



# UNIVERSIDAD SANTO TOMÁS

## PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA

---

### SECCIONAL TUNJA

---

VIGILADA MINEDUCACIÓN - SNIES 1732



Acreditación Institucional  
**Internacional**  
OTORGADA POR EL IAC CINDA ACUERDO 55 DEL 9 DE MAYO-VIGENCIA 5 AÑOS



Res. MEN No. 01456 del 29 de enero de 2006

Vigencia por seis años





**UNIVERSIDAD SANTO TOMÁS**  
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA  
**SECCIONAL TUNJA**

VIGILADA MINEDUCACIÓN - SNIES 1732

**Faculty:** Systems engineer

**Course:** Deep Learning

**Topic:** Computer vision

---

**Professor:** Luis Fernando Castellanos Guarin

**Email:** [Luis.castellanosg@usantoto.edu.co](mailto:Luis.castellanosg@usantoto.edu.co)

**Phone:** 3214582098

# CONTENIDO

- Sentido de la visión en humanos
  - Lóbulo occipital
  - Corteza visual
  - Interpretando escenas
- Visión artificial
  - Historia
  - Que datos tiene una imagen?
  - Técnicas y niveles en las imágenes
  - En funcionamiento

¡Siempre  
hacia lo alto!



*¿Por qué estudiar e  
investigar en el campo de la  
visión artificial?*



*¡Siempre  
hacia lo alto!*



# El sentido de la visión

**El sentido de la visión es uno de los grandes prodigios de la Naturaleza para la mayoría de los seres vivos.**

En fracciones de segundos, podemos identificar objetos dentro de nuestro campo de visión, sin siquiera detenernos a pensar en ello.

Pero no sólo podemos nombrar estos objetos que observamos, sino que también podemos percibir su profundidad, distinguir perfectamente sus contornos, y separarlos de sus fondos.

Los ojos captan datos de **píxeles**, pero el cerebro transforma esa información en características más significativas - **líneas, curvas y formas** - que podrían a una persona o una manzana al mismo tiempo.



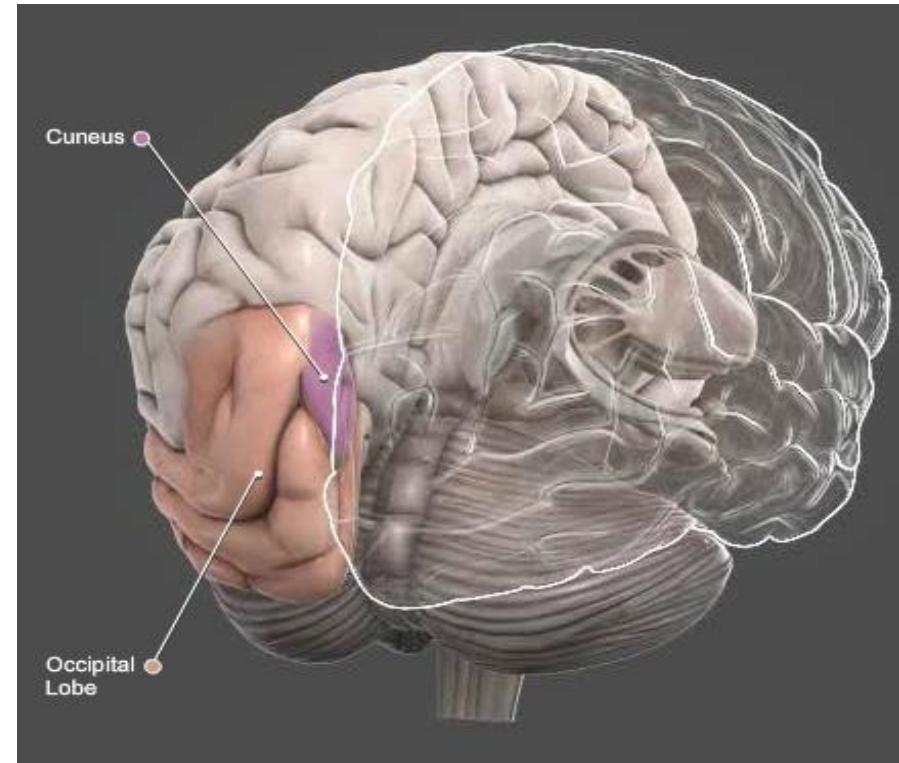


# El sentido de la visión en humanos

## Lóbulo occipital

El lóbulo occipital está situado en la zona posterior del cerebro, por detrás de los lóbulos parietal y temporal. Se encarga del procesamiento visual.

Corteza visual o “*Córtex visual*”. El lóbulo occipital contiene el área visual primaria, secundaria y de asociación.

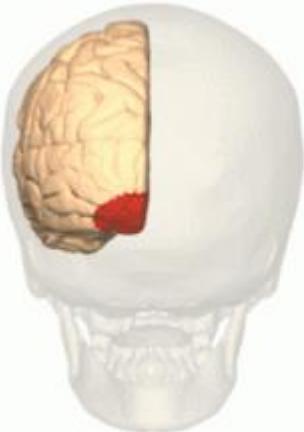


*Tomada de <http://www.g2conline.org>*



# El sentido de la visión en humanos

## Lóbulo occipital



Corteza visual primaria V1 o áreas de Brodmann BA17



Corteza visual, citoarquitectura de la corteza cerebral con sus áreas de Brodmann BA17, BA18 y BA19

### corteza visual (cerebro humano)

**V1 o área visual primaria** situada en la región más posterior del lóbulo occipital. V1 tiene un muy bien definido mapa de la información espacial en visión. Lesiones en ésta zona provocan ceguera cortical, es decir, la persona no ve el estímulo debido a que éste no es procesado en el área cortical correspondiente.

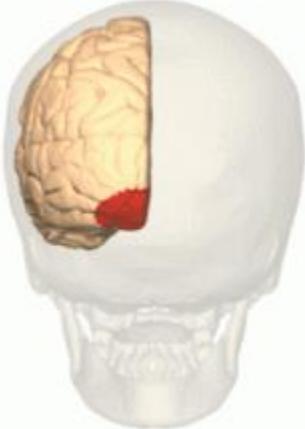
Es quien recibe la información visual directamente del ojo. Las neuronas de V1 se ocupan de características visuales de bajo nivel, tales como pequeños segmentos de contorno, componentes de pequeña escala del movimiento, disparidad binocular (es la forma de percibir profundidad y relieve) , e información básica de contraste y color.

V1 luego alimenta de información a otras áreas, como V2, V4 y V5. Cada una de estas áreas se ocupa de los aspectos más específicos o abstractas de la información. Por ejemplo, las neuronas en V4 se ocupan de objetos de mediana complejidad, tales como formas de estrellas en diferentes colores.



# El sentido de la visión en humanos

## Lóbulo occipital



Corteza visual primaria V1 o áreas de Brodmann BA17



Corteza visual, citoarquitectura de la corteza cerebral con sus áreas de Brodmann BA17, BA18 y BA19

### corteza visual (cerebro humano)

#### Área visual terciaria o V3, V4 y V5:

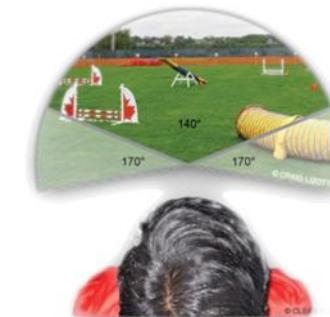
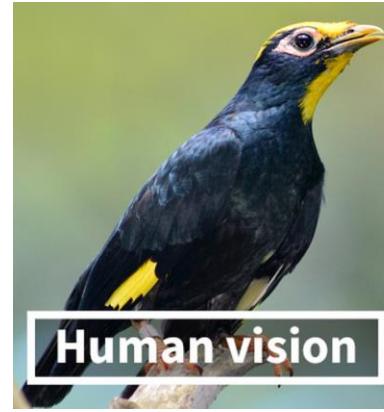
Reciben aferencias tanto de V1 como de V2.

- Las células de V3 son sensibles a la orientación y la disparidad binocular (permite una visión tridimensional de casi la totalidad del espacio visual).
- V4 participa en el análisis del color y la forma de los estímulos visuales.
- V5 colabora en la percepción del movimiento.

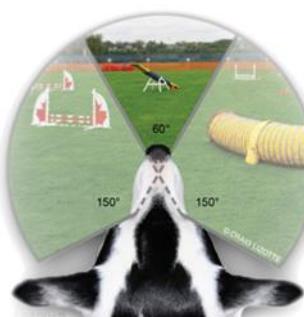


# El sentido de la visión

La corteza visual de los **animales** es el más potente sistema de procesamiento visual que conocemos, por lo que suena lógico inspirarse en ella para crear una variante de **redes neuronales artificiales (RNA)** que ayude a identificar imágenes, usando **DEEP LEARNING**.



Human vision



Dog vision

¡Siempre  
hacia lo alto!

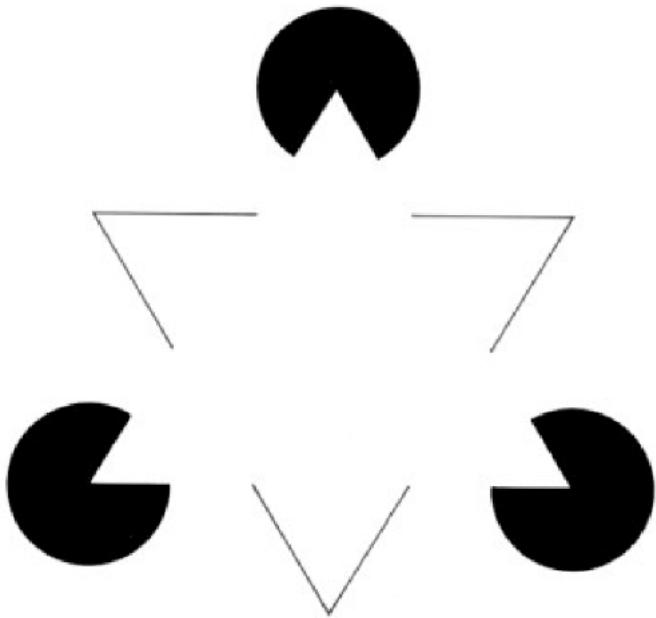


# El sentido de la visión

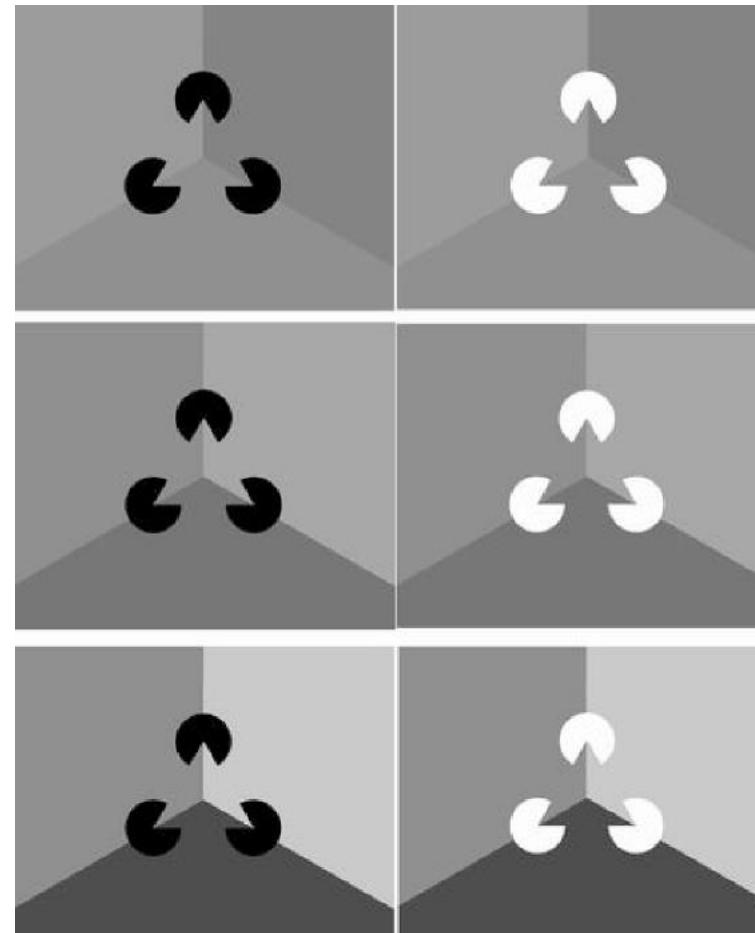
## Lóbulo occipital

### Interpretando escenas

Triángulo de Kaniza



En la figura se percibe un triángulo equilátero blanco, pero de hecho no existe ninguno. Este efecto es conocido como contorno subjetivo.



¡Siempre  
hacia lo alto!



# El sentido de la visión

Lóbulo occipital

## Interpretando escenas

Efectos visuales en imágenes que engañan al cerebro pensando en 3D y son escena 2D.



¡Siempre  
hacia lo alto!

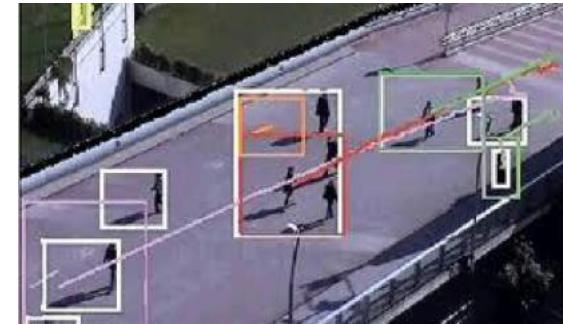


# Visión Artificial

La visión artificial tiene como finalidad la extracción de información del mundo físico a partir de imágenes digitales, utilizando para ellos un “computador”.

Tiene aplicaciones en multitud de campos:

- Seguridad.
- Salud.
- Educación.
- Industria.
- Cultivos.
- .....



¡Siempre  
hacia lo alto!



# Visión Artificial

## Historia

- 1966: Minsky assigns computer vision as an undergrad summer Project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- **2010's -> present: Deep Learning – tensorflow - Opencv**

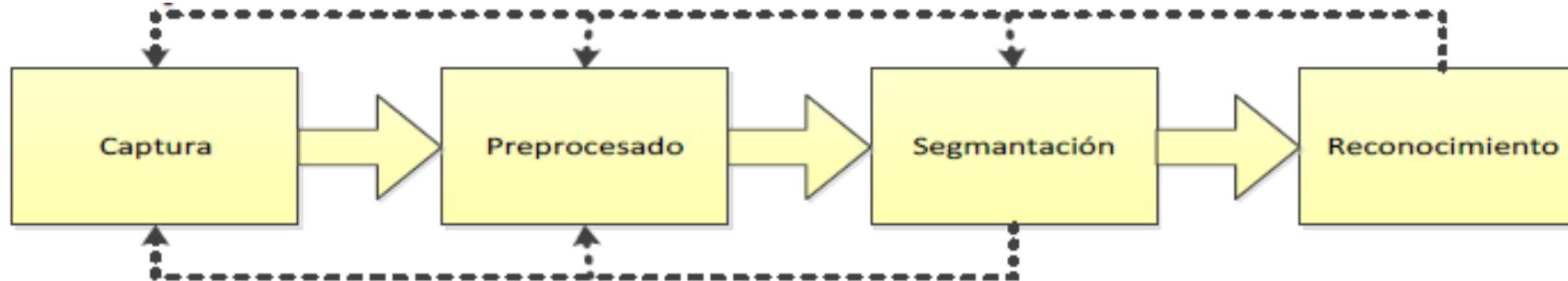


iSiempre  
hacia lo alto!



# Visión Artificial

## Pasos



- **Captura:** Adquisición de imágenes digitales mediante algún tipo de sensor, típicamente un CCD.
- **Pre-procesado:** Preparar la imagen eliminando las partes no útiles y/o realzando las partes interesantes.
- **Segmentación:** Aislar los elementos de interés para su posterior interpretación.
- **Reconocimiento:** Distinguir los diferentes objetos segmentados en función de sus características.

En muchas ocasiones es necesario una realimentación entre las fases.

¡Siempre  
hacia lo alto!



# Visión Artificial

Datos en una imagen

x =	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	
y =	41	210	209	204	202	197	247	143	71	64	80	84	54	54	57	58
	42	206	196	203	197	195	210	207	56	63	58	53	53	61	62	51
	43	201	207	192	201	198	213	156	69	65	57	55	52	53	60	50
	44	216	206	211	193	202	207	208	57	69	60	55	77	49	62	61
	45	221	206	211	194	196	197	220	56	63	60	55	46	97	58	106
	46	209	214	224	199	194	193	204	173	64	60	59	51	62	56	48
	47	204	212	213	208	191	190	191	214	60	62	66	76	51	49	55
	48	214	215	215	207	208	180	172	188	69	72	55	49	56	52	56
	49	209	205	214	205	204	196	187	196	86	62	66	87	57	60	48
	50	208	209	205	203	202	186	174	185	149	71	63	55	55	45	56
	51	207	210	211	199	217	194	183	177	209	90	62	64	52	93	52
	52	208	205	209	209	197	194	183	187	187	239	58	68	61	51	56
	53	204	206	203	209	195	203	188	185	183	221	75	61	58	60	60
	54	200	203	199	236	188	197	183	190	183	196	122	63	58	64	66
	55	205	210	202	203	199	197	196	181	173	186	105	62	57	64	63



RGB (**red, green, blue**) o RVA (rojo, verde y azul) es la composición del color en términos de la intensidad de los colores primarios de la luz.

¡Siempre  
hacia lo alto!



# Visión Artificial

Existen tres técnicas y niveles para el tratamiento de las imágenes:

- Alto nivel: procesamiento de características
- Nivel medio: imágenes a características
- Bajo nivel: imágenes a imágenes



¡Siempre  
hacia lo alto!



# Visión Artificial

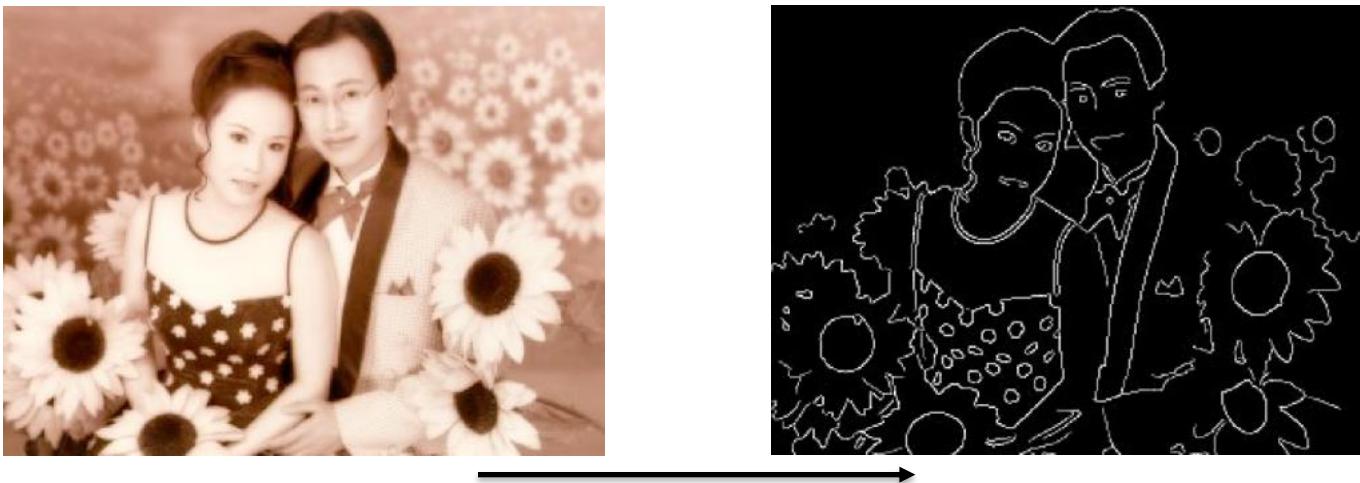
Nivel Bajo

Sharpening /afilar



blurring

Canny (detectar bordes)



¡Siempre  
hacia lo alto!



# Visión Artificial

Nivel Medio



edge image



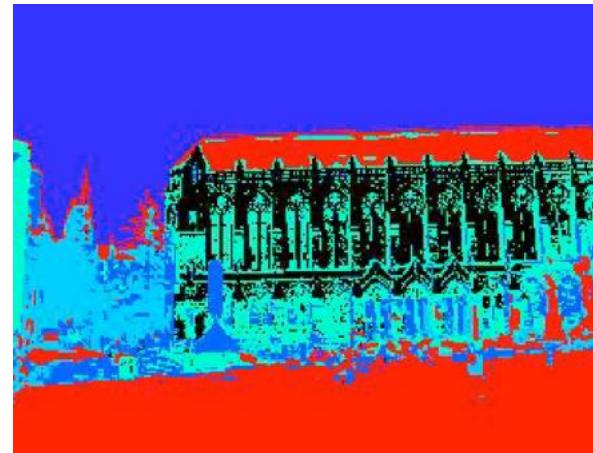
original color image

Data  
structure



circular arcs and line segments

K-means  
clustering



regions of homogeneous color

¡Siempre  
hacia lo alto!



# Visión Artificial

Nivel Alto

Reconocimiento de objetos

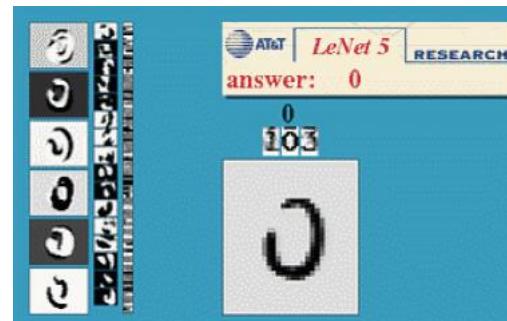


Reconocimiento de caracteres

4YCH428

4YCH428

4YCH428



¡Siempre  
hacia lo alto!



# Visión artificial

En funcionamiento

## Detección y reconocimiento de caras



## Efectos especiales



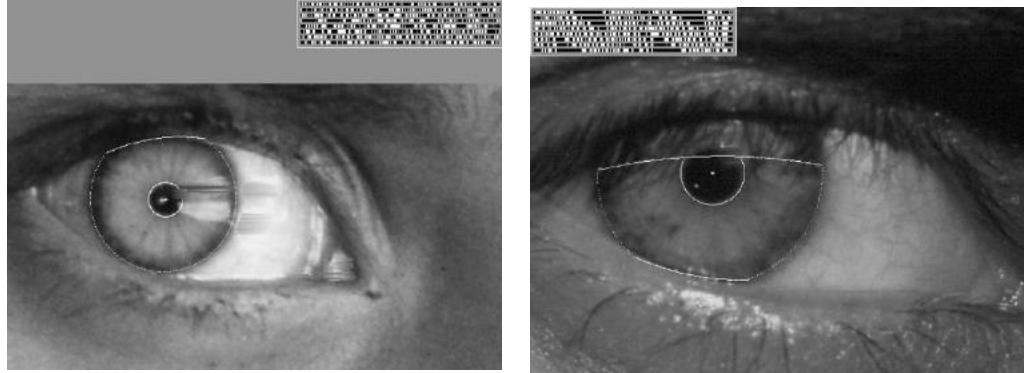
¡Siempre  
hacia lo alto!



# Visión artificial

En funcionamiento

Datos biométricos a partir de imágenes



Sistemas de seguridad



¡Siempre  
hacia lo alto!



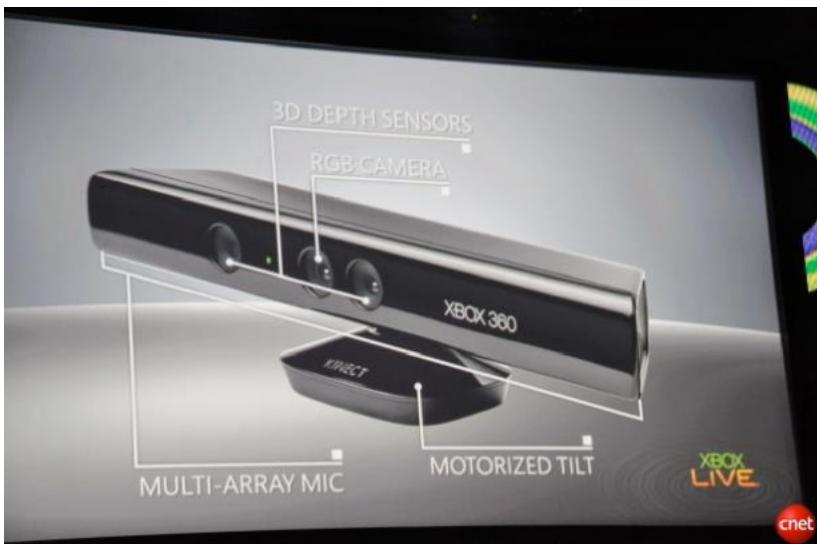
# Visión artificial

En funcionamiento

## Retransmisiones deportivas



## Juegos interactivos: Kinect , HTC vive



¡Siempre  
hacia lo alto!



## Software que vamos a necesitar

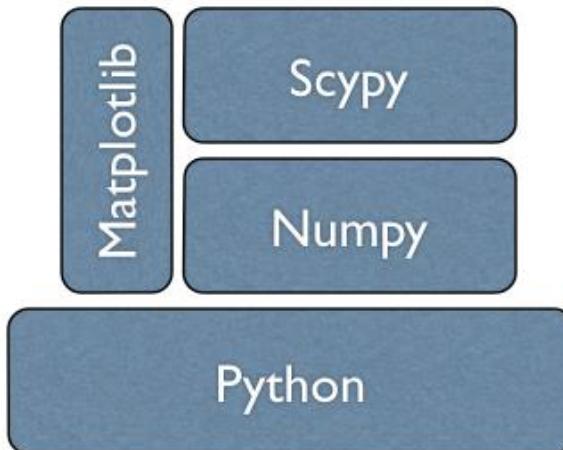
Básico

# Python

**Scipy:** librerías de alto nivel

**Numpy:** arrays y operaciones con matrices

**Matplotlib**: visualización, gráficas



¡Siempre  
hacia lo alto!



## Software que vamos a necesitar

Especializado

### OpenCV

#### Historia

- Librería desarrollada inicialmente por Intel en 1999
- Mantenido en la actualidad por la empresa de robótica Willow Garage



#### Características principales

- Proporciona una librería open source bien testeada y optimizada para procesamiento de imagen, visión y operaciones geométricas
- Escrita en C, asegurándose que la implementación es rápida y portable
- Compilado para múltiples plataformas, incluyendo plataformas embebidas y GPUs
- Wrappers para distintos lenguajes, incluyendo Python y C++



## Software que vamos a necesitar

Especializado

### Tensorflow



TensorFlow

#### Historia

TensorFlow es el sistema de aprendizaje automático de segunda generación de Google Brain, liberado como software de código abierto en 9 de noviembre del 2015.

#### Características principales

- La API esta diseñada para Python, aunque hay acceso al lenguaje subyacente C++
- Puede funcionar en una sola CPUs o con GPUs, así como dispositivos móviles (tensorflow lite)
- Usada para diagnósticos médicos.
- Procesamiento de imágenes

¡Siempre  
hacia lo alto!



# Software que vamos a necesitar

Especializado

## Tensorflow

<https://deepdreamgenerator.com/>



TensorFlow



Se trata de un programa de visión artificial creado por el ingeniero de Google Alexander Mordvintsev, que utiliza DEEP LEARNING con una red neuronal convolucional (CNN) para encontrar y mejorar patrones en imágenes mediante pareidolia algorítmica, creando así una apariencia alucinógena, similar a un sueño, creando imágenes deliberadamente sobreprocesadas

<https://edoc.hu-berlin.de/bitstream/handle/18452/19403/Spratt%20-%20final.pdf>

¡Siempre  
hacia lo alto!



UNIVERSIDAD SANTO TOMÁS  
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA  
SECCIONAL TUNJA

VIGILADA MINEDUCACIÓN - SNIES 1732

¡Siempre  
hacia lo alto!

**USTATUNJA.EDU.CO**

