

Tema 5.1

Contexto de las redes convolucionales

Deep Learning

Máster Oficial en Ingeniería Informática

Universidad de Sevilla

Contenido

- Contexto
- Un poco de historia
- Uso de redes convolucionales

Contexto

- **Deep learning** = aprender representaciones jerárquicamente
- Es **profundo** si tiene más de una etapa de transformación de características no lineal
- Modelos con 2 capas no son Deep (**Shallow**)
- Red neuronal profunda: más de 1 capa oculta.

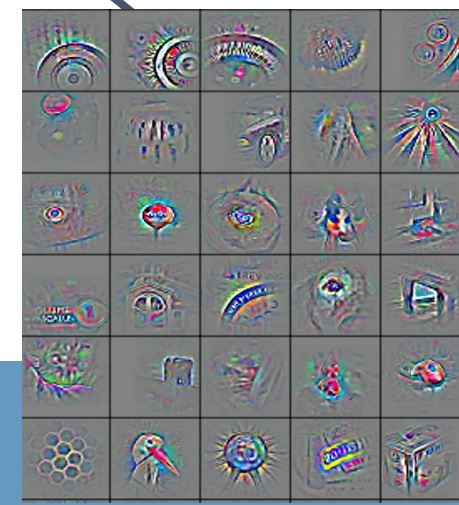
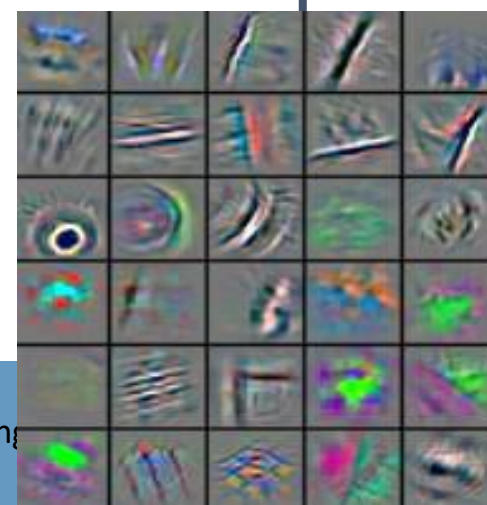


Low-level
feature

Mid-level
feature

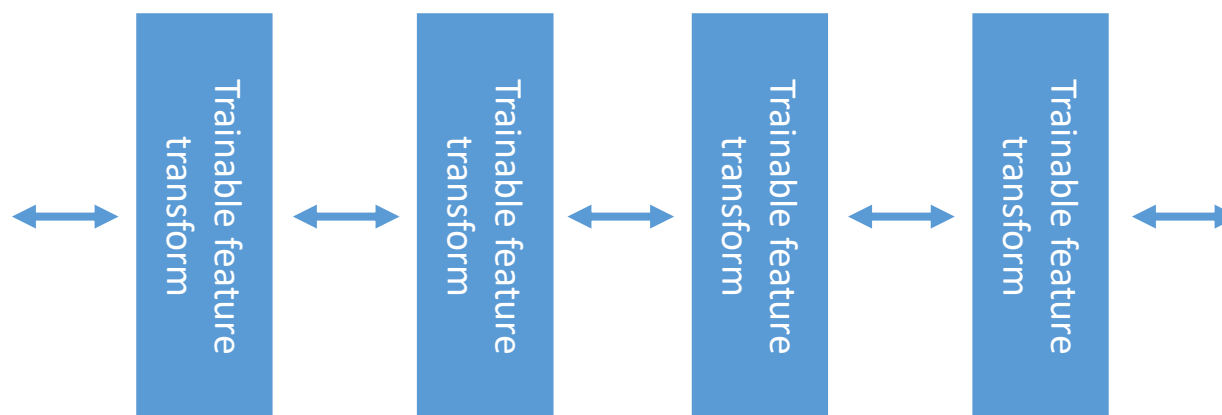
High-level
feature

Trainable
classifier



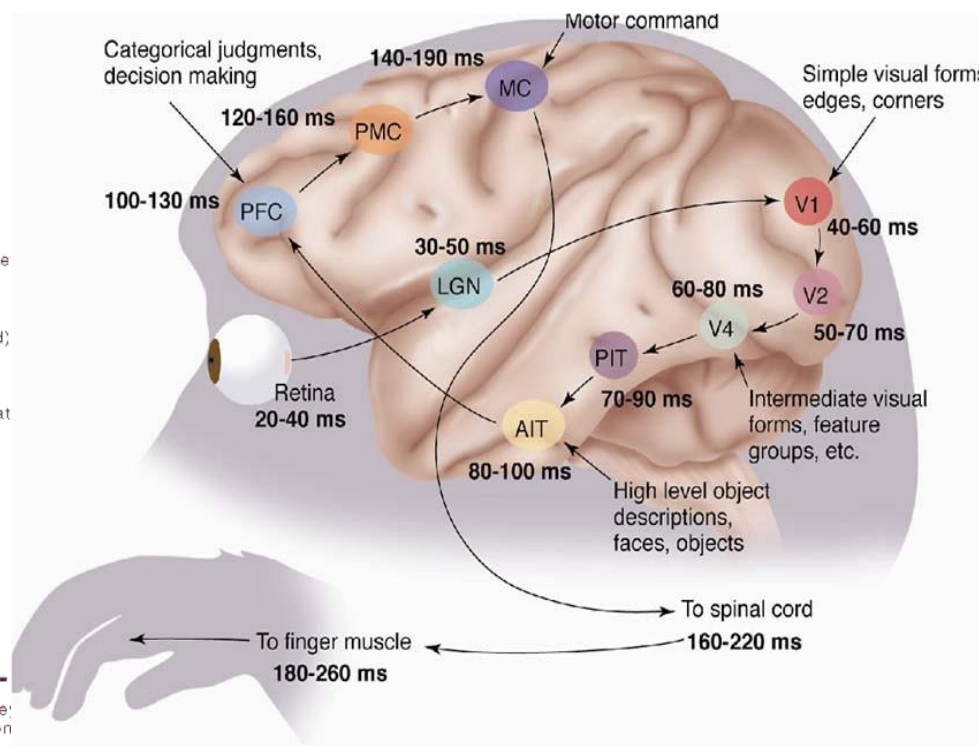
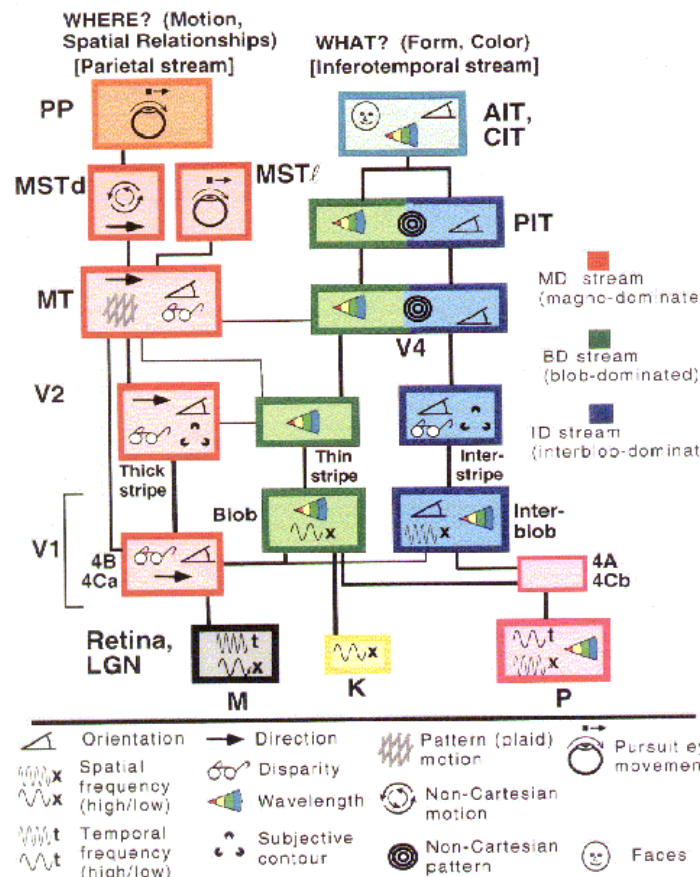
Contexto

- **Jerarquía de representaciones** con nivel creciente de abstracción: cada etapa es un tipo de transformación de características entrenable.
- Reconocimiento de imágenes:
 - Pixel → borde → textura → detalles → parte → objeto
- Texto:
 - Carácter → palabra → grupo de palabras → cláusula → frase → historia
- Habla:
 - Ejemplo → banda espectral → sonido → ... → fonética → palabra



Contexto

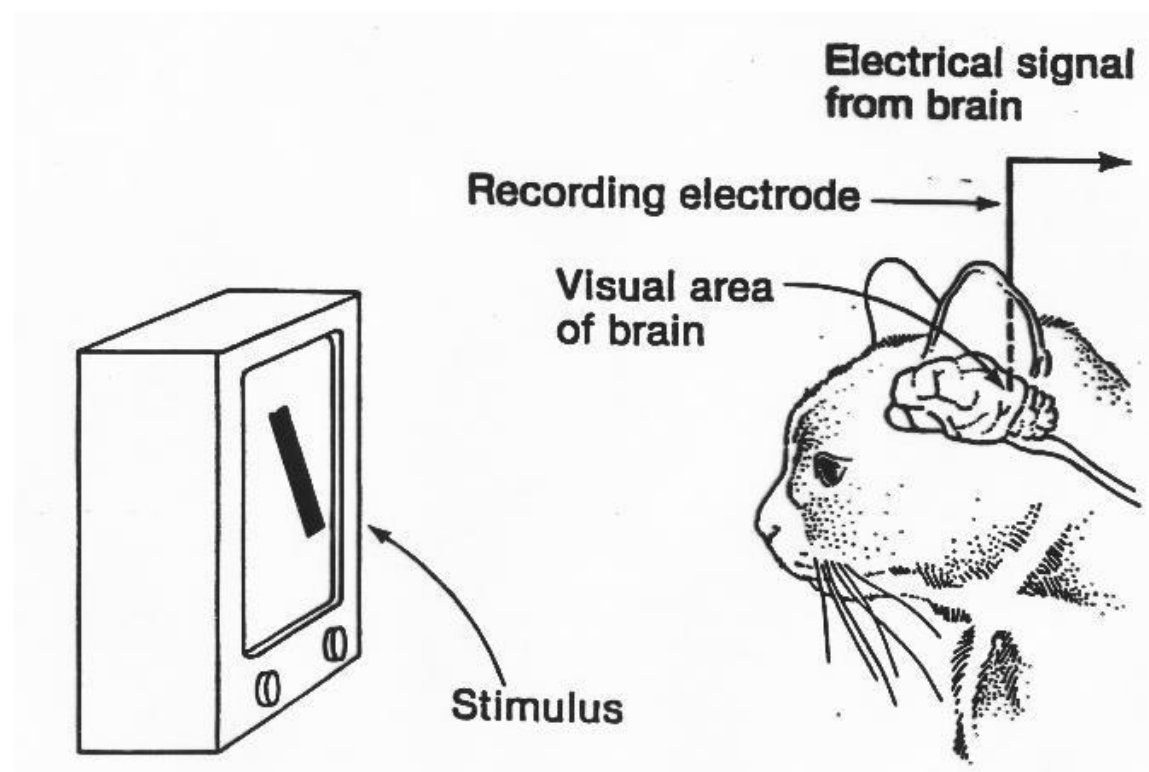
- Inspiración: el **córtex visual** de los mamíferos es jerárquico
- La vía ventral visual en el córtex visual tiene múltiples etapas, con múltiples representaciones: Retina - LGN - V1 - V2 - V4 - PIT - AIT



[Gallant & Van Essen]

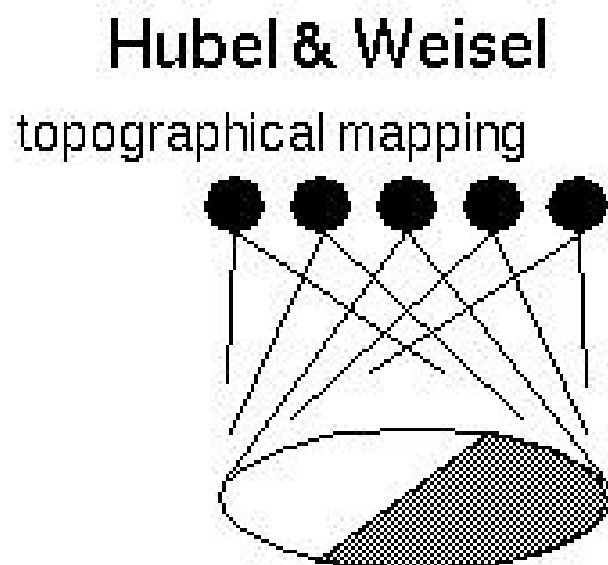
Un poco de historia

- El experimento [Hubel & Wiesel](#) 1959, 1962, 1968

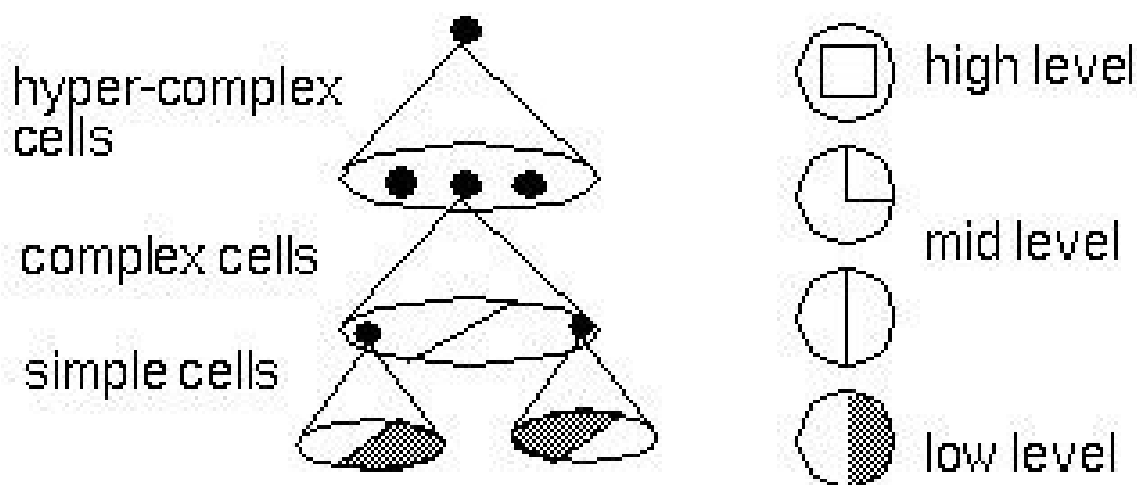


Un poco de historia

- El experimento [Hubel & Wiesel](#) 1959, 1962, 1968

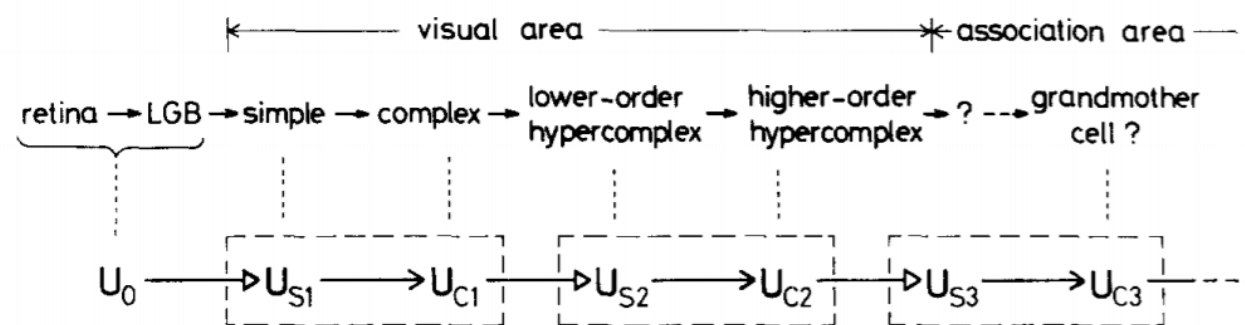
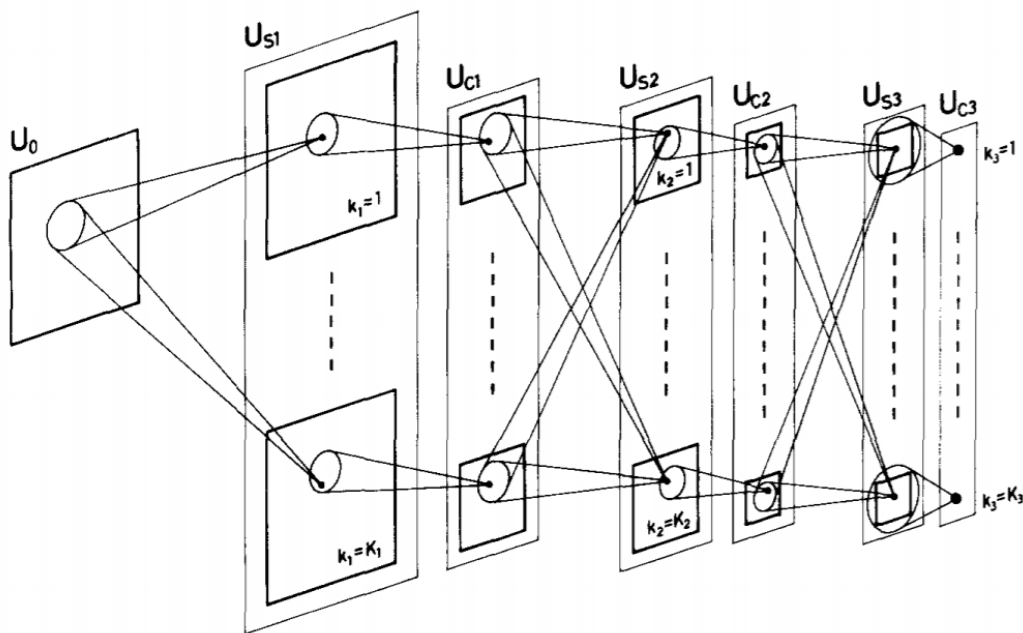


featural hierarchy



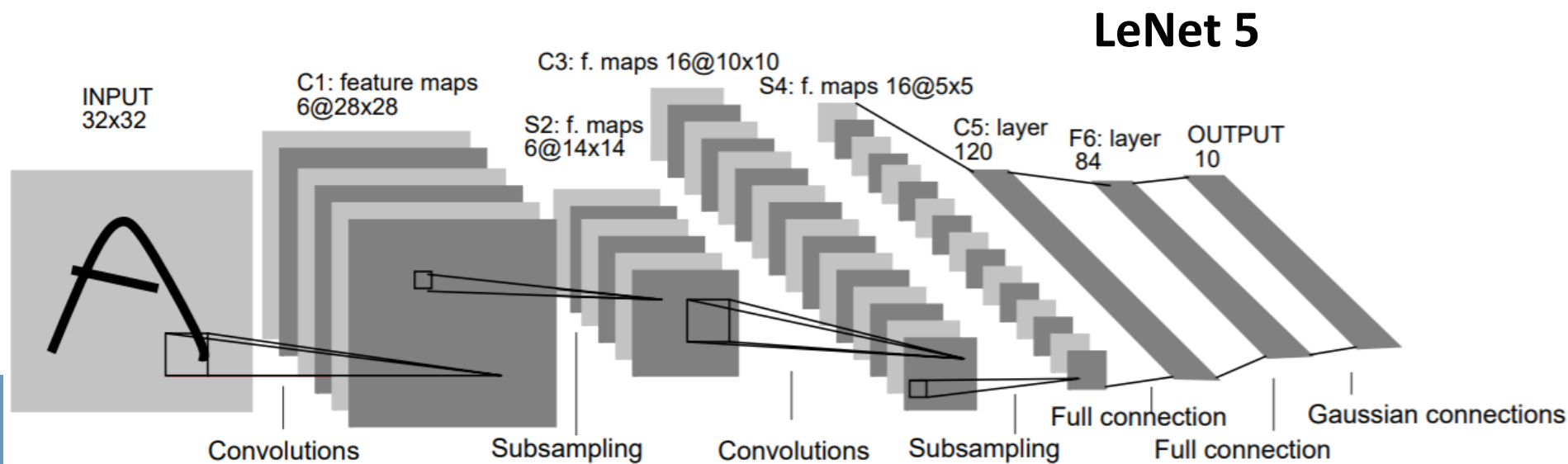
Un poco de historia

- Primera implementación: **Neocognitron** [[Fukushima 1980](#)]
 - Capas de células simples (parámetros modificables) y capas de células complejas (pooling, no modificable).



Un poco de historia

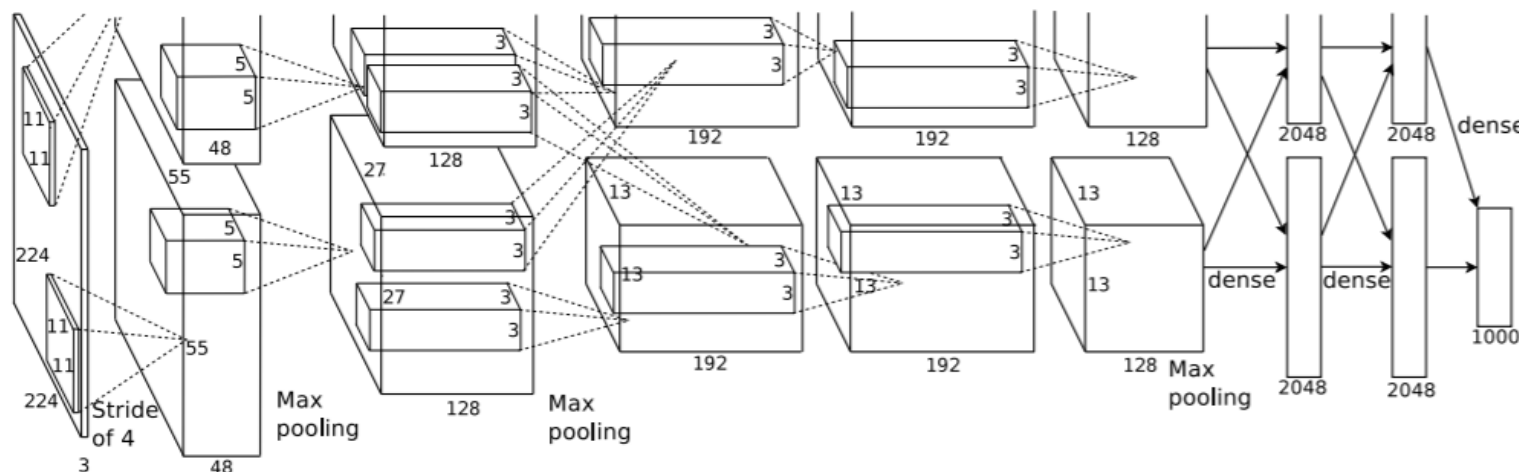
- Siguiente hito: Gradient-based learning applied to document recognition [[LeCun, Bottou, Bengio, Haffner 1998](#)]
 - Reconocimiento de caracteres escritos a mano.
 - Ya primeros experimentos en [1993](#).
 - Entrenado con **backpropagation**.



Un poco de historia

- Siguiente hito: **ImageNet** Classification with Deep Convolutional Neural Networks [[Krizhevsky, Sutskever, Hinton 2012](#)]. **Diferencias:**
 - Entrenado sobre ImageNet (10^6 imágenes, 1000 categorías)
 - Usa ReLU (LeNet 5 usa tanh) y dropout
 - Más profundo
 - Uso de GPUs (6 días)

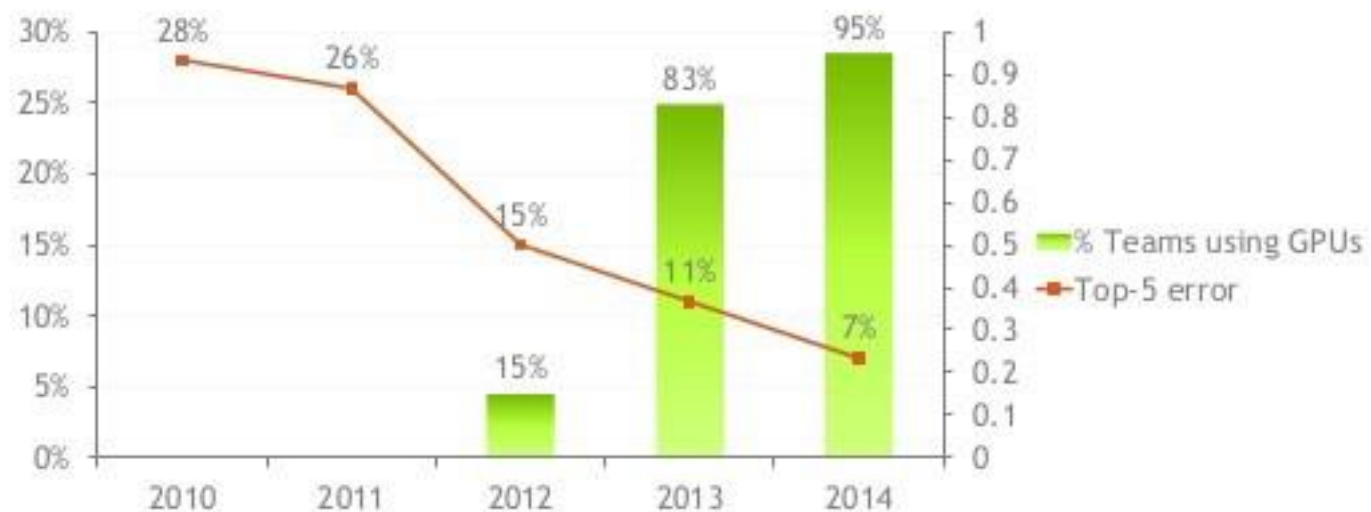
AlexNet



Un poco de historia

- Tendencia**

ImageNet: results for 2010-2014



Rank	Method	Top 1 Accuracy	Top 5 Accuracy	Number of params	Extra Training Data	Paper Title	Year	Paper	Code
1	FixResNeXt-101 32x48d	86.4%	98.0%	829M	✓	Fixing the train-test resolution discrepancy	2019		
2	ResNeXt-101 32x48d	85.4%	97.6%	829M	✓	Exploring the Limits of Weakly Supervised Pretraining	2018		
3	ResNeXt-101 32x32d	85.1%	97.5%	466M	✓	Exploring the Limits of Weakly Supervised Pretraining	2018		

Uso de redes convolucionales

Traffic Sign Recognition
(GTSRB) 99.2% accuracy

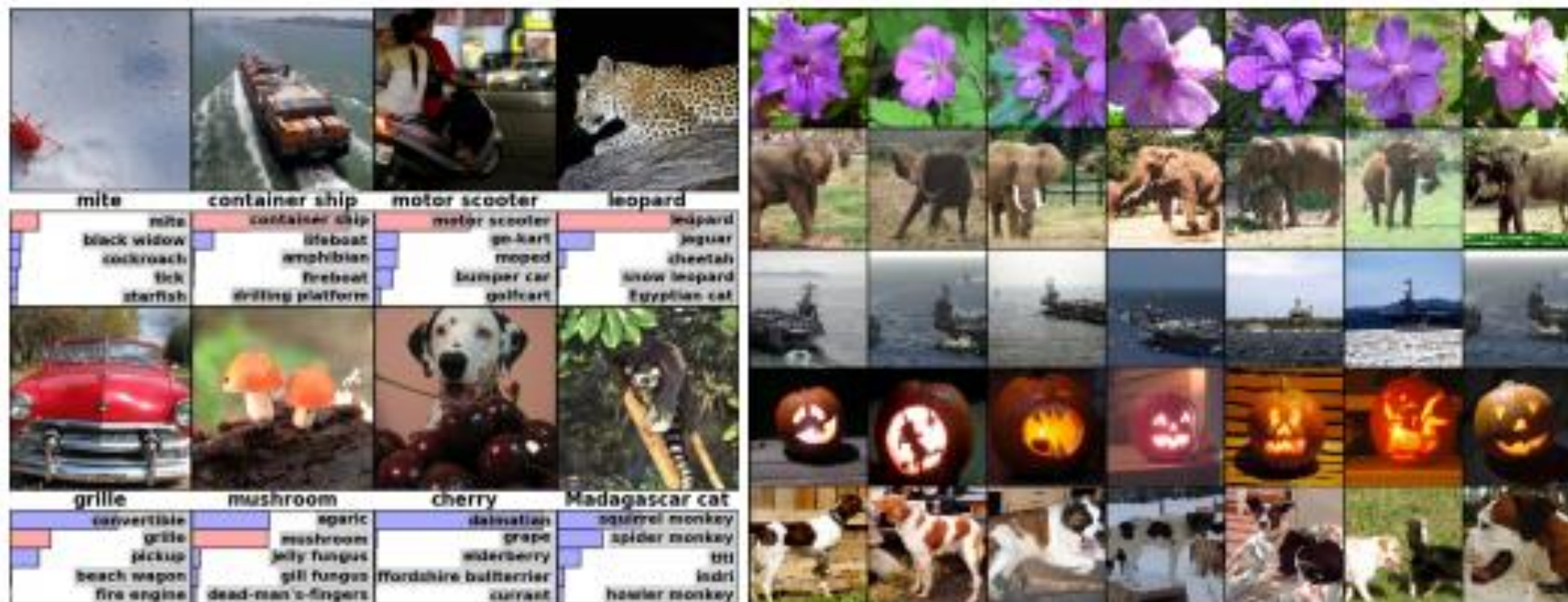


House Number Recognition
(Google) 94.3% accuracy



Uso de redes convolucionales

- Clasificación y detección de objetos en imágenes



Uso de redes convolucionales

- Localización de objetos en imágenes



Uso de redes convolucionales

- Subtitulado de imágenes (caption)



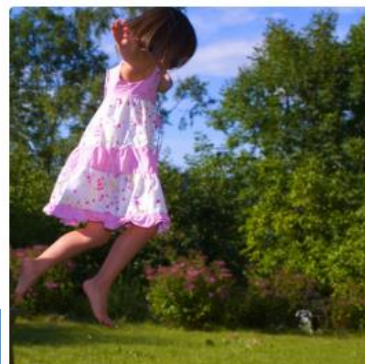
"man in black shirt is playing guitar."



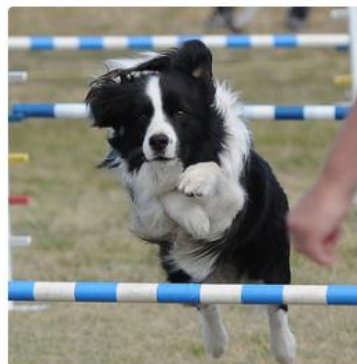
"construction worker in orange safety vest is working on road."



"two young girls are playing with lego toy."



"girl in pink dress is jumping in air."



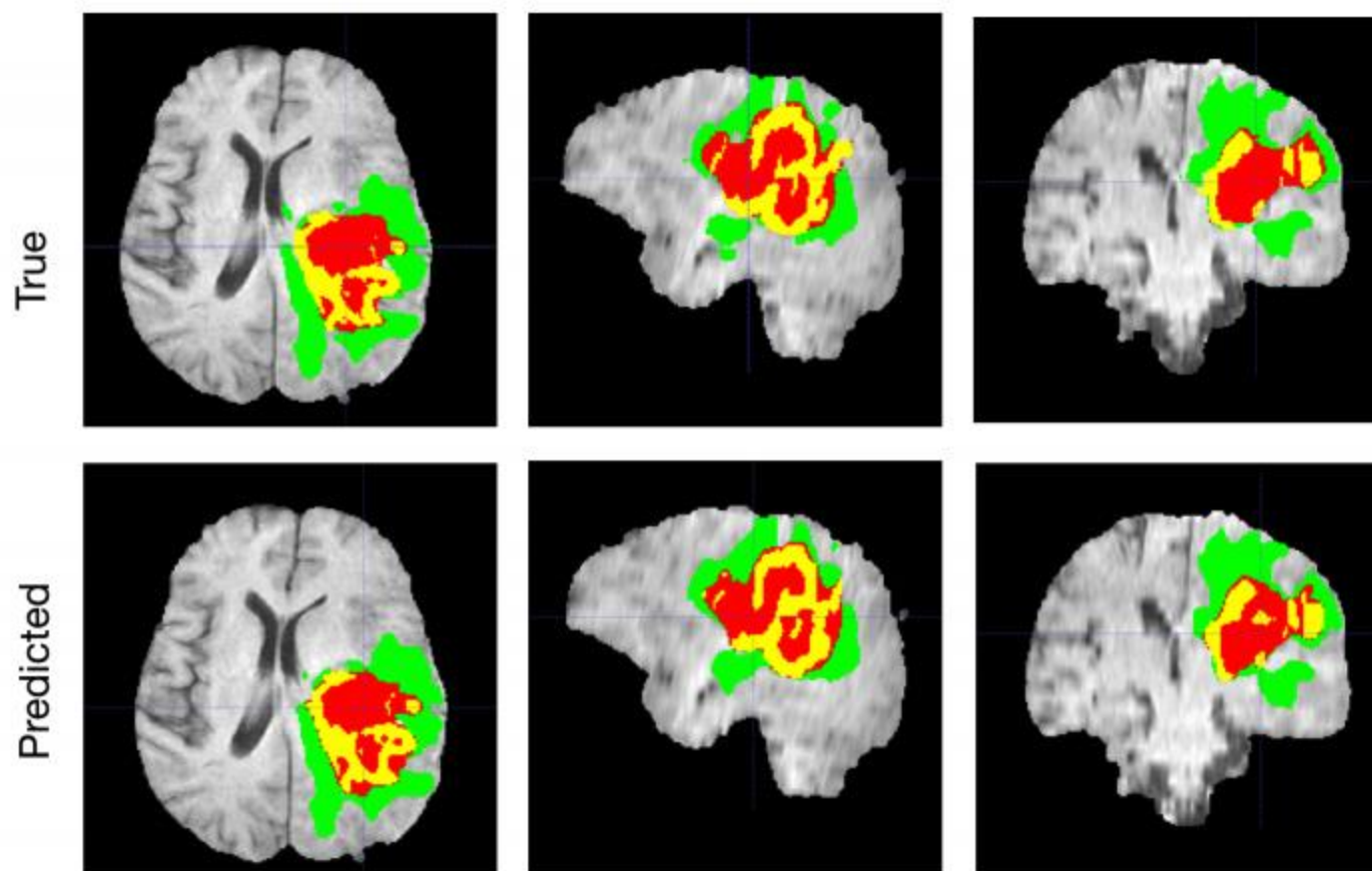
"black and white dog jumps over bar."



"young girl in pink shirt is swinging on swing."

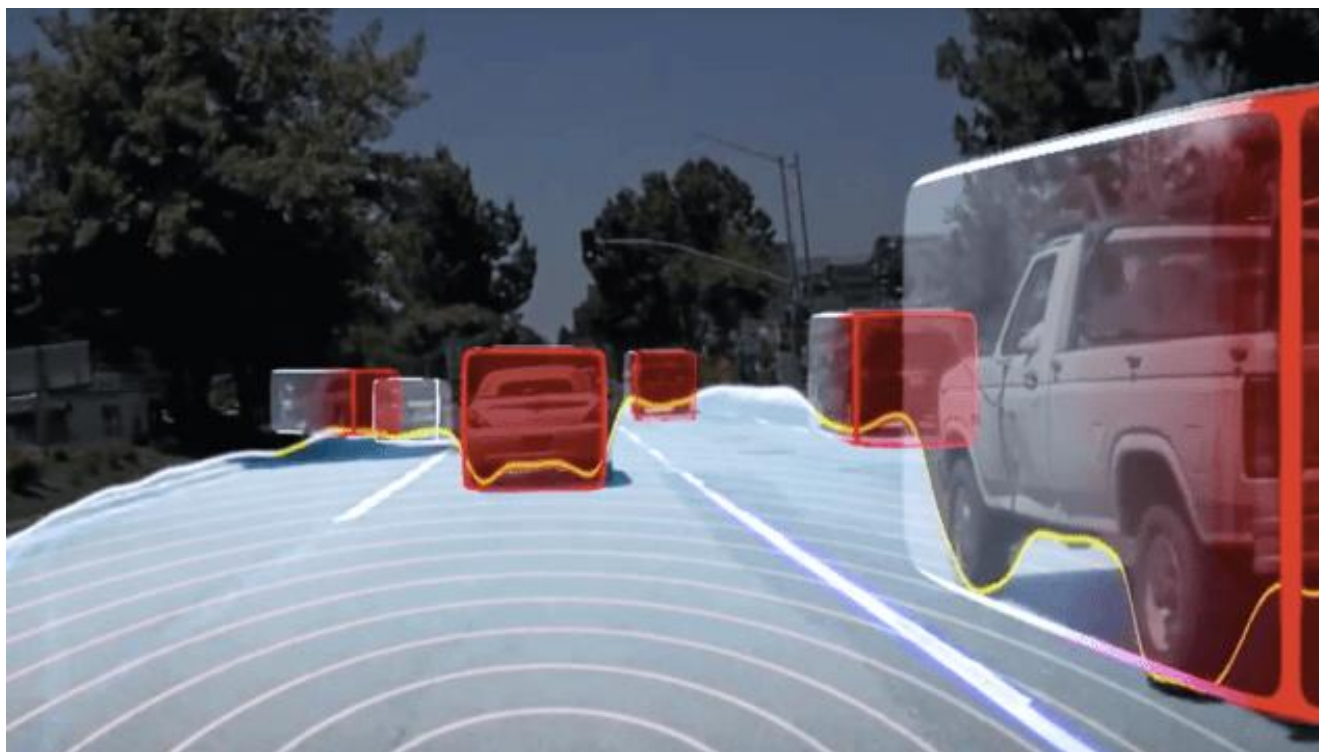
Uso de redes convolucionales

- Segmentación de tumores



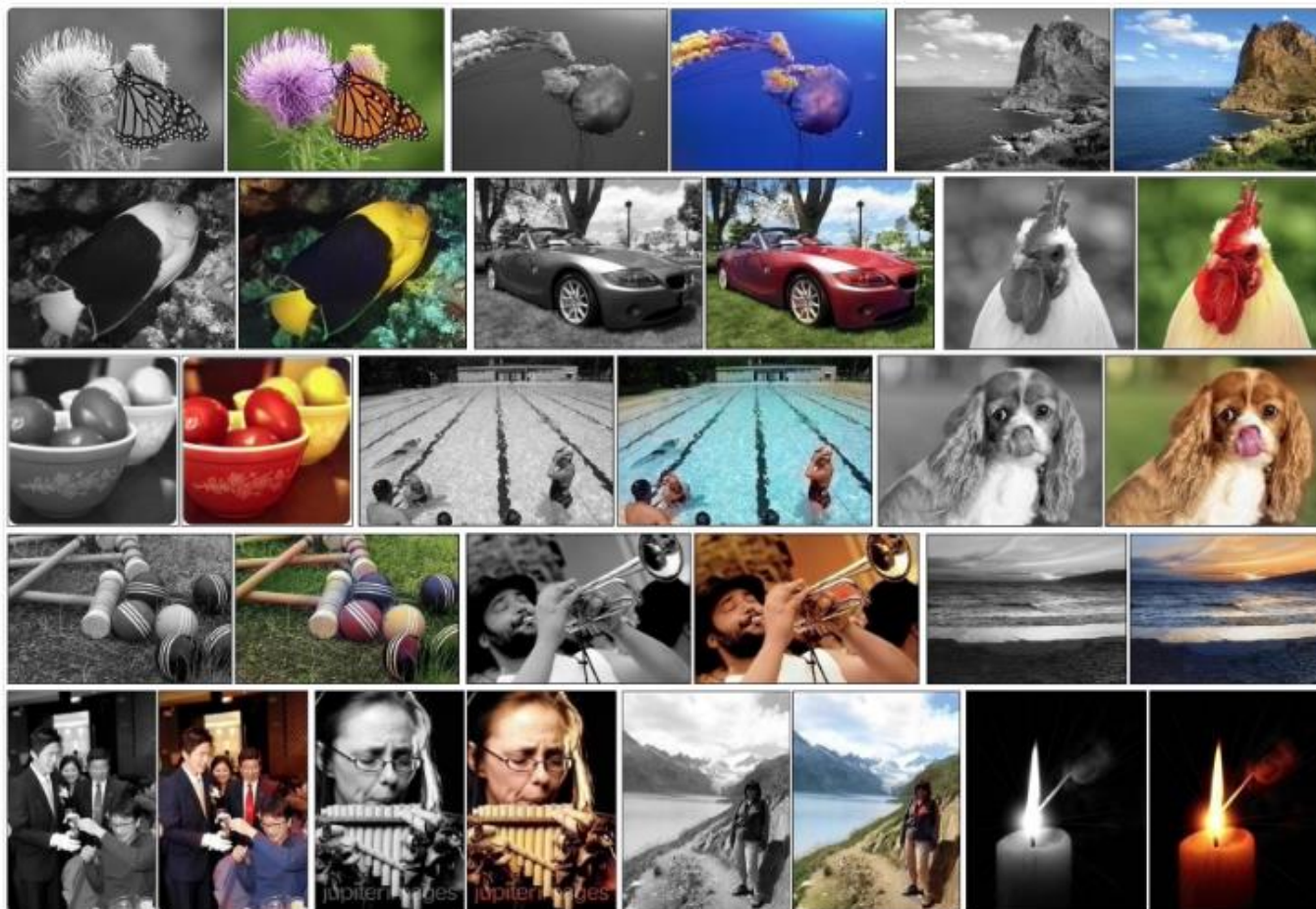
Uso de redes convolucionales

- Conducción autónoma (segmentación, localización obstáculos)



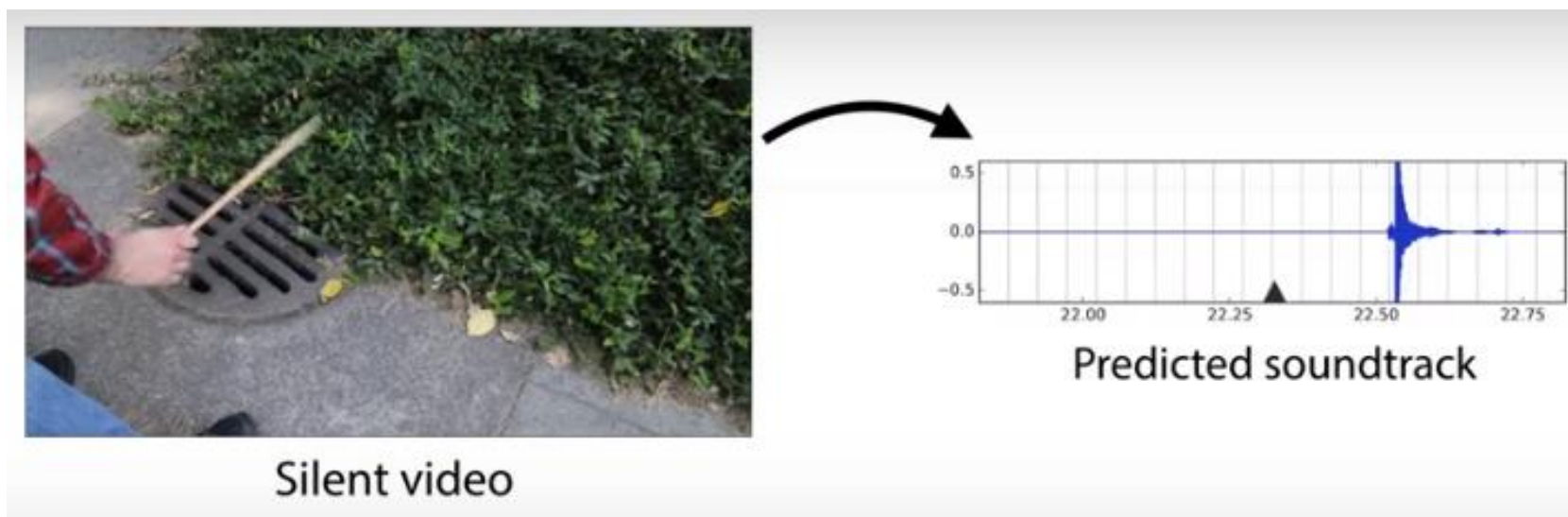
Uso de redes convolucionales

- Colorización de imágenes en blanco y negro



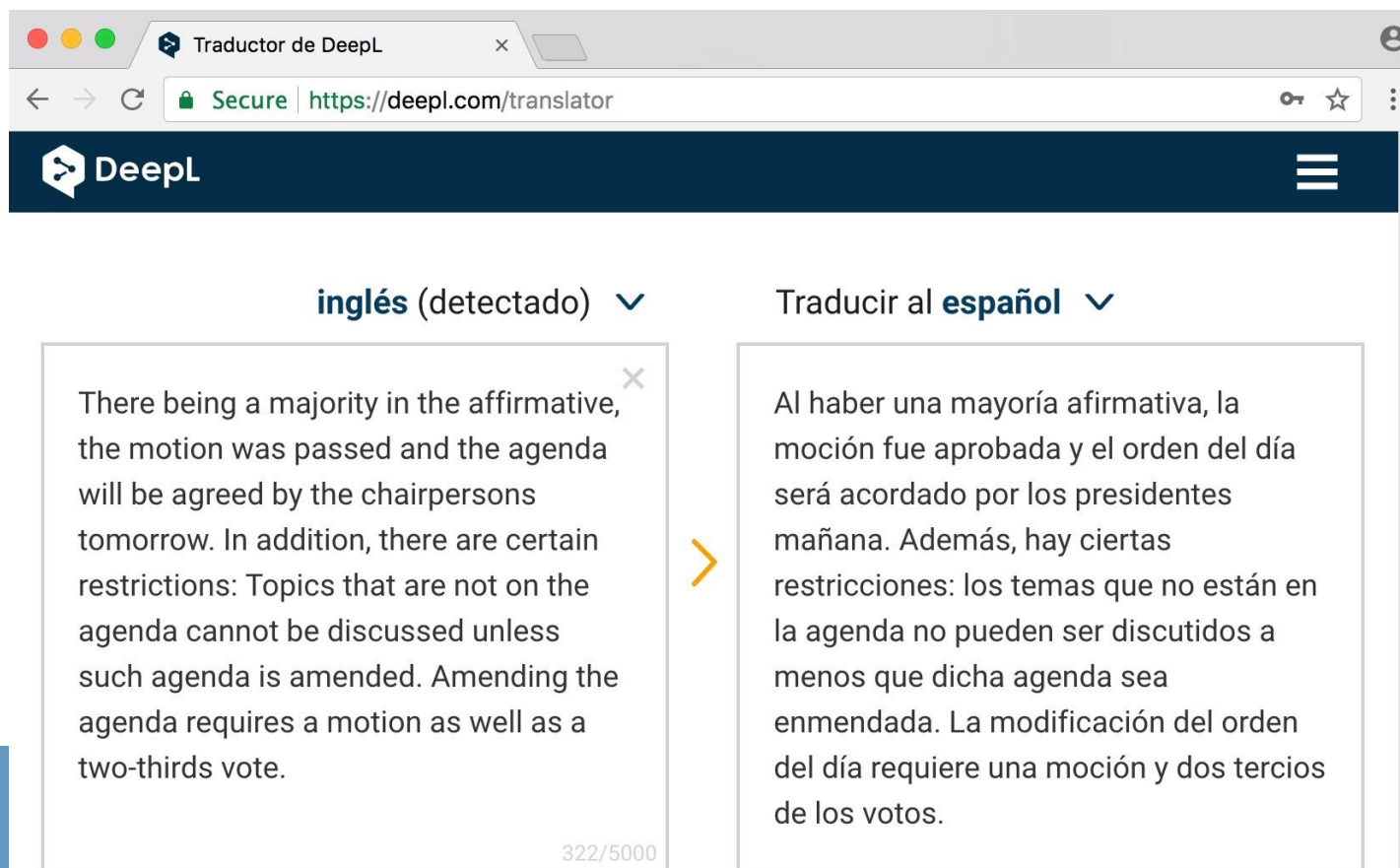
Uso de redes convolucionales

- Añadir audio automático a vídeos sin sonido



Uso de redes convolucionales

- También son buenas con otros tipos de datos, como texto o audio



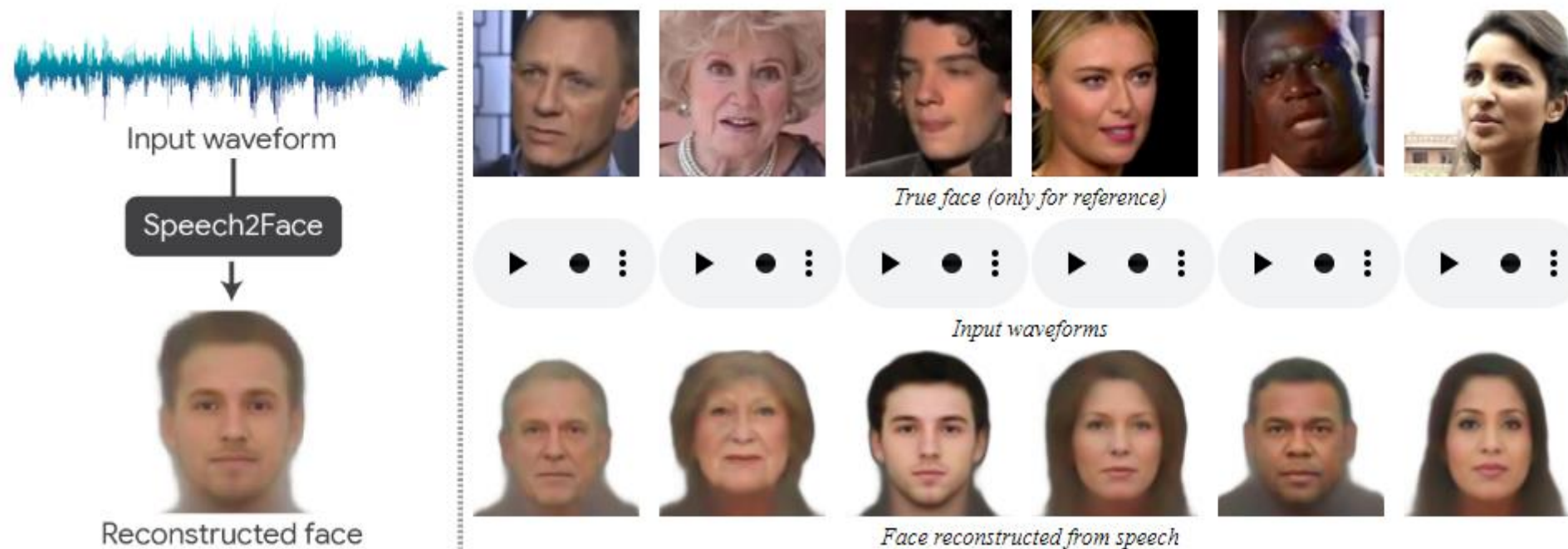
Uso de redes convolucionales

- También son buenas con otros tipos de datos, como texto o audio



Uso de redes convolucionales

- O poner cara a una voz



Uso de redes convolucionales

- También las ConvNets son buenas generando, por ejemplo texto

PANDARUS:

Alas, I think he shall be come approached and the day
When little strain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:

They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:

Well, your wit is in the care of side and that.

Second Lord:

They would be ruled after this chamber, and
my fair nudes begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

Clown:

Come, sir, I will make did behold your worship.

VIOLA:

I'll drink it

Machine learning Mastery
Machine Learning Mastery
Machine Learning Mastery

Uso de redes convolucionales

- Aplicando un estilo a imágenes



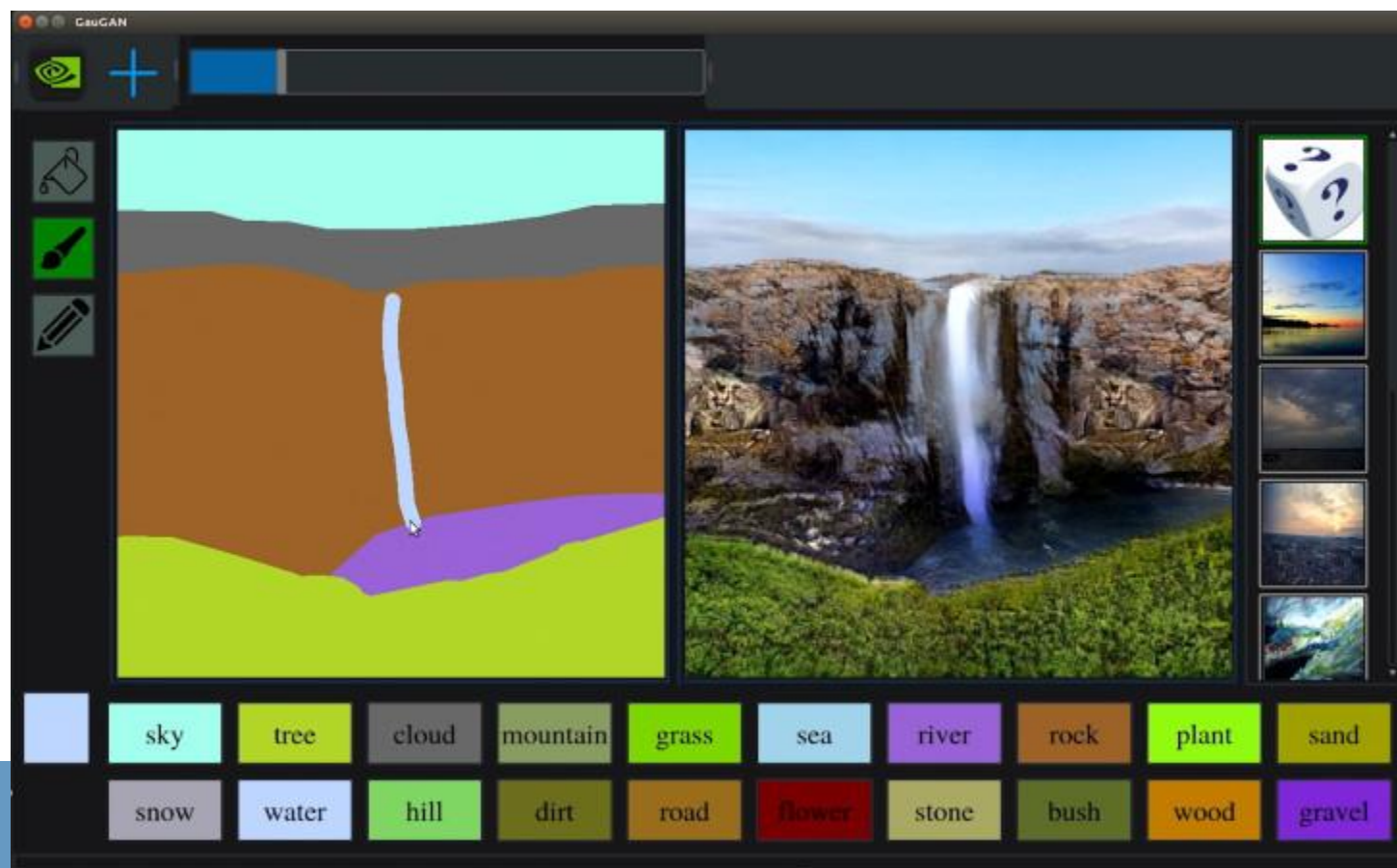
Uso de redes convolucionales

- Generando caras artificiales



Uso de redes convolucionales

- Paisajes fotorealísticos desde un dibujo



Uso de redes convolucionales

- O en aprendizaje por refuerzo, como ganar a StarCraft (AlphaStar) y al GO (AlphaGo)



Recapitulando

- La **clave** del **Deep Learning**: representación jerarquizada de características
- Inspirado en el **córtex visual** de los mamíferos
- **Redes convolucionales** permiten representar esas características
 - LeNet (años 90)
 - AlexNet (2012)
- Hay muchas aplicaciones de las redes convolucionales (y seguro sabéis de muchos más)