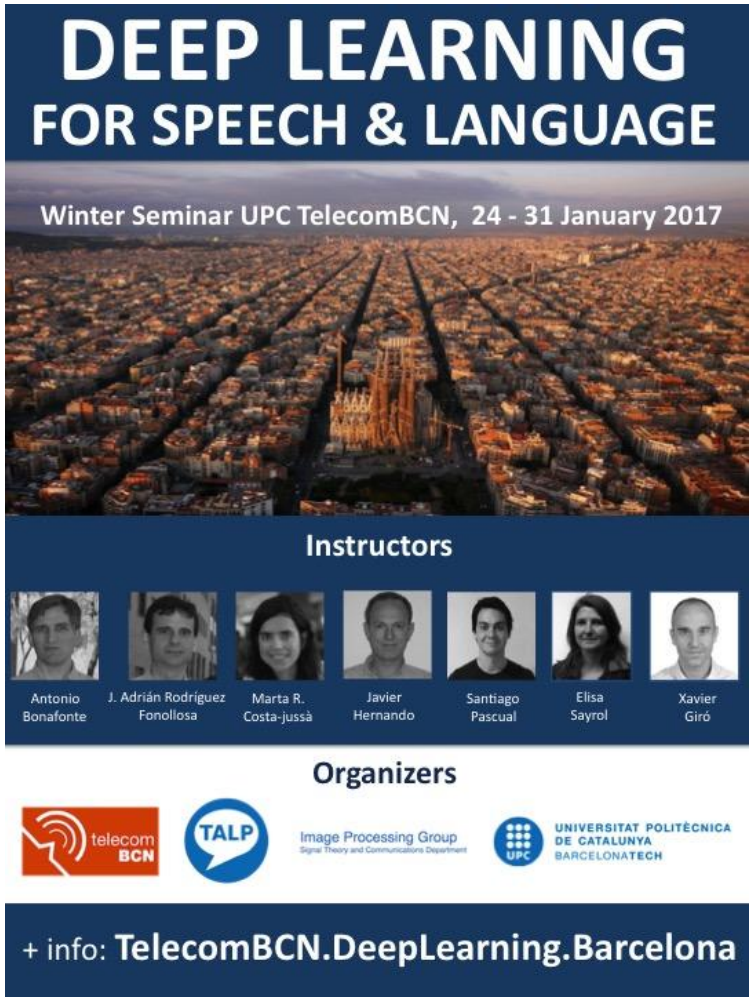



[course site]

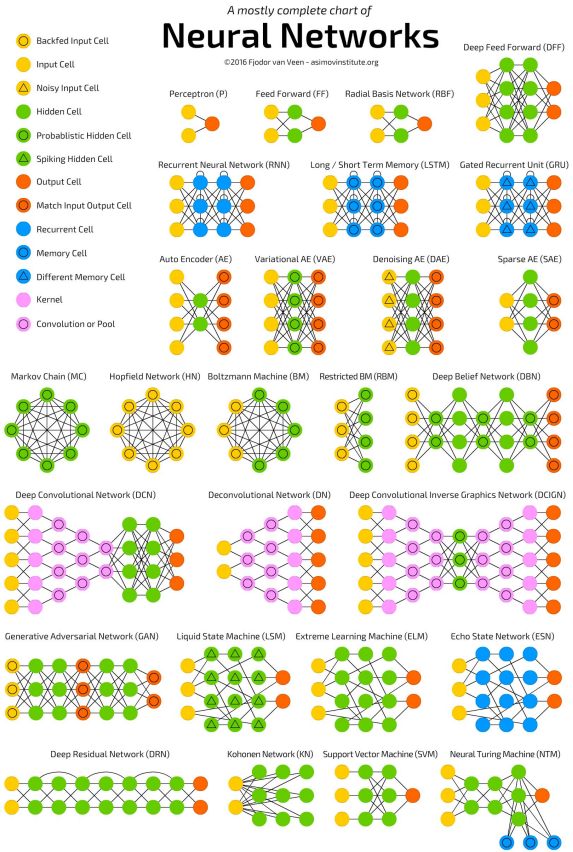
Day 2 Lecture 6

# Advanced Deep Architectures

Xavier Giró-i-Nieto

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Department of Signal Theory and Communications
*Image Processing Group*

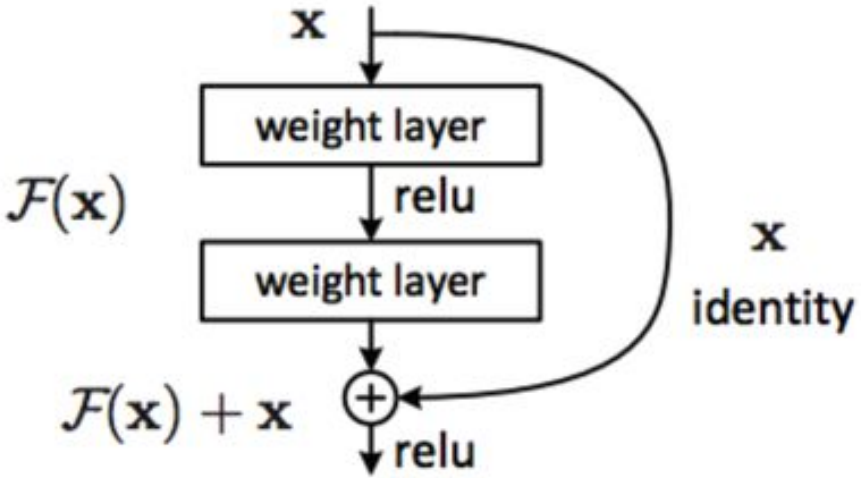
# The Full Story




F. Van Veen, [“The Neural Network Zoo”](#) (2016)

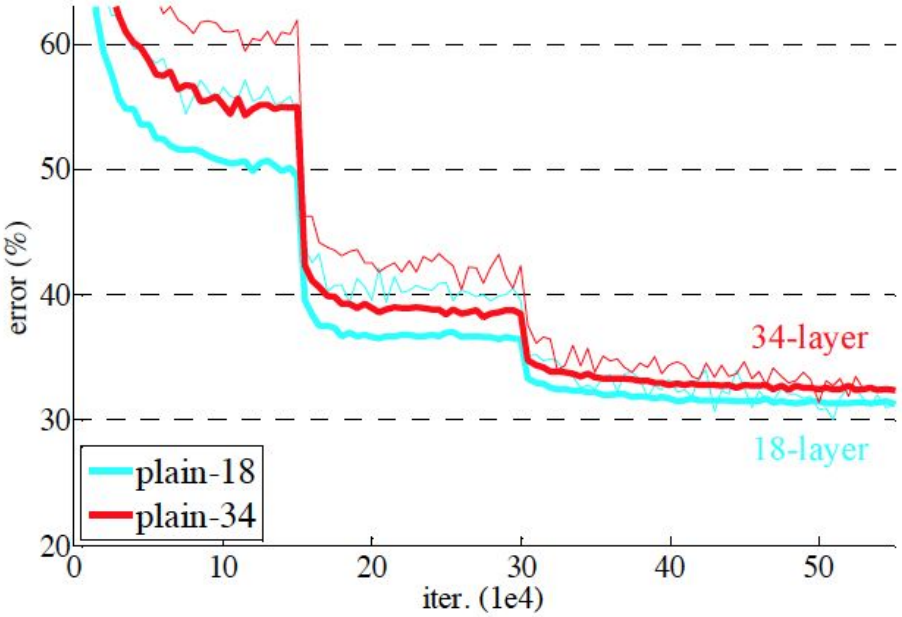
# Deep Residual Networks

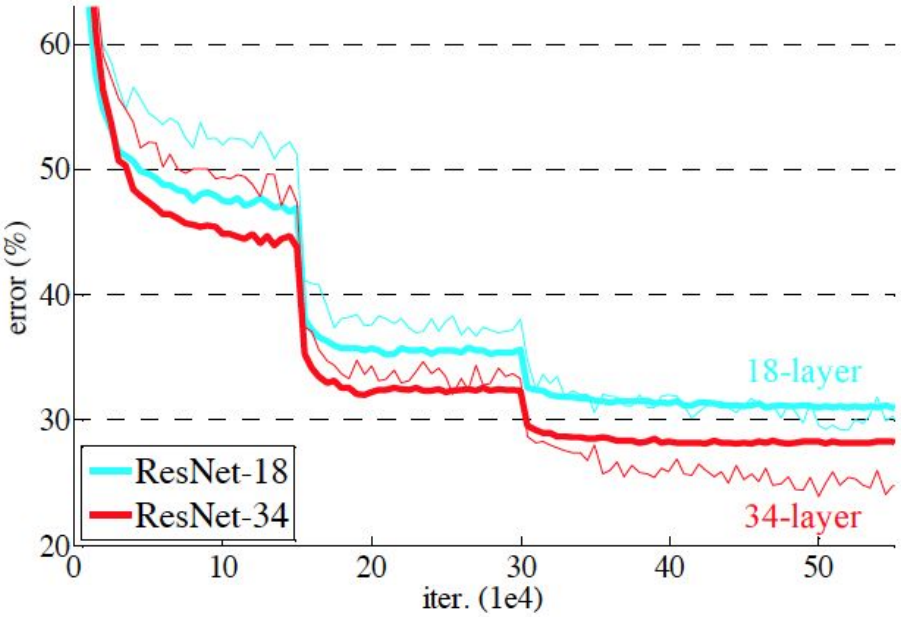
Residual learning: reformulate the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions




He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." CVPR 2016 [slides]

# Deep Residual Networks




He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." CVPR 2016 [slides]

# Deep Residual Networks

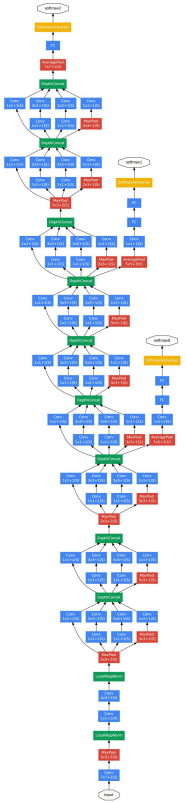
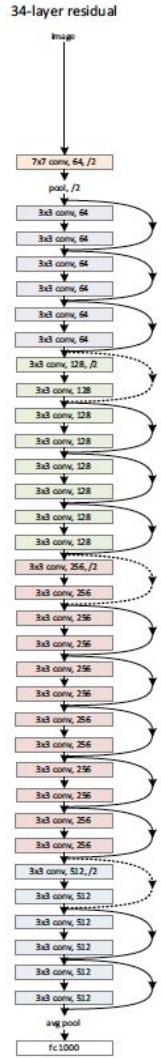
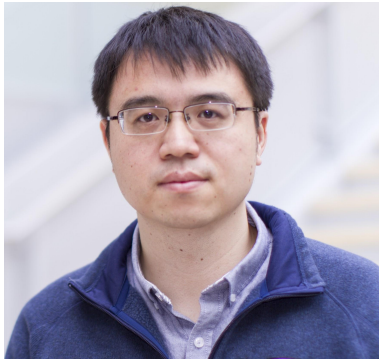


image
conv-64
conv-64
maxpool
conv-128
conv-128
maxpool
conv-256
conv-256
conv-256
conv-256
maxpool
conv-512
conv-512
conv-512
conv-512
maxpool
conv-512
conv-512
conv-512
conv-512
maxpool
FC-4096
FC-4096
FC-1000
softmax

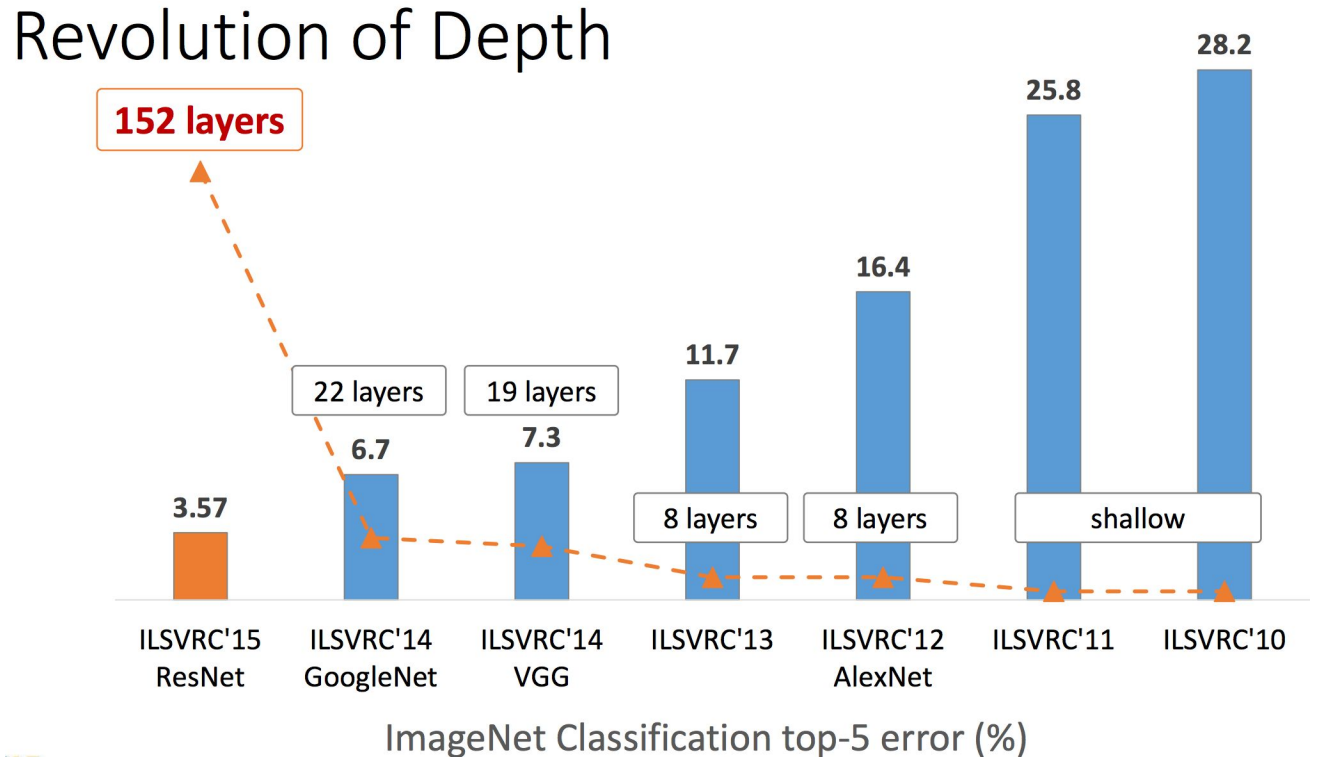



Microsoft Research

3.6% top 5 error…
with 152 layers !!

# Deep Residual Networks




He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." CVPR 2016 [slides]

# Deep Residual Networks

## ResNext

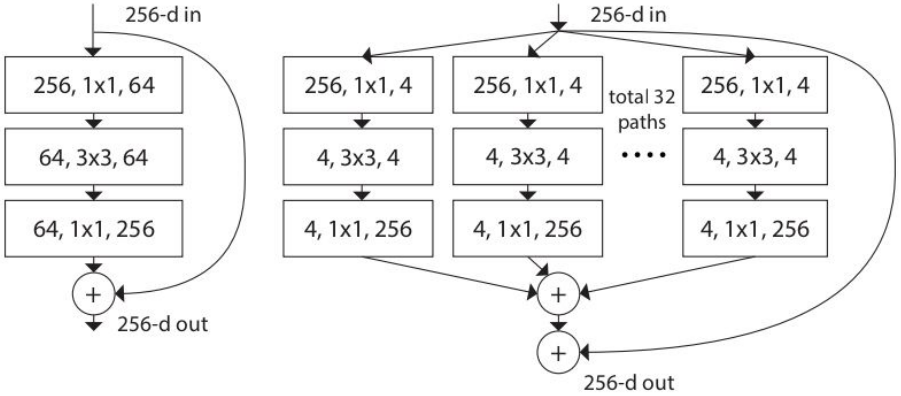



Figure 1. **Left**: A block of ResNet [13]. **Right**: A block of ResNeXt with cardinality = 32, with roughly the same complexity. A layer is shown as (# in channels, filter size, # out channels).

| | 224×224 | | 320×320 / 299×299 | |
|---|---|---|---|---|
| | top-1 err | top-5 err | top-1 err | top-5 err |
| ResNet-101 [13] | 22.0 | 6.0 | - | - |
| ResNet-200 [14] | 21.7 | 5.8 | 20.1 | 4.8 |
| Inception-v3 [38] | - | - | 21.2 | 5.6 |
| Inception-v4 [36] | - | - | 20.0 | 5.0 |
| Inception-ResNet-v2 [36] | - | - | 19.9 | 4.9 |
| ResNeXt-101 (**64** × **4d**) | 20.4 | 5.3 | **19.1** | **4.4** |

Table 5. State-of-the-art models on the ImageNet-1K validation set (single-crop testing). The test size of ResNet/ResNeXt is 224×224 and 320×320 as in [14] and of the Inception models is 299×299.

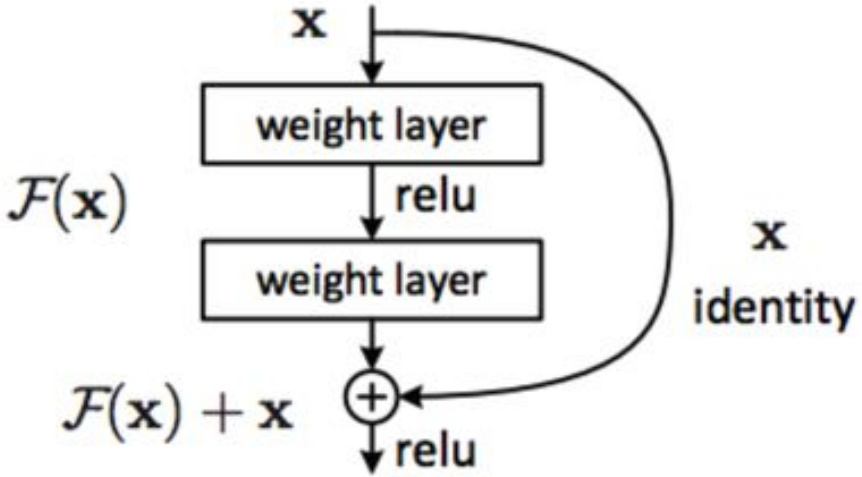
Xie, Saining, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. "Aggregated residual transformations for deep neural networks." arXiv preprint arXiv:1611.05431 (2016).

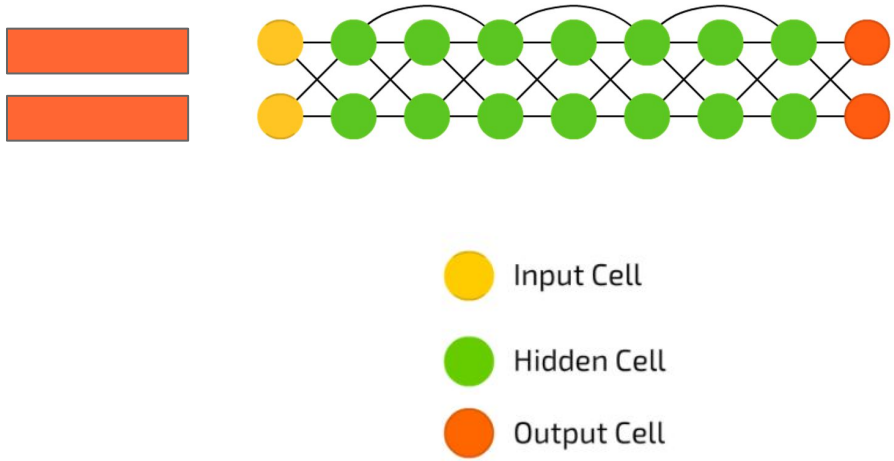
# Deep Residual Networks

## Cross-residuals

Xie, Saining, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. "Aggregated residual transformations for deep neural networks." arXiv preprint arXiv:1611.05431 (2016).

# Deep Residual Networks




F. Van Veen, "The Neural Network Zoo" (2016)

# Skip connections

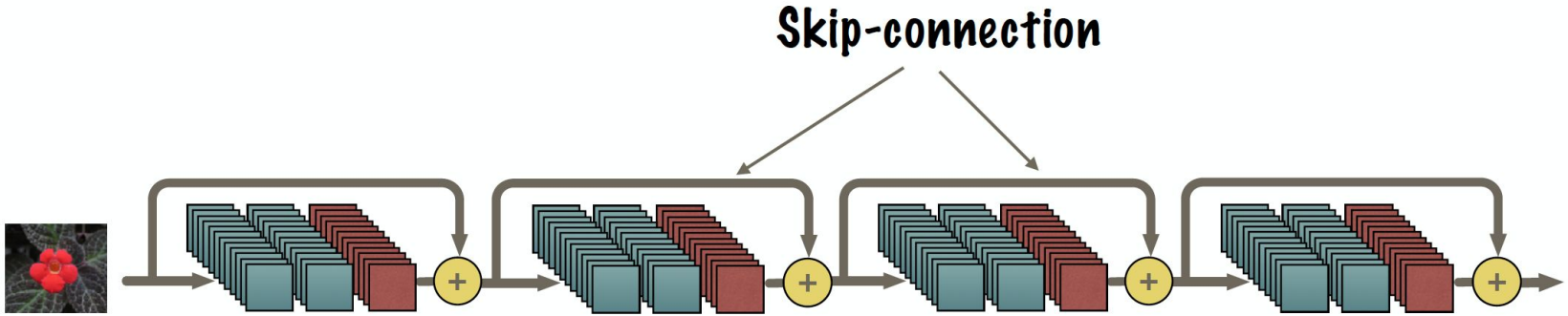
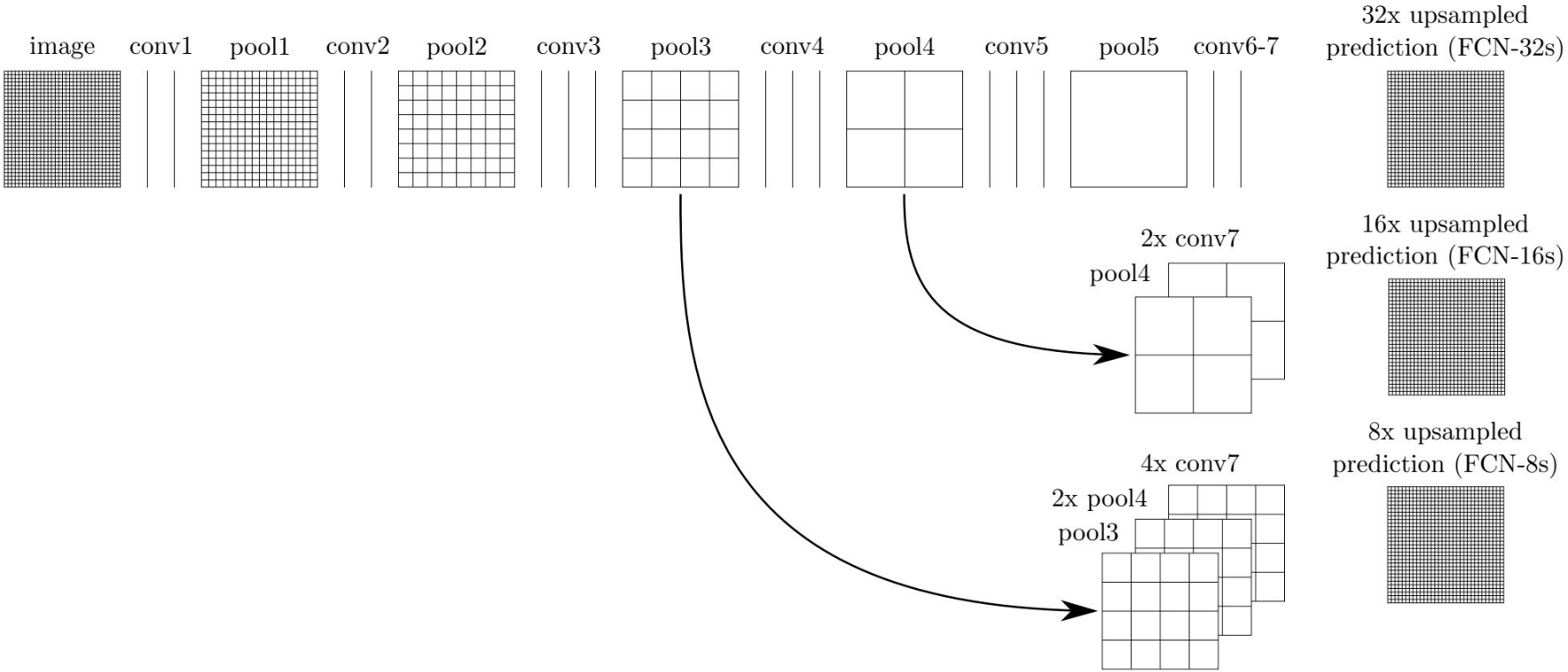


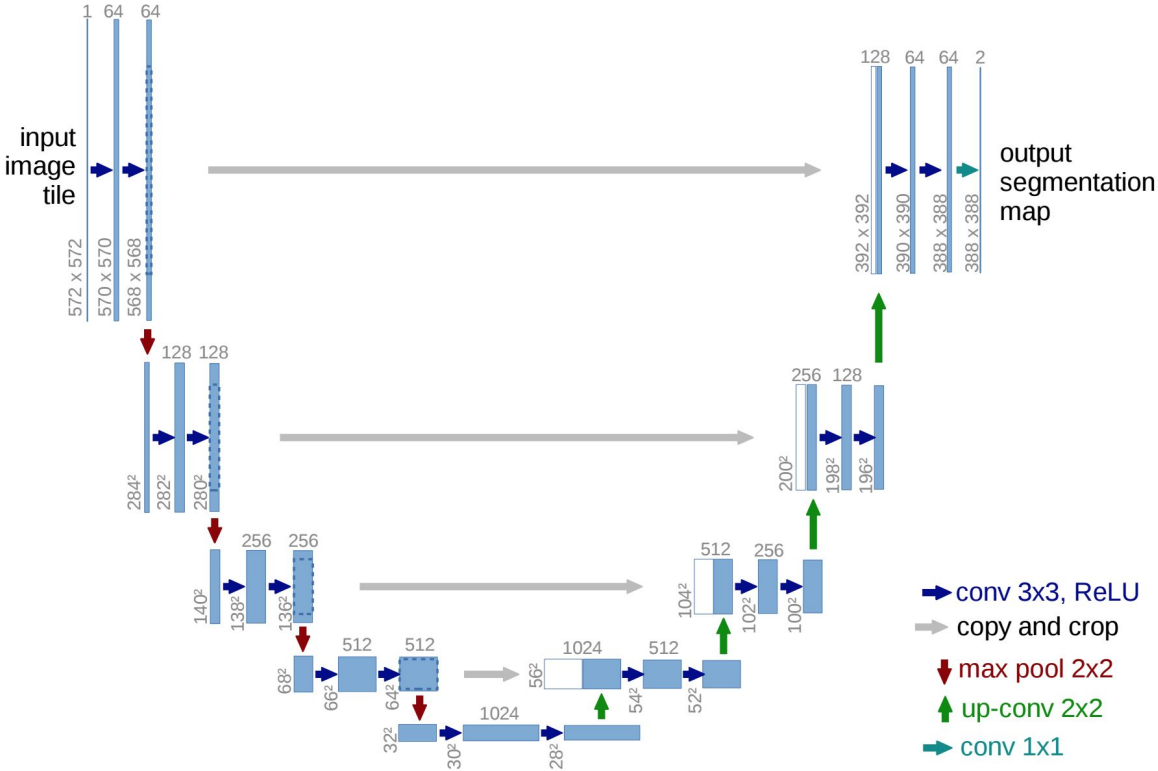

Figure: Kilian Weinberger

# Skip connections



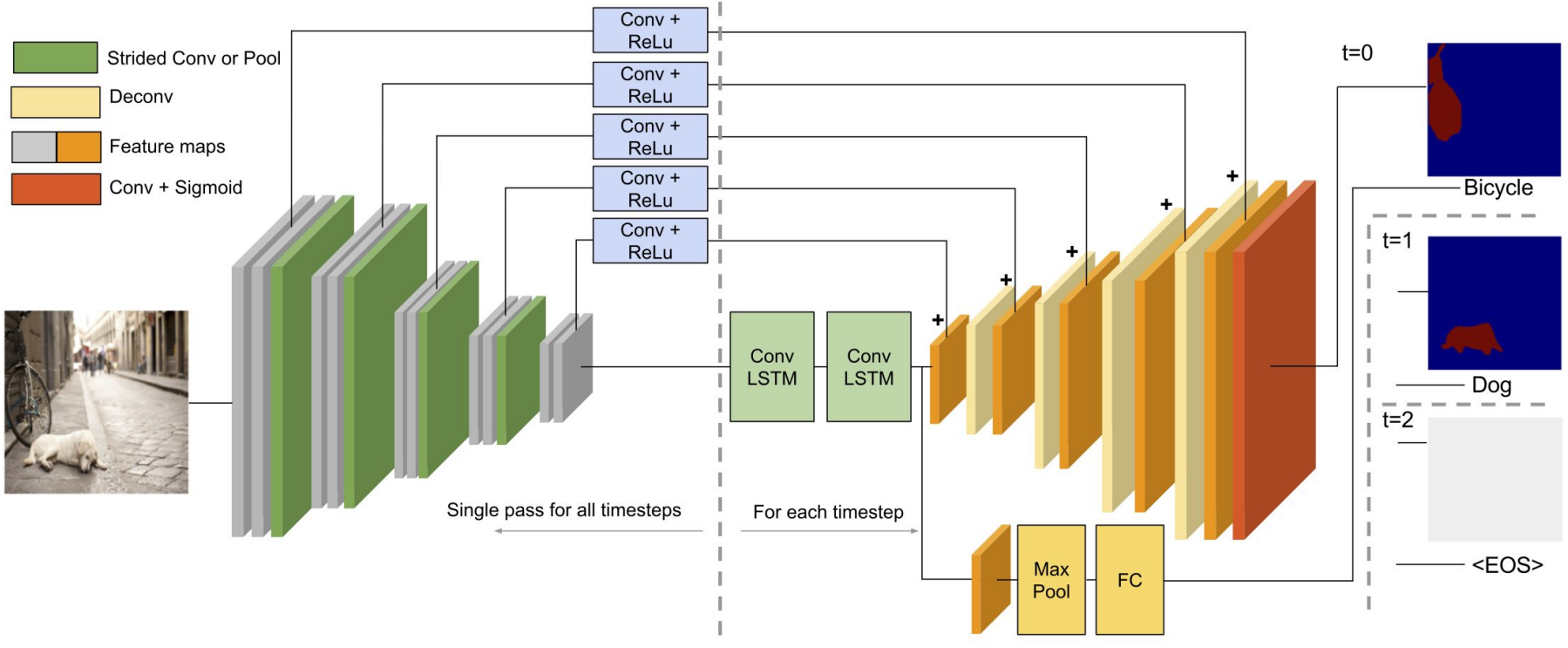

Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." CVPR 2015 & PAMI 2016.

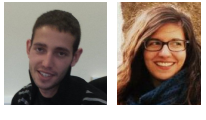
# Skip connections




Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234-241. Springer International Publishing, 2015

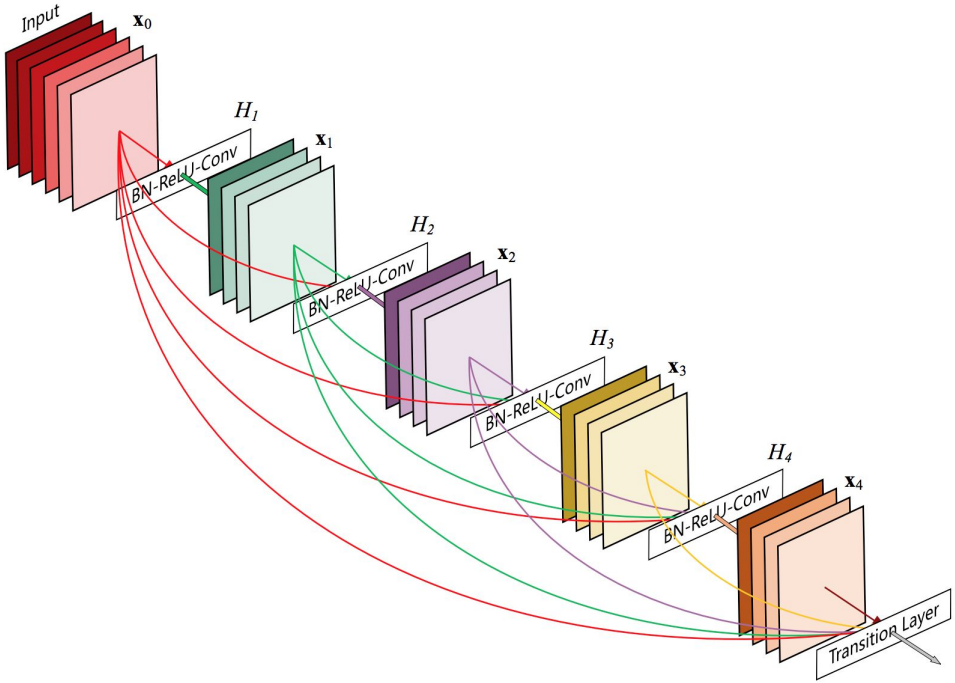
# Skip connections



# Dense connections

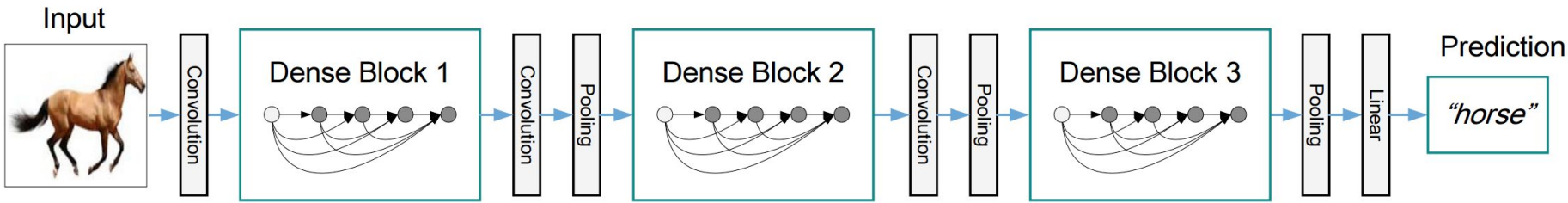
Connect every layer to every other layer of the same filter size.




**Dense Block** of 5-layers
with a growth rate of k=4

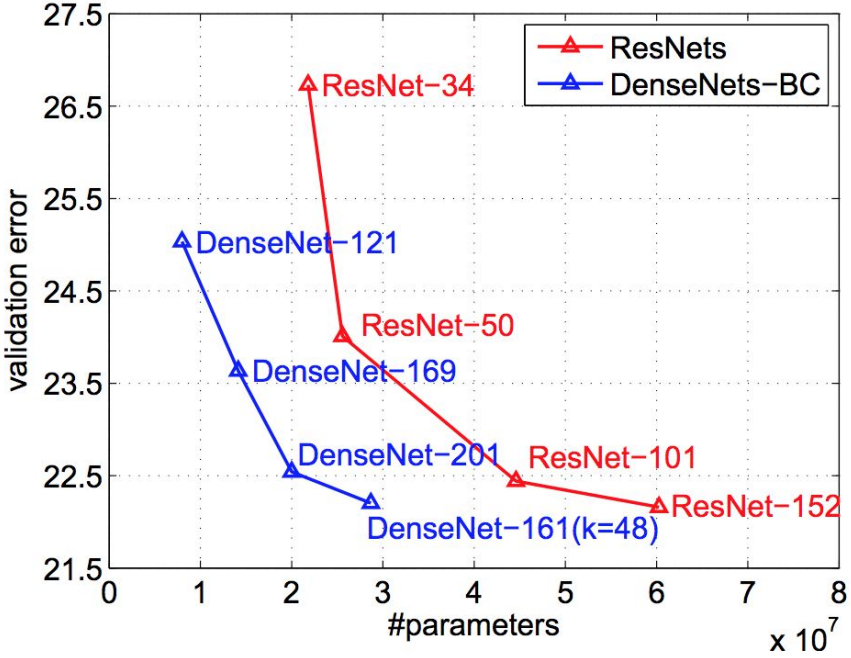
Huang, Gao, Zhuang Liu, Kilian Q. Weinberger, and Laurens van der Maaten. "Densely connected convolutional networks." *arXiv preprint arXiv:1608.06993* (2016). [code]

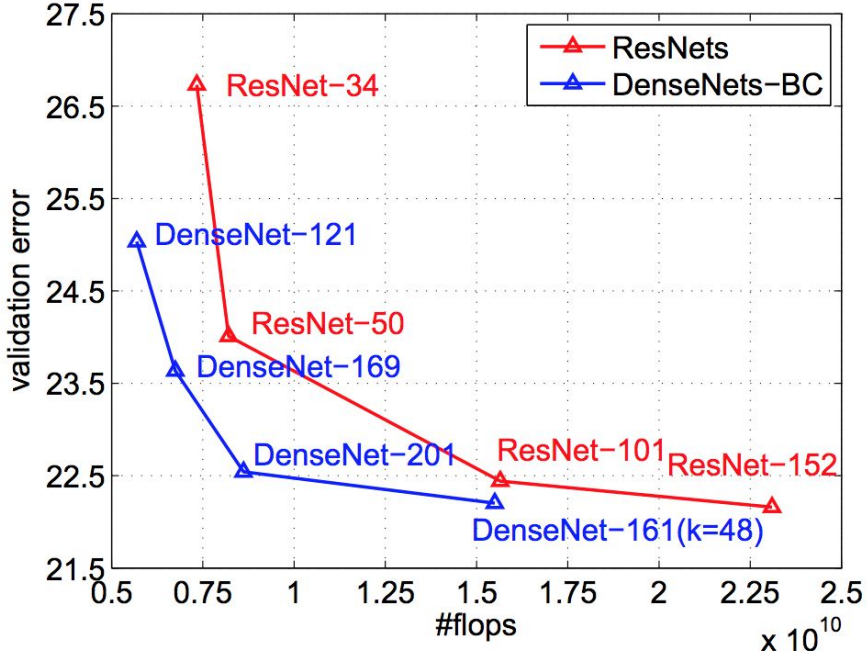
# Dense connections




Huang, Gao, Zhuang Liu, Kilian Q. Weinberger, and Laurens van der Maaten. "Densely connected convolutional networks." *arXiv preprint arXiv:1608.06993* (2016). [code]

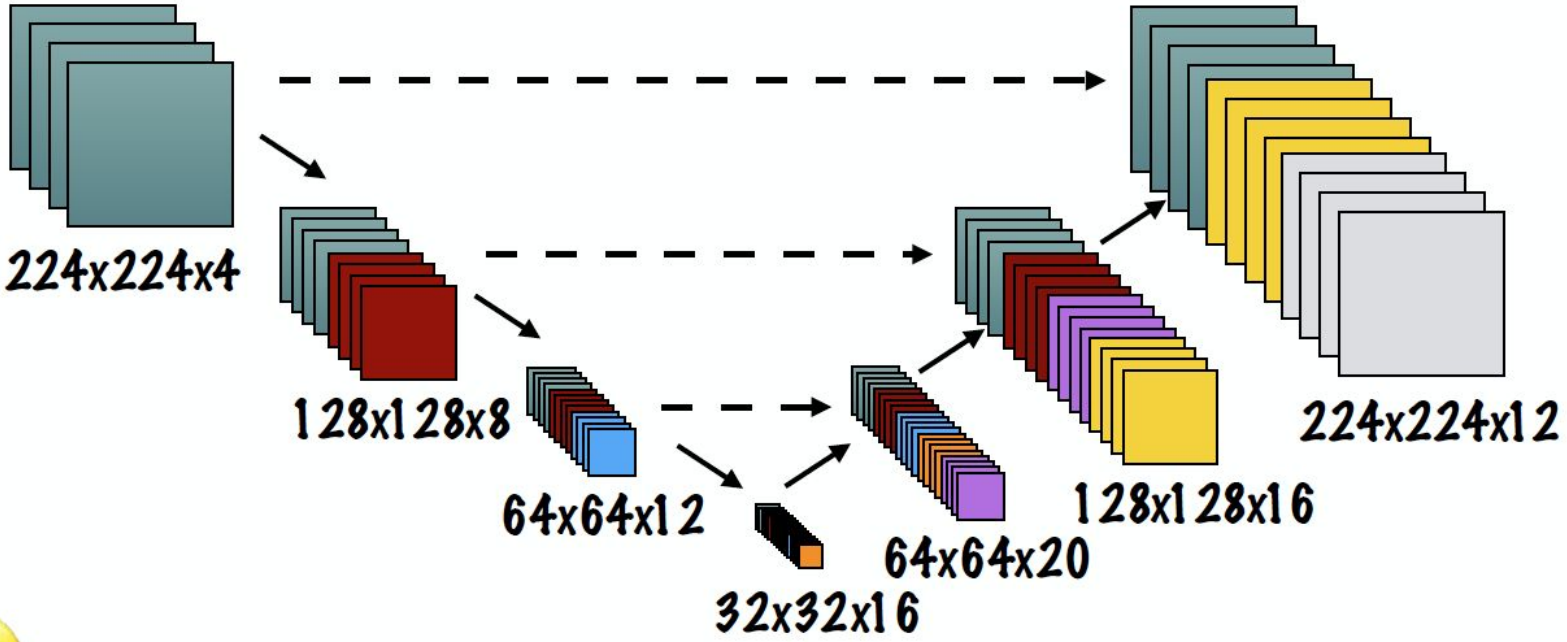
# Dense connections




Huang, Gao, Zhuang Liu, Kilian Q. Weinberger, and Laurens van der Maaten. "Densely connected convolutional networks." *arXiv preprint arXiv:1608.06993* (2016). [code]

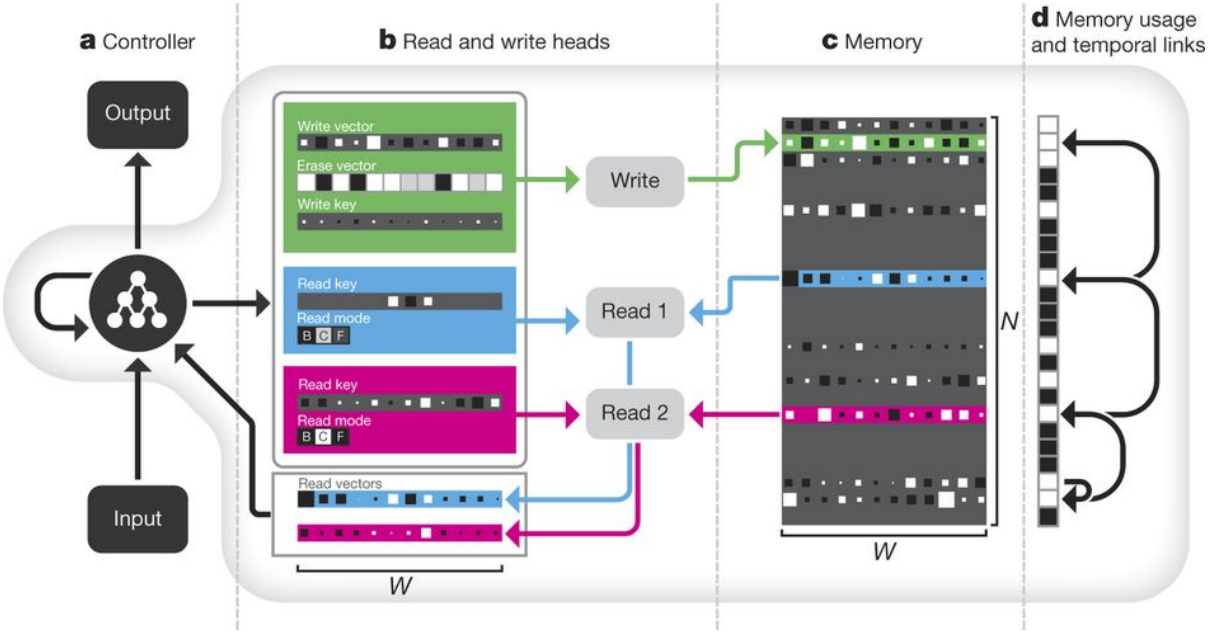
# Dense connections




Jégou, Simon, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio. "The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation." arXiv preprint arXiv:1611.09326 (2016). [code] [slides]

# Differentiable Neural Computers (DNC)

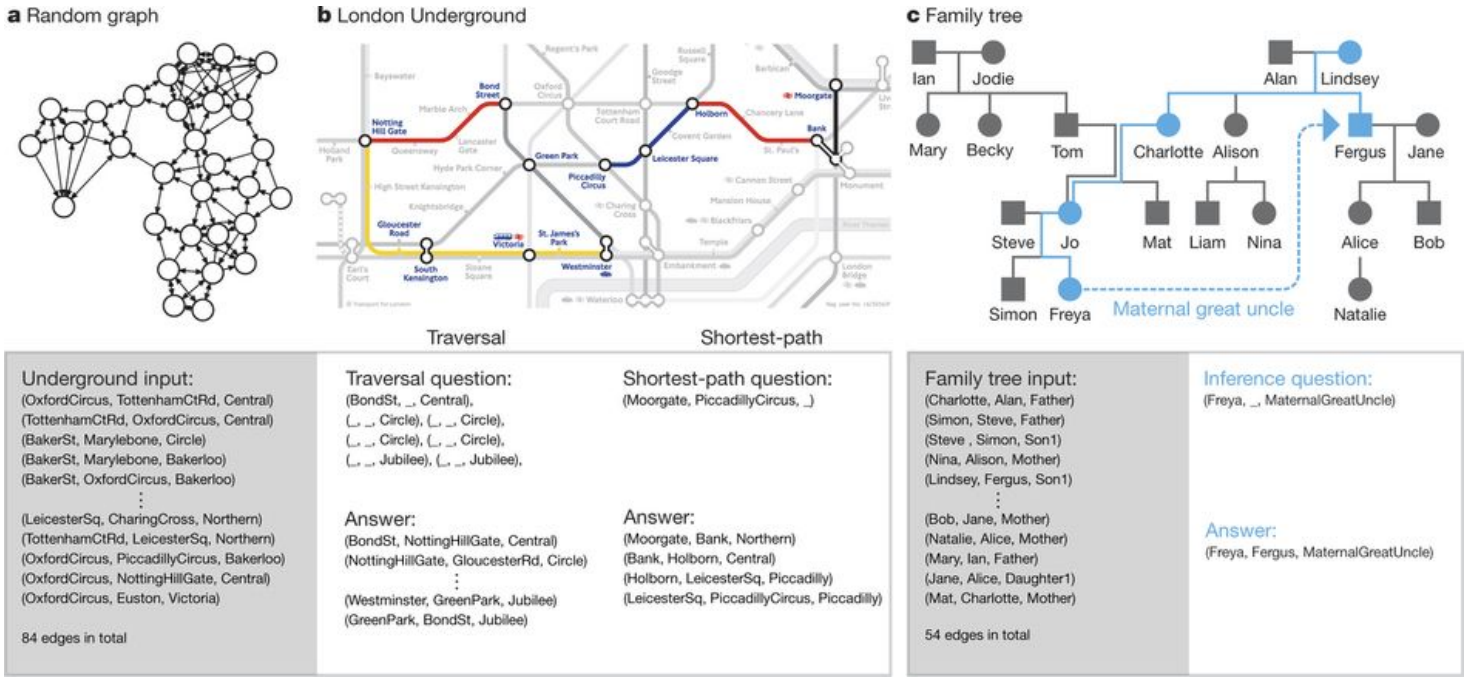
Add a trainable external memory to a neural network.




Graves, Alex, Greg Wayne, Malcolm Reynolds, Tim Harley, Ivo Danihelka, Agnieszka Grabska-Barwińska, Sergio Gómez Colmenarejo et al. "Hybrid computing using a neural network with dynamic external memory." Nature 538, no. 7626 (2016): 471-476. [Post by DeepMind]

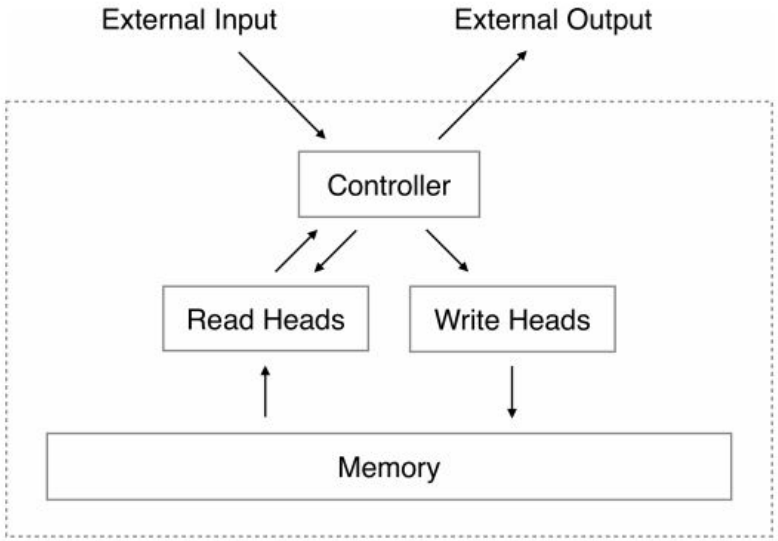
# Differentiable Neural Computers (DNC)

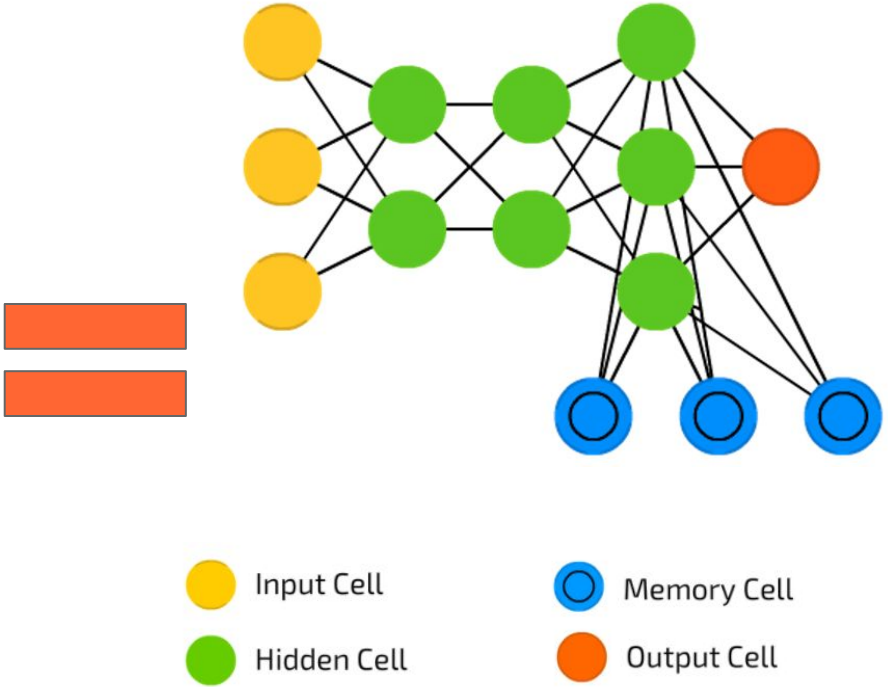
DNC can solve tasks reading information from a trained memory.




Graves, Alex, Greg Wayne, Malcolm Reynolds, Tim Harley, Ivo Danihelka, Agnieszka Grabska-Barwińska, Sergio Gómez Colmenarejo et al. "Hybrid computing using a neural network with dynamic external memory." Nature 538, no. 7626 (2016): 471-476. [Post by DeepMind]

# Differentiable Neural Computers (DNC)



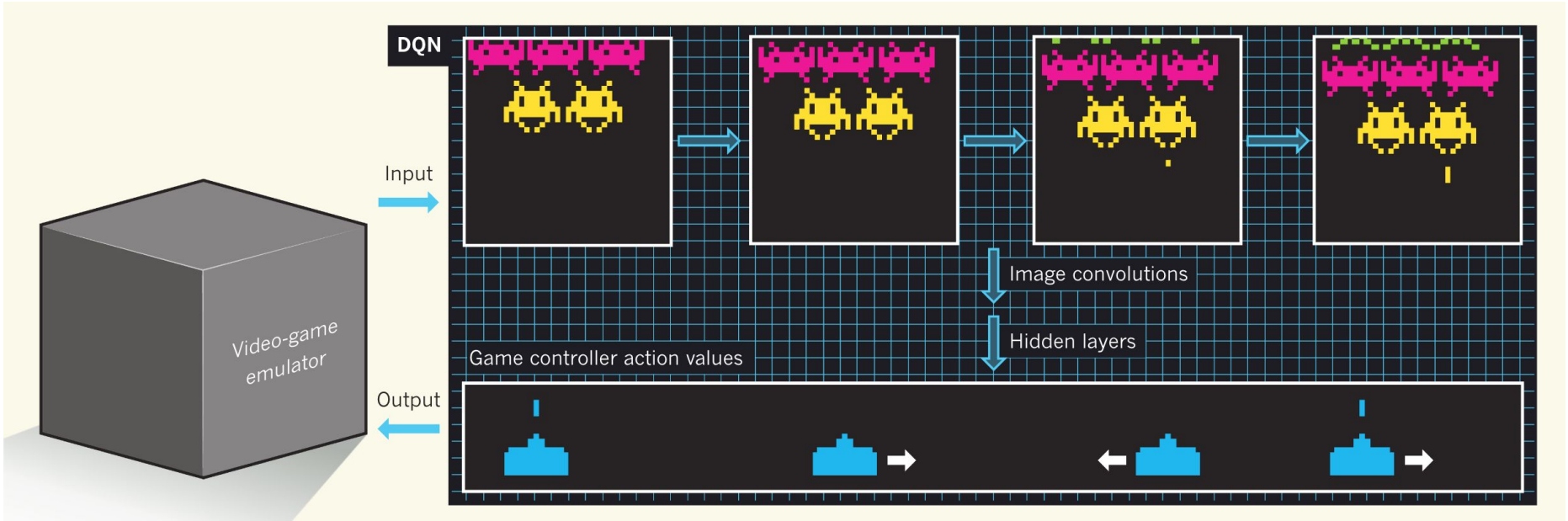

Graves, Alex, Greg Wayne, and Ivo Danihelka. "Neural turing machines." *arXiv preprint arXiv:1410.5401* (2014). [slides] [code]

F. Van Veen, "The Neural Network Zoo" (2016)

# Reinforcement Learning (RL)


Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602 (2013).

# Reinforcement Learning (RL)




Bernhard Schölkkopf, "Learning to see and act" Nature 2015.

# Reinforcement Learning (RL)

**Artificial intelligence (AI)**

## Google buys UK artificial intelligence startup Deepmind for £400m

Google makes its biggest EU purchase yet with the technology that aims to make computers think like humans

Samuel Gibbs

Monday 27 January 2014 13.23 GMT

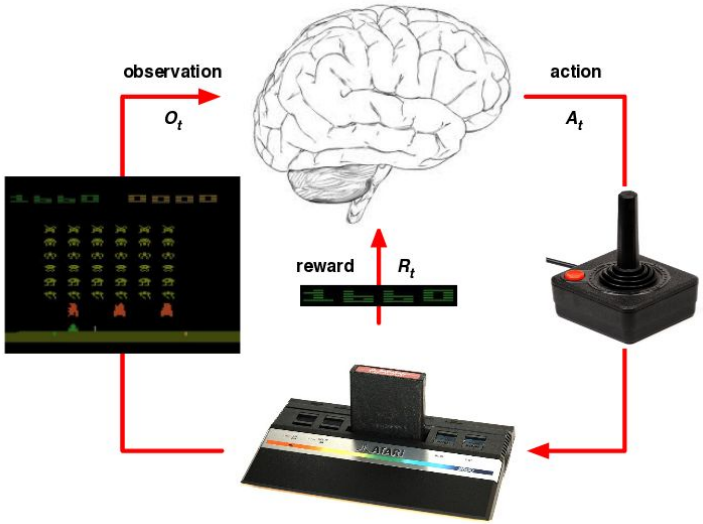
This article is 2 years old

1046 186

https://www.theguardian.com/technology/2014/jan/27/google-acquires-uk-artificial-intelligence-startup-deepmind

# Reinforcement Learning (RL)

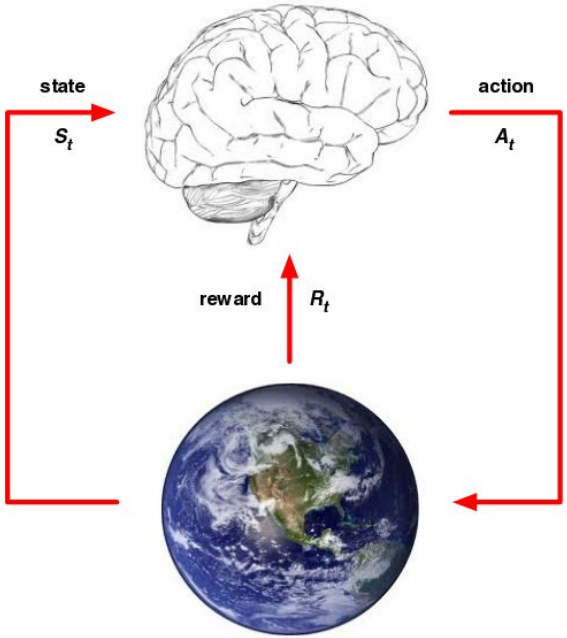
An agent that is a **decision-maker interacts** with the environment and learns through **trial-and-error**




Slide credit: UCL Course on RL by David Silver

# Reinforcement Learning (RL)

An agent that is a **decision-maker interacts** with the environment and learns through **trial-and-error**



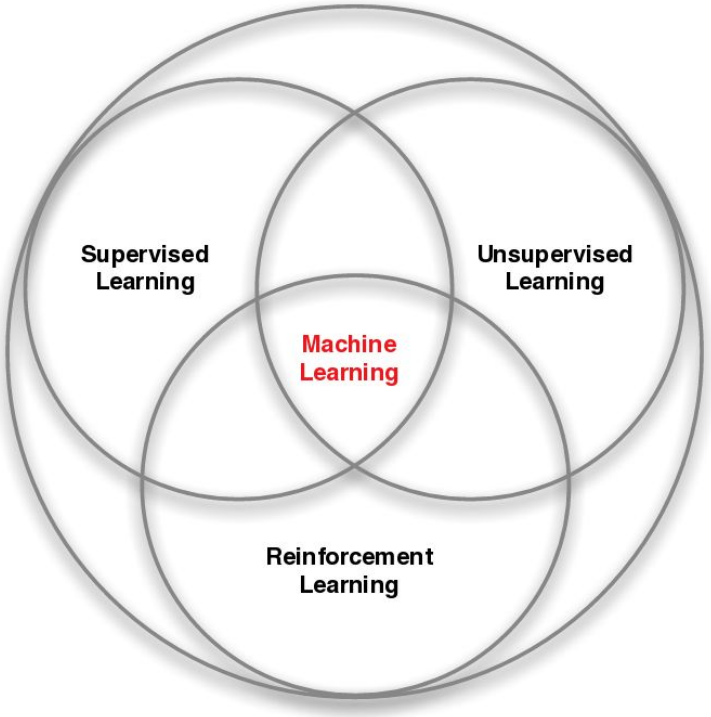

We model the decision-making process through a **Markov Decision Process**

Slide credit: UCL Course on RL by David Silver

# Reinforcement Learning (RL)

**Reinforcement Learning**

- There is no supervisor, only reward signal
- Feedback is delayed, not instantaneous
- Time really matters (sequential, non i.i.d data)




Slide credit: UCL Course on RL by David Silver

# Reinforcement Learning (RL)

**What is Reinforcement Learning ?**

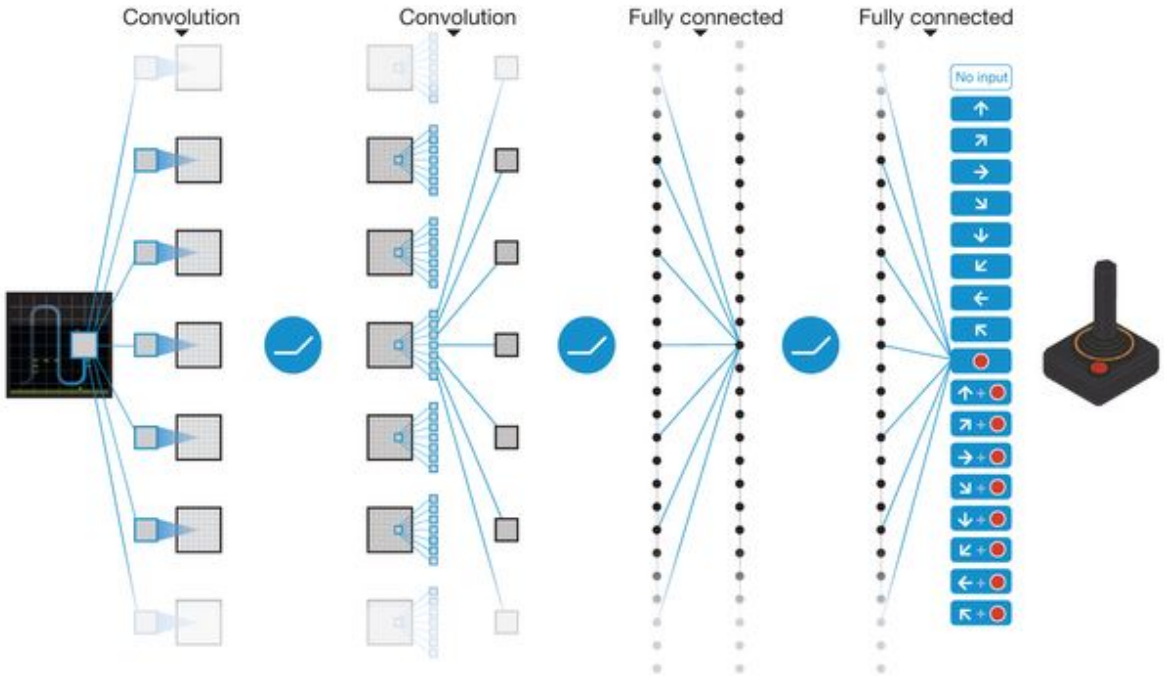
“a way of programming agents by reward and punishment without needing to specify how the task is to be achieved”

[Kaelbling, Littman, & Moore, 96]

Slide credit: Míriam Bellver

# Reinforcement Learning (RL)

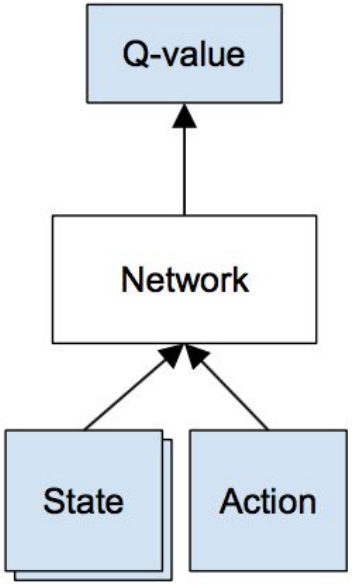
## Deep Q-Network (DQN)



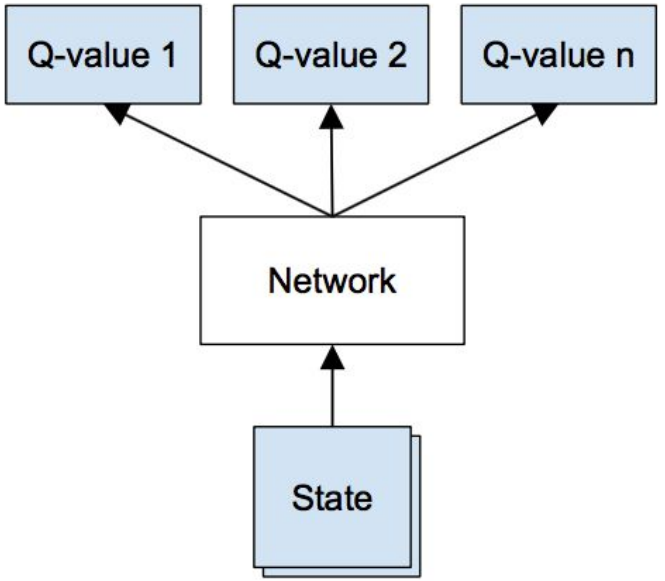

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves et al. "Human-level control through deep reinforcement learning." *Nature* 518, no. 7540 (2015): 529-533.

# Reinforcement Learning (RL)

## Deep Q-Network (DQN)




Naive DQN

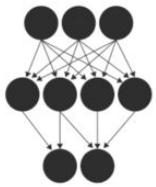



Refined DQN

Source: Tambet Matiisen, Demystifying Deep Reinforcement Learning (Nervana)

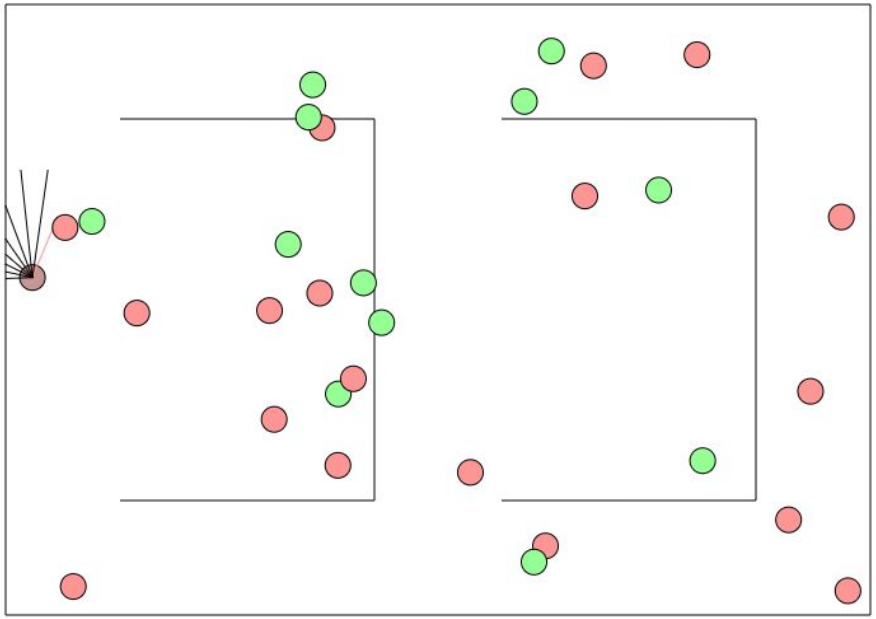
# Reinforcement Learning (RL)

## Deep Q-Network (DQN)



ConvNetJS

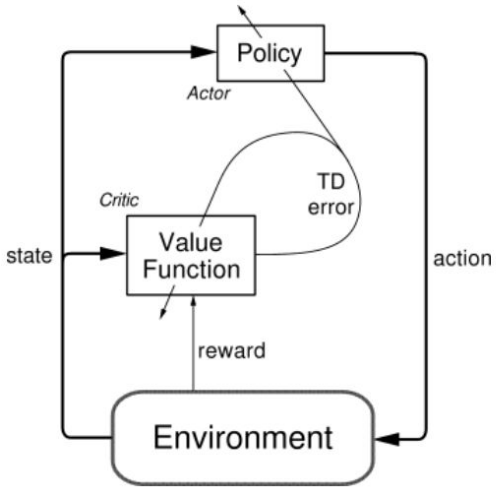
Deep Learning in your browser

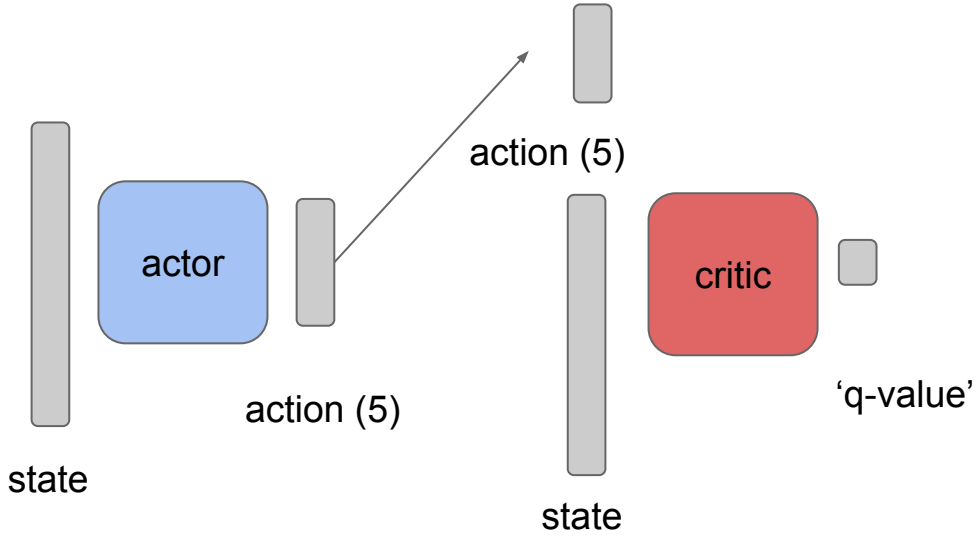


Andrej Karpathy, [“ConvNetJS Deep Q Learning Demo”]

# Reinforcement Learning (RL)

## Actor-Critic algorithm




actor performs an action

critic assesses how good the action was, and the gradients are used to train the actor and the critic

Slide credit: Míriam Bellver

Grondman, Ivo, Lucian Busoniu, Gabriel AD Lopes, and Robert Babuska. "A survey of actor-critic reinforcement learning: Standard and natural policy gradients." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42, no. 6 (2012): 1291-1307.

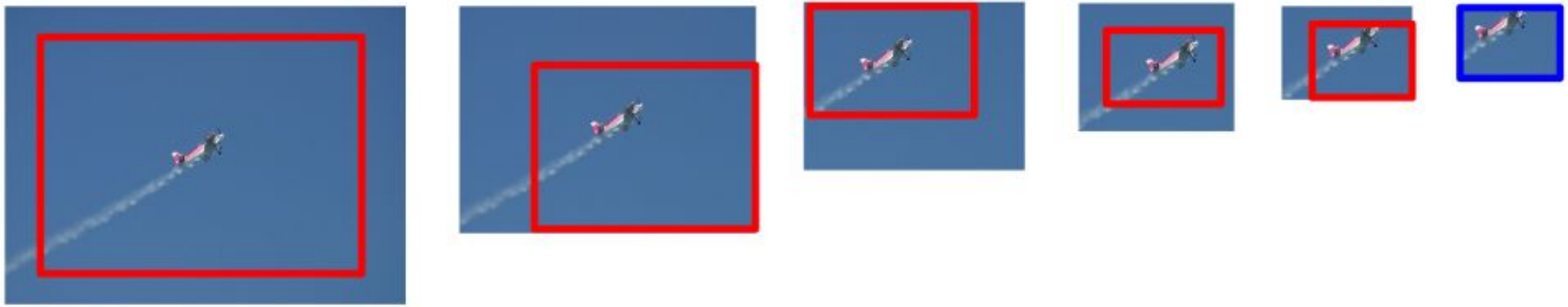
# Reinforcement Learning (RL)


AlphaGo

Google DeepMind

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. and Dieleman, S., 2016. Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), pp.484-489

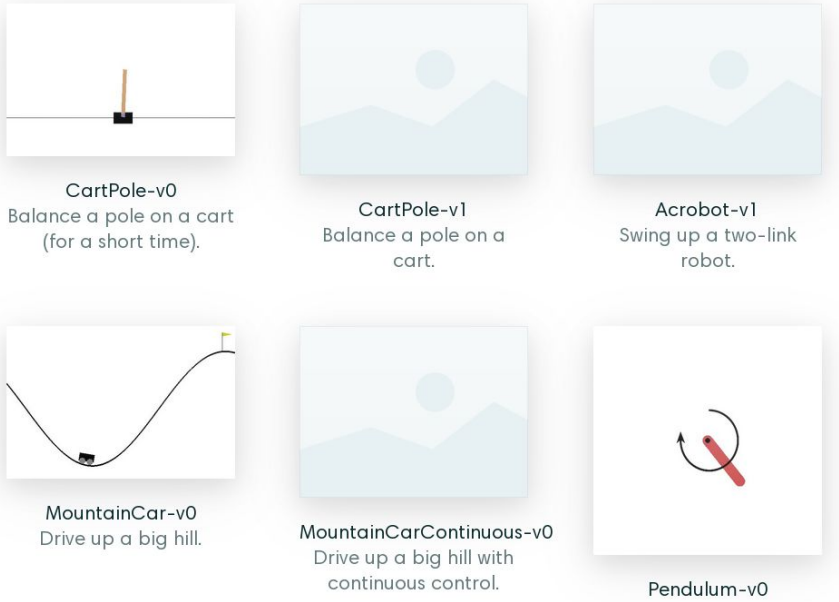
# Reinforcement Learning (RL)



Miriam Bellver, Xavier Giro-i-Nieto, Ferran Marques, and Jordi Torres. "Hierarchical Object Detection with Deep Reinforcement Learning." In Deep Reinforcement Learning Workshop (NIPS). 2016.

# Reinforcement Learning (RL)

OpenAI Gym + keras-rl



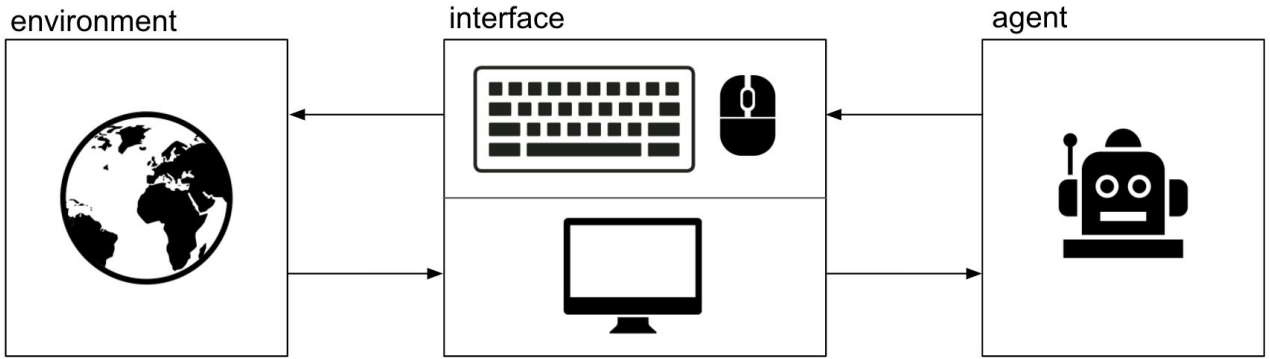

+

**keras-rl**

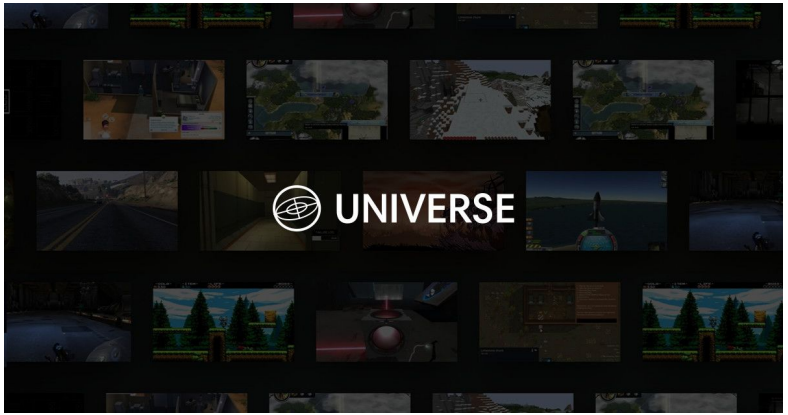
keras-rl implements some state-of-the art deep reinforcement learning algorithms in Python and seamlessly integrates with the deep learning library Keras. Just like Keras, it works with either Theano or TensorFlow, which means that you can train your algorithm efficiently either on CPU or GPU. Furthermore, keras-rl works with OpenAI Gym out of the box.

Slide credit: Míriam Bellver

# Reinforcement Learning (RL)




OpenAI

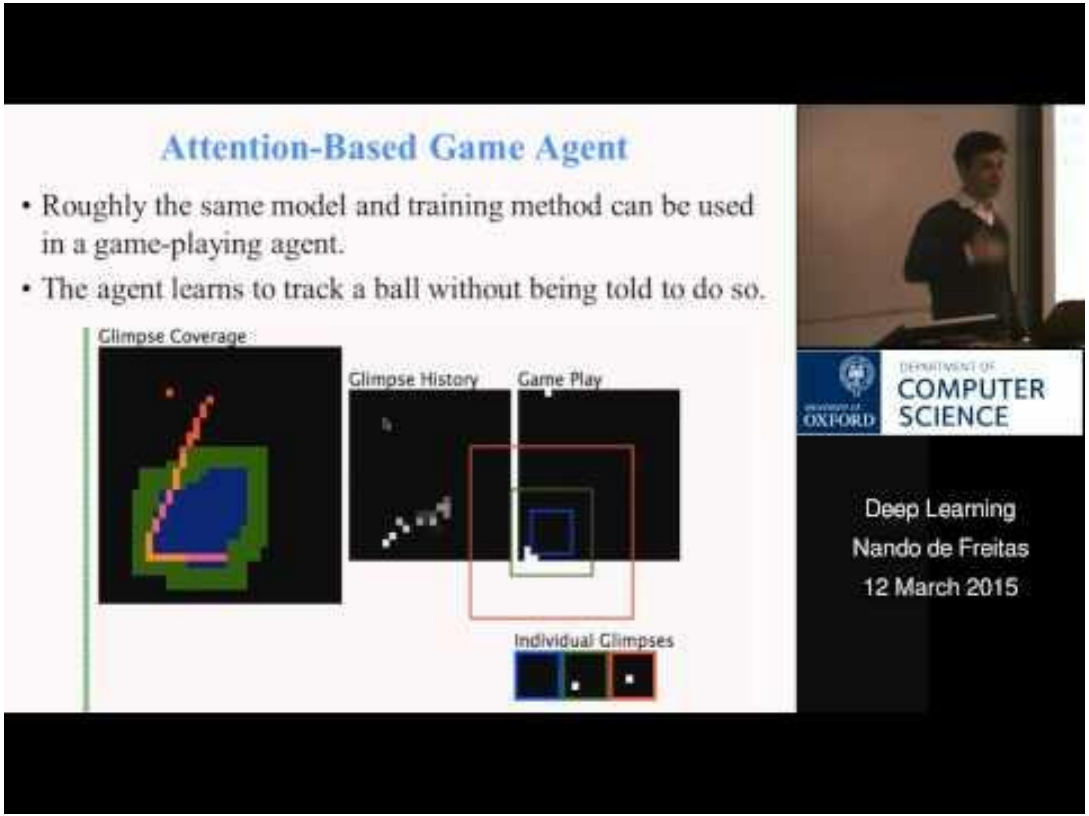



OpenAI Universe environment

# Reinforcement Learning (RL)


Deep Learning TV, [“Reinforcement learning - Ep. 30”]

# Reinforcement Learning (RL)




Nando de Freitas, [“Deep Reinforcement Learning - Policy search”] (University of Oxford)

# Reinforcement Learning (RL)


Davdi Silver, [“Reinforcement Learning Course”](#) (Google DeepMind)”

# Adversarial Networks

Generative Adversarial Networks (GANs)

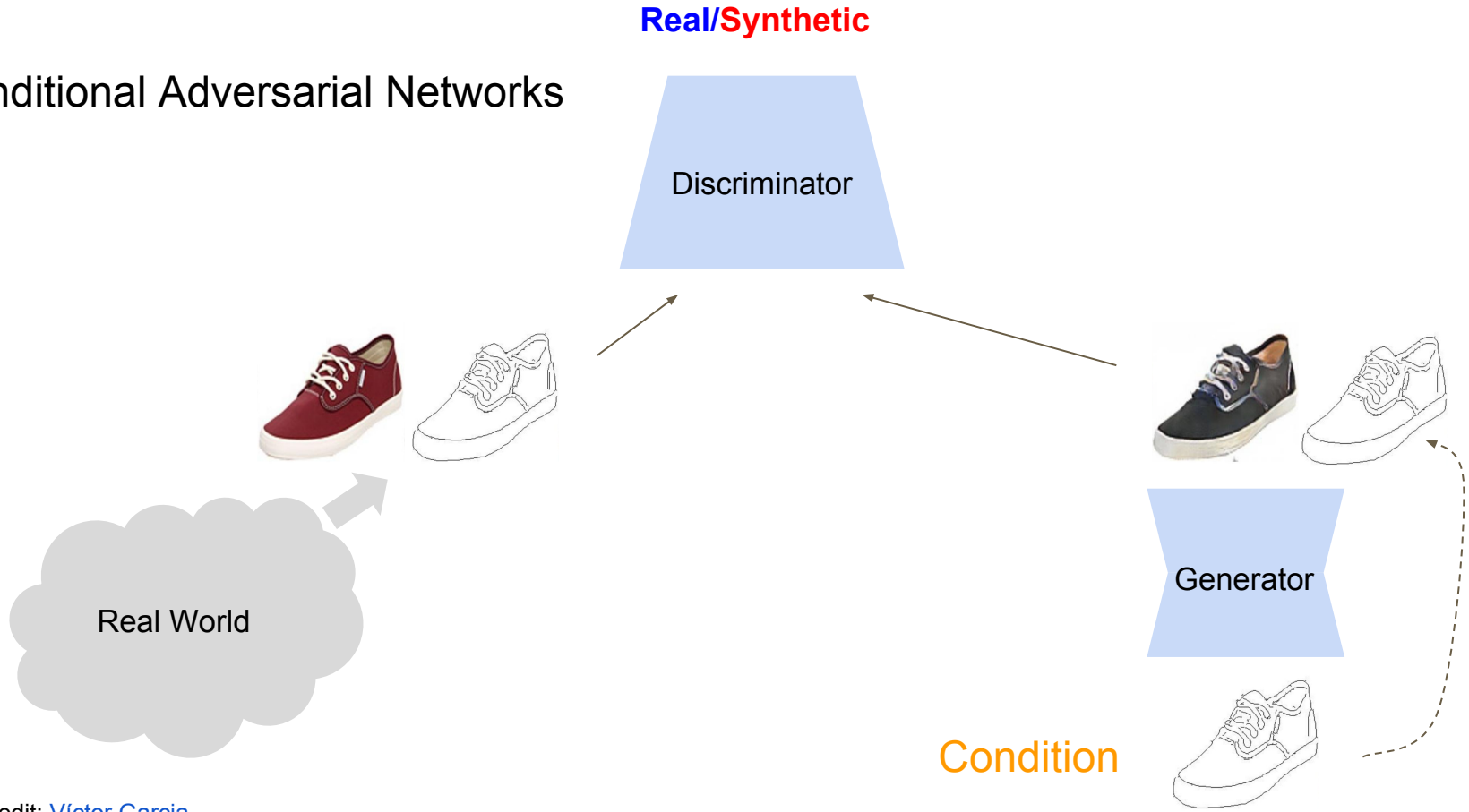
More details in D2L5.




Slide credit: Víctor Garcia

# Adversarial Networks

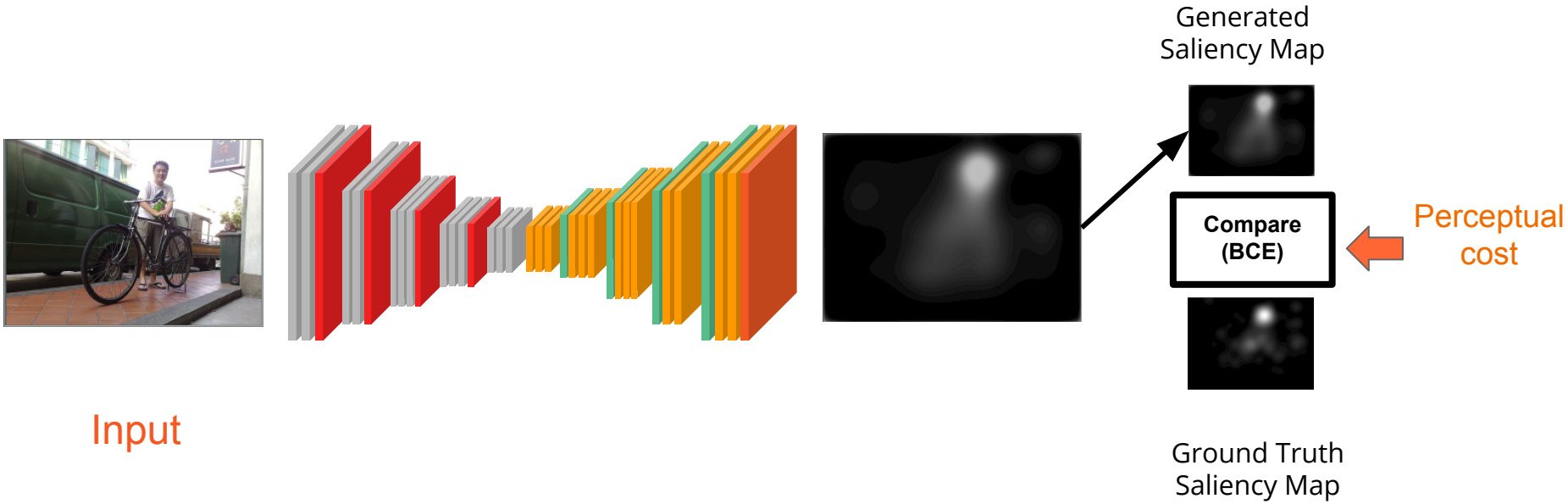
Conditional Adversarial Networks

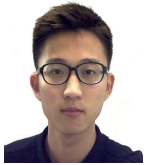



Slide credit: Víctor Garcia

# Adversarial Networks

A computer vision problem such as visual saliency prediction...



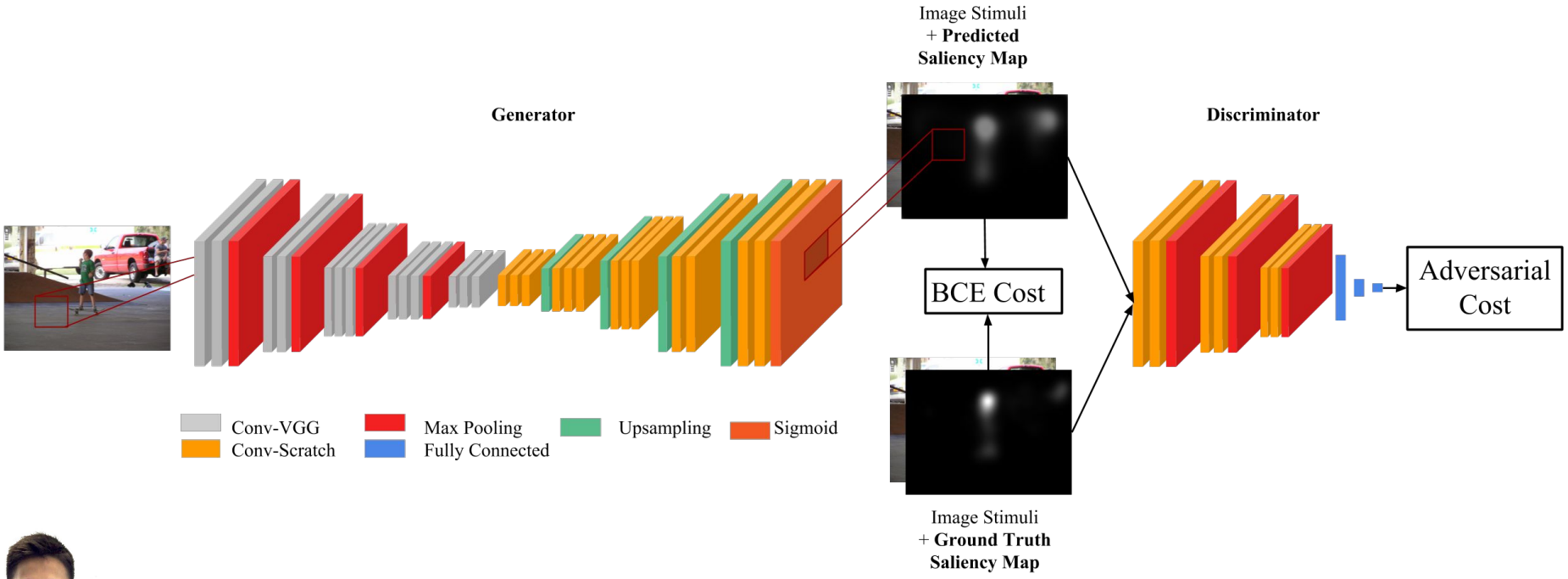
Junting Pan, Cristian Canton, Kevin McGuinness, Noel E. O'Connor, Jordi Torres, Elisa Sayrol and Xavier Giro-i-Nieto. "SalGAN: Visual Saliency Prediction with Generative Adversarial Networks." arXiv. 2017.

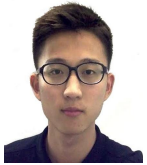
# Adversarial Networks

...can benefit from adding an adversarial loss:

Perceptual cost  Adversarial cost

$$\alpha \cdot \mathcal{L}_{BCE} - \log D(I, \hat{S}),$$



Junting Pan, Cristian Canton, Kevin McGuinness, Noel E. O'Connor, Jordi Torres, Elisa Sayrol and Xavier Giro-i-Nieto. "SalGAN: Visual Saliency Prediction with Generative Adversarial Networks." arXiv. 2017.

# Adversarial Networks

$$= \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{HR} - G_{\theta_G}(I^{LR})_{x,y})^2$$

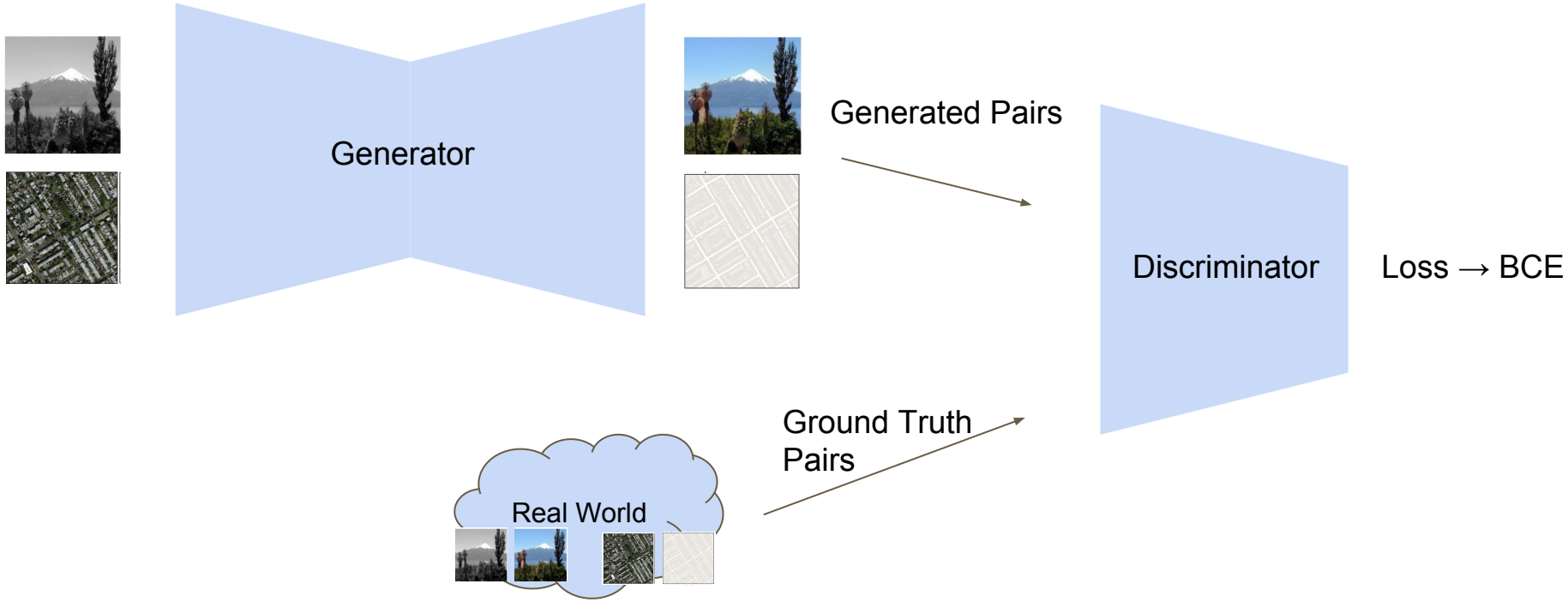
Generator

Discriminator

**Loss2 GAN**
**{Binary Crossentropy}**
**1/0**

Víctor Garcia and Xavier Giró-i-Nieto (work under progress)

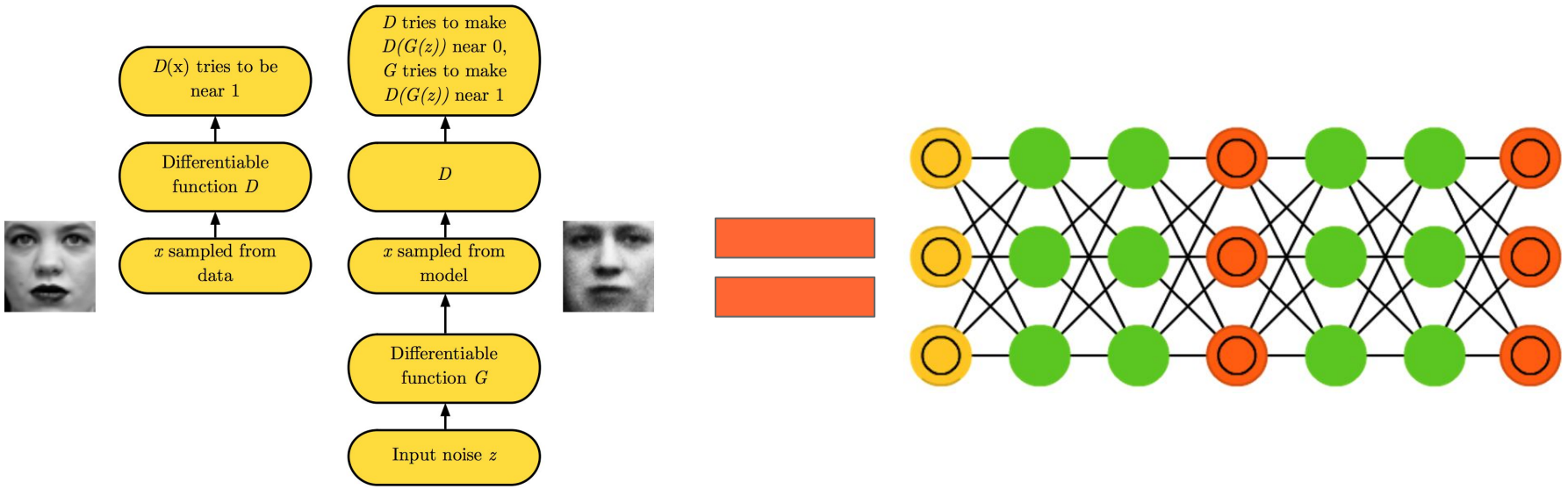
# Adversarial Networks




Slide credit: Víctor Garcia

Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-image translation with conditional adversarial networks." arXiv preprint arXiv:1611.07004 (2016).

# Adversarial Networks

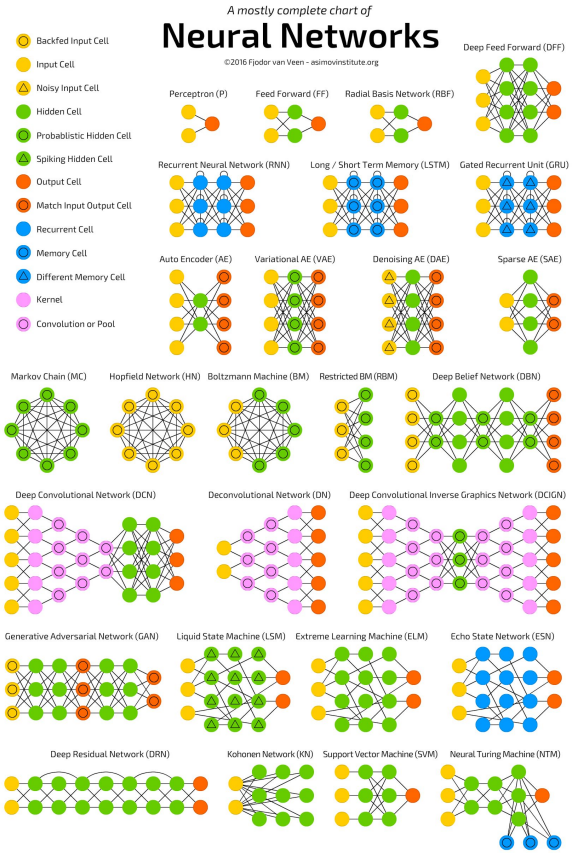



Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." NIPS 2014

Goodfellow, Ian. "NIPS 2016 Tutorial: Generative Adversarial Networks." *arXiv preprint arXiv:1701.00160* (2016).

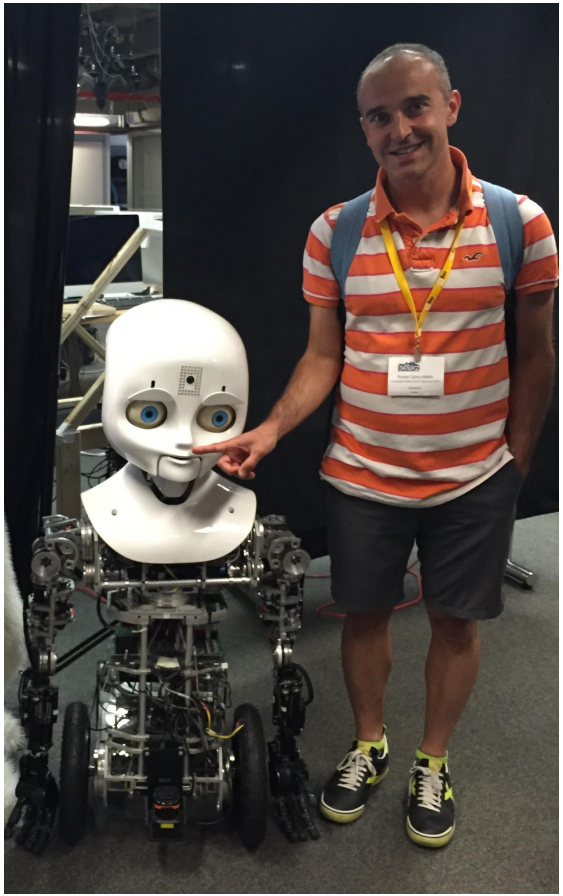
F. Van Veen, "The Neural Network Zoo" (2016)

# The Full Story




F. Van Veen, [“The Neural Network Zoo”](#) (2016)

# Thanks ! Q&A ?



Follow me at

/ProfessorXavi

@DocXavi

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Department of Signal Theory and Communications
*Image Processing Group*

https://imatge.upc.edu/web/people/xavier-giro