



Bootcamp

Inteligencia Artificial

Nivel (Intermedio)
Docente: Víctor Viera Balanta
Fecha 21/08/2024

UT TALENTOTECH

Tabla de contenidos

- 1
- 2
- 3

Recapitular

VISIÓN ARTIFICIAL

Tiempo de Proyecto

Clasificación de Texto con DeepLearning

Recordando Machine Learning Naive Bayes

Profesor: Víctor Viera Balanta

UT TALENTOTECH



Thomas Bayes

ScienceProg

Teorema de Bayes

el modelo fue creado por el matemático inglés, Thomas Bayes (1701 – 1761),

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Probabilidad Condicionada

Naïve Bayes Classifier

Clasificador Bayesiano ingenuo

Filtrado de Spam clásico

$$P(\text{palabra}|\text{categoria}) = \frac{P(\text{categoria}|\text{palabra}) \times P(\text{palabra})}{P(\text{categoria})}$$



$$P(\text{categoria}|\text{palabra}_1, \text{palabra}_2, \dots, \text{palabra}_n) = P(\text{categoria}) \times \prod_i P(\text{palabra}_i|\text{categoria})$$

UT TALENTOTECH

DeepLearning-Natural Language Processing(recordando Machine Learning)

Clasificador Bayesiano ingenuo
Filtrado de Spam clásico

Se utiliza el método entrenar

[**"Quedamos mañana lunes para ir al cine"**, "nospam"]

[**"precios de productos rebajados compra aquí"**, "spam"]

Recordar el código
de Naive Bayes

DeepLearning-Natural Language Processing(recordando Machine Learning)



Plataforma de extremo a extremo enfocada en el aprendizaje automático

En mayo de 2016 Google anunció TPU (Tensor Processing Unit). Se trata de una construcción ASIC específica para el aprendizaje automático y adaptada para TensorFlow.



Puentes Digitales

DeepLearning-Natural Language Processing(recordando Machine Learning)

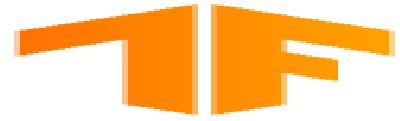


TensorFlow

¿Qué es TensorFlow?

TensorFlow es una biblioteca de software de código abierto para computación numérica, que utiliza gráficos de flujo de datos. Los nodos en las gráficas representan operaciones matemáticas, mientras que los bordes de las gráficas representan las matrices de datos multidimensionales (tensores) comunicadas entre ellos.

DeepLearning-Natural Language Processing(recordando Machine Learning)



TensorFlow

imdb_reviews

Gran conjunto de datos de revisión de películas. Este es un conjunto de datos para la clasificación de sentimientos binarios que contiene sustancialmente más datos que los conjuntos de datos de referencia anteriores. Proporcionamos un conjunto de 25 000 reseñas de películas altamente polarizadas para capacitación y 25 000 para pruebas. También hay datos adicionales sin etiquetar para su uso.

DeepLearning-Natural Language Processing(recordando Machine Learning)

imdb_reviews



IMDB Dataset of 50K Movie Reviews

Large Movie Review Dataset

[Data Card](#) [Code \(1144\)](#) [Discussion \(9\)](#) [Suggestions \(1\)](#)

About Dataset

IMDB dataset having 50K movie reviews for natural language processing or Text analytics.

This is a dataset for binary sentiment classification containing substantially more data than previous benchmark datasets. We provide a set of 25,000 highly polar movie reviews for training and 25,000 for testing. So, predict the number of positive and negative reviews using either classification or deep learning algorithms.

For more dataset information, please go through the following link,

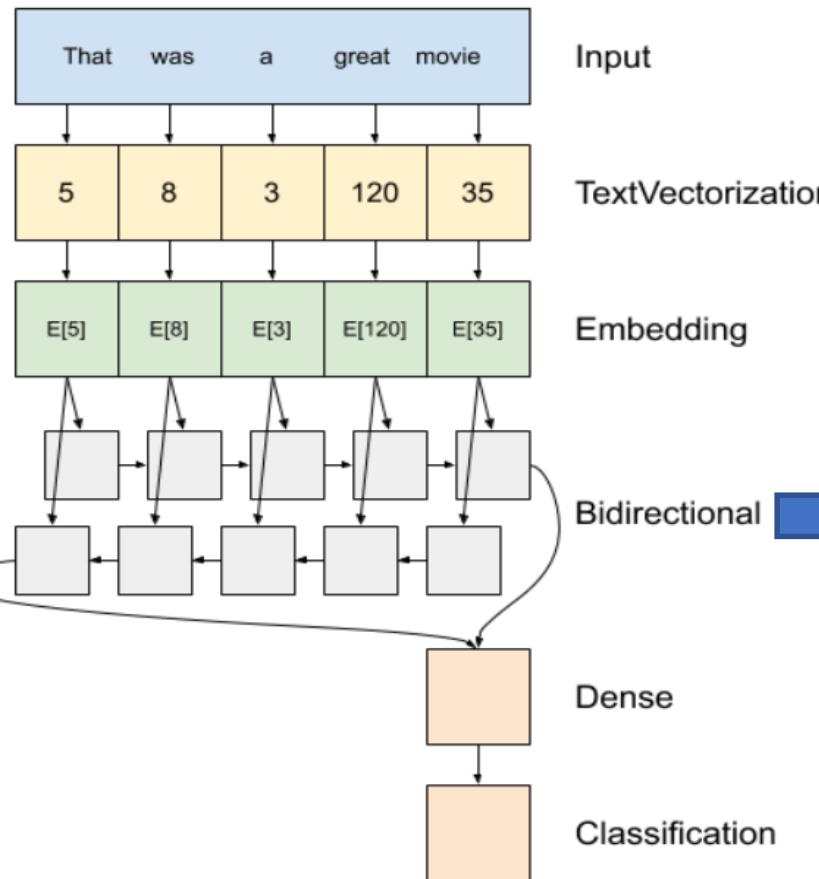
<http://ai.stanford.edu/~amaas/data/sentiment/>

<https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>

Profesor: Víctor Viera Balanta

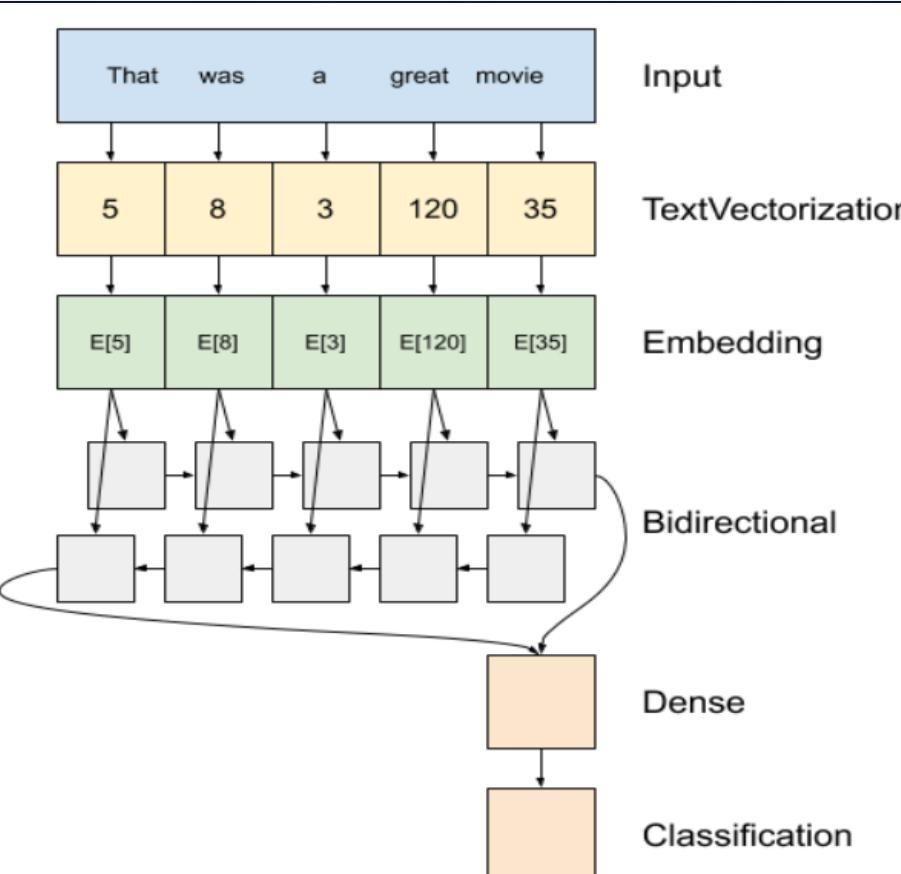
UT TALENTOTECH

DeepLearning-Natural Language Processing



- Este modelo se puede construir como un `tf.keras.Sequential`.
 - La primera capa es la `encoder` índices de tokens.
 - capa de incrustación almacena un vector por palabra convierte las secuencias de índices de palabras en secuencias de vectores
 - El `tf.keras.layers.Bidirectional` envoltura también se puede utilizar con una capa RNN. Esto propaga la entrada hacia adelante y hacia atrás a través de la capa RNN y luego concatena la salida final.
 - Una red neuronal recurrente (RNN) procesa la entrada de secuencia iterando a través de los elementos.
 - `tf.keras.layers.Dense` capa.
 - la RNN ha convertido la secuencia para un único vector

DeepLearning-Natural Language Processing-Primera Arquitectura



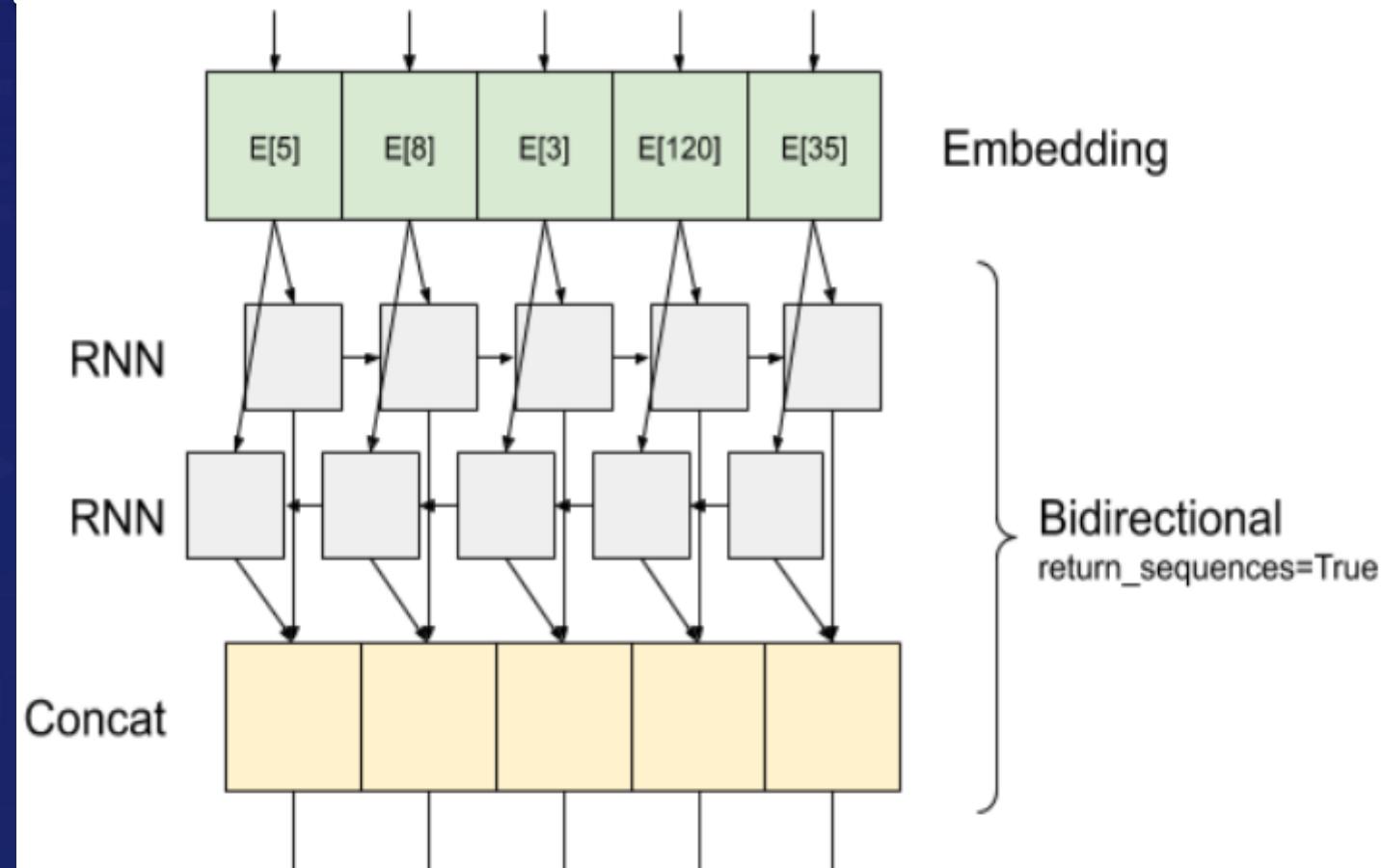
```
model = tf.keras.Sequential([
    encoder,                                     → Texto a Secuencia Numérica
    tf.keras.layers.Embedding(
        input_dim=len(encoder.get_vocabulary()),
        output_dim=64,                                → Vectores 64 dimensión
        # Use masking to handle the variable sequence lengths
        mask_zero=True),
    tf.keras.layers.Bidirectional(tf.keras.layers.LSTM(64)),
    tf.keras.layers.Dense(64, activation='relu'),
    tf.keras.layers.Dense(1)                         → Aplica una capa LSTM
                                                    bidireccional a las
                                                    incrustaciones
])
```

]) La capa de salida, con una sola neurona, clasificación binaria.

Annotations pointing to specific code snippets:

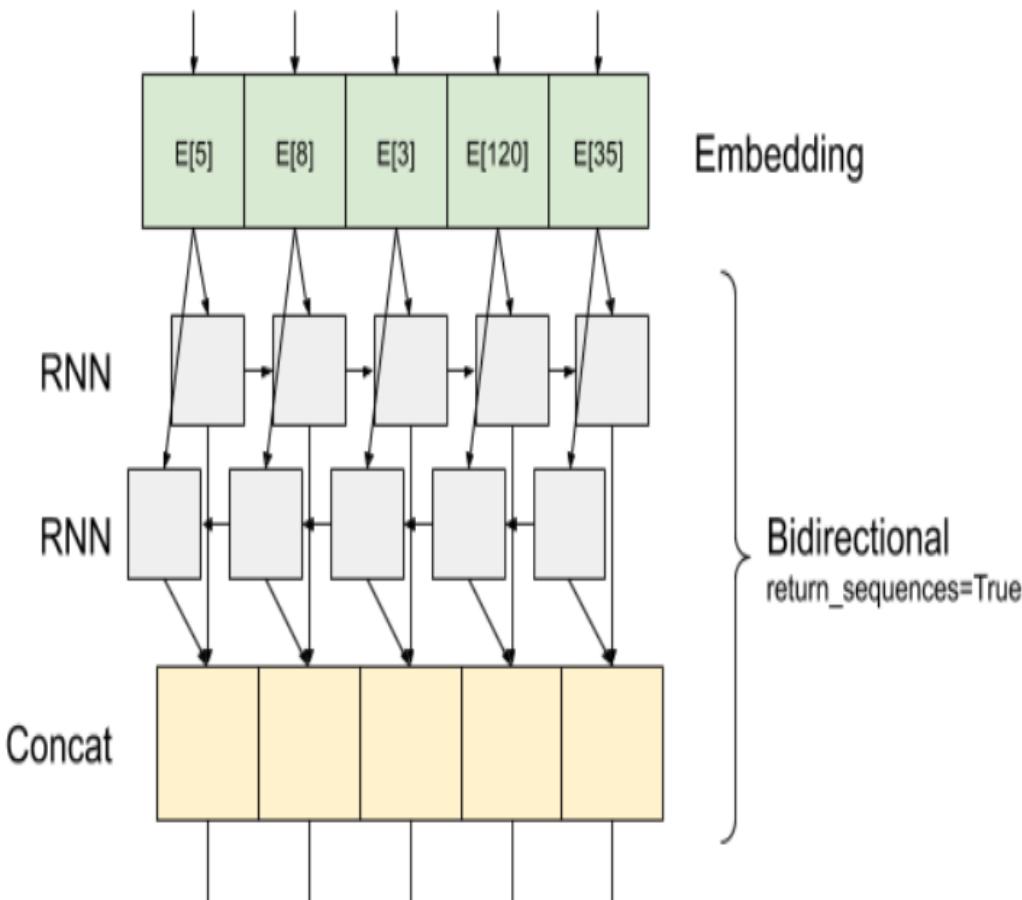
- An arrow points from the word "encoder" to the first line of the code, labeled "Texto a Secuencia Numérica".
- An arrow points from the parameter "output_dim=64" to the text "Vectores 64 dimensión".
- An arrow points from the parameter "mask_zero=True" to the text "Ceros ignorados".
- An arrow points from the word "Dense" to the line "tf.keras.layers.Dense(64, activation='relu')", labeled "Aplica una capa LSTM bidireccional a las incrustaciones".
- An arrow points from the parameter "activation='relu'" to the text "capa densa completamente conectada con 64 neuronas".

DeepLearning-Natural Language Processing-Otra Arquitectura



El uso de un RNN con `return_sequences=True` es que la salida todavía tiene 3 ejes, como la entrada, por lo que se puede pasar a otra capa RNN:

DeepLearning-Natural Language Processing-Otra Arquitectura



```
model = tf.keras.Sequential([
    encoder,
    tf.keras.layers.Embedding(len(encoder.get_vocabulary()), 64, mask_zero=True),
    tf.keras.layers.Bidirectional(tf.keras.layers.LSTM(64, return_sequences=True)),
    tf.keras.layers.Bidirectional(tf.keras.layers.LSTM(32)),
    tf.keras.layers.Dense(64, activation='relu'),
    tf.keras.layers.Dropout(0.5),
    tf.keras.layers.Dense(1)
])
```

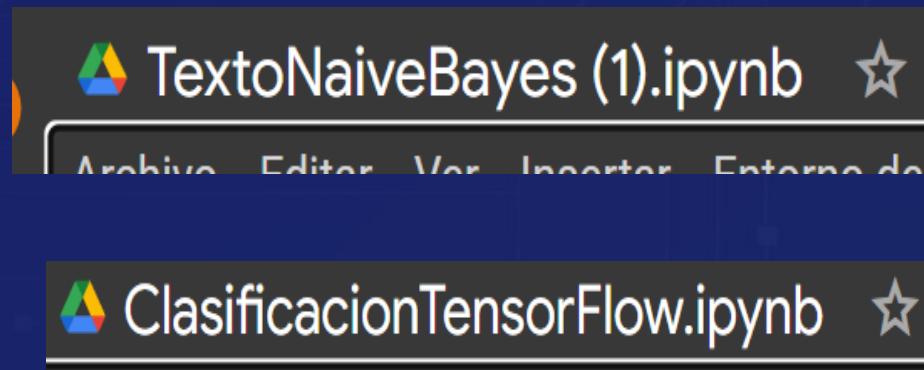
50% de las neuronas se desactivarán aleatoriamente

Cada número entero en la secuencia se mapea a un vector de 64 dimensiones.

Aplica una capa LSTM (64uni)bidireccional a las incrustaciones.

Aplica una capa LSTM (32uni)bidireccional a las incrustaciones.

DeepLearning-Natural Language Processing-Código Fuente



UT TALENTOTECH



Visión Artificial

Profesor: Víctor Viera Balanta

UT TALENTOTECH

Tratamiento digital de la imagen



Visión Artificial



Visión Artificial

¿Cómo funciona una cámara digital?



<https://youtu.be/UD3mMApdQP8?si=j3ufviQFughpO3ve>

Cómo funciona la cámara de fotos digital Discovery

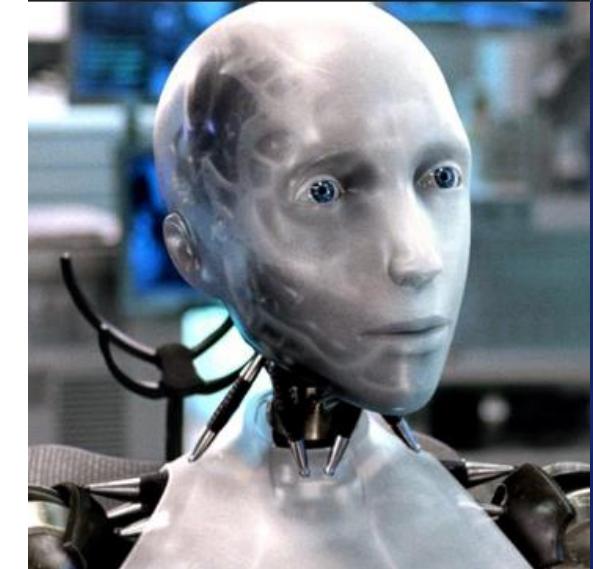
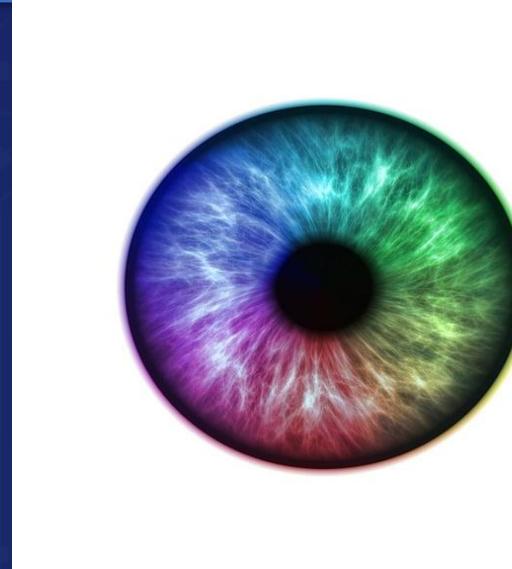
<https://youtu.be/t48UUwDvBv8?si=Im-xhHGORDznxmgG>

Profesor: Víctor Viera Balanta

UT TALENTOTECH

Visión Artificial

Campo de estudio es interdisciplinario, y persigue el objetivo de obtener sistemas capaces de interpretar información valiosa derivada de imágenes, videos y otros tipos de captura.



UT TALENTOTECH

Visión Artificial

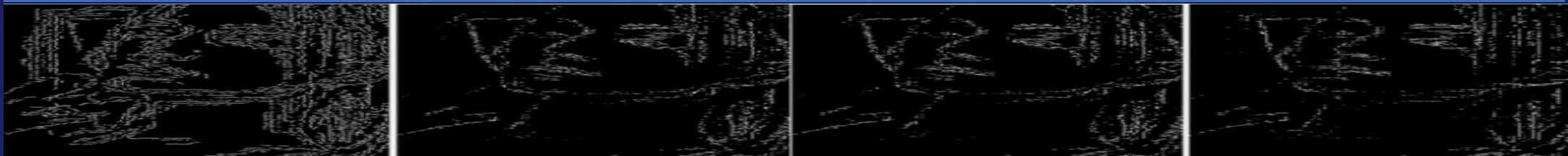
En 1959 se diseña el primer sistema que permite escanear imágenes para convertirlas en matrices de números. Nace la imagen digital.



Visión Artificial

En los años 70 surgieron los primeros algoritmos de extracción de características a partir de estas matrices numéricas denominadas imágenes digitales.

Durante esta década comenzaron las técnicas de extracción de bordes, estimación del movimiento y detección de líneas y formas básicas.



TALENTO TECH

Visión Artificial

Surgieron en consecuencia los primeros detectores de caracteres (OCRs por sus siglas en inglés de Reconocimiento Óptico de Caracteres) que permitían leer el texto contenido en imágenes, originalmente como apoyo a las personas con limitaciones en la visión.



Visión Artificial

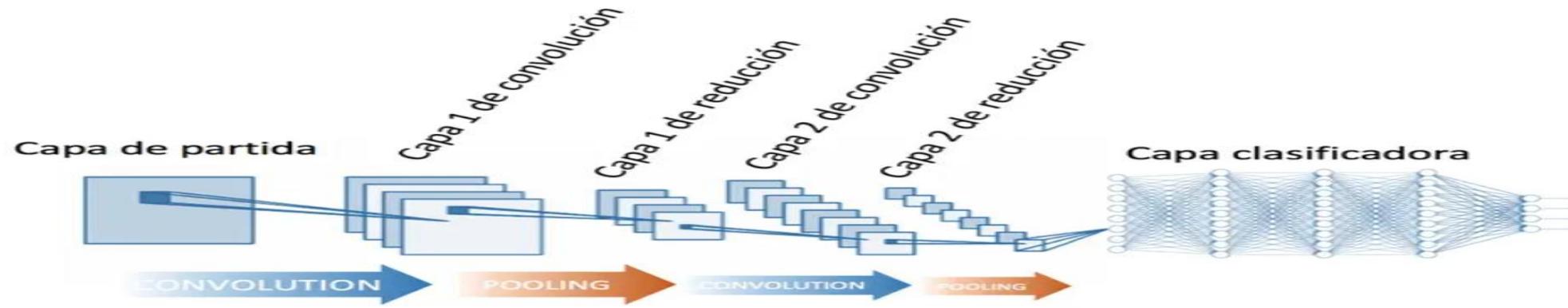
los 80, estos métodos fueron complicándose a medida que se aplicaba el análisis matemático a estas matrices de números, y se comenzó a trabajar con el concepto de textura, de sombreado y de contornos complejos.



UT TALENTOTECH

Visión Artificial

Los 80's esta fue la época en la que se presentaron por primera vez los primeros modelos de redes neuronales complejas destinadas a la visión artificial, tales como las Redes Neuronales Convolucionales (CNN).



UT TALENTOTECH

Visión Artificial

En los años 90, el foco se puso en los dispositivos de captura. La calibración espacial de cámaras, así como la extracción de información de profundidad a base de dos cámaras como en las cámaras estereoscópicas centraron toda la atención.

La fotogrametría y la reconstrucción de escenas ganaron terreno.



Visión Artificial

Gracias a la gran expansión de internet durante esta década, la red se llenó de imágenes digitales, lo que sería el campo de cultivo sobre el que crecerían las bases de datos para entrenar los grandes modelos de aprendizaje automático posteriores.



UT TALENTOTECH

Visión Artificial

Ya en el siglo XXI aparecen las grandes competiciones de reconocimiento de imágenes como PascalVOC o ImageNet. También comienza a crecer enormemente el sector del vehículo autónomo, siempre sustentado en los últimos algoritmos de visión artificial.

The **PASCAL** Visual Object Classes Homepage

Visión Artificial

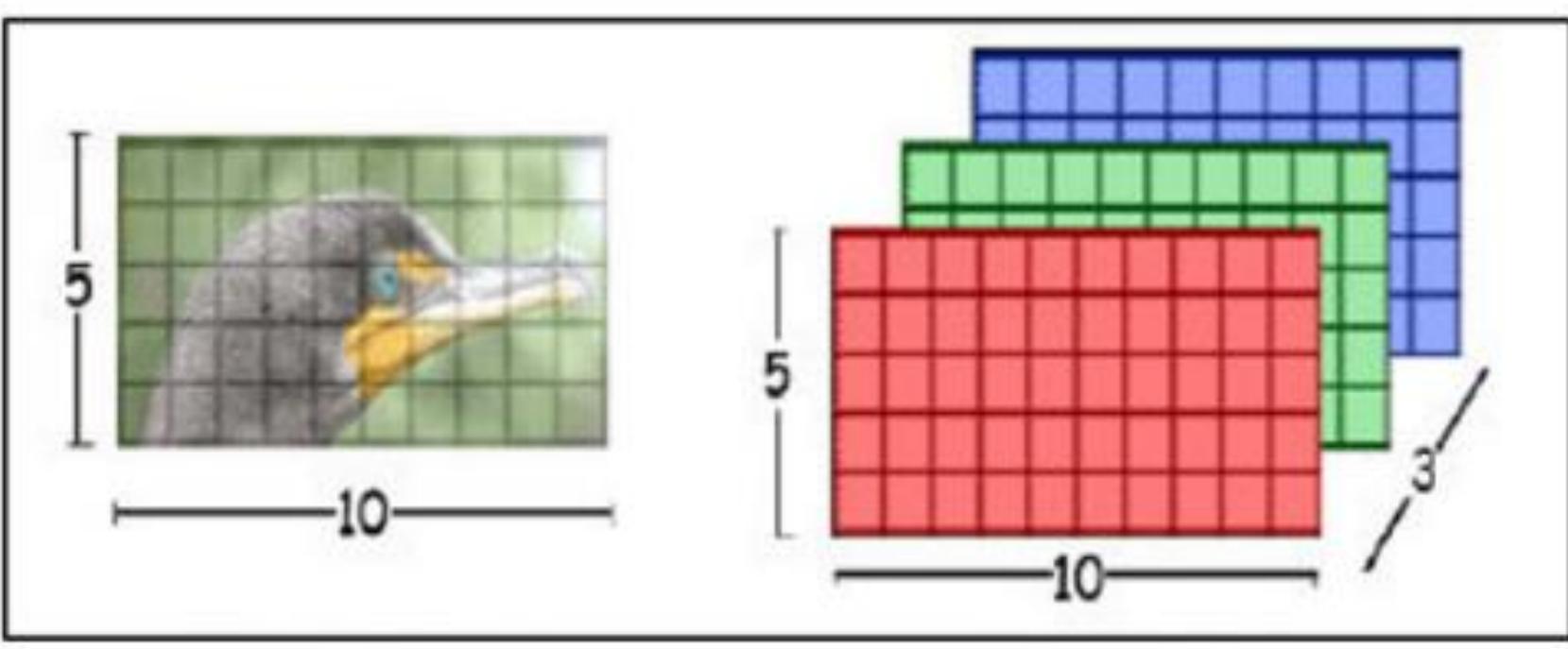
El sector de la videovigilancia comienza a usar en masa algoritmos clásicos de detección de personas y segmentación de espacios.



Visión Artificial

En 2012, cuando gracias al empleo de las GPUs de NVIDIA para ejecutar Redes Neuronales Convolucionales, que explotaría esta disciplina. A partir de este momento, el empleo de redes neuronales de aprendizaje profundo se ha extendido por todos los sectores que emplean visión artificial, y cientos de miles de empresas de todos los tamaños se han especializado en esta tecnología.

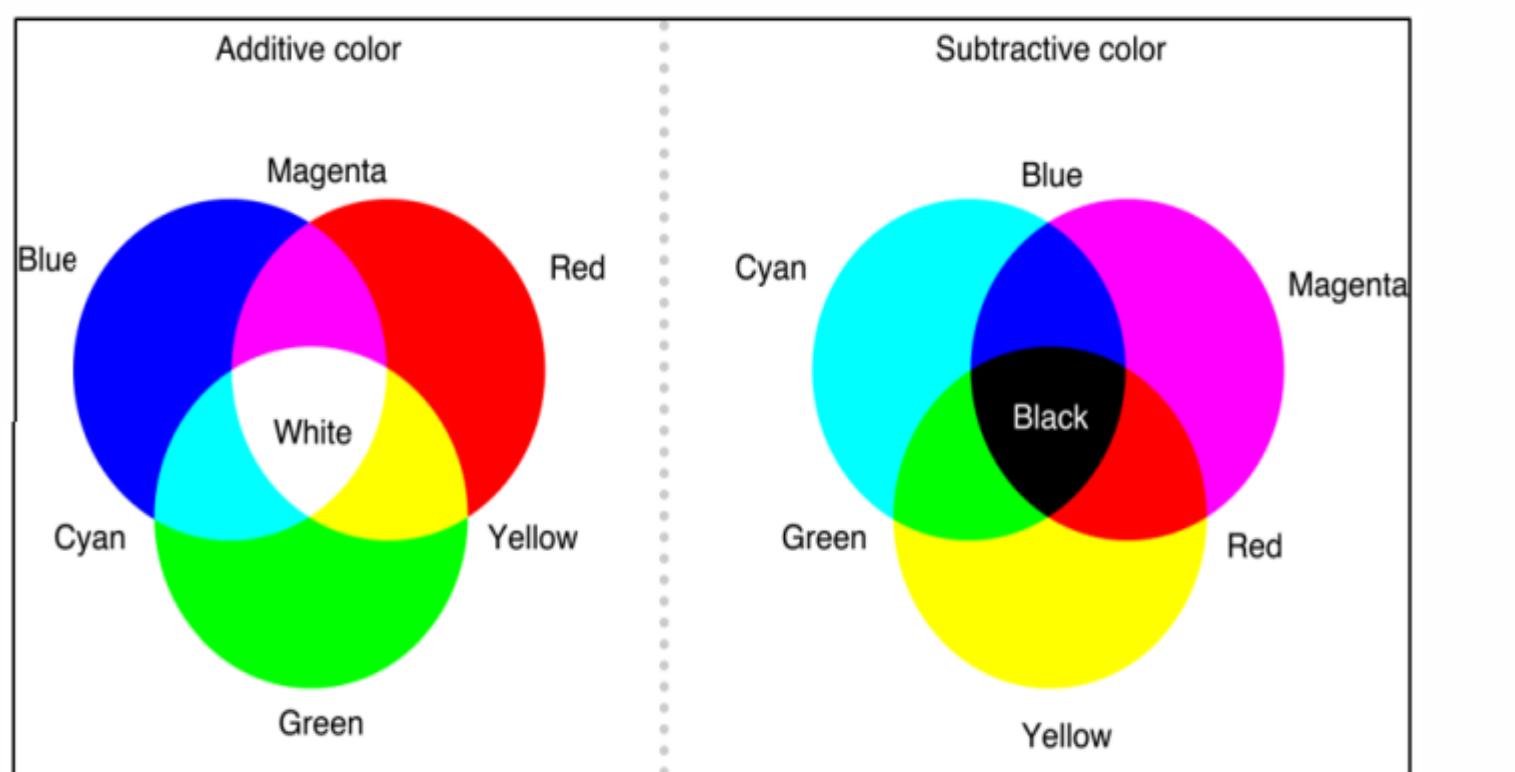
Visión Artificial



Profesor: Víctor Viera Balanta



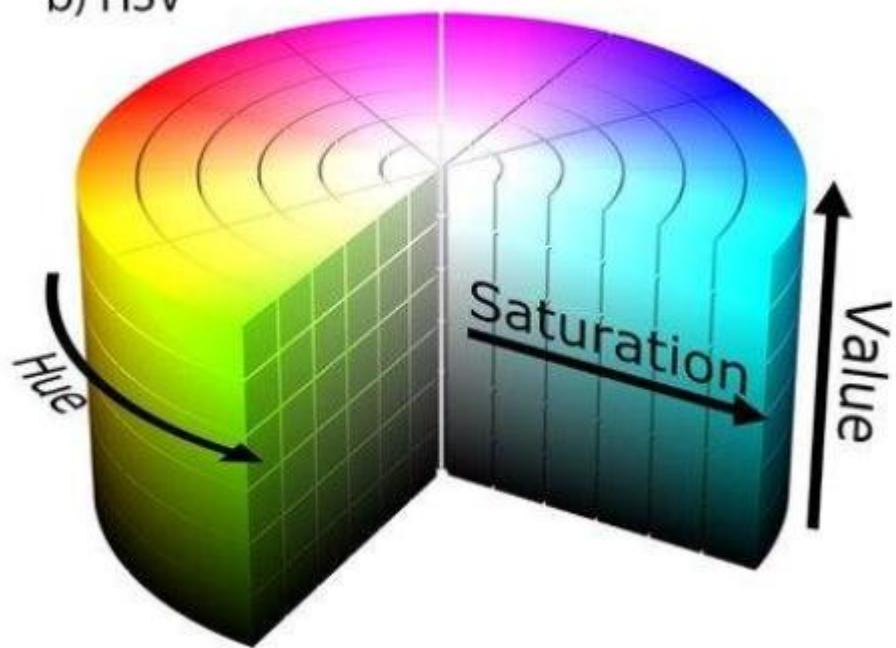
Visión Artificial



Comparación entre los colores primarios y su unión en los colores luz (izq) y pigmento (dcha.)

Visión Artificial -humano

b) HSV



Federico Alejandro Vazquez Sarauollo
copyright

Profesor: Víctor Viera Balanta

valor hue (H), es el responsable del código de color que observamos. 1 grado a 360 grados. 360 posibles valores de color.

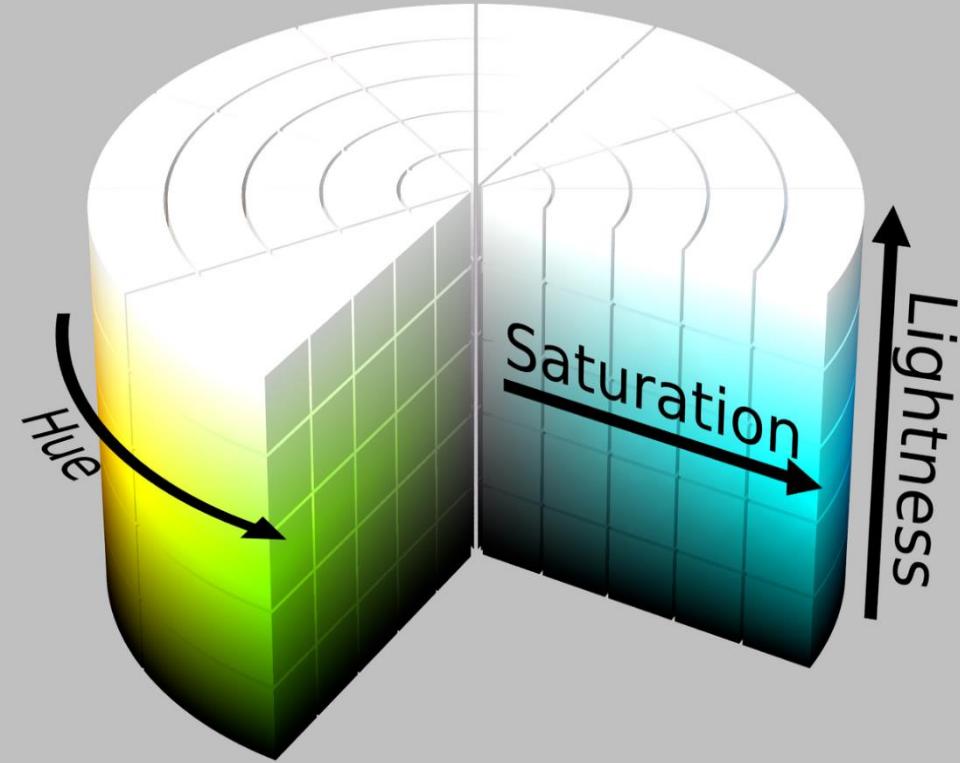
Los rojos se encuentran entre 0 y 60 grados

- Los amarillos se encuentran entre 61 y 120 grados
- Los verdes se sitúan entre 121 y 180 grados
- Los cianes están entre 181 y 240 grados
- Los azules están entre 241 y 300 grados
- magenta están entre 301 y 360 grados

(S) Saturación, cantidad e Gris (0 a 100, 100 sin gris)

(V) Value la intensidad con la que se presenta el color. varía entre 0 y 100, 0 (negro)

Visión Artificial-humano



valor hue (H), es el responsable del código de color que observamos. 1 grado a 360 grados. 360 posibles valores de color.

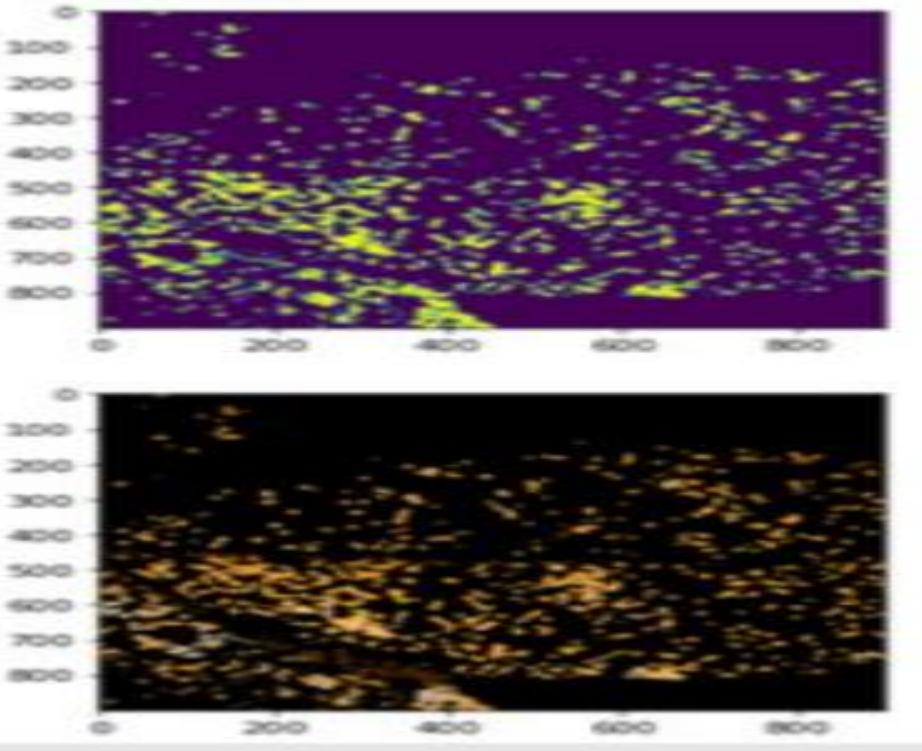
Los rojos se encuentran entre 0 y 60 grados

- Los amarillos se encuentran entre 61 y 120 grados
- Los verdes se sitúan entre 121 y 180 grados
- Los cianes están entre 181 y 240 grados
- Los azules están entre 241 y 300 grados
- magenta están entre 301 y 360 grados

(S) Saturación, cantidad e Gris (0 a 100, 100 sin gris)

(L)Luminosidad 100, 100 (totalmente blanco)

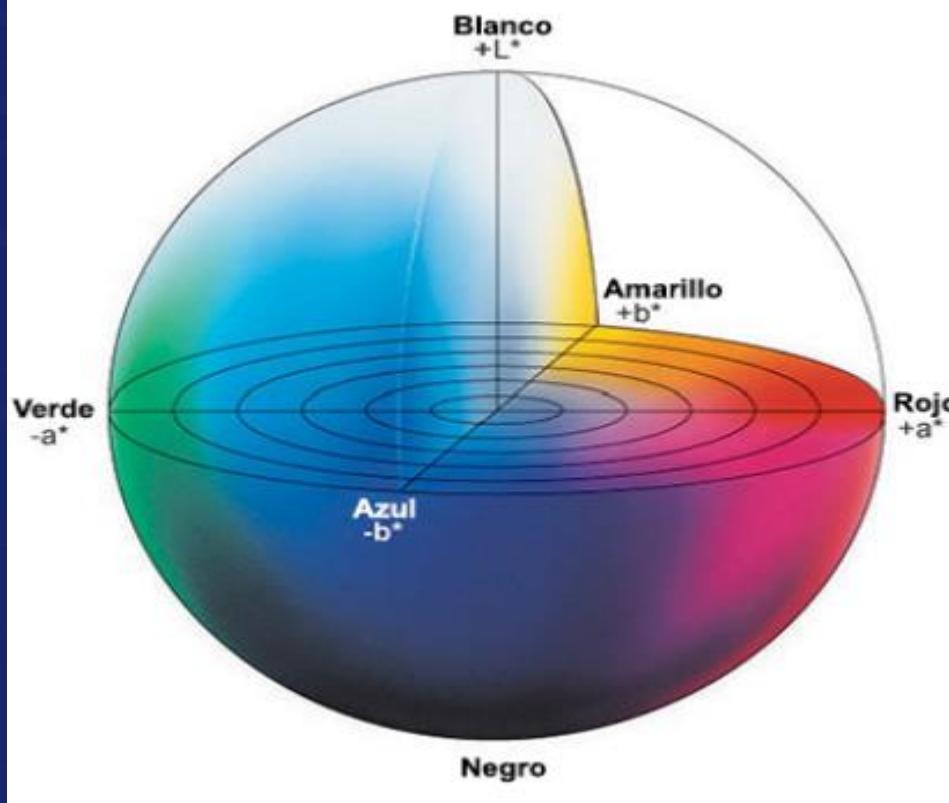
Visión Artificial-humano



Detección de enfermedades en plantas. Por el color de las manchas.

Visión Artificial

El Espacio de Color CIE L*A*B*

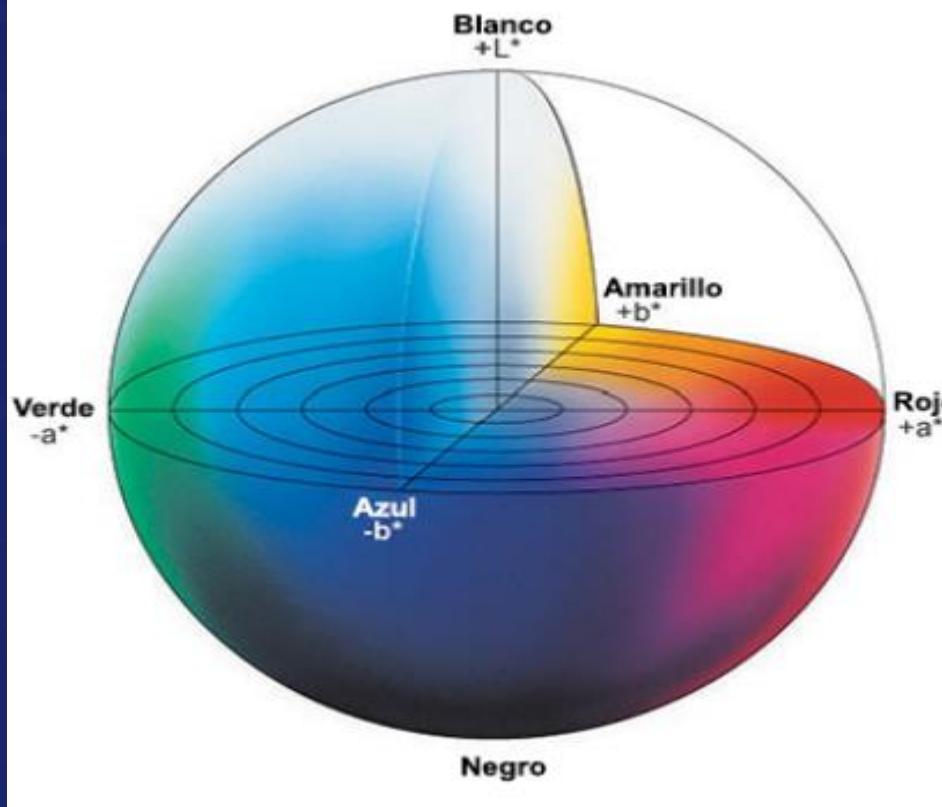


Espacio de color será el último de los que analizaremos y trabajaremos en esta asignatura. LAB surge en 1940 de mano de Richard Hunter, tratando de emular la percepción humana del color. Fue el organismo CIE (Comisión Internacional de la Iluminación)

quién lo reeditó treinta años más tarde y por el que recibe la denominación de CIELAB.

Visión Artificial

El Espacio de Color CIE L*A*B*



Profesor: Víctor Viera Balanta

L, Lightness. La luminosidad con valores entre el 0 (sin luz, o negro absoluto) y 100 (la mayor luz disponible, es decir, blanco puro)

a, es el valor de rojo/verde. Este valor va de -128 hasta 127 (para completar los 255 posibles valores). Los valores negativos de a indican tonos verdes y los valores positivos indican tonos rojos.

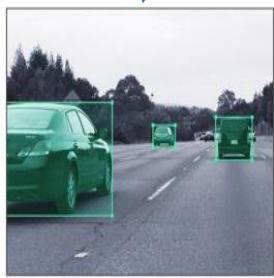
b, es el valor de azul/amarillo. Se comporta igual que el canal a, siendo los valores negativos colores azules y los valores positivos colores amarillos.

Visión Artificial

Clasificación

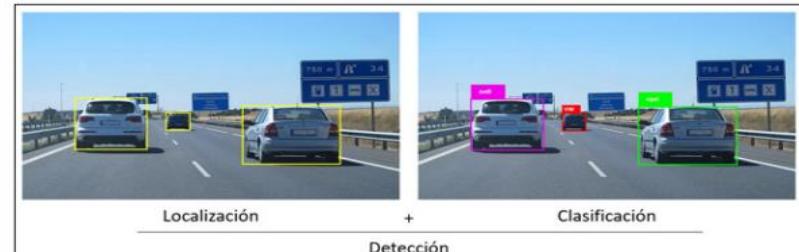


Localización



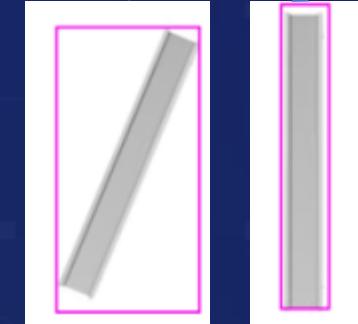
Detección

Clasificación



Localización

Segmentación



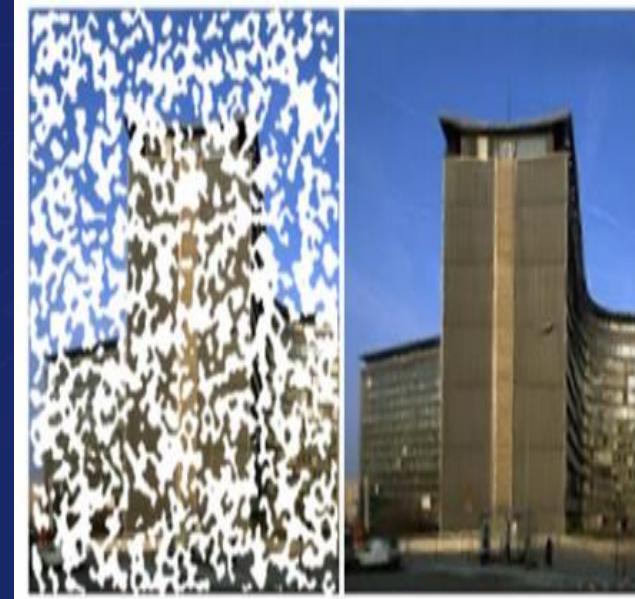
Visión Artificial

Descripción de imágenes



"man in black shirt is playing guitar."

Restauración de imágenes



Visión Artificial

Análisis de Movimiento



Reconocimiento Facial

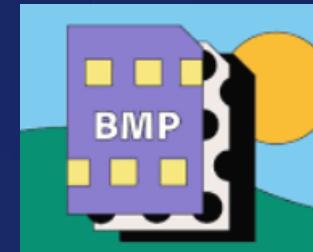


Visión Artificial

Formato de imagen sin comprimir

RAW. Es el formato con el que capturan imágenes las cámaras digitales. No tiene ningún tipo de compresión, lo que implica que las imágenes ocupan más que en ningún otro método.

BMP. Este formato sin compresión de Microsoft permite almacenar la información sin comprimir de cualquier imagen de forma estandarizada.



Visión Artificial



Formato de imagen con compresión

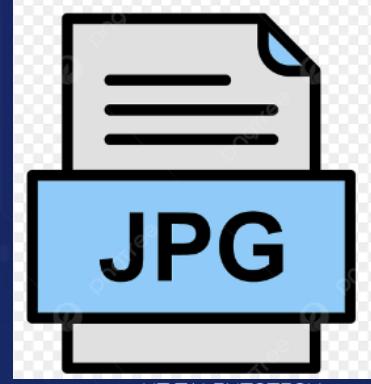
PNG. Portable Network Graphics es un formato que permite mantener los mismos 16 millones de colores una vez descomprimida la imagen, por lo que esta no pierde calidad. Además, añade un cuarto canal al formato RGB para permitir indicar **transparencias**, por lo que pueden llegar a ser imágenes muy pesadas.



Visión Artificial

Formato de imagen con compresión

JPG. Este formato estándar así llamado por las siglas de **Joint Photographic Experts Group** realiza una gran compresión. Convierte el conjunto de pixeles de la imagen en una serie de bloques para componer un mosaico, que a la hora de descomprimirse puede mostrar pérdidas de calidad.



UT TALENTOTECH

Visión Artificial

Formato de imagen con compresión

TIFF. Este formato con una baja compresión fue diseñado para almacenar los artes finales previos a la impresión de diversos elementos gráficos, y tiene como gran ventaja la capacidad de almacenar capas. Es ampliamente usado en el mundo del diseño gráfico y de la impresión, pero no así en la captura de imagen y video digital orientados a sistemas de visión artificial.



Visión Artificial

GIF: Este formato de imagen ofrece 256 colores, o lo que es lo mismo, trata de comprimir los tres canales RGB en un solo canal de 8 bits. En caso de que la imagen tenga menos de 256 colores, la calidad de esa no se verá afectada. En caso contrario, perderemos información que puede ser muy valiosa. Este formato es tan usado en animaciones porque el color es menos relevante en ellas, y permite almacenar todos los fotogramas necesarios en un formato de archivo muy poco pesado.



UT TALENTOTECH



TIC

Visión Artificial

Profesor: Víctor Viera Balanta



tecnal:
colombia



TIC

¡Gracias!

UT TALENTOTECH

