

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



## یادگیری عمیق

جلسه ۵

# کاربردهای یادگیری عمیق

## Applications of Deep Learning

کاظم فولادی قلعه  
دانشکده مهندسی، پردیس فارابی  
دانشگاه تهران

<http://courses.fouladi.ir/deep>

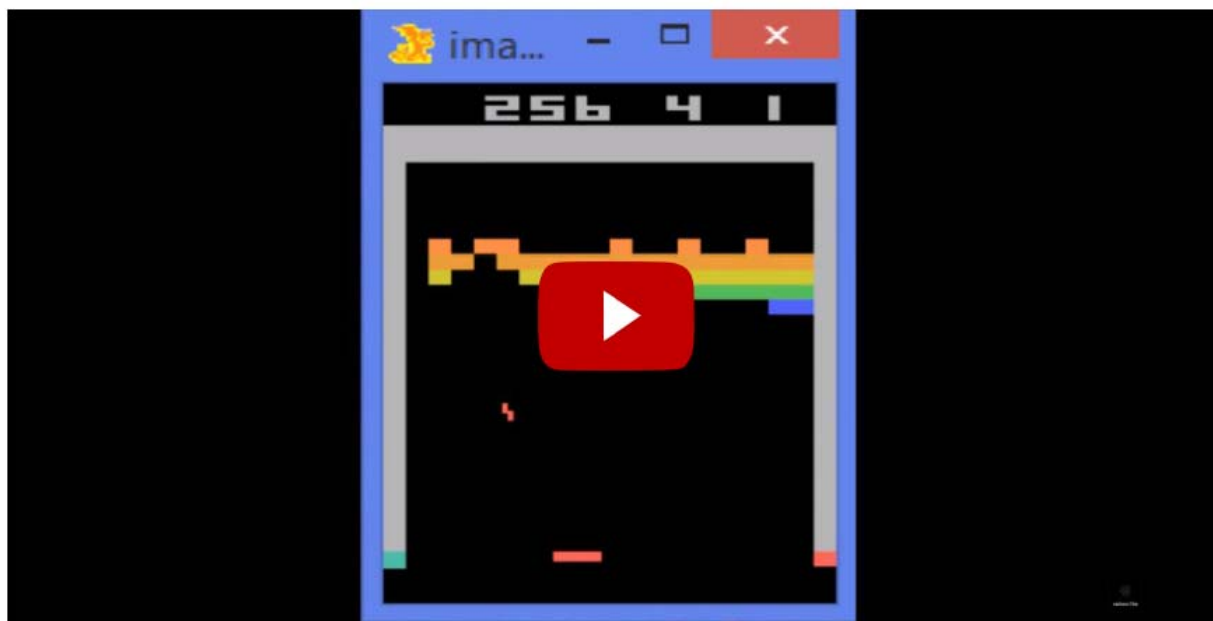
کاربردهای یادگیری عمیق

۱

# کاربردهای نمونه

# Deep Learning for Games

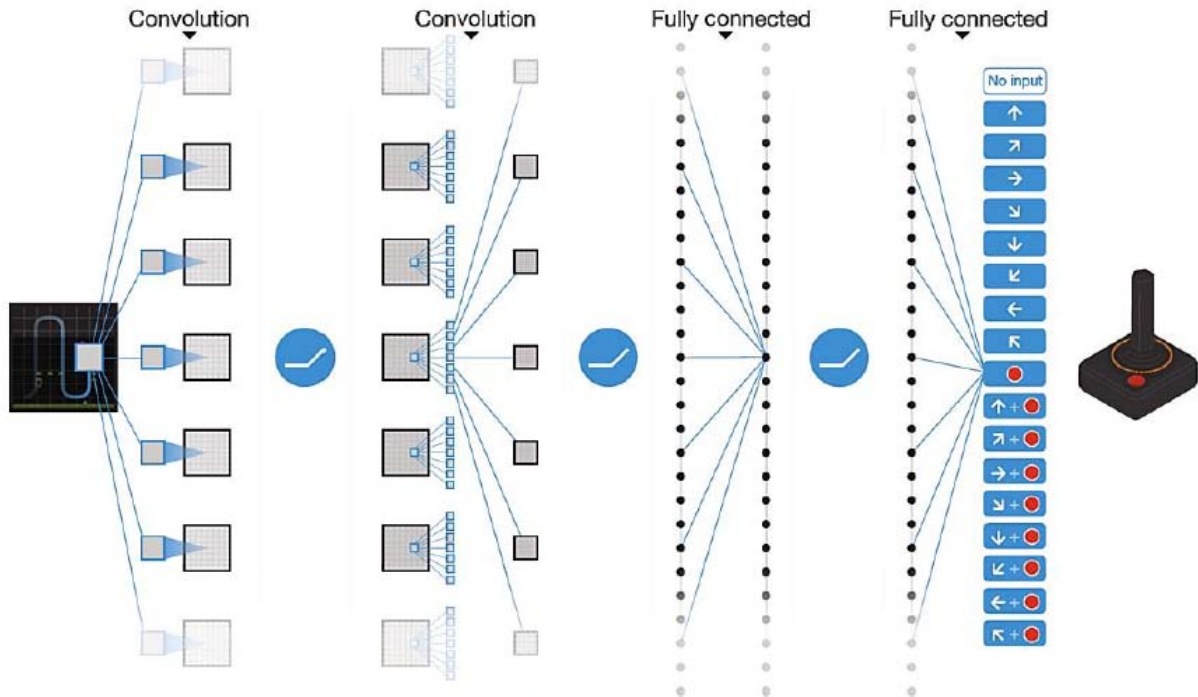
DeepMind Deep Q-Learning



Outperforms humans in over 30 Atari games just by receiving the pixels on the screen with the goal to maximize the score (Reinforcement Learning)

# Deep Learning for Games

## DeepMind Deep Q-Learning



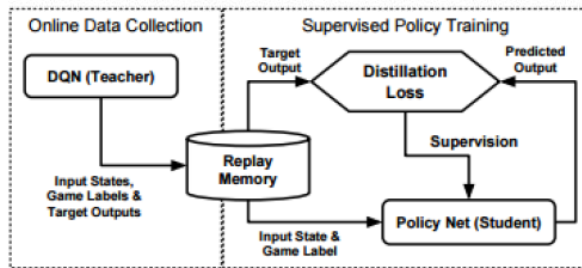
Deep Q-Learning (DQN) is a model-free approach to reinforcement learning using deep networks in environments with discrete action choices

# Deep Learning - Basics

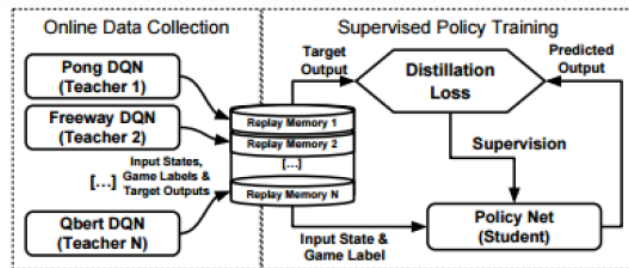
## DeepMind Deep Q-Learning



Policy distillation: Extracts the learned state (*policy*) of a reinforcement learning agent (*teacher*) and **trains a new network (*student*) that performs at the expert level** while being dramatically smaller and more efficient.



Single-task policy distillation



Multi-task policy distillation

# Deep Learning for Games

Open Environments – Deep Mind Lab



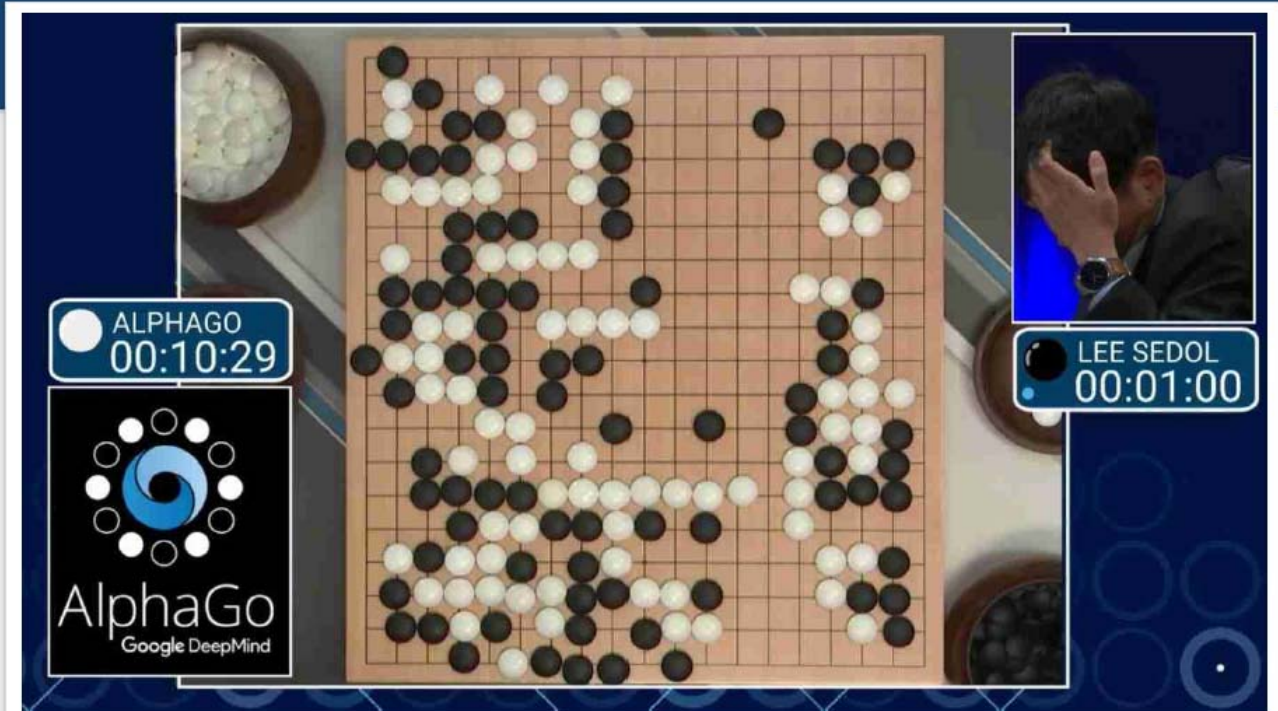
# Deep Learning for Games

Open Environments – Open AI Universe



# Deep Learning for Games

DeepMind AlphaGo



History is made: Google's **AlphaGo wins the match** against Go champion Lee Sedol



# Deep Learning for Games

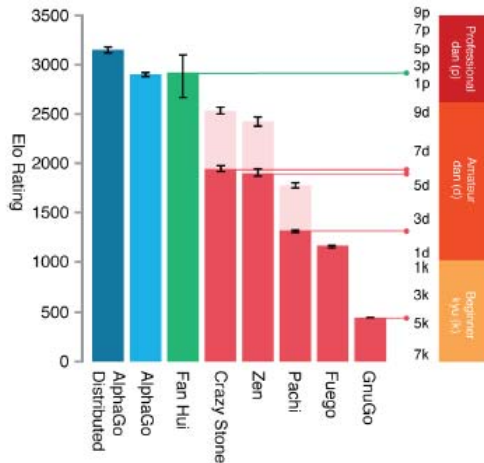
## DeepMind AlphaGo



Demis Hassabis

@demishassabis

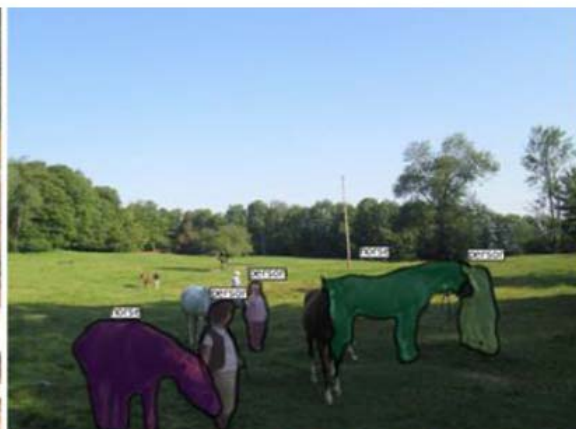
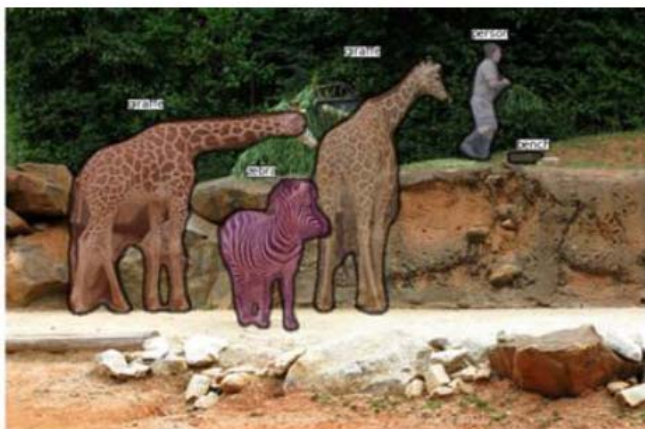
Just been told 60m viewers watched the first match online just in China! 100m+ worldwide inc. TV, 3300 news articles just in Korea. Amazing!



AlphaGo could learn the game by examining thousands of human Go moves, and then it could master the game by **playing itself over and over and over again**. The result is a system of unprecedented beauty.

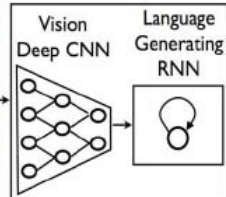
# Deep Learning in Computer Vision

## Image Segmentation



# Deep Learning in Computer Vision

## Image Captioning



**A group of people shopping at an outdoor market.**

**There are many vegetables at the fruit stand.**

Neural Image Caption Generator **generates fitting natural-language captions only based on the pixels** by combining a vision CNN and a language-generating RNN.



A close up of a child holding a stuffed animal



Two pizzas sitting on top of a stove top oven



A man flying through the air while riding a skateboard

# Deep Learning in Computer Vision

## Image Captioning v2



### Human captions from the training set



### Automatically captioned



# Deep Learning in Computer Vision

## Image Compression



Image compression with Residual Gated Recurrent Unit (Residual GRU).  
Left: Original (1419KB PNG), Center: JPEG (33KB), Right: Residual GRU (24KB).  
=> 25% smaller for comparable image quality.

# Deep Learning in Computer Vision

## Image Localization



Photo CC-BY-NC by edwin.11



Photo CC-BY-NC by stevekc

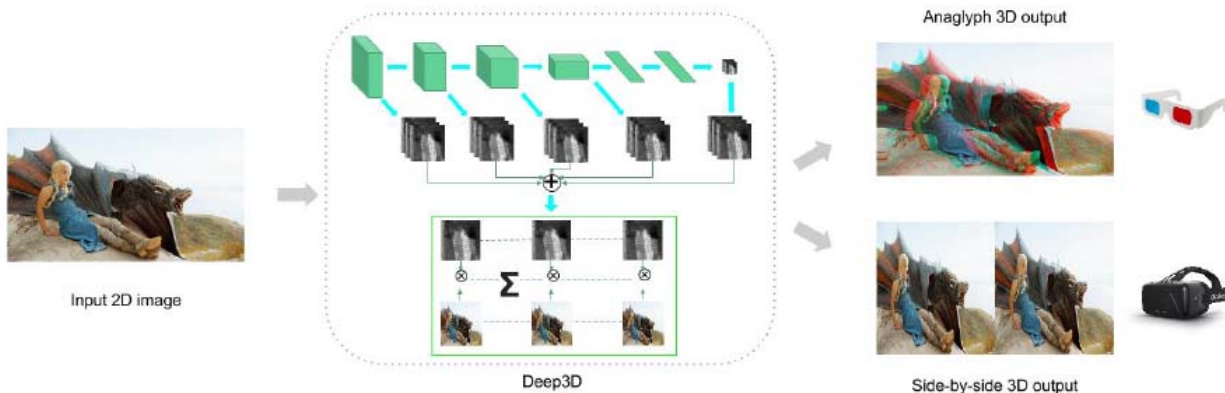


PlaNet is able to **determine the location of almost any image** with superhuman ability.

# Deep Learning in Computer Vision

## Image Transformation – 2D-to-3D

Deep3D can automatically convert image/video from 2D-to-3D with Convolutional Neural Networks. It learns to infer 3D representations of the world based on training set of 3D movies.



# Deep Learning in Computer Vision

## Image Sharpening



DCGAN architecture to **upscale and sharpen an image** with features that are plausible based on the dataset that was used to train the neural net.



# Deep Learning in Computer Vision

## Image Completion



Image completion with deep convolutional generative adversarial networks (DCGAN). The centers of these images are being automatically generated.

# Deep Learning in Computer Vision

## Image Transformation – Adding features



Performs high-level **semantic transformations** on images like "make older/younger", "make bespectacled", "add smile".

# Deep Learning in Computer Vision

## Image Colorization



Given a grayscale photograph as input, this Convolutional Neural Network tackles the problem of **hallucinating a plausible color version of the photograph**.

# Deep Learning in Computer Vision

## Image-to-Image Translation

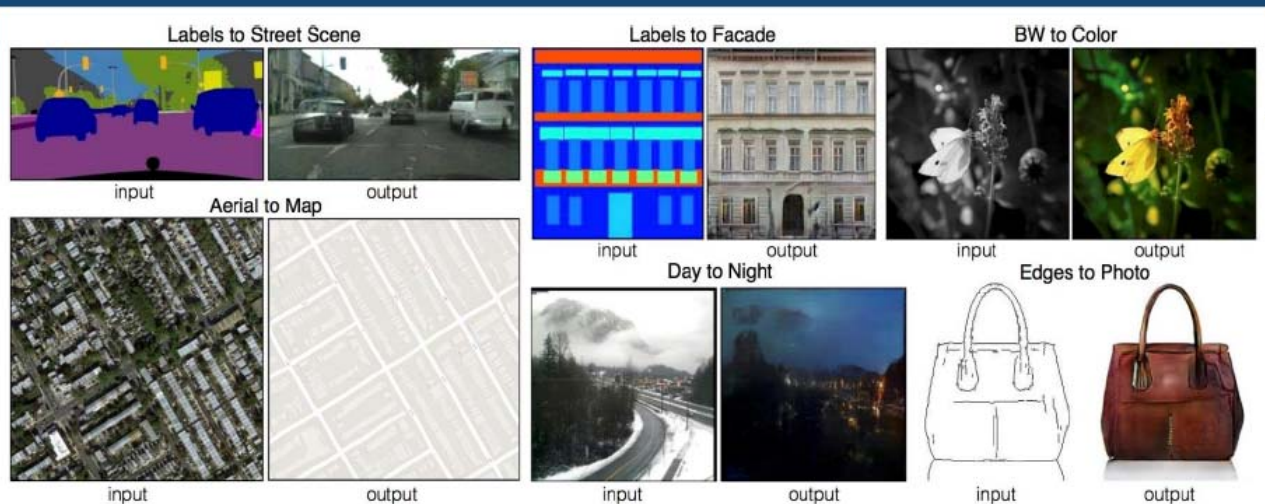
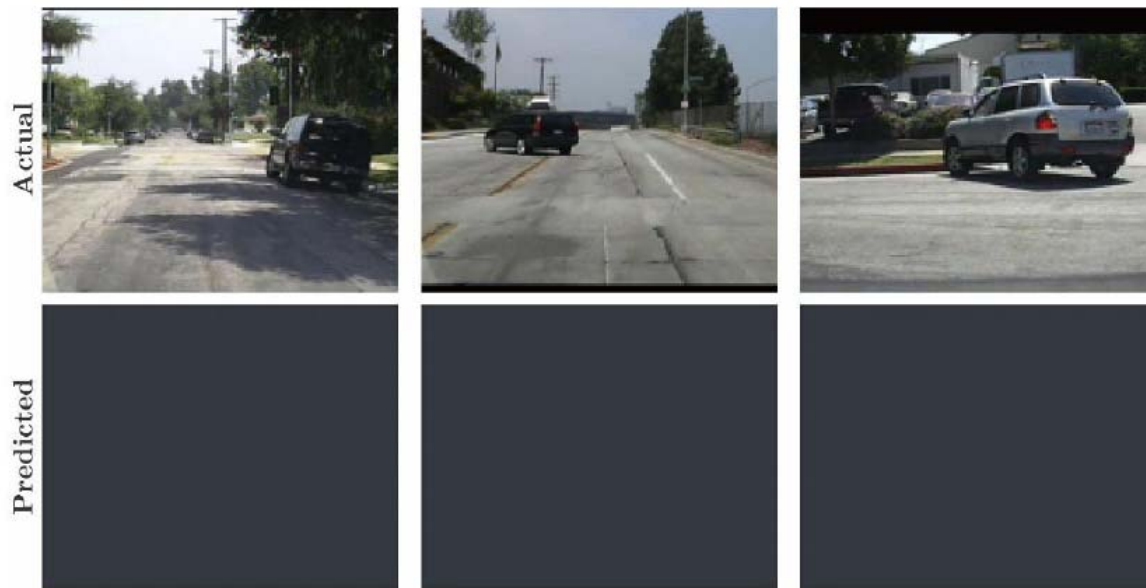


Image-to-Image Translation with Conditional Adversarial Nets. This approach can be applied as a **generic solutions to any Image-to-Image translation** problem such as synthesizing photos from label maps, reconstructing objects from edge maps, and colorizing images.

# Deep Learning in Computer Vision

## Video Sequence Prediction



PredNet - a deep convolutional recurrent neural network that **predicts the future frames in a video sequence**. These networks are able to robustly learn to predict the movement of synthetic (rendered) objects.

# Deep Learning in Computer Vision

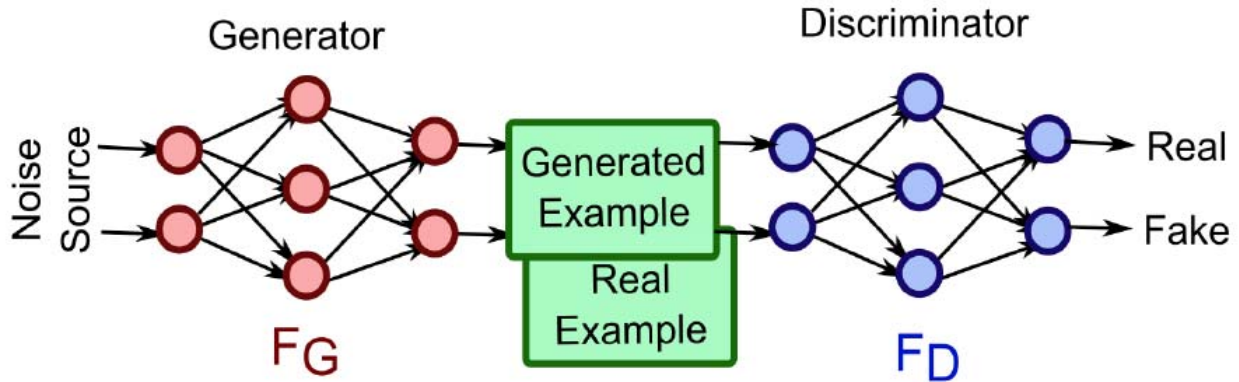
## Image Generation – Bedrooms



These neural networks are learning what the visual world looks like!

# Deep Learning - Basics

## Generative Adversarial Networks



Generative Adversarial Networks (GANs) consist of any two networks with **one tasked to generate content and the other has to judge content.**

# Deep Learning in Computer Vision

## Image Generation – Album Covers





# Deep Learning in Computer Vision

## Image Generation – Fine Art



# Deep Learning in Computer Vision

## Image Generation – Plug & Play Generative Networks



redshank

ant

monastery



# Deep Learning in Computer Vision

## Image Generation – Faces



**Generating realistic faces** based on a selected person's identity, emotion, and orientation with deconvolution network. You give the network the parameters of the thing you want to draw and it does it.

# Deep Learning in Computer Vision

## LipNet - Sentence-level Lipreading



LipNet achieves 93.4% accuracy, outperforming experienced human lipreaders and the previous 79.6% state-of-the-art accuracy.

# Deep Learning in Computer Vision

## Image Generation – From Descriptions

**Text descriptions (content)**   **Images (style)**

The bird has a **yellow breast** with **grey** features and a small beak.

This is a large **white** bird with **black** wings and a **red** head.

A small bird with a **black** head and **wings** and features grey wings.

This bird has a **white** breast, brown and white coloring on its head and wings, and a thin pointy beak.

A small bird with **white** base and **black** stripes throughout its belly, head, and feathers.

A small sized bird that has a cream belly and a short pointed bill.

This bird is **completely red**.



this small bird has a pink breast and crown, and black primaries and secondaries.



the flower has petals that are bright pinkish purple with white stigma



this magnificent fellow is almost all black with a red crest, and white cheek patch.



this white and yellow flower have thin white petals and a round yellow stamen



Deep architecture and GAN formulation to **translate visual concepts from characters to pixels**. We demonstrate the capability of our model to generate plausible images of birds and flowers from detailed text descriptions.

# Deep Learning in Computer Vision

## Image Generation - Handwriting

This is an impressive demo of a recurrent neural network.

This is an impressive demo of a recurrent neural network.

This is an impressive demo of a recurrent neural network.

This LSTM recurrent neural network is able to **generate highly realistic cursive handwriting in a wide variety of styles**, simply by predicting one data point at a time.

# Deep Learning in Computer Vision

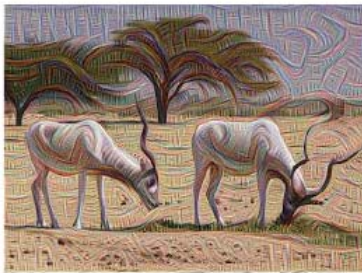
## DeepDream – Inceptionism



Inceptionism helps to **understand and visualize** what a neural network has learned during training



“No picture of a dumbbell is complete without a muscular weightlifter”



Let the network **over-interpret whatever it detects** in a selected layer (*e.g. edges*)

# Deep Learning in Computer Vision

## DeepDream – Inceptionism



Leaves



Birds & Insects

By choosing higher level layers, more sophisticated features or even **whole objects** tend to emerge.





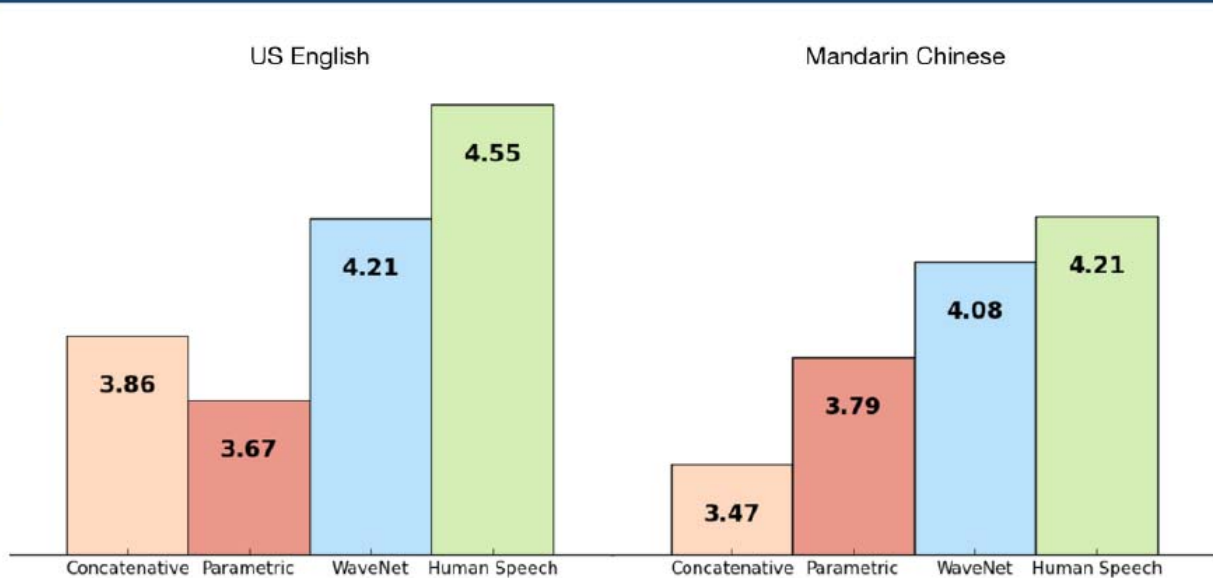
# Deep Learning in Computer Vision

Style Transfer – morph images into paintings



# Deep Learning in Audio Processing

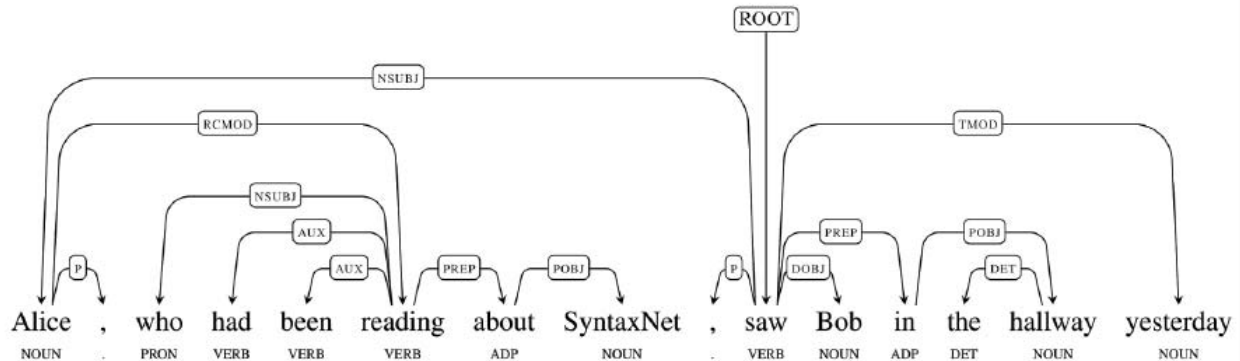
## Sound Generation



DeepMind's **WaveNet** is able to generate speech which mimics any human voice and which sounds more natural than the best existing Text-to-Speech systems, reducing the gap with human performance by over 50%.

# Deep Learning in NLP

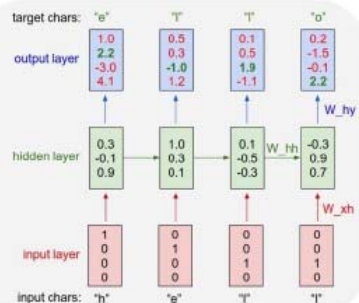
## Syntax Parsing



SyntaxNet (Parsey McParseface) tags each word with a part-of-speech tag, and it determines the syntactic relationships between words in the sentence with an **94% accuracy** compared to a human performance at 96%.

# Deep Learning in NLP

## Generating Text



To train the RNN, insert characters sequentially and predict the probabilities of the next letter. Backpropagate error and update RNN's weights to increase the confidence of the correct letter (green) and decrease the confidence of all other letters (red).

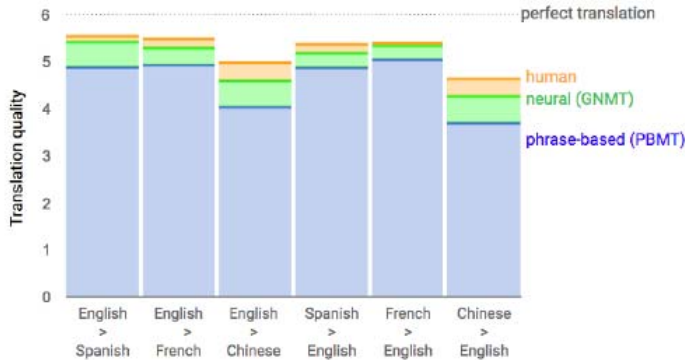
The emperor travelled back to [[Antioch, Perth, October 25|21]] to note, the Kingdom of Costa Rica, unsuccessful fashioned the [[Thrales]], [[Cynth's Dajoard]], known in western [[Scotland]], near Italy to the conquest of India with the conflict. Copyright was the succession of independence in the slop of Syrian influence that was a famous German movement based on a more popular servicious, non-doctrinal and sexual power post. Many governments recognize the military housing of the [[Civil Liberalization and Infantry Resolution 265 National Party in Hungary]], that is sympathetic to be to the [[Punjab Resolution]] (PJS)[<http://www.humah.yahoo.com/guardian.cfm/7754800786d17551963s89.htm>]

Trained on structured Wikipedia markdown. Network learns to spell English words completely from scratch and copy general syntactic structures.



# Deep Learning in NLP

## Neural Machine Translation



Googles Neural Machine Translation system based on a deep LSTM seq2seq architecture **reduces translation errors by more than 55%-85%** compared to traditional approaches.

<b>Input sentence:</b>	<b>Translation (PBMT):</b>	<b>Translation (GNMT):</b>	<b>Translation (human):</b>
李克強此行將啟動中加總理年度對話機制，與加拿大總理杜魯多舉行兩國總理首次年度對話。	Li Keqiang premier added this line to start the annual dialogue mechanism with the Canadian Prime Minister Trudeau two prime ministers held its first annual session.	Li Keqiang will start the annual dialogue mechanism with Prime Minister Trudeau of Canada and hold the first annual dialogue between the two premiers.	Li Keqiang will initiate the annual dialogue mechanism between premiers of China and Canada during this visit, and hold the first annual dialogue with Premier Trudeau of Canada.

# Deep Learning in NLP

## Neural Conversational Model



Neural Chatbot: **Predicts the next sentence** given the previous sentences in a conversation.

The LSTM remembers facts, **understands contexts** and performs common sense reasoning in the trained domain.

Human: what is the purpose of life ?

Machine: to serve the greater good .

Human: what is the purpose of living ?

Machine: to live forever

Human: what is the purpose of existence?

Machine: to find out what happens when we get to the planet earth

Human: where are you now ?

Machine: i 'm in the middle of nowhere .

Human: what is the purpose of dying ?

Machine: to have a life

Human: what is the purpose of being intelligent ?

Machine: to find out what it is.

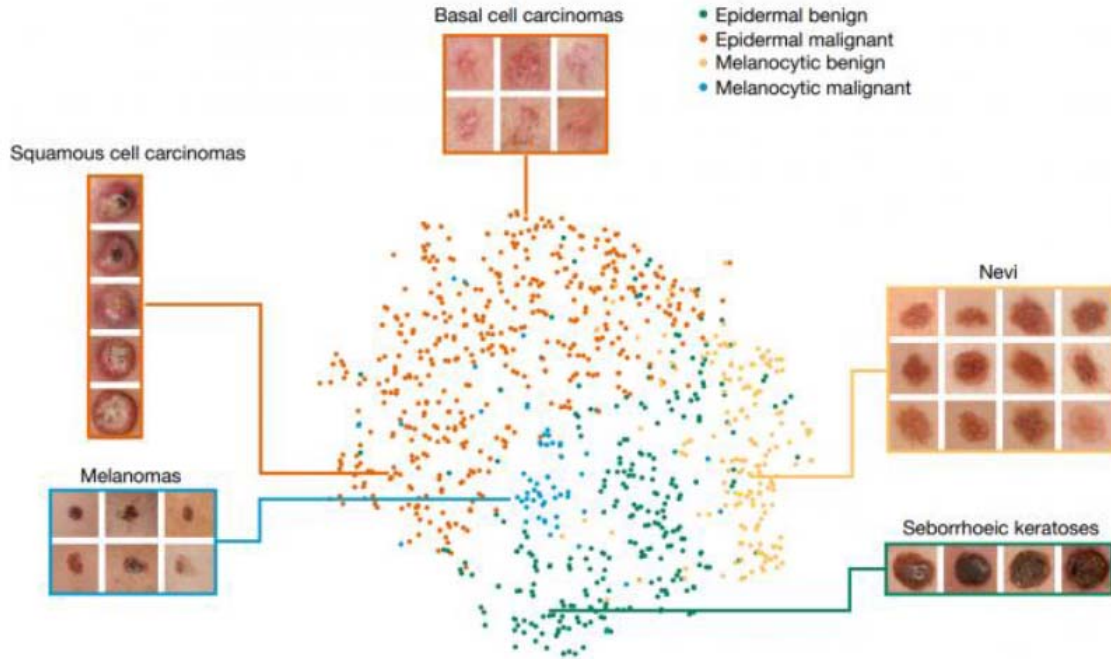
Human: what is the purpose of emotions ?

Machine: I don 't know .

Philosophical Q&A – Network trained with OpenSubtitles (movie subtitles)

# Deep Learning in Medicine

## Skin Cancer Diagnoses



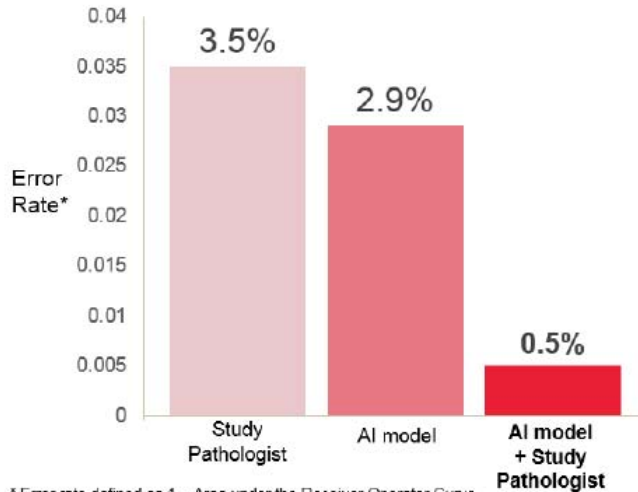
The CNN performed just as well as almost two dozen veteran dermatologists in deciding whether a lesion needed further medical attention.



# Deep Learning in Medicine

## Breast Cancer Diagnoses

(AI + Pathologist) > Pathologist



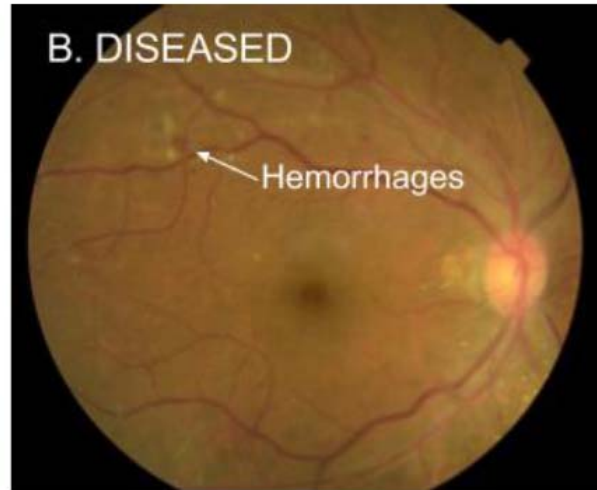
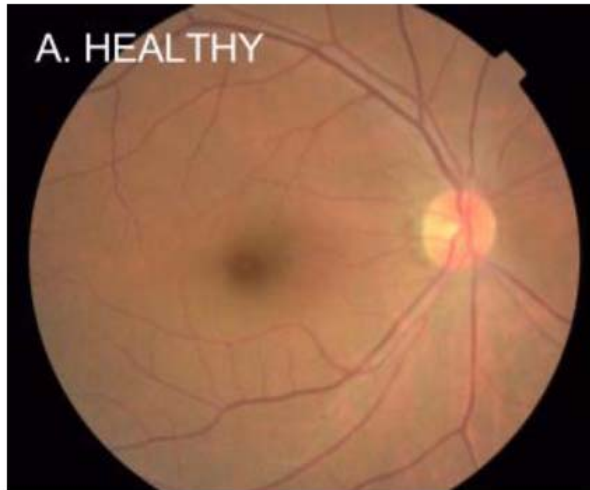
\* Error rate defined as 1 – Area under the Receiver Operator Curve

\*\* A study pathologist, blinded to the ground truth diagnoses, independently scored all evaluation slides.

Deep Learning **drops error rate for breast cancer Diagnoses by 85%**. Researchers trained their models with millions of labeled images to find the probability that a patch contains cancer, eventually creating tumor probability heatmaps.

# Deep Learning in Medicine

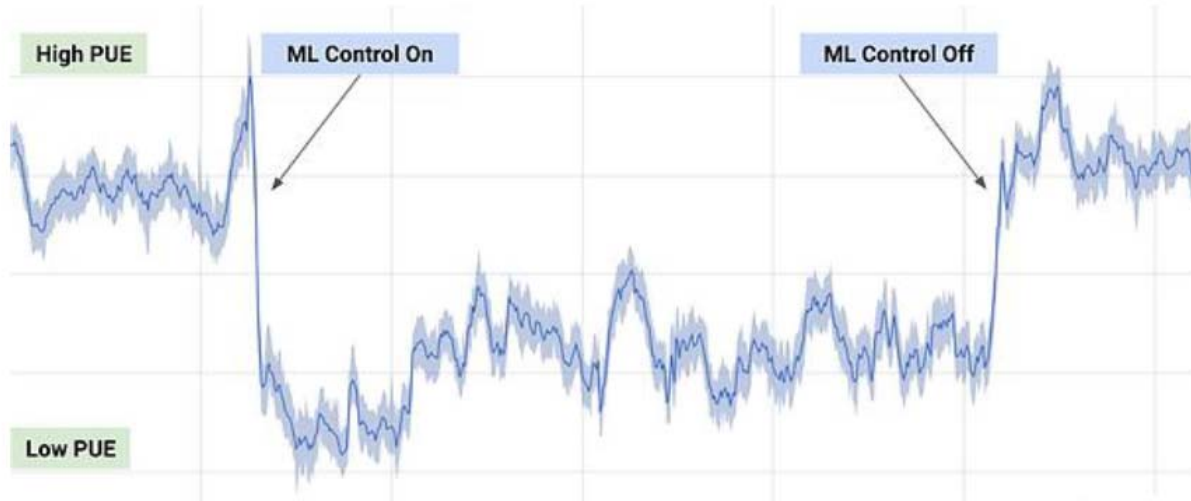
Detection of diabetic eye disease



Their deep learning algorithm performed better than the median board-certified ophthalmologist in assessing signs of diabetic retinopathy

# Deep Learning in Science

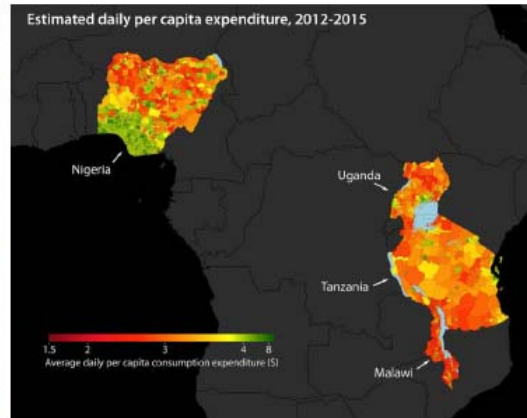
## Saving Energy



DeepMind AI **reduces data center cooling bill by 40%** using a system of neural networks trained on different operating scenarios and parameters within Google's data centers.

# Deep Learning in Science

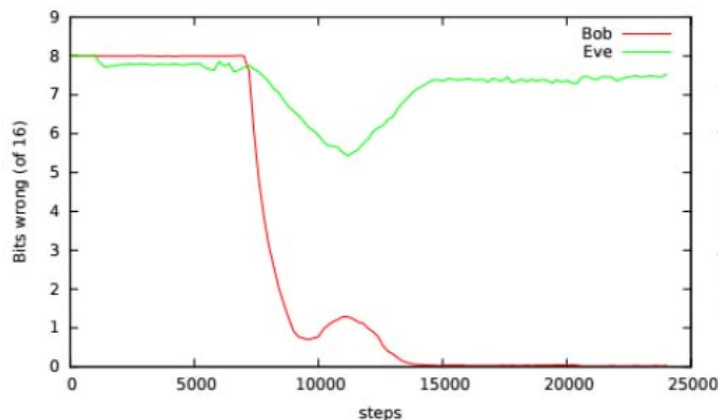
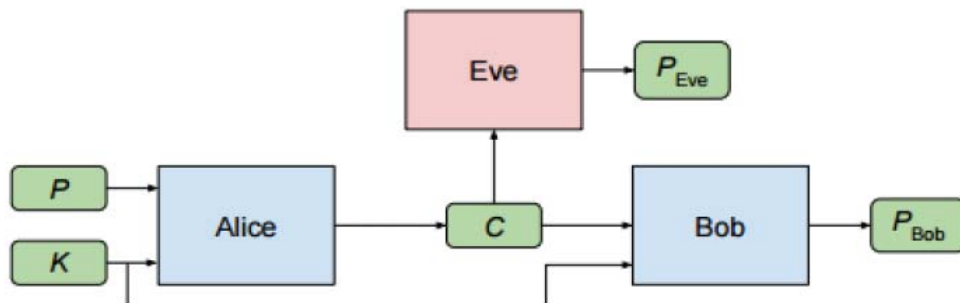
## Mapping Poverty



Combining satellite imagery and machine learning to predict poverty. A **deep-learning algorithm that can recognize signs of poverty in satellite images** – such as condition of roads – by sorting through a million images to accurately identify economic conditions in five African countries.

# Deep Learning in Cryptography

Learning to encrypt and decrypt communication



This end-to-end adversarially trained architecture **learned how to perform forms of encryption and decryption**, and also how to apply these operations selectively in order to meet confidentiality goals.

# Deep Learning in Robotics

Learning skills from shared experiences



Google researchers tasked robots with trying to move their arms to goal locations, or reaching to and opening a door. Each robot has a copy of a neural network that allows it to estimate the value of taking a given action in a given state. Through a **trial-and-error process these robots are able to acquire new skills.**

# Deep Learning in Google Products



RankBrain (Search): In few months, RankBrain has become the **third-most important signal** contributing to the search result.



Speech Recognition (Google Assistant): 30% reduction in Word Error Rate for English. **Biggest single improvement in 20 years** of speech research.



Photo Search (Google Photos): Error rate of just 5% which is **as good as humans performing** the same task. Also, superhuman performance in face recognition.

کاربردهای یادگیری عمیق

# ۲

## ابزارهای یادگیری عمیق



# Deep Learning - Tools

Its all Open Source



theano

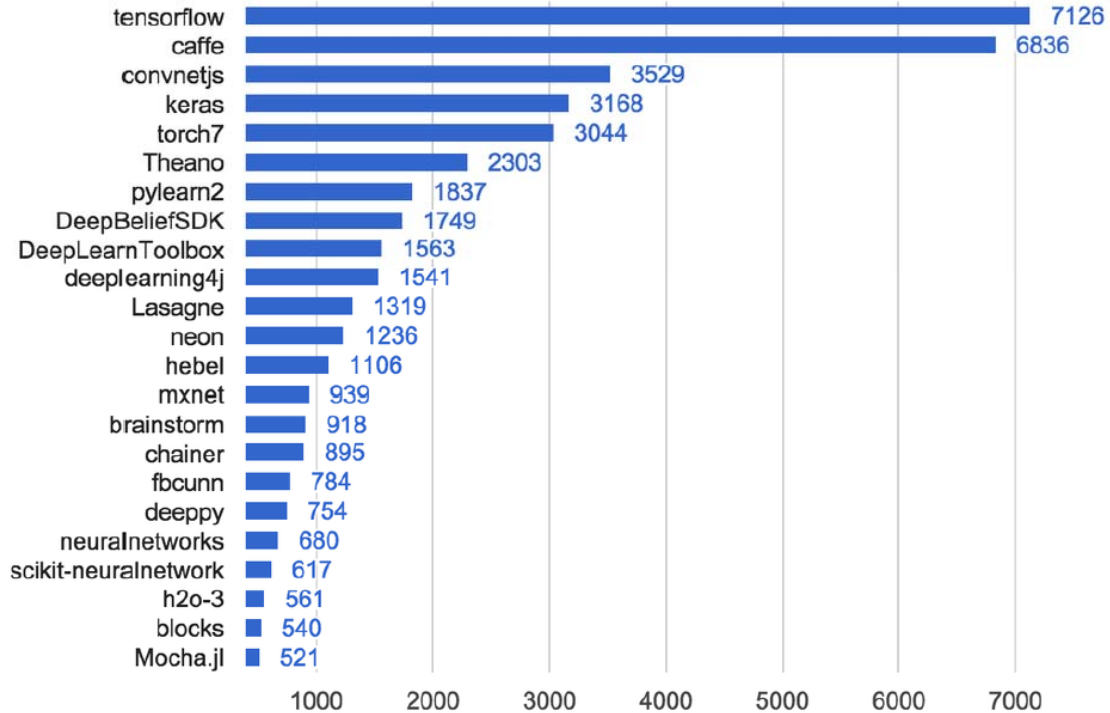


Caffe



# Deep Learning - Tools

Its all Open Source



# Deep Learning - Tools

Computing is affordable



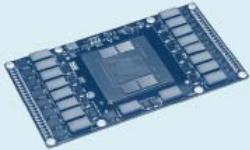
AWS EC2 GPU Spot Instance: *g2.2xlarge* - *\$0.0782 per Hour*



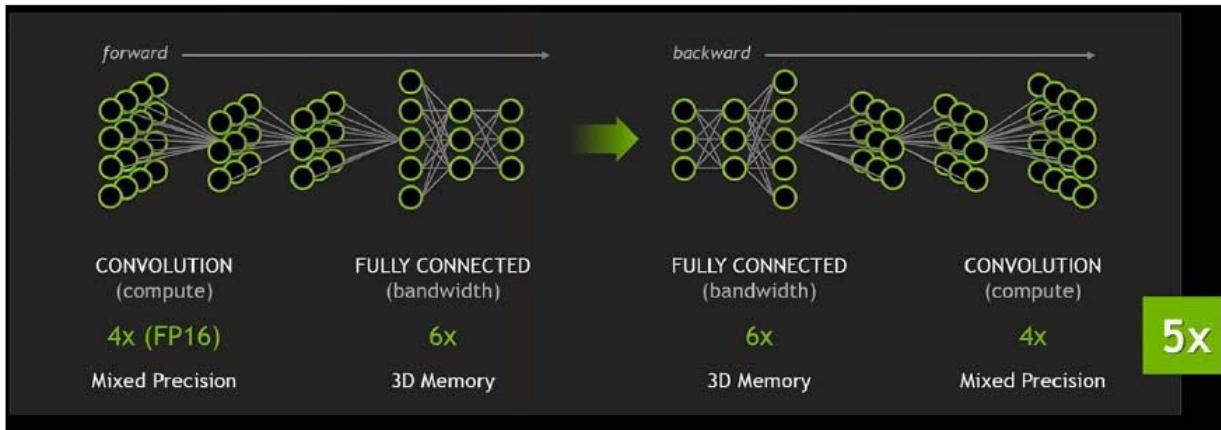
The DIGITS DevBox combines the world's best hardware (4 GPUs), software, and systems engineering for deep learning in a powerful solution that can fit under your desk. *Cost: \$15k*

# Outlook

## NVIDIA Pascal



NVIDIA's Pascal GPU architecture will **accelerate deep learning applications** up to 10X beyond the speed of its current-generation Maxwell processors.



کاربردهای یادگیری عمیق

۳

# دورنمای یادگیری عمیق

# Outlook

## Goal-based AI

### Traditional Programming



### Machine Learning



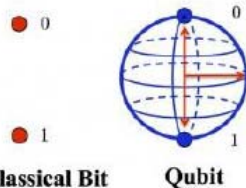
### Goal-based AI



# Outlook

## Artificial Quantum Intelligence

Quantum Artificial Intelligence Lab is a joint initiative of NASA and Google to study how **quantum computing might advance machine learning**. This type of computing may provide the most creative and parallelized problem-solving process under the known laws of physics.



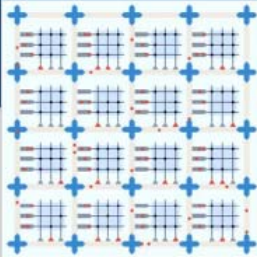
Quantum computers handle what are called **quantum bits** or qubits that can readily have a **value of one or zero or anything in between**.



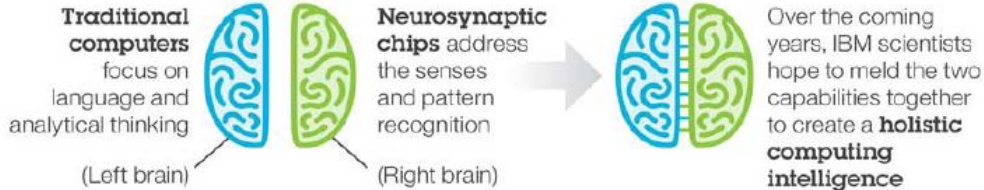
Quantum computing represents a paradigm shift, a radical change in the way we do computing and at a scale that has unimaginable power – *Eric Ladizinsky (Co-founder D-Wave)*

# Outlook

## Neuromorphic Chips



IBM TrueNorth is a **brain-inspired computer chip** that implements networks of integrate-and-fire spiking artificial neurons and uses only a tiny 70 mw of power – **orders of magnitude less energy** than traditional chips. The system is designed to be able to run deep-learning algorithms.



1 million  
Programmable  
Neurons



256 million  
Programmable  
Synapses



4096  
Neurosynaptic  
Cores



# Outlook

## The Enabler



Many of the **biggest problems facing humanity** today, like curing diseases or addressing climate change, would be **vastly easier** with the help of AI.



The **Big Bang for Self-Driving Cars** (*10-15 years*). Fully autonomous taxi systems will change the paradigm of the need to own a car.



AI will fuel a **medical revolution** (*5-10 years*) by enabling far more efficient drug discovery, diagnoses and research.

# Takeaways



Machines that **learn to represent the world** from experience.



Deep Learning is **no magic!** Just statistics in a black box, but exceptional effective at learning patterns.



We haven't figured out **creativity** and **human-empathy**.



Transitioning from research to consumer products. Will make the tools you use every day **work better, faster and smarter**.

کاربردهای یادگیری عمیق

۴

منابع

## منبع اصلی

