Deep Learning

Ali Ghodsi

Deep Learning

Deep learning attempts to learn representations of data with multiple levels of abstraction. Deep learning usually refers to a set of algorithms and computational models that are composed of multiple processing layers. These methods have significantly improved the state-of-the-art in many domains including, speech recognition, classification, pattern recognition, drug discovery, and genomics.

Success Stories

Deep Learning Machine Teaches Itself Chess in 72 Hours, Plays at International Master Level. An artificial intelligence machine

plays chess by evaluating the board rather than using brute force to work out every possible move.

Success Stories

- Google AI algorithm masters ancient game of Go
- **Deep-learning software defeats human professional for first time.**

Face2Face

- Face2Face: Real-time Face Capture and Reenactment of RGB Videos

Justus Thies

Michael Zollhöfer

Marc Stamminger

Christian Theobalt

Matthias Nießner

Computer Vision and Pattern Recognition (CVPR), IEEE, June 2016

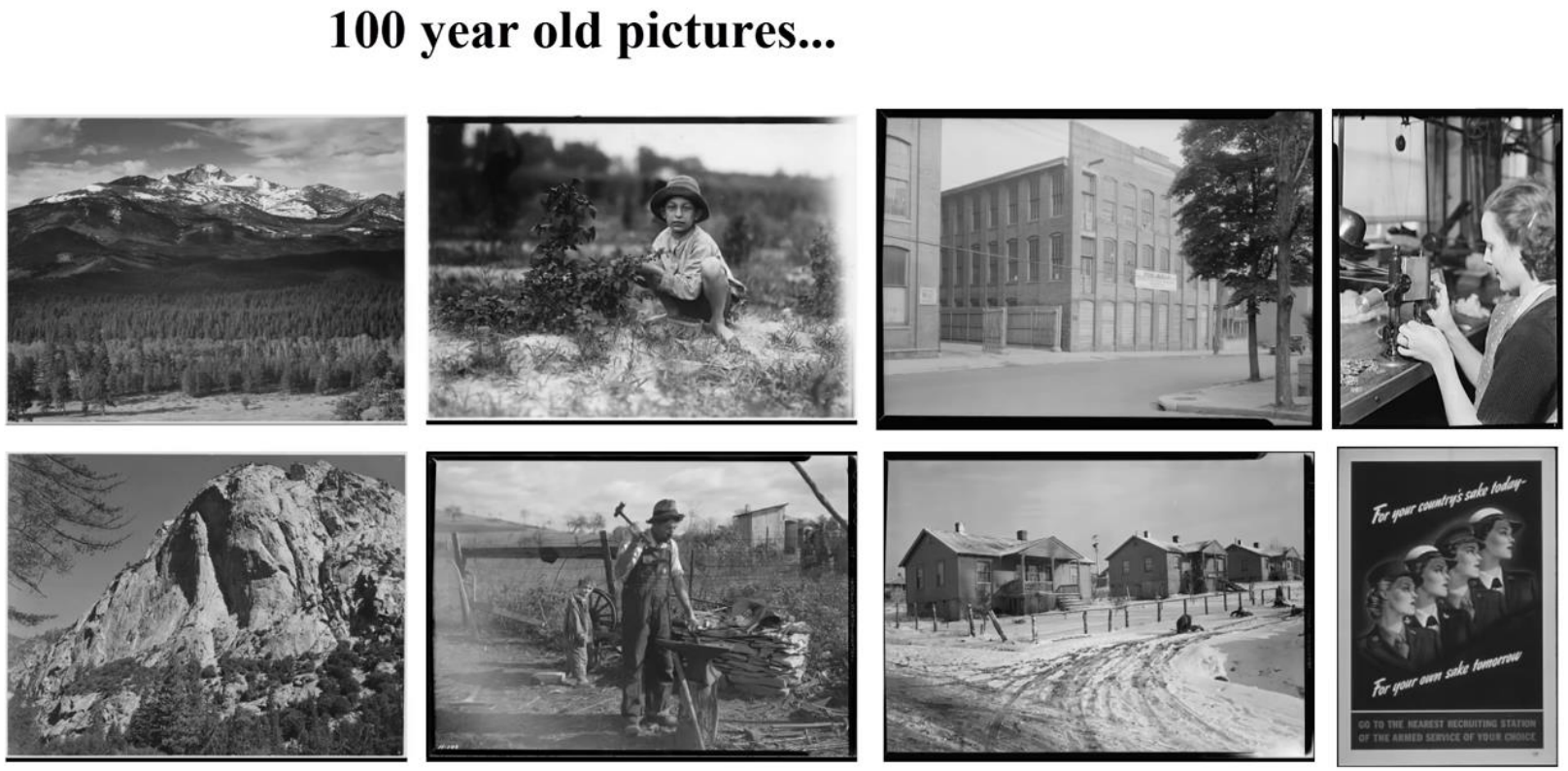
Face2Face: Real-time Face Capture and Reenactment of RGB Videos

Justus Thies[1], *Michael Zollhöfer*[2], *Marc Stamminger*[1], *Christian Theobalt*[2], *Matthias Nießner*[3]

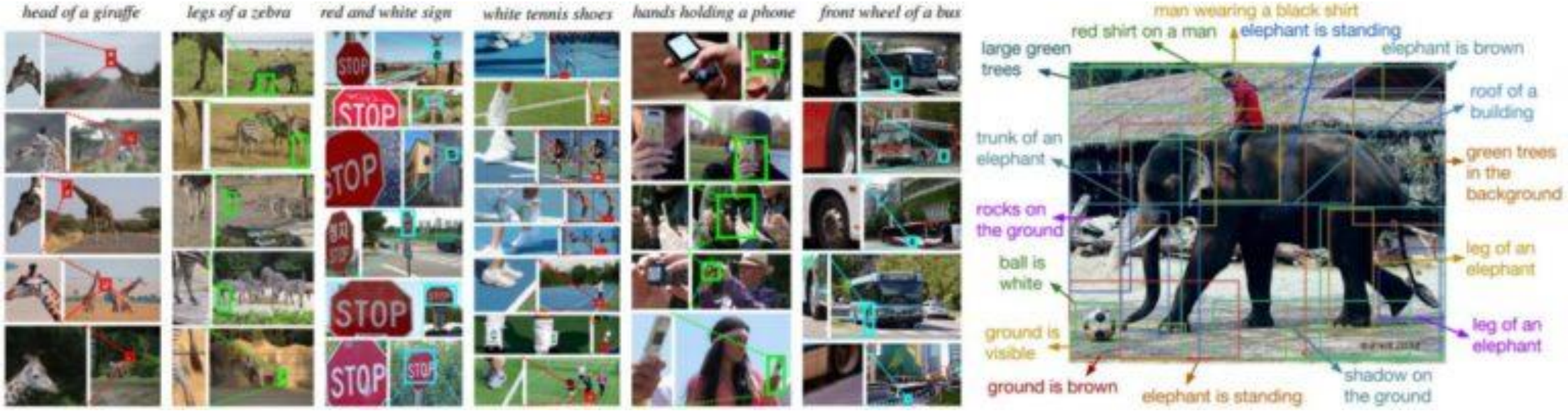
[1]University of Erlangen-Nuremberg
[2]Max-Planck-Institute for Informatics
[3]Stanford University

CVPR 2016 (Oral)

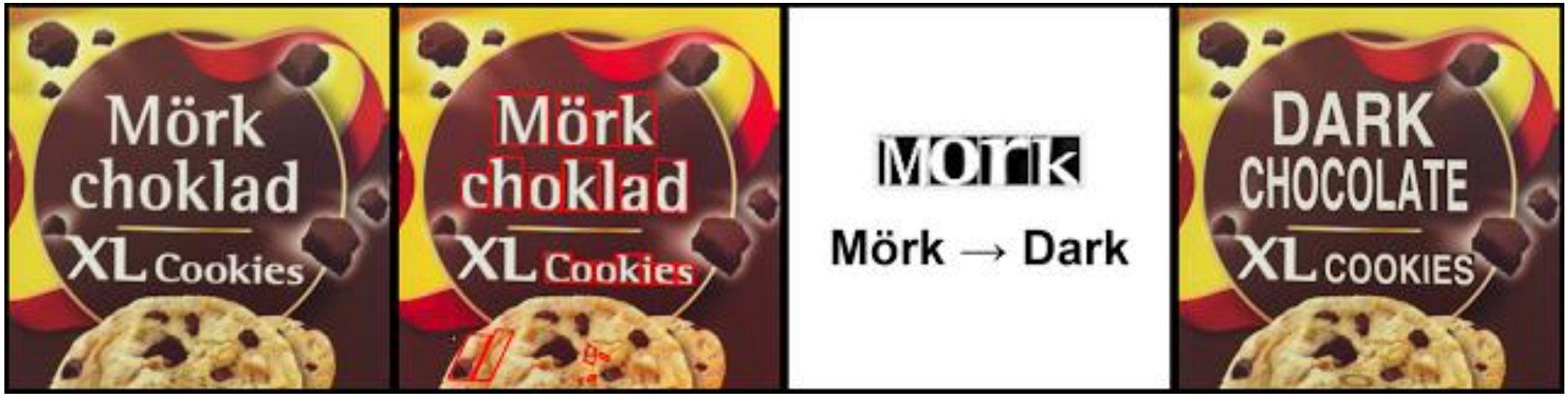
Restore colors in B&W photos and videos



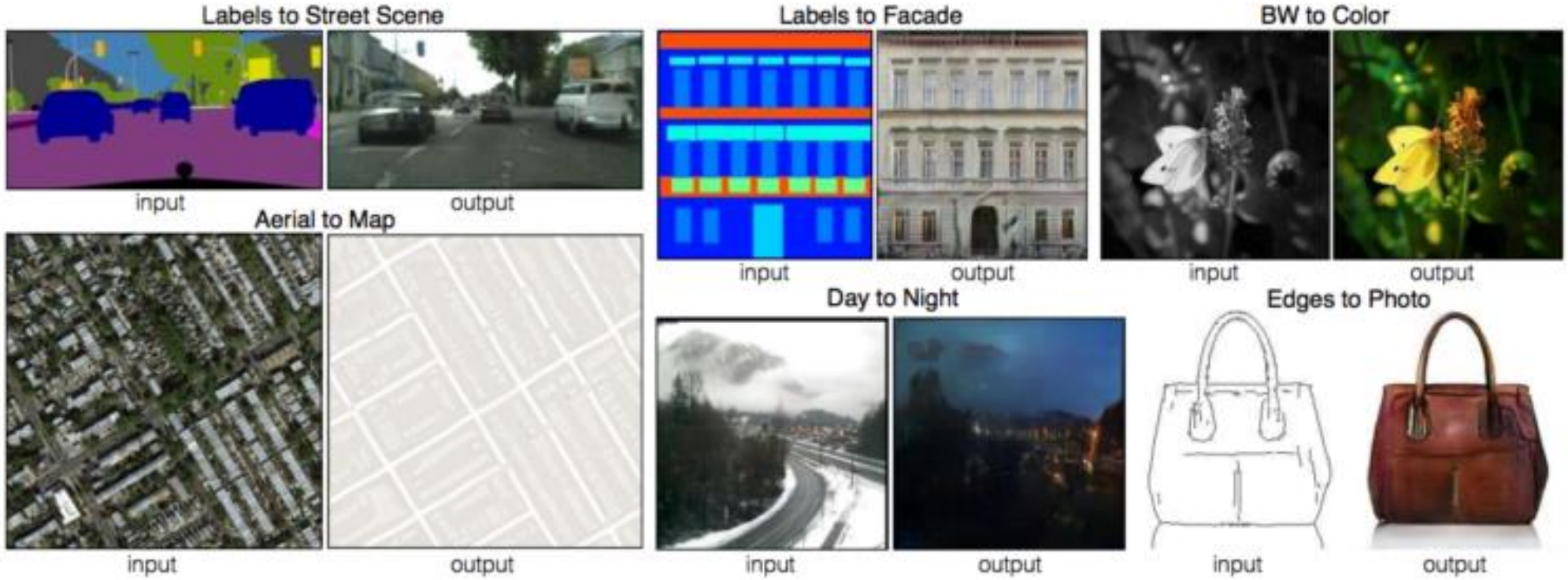
Describing photos



Translation

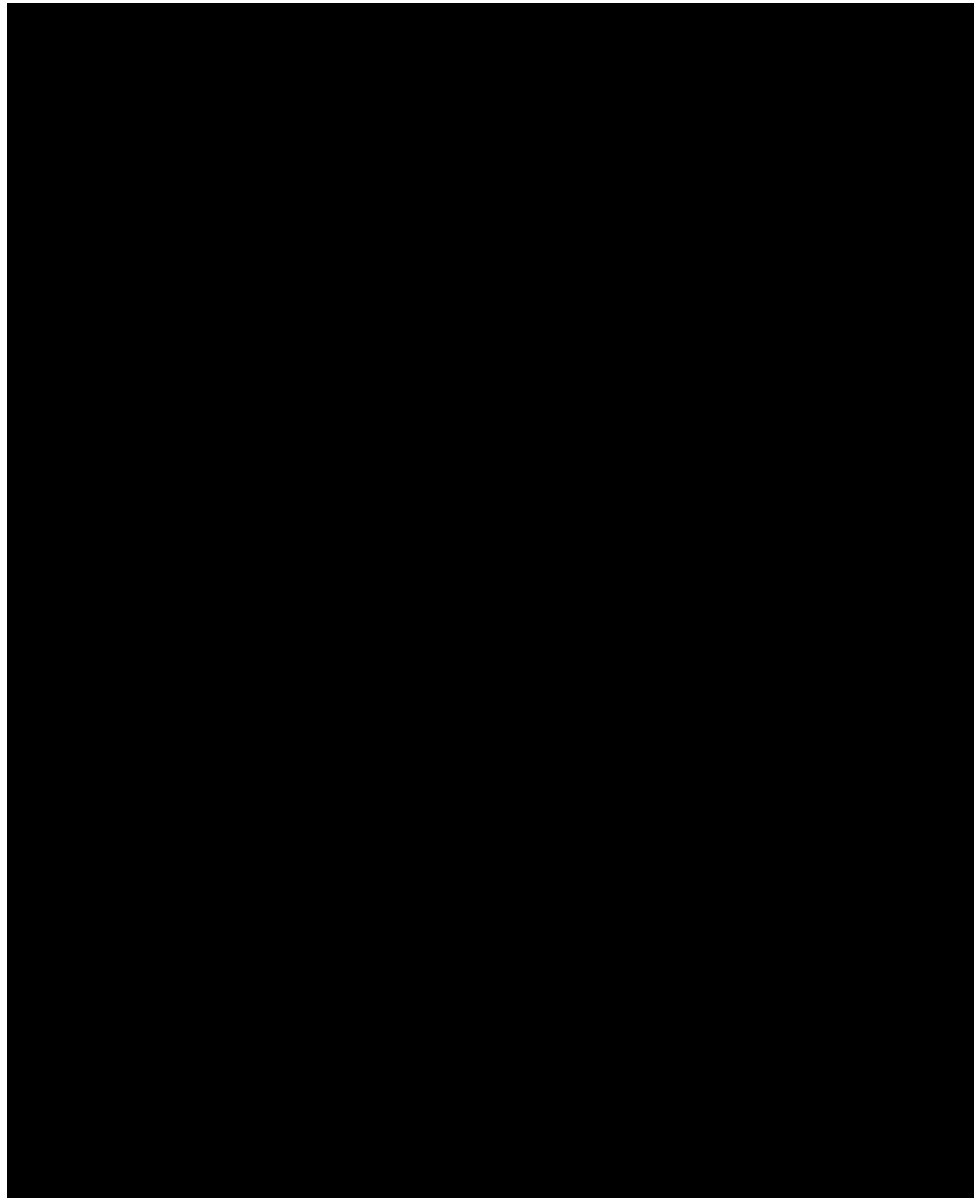


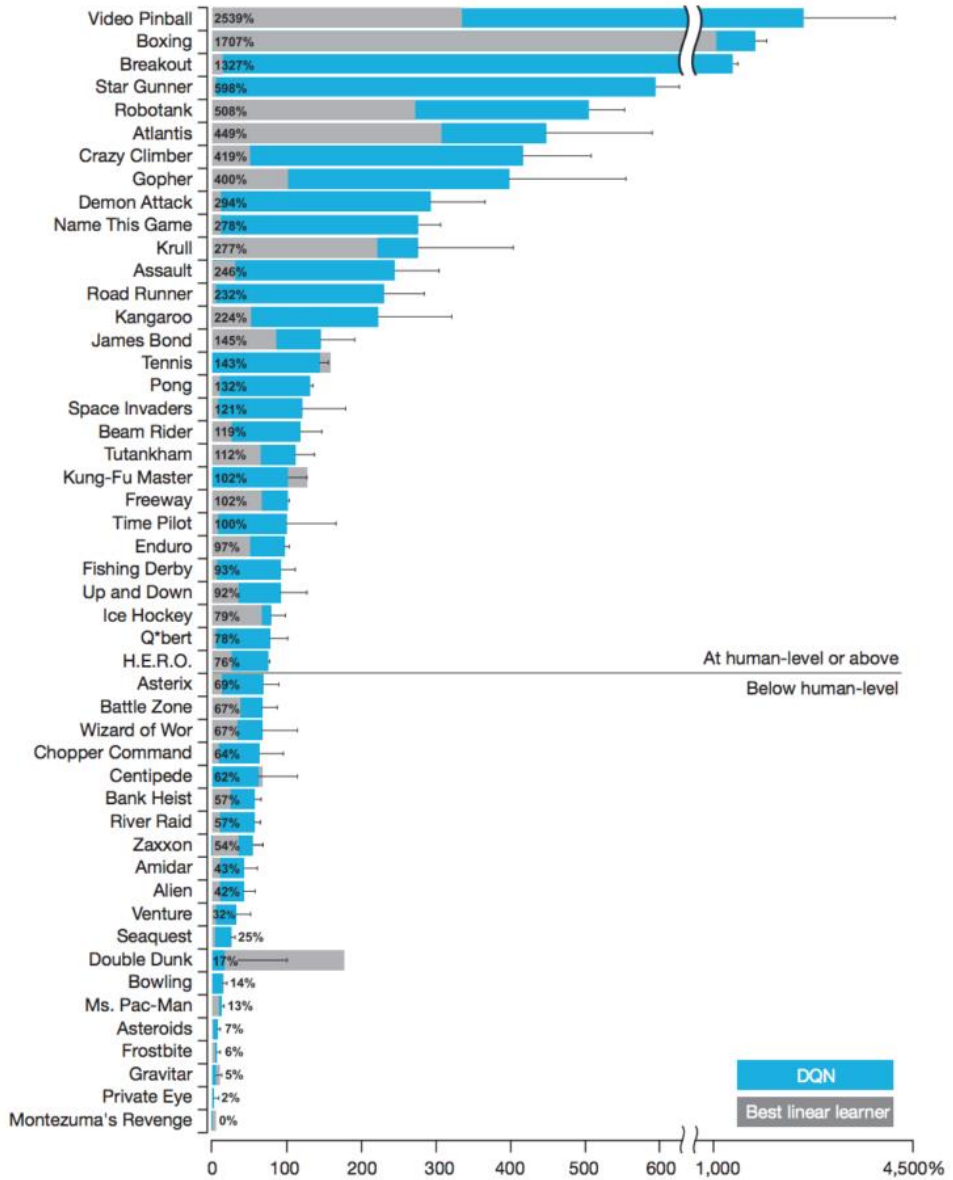
Create new images



Create new images

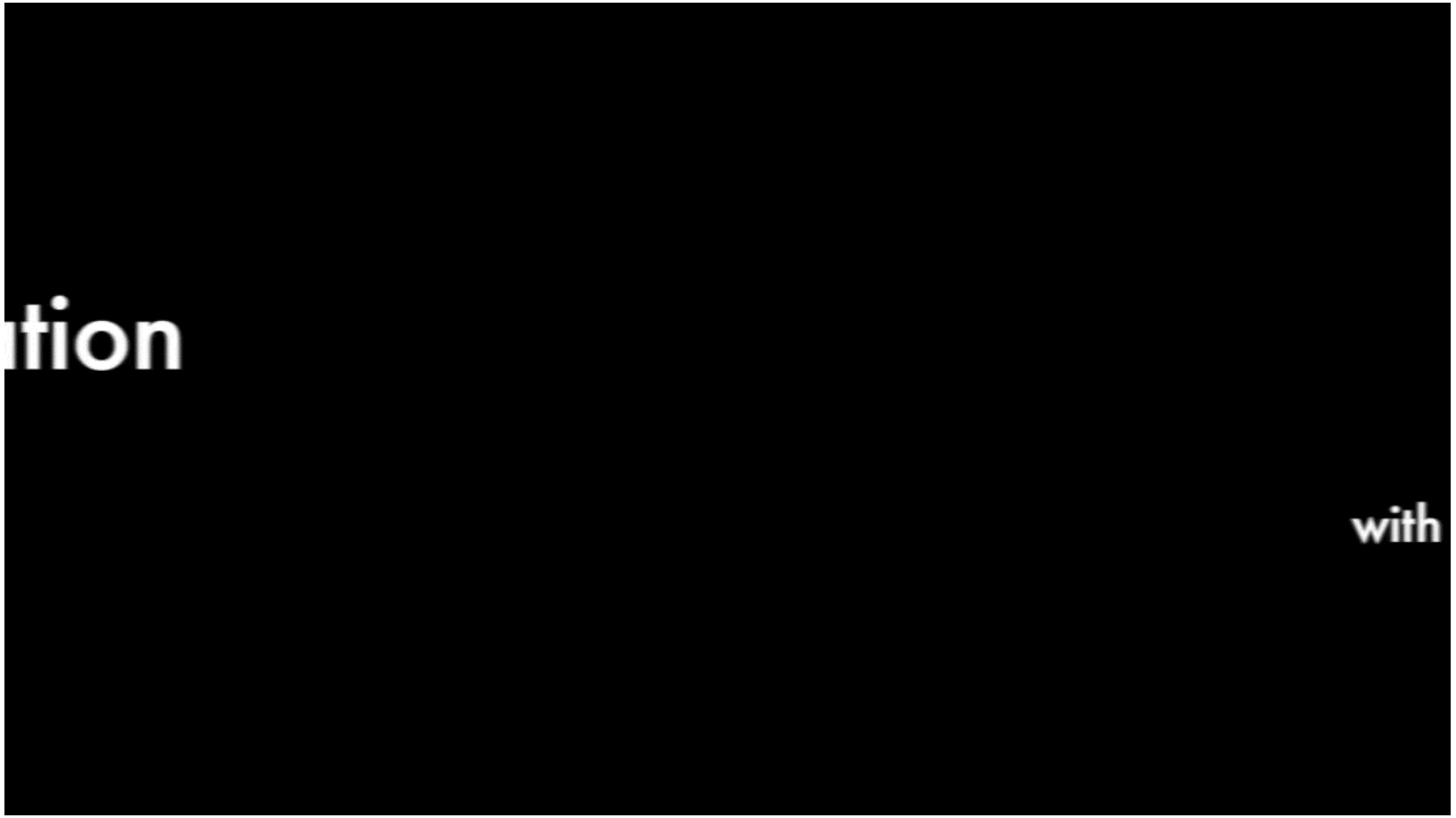
Winning Atari Breakout



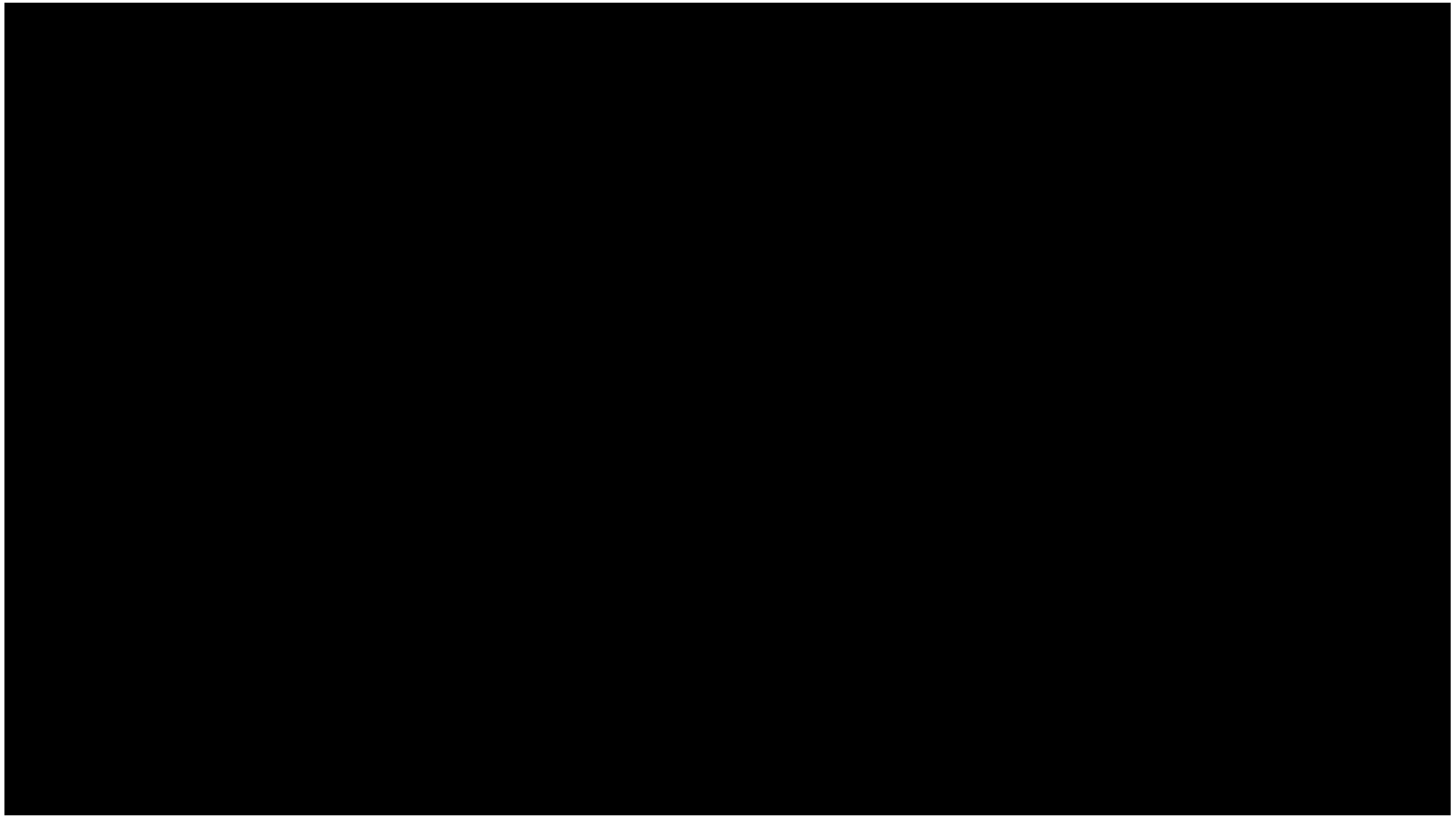

Video Pinball 2539%
Boxing 1707%
Breakout 1327%
Star Gunner 598%
Robotank 508%
Atlantis 449%
Crazy Climber 419%
Gopher 400%
Demon Attack 294%
Name This Game 278%
Krull 277%
Assault 246%
Road Runner 232%
Kangaroo 224%
James Bond 145%
Tennis 143%
Pong 132%
Space Invaders 121%
Beam Rider 119%
Tutankham 112%
Kung-Fu Master 102%
Freeway 102%
Time Pilot 100%
Enduro 97%
Fishing Derby 93%
Up and Down 92%
Ice Hockey 79%
Q*bert 78%
H.E.R.O. 76%
Asterix 69%
Battle Zone 67%
Wizard of Wor 67%
Chopper Command 64%
Centipede 62%
Bank Heist 57%
River Raid 57%
Zaxxon 54%
Amidar 43%
Alien 42%
Venture 32%
Seaquest 25%
Double Dunk 17%
Bowling 14%
Ms. Pac-Man 13%
Asteroids 7%
Frostbite 6%
Gravitar 5%
Private Eye 2%
Montezuma's Revenge 0%
At human-level or above
Below human-level
DQN
Best linear learner
0
100
200
300
400
500
600
1,000
4,500%

Self-driving cars

Music composition



LipNet



Success Stories

Word2vec , Mikolov, 2013.

$$\text{king} \quad \text{man} + \text{woman} = \text{queen}$$

Success Stories

- day + night =

- flying + sailing =

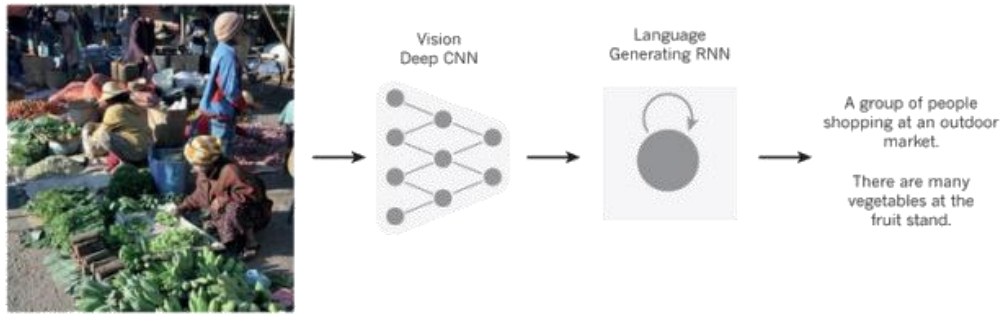
- bowl + box =

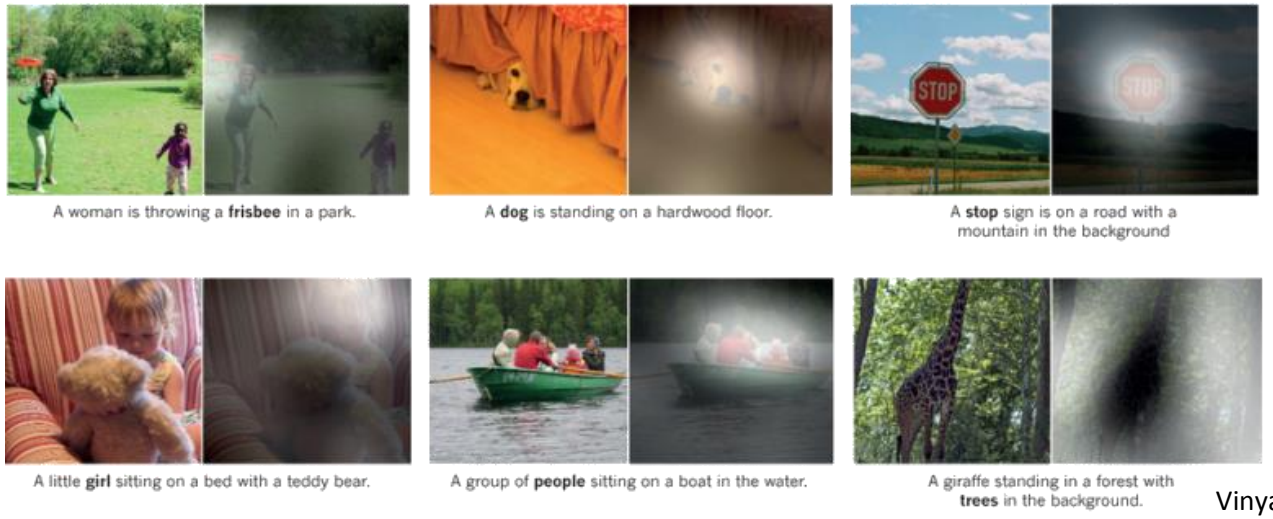
- box + bowl =

Nearest Image

(Kiros, Salakhutdinov, Zemel, TACL 2015)

Success Stories

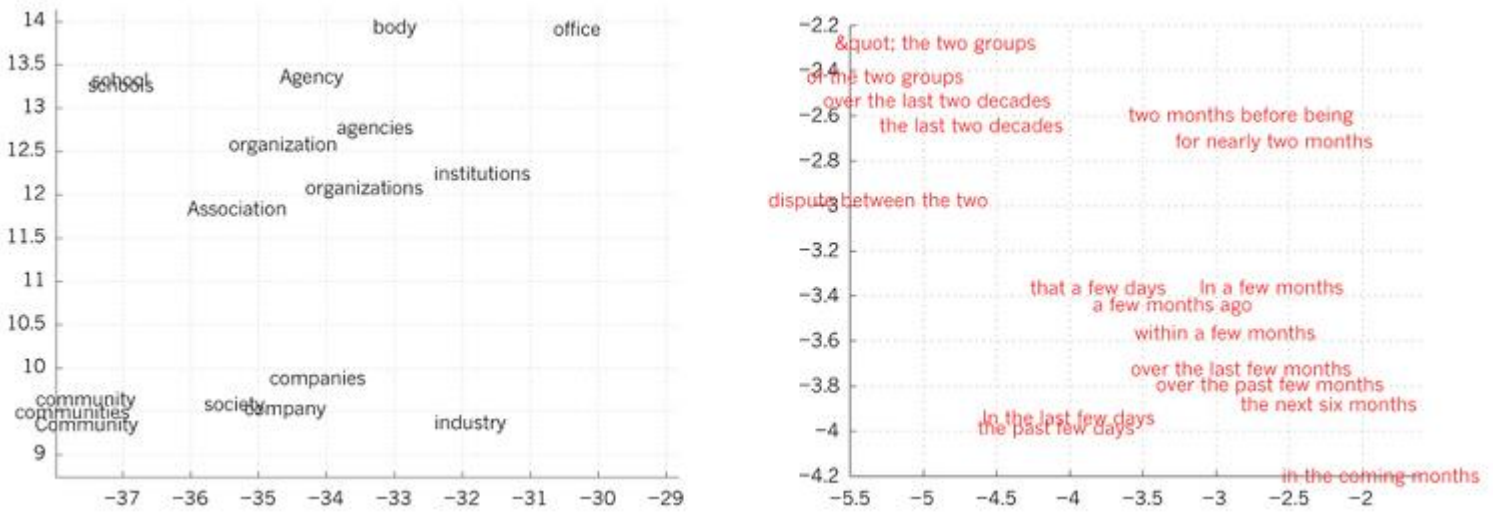




Vinyals et. al 2014

Captions generated by a recurrent neural network.

Success Stories



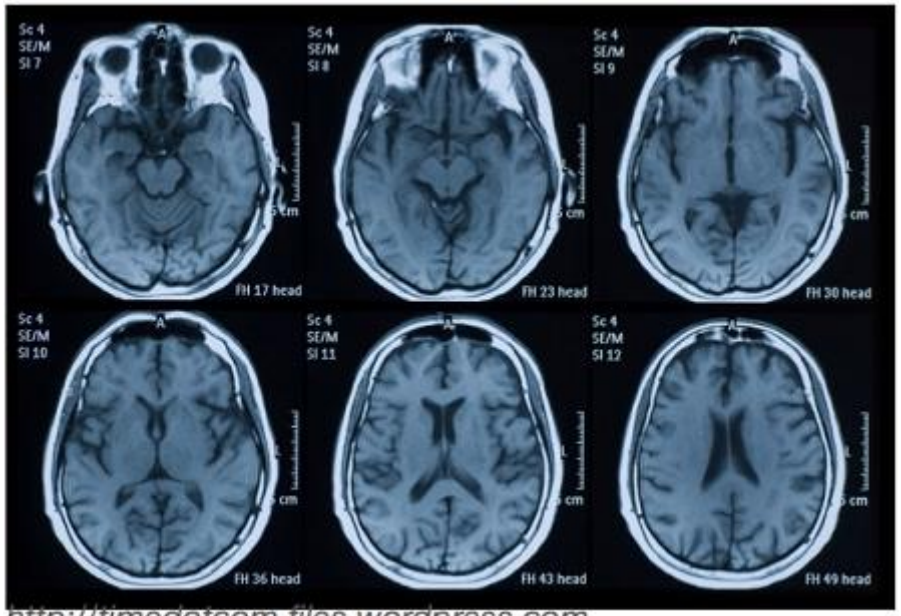
Credit: LeCun, et. al., 2015, Nature

On the left is an illustration of word representations learned for modelling language, non-linearly projected to 2D for visualization using the t-SNE algorithm. On the right is a 2D representation of phrases learned by an English-to-French encoderdecoder recurrent neural network. One can observe that semantically similar words or sequences of words are mapped to nearby representations.

Success Stories

PayPal is using deep learning via H2O, an open source predictive analytics platform, to help prevent fraudulent purchases and payment transactions.

Success Stories



http://timedotcom.files.wordpress.com

New startup Enlitic is using deep learning to process X-rays, MRIs, and other medical images to help doctors diagnose and treat complicated diseases. Enlitic uses deep learning algorithms that "are suited to discovering the subtle patterns that characterize disease profiles."

Success Stories

Credit: Hansen, 2014

AlchemyVisions Face Detection and Recognition service is able to distinguish between look-alikes such as actor Will Ferrell and Red Hot Chili Peppers drummer, Chad Smith.

Recovering sound waves from the vibrations

Davis, A., Rubinstein, M., Wadhwa, N., Mysore, G., Durand, F., and Freeman, W. T.(2014). The visual microphone: Passive recovery of sound from video. ACM Transac-tions on Graphics (Proc. SIGGRAPH), 33(4), 79:179:10.

The visual microphone

Demos

http://deeplearning.net/demos/

Tentative topics

- Feedforward Deep Networks
- Optimization and regularization for Training Deep Models
- Convolutional Networks
- Sequence Modeling: Recurrent Nets
- Auto-Encoders
- Representation Learning
- Deep Generative Models (moment matching networks)
- Generative Adversarial Networks (GANs)
- Deep learning for Natural Language Processing
- Variational Autoencoders
- Attention model, Highway network, residual network, sequence-to-sequence models

Tentative Marking Scheme

Group Project 50%
Data Challenge 30%
Paper presentation / Paper critiques 20%

Communication

All communication should take place using the Piazza discussion board. You will be sent an invitation to your UW

email address. It will include a link to a web page where you may complete the enrollment process.

Note 1: You cannot borrow part of an existing thesis work, nor can you re-use a project from another course as your final project.

Note 2: We will use wikicoursenote for paper critiques and possibly for course note (details in class).

History, McCulloch and Pitts network

1943

The first model of a neuron was invented by McCulloch (physiologists) and Pitts (logician).

The model had two inputs and a single output.

A neuron would not activate if only one of the inputs was active.

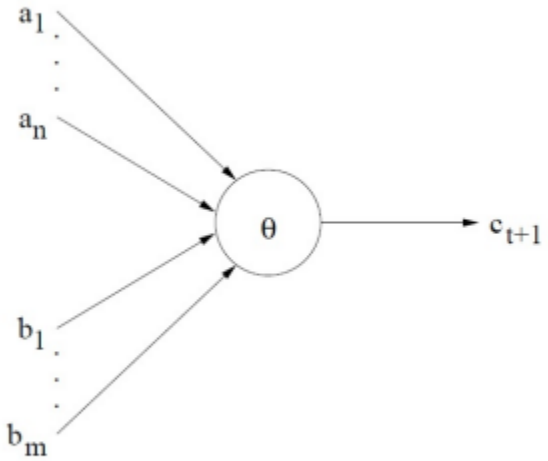
The weights for each input were equal, and the output was binary.

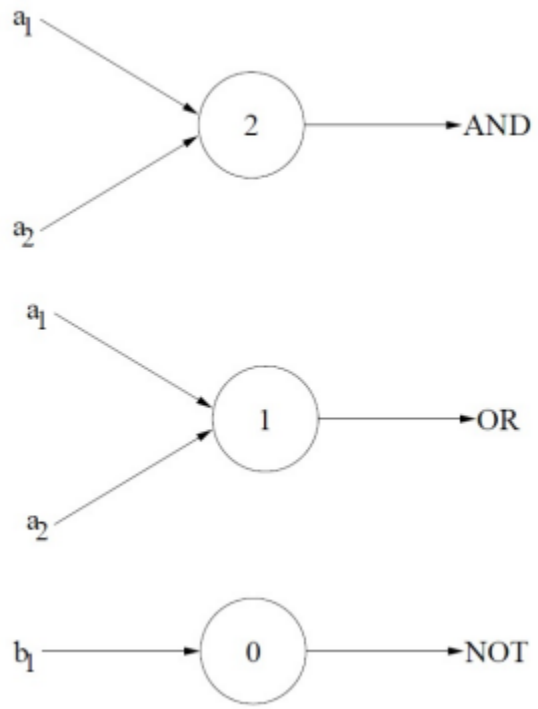
Until the inputs summed up to a certain threshold level, the output would remain zero.

The McCulloch and Pitts' neuron has become known today as a logic circuit.

History, McCulloch and Pitts network (MPN)

1943



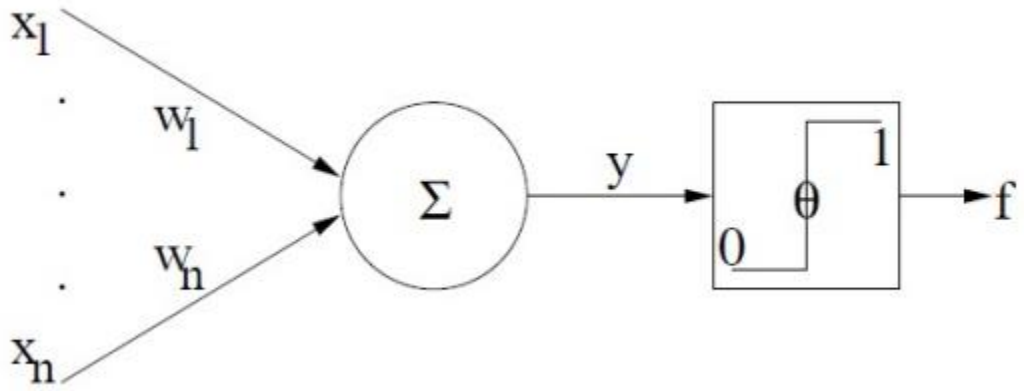


logic functions can be modeled by a network of MP-neurons

History, Perceptron

1958

The perceptron was developed by Rosenblatt (physiologist).



Credit:

Perceptron, the dream

1958

Rosenblatt randomly connected the perceptrons and changed the weights in order to achieve "learning."

Based on Rosenblatt's statements in a press confrence in 1958, The New York Times reported the perceptron to be 'the embryo of an electronic computer that [the Navy] expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.'

MPN vs. Perceptron

Apparently McCulloch and Pitts' neuron is a better model for the electrochemical process inside the neuron than the perceptron.

But perceptron is the basis and building block for the modern neural networks.

History, optimization

1960

Widrow and Hoff proposed a method for adjusting the weights. They introduced a gradient search method based on minimizing the error squared (Least Mean Squares).

In the 1960's, there were many articles promising robots that could think.

It seems there was a general belief that perceptrons could solve any problem.

History, shattered dream

1969

Minsky and Papert published their book *Perceptrons*. The book shows that perceptrons could only solve linearly separable problems.

They showed that it is not possible for perseptron to learn an XOR function.

After Perceptrons was published, researchers lost interest in perceptron and neural networks.

History, optimization

1969

Arthur E. Bryson and Yu-Chi Ho described proposed Backpropagation as a multi-stage dynamic system optimization method. (Bryson, A.E.; W.F. Denham; S.E. Dreyfus. Optimal programming problems with inequality constraints. I: Necessary conditions for extremal solutions. AIAA J. 1, 11 (1963) 2544-2550)

1972

Stephen Grossberg proposed networks capable of learning XOR function.

History

1974

Backpropagation was reinvented / applied in the context of neural networks by Paul Werbos, David E. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams.

Back propagation allowed perceptrons to be trained in a multilayer configuration.

History

1980s

The filed of artificial neural network research experienced a resurgence.

2000s

Neural network fell out of favor partly due to BP limitations.

Backpropagation Limitations

It requires labeled training data.

It is very slow in networks with multiple layers (doesn't scale well).

It can converge to poor local minima.

History, optimization

But has returned again in the 2010s, now able to train much larger networks using huge modern computing power such as GPUs. For example, in 2013 top speech recognisers now use backpropagation-trained neural networks.