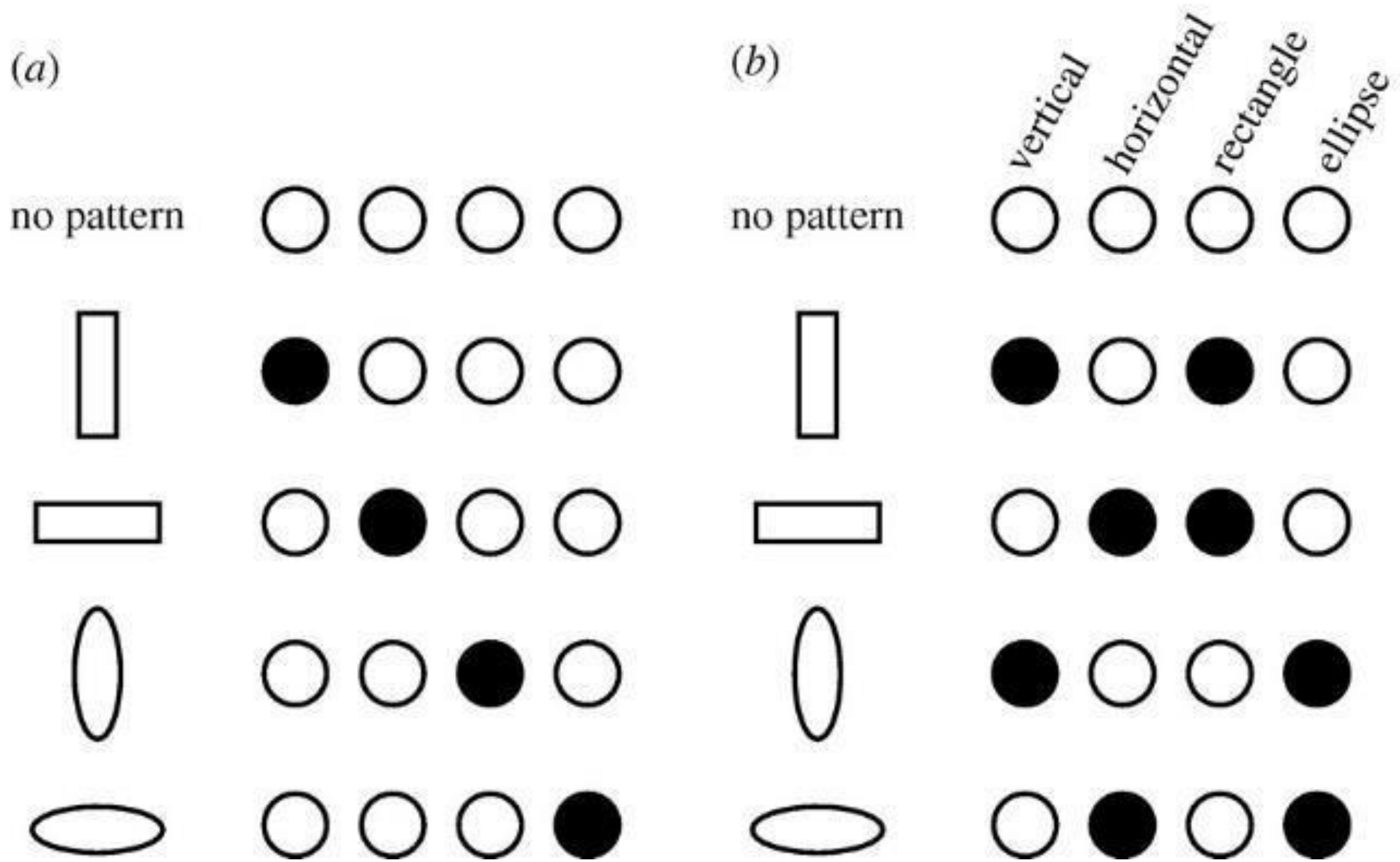
Distributed Representations Toy Example

- Local vs Distributed

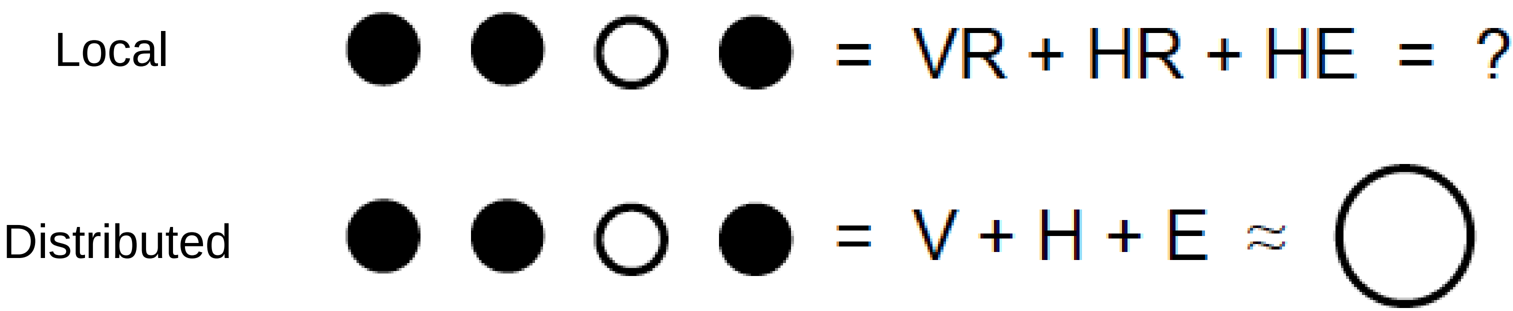


Distributed Representations Toy Example

- Can we interpret each dimension?

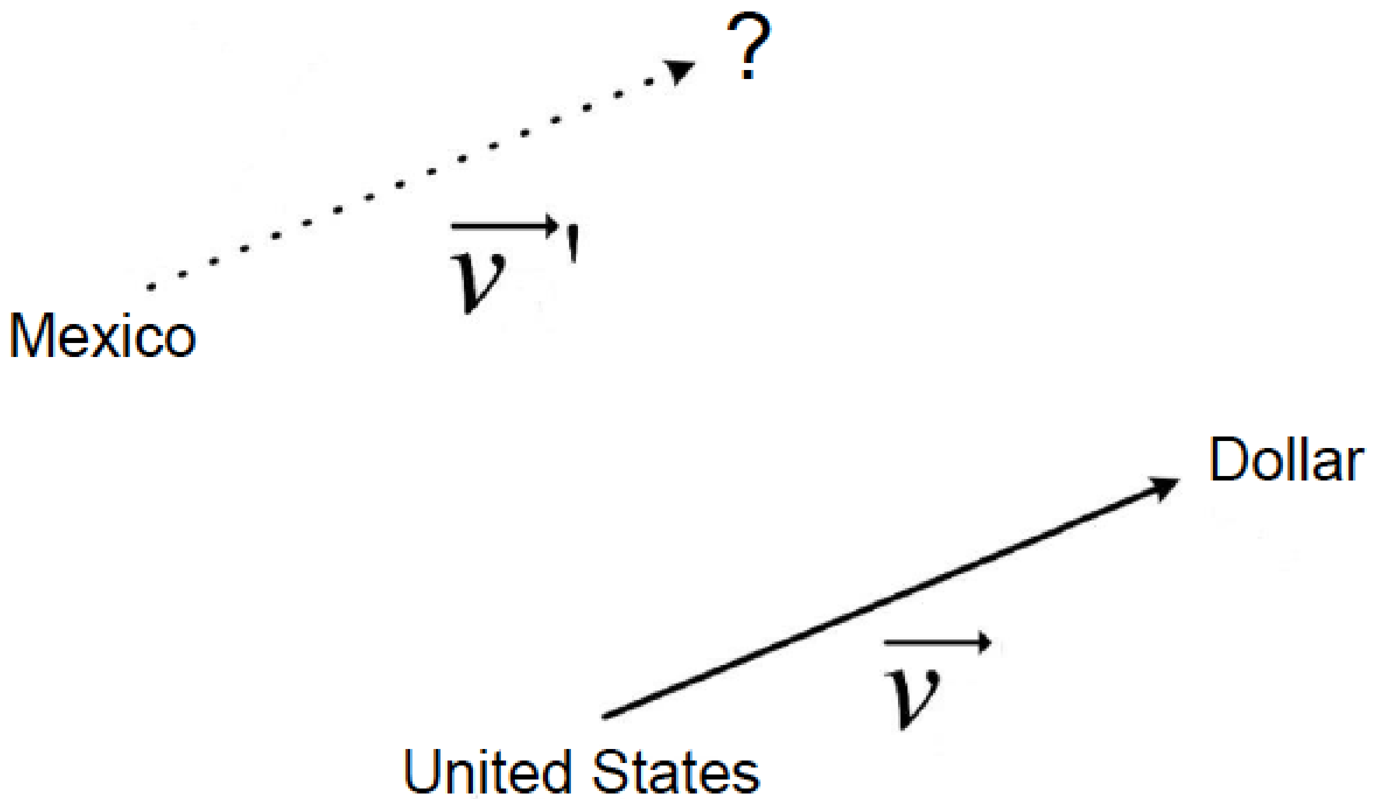


Power of distributed representations!



Power of distributed representations!

- United States:Dollar :: Mexico:?



ThisPlusThat.me

the matrix - thoughtful + dumb | **Search** | How it Works

mbiguated into *+1 the_matrix -1 thoughtful +1 dumb* in 0.0 seconds from ip-10-32-114-31

FILM, W FILM, NETFLIX TITLE,

Blade II

Blade II is a 2002 American vampire superhero action film base
Marvel Comics character Blade. It is the sequel of the first film a
part of the Blade film series. It was written by David S. Goyer, w
previous film. Guillermo del Toro was signed in to d...

Horror Film

Image Credit: http://insightdatascience.com/blog/thisplusthat_a_search_engine_that_lets_you_add_words_as_vectors.html

So what *is* Deep (Machine) Learning?

- A few different ideas:

- (Hierarchical) Compositionality
  - Cascade of non-linear transformations
  - Multiple layers of representations

- End-to-End Learning
  - Learning (goal-driven) representations
  - Learning to feature extraction

- Distributed Representations
  - No single neuron “encodes” everything
  - Groups of neurons work together

Benefits of Deep/Representation Learning

- (Usually) Better Performance
  - Caveats: given enough data, similar train-test distributions, non-adversarial evaluation, etc, etc.

- New domains without "experts"
  - RGBD/Lidar
  - Multi-spectral data
  - Gene-expression data
  - Unclear how to hand-engineer

“Expert” intuitions can be misleading

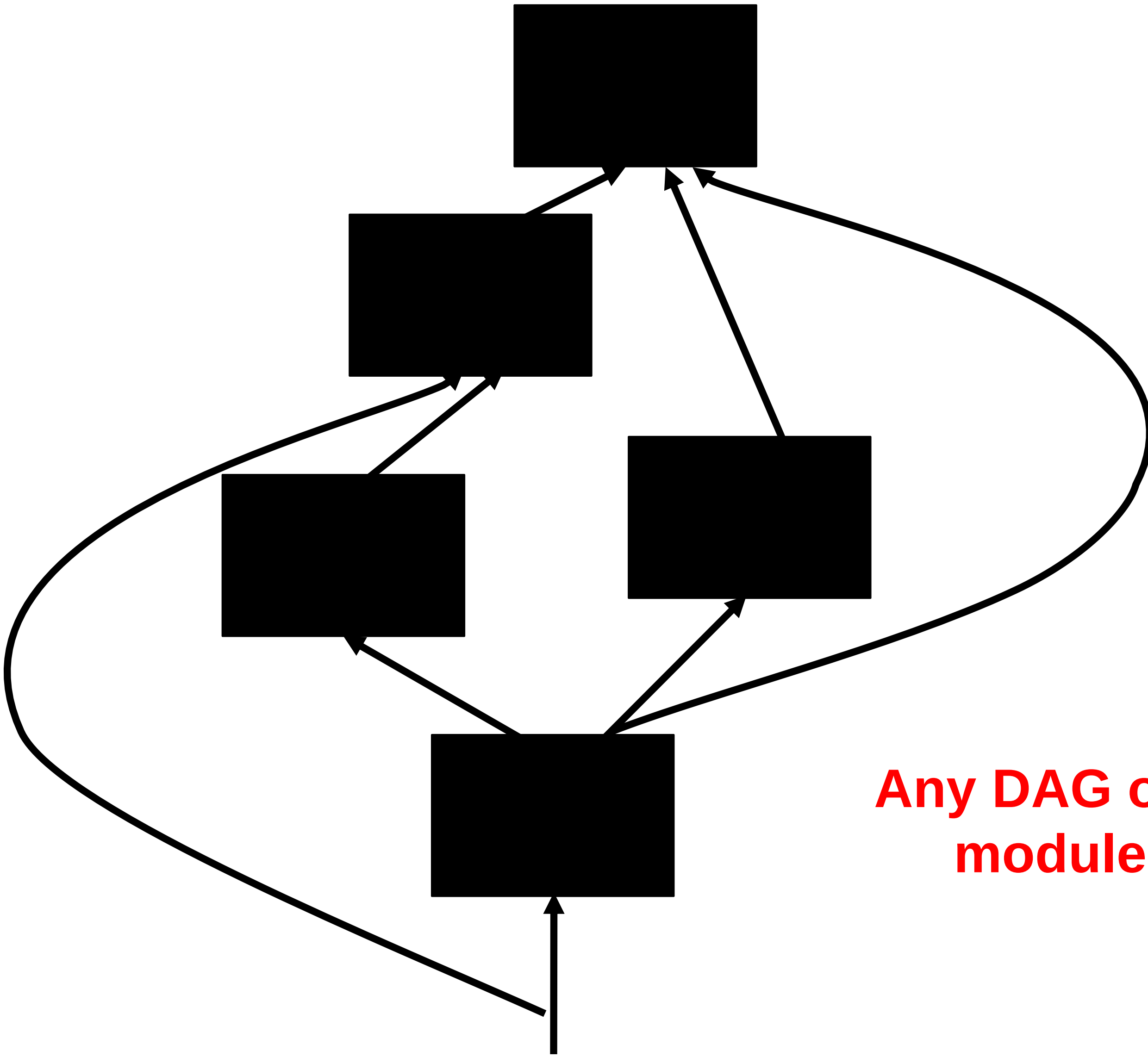
- *“Every time I fire a linguist, the performance of our speech recognition system goes up”*
  - Fred Jelinik, IBM ’98

- *“Because gradient descent is better than you”*
  - *Yann LeCun,* CVPR ‘13

Benefits of Deep/Representation Learning

- Modularity!
- Plug and play architectures!

Differentiable Computation Graph



Any DAG of differentialble modules is allowed!

Problems with Deep Learning

- Problem#1: Lack of a formal understanding
  - Non-Convex! Non-Convex! Non-Convex!
    - Depth>=3: most losses non-convex in parameters
  - Worse still, existing intuitions from classical statistical learning theory don't seem to carry over.
  - Theoretically, we are stumbling in the dark here

- Standard response #1
  - "Yes, but this just means there's new theory to be constructed"
  - "All interesting learning problems are non-convex"
    - For example, human learning
      - Order matters → wave hands → non-convexity

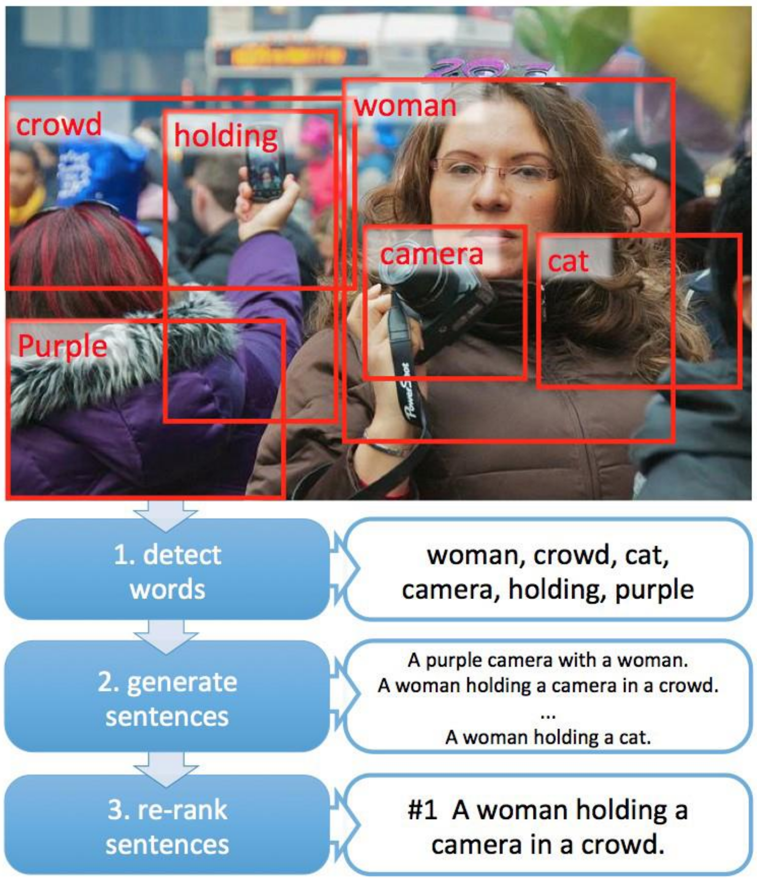
- Standard response #2
  - "Yes, but it often works!"

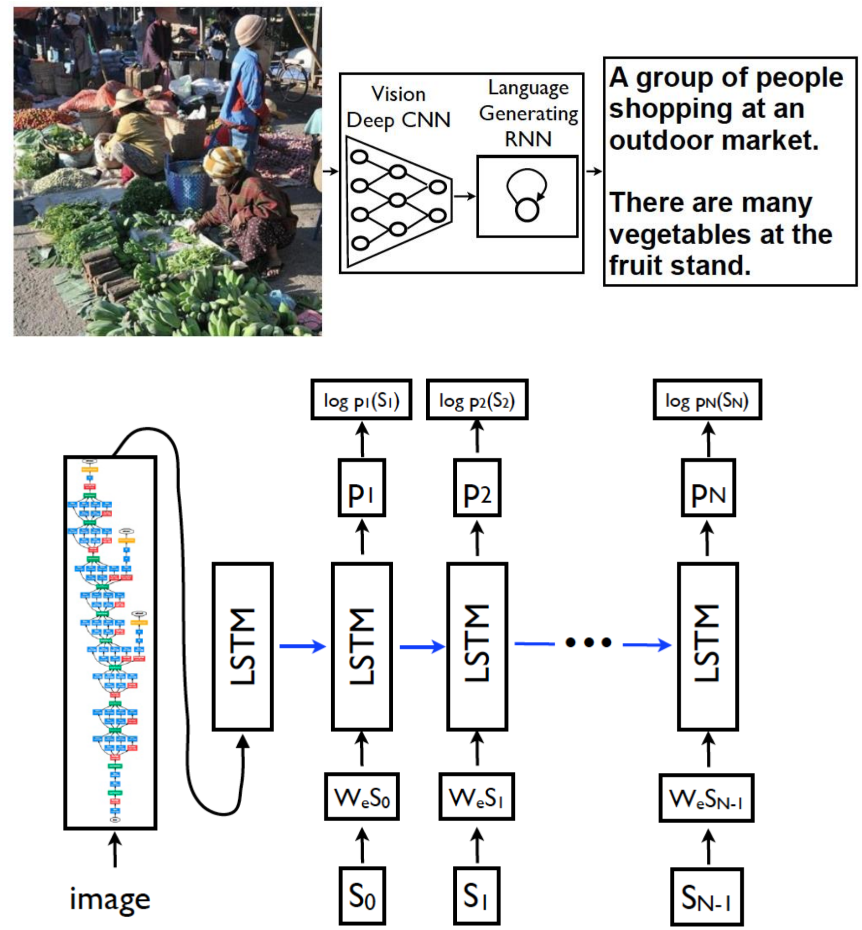
Problems with Deep Learning

- Problem#2: Lack of interpretability
  - Hard to track down what's failing
  - Pipeline systems have "oracle" performances at each step
  - In end-to-end systems, it's hard to know why things are not working

Problems with Deep Learning

- Problem#2: Lack of interpretability





[Fang et al. CVPR15]

[Vinyals et al. CVPR15]

Pipeline

End-to-End

Problems with Deep Learning

- Problem#2: Lack of interpretability
  - Hard to track down what's failing
  - Pipeline systems have "oracle" performances at each step
  - In end-to-end systems, it's hard to know why things are not working

- Standard response #1
  - Tricks of the trade: visualize features, add losses at different layers, pre-train to avoid degenerate initializations…
  - "We're working on it"

- Standard response #2
  - "Yes, but it often works!"

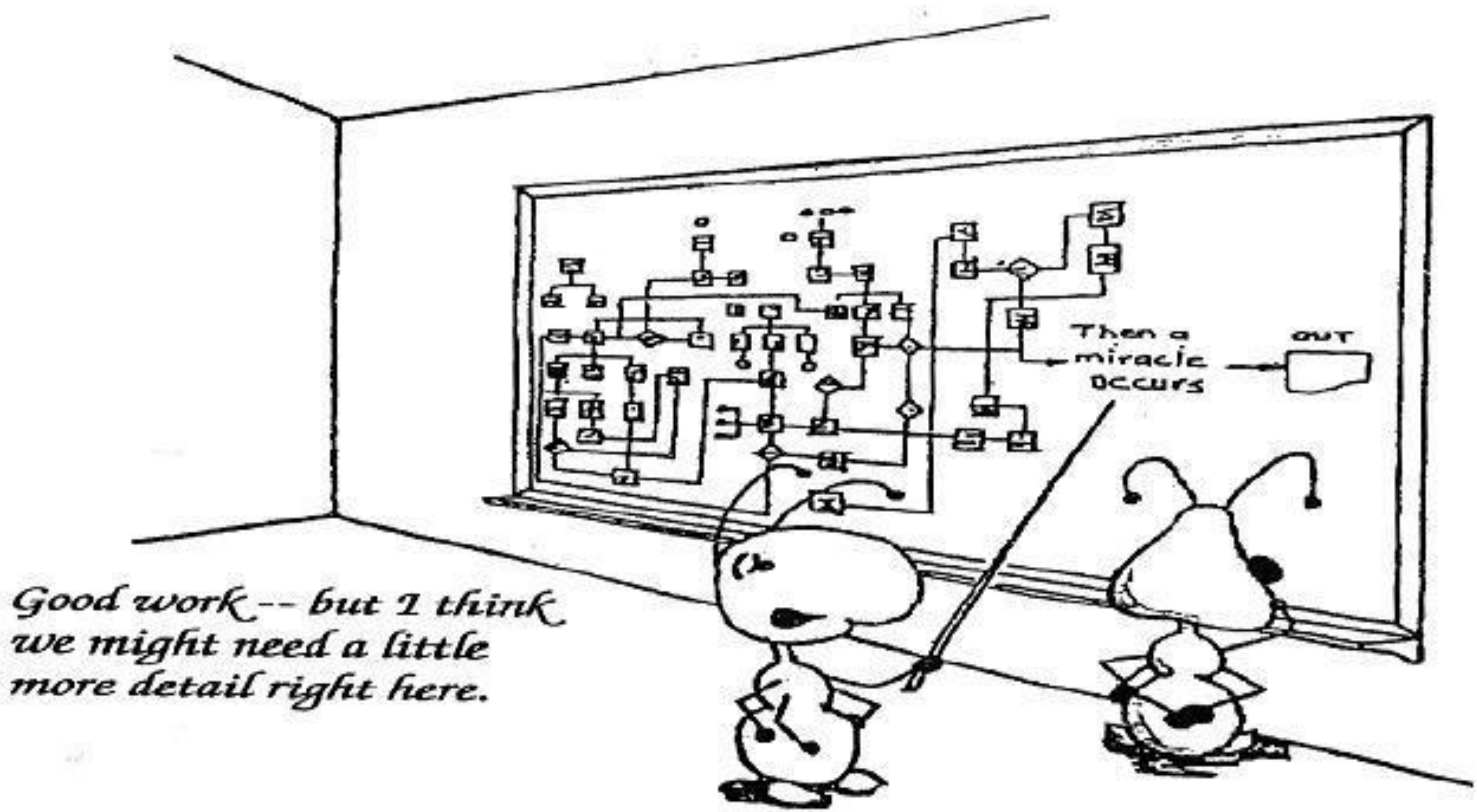
Problems with Deep Learning

- Problem#3: Lack of easy reproducibility
  - Direct consequence of stochasticity & non-convexity
    - different initializations → different local minima

- Standard response #1
  - It's getting much better
  - Standard toolkits/libraries/frameworks now available
  - PyTorch, TensorFlow, MxNet…

- Standard response #2
  - "Yes, but it often works!"

Yes it works, but how?



Three scientists who kickstarted an AI revolution by studying the learning abilities of large artificial neural networks have been awarded the most prestigious accolade in computer science: the $1 million Turing Award.

www.nature.com › review articles - Dịch trang này

Deep learning | Nature

27 thg 5, 2015 - New learning algorithms and architectures that are currently being developed for **deep neural networks** will only accelerate this progress.

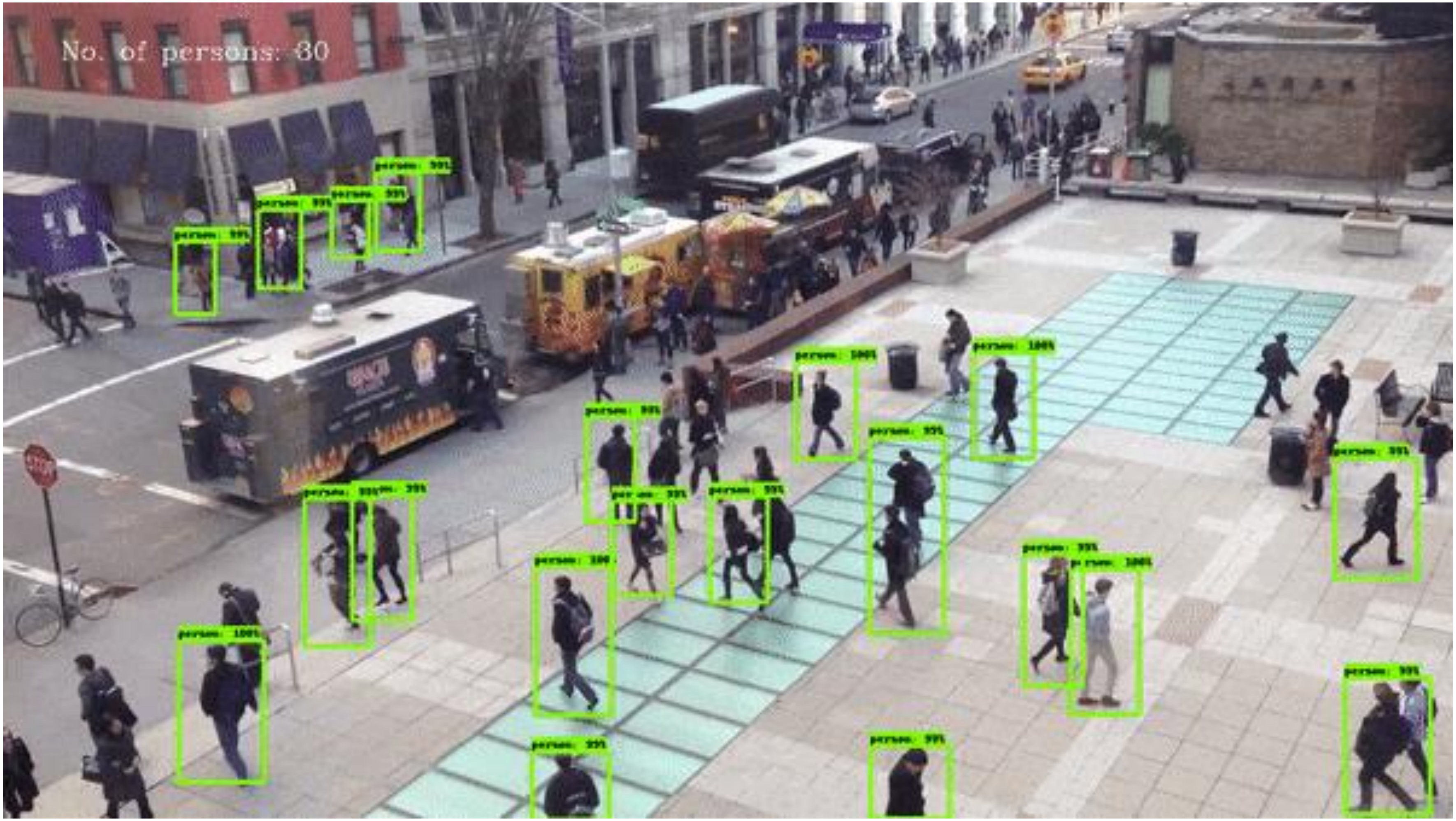
viết bởi Y LeCun - 2015 - Trích dẫn 27896 bài viết - Bài viết có liên quan

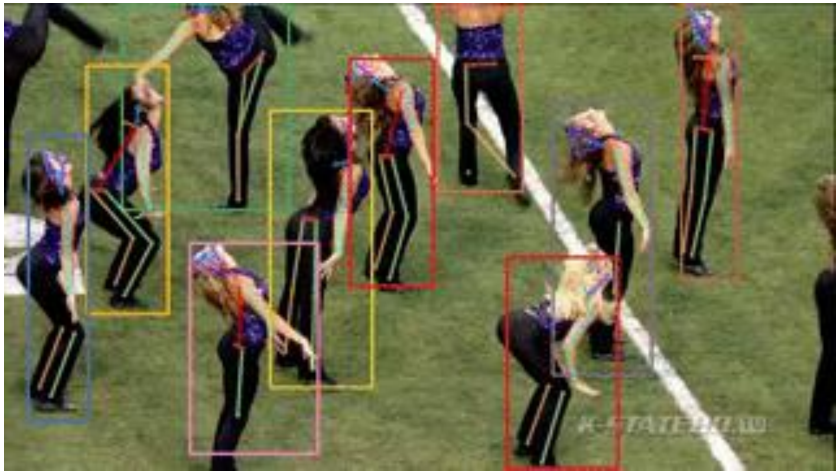
Deep Learning

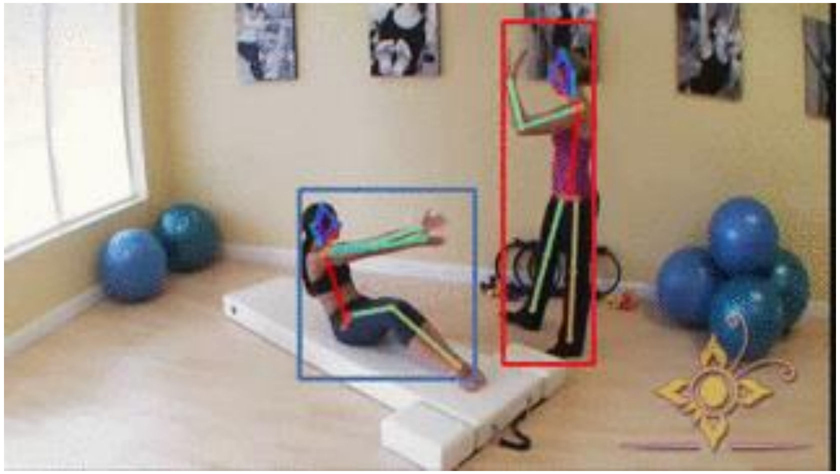
- Deep Supervised Learning
  - FNN (Feed - Forward Neural Network)
  - CNN (Convolutional Neural Network)
  - RNN (Recurrent Neural Network)
- Deep Unsupervised Learning
- Deep Semisupervised Learning
- Deep Reinforcement Learning

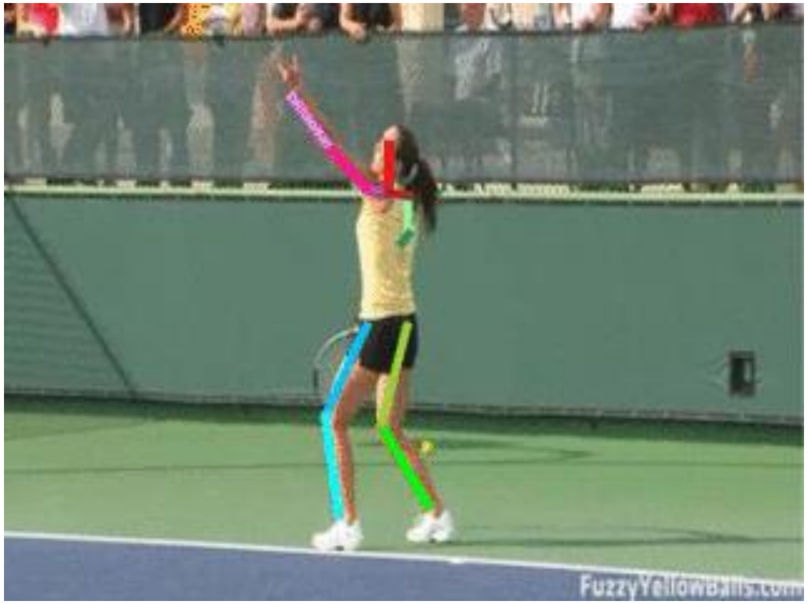

flickr
You Tube
.com
yelp
flickr
You Tube
yelp
flickr
.com
flickr
yelp
You Tube
yelp
.com
yelp
You Tube
.com
.com
yelp
You Tube
.com

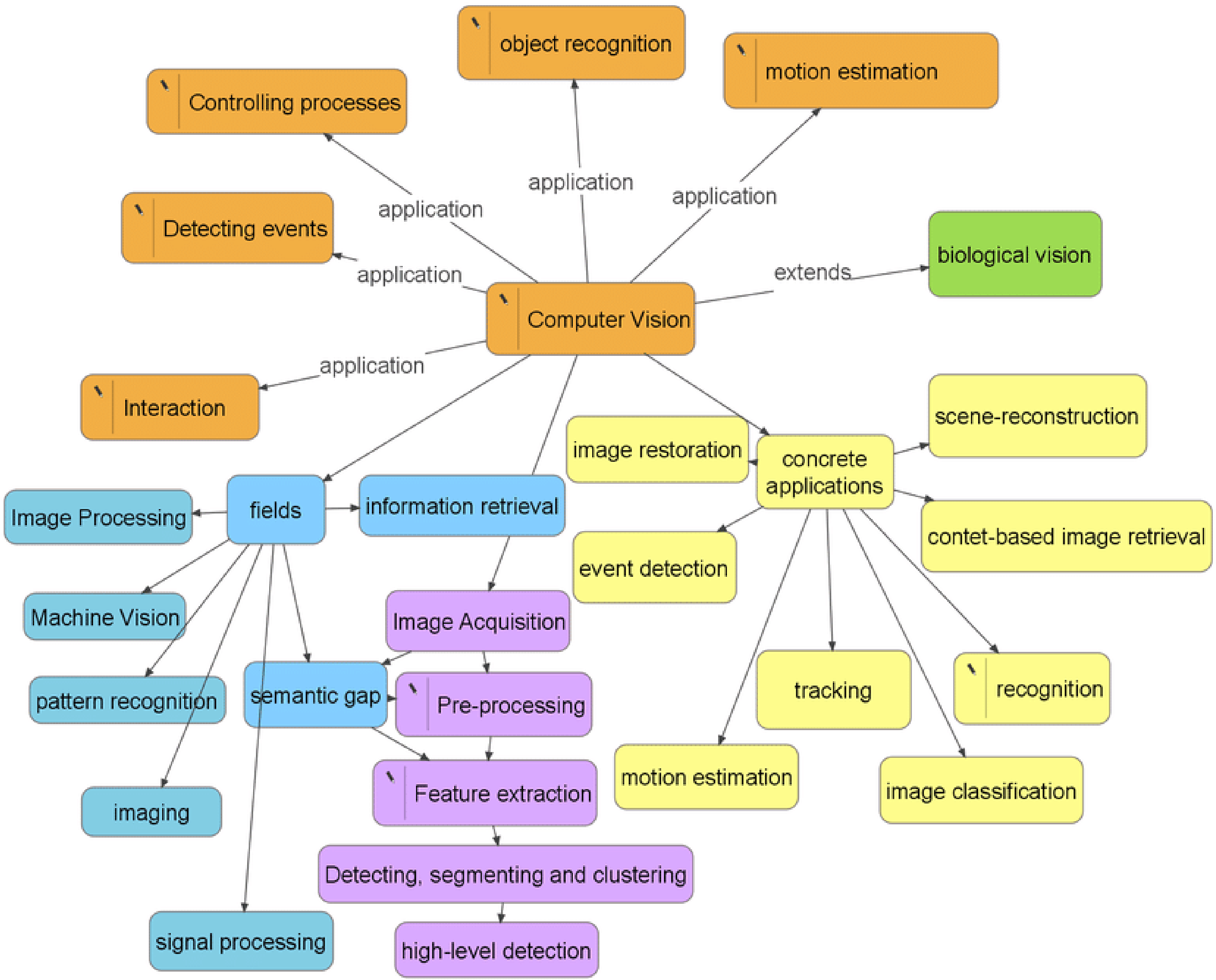

GUIDE
JULY 31


No. of persons: 30






FuzzyYellowBalls.com


object recognition
motion estimation
Controlling processes
application
application
application
Detecting events
application
biological vision
extends
Computer Vision
application
Interaction
scene-reconstruction
image restoration
concrete applications
Image Processing
fields
information retrieval
contet-based image retrieval
event detection
Machine Vision
Image Acquisition
tracking
recognition
pattern recognition
semantic gap
Pre-processing
motion estimation
image classification
imaging
Feature extraction
Detecting, segmenting and clustering
signal processing
high-level detection

Voice Assistants

https://www.softwaretestinghelp.com/

Framework Suggestion

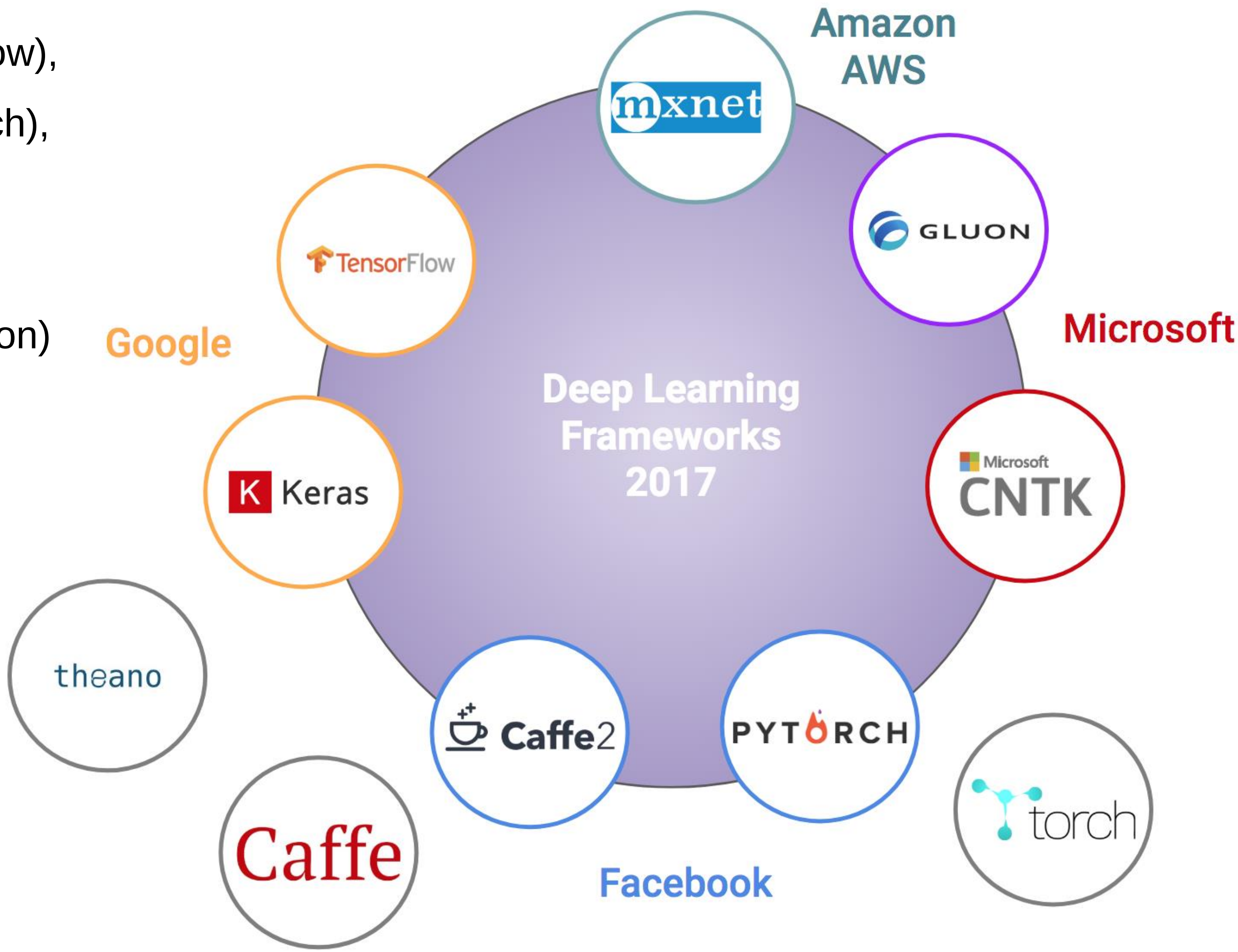
Google (Keras, TensorFlow),

Facebook (Caffe2, Pytorch),

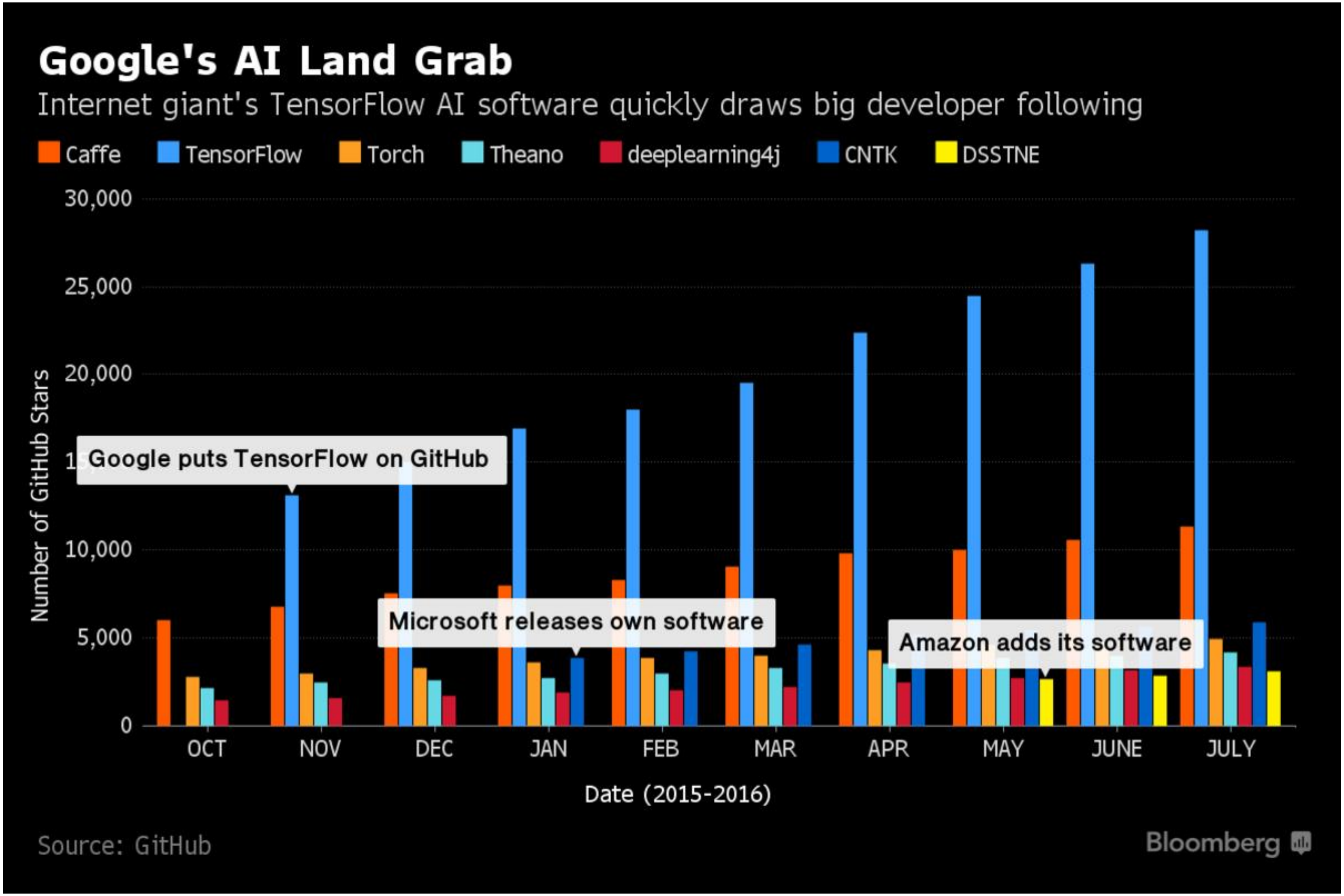
Microsoft (CNTK),

Amazon (Mxnet),

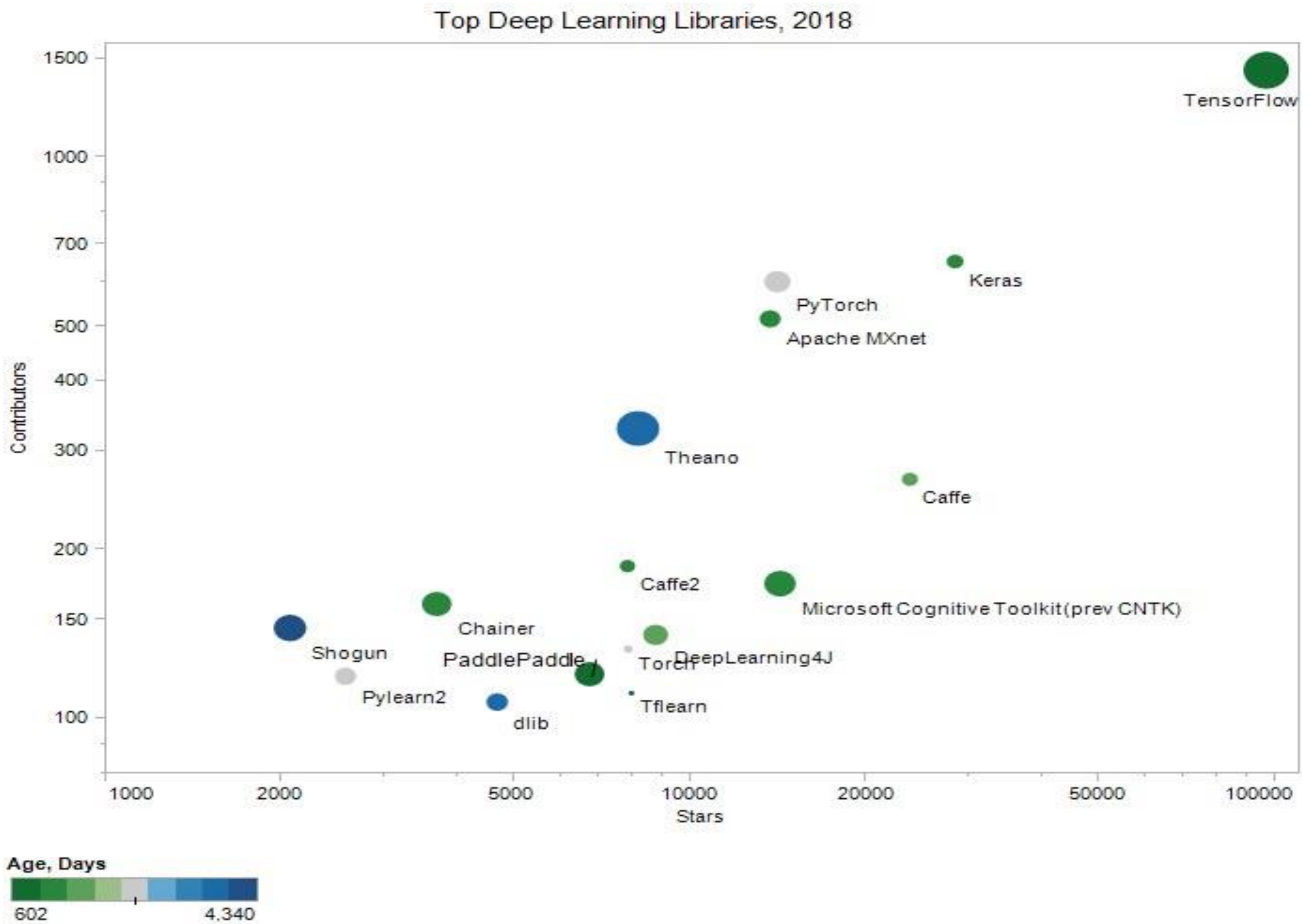
Microsoft & Amazon (Gluon)



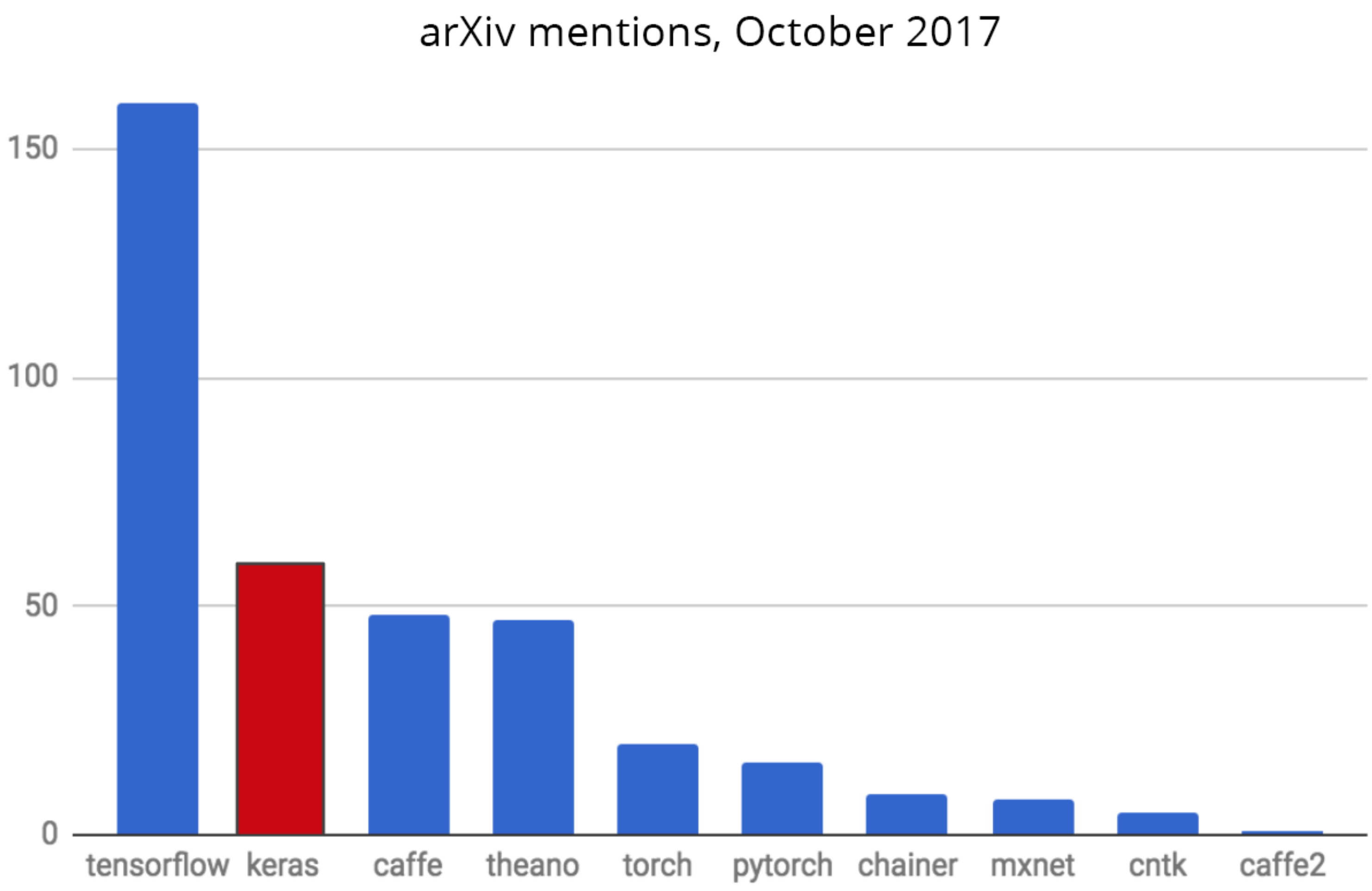
https://towardsdatascience.com/battle-of-the-deep-learning-frameworks-part-i-cff0e3841750



https://blog.paperspace.com/which-ml-framework-should-i-use/



https://blog.paperspace.com/which-ml-framework-should-i-use/



https://keras.io/why_keras/

Framework Suggestion

In general, a good deep learning library should have the following characteristics:

- ✓ Supports computation with GPUs and distributed systems. This is paramount because training deep learning models requires very strong computation.
- ✓ Support for popular programming languages: C / C ++, Python, Java, R, ...
- ✓ Can run on multiple operating systems.
- ✓ The time from concept to building and training of models is short.
- ✓ Can run on browsers and mobile devices.
- ✓ Capable of helping programmers intervene deeply in the model and create complex models.
- ✓ Contains many model zoo, ie popular deep learning models that have been trained.
- ✓ Support automatic backpropagation calculation.
- ✓ There is a large community of questions and answers.