

# ¡Hola!

Soy Marina Martínez-García

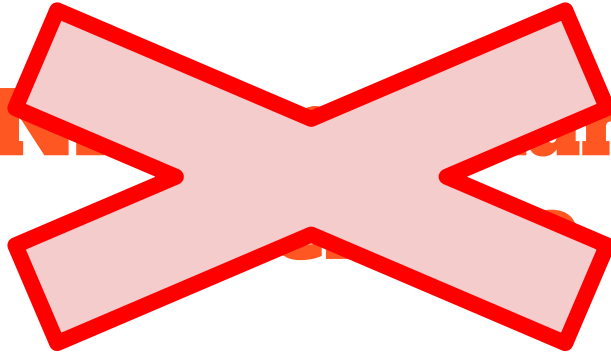
Professora del Dept. Matemàtiques Universitat  
Jaume I, Castelló, Espanya

[martigar@uji.es](mailto:martigar@uji.es)



# **De la Biología al Machine Learning y viceversa**

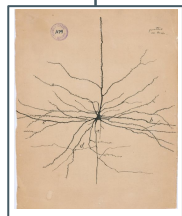
**Las ANNs se ven como un**



# **De la Biología al Machine Learning**

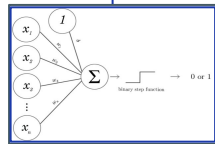
1906

Ramon y Cajal



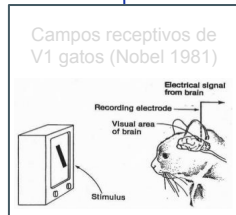
1943

Perceptron



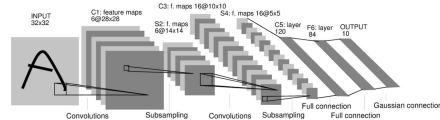
1959

H&W



1998

CNN



Lentes capaces de distinguir células

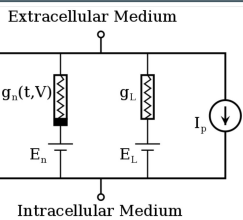
Óptica

1877

Lapicque: Modelo de neurona

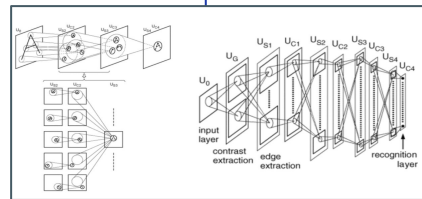
I&F

1907



H&H

1952



Neocognitron

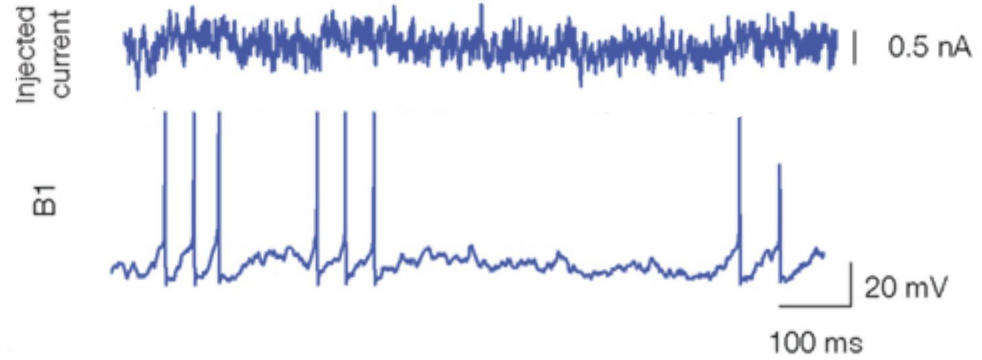
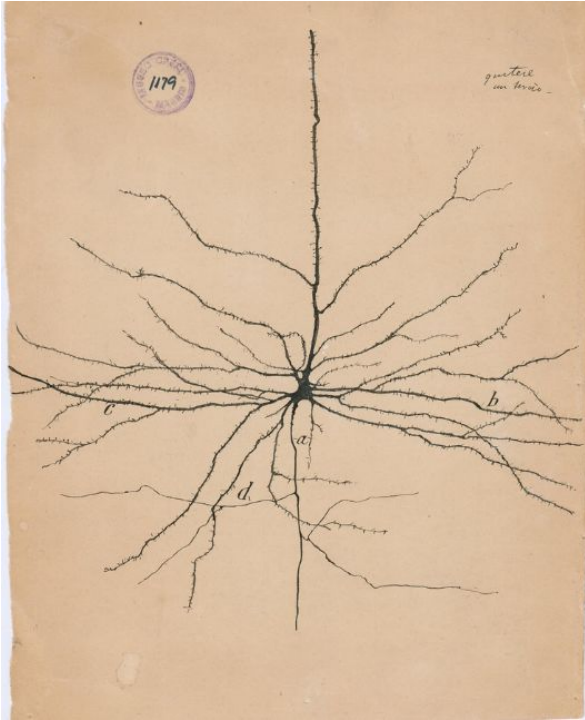
1980



J. Aniston

2005

# Nace una nueva ciencia



[ref](#)

## Historias cruzadas

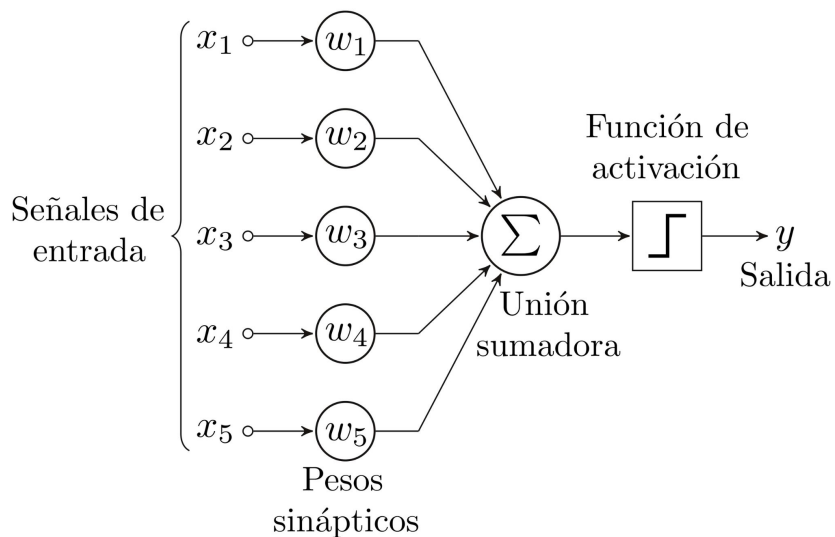
En 1907 Lapicque, se considera el padre de los primeros modelos [neurales](#) llamados *integrate-and-fire*: estos modelos "acumulan el voltaje" hasta que llega a un nivel (se considera que la neurona ha hecho un spike) y se restablece el voltaje inicial. Estos modelos se siguen usando a día de hoy. [Simulador](#)



1437 PARIS. — La Sorbonne, laboratoire de physiologie, M. Lapicque (Electricité)

# Historias cruzadas

En 1943, McCulloch, et al publican *A logical calculus of the ideas immanent in nervous activity*, El primer perceptrón: aprender a clasificar patrones (falta ref)





# Historias cruzadas

1952 A. Huxley y A. Hodgkin modelaron la neurona como circuito eléctrico complejo:

[Simulador](#)

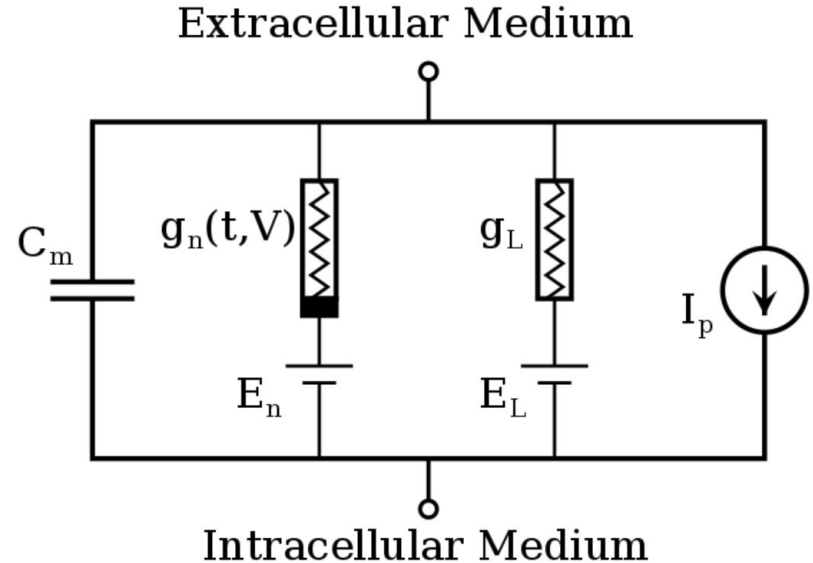
$$I = C_m \frac{dV_m}{dt} + \bar{g}_K n^4 (V_m - V_K) + \bar{g}_{Na} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l)$$

Donde,

$$\frac{dn}{dt} = \alpha_n(V_m)(1 - n) - \beta_n(V_m)n$$

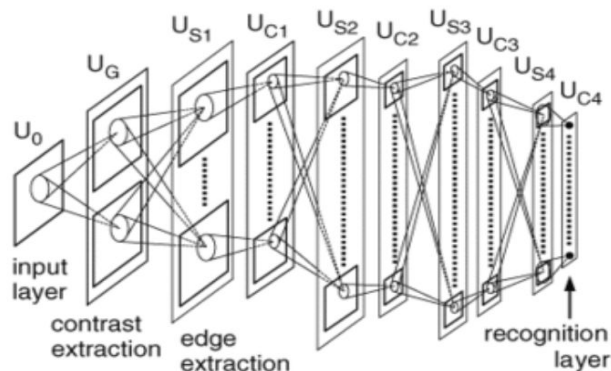
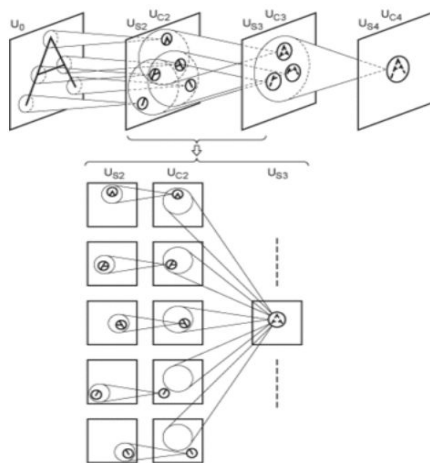
$$\frac{dm}{dt} = \alpha_m(V_m)(1 - m) - \beta_m(V_m)m$$

$$\frac{dh}{dt} = \alpha_h(V_m)(1 - h) - \beta_h(V_m)h$$



# Historias cruzadas

Ya en los 80 Fukushima, K., 1980. publicaron el [Neocognitron](#). Vemos que tanto los modelos biológicamente plausibles como las ANN van mejorando y aumentando su complejidad.



# Campos receptivos

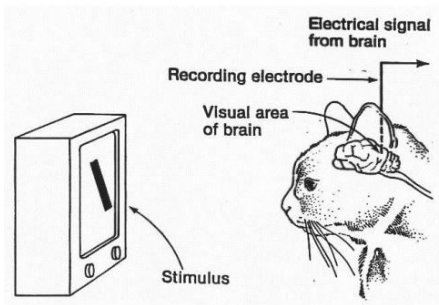
El campo receptivo de una neurona sensorial es la región particular del espacio sensorial (campo visual, sonido, superficie del cuerpo ...) en el que un estímulo provoca cambios en la tasa de disparo de la neurona. Se han identificado campos receptivos para las neuronas del sistema auditivo, el sistema somatosensorial, y el sistema visual.





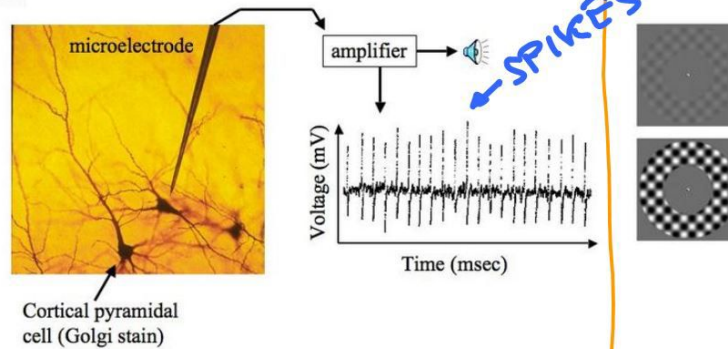


# Campos receptivos



RESPUESTA (DE UNA NEURONA)

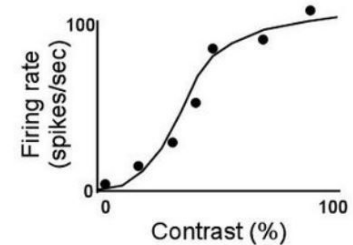
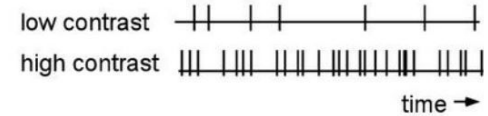
## Single-cell electrophysiology



Microscope pictures neuron with microelectrode

TASA

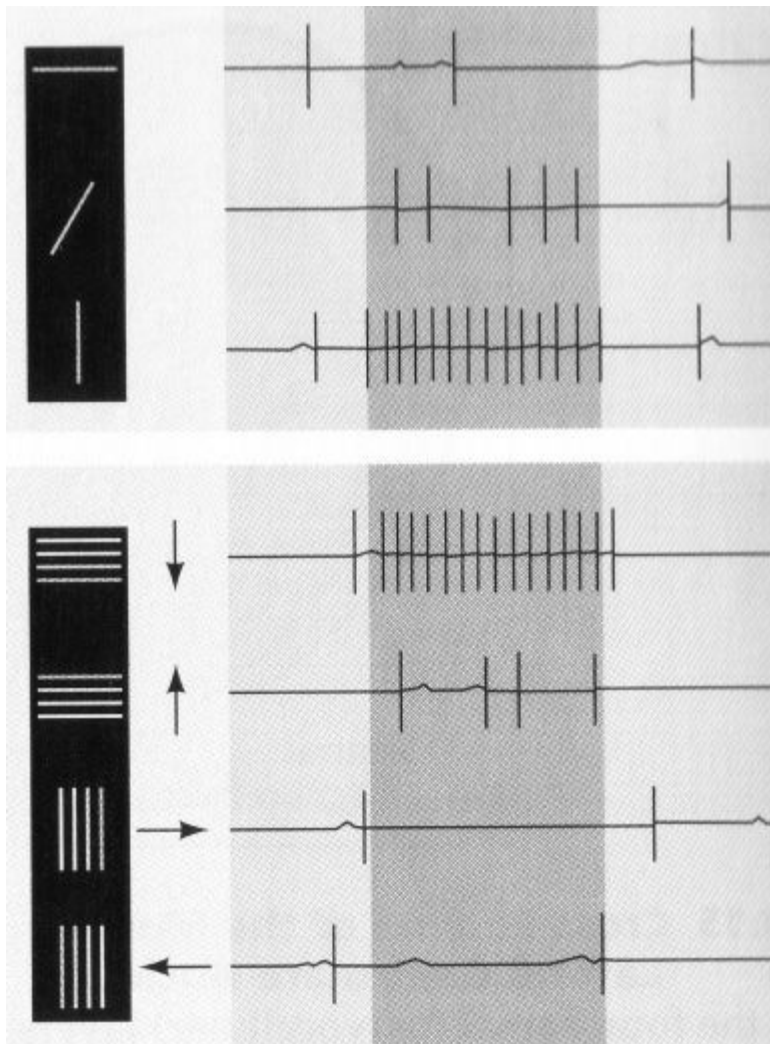
Firing rate: visual responses



Nota: ¿la función de respuesta que vemos se "parece" a la función de activación del perceptrón?

# Campos receptivos

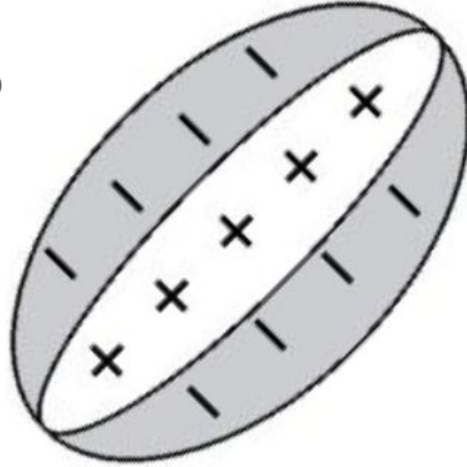
En los años 80 Daugman y Marcelja propusieron que los campos receptivos de las células del córtex visual primario (V1) podían ser modelados con un funciones Gabor-2D.





# Campos receptivos

- Localizados en el espacio
- Orientacion
- Ancho de banda



Oriented  
receptive field  
of a neuron in  
primary visual  
cortex (V1)



# Campos receptivos : Gabor

Siguiendo los pasos de Daugman y Marcelja, para modelar los campos receptivos de las neuronas se usa un coseno y una función gaussiana. Un Gabor es el producto de la dos

$$Gabor(x, y) = \begin{cases} \text{cosinus}(\lambda, \varphi) & \cos\left(2\pi\frac{x}{\lambda} + \varphi\right) \\ \text{gaussiana}(\gamma, \sigma) & \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right) \end{cases}$$

Resulta una función que depende de las siguientes variables  
( $x, y, \lambda, \varphi, \gamma, \sigma$ ).

# Campos receptivos

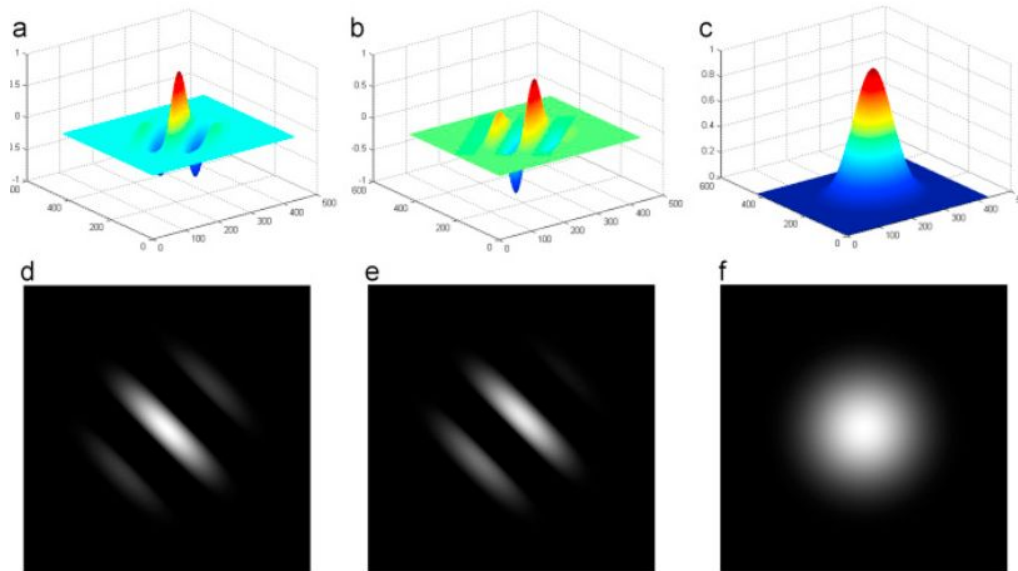
Los campos receptivos son sensibles a diferentes direcciones. Esto se modela mediante matrices de rotación,  $\sigma$  el ángulo de giro, entonces la matriz tiene la siguiente expresión:

$$\text{rotación : } \begin{cases} x' = x \cos(\theta) + y \sin(\theta) \\ y' = -x \sin(\theta) + y \cos(\theta) \end{cases}$$

Finalmente, juntaremos el filtro y la rotación, los Gabor tendrán la siguiente expresión:

$$G(\lambda, \theta, \gamma, \sigma, \varphi)(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cdot \cos\left(2\pi \frac{x'}{\lambda} + \varphi\right)$$

# Campos receptivos

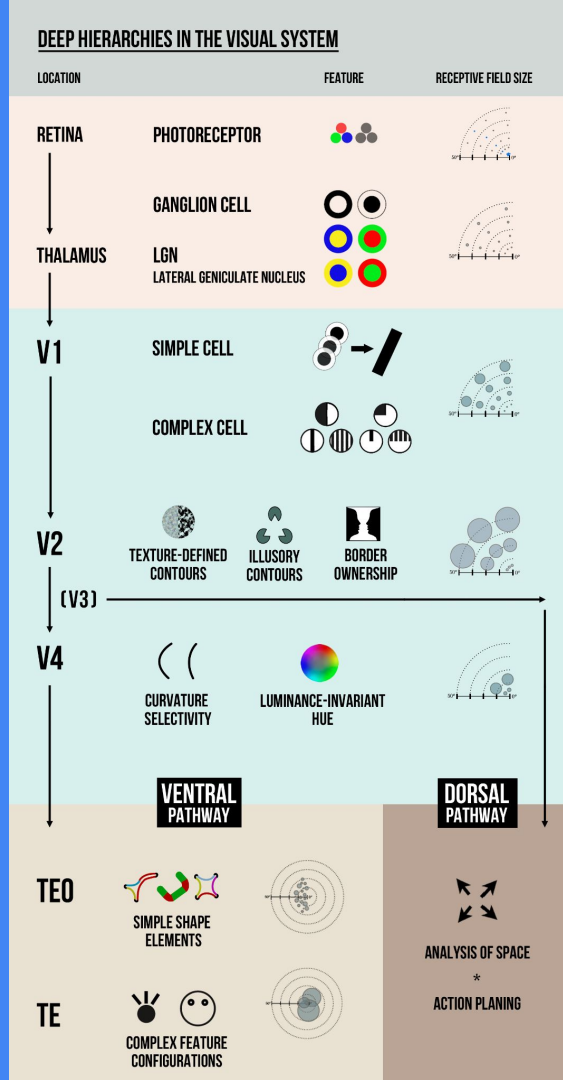


[Simulador](#)  
(sobre imágenes)

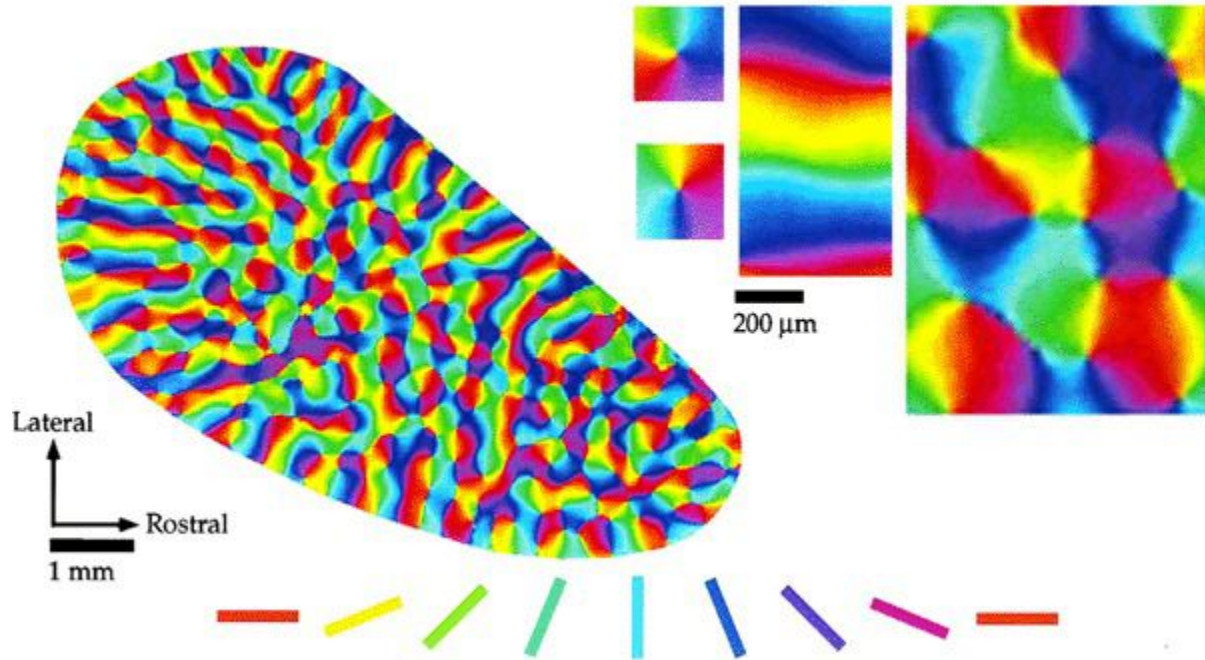
¿Y si en lugar de campos receptivos los llamamos filtros.....?

# Sistema visual

¿Cómo funciona sistema visual humano?



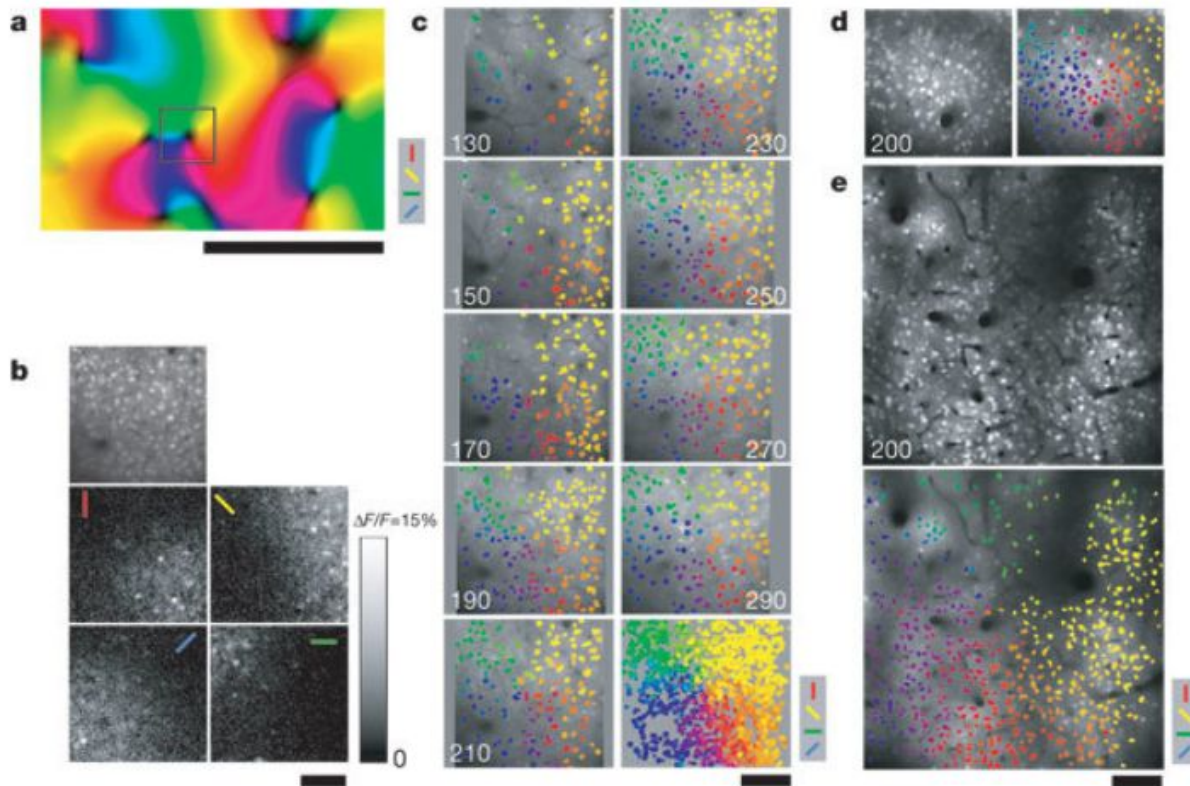
# Pinwheels



# Retinotopia

[Ohki et al](#)  
2006

[Más](#)

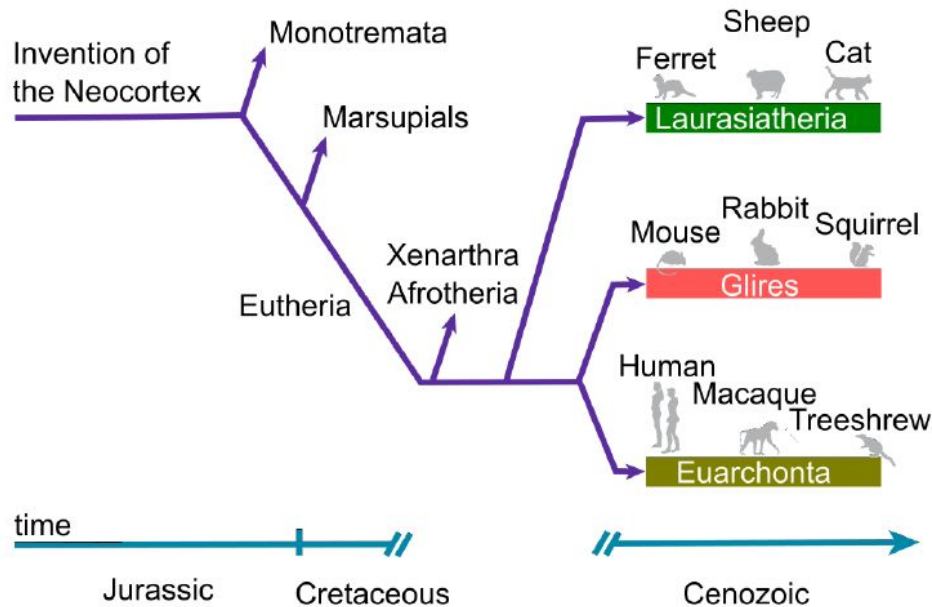


# ¿Por qué puede pasar? ¿Qué lo puede explicar?

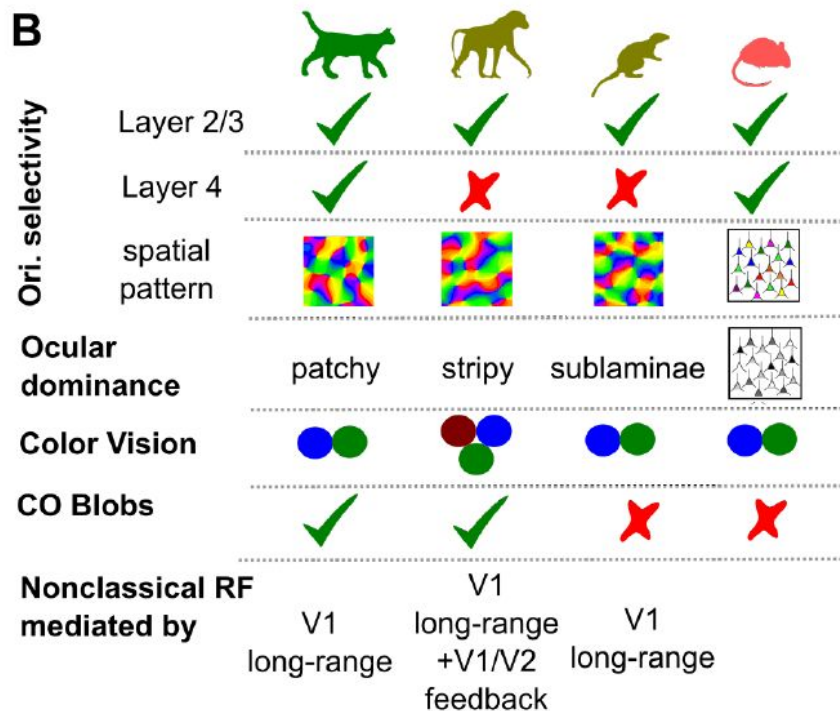
- Aproximaciones NO Funcionales:
  - Interleaved Lattices and Moire (D. Ringach)
  - Local interaction at LGN implies patterns in V1 (M. Kaschube)
  
- Aproximaciones Funcional (Machine learning):
  - Information theory, e.g. redundancy reduction (A. Hyvarinen)

# ¿Está cada especie optimizando funciones distintas?

**A**



**B**

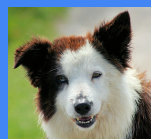




# Invarianzas

$G$  un grup (rotacions) actuant sobre el dataset  $X$ :

$F: X \rightarrow Y$  *invariant* si  $F(g \cdot x) = F(x) \quad \forall g \in G, x \in X$



Gos

# Invarianzas

Los humanos (y los animales en general) somos invariantes a muchos cambios en los estímulos. Pensando en imágenes, somos invariantes a :

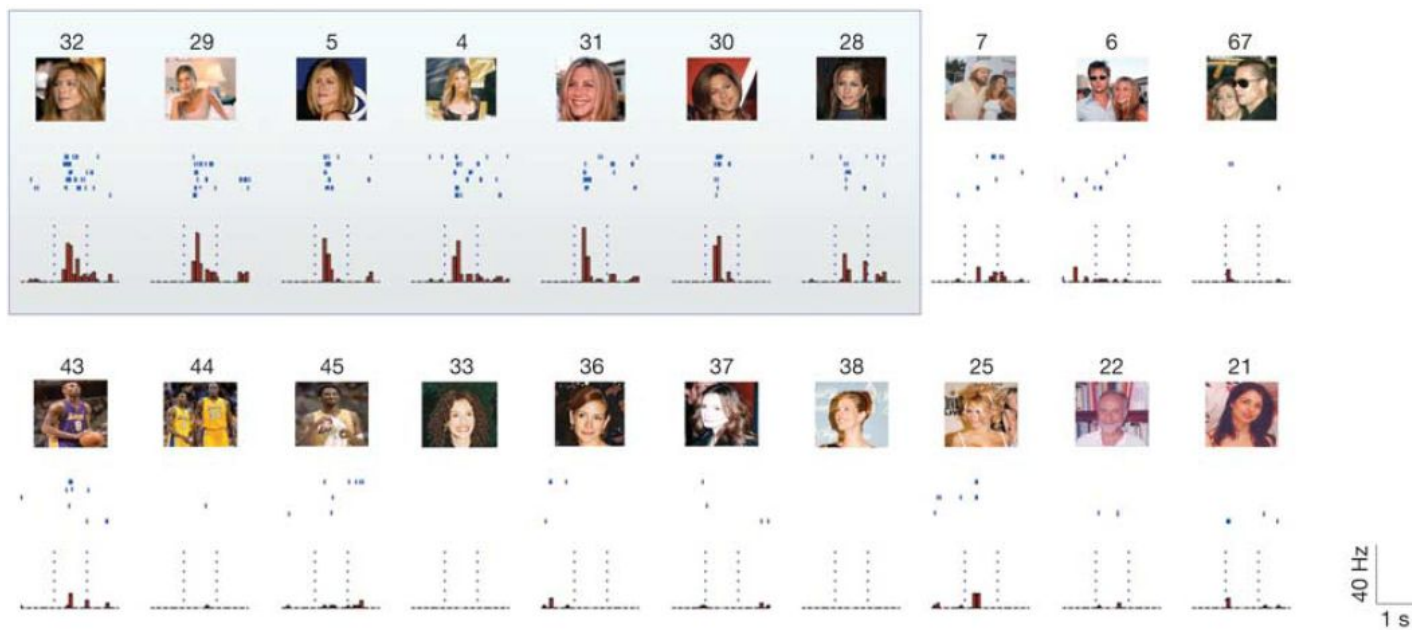
- Traslaciones,
- Cambio de escala,
- Rotaciones, iluminación
- Contraste .....

De alguna manera nuestros campos receptivos se combinan o adaptan para serlo. Si aprendemos de nuestras neuronas podremos obtener algoritmos invariantes a estos aspectos.

Sin usar el data augmentation.....

# Invarianzas

Hasta tal punto que en 2005 se vieron neuronas "invariantes" a Jennifer Aniston



# Invarianzas

Hay muchos motivos para ir en este sentido:

- Incrementar el conocimiento del cerebro: desde entender la enfermedades neurológicas a cómo mejorar las pantallas de TV .
- Mejorar el funcionamiento/entrenamiento de ML: porque cerebro no tienes uno de los grandes problemas de las ANN: Adversarial attacks: [Demo](#)

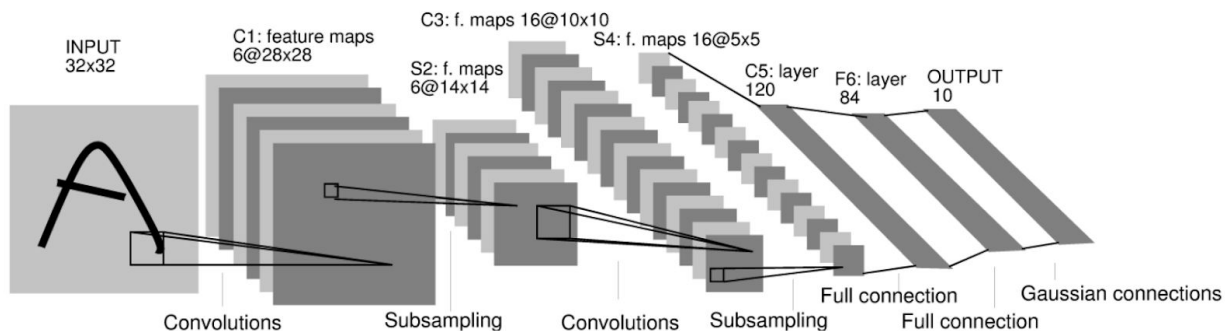
# A translación: ¡Convolución !

La convolución es una operación que nos puede ayudar con la invarianza a traslaciones. Se puede mostrar que el operador de convolución conmuta con respecto a la translación. [Ref](#)

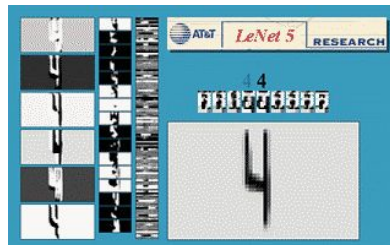
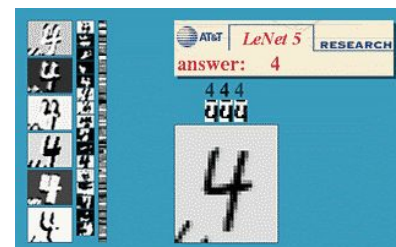
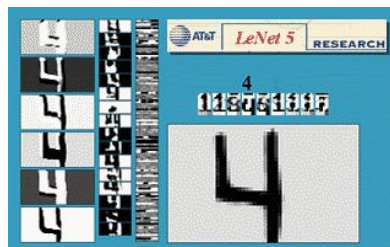
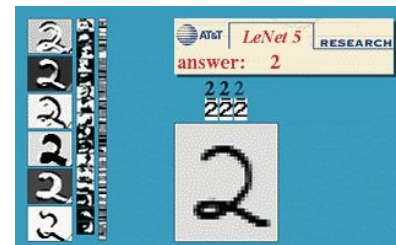
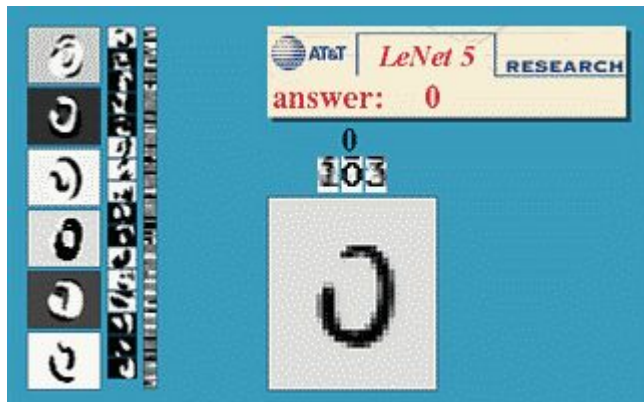
# **Del Machine Learning a la Biología**

# Inv a transacciones Convolutional NN

Estos trucos (y muchos otros como el backpropagation) los usaron a LeCun and cia. entre 1989-1998. Llegando a generar la LeNet5, que dio un gran salto en cuanto a los resultados de clasificación y reconocimiento de dígitos manuscritos dando lugar a las CNN



# CNN Ref





# Filtros CNN

¿Cómo son los filtros resultantes de entrenar estas redes?

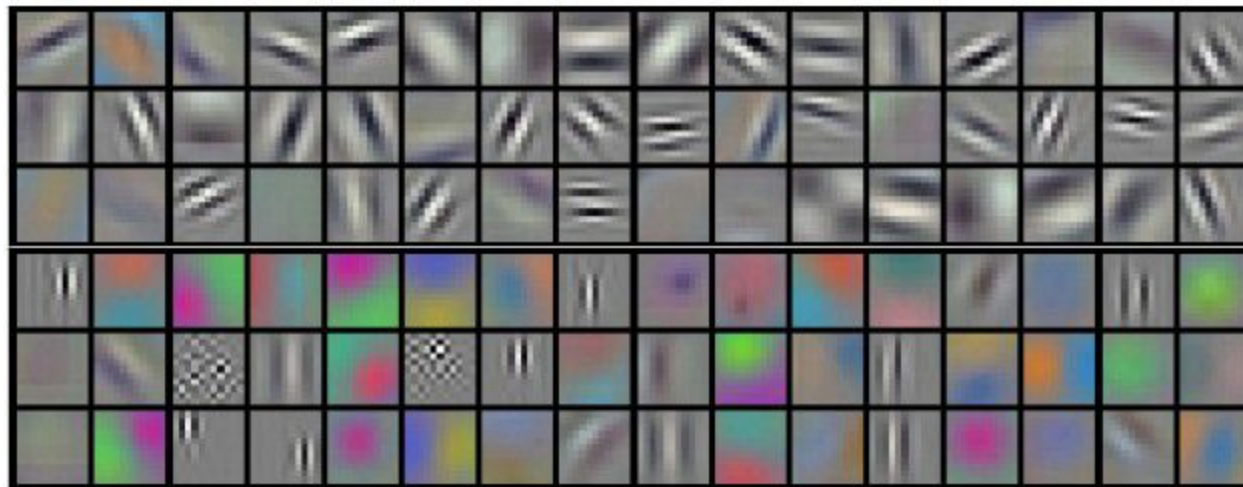


Figura: Example lters learned by Krizhevsky et al. Each of the 96 filters shown here is of size  $[11 \times 11 \times 3]$ , and each one is shared by the  $55 \times 55$  neurons in one depth slice. [Ref](#)

# Filtros CNN: colores opuestos



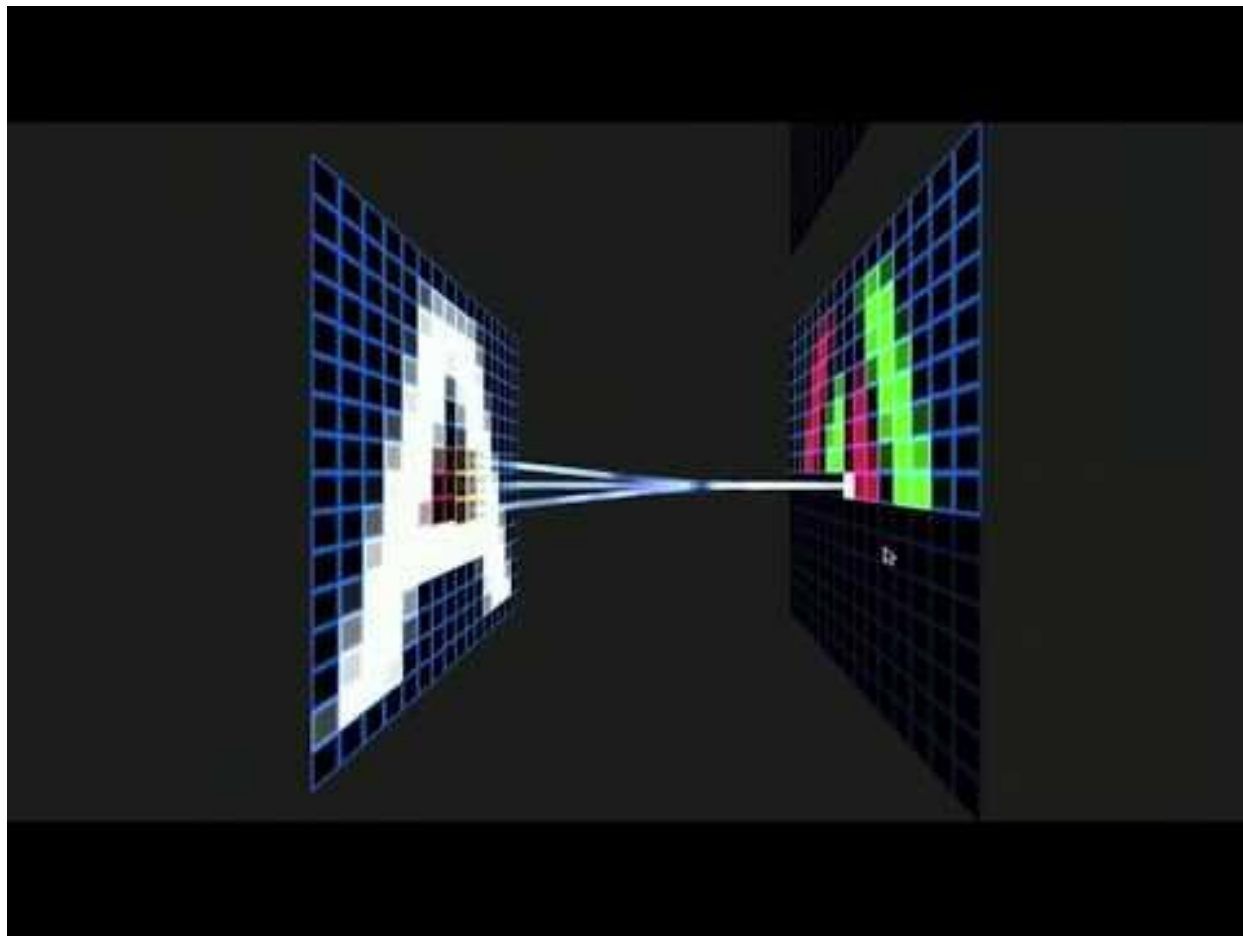
	
verde	rojo
	
verde amarillo	rojo púrpura
	
amarillo	púrpura o morado
	
ámbar	violeta
	
naranja	azul
	
bermellón	azul verde

# Filtros CNN

Se parecen mucho a los filtros de Gabor, ¿no?

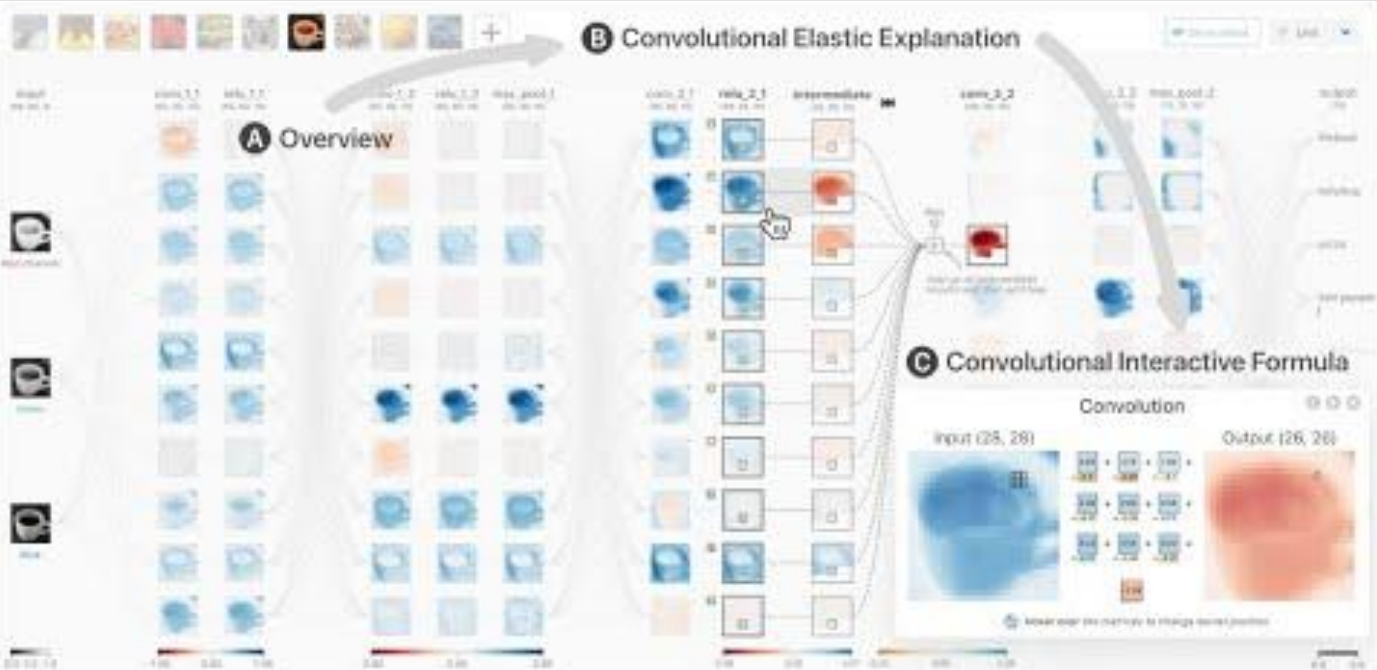
¿Qué implica?

¿Son estos filtros la solución óptima para el algoritmo de backpropagation como lo son para el cerebro? ¿Son solución al mismo problema o son problemas distintos?



# CNN EXPLAINER

Learn Convolutional Neural Network (CNN) in your browser!



Type: Perceptron  
Data Set: MNIST  
Hidden Neurons: 2000  
Synapses: 1191000  
Synapses shown: 2%  
Learning: WCor

0 1 2 3 4 5 6 7 8 9



[www.cybercontrols.org](http://www.cybercontrols.org)

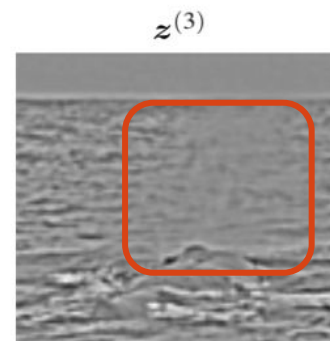
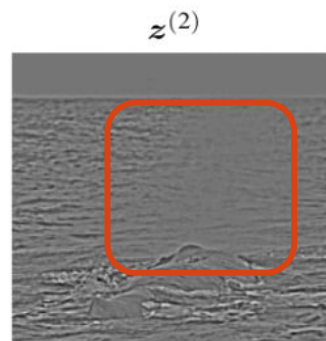
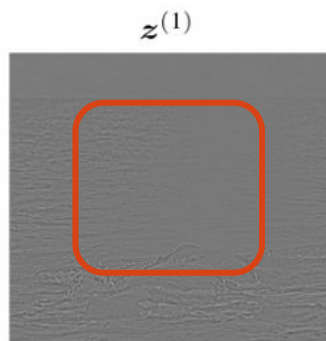
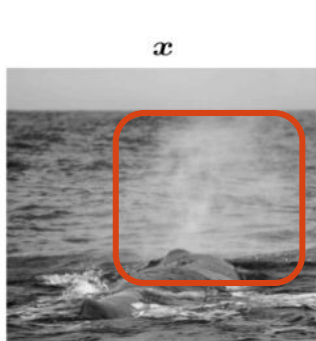
# Bio inspiración: Normalización Divisiva

La normalización divisiva fue motivada originalmente por las propiedades observadas del sistema visual, donde se utilizó para explicar las no linealidades en las respuestas de las neuronas. Estas neuronas se adaptan al entorno cercano:

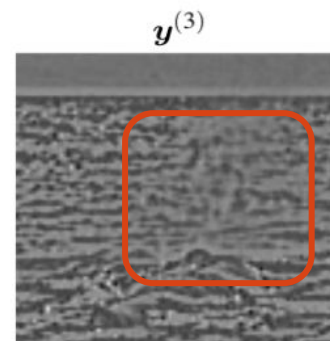
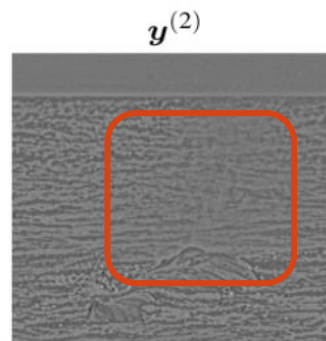
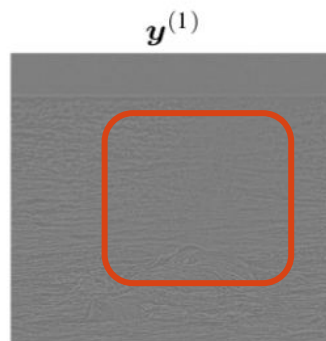
$$y = \frac{z}{\beta + \gamma z^\alpha}$$

Actualmente tiene muchas aplicaciones en ML: siendo una función de activación común para las ANN, su uso para la compresión de imágenes o para generar invarianzas respecto al contraste de una imagen.

# Normalización Divisiva



$$y_i = z_i / f_c$$

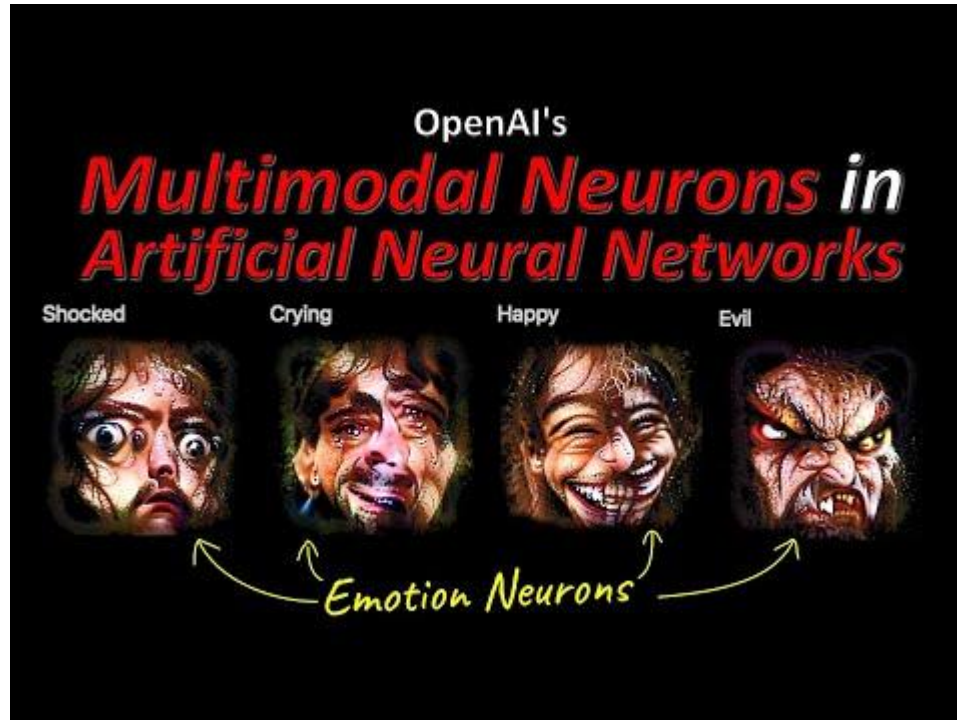




# Jpg vs DN



# Multimodal Neurons: como las Jennifer Aniston cells



[Ref](#)

# Capsule neural network

Algunos intentos no muy fructíferos

Sí construyendo ANN invariantes a la traslación/ contraste hemos dado este paso de gigante,

¿Qué pasaría si  
consiguieses mos ANN  
invariantes a otros  
aspectos?

(como el cerebro humano lo es)

Vamos a ver algunos ejemplos de cómo el ML nos ha ayudado a "entender" un poco más cómo puede funcionar el cerebro, como/ porque la naturaleza ha llegado al cerebro "óptimo".

# XAI

Para ello necesitamos un ML "entendible" o un o como se dice ahora "XAI" ("Explainable Artificial Intelligence"). En estos momentos hay fuerte corriente para "ver" cómo funciona un ANN por dentro y entender el porqué de ello.

**Black Box**



**Gray Box**



**Glass Box**



# Representación de imágenes

Muchos autores se han preguntado cómo podemos descomponer imágenes  $I(x, y)$  en unas funciones básicas para usarlas como filtros de las ANN:

$$I(x, y) = \sum_i a \phi_i(x, y)$$

Y cómo estas funciones se relacionan con la estadística de las imágenes naturales. La hipótesis de Barlow dice que los componentes del código han de ser estadísticamente independiente, Esto tiene muchas implicaciones en el ML.

# Representación de imágenes

En este artículo (Nature 1996), sus autores propusieron un método de ML para "aprender" estas siguiendo la hipótesis de Barlow.

## **Emergence of simple-cell receptive field properties by learning a sparse code for natural images**

**Bruno A. Olshausen\* & David J. Field**

Department of Psychology, Uris Hall, Cornell University, Ithaca  
New York 14853, USA

¡¡ No supervisado !!



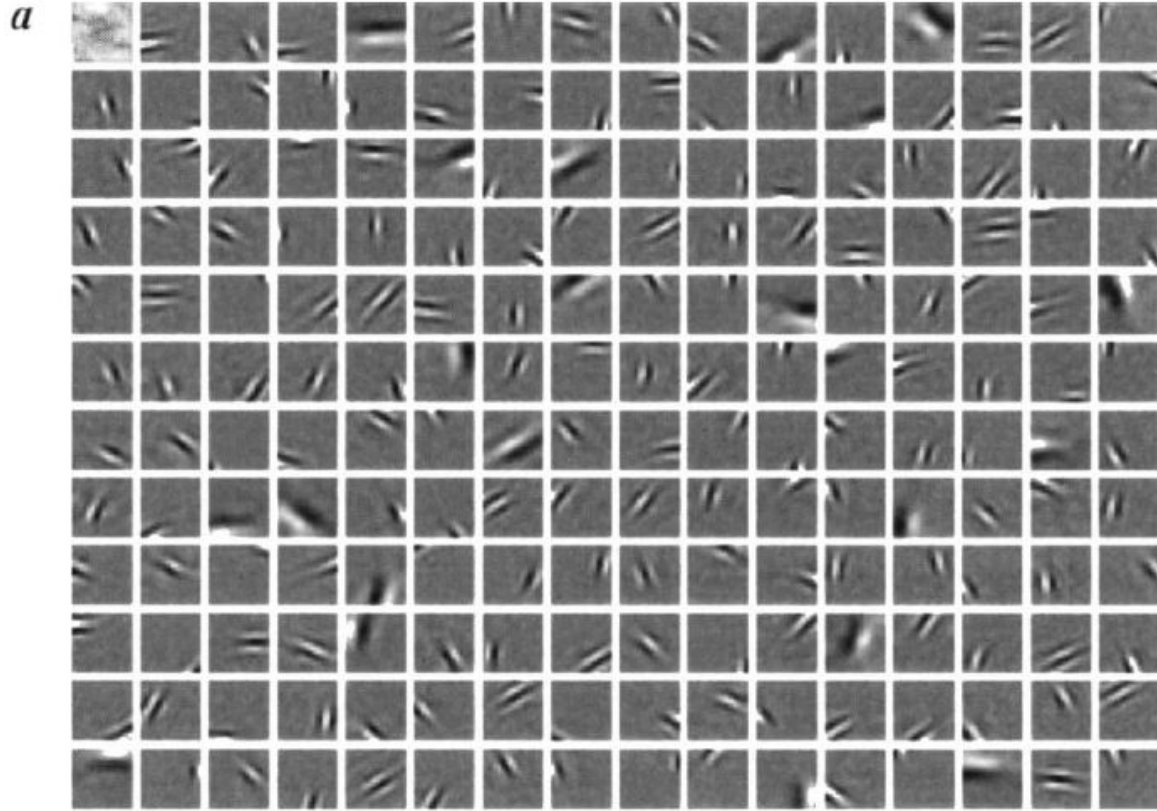


Figure: 192 funciones básicas después de 4000 iteraciones del algoritmo de minimización donde las entradas son imágenes naturales

# Representación de imágenes: ICA (PCA de orden superior)

A partir del 2000 Hyvärinen (y colaboradores) empezaron a estudiar el método [Independent Component Analysis](#) y aplicarlo a [imágenes](#).

Este método consiste en transformar linealmente ( $W$ ) los datos observados  $x$  como

$$s = Wx,$$

tal que  $s$  es un vector de componentes estadísticamente independientes. Esta "independencia" se "mide" con alguna función de contenido de información, como la Mutual Information.

Con técnicas de ML hemos visto cómo los gabor se obtienen va un proceso iterativo de aprendizaje. ¿Que quiere decir esto? ¿Podemos pensar en el cerebro como una ANN que está minimizado una función objetivo?

¿El cerebro humano ha evolucionado para minimizar el esfuerzo en codificar la información visual?

En ML nos puede ayudar a explicar un posible porqué al funcionamiento del cerebro

# Textura- forma

Otro ejemplo, mejorando las ANN  
haciéndolas más humanas.

La especie humana se “fija” más en  
la forma que en la textura de los  
objetos.

¿Y los sistemas actuales?

¿Para qué y porqué es útil ser más  
“humano” en este caso?

---

# Texture - shape Geirhos 2019

*shape hypothesis* LeCun et al. (2015): Intermediate CNN layers recognise “parts of familiar objects, and subsequent layers [...] detect objects as combinations of these parts” (p. 436).

*texture hypothesis*: in contrast to the common assumption, object textures are more important than global object shapes for CNN object recognition.

# Texture - shape Geirhos 2019



(a) Texture image

81.4% **Indian elephant**  
10.3% indri  
8.2% black swan



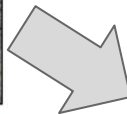
(b) Content image

71.1% **tabby cat**  
17.3% grey fox  
3.3% Siamese cat



(c) Texture-shape cue conflict

63.9% **Indian elephant**  
26.4% indri  
9.6% black swan



- Using style transfer
- Human (97) vs CNN
- Sigue siendo una gato para los sujetos

Remarkably, networks with a higher shape bias are inherently more robust to many different image distortions and reach higher performance on classification and object recognition tasks. Los seres humanos somos “más” invariantes y más robustos....

# Texture- shape Geirhos 2019



(a) Texture image  
81.4% **Indian elephant**  
10.3% indri  
8.2% black swan



(b) Content image  
71.1% **tabby cat**  
17.3% grey fox  
3.3% Siamese cat



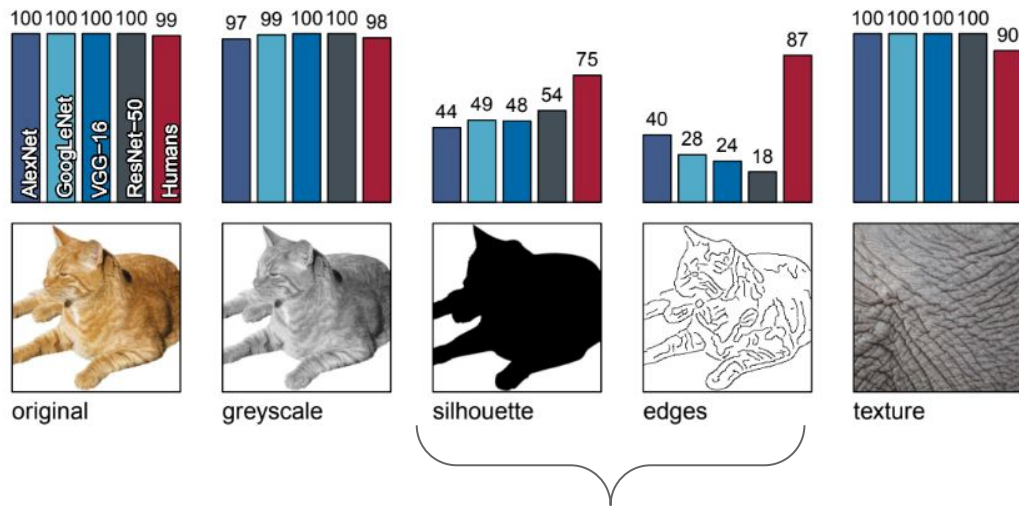
(c) Texture-image  
63.9% **Indian elephant**  
26.4% indri  
9.6% black swan

- Using style transfer
- Human (97) vs CNN

- Sigue siendo una gato para los sujetos

Remarkably, networks with a higher shape bias are inherently more robust to many different image distortions and reach higher performance on classification and object recognition tasks. Los seres humanos somos “más” invariantes y más robustos....

# Texture - shape Geirhos 2019



No hay textura:  
problema de entrenamiento? sacamos  
martillo de Tor y le metemos más y más  
en training...¿?



# **Búsqueda Bibliográfica**

# En que me fijo

**Partial success in closing the gap between  
human and machine vision**

Robert Geirhos<sup>1,2§</sup>

Kantharaju Narayanappa<sup>1</sup>

Benjamin Mitzkus<sup>1</sup>

Tizian Thieringer<sup>1</sup>

Matthias Bethge<sup>1\*</sup>

Felix A. Wichmann<sup>1\*</sup>

Wieland Brendel<sup>1\*</sup>

<sup>1</sup>University of Tübingen

<sup>2</sup>International Max Planck Research School for Intelligent Systems

\*Joint senior authors

<sup>§</sup>To whom correspondence should be addressed: robert.geirhos@uni-tuebingen.de

- Título: sin preguntas, key words en él
- ¿Quién?
  - Mira la primera y última posición
  - Al final en un tema l@s autor@s se repiten
- ¿De donde son? Qué uni? o centro
  - no es lo mismo un sitio que otro
- Revista? está indexada? o es arxiv? o es un congreso?
  - Hay congresos muy buenos
  - Ojo hay corriente que publica solo en arxiv
  - Mejor si son revistas INDEXADAS

# Como buscar

- Key word
- <https://typeset.io/>
- <https://chatdoc.com/>
- <https://researchrabbitapp.com/home>
- <https://www.explainpaper.com/>

**link1**

**¡Gracias!**