
What, so everyone's supposed to sleep every single night now.
You realize that nighttime makes up half of all time?
[adult swim]


(Hinton 2006)
How the brain works: deep learning

Topics I'll be presenting today

1. How to start with Deep Learning?
 - Courses
 - Coding Packages
 - Logistics
 - Simulation Environments

2. Computer Vision

 - ImageNet Challenge
 - YOLO
 - Visual attention and saliency
 - Image Captioning
 - Feature Visualization for Deep Networks

Topics I'll be presenting today

3. Generative Models

- Generative Adversarial Networks

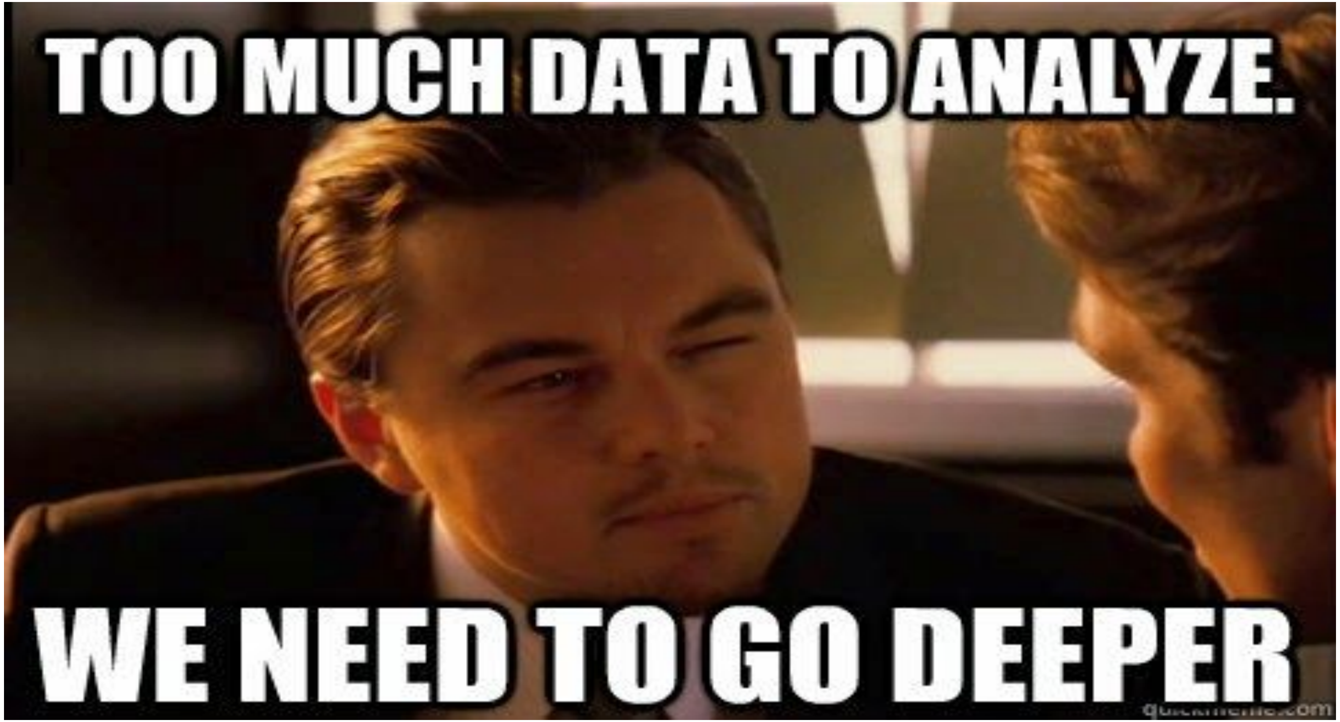
4. Deep Learning Models and Architectural breakthroughs

- Capsule networks

5. Reinforcement Learning

- Deep RL
- Inverse Reinforcement Learning (IRL)

And LOTS of this!



HOW TO START WITH DEEP LEARNING

COURSES:

[Coursera] Neural Networks for Machine Learning — Geoffrey Hinton 2016

Deeplearning.ai course playlist - Andrew Ng - 4 courses

CS 294: Deep Reinforcement Learning, Fall 2017

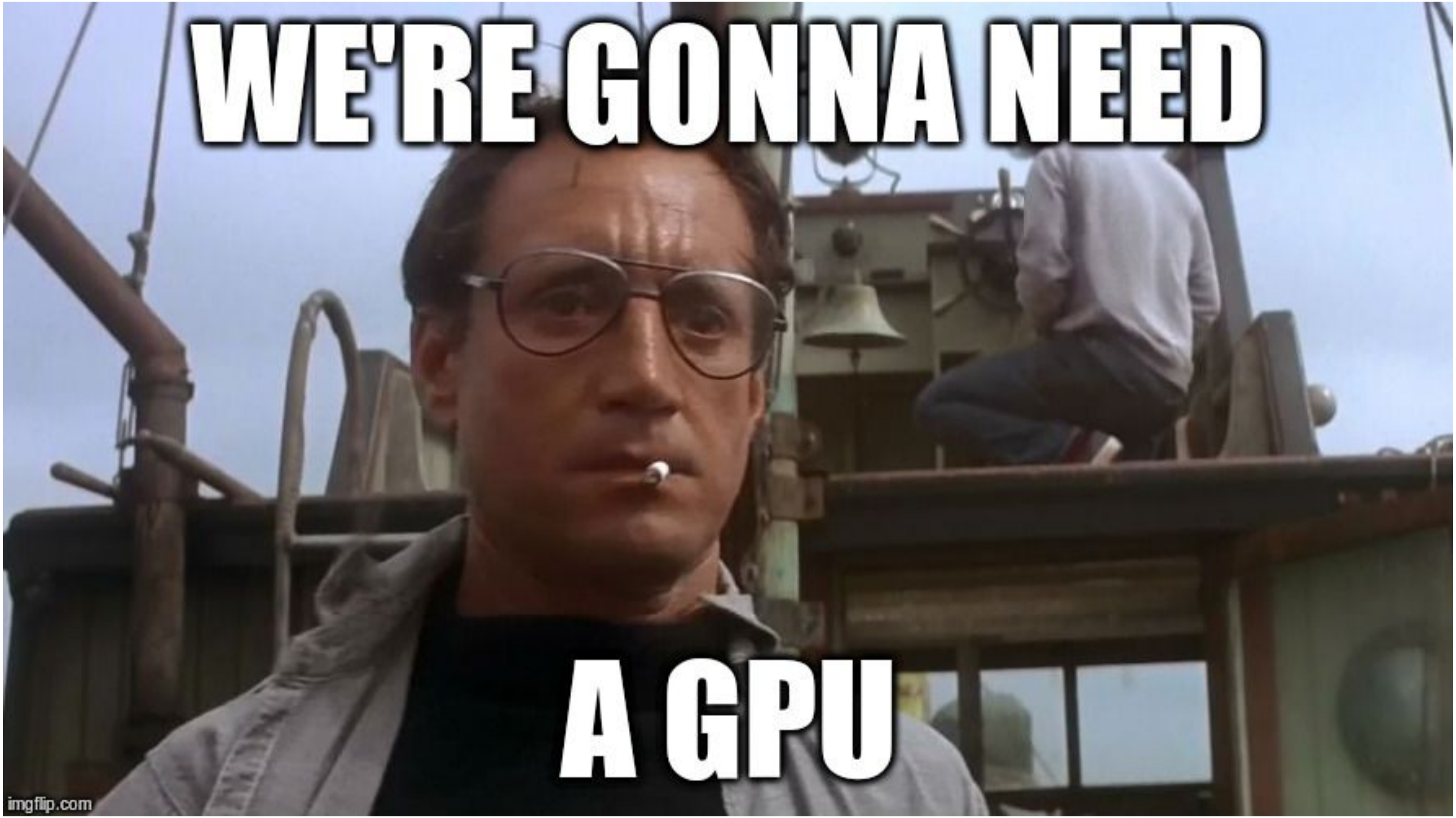
Papers:

Arxiv Sanity Preserver

HOW TO START WITH DEEP LEARNING

CODING PLATFORMS:

1. Tensorflow - static computation graphs, resources
2. Keras - Front end library
3. PyTorch - DCG, modular,imperative programming
4. Caffe - C++ based


WE'RE GONNA NEED
A GPU
imgflip.com

HOW TO START WITH DEEP LEARNING

LOGISTICS:

Cloud: Cloud AI - Google Cloud

AWS Deep Learning in the Cloud

Desktop: Deep learning setup for Ubuntu 16.04 -- with CUDA installation

Datasets: Kaggle

Datasets @ Deeplearning.net

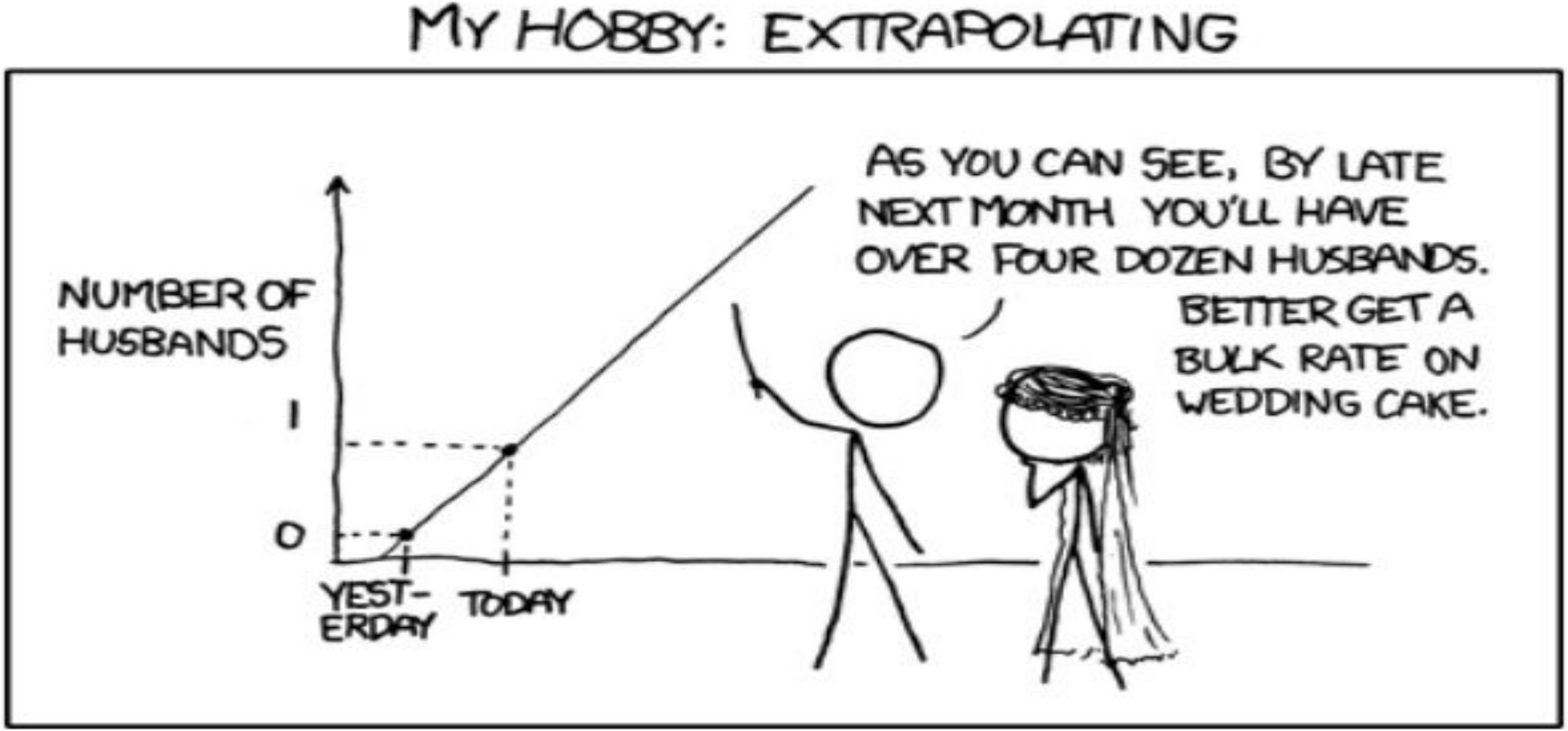
Open Data for Deep Learning

HOW TO START WITH DEEP LEARNING

SIMULATION ENVIRONMENTS (ESPECIALLY FOR REINFORCEMENT LEARNING PROBLEMS)

1. OpenAI Gym
2. Carla - Open-source simulator for autonomous driving research
3. AirSim - Open source simulator based on Unreal Engine for autonomous vehicles from Microsoft AI & Research

"Generalizing"



Research groups to look out for

1. Google Brain and DeepMind
2. OpenAI
3. FAIR
4. UToronto Machine Learning
5. MILA
6. Baidu AI

Leading deep learning researchers







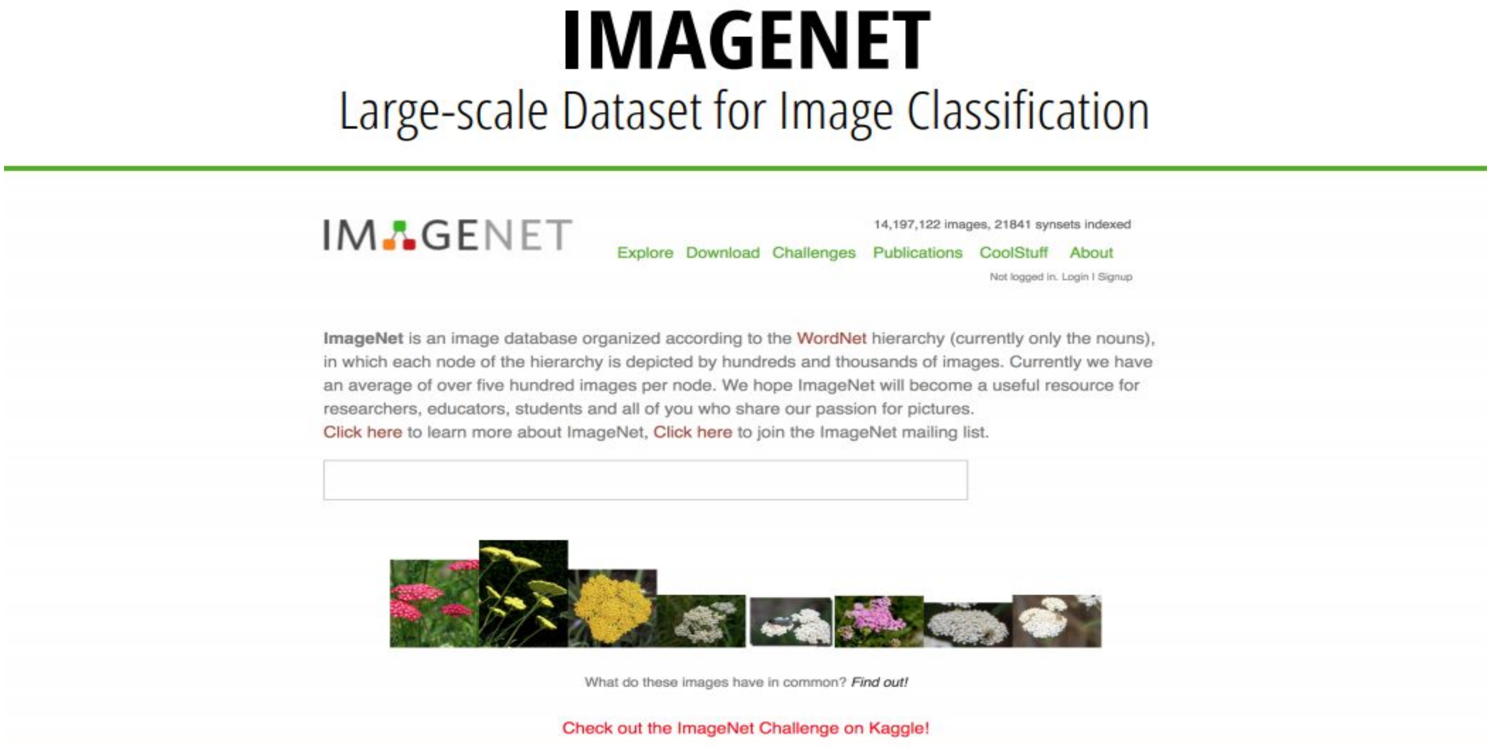


SUPPORT VECTOR MACHINES
REPEA POWE LAWS
END DUALITy GAP
Map Reduce
Map Reuse
Map Recycle
BAYSIANS AGAINST DISCRIMINATION
FREE VARIABLES!
BAN GENETIC ALGORITHMS

Let's move to V-I-S-I-O-N

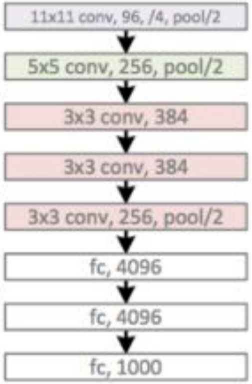
ImageNet

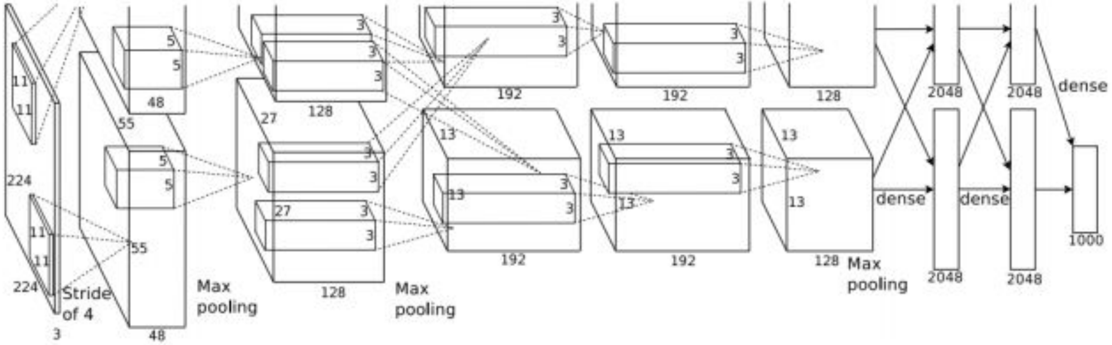
Challenge



CONVNET ARCHITECTURES

AlexNet (2012): The Return of Convolutional Neural Networks





Repopularized convolutional networks by winning the Imagenet Challenge in 2012 [Krizhevsky et al. 2012].

Error of 16% vs 26% for second place!

7 Hidden Layers, 650,000 neurons, 60 Million parameters


COMPUTER VISION
IS EVERYWHERE

Yolo - you only look once

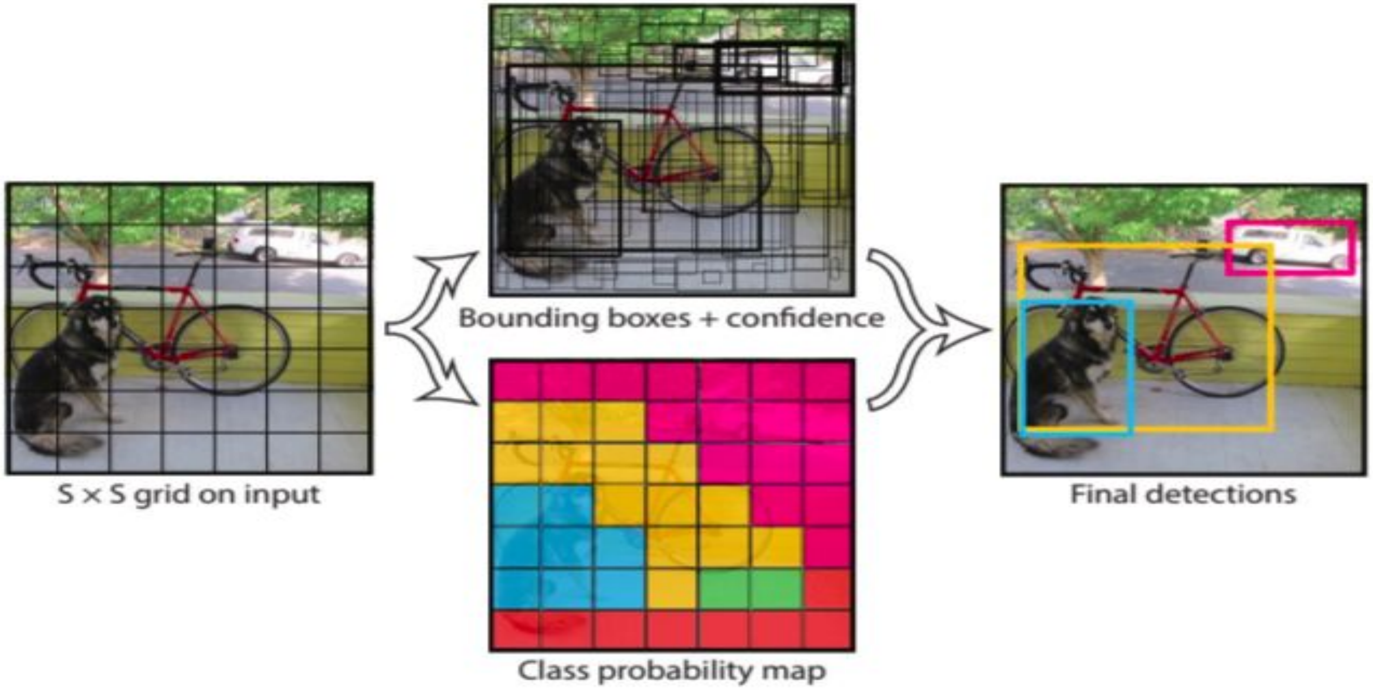
YOLO

- Different types of anchor-boxes
- Divide images into a grid and input each cell to the NN
- Make predictions (during testing) for each of them
- Use non-max suppression for finding outputs

Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.

Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." *arXiv preprint arXiv:1612.08242* (2016).

Yolo - framework



Yolo - neural net arch

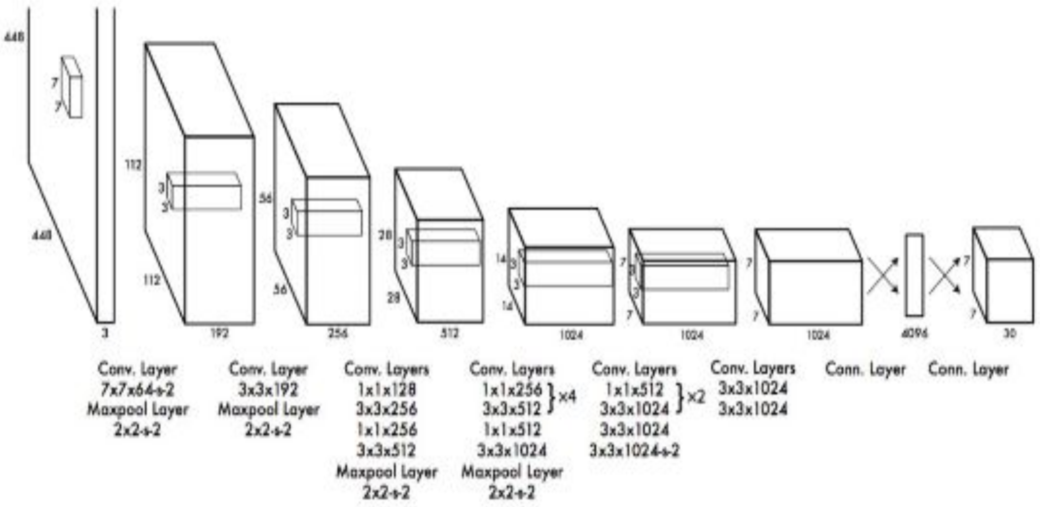


Figure 3: The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

Visual attention and saliency a.k.a where to look?

Liu, Nian, et al. "Predicting eye fixations using convolutional neural networks." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.

Predicting eye fixations using CNN - CVPR 2015

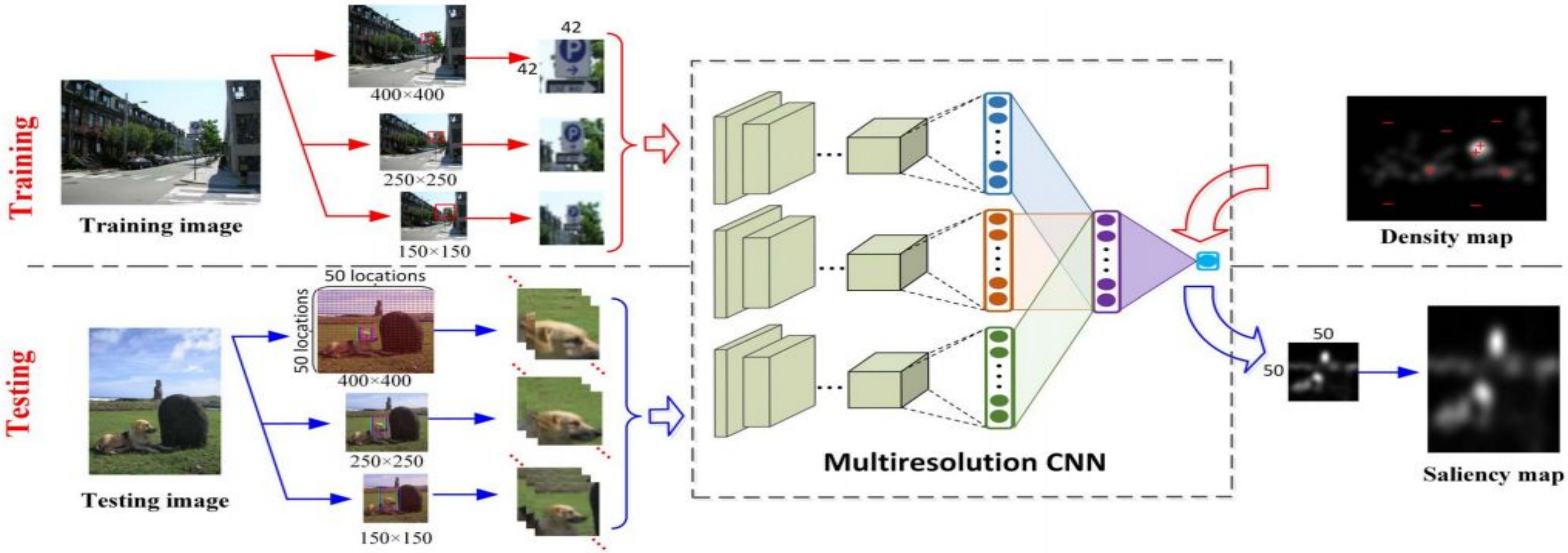


Figure 1: Diagram of our Mr-CNN based model. First, the given image is rescaled to three scales, i.e. 150×150, 250×250 and 400×400, then 42×42 sized image regions with the same center locations are extracted from the rescaled image duplicates as inputs to the Mr-CNN. We extract fixation and non-fixation image regions to train the Mr-CNN. When testing, we just evenly sample 50×50 locations per image to estimate their saliency values to reduce computation cost. The obtained down-sampled saliency map is rescaled to the original size to achieve the final saliency map.

Image captioning

man in black shirt is playing guitar.

construction worker in orange safety vest is working on road.

two young girls are playing with lego toy.

boy is doing backflip on wakeboard.

(from Andrej Karpathy, Li Fei-Fei, Deep Visual-Semantic Alignments for Generating Image Description, CVPR, 2015.)

Explain Images with Multimodal RNNs - Baidu Research

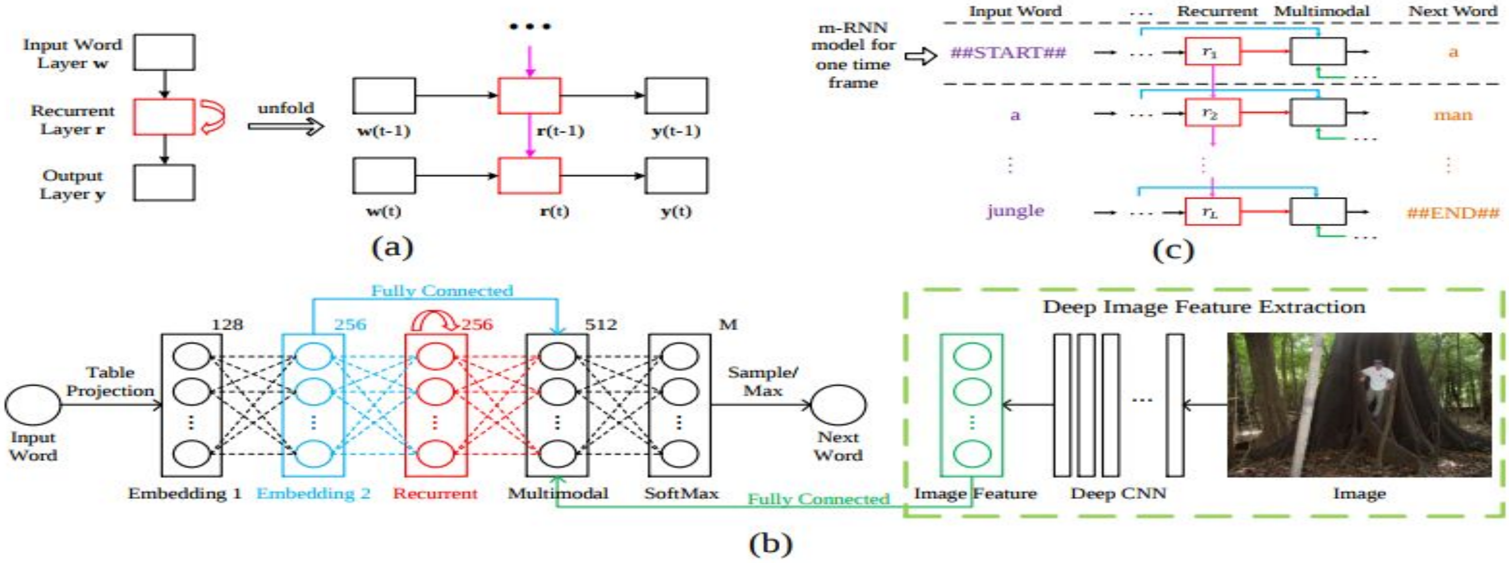


Figure 2: Illustration of the simple Recurrent Neural Network (RNN) and our multimodal Recurrent Neural Network (m-RNN) architecture. (a). The simple RNN. (b). Our m-RNN model. The input of our model is an image and its corresponding sentences (e.g. the sentence for the shown image is: *a man at a giant tree in the jungle*). The model will estimate the probability distribution of the next word given previous words and the image. This architecture is much deeper than the simple RNN. (c). The illustration the unfolded m-RNN. The model parameters are shared for each temporal frame of the m-RNN model.

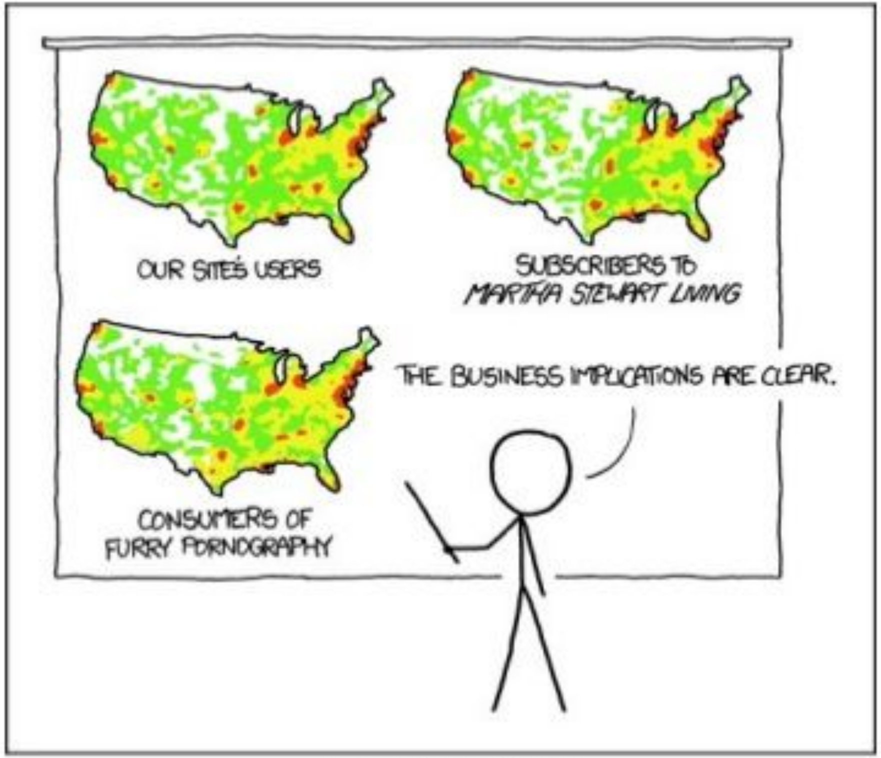
Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." *European conference on computer vision*. Springer, Cham, 2014.

FEATURE VISUALIZATION FOR DEEP NETWORKS

Understanding representations in deep nets - why?

- If we don’t understand what the network is learning, we can’t tune it to what we want it to learn. A black-box genius is of very little use.
- Explainable AI: AI systems are exceeding human level intelligence at complex tasks. Lack of transparency can be a major drawback.
- If the features learnt by the system are not understood, the network will be even more prone to adversarial attacks.

Visualization



PET PEEVE #208:
GEOGRAPHIC PROFILE MAPS WHICH ARE
BASICALLY JUST POPULATION MAPS

PAPERS TO READ

- Karel Lenc, Andrea Vedaldi, Understanding image representations by measuring their equivariance and equivalence, CVPR, 2015. [Paper]
- Anh Nguyen, Jason Yosinski, Jeff Clune, Deep Neural Networks are Easily Fooled:High Confidence Predictions for Unrecognizable Images, CVPR, 2015. [Paper]
- Aravindh Mahendran, Andrea Vedaldi, Understanding Deep Image Representations by Inverting Them, CVPR, 2015. [Paper]
- Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, Antonio Torralba, Object Detectors Emerge in Deep Scene CNNs, ICLR, 2015. [arXiv Paper]
- Alexey Dosovitskiy, Thomas Brox, Inverting Visual Representations with Convolutional Networks, arXiv, 2015. [Paper]
- Matthrew Zeiler, Rob Fergus, Visualizing and Understanding Convolutional Networks, ECCV, 2014. [Paper]


DEEP LEARNING,
SO HOT RIGHT NOW
imgflip.com

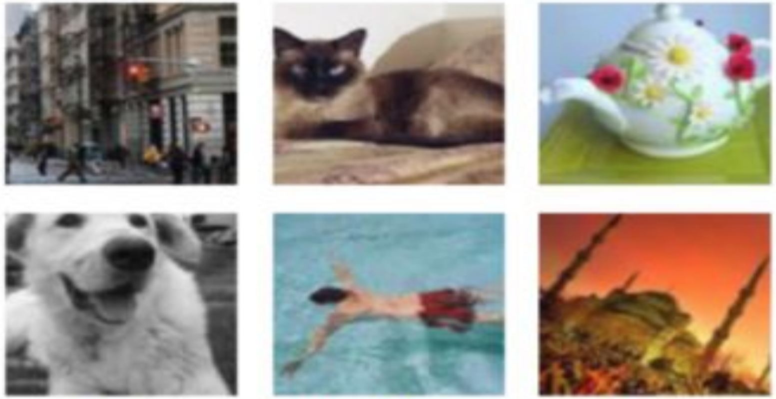
GENERATIVE MODELS

https://github.com/soumith/ganhacks

Generative Models

Given training data, generate new samples from the same distribution.

Training data
$p_{data}(x)$

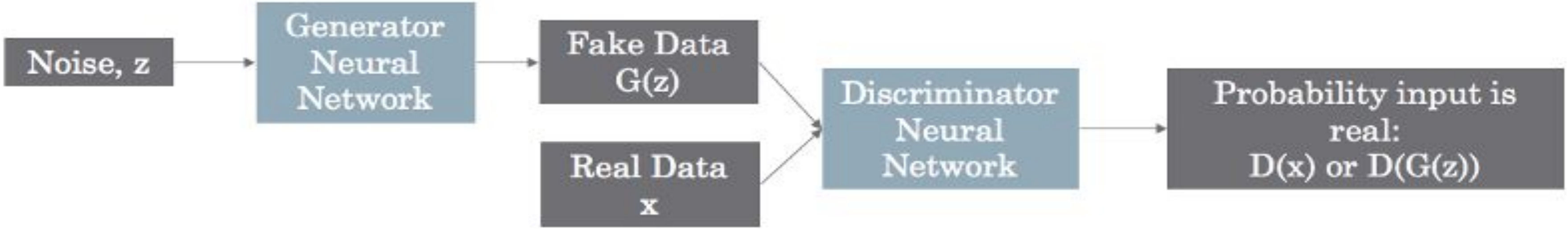


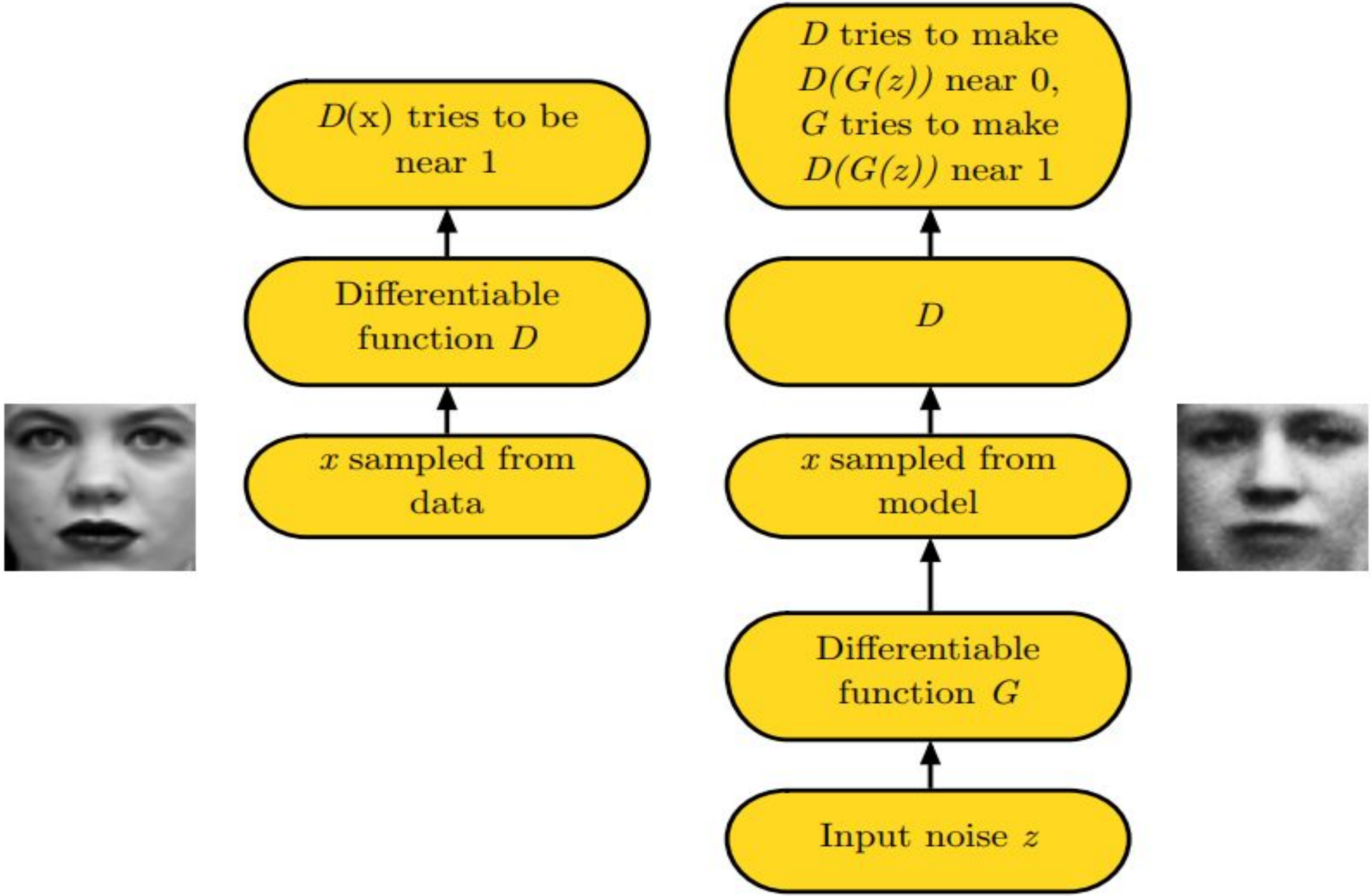
Generated Samples
$p_{model}(x)$

We want to learn $p_{model}(x)$ similar to $p_{data}(x)$.

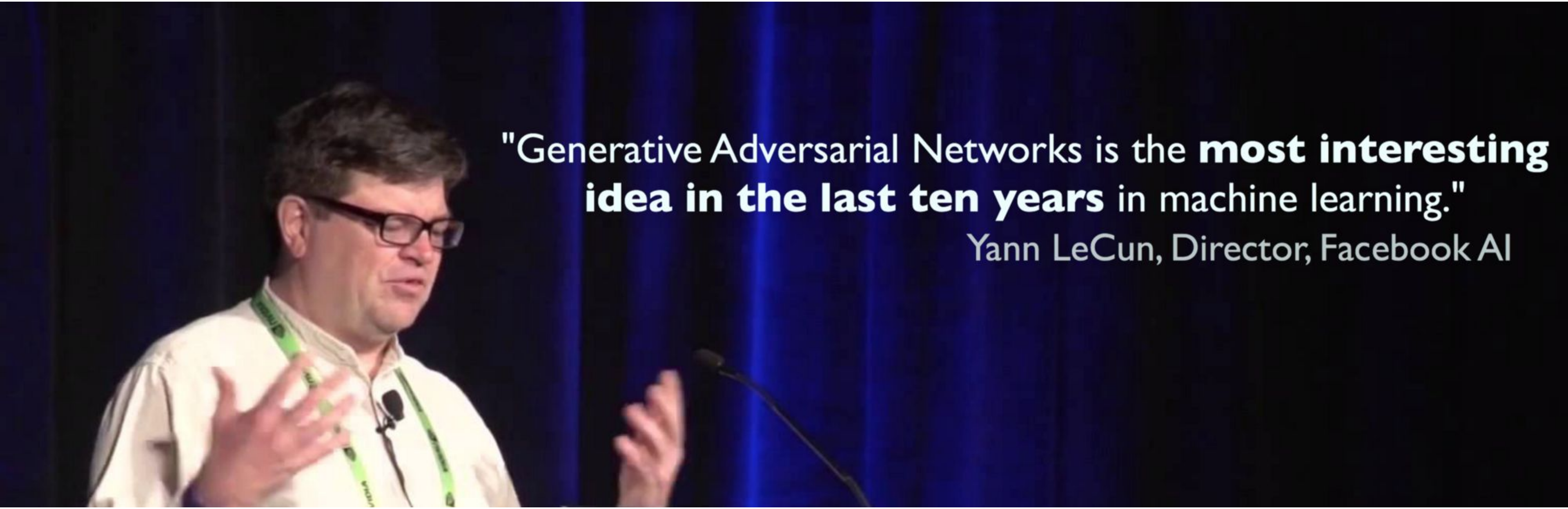
GENERATIVE ADVERSARIAL NETWORKS - GANs

- For each mistake on a fake image, the discriminator gets penalized and the generator gets a rewarded.
- Adversarial—the discriminator's loss is the generator's gain
- competition leads to mutual improvement.



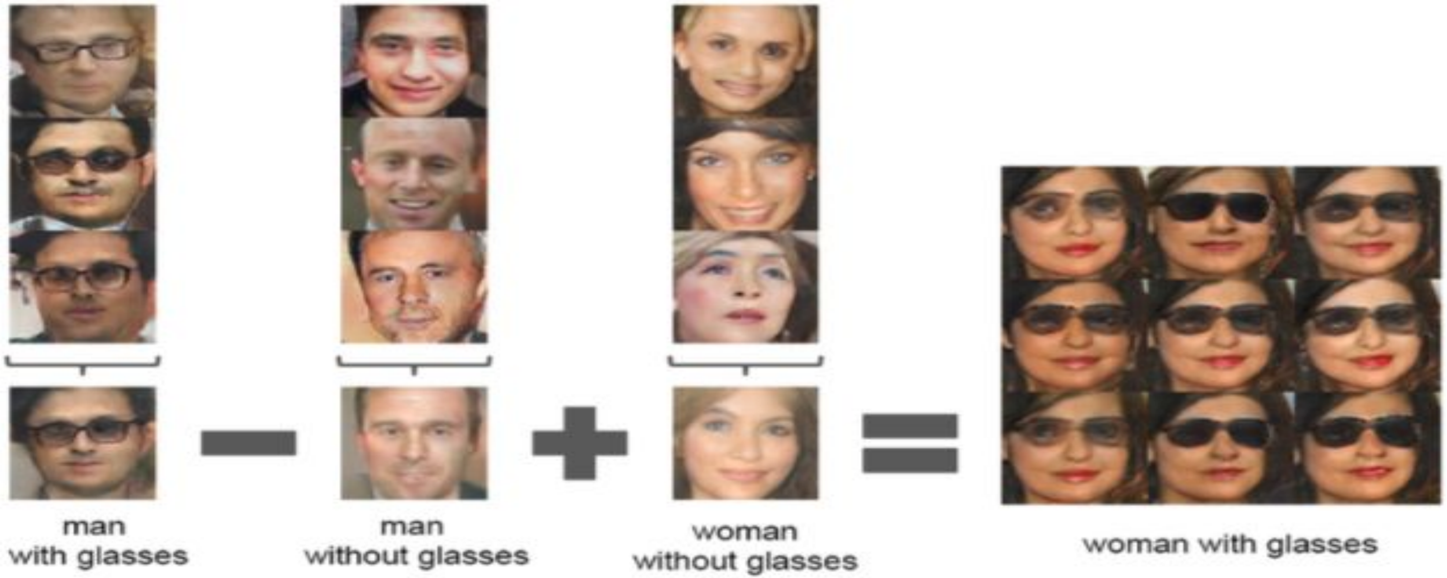


(Goodfellow 2016)


"Generative Adversarial Networks is the most interesting idea in the last ten years in machine learning."
Yann LeCun, Director, Facebook AI

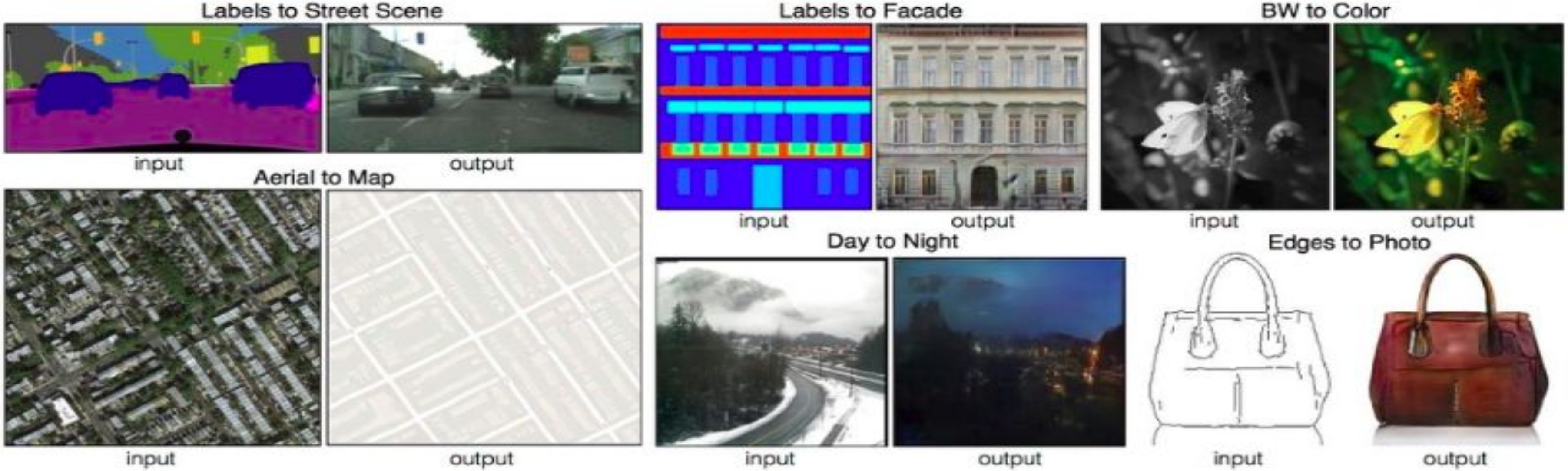
Deep Convolutional GANs

Fun Operations



A. Radford, L.Metz, and S. Chintala. *Unsupervised representation learning with deep convolutional generative adversarial networks*. ICLR 2016.

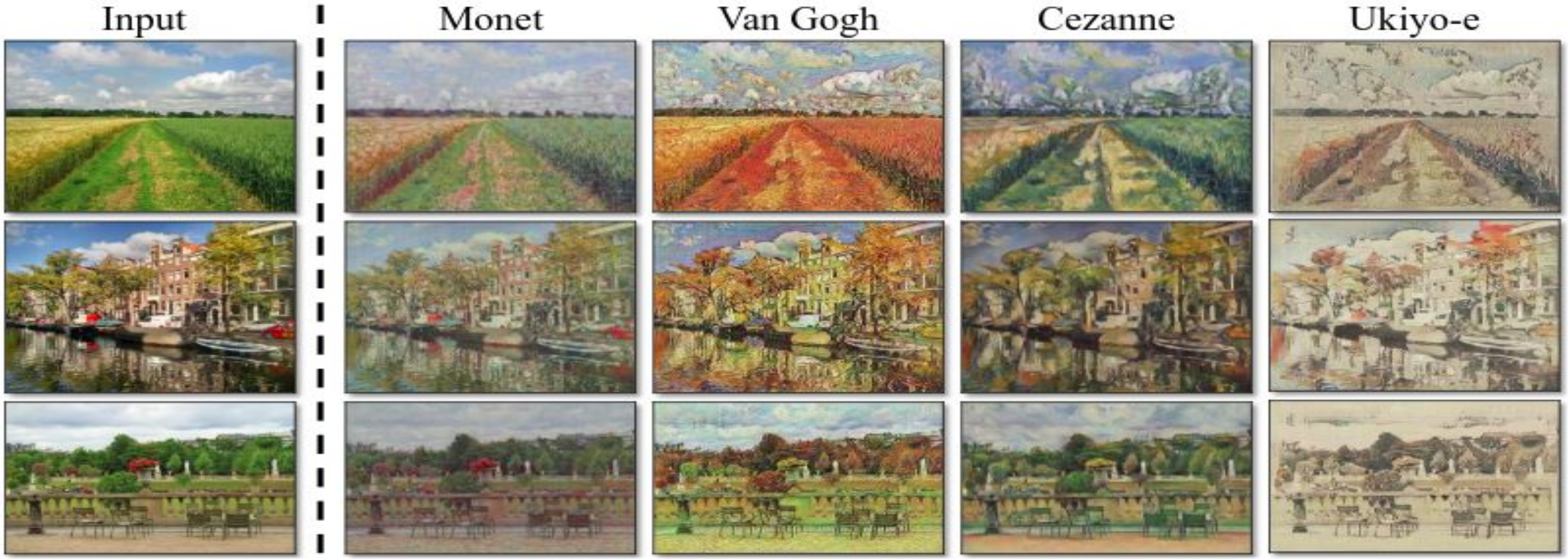
Image-to-Image Translation



Popularly known as Pix2Pix.

P. Isola, J. Zhu, T. Zhou, A. Efros. *Image-to-Image Translation with Convolutional Adversarial Networks. CVPR 2017*

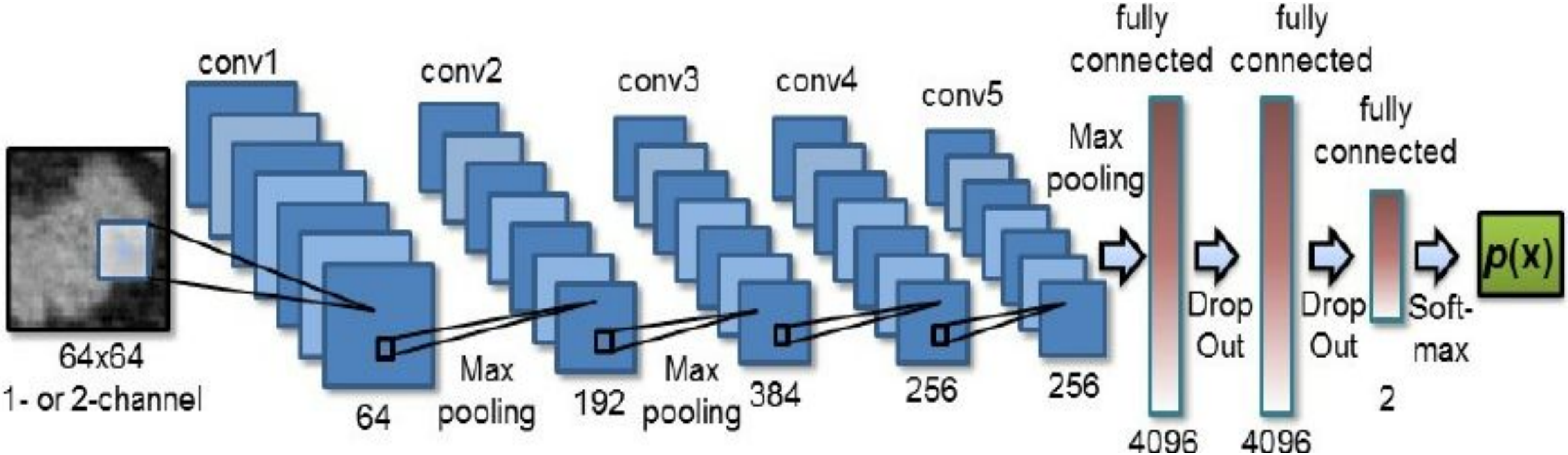
Unpaired Image-to-Image Translation



J. Zhu, T. Park, P. Isola, A. Efros. *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks.* In ICCV 2017.


GENERATOR
DISCRIMINATOR
risovach.ru

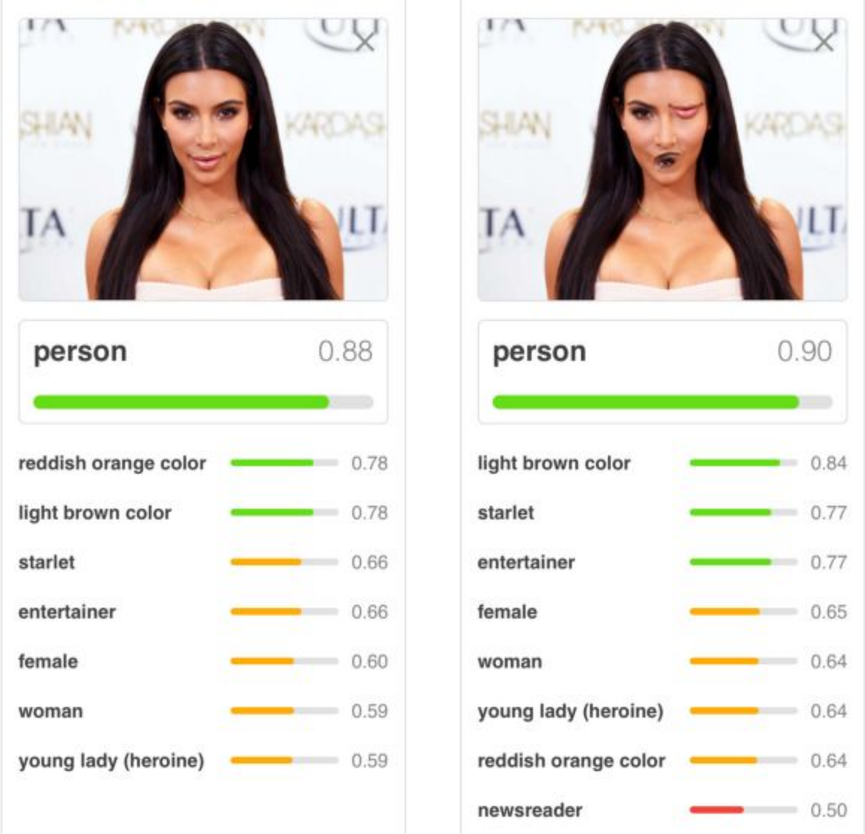
CNN architecture



EQUIVARIANCE VS INVARIANCE

- The pooling operation tries to make the NN invariant to small changes in viewpoint.
- This leads to a complication that spatial information being lost during pooling, hence making a CNN dumb.
- Equivariance means that changes in viewpoints will correspond to a change in the firing of the neurons, hence learning richer representations of the data.

CAPSULE NETWORKS - HINTON

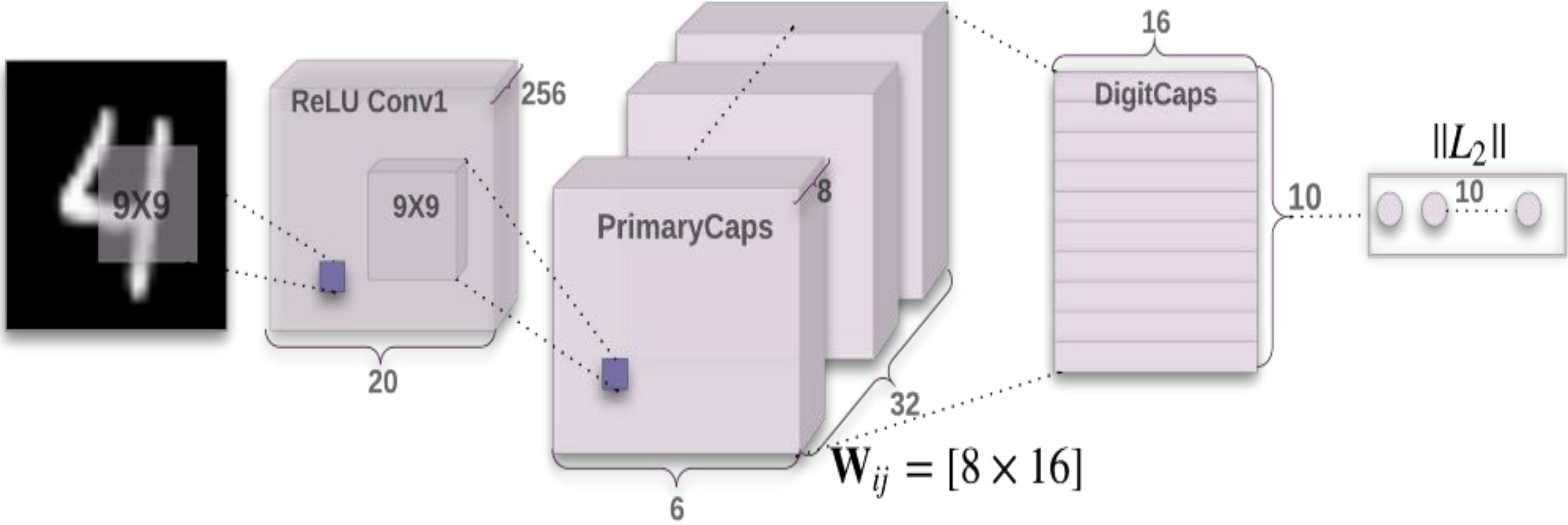


Hinton: "The pooling operation used in convolutional neural networks is a big mistake and the fact that it works so well is a disaster."

Capsule networks

- CNNs do routing by pooling
- A nested layer inside a layer can be called a ‘capsule’
- State of the art for MNIST
- Introduces dynamic routing in between convolutional layers
- Layer-based squashing - instead of applying ReLU to each neuron separately, we apply it to the grouped neurons as a vector (capsule)
- Capsule networks essentially encodes important routing information in the capsules, and uses this info during the forward pass

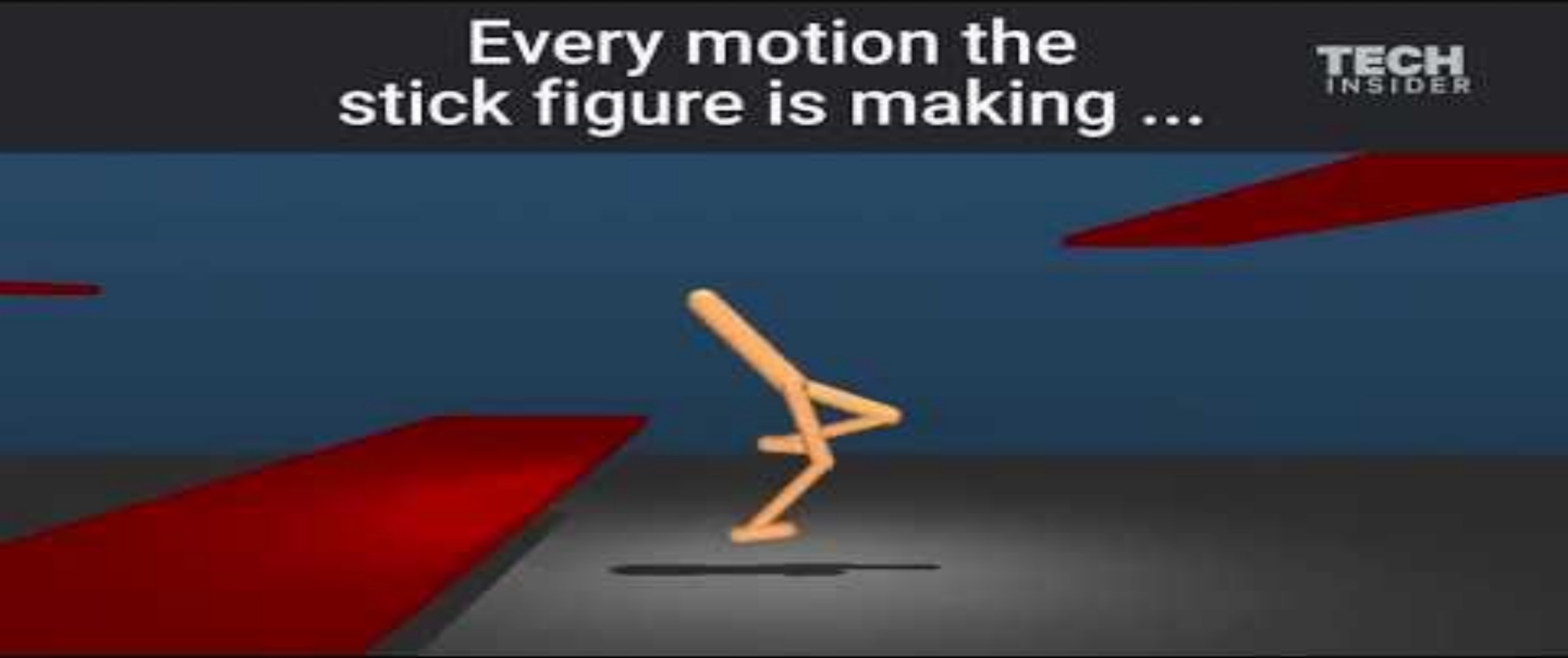
CAPSULE NETWORK ARCHITECTURE



https://arxiv.org/pdf/1710.09829.pdf


ONE DOES NOT SIMPLY
"PICK" HYPERPARAMETERS
imgflip.com

REINFORCEMENT LEARNING


Every motion the
stick figure is making ...
TECH
INSIDER

REINFORCEMENT LEARNING BASICS

Basic idea:

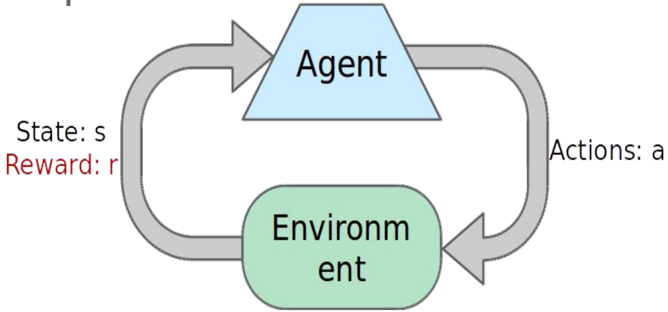
Receive feedback in the form of rewards

Agent's utility is defined by the reward function

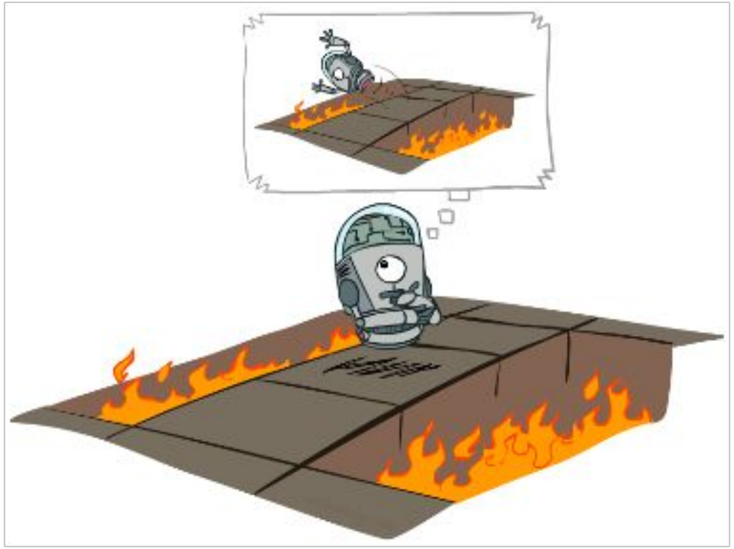
Must (learn to) act so as to maximize expected rewards

All learning is based on observed

samples of outcomes!



Offline (MDPs) vs. Online (RL)

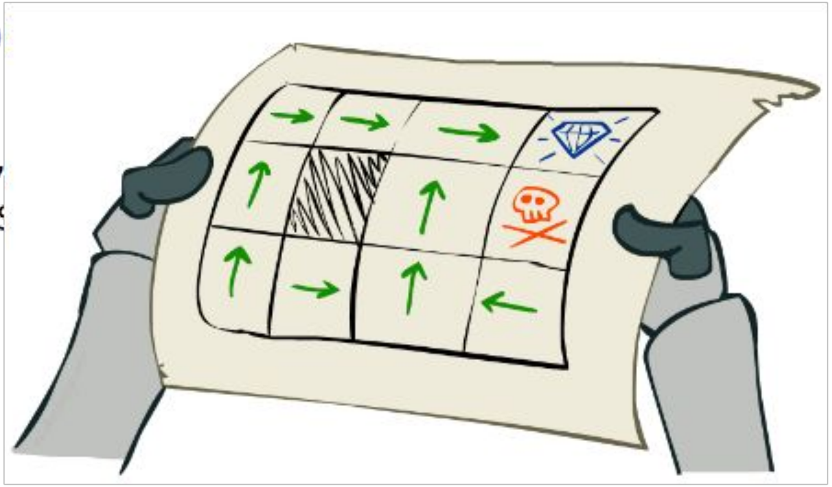


Offline
Solution

Online
Learning

Passive Reinforcement Learning

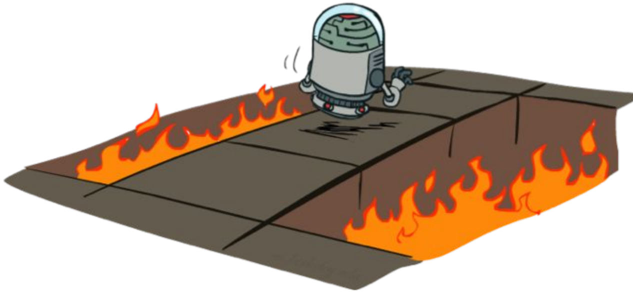
- Simplified task: policy evaluatio
  - Input: a fixed policy $\pi(s)$
  - You don't know the transitions T(s,
  - You don't know the rewards R(s,a,
  - Goal: learn the state values



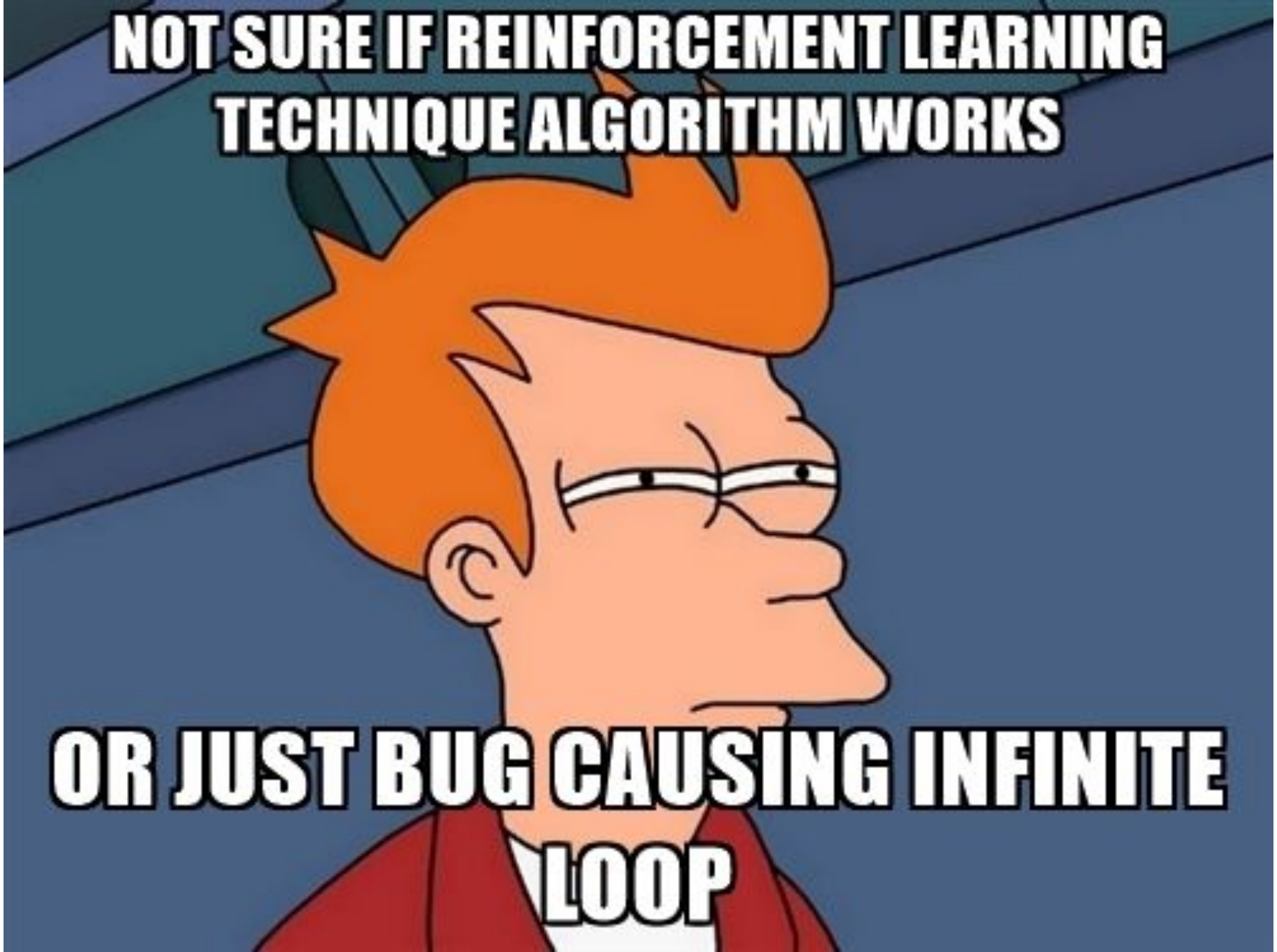
- In this case:
  - Learner is "along for the ride"
  - No choice about what actions to take
  - Just execute the policy and learn from experience
  - This is NOT offline planning! You actually take actions in the world.

Active Reinforcement Learning

- Full reinforcement learning: optimal policies (like value iteration)
  - You don't know the transitions T(s,a,s')
  - You don't know the rewards R(s,a,s')
  - You choose the actions now
  - Goal: learn the optimal policy / values

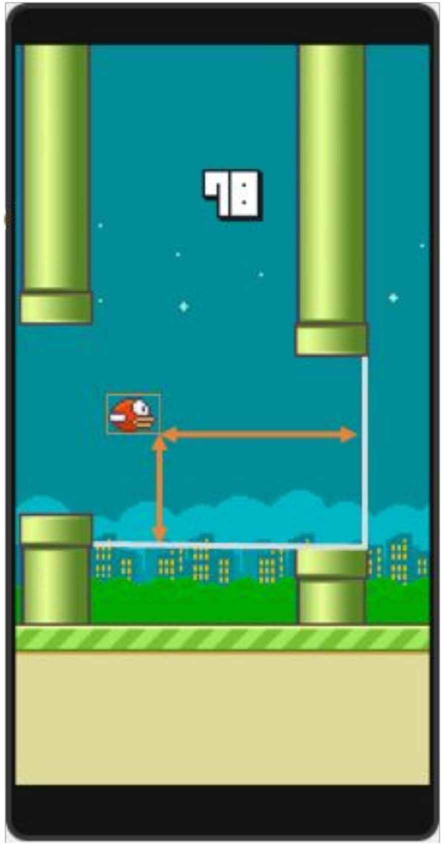


- In this case:
  - Learner makes choices!
  - Fundamental tradeoff: exploration vs. exploitation
  - This is NOT offline planning! You actually take actions in the world and find out what happens…


NOT SURE IF REINFORCEMENT LEARNING TECHNIQUE ALGORITHM WORKS
OR JUST BUG CAUSING INFINITE LOOP

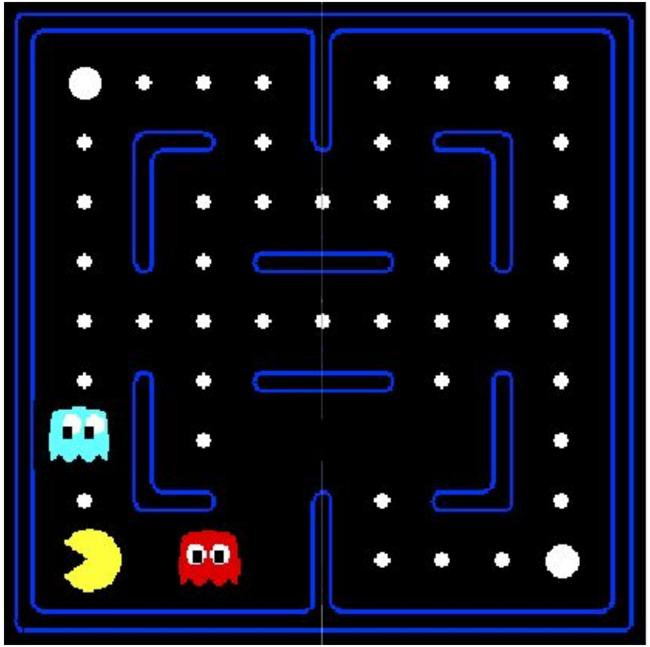
AI in the News – Flappy Bird RL

- State space
  - Discretized vertical distance from lower pipe
  - Discretized horizontal distance from next pair
  - Life: Dead or Living
- Actions
  - Click
  - Do nothing
- Rewards
  - +1 if Flappy Bird still alive
  - -1000 if Flappy Bird is dead
- 6-7 hours of Q-learning



Feature-Based Representations

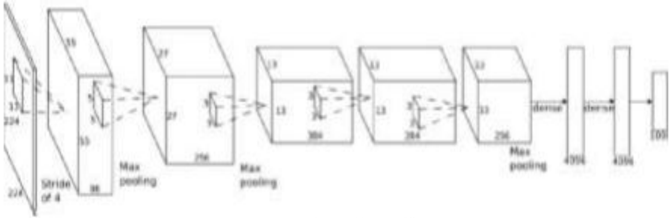
- Solution: describe a state using a vector of features (properties)
  - Features are functions from states to real numbers (often 0/1) that capture important properties of the state
  - Example features:
    - Distance to closest ghost
    - Distance to closest dot
    - Number of ghosts
    - 1 / (dist to dot)^2
    - Is Pacman in a tunnel? (0/1)
    - …… etc.
    - Is it the exact state on this slide?
  - Can also describe a q-state (s, a) with features (e.g. action moves closer to food)



when you keep making new email accounts to get free trials

Imitation Learning

$\mathbf{o}_t$



$\pi_\theta(\mathbf{a}_t|\mathbf{o}_t)$

$\mathbf{a}_t$

$\mathbf{o}_t$ $\mathbf{a}_t$ → training data → supervised learning → $\pi_\theta(\mathbf{a}_t|\mathbf{o}_t)$

Images: Bojarski et al. '16, NVIDIA

Where does the reward function come from?

Computer Games

reward

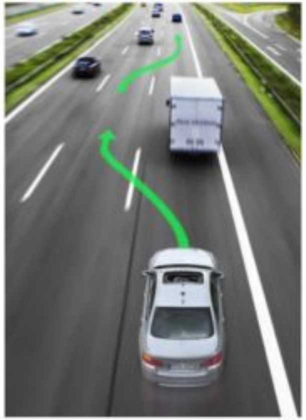
Mnih et al. '15

Real World Scenarios

robotics

dialog

autonomous driving



what is the reward?
often use a proxy

frequently easier to provide expert data

Inverse reinforcement learning: infer reward function from roll-outs of expert policy

A bit more formally

"forward" reinforcement learning

given:

states $\mathbf{s} \in \mathcal{S}$, actions $\mathbf{a} \in \mathcal{A}$

(sometimes) transitions $p(\mathbf{s}'|\mathbf{s}, \mathbf{a})$

reward function $r(\mathbf{s}, \mathbf{a})$

learn $\pi^\star(\mathbf{a}|\mathbf{s})$

inverse reinforcement learning

given:

states $\mathbf{s} \in \mathcal{S}$, actions $\mathbf{a} \in \mathcal{A}$

(sometimes) transitions $p(\mathbf{s}'|\mathbf{s}, \mathbf{a})$

samples $\{\tau_i\}$ sampled from $\pi^\star(\tau)$

learn $r_\psi(\mathbf{s}, \mathbf{a})$

reward parameters

...and then use it to learn $\pi^\star(\mathbf{a}|\mathbf{s})$

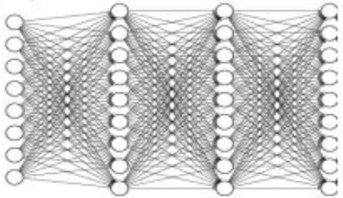
linear reward function:

$r_\psi(\mathbf{s}, \mathbf{a}) = \sum_i \psi_i f_i(\mathbf{s}, \mathbf{a}) = \psi^T \mathbf{f}(\mathbf{s}, \mathbf{a})$

neural net reward function:

$\mathbf{s}$

$\mathbf{a}$



$r_\psi(\mathbf{s}, \mathbf{a})$

parameters ψ

IRL as adversarial optimization

Guided Cost Learning

ICML 2016

minimized maximized

reward function

robot attempt

human demonstrations

learns distribution $p(\tau)$ such that demos have max likelihood

$p(\tau) \propto \exp(r(\tau))$ (MaxEnt model)

$$D(\tau) = \frac{\frac{1}{Z}\exp(r(\tau))}{\frac{1}{Z}\exp(r(\tau)) + \pi(\tau)}$$

actually the same thing!

Generative Adversarial Imitation Learning

Ho & Ermon, NIPS 2016

False True

classifier

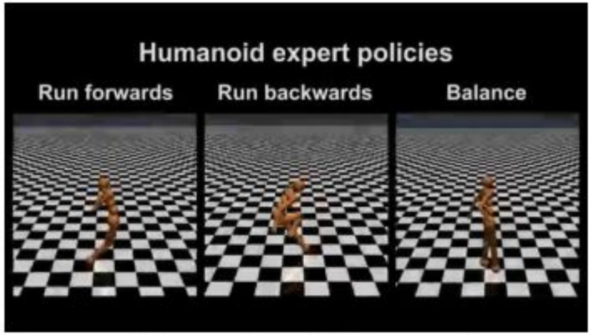
robot attempt

human demonstrations

$D(\tau) =$ probability τ is a demo

use $\log D(\tau)$ as "reward"

$D(\tau) =$ some classifier



Hausman, Chebotar, Schaal, Sukhatme, Lim

Merel, Tassa, TB, Srinivasan, Lemmon, Wang, Wayne, Heess


If I let you make me nervous, then we can't get schwifty.

PAPERS TO READ - IRL

- Abbeel & Ng ICML ’04. Apprenticeship Learning via Inverse Reinforcement Learning. Good introduction to inverse reinforcement learning
- Ziebart et al. AAAI ’08. Maximum Entropy Inverse Reinforcement Learning. Introduction to probabilistic method for inverse reinforcement learning

  Modern Papers: Finn et al. ICML ’16.
- Guided Cost Learning. Sampling based method for MaxEnt IRL that handles unknown dynamics and deep reward functions Wulfmeier et al. arXiv ’16.
- Deep Maximum Entropy Inverse Reinforcement Learning. MaxEnt inverse RL using deep reward functions Ho & Ermon NIPS ’16.
- Generative Adversarial Imitation Learning. Inverse RL method using generative adversarial networks

META LEARNING SHARED HIERARCHIES

Kevin Frans
Henry M. Gunn High School
Work done as an intern at OpenAI
kevinfrans2@gmail.com

Jonathan Ho, Xi Chen, Pieter Abbeel
UC Berkeley, Department of Electrical
Engineering and Computer Science

John Schulman
OpenAI

ABSTRACT

We develop a metalearning approach for learning hierarchically structured policies, improving sample efficiency on unseen tasks through the use of shared primitives—policies that are executed for large numbers of timesteps. Specifically, a set of primitives are shared within a distribution of tasks, and are switched between by task-specific policies. We provide a concrete metric for measuring the strength of such hierarchies, leading to an optimization problem for quickly reaching high reward on unseen tasks. We then present an algorithm to solve this problem end-to-end through the use of any off-the-shelf reinforcement learning


SAY DEEP LEARNING
ONE MORE TIME