



¿Qué es el Aprendizaje de Máquinas?

Iván Jourdan



VOL. LIX. No. 236.]

[October, 1950]

M I N D
A QUARTERLY REVIEW
OF
PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND
INTELLIGENCE

By A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, ‘Can machines think ?’ This should begin with definitions of the meaning of the terms ‘machine’ and ‘think’. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words ‘machine’, and ‘think’ are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, ‘Can machines think ?’ is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the ‘imitation game’. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either ‘X is A and Y is B’ or ‘X is B and Y is A’. The interrogator is allowed to put questions to A and B thus :

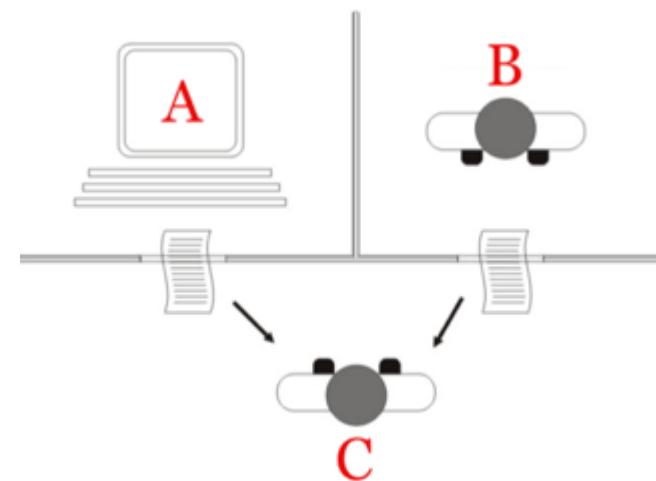
C : Will X please tell me the length of his or her hair ?

Now suppose X is actually A, then A must answer. It is A’s

¿Qué es Inteligencia Artificial?

Alan. M. Turing 1950

I Propose to consider the question
‘Can machine think?’

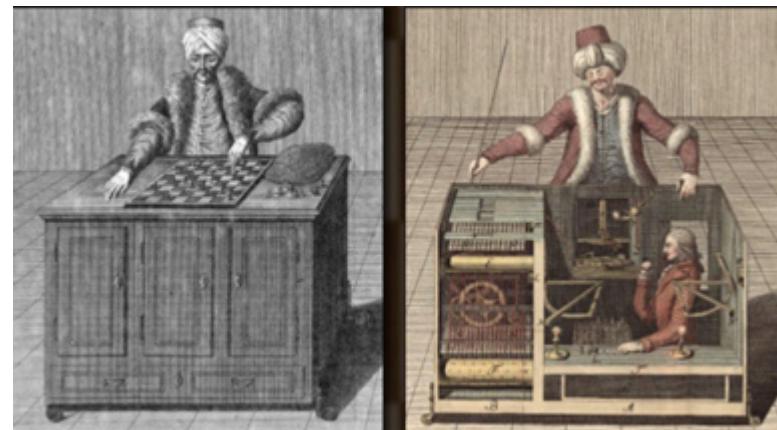


- ¿Qué es una máquina?
- ¿Qué es pensar?



¿Qué es Inteligencia Artificial?

- Modelos del comportamiento humano
máquinas que se comportan externamente como un humano



The Turk, 1770

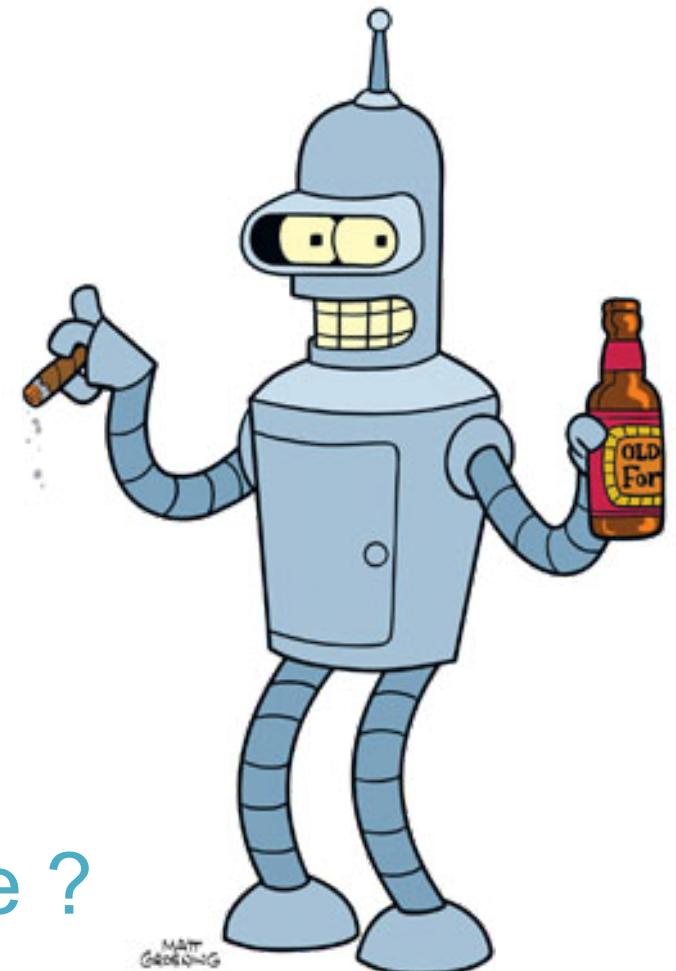


Star Wars Chess, 1993



¿Qué es Inteligencia Artificial?

- Modelos del comportamiento humano
máquinas que se comportan externamente como un humano
- Modelos que se comportan inteligentemente o poseen inteligencia
máquinas inteligentes.

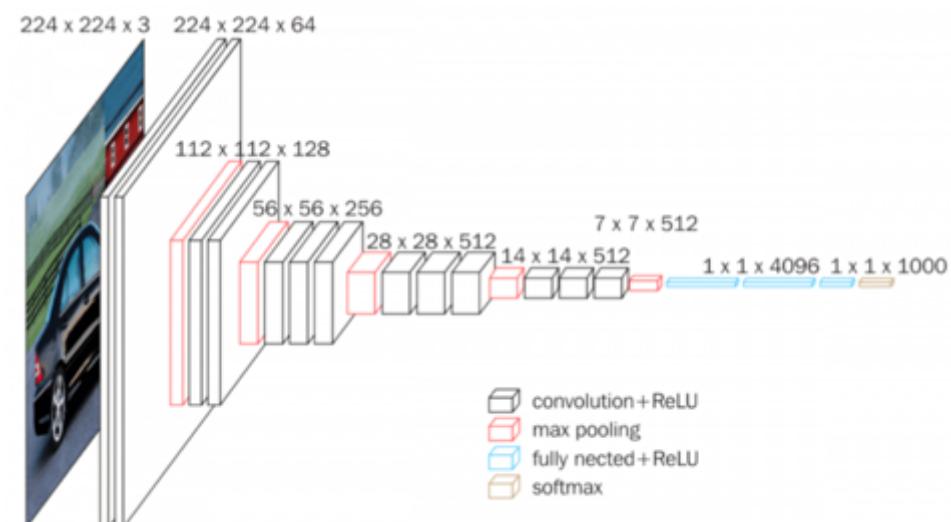
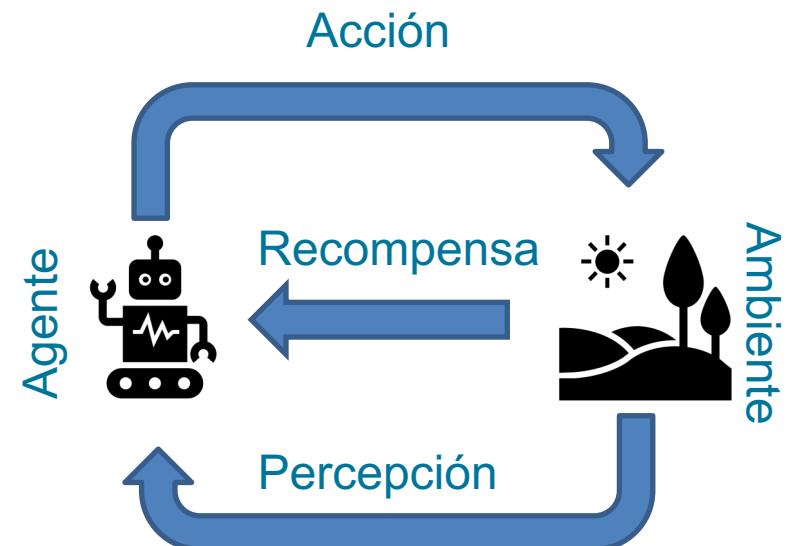


¿ Qué es inteligencia ?

¿ Puede un sistema identificarse ?

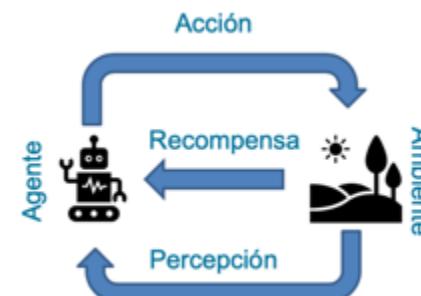
¿Qué es Inteligencia Artificial?

- Modelos del comportamiento humano
máquinas que se comportan externamente como un humano
- Modelos que se comportan inteligentemente o poseen inteligencia
máquinas inteligentes.
- Modelos de los procesos del pensamiento humano
máquinas que internamente procesan de la manera en que lo hacemos los humanos

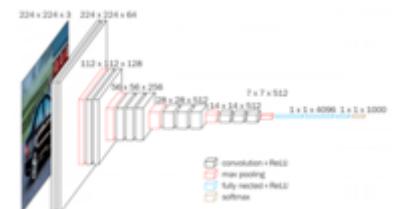


¿Qué es Inteligencia Artificial?

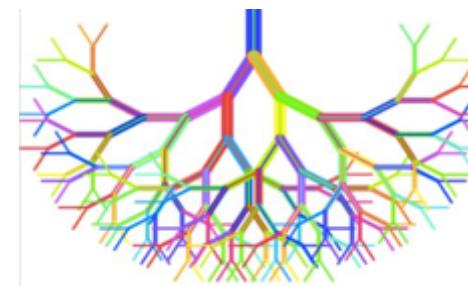
- Modelos del comportamiento humano
máquinas que se comportan externamente como un humano
- Modelos que se comportan inteligentemente o poseen inteligencia
máquinas inteligentes.
- Modelos de los procesos del pensamiento humano
máquinas que internamente procesan de la manera en que lo hacemos los humanos
- Modelos con comportamiento racional.
máquinas que devuelven resultados esperables, es decir, que hacen bien las cosas que tiene que hacer.



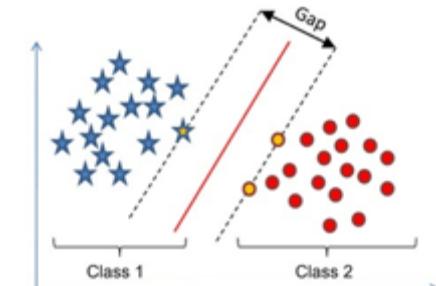
Deep Learning



Random Forest



SVM



K-means



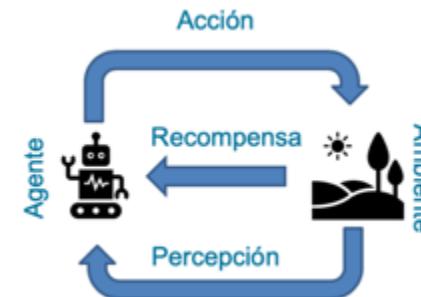
Self –Organizing Maps



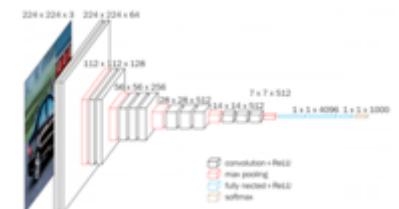


Aprendizaje de Máquinas (Machine Learning)

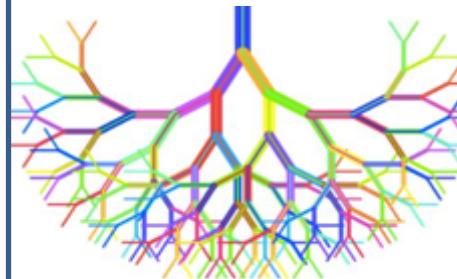
- Modelos computacionales de los procesos del pensamiento humano
máquinas que internamente procesan de la manera en que lo hacemos los humanos
- Modelos computacionales con comportamiento racional.
máquinas que devuelven resultados esperables, es decir, que hacen bien las cosas que tiene que hacer.



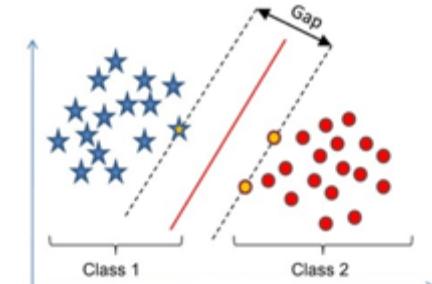
Deep Learning



Random Forest



SVM



K-means



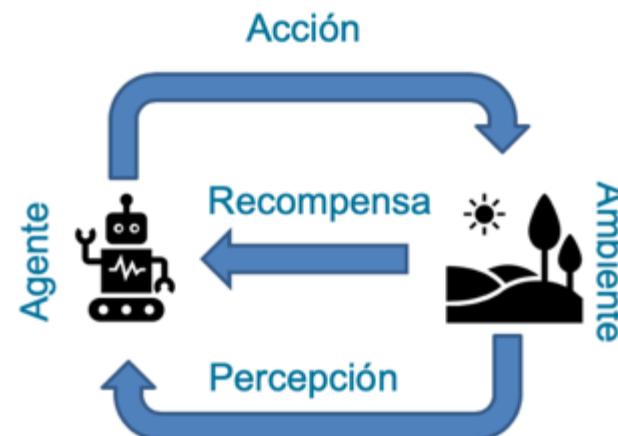
Self –Organizing Maps



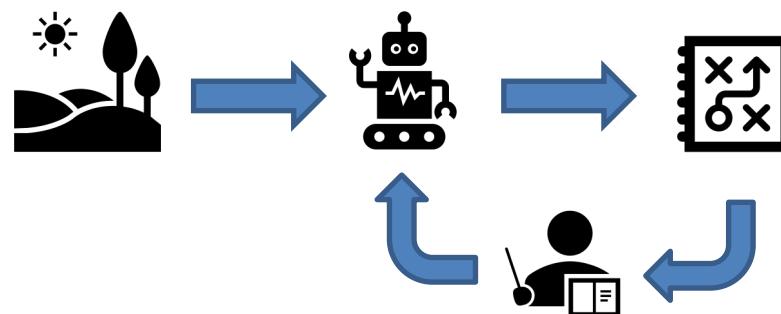


Aprendizaje

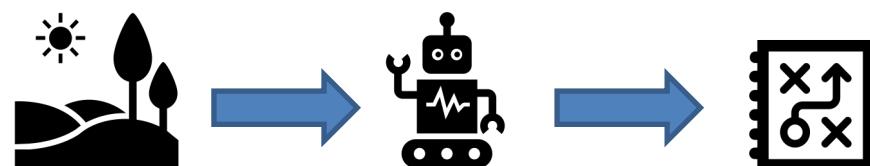
Aprendizaje por refuerzos



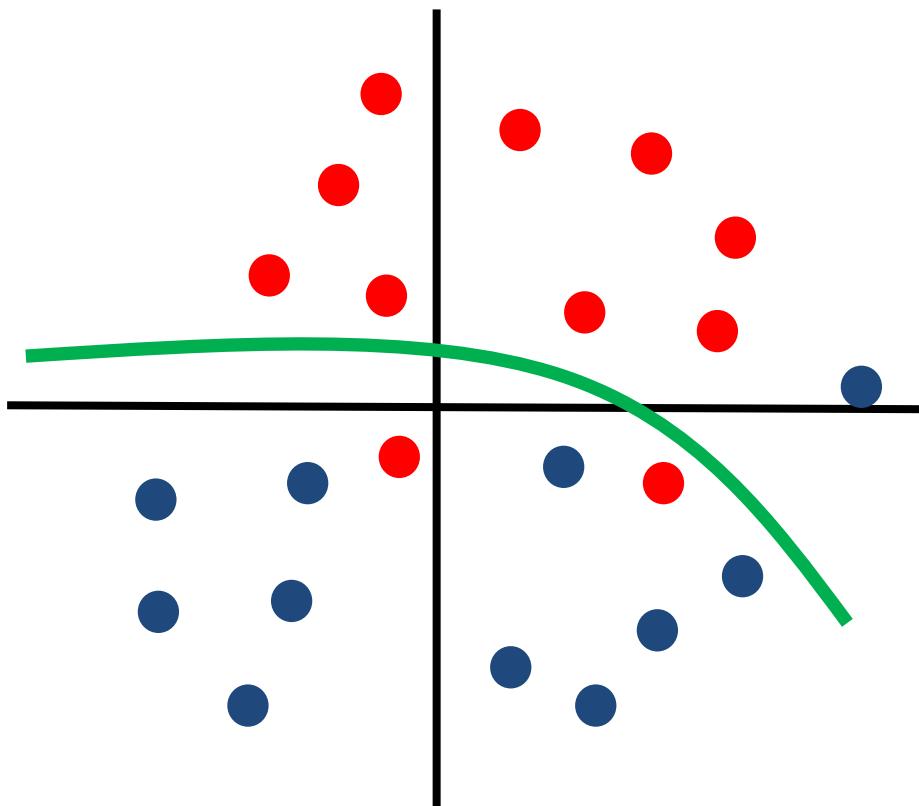
Aprendizaje Supervisado



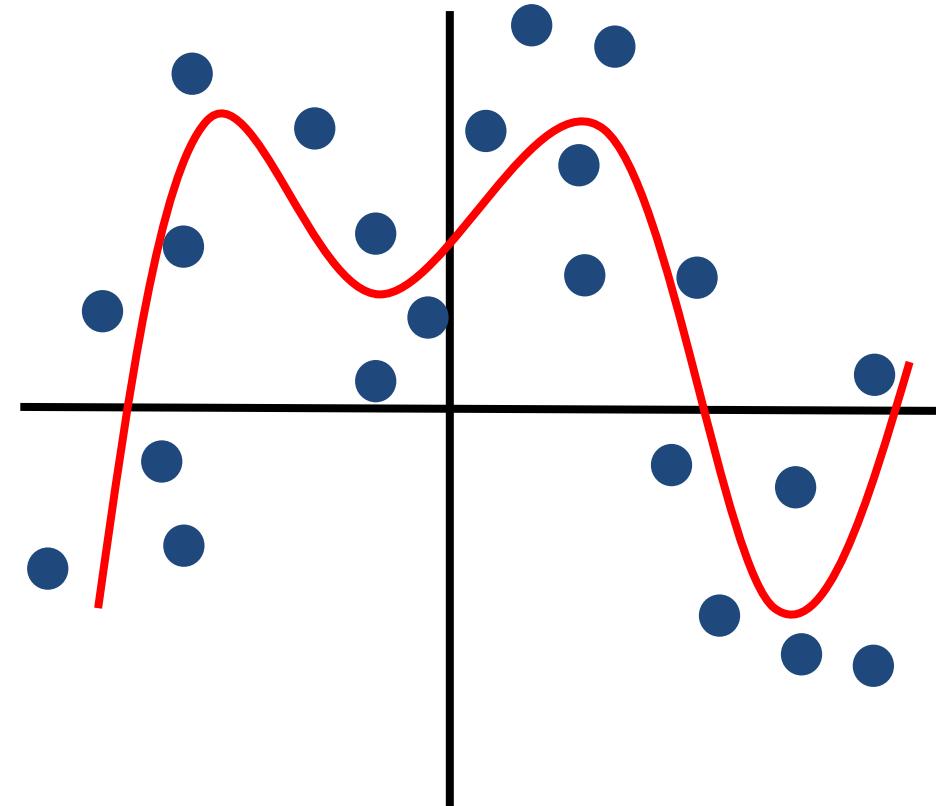
Aprendizaje No Supervisado



¿Qué problemas resuelve?



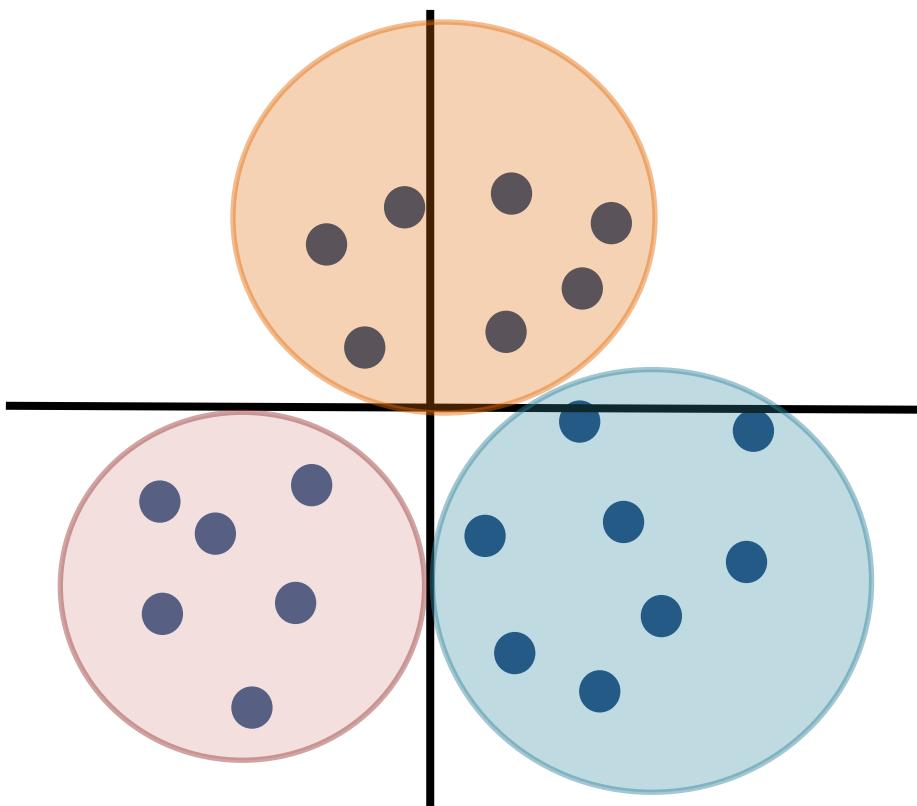
Clasificación



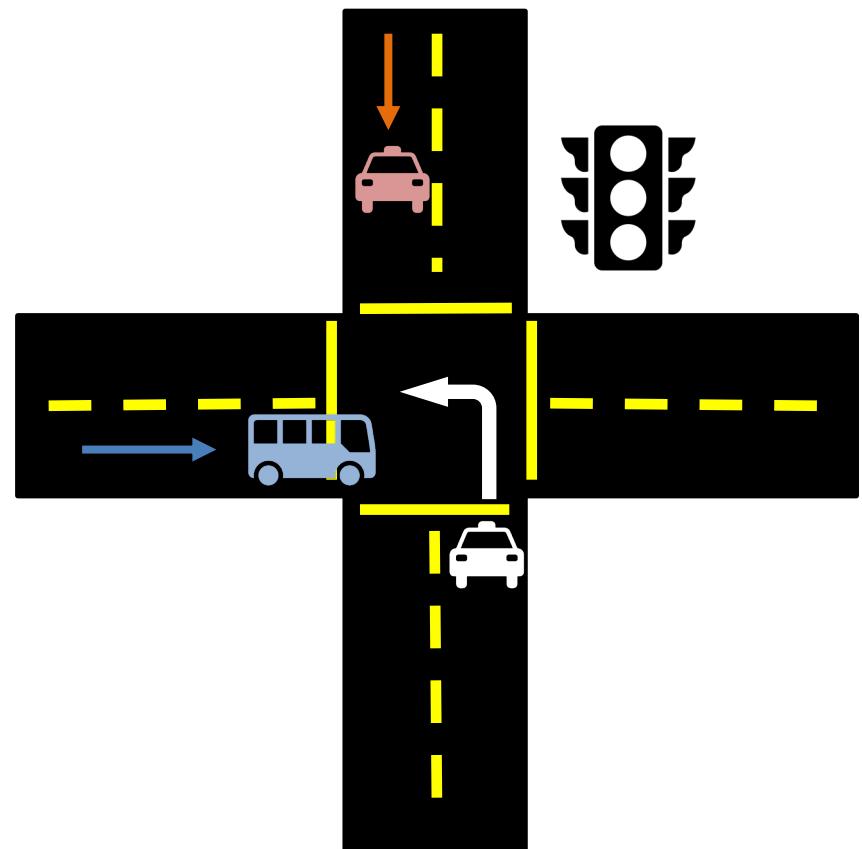
Regresión



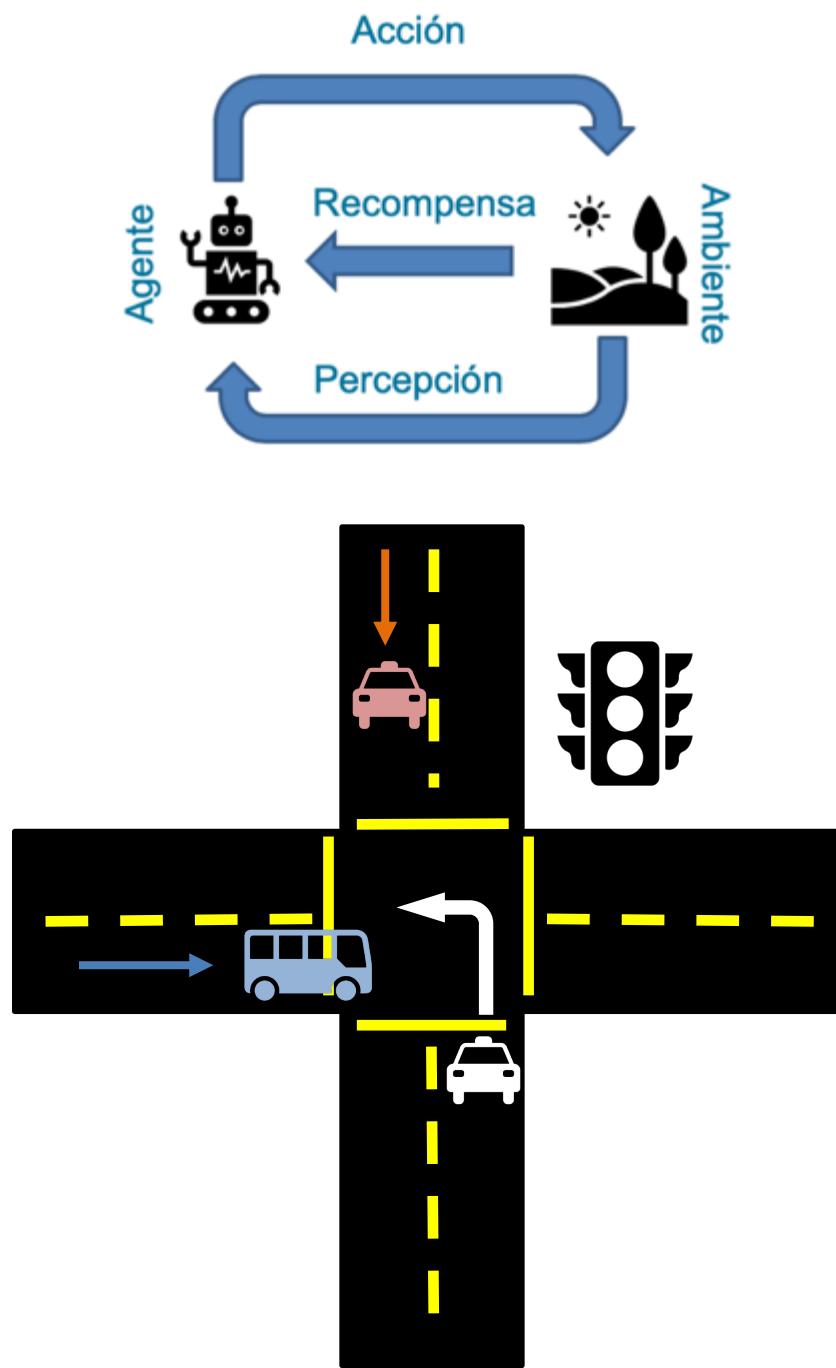
¿Qué problemas resuelve?



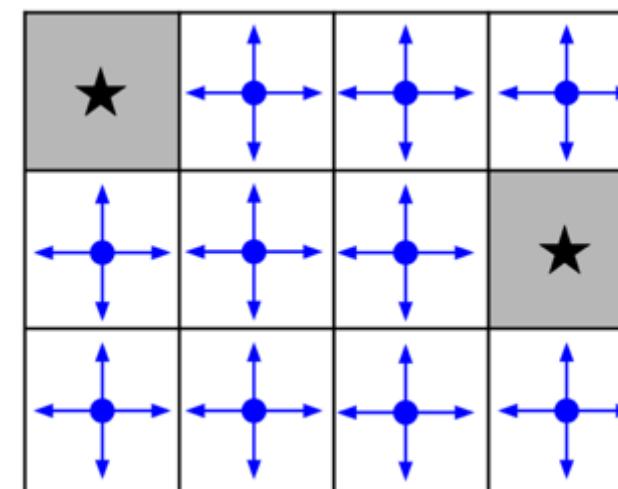
Clustering



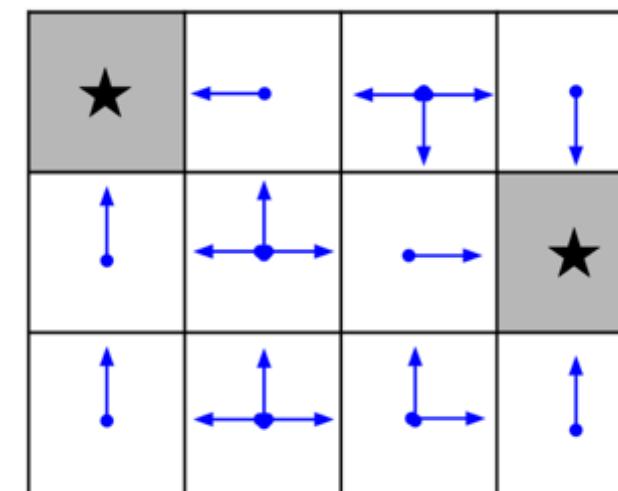
Toma de decisión



Aprendizaje por refuerzo



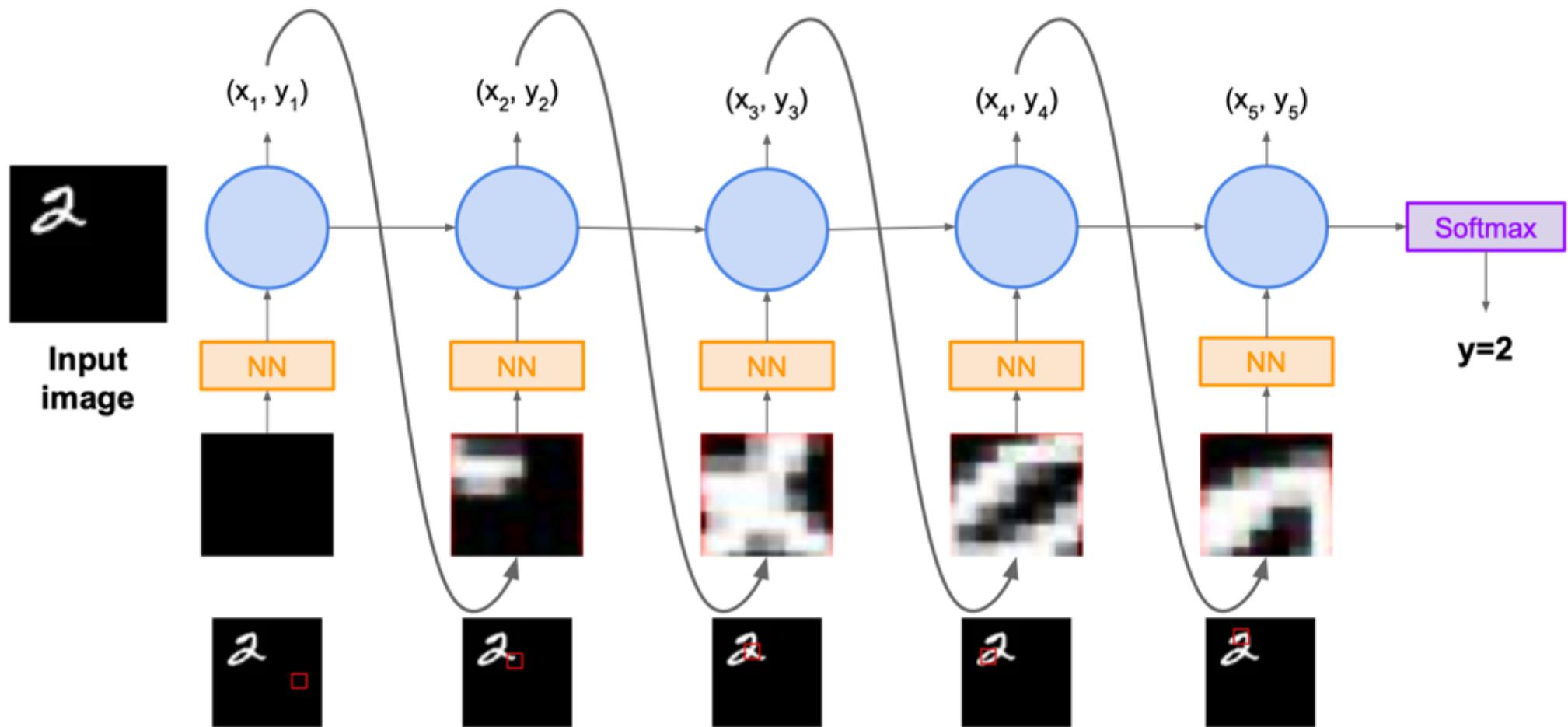
Política aleatoria



Política óptima

Reconocimiento óptico de caracteres (OCR) (2014)

NN: Red Neuronal.

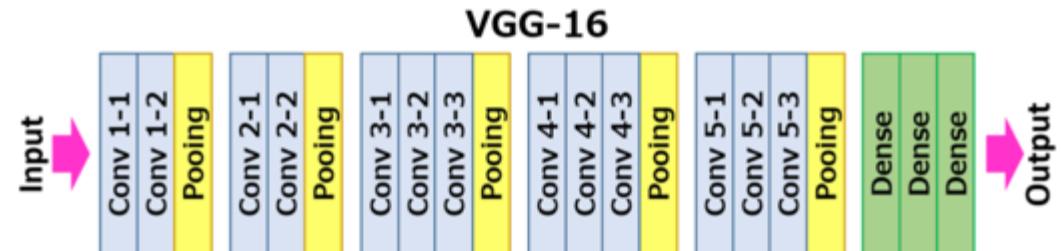


"Recurrent Models of Visual Attention" V Mnih, N Heess, A Graves K Kavukcuoglu 2014 Google DeepMind
<https://arxiv.org/abs/1406.6247>
Imagen: CS231n: Convolutional Neural Networks for Visual Recognition – Stanford University

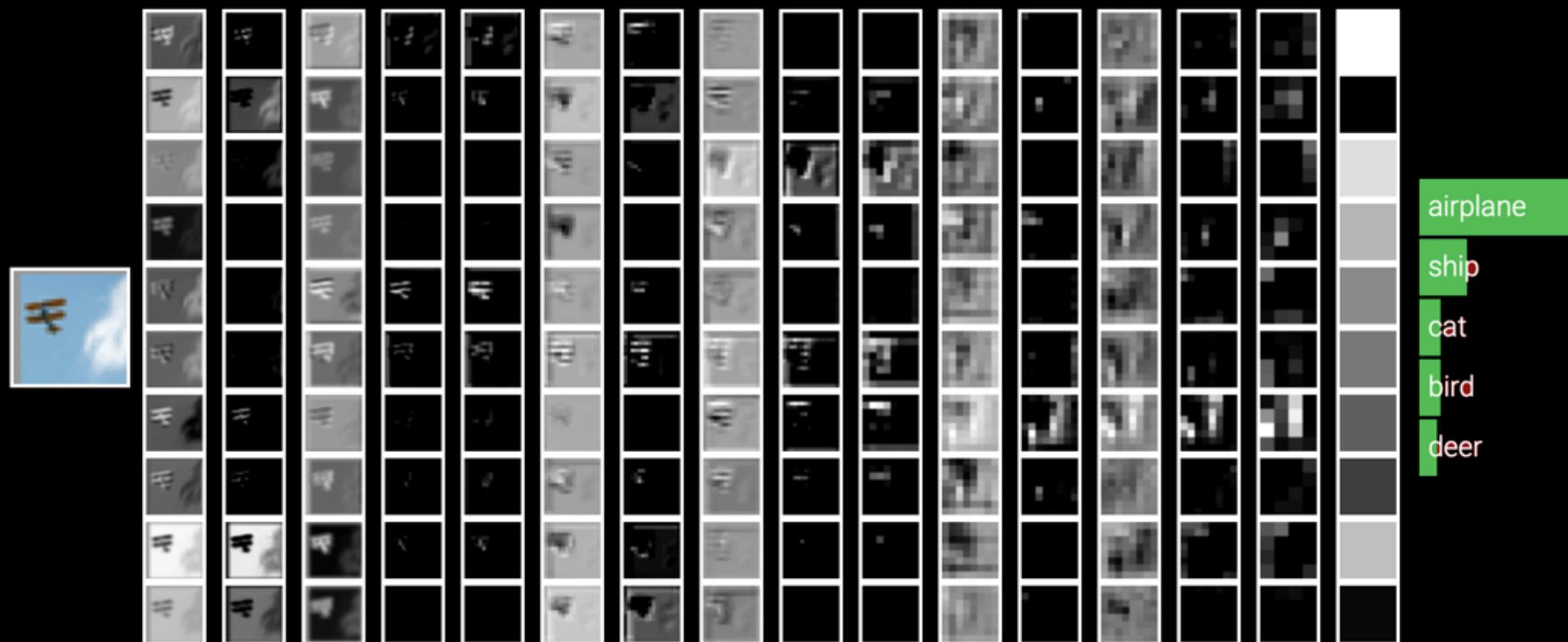
Clasificación de Imágenes – Aprendizaje Profundo

VGG-16 y VGG-19 (2014)

VERY DEEP CONVOLUTIONAL NETWORKS FOR
LARGE-SCALE IMAGE RECOGNITION
Karen Simonyan* & Andrew Zisserman
Visual Geometry Group, Department of Engineering
Science, University of Oxford
<https://arxiv.org/pdf/1409.1556.pdf>



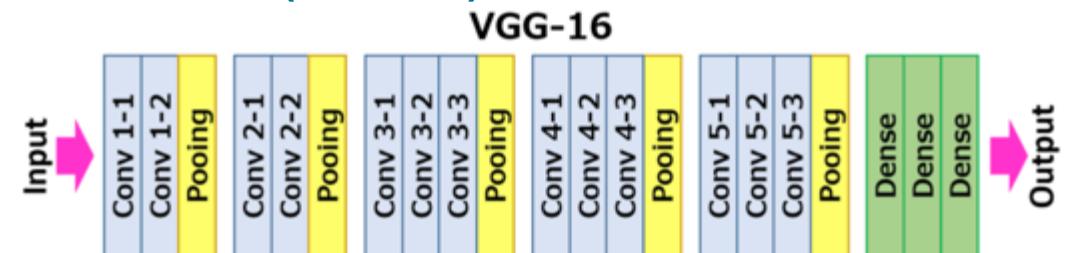
Imágenes extraídas de <http://cs231n.stanford.edu>



Clasificación de Imágenes – Aprendizaje Profundo

VGG-16 y VGG-19 (2014)

VERY DEEP CONVOLUTIONAL NETWORKS FOR
LARGE-SCALE IMAGE RECOGNITION
Karen Simonyan* & Andrew Zisserman
Visual Geometry Group, Department of Engineering
Science, University of Oxford
<https://arxiv.org/pdf/1409.1556.pdf>



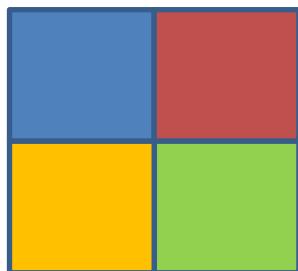
Imágenes extraídas de <http://cs231n.stanford.edu>





Una imagen es una matriz, o un vector

2x2



4



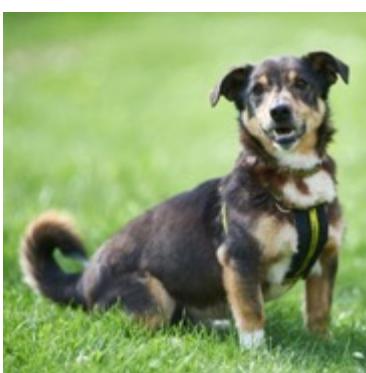
32 x 32



1024



800 x 800



640.000

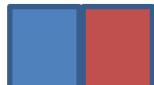




Maldición de la Dimensión

Supongamos que se busca clasificar un punto $[x,y]$, como perteneciente, o no, al círculo de radio 1:

$$A_1 = \left\{ x, y \in [-1, 1] / \sqrt{x^2 + y^2} \leq 1 \right\}$$



$$P(A_1) = 0.78$$

Si aumentamos la dimensión:



$$P(A_1) = 0.52$$

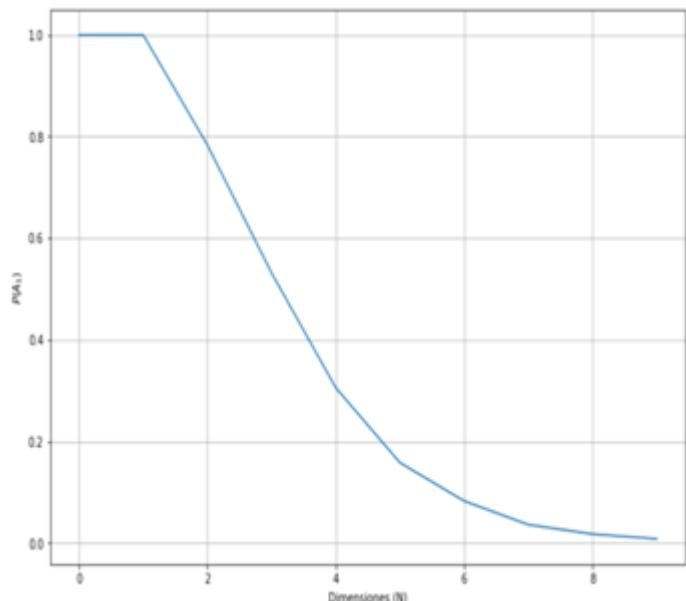
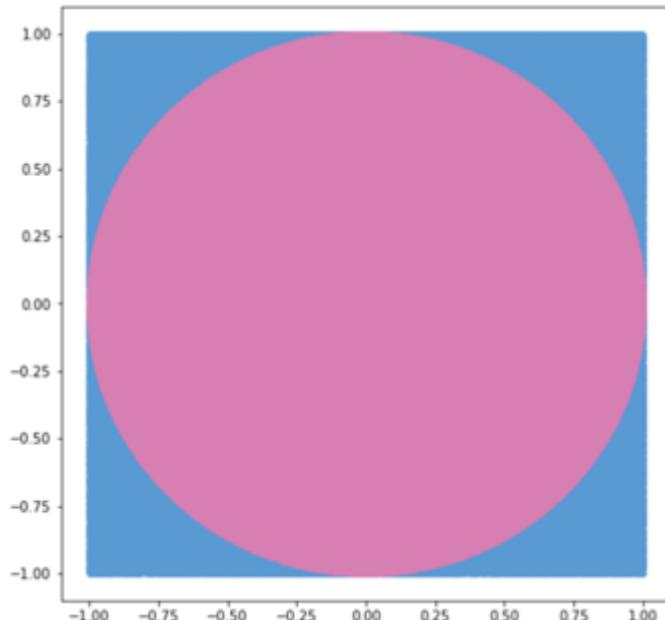


$$P(A_1) = 0.3$$



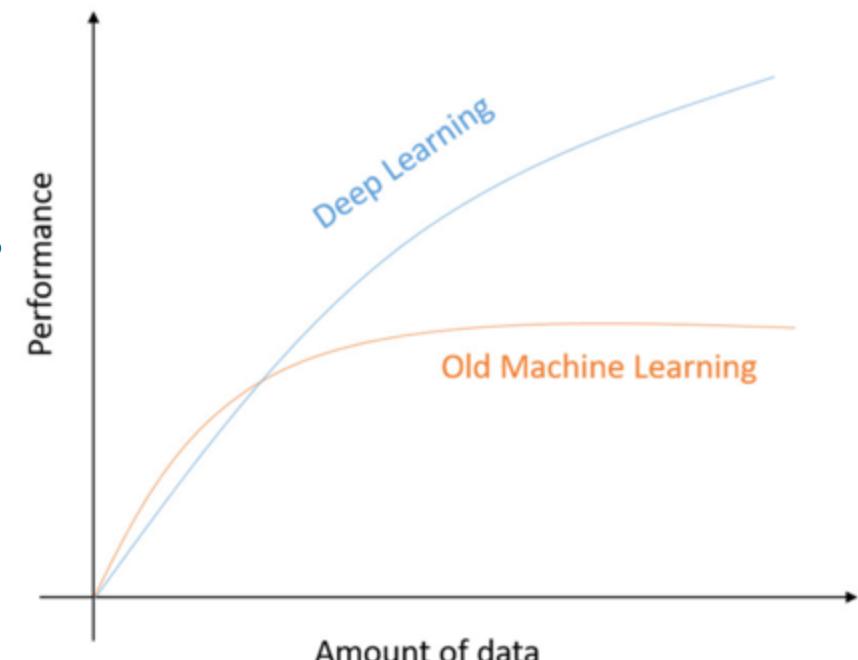
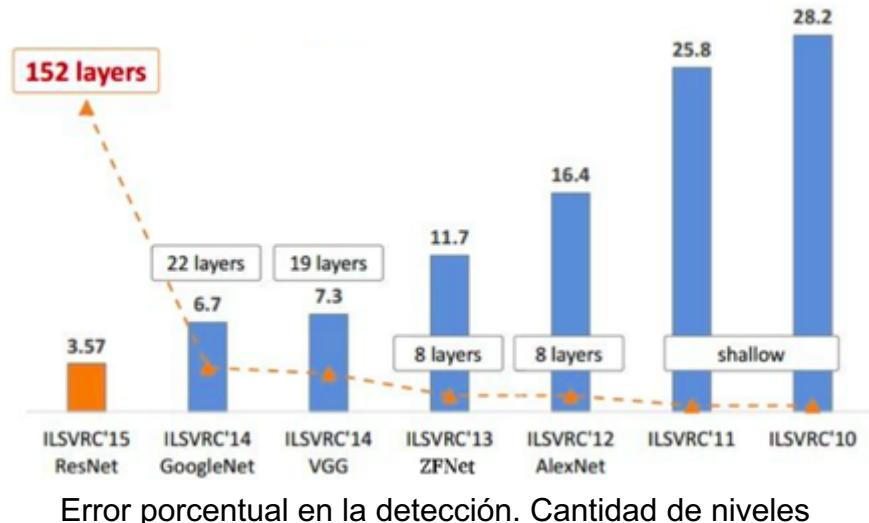
$$P(A_1) = 0$$

¿Entonces? ¿Porqué funciona?



Aprendizaje Profundo (Deep Learning)

- Modelos con gran cantidad de parámetros.
- Procedimientos y técnicas de entrenamiento
- Matemática y Teo de Optimización
- Hardware dedicado y placas de video
- **DATOS** de entrenamiento Pero muchos
- Tiempo y Energía Eléctrica.





Conjunto de Datos de Entrenamiento

MNIST Data Base (30.000 patrones)

<http://yann.lecun.com/exdb/mnist/>

0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9

Free Music Archive (343 días de música)

<https://freemusicarchive.org>

Common Objects in Context (COCO)

Microsoft (2.500.000 de fotos)

<http://cocodataset.org/#home>



Enron Email Dataset (500.000 correos)

<https://www.cs.cmu.edu/~enron/>

Twitter - Facebook

Utilizadas para análisis de sentimientos

Nuestros Proyectos

¿Necesitan ese volumen de datos?

Por suerte, NO

1888 – Santiago Ramón y Cajal. Neurociencia

Demuestra la unidad neuronal y la relación de contigüidad entre ellas.

Queda invalidada la hipótesis de continuidad o de grilla.

S.R. Cajal, Estructura de los centros nerviosos de las aves, Rev. Trimb.
Histol. Norm. Patol. 1 (1888) 1–10.



1928 –

Transmission of Information¹

By R. V. L. HARTLEY

SYNOPSIS: A quantitative measure of "information" is developed which is based on physical as contrasted with psychological considerations. How the rate of transmission of this information over a system is limited by the

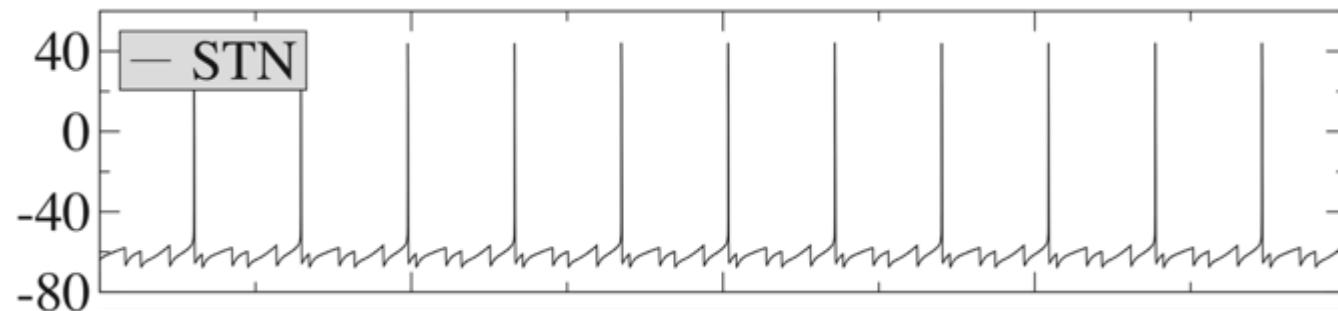
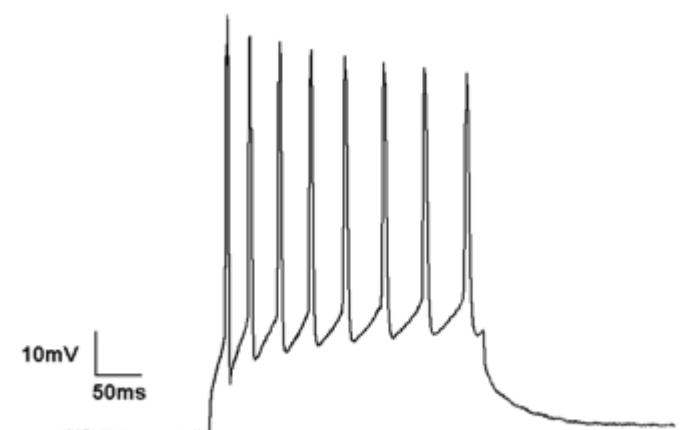
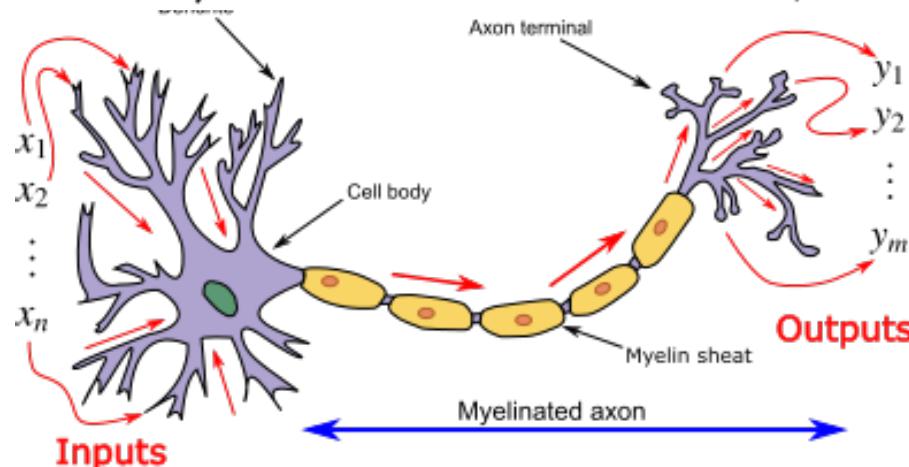


1943

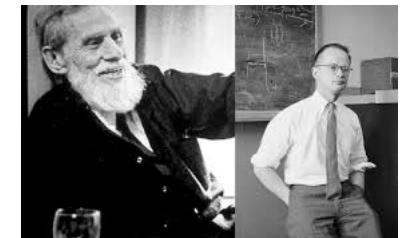
A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. McCULLOCH and WALTER H. PITTS

Because of the “all-or-none” character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of

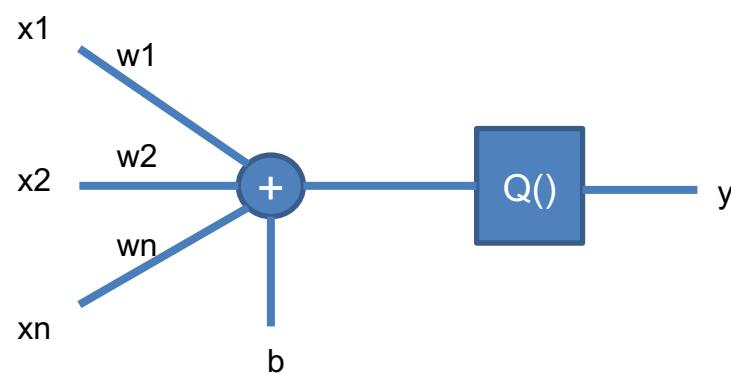
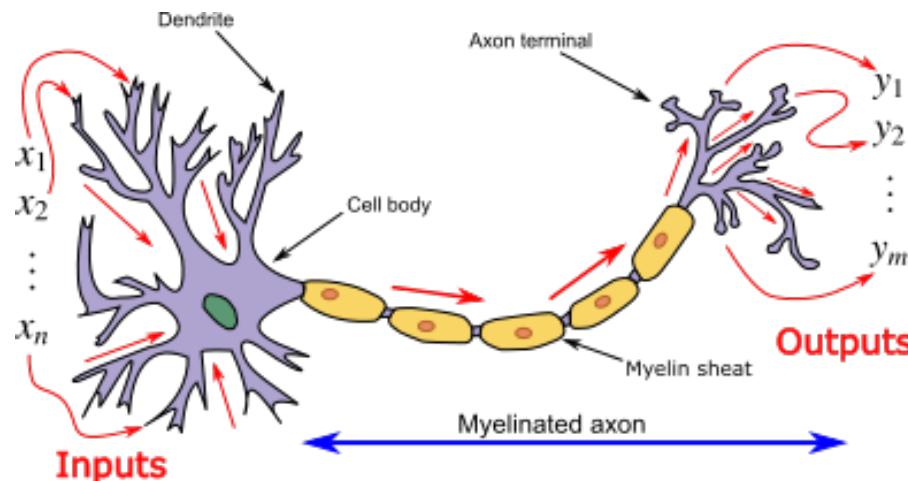


Rubchinsky et al 2003





1943 – Warren McCulloch y Walter Pitts. Modelo matemático de la neurona



$$y = Q \left(\sum_j w_j x_j + b \right)$$

Donde $Q()$ es una función de activación:

$$Q(v) = \begin{cases} 1 & \text{si } v > \theta \\ 0 & \text{si } v \leq \theta \end{cases}$$

W son los pesos y **b** está relacionado con el umbral de activación

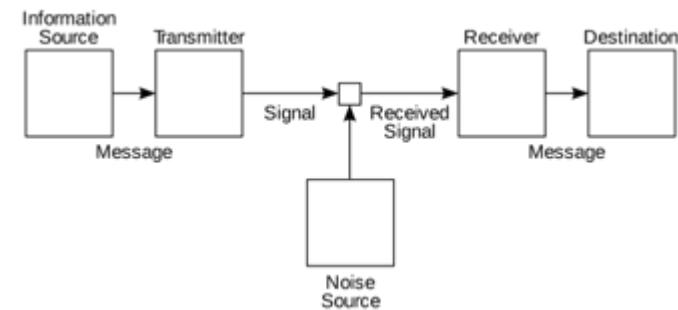
1948 – Claude Shannon: A Mathematical Theory of Communication

The fundamental problem of communication is that of **reproducing at one point either exactly or approximately a message selected at another point**. Frequently the messages have meaning . . . These semantic aspects of communication are irrelevant to the engineering problem.

– A Mathematical Theory of Communication, Shannon (1948)



- Modelo matemático de fuente y de canal
- Medidas de información
- Límites en la compresión y la transmisión
- Bits como interface entre fuente y canal



Aplicaciones:

Inferencia estadística, procesamiento de lenguaje natural, criptografía, selección de modelos, reconocimiento de patrones, visión de computadoras, computación cuántica, astronomía, genética entre otras.

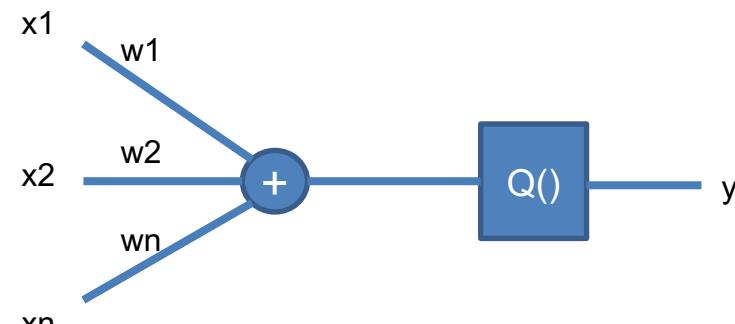
1950 – Alan Turing: Turing Test

1956 – Jhon Mc Carthy : Dartmouth College Workshop.
 (Propuesto por Claude Shannon)
 Nace AI como campo científico

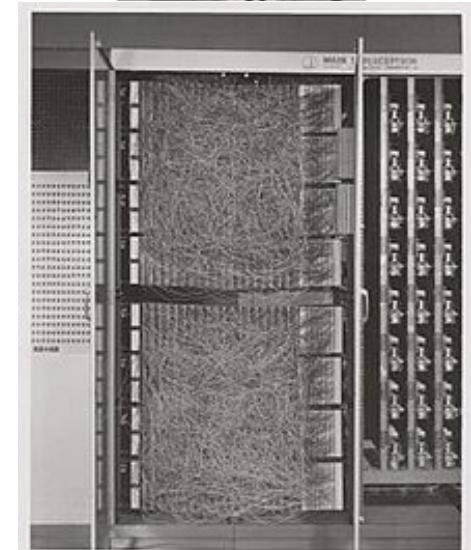


1962 – Frank Roseblatt : Perceptron.

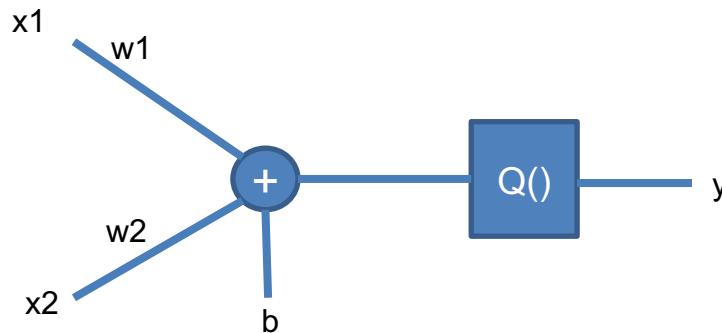
Hardware de aprendizaje supervisado
 para clasificación (lineal) binaria



Aprendizaje \rightarrow $w_j(n + 1) = w_j(n) + \rho(V_R - y_i)x_{j,k}$



1962 – Frank Roseblatt : Perceptron.



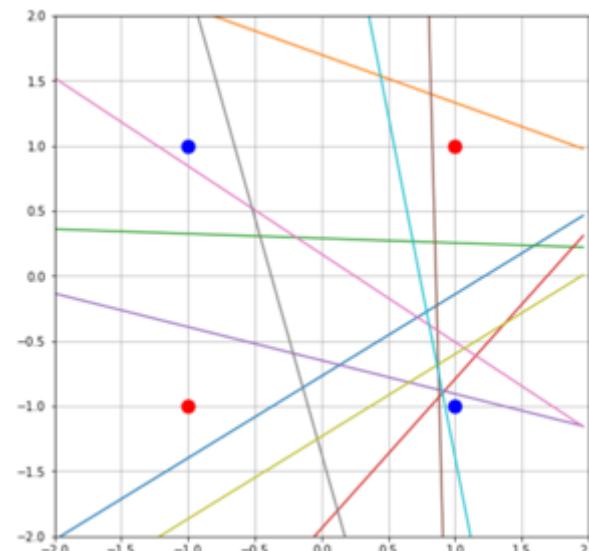
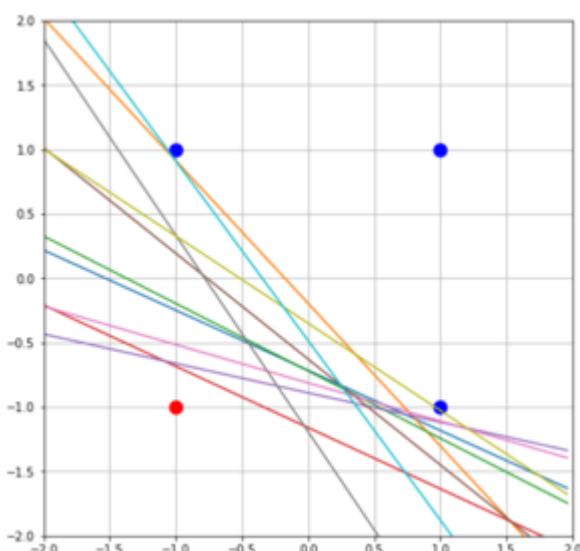
$$y = Q \left(\sum_j w_j x_j + b \right)$$

$$Q(v) = \begin{cases} 1 & \text{si } v > \theta \\ -1 & \text{si } v \leq \theta \end{cases}$$

$$e(n) = V_R - y_i(n)$$

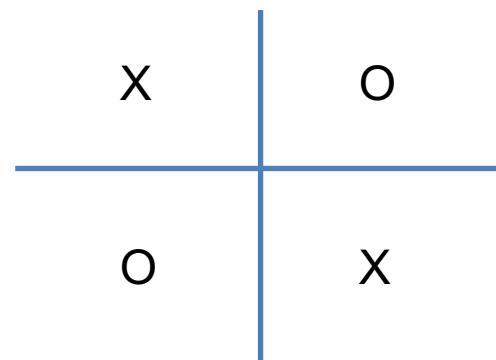
$$w_j(n+1) = w_j(n) + \rho(V_R - y_i(n))x_{j,k} \quad \rho < 1$$

Hiperplano de clasificación
 $w_1x_1 + w_1x_1 + b = 0$



1963 – Vladimir Vapnik y Alexey Chervoneskis
Support Vector Machines (versión lineal)

1969 – Marvin Minsky y S.A. Papert:
Perceptrons. an introduction to computational geometry (MIT Press)



1973 – Reporte de James Lighthill
Plantea que la explosión combinatoria hace que no se puedan abordar problemas de la realidad

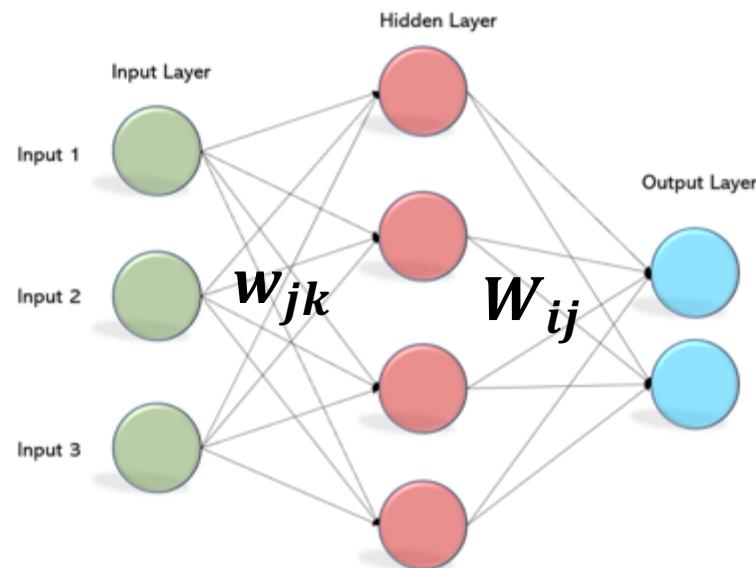
PRIMER INVIERNO



1982 – John Hopfield : Red de Hopfield – Memoria asociativa



1986 – David Rumelhart: Redes multilayer y Back-Propagation



$$o_i = Q \left(\sum_j w_{ij} h_j \right)$$

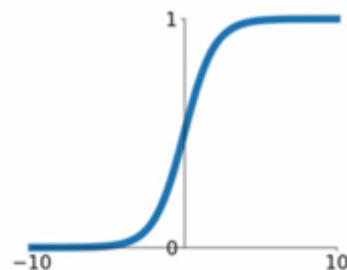


Función de Activación Q

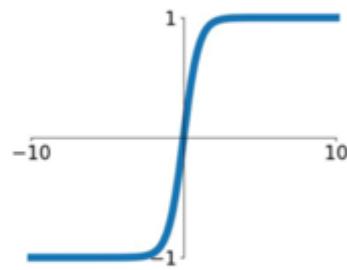
$$o_i = \sigma \left(\sum_i W_{ij} h_j \right)$$

Sigmoid

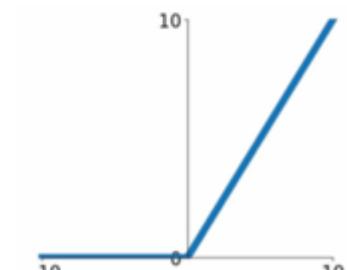
$$\sigma(x) = \frac{1}{1+e^{-x}}$$

**tanh**

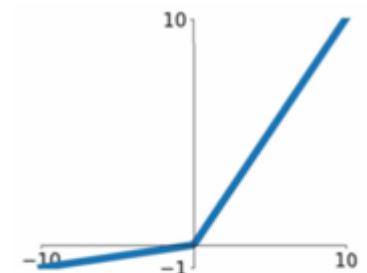
$$\tanh(x)$$

**ReLU**

$$\max(0, x)$$

**Leaky ReLU**

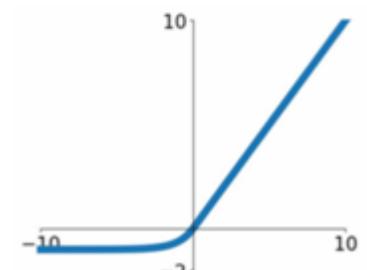
$$\max(0.1x, x)$$

**Maxout**

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



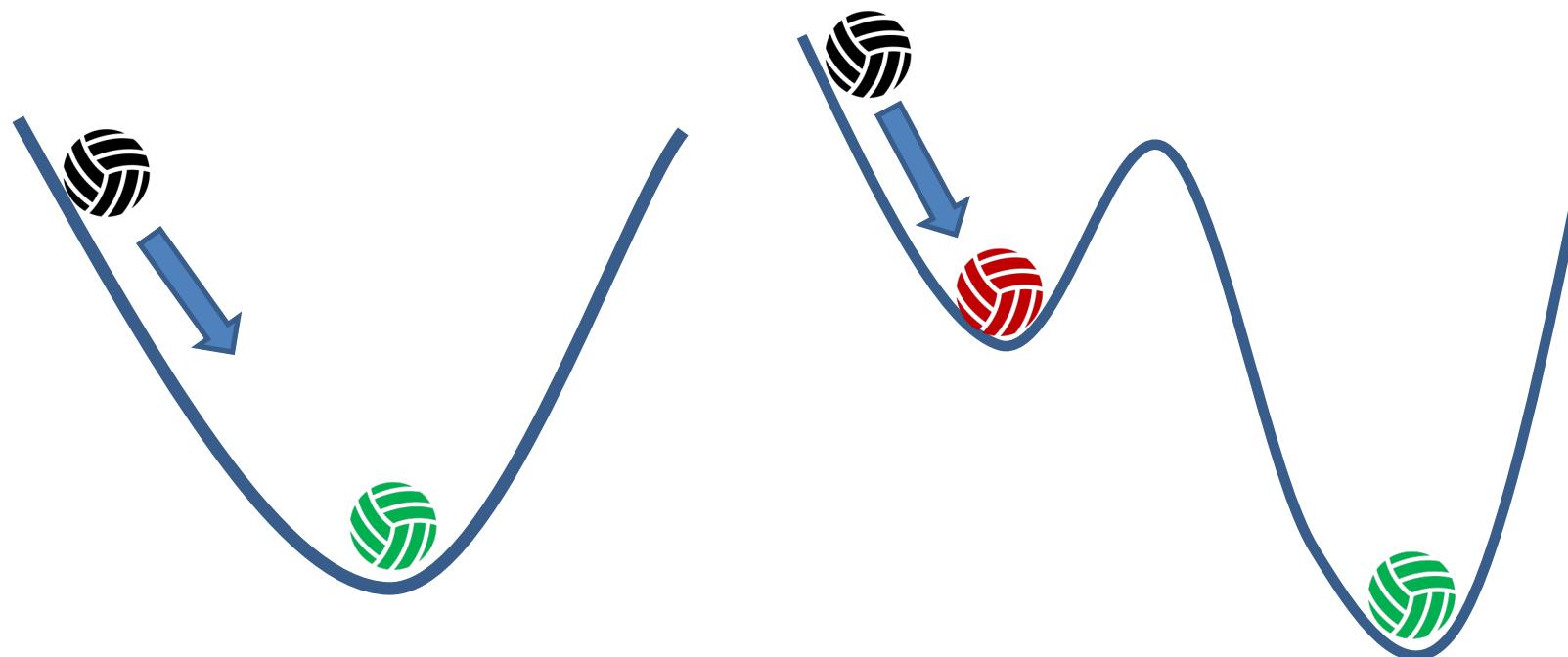
Aprendizaje - Gradiente descendente

$$O_i = Q \left(\sum_i W_{ij} h_j \right)$$

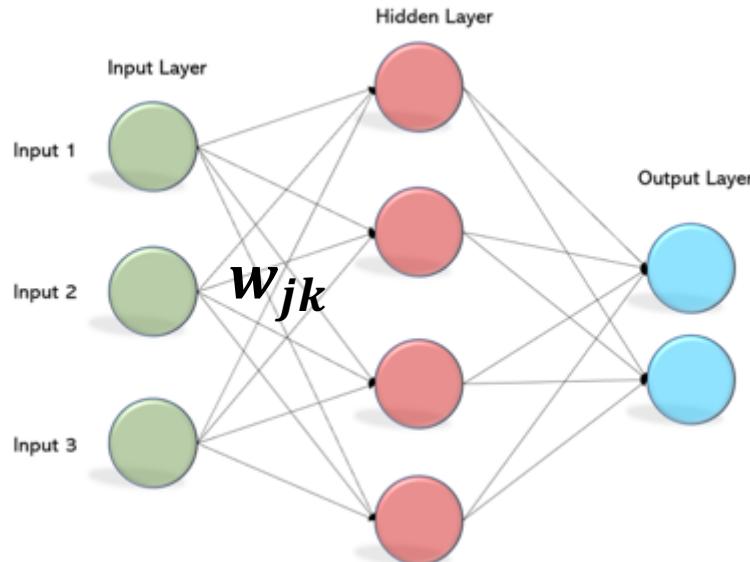
$$E(W_{ij}) = \frac{1}{2} \sum_i (V_R - O_i)^2$$

$$W_{ij}^* = \min E(W_{ij}) \forall W_{ij}$$

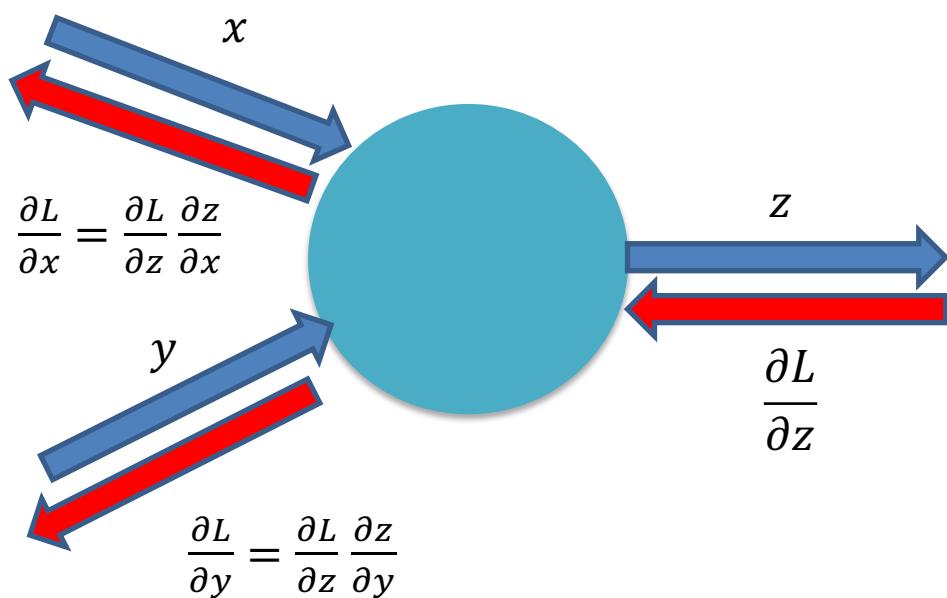
$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}}$$



Backpropagation



$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}}$$



Regla de la cadena
+
Algunas condiciones

Método rápido de optimización

Teorema de Aproximación Universal.

Sea $\phi : \mathbb{R} \rightarrow \mathbb{R}$ una función no constante, acotada y continua, a la que se la denomina función de activación. Sea $U_m : [0, 1]^m$ el hipercubo de dimensión m , y sea $C(U_m)$ el espacio de funciones reales continuas sobre U_m .

Entonces, dado cualquier $\epsilon > 0$ y cualquier función $f \in C(U_m)$, existe un entero N , constantes $\alpha_i, \beta_i \in \mathbb{R}$ y un vector $w_i \in \mathbb{R}^m$ para $i \in 1, \dots, N$, de manera tal que se define:

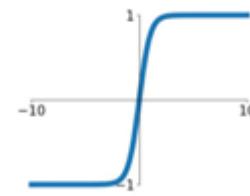
$$F(x) = \sum_{i=1}^N \alpha_i \phi(w_i^T x + \beta_i)$$

como una aproximación de la función f , tal que $|F(x) - f(x)| < \epsilon$

1989 - George Cybenko para funciones de activación sigmoideas.

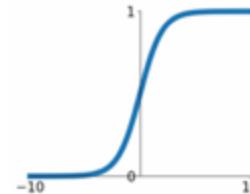
1991 - Kurt Hornik demuestra que es suficiente la arquitectura feedforward.

tanh
 $\tanh(x)$



Sigmoid

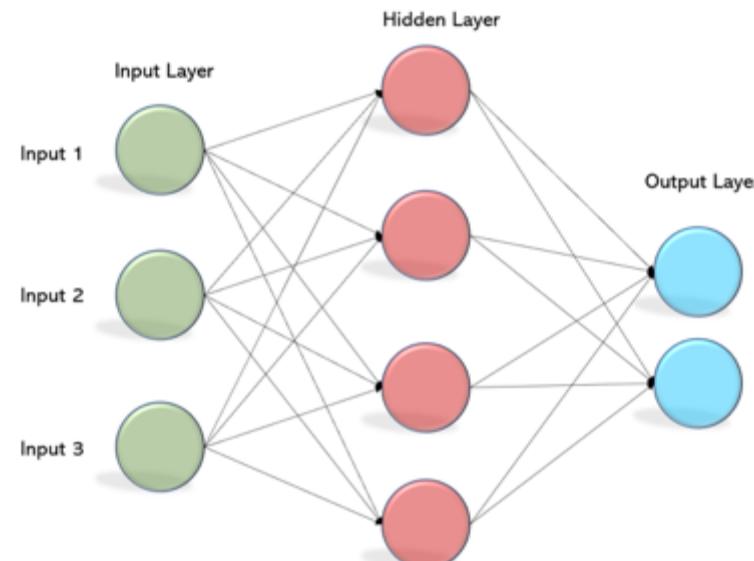
$$\sigma(x) = \frac{1}{1+e^{-x}}$$

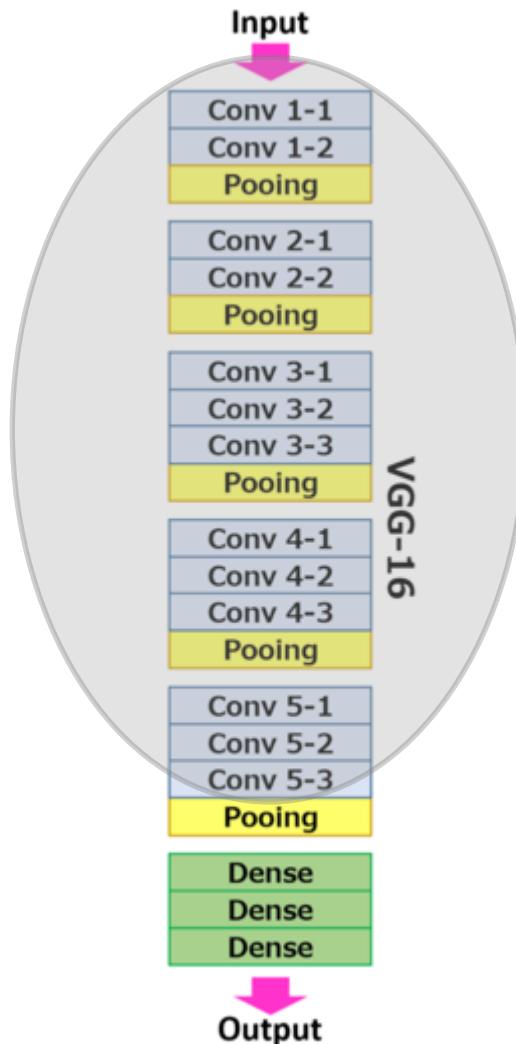


Se puede representar
cualquier función continua
con un perceptron



Dense Network Fully Connected Network





Redes Neuronales Convolucionales

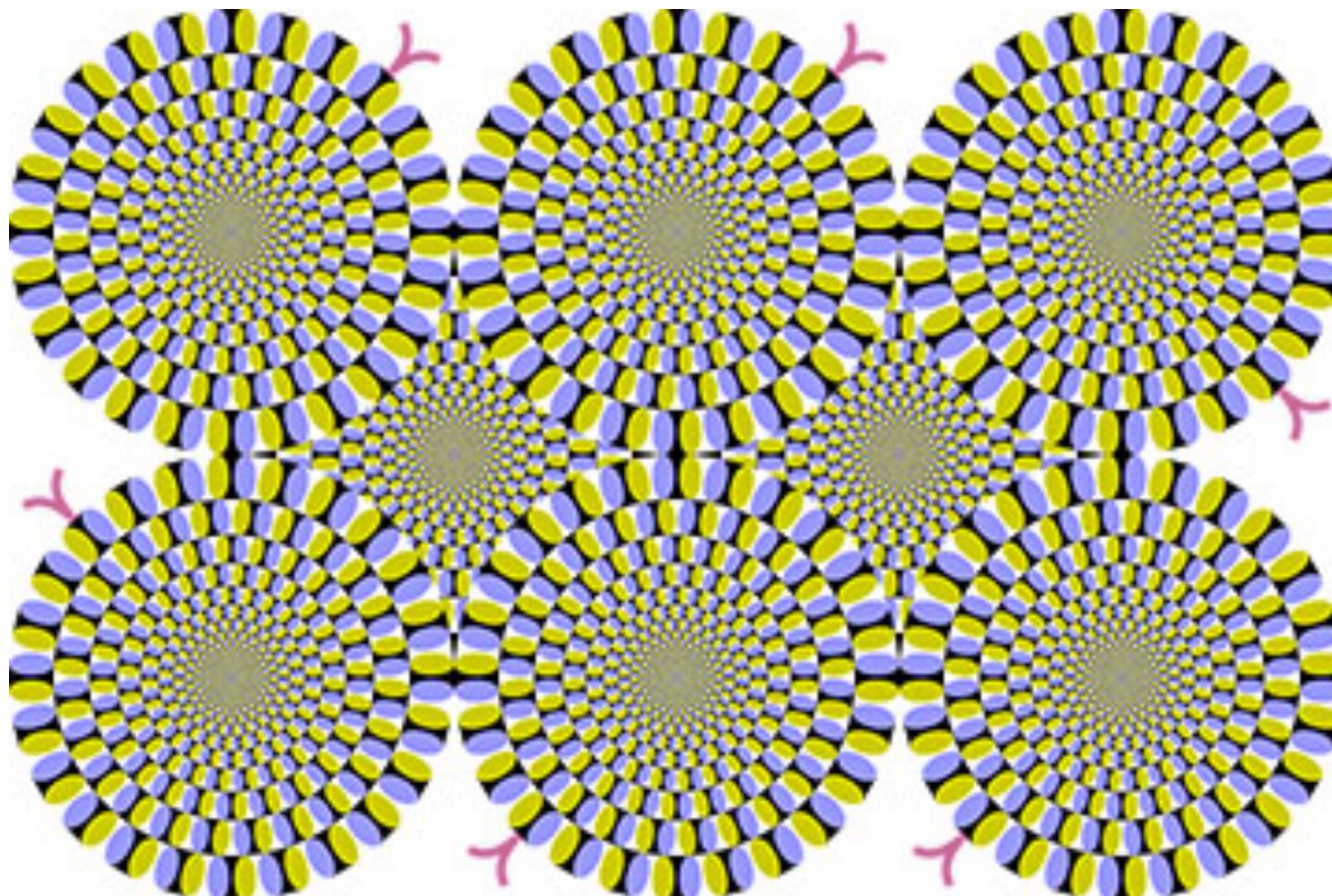


Procesamiento Visual – Campo visual



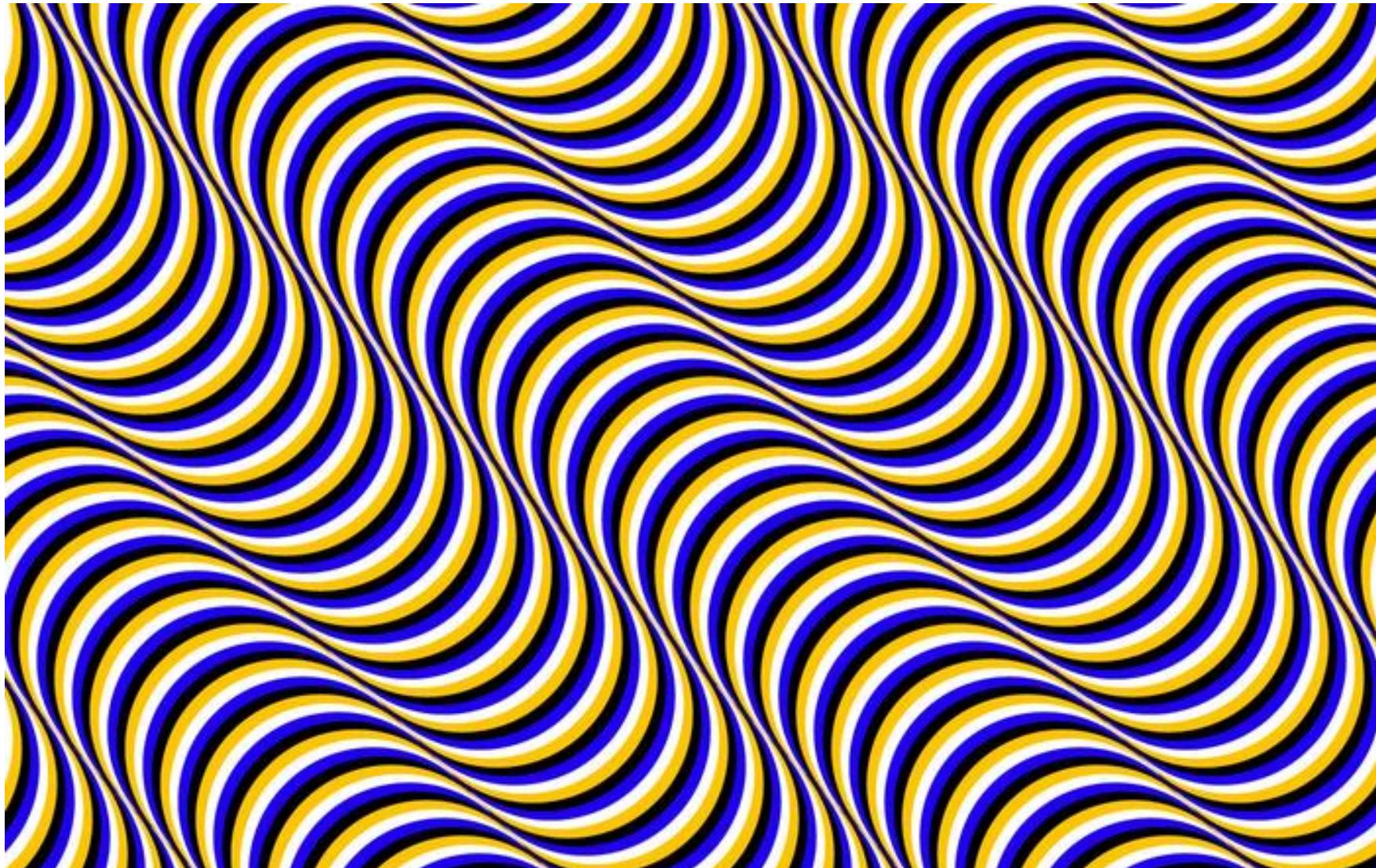


Procesamiento Visual





Procesamiento Visual





Procesamiento Visual





J. Physiol. (1959) 148, 574-591

RECEPTIVE FIELDS OF SINGLE NEURONES IN
THE CAT'S STRIATE CORTEX

BY D. H. HUBEL* AND T. N. WIESEL*

From the Wilmer Institute, The Johns Hopkins Hospital and
University, Baltimore, Maryland, U.S.A.

J. Physiol. (1962), 160, pp. 106-154

With 2 plates and 20 text-figures

Printed in Great Britain

RECEPTIVE FIELDS, BINOCULAR INTERACTION
AND FUNCTIONAL ARCHITECTURE IN
THE CAT'S VISUAL CORTEX

BY D. H. HUBEL AND T. N. WIESEL

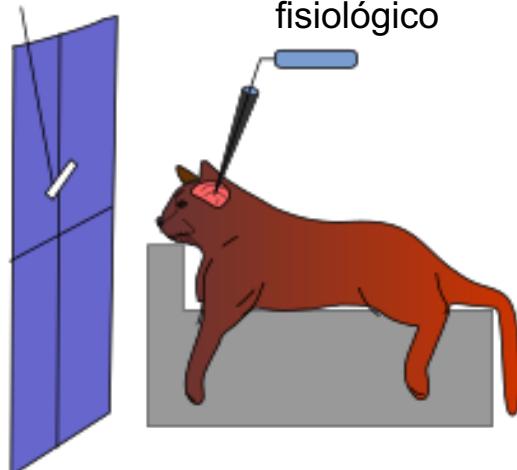
From the Neurophysiology Laboratory, Department of Pharmacology
Harvard Medical School, Boston, Massachusetts, U.S.A.

(Received 31 July 1961)

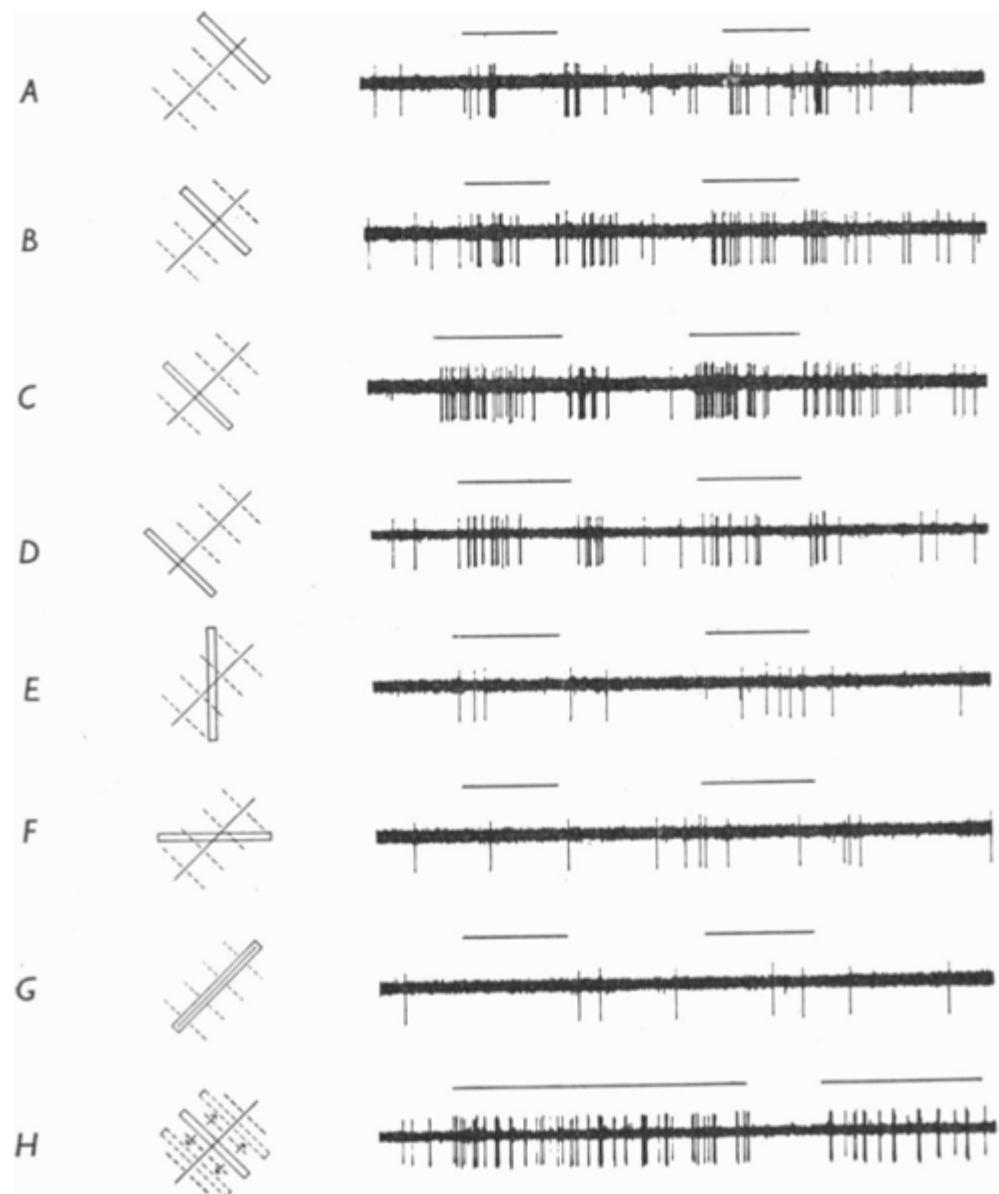
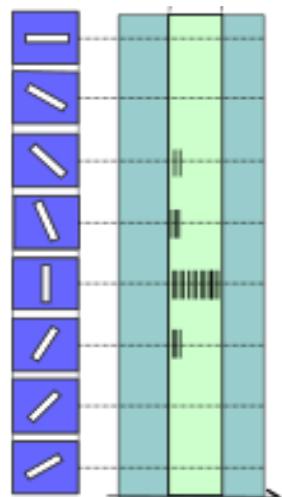
What chiefly distinguishes cerebral cortex from other parts of the central nervous system is the great diversity of its cell types and interconnexions. It would be astonishing if such a structure did not profoundly modify the response patterns of fibres coming into it. In the cat's visual cortex, the receptive field arrangements of single cells suggest that there is indeed a degree of complexity far exceeding anything yet seen at lower levels in the visual system.

Procesamiento Visual

Barra lumínica Registro electrofisiológico



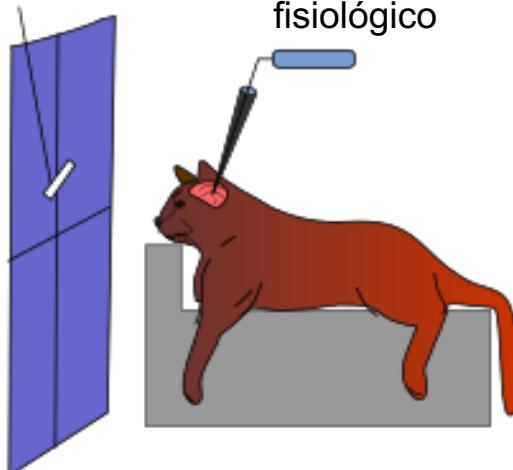
Respuesta a estímulos



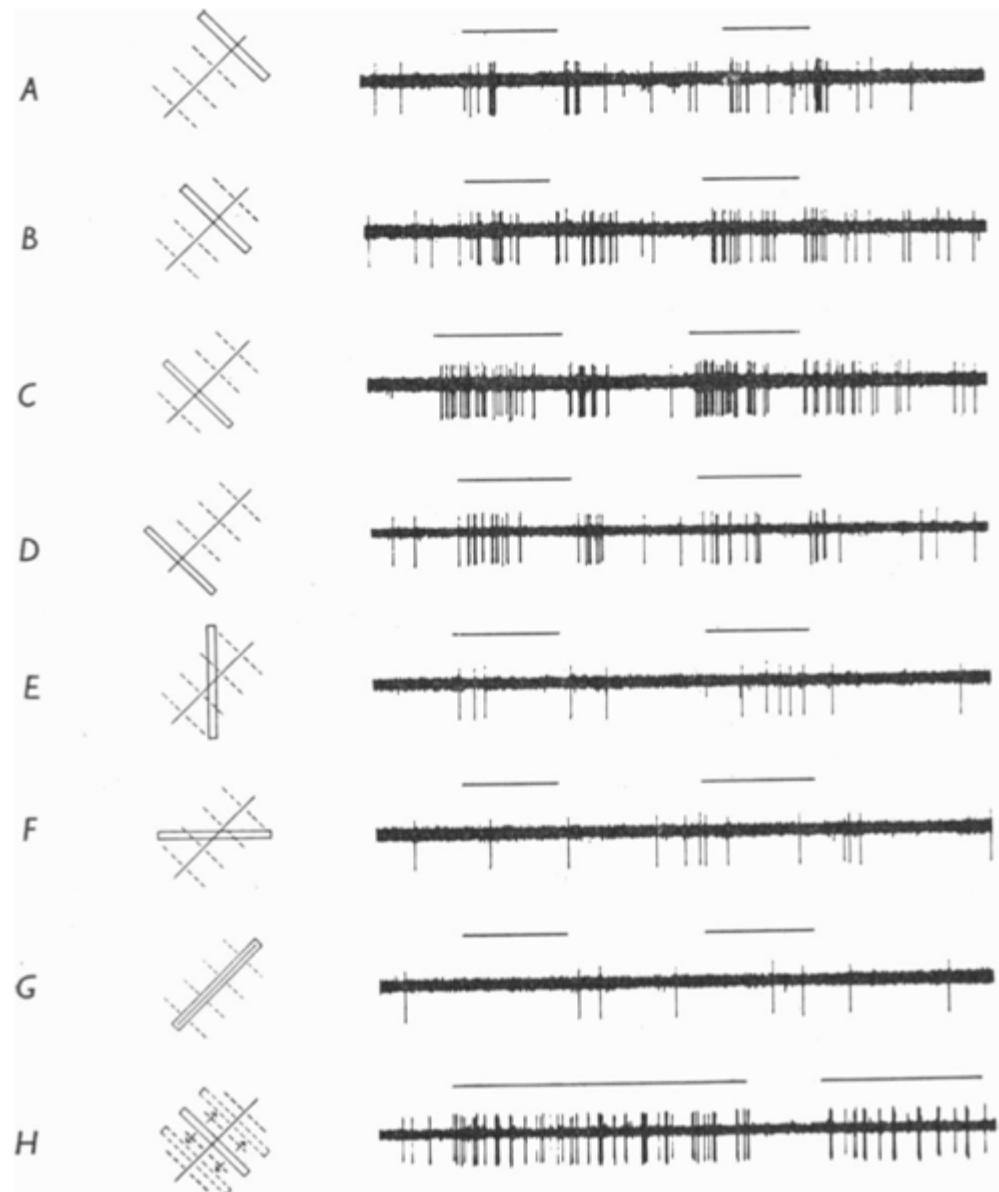
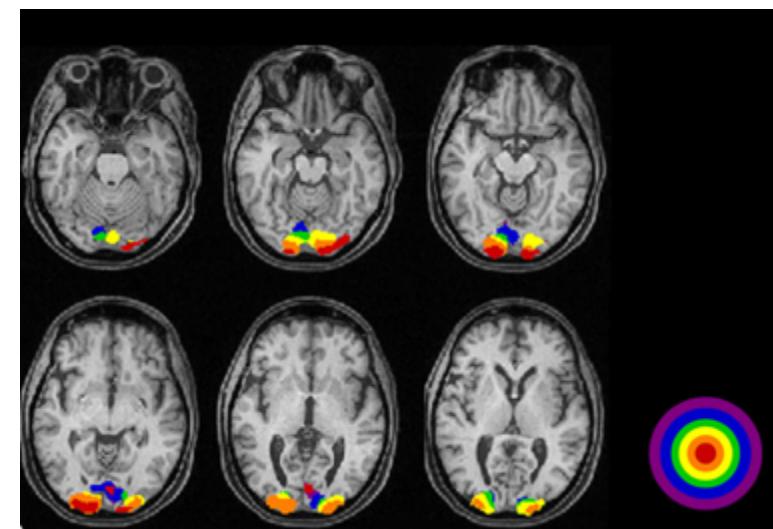
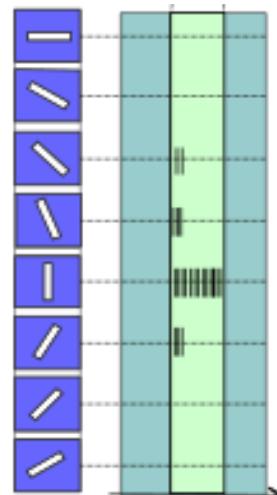


Procesamiento Visual

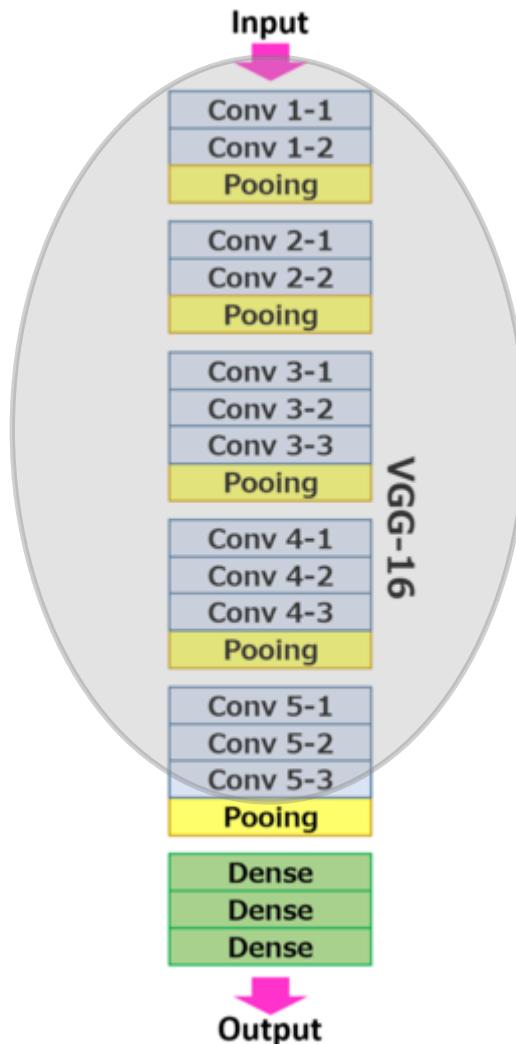
Barra lumínica Registro electrofisiológico



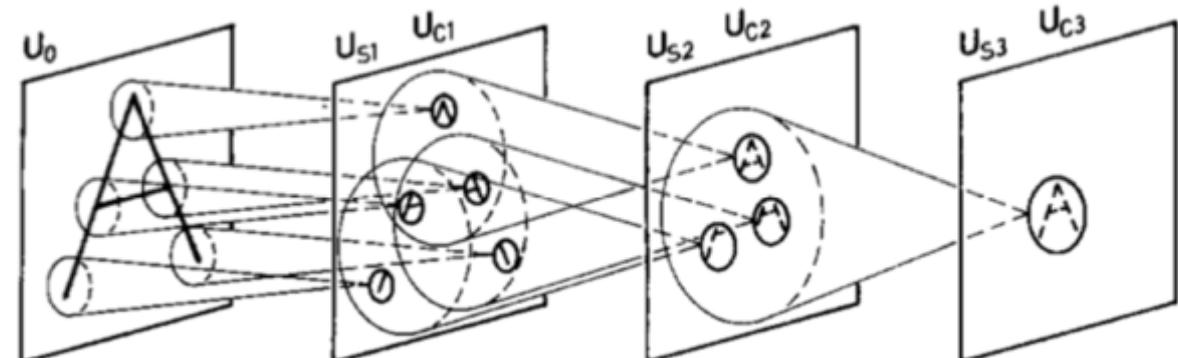
Respuesta a estímulos



Redes Neuronales Convolucionales



1980 – Kunihiko Fukushima: Neocognitron, una red convolucional no supervisada basada en el sistema visual biológico.





Redes Neuronales Convolucionales - 1989

Object Recognition with Gradient-Based Learning

Yann LeCun, Patrick Haffner, Léon Bottou, and Yoshua Bengio

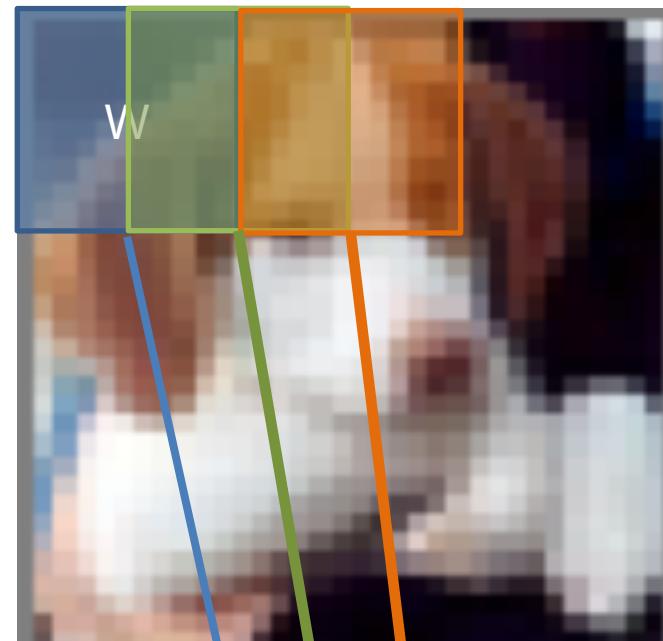
AT&T Shannon Lab, 100 Schulz Drive, Red Bank NJ 07701, USA,
yann@research.att.com
<http://www.research.att.com/~yann>

Abstract. Finding an appropriate set of features is an essential problem in the design of shape recognition systems. This paper attempts to show that for recognizing simple objects with high shape variability such as handwritten characters, it is possible, and even advantageous, to feed the system directly with minimally processed images and to rely on learning to extract the right set of features. Convolutional Neural Networks are shown to be particularly well suited to this task. We also show that these networks can be used to recognize multiple objects without requiring explicit segmentation of the objects from their surrounding. The second



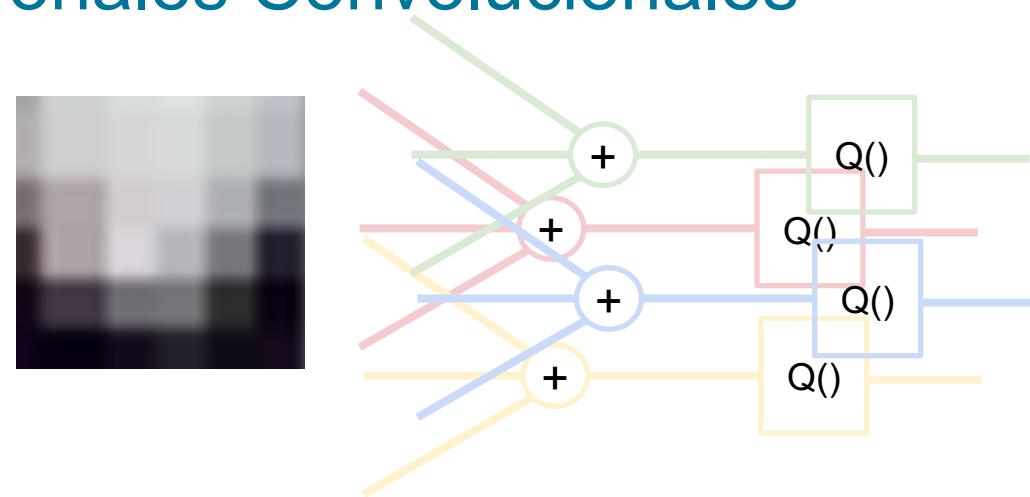
Redes Neuronales Convolucionales

W: Campo visual



5	2	0	1
1	3	0	0
3			

Redes Neuronales Convolucionales



Imagen

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

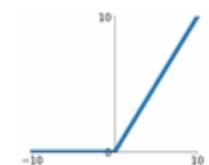
Kernel convolucional (W)

<table border="1"><tbody><tr><td>1</td><td>0</td></tr><tr><td>0</td><td>-1</td></tr></tbody></table>	1	0	0	-1	<table border="1"><tbody><tr><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td></tr></tbody></table>	1	1	0	0
1	0								
0	-1								
1	1								
0	0								
<table border="1"><tbody><tr><td>0</td><td>-1</td></tr><tr><td>1</td><td>1</td></tr></tbody></table>	0	-1	1	1	<table border="1"><tbody><tr><td>0</td><td>0</td></tr><tr><td>1</td><td>0</td></tr></tbody></table>	0	0	1	0
0	-1								
1	1								
0	0								
1	0								

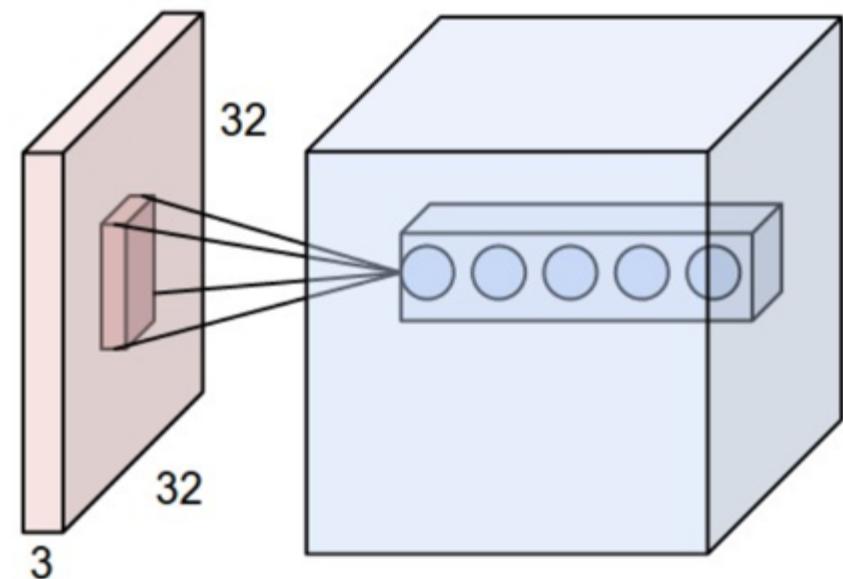
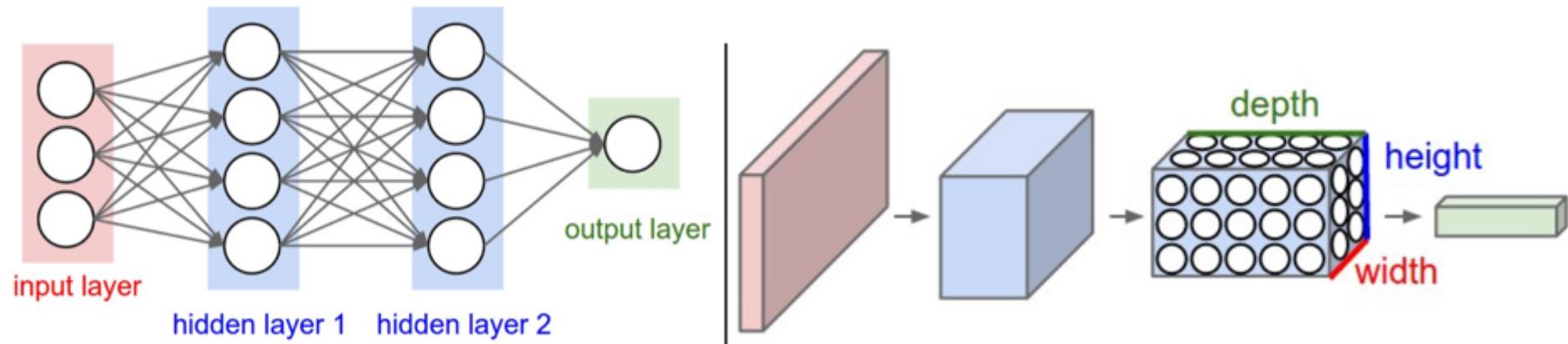
Convolución

-5	6
1	3

ReLU
 $\max(0, x)$



Redes Neuronales Convolucionales



2012

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky

University of Toronto

kriz@cs.utoronto.ca

Ilya Sutskever

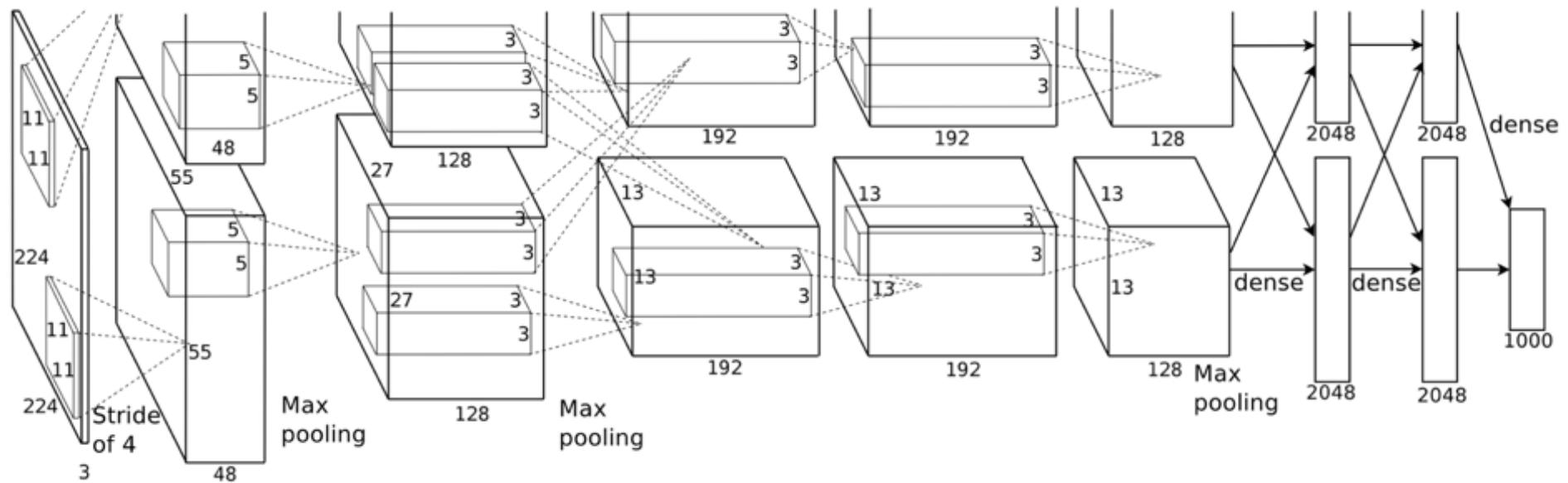
University of Toronto

ilya@cs.utoronto.ca

Geoffrey E. Hinton

University of Toronto

hinton@cs.utoronto.ca



ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky

University of Toronto

kriz@cs.utoronto.ca

Ilya Sutskever

University of Toronto

ilya@cs.utoronto.ca

Geoffrey E. Hinton

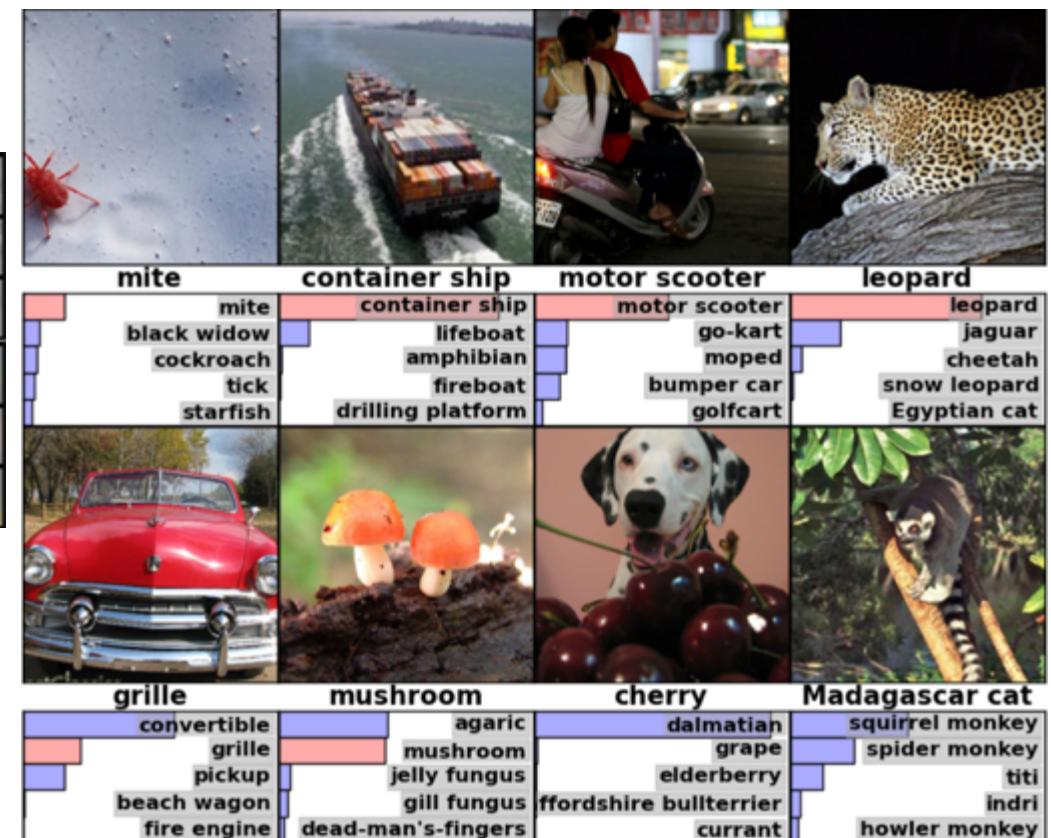
University of Toronto

hinton@cs.utoronto.ca

Kernel convolucional (W)

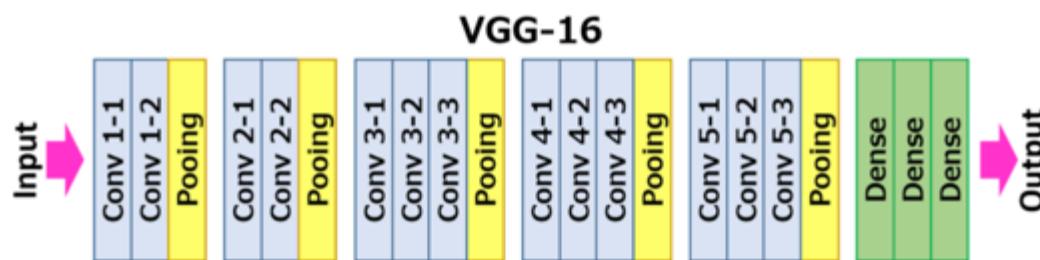


11x11x3



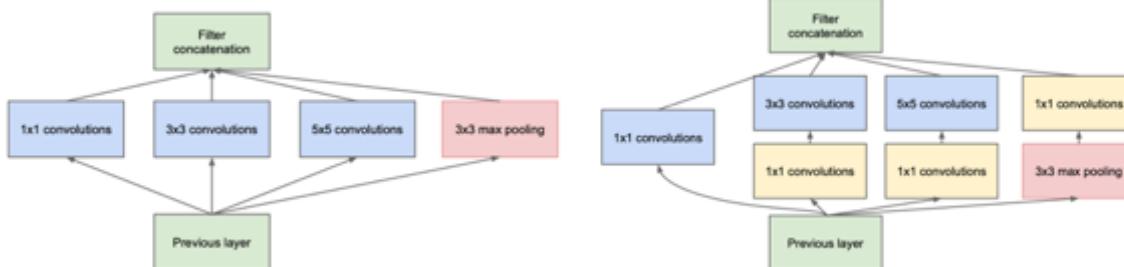
2014 – Se consolida el aprendizaje profundo

VGG [Simonyan & Zisserman 2014]



(a) Siberian husky

Inception [Szegedy et al 2014]
(Google)



(a) Inception module, naïve version

(b) Inception module with dimension reductions

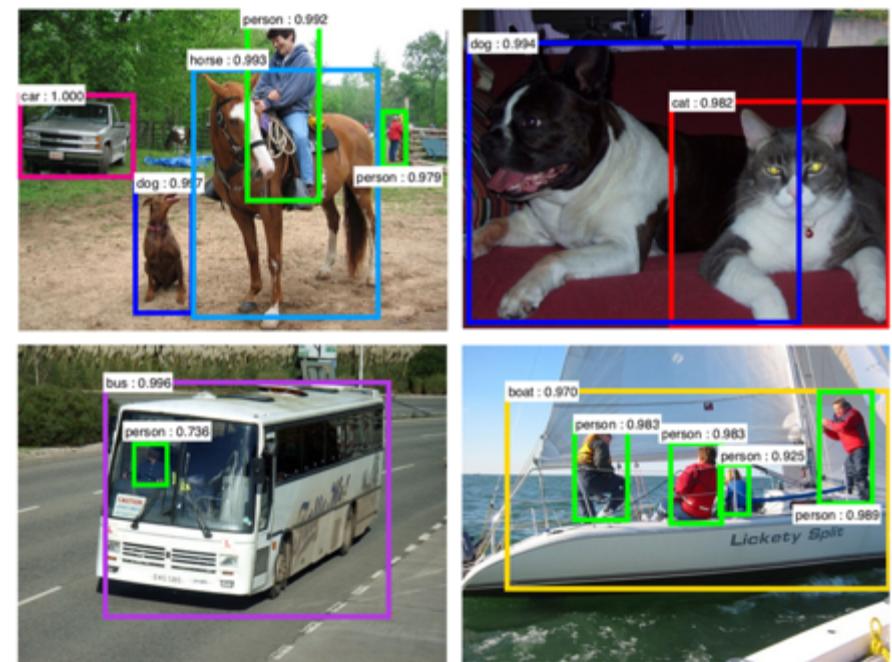
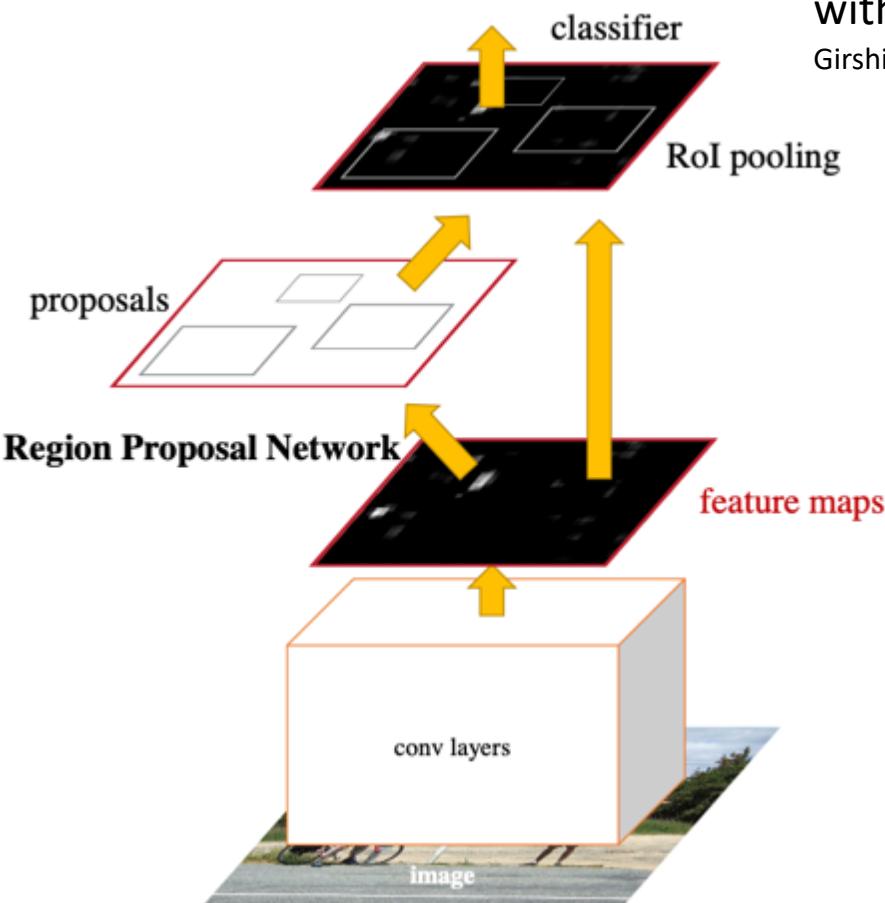


(b) Eskimo dog

Region Proposal Networks (R-CNN)

2015 “Fast R-CNN,” R. Girshick, in *IEEE International Conference on Computer Vision (ICCV)*.

2016 “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks” [https://aShaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun \(Microsoft\) rxiv.org/pdf/1506.01497.pdf](https://aShaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun (Microsoft) rxiv.org/pdf/1506.01497.pdf)

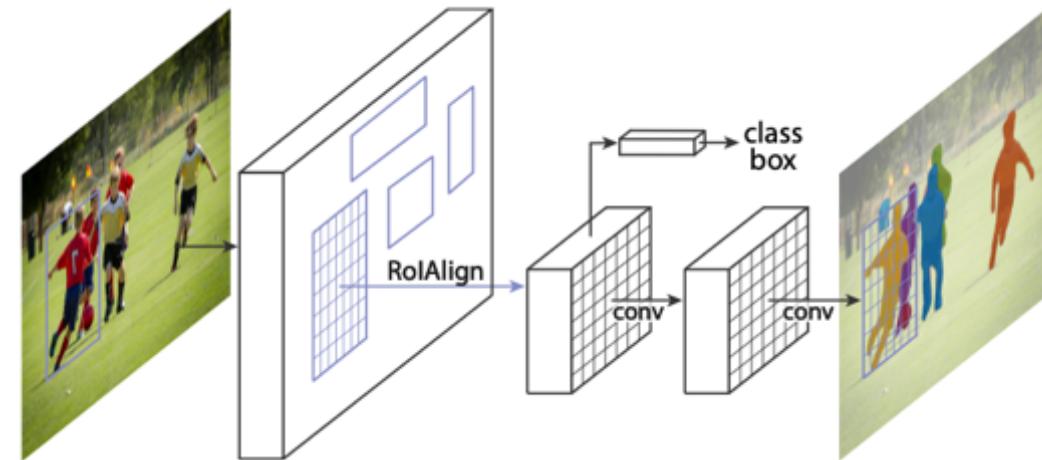


Region Proposal Networks R-CNN

2017 MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications
Andrew G. Howard et al (Google)

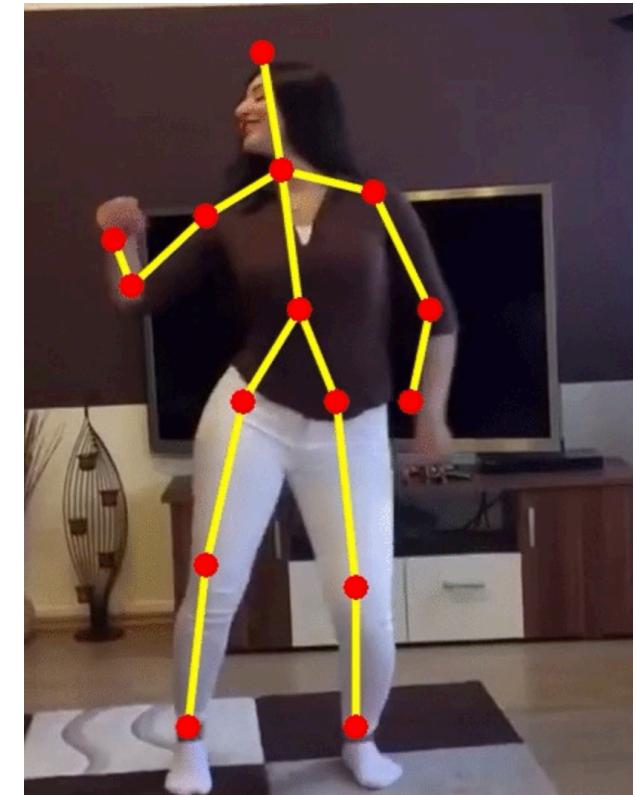
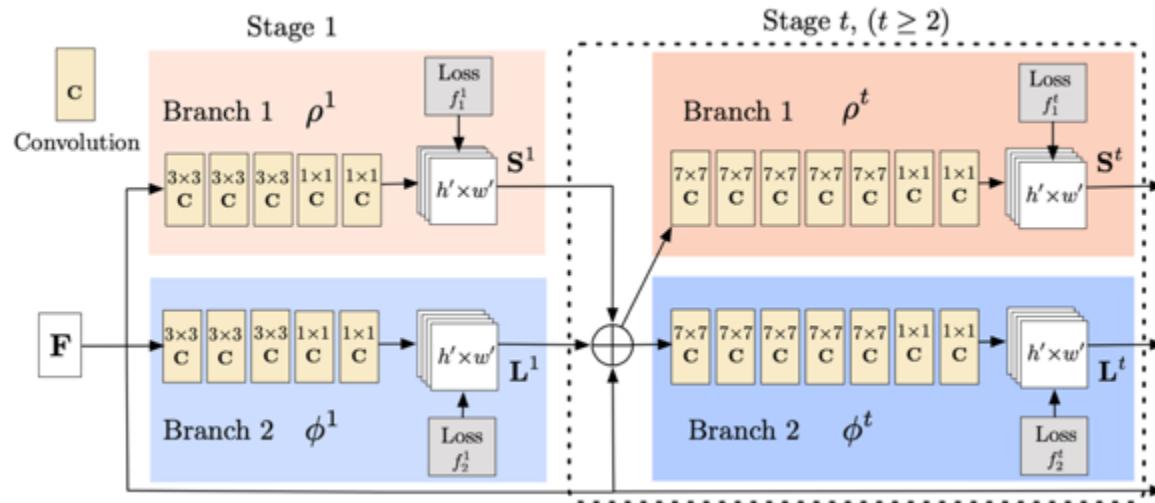
Framework Resolution	Model	mAP	Billion Mult-Adds	Million Parameters
SSD 300	deeplab-VGG	21.1%	34.9	33.1
	Inception V2	22.0%	3.8	13.7
	MobileNet	19.3%	1.2	6.8
Faster-RCNN 300	VGG	22.9%	64.3	138.5
	Inception V2	15.4%	118.2	13.3
	MobileNet	16.4%	25.2	6.1
Faster-RCNN 600	VGG	25.7%	149.6	138.5
	Inception V2	21.9%	129.6	13.3
	MobileNet	19.8%	30.5	6.1

2018 Mask R-CNN Kaiming He Georgia Gkioxari Piotr Dollár Ross Girshick
Facebook AI Research (FAIR) <https://arxiv.org/abs/1703.06870>

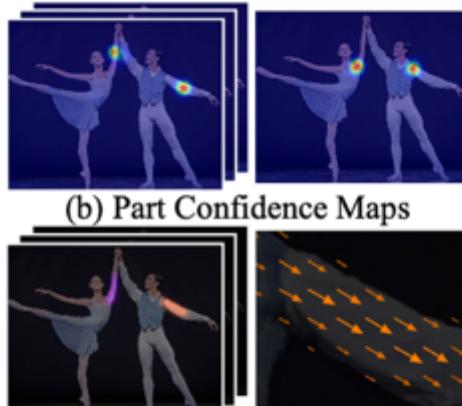


Detección de Pose. (2017)

Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields Zhe Cao Tomas Simon Shih-En Wei Yaser Sheikh The Robotics Institute, Carnegie Mellon University <https://arxiv.org/pdf/1611.08050.pdf>



F : Visual Goemetry Group VGG-19



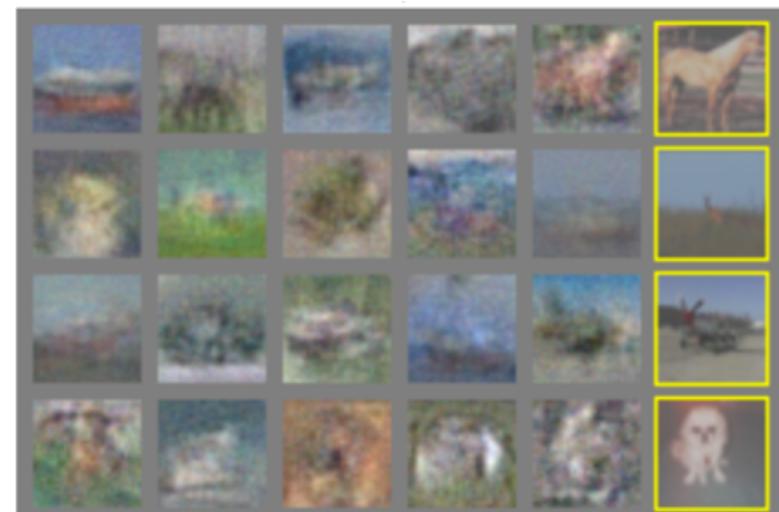
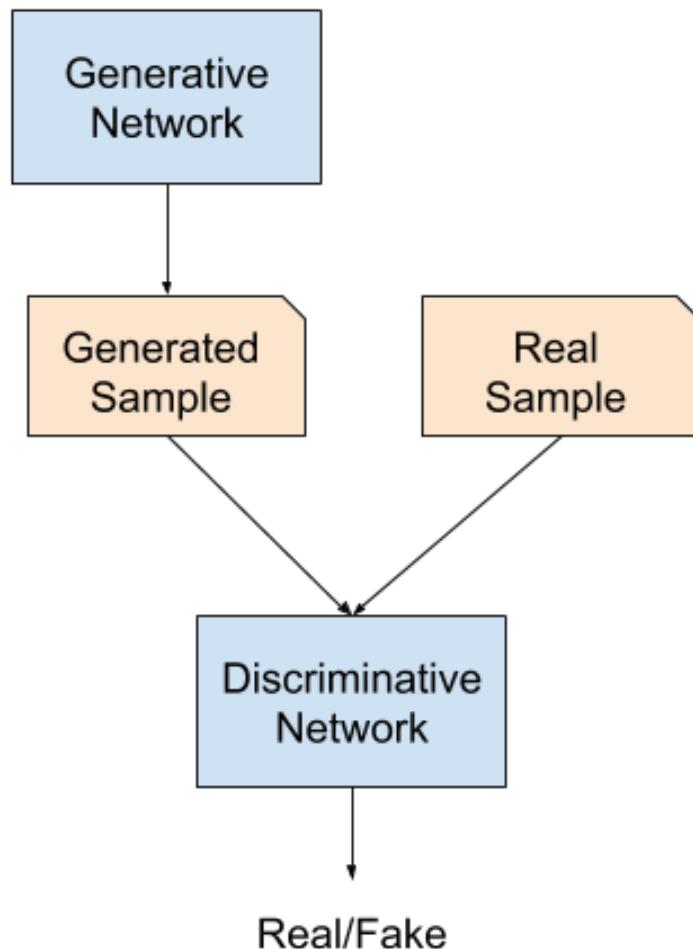


2014 – Generative Adversarial Networks

Ian J. Goodfellow, Jean Pouget-Abadie*, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair†, Aaron Courville, Yoshua Bengio‡

Département d'informatique et de recherche opérationnelle

Université de Montréal





2019 - A Style-Based Generator Architecture for Generative Adversarial Networks

Tero Karras Samuli Laine Timo Aila (NVIDIA) <https://arxiv.org/abs/1812.04948>

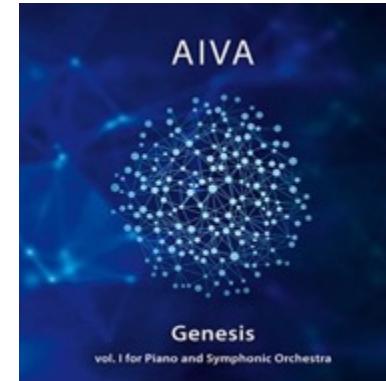
Source B aspectos de alto nivel, pose, estilo de cabello, forma de la cara, anteojos





Generative Adversarial Networks

2016 – AIVA primer máquina declarada oficialmente compositor



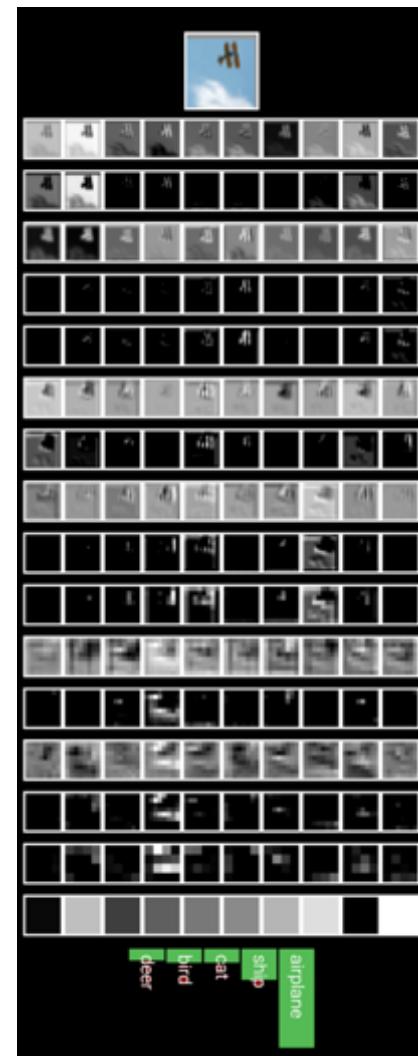
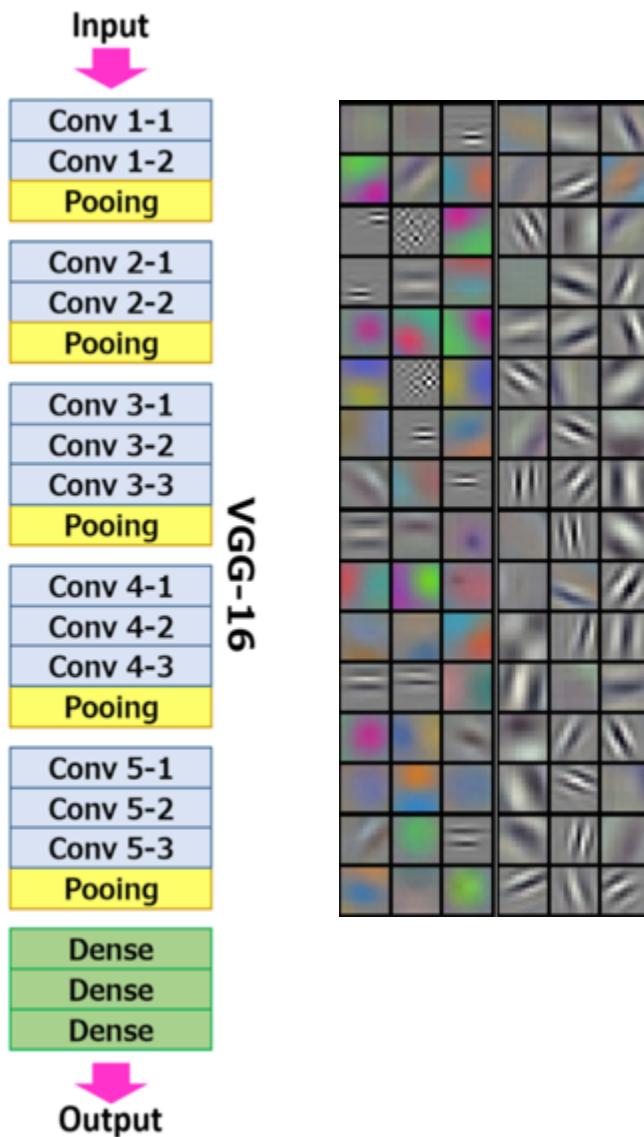
Fuente

Aplicaciones:

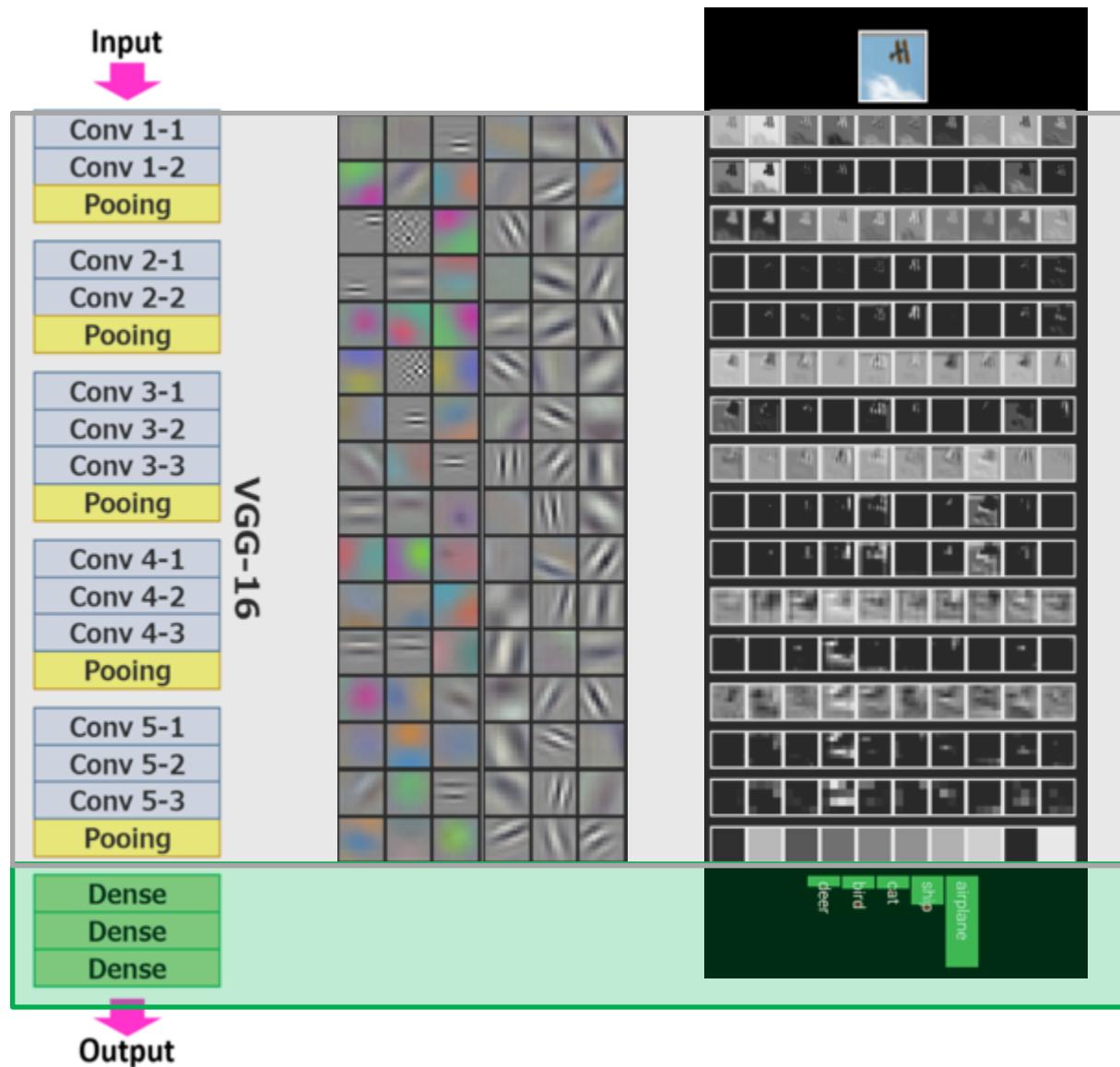
- Arte
- Habla
- Medicina
- Robótica
- Deepfake Detection



Nuestros Proyectos ¿Necesitan ese volumen de datos?



Transferencia del Aprendizaje



Se mantienen invariantes
los kernels o filtros que
determinan el feature map

Reentrena el clasificador

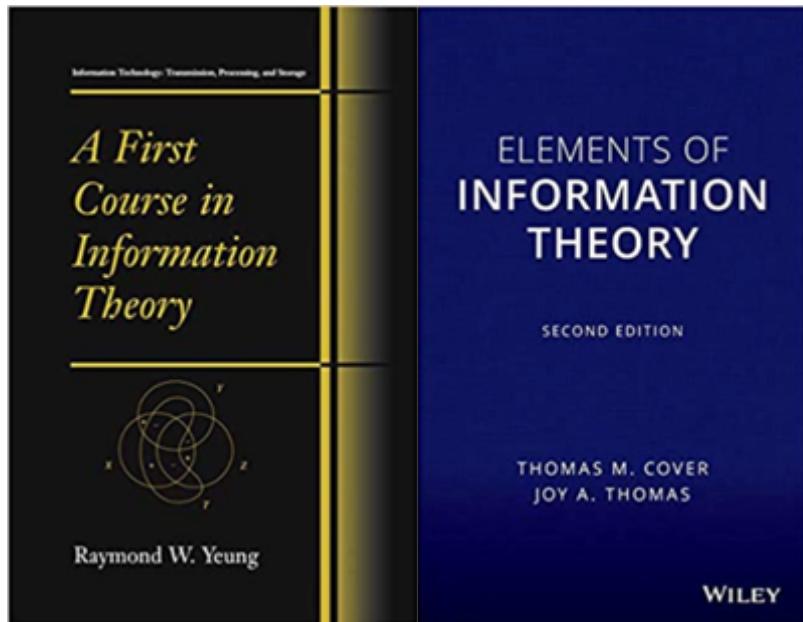
<https://pytorch.org>

<https://www.tensorflow.org/>

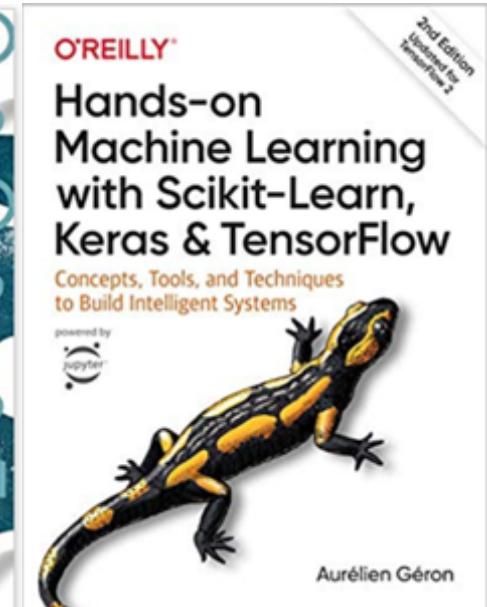
<https://keras.io>

<https://scikit-learn.org/>





<https://arxiv.org>



GRACIAS