

Visión por Computador basado en Aprendizaje Profundo

Introducción al curso

Mateo Tobón Henao
mtobonh@unal.edu.co

Juan Carlos Aguirre Arango
jucaguirrear@unal.edu.co



Universidad Nacional de Colombia - Sede Manizales
Facultad de Ingeniería y Arquitectura
Departamento de Ingeniería Eléctrica, Electrónica y Computación

12 de agosto de 2022

Tabla de Contenido

- 1 Visión por Computador: ¿Qué y para qué?
- 2 Enfoque del Curso
- 3 Estructura del Curso
- 4 Recursos y Referencias



Tabla de Contenido

1 Visión por Computador: ¿Qué y para qué?

2 Enfoque del Curso

3 Estructura del Curso

4 Recursos y Referencias



¿Qué es Visión por Computador?



¿Dónde están los computadores portátiles?, ¿Dónde está la botella de agua?

créditos a: daryl mitchell, Licencia [CC BY-SA 2.0](#)



¿Qué está mal en esta imagen?

créditos a: [Max Pixel](#)



¿Qué es Visión por Computador?



¿Dónde están los computadores portátiles?, ¿Dónde está la botella de agua?

créditos a: daryl mitchell, Licencia [CC BY-SA 2.0](#)



¿Qué está mal en esta imagen?

créditos a: [Max Pixel](#)

¿Puede un computador responder estas preguntas?

¿Qué es la Visión por Computador?

Visión por Computador

La visión por computador (VC) es un campo de la inteligencia artificial (IA) que permite a los computadores y sistemas extraer información relevante de imágenes digitales, videos y otras entradas visuales; con el fin de ver, observar y comprender estas del mismo modo que los seres humanos.



¿Qué es la Visión por Computador?

Visión por Computador

La visión por computador (VC) es un campo de la inteligencia artificial (IA) que permite a los computadores y sistemas extraer información relevante de imágenes digitales, videos y otras entradas visuales; con el fin de ver, observar y comprender estas del mismo modo que los seres humanos.

Otras Definiciones

- “Construcción de descripciones explícitas y significativas de objetos físicos a partir de imágenes” [Ballard and Brown, 1982]
- “Propiedades computacionales del mundo 3D a partir de una o más imágenes digitales” [Trucco and Verri, 1998]
- “Tomar decisiones útiles a partir de imágenes que capturen objetos y escenas del mundo real” [Stockman and Shapiro, 2001]

Aplicaciones de VC



(a) Conducción Autónoma [Tesla, 2016]



(b) Analítica de Deportes [Langhee, 2020]



(c) Análisis de Imágenes Médicas
[Intel, 2022]



(d) Agricultura [AI, 2021]



Más Aplicaciones de VC

- Retail ([Amazon Go, Virtual Try-on](#))
- Cuidado de la salud ([Detector de pérdida de sangre, DermLens](#))
- Agricultura ([SlantRange, Reconocimiento Facial de Ganadería](#))
- Censado Remoto ([Uso de la tierra](#))
- Realidad Aumentada ([Gestión de almacenes y empresas](#))
- Monitoreo de Estructuras ([Inspección de puentes y reconstrucción 3D con drones](#))



Tabla de Contenido

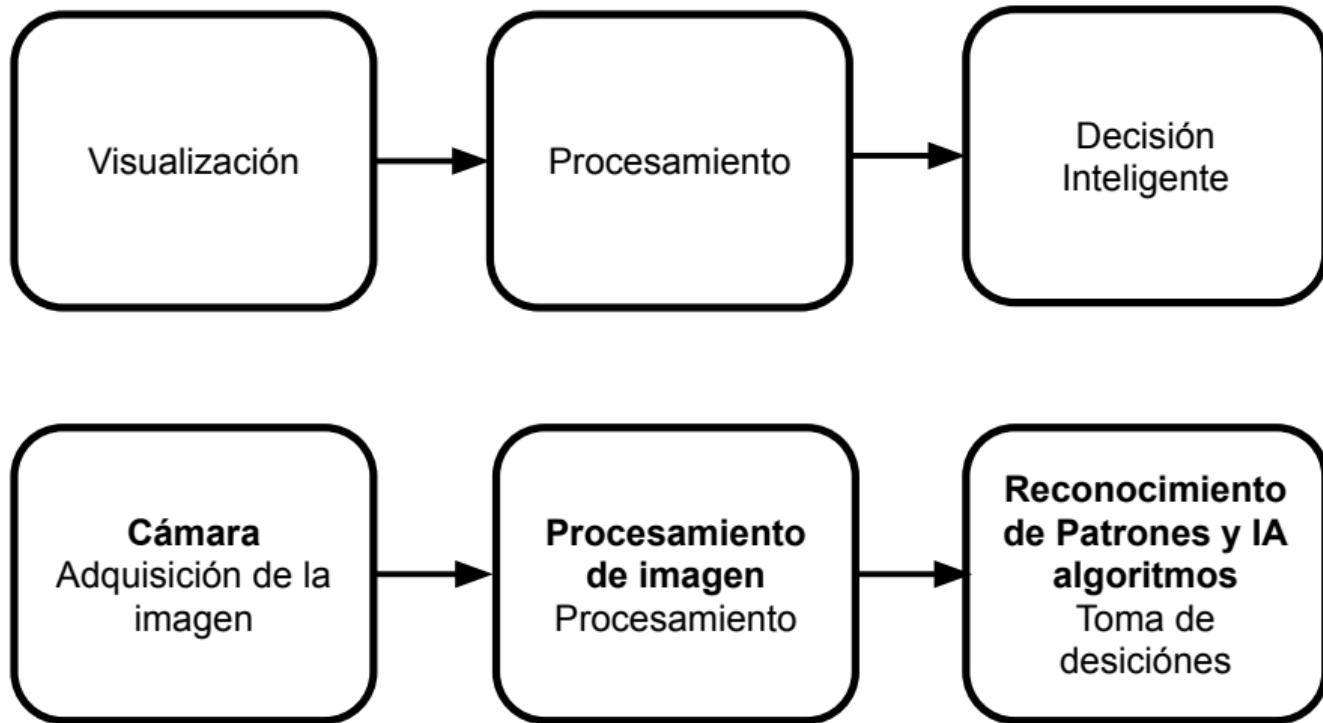
1 Visión por Computador: ¿Qué y para qué?

2 Enfoque del Curso

3 Estructura del Curso

4 Recursos y Referencias

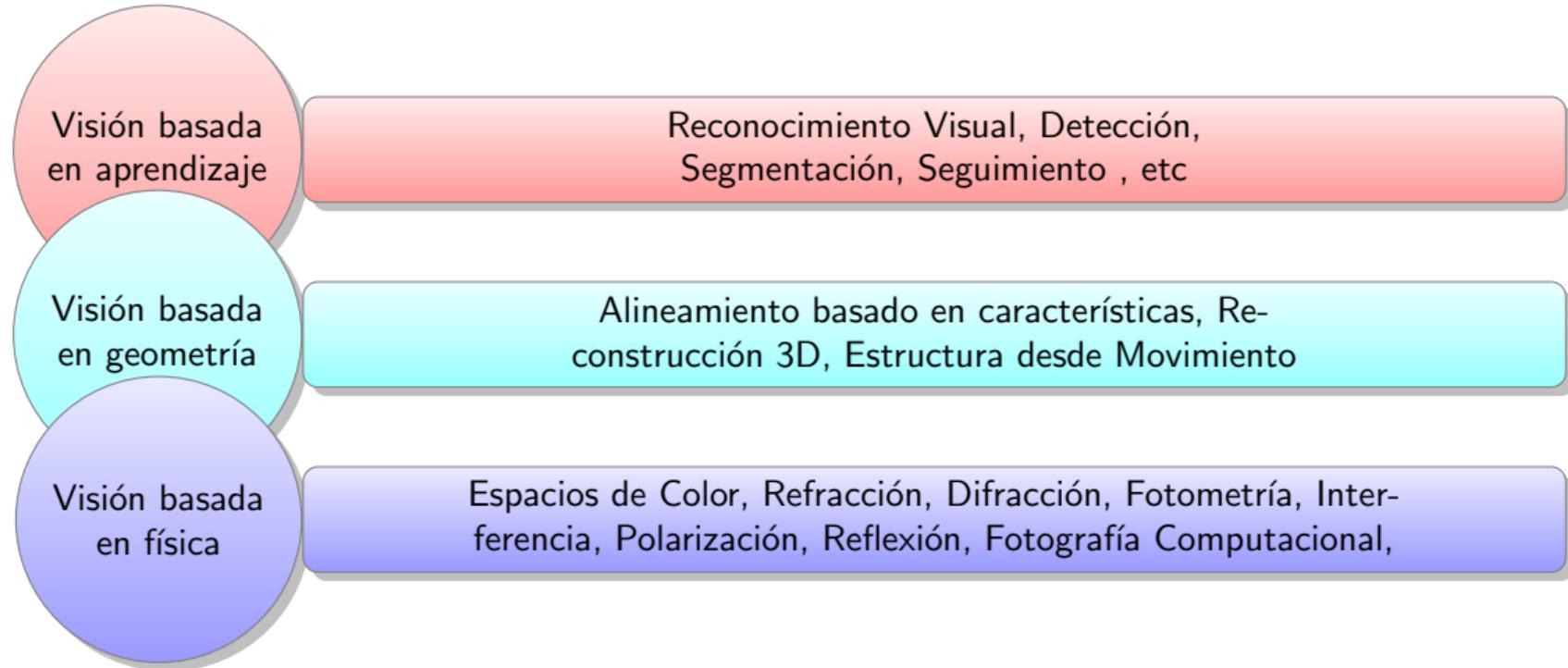




- Visión por Computador
 - Entrada: Imagen
 - Salida: Interpretación
- Procesamiento Digital de Imágenes
 - Entrada: Imagen
 - Salida: Imagen
- Computación Gráfica
 - Entrada: Modelo
 - Salida: Imagen



Tópicos



Esquema General I

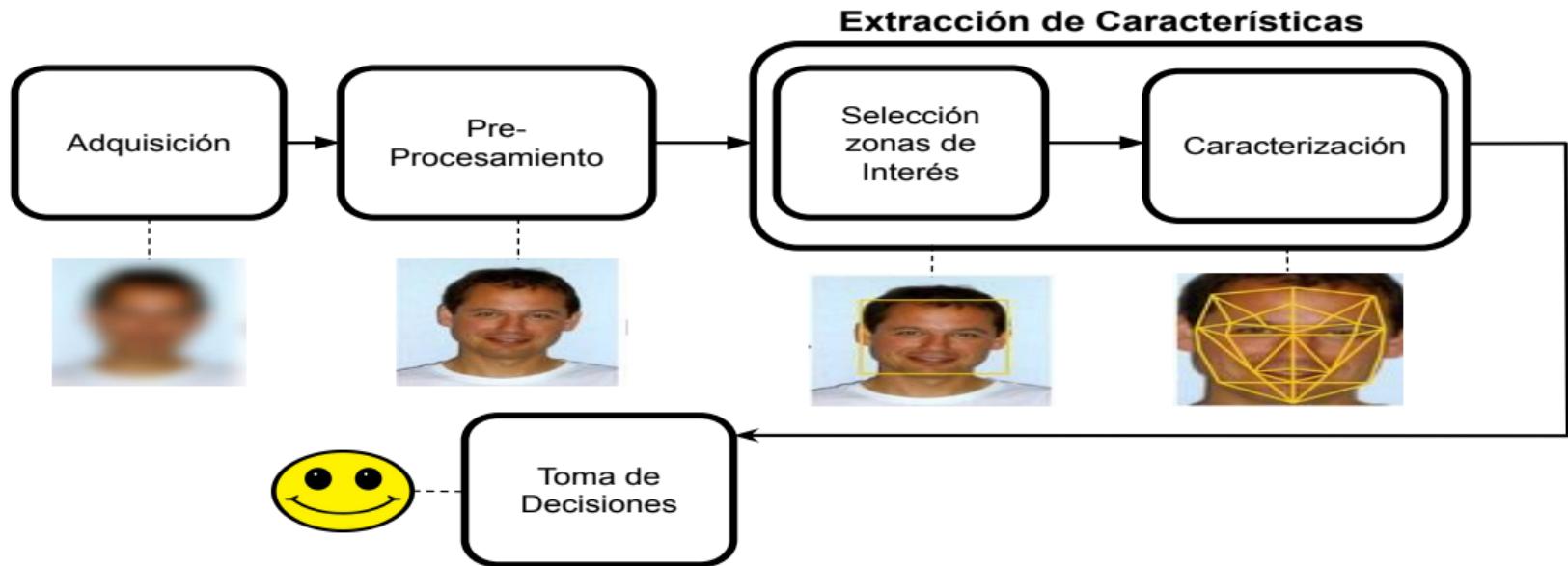


Figura: Clásico

Esquema General II

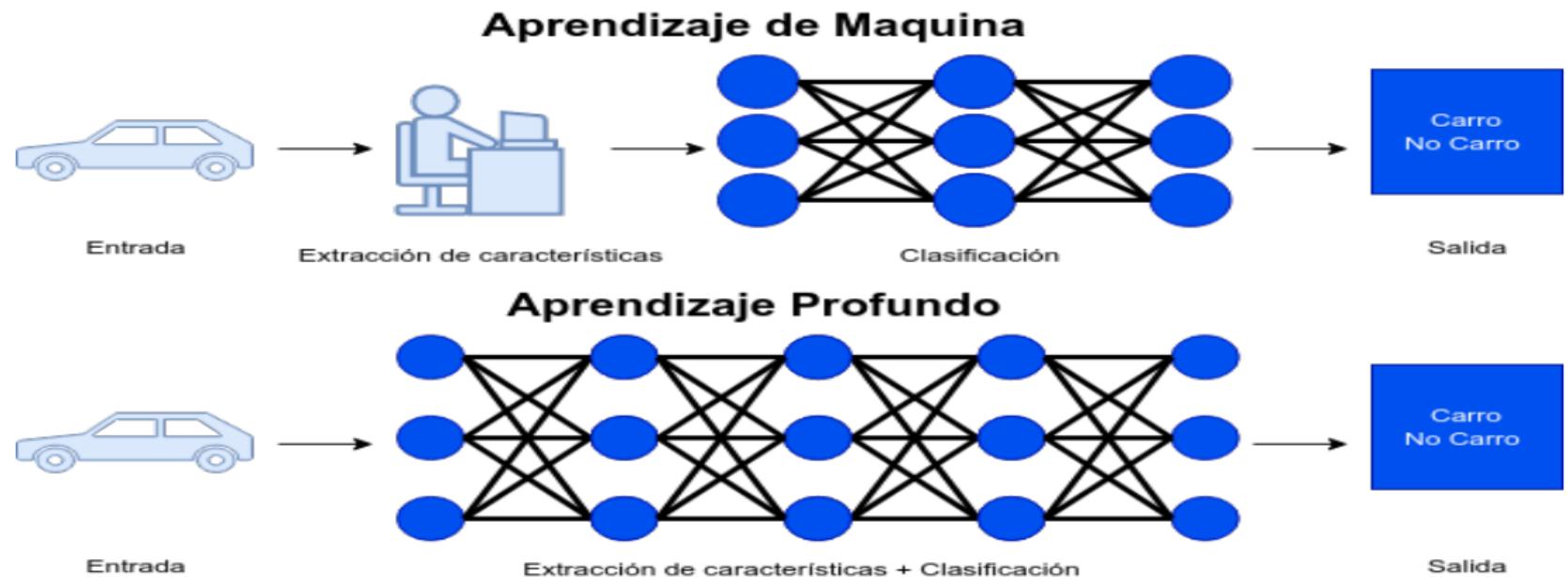
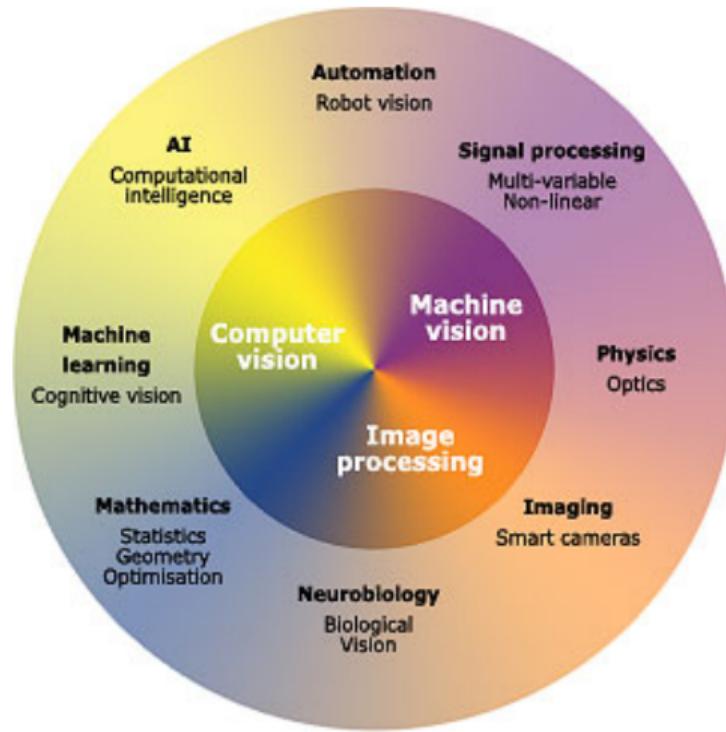


Figura: Aprendizaje Profundo



Disciplinas Relacionadas con VC



¿Por qué es difícil? I



(a) Varianza intra clase
Créditos: [Pexels](#), Licencia CC0



(b) Occlusiones
Créditos: [Pexels](#), Licencia CC0

¿Por qué es difícil? II



(c) Iluminación [Szeliski, 2022]



(d) Diferentes Escalas
Créditos: [Pipsels](#), Licencia CC0

¿Por qué es difícil? III



(e) Estructuras no Rígidas
Créditos: [Pixnio](#), Licencia CC0



(f) Diferentes Perspectivas
Créditos: [PublicDomainPictures.net](#), Licencia CC0 Dominio Público



Tabla de Contenido

1 Visión por Computador: ¿Qué y para qué?

2 Enfoque del Curso

3 Estructura del Curso

4 Recursos y Referencias



Contenido del curso I

Segmento 1: Conceptos Preliminares

- Introducción a la visión por computador
- Repaso programación orientada a objetos
- Repaso notación matemática, librerías de Numpy y Matplotlib

Segmento 2: Conceptos básicos de imágenes

- Formación de imágenes: Modelo de cámara y lentes
- Representación de imágenes en la computadora y espacios de color
- Filtros morfológicos
- Registro de imágenes



Contenido del curso II

Segmento 3: Conceptos básicos de redes neuronales

- Contextualización: Inteligencia artificial y visión por computador
- Perceptrón
- Perceptrón multicapa
- Fundamentos de Tensorflow y fuentes de datos
- Propagación hacia atrás y gradiente descendiente

Segmento 4: Teoría de redes neuronales convolucionales

- Introducción a las redes convolucionales
- Entrenamiento de redes convolucionales
- Técnicas de regularización
- Transferencia de aprendizaje
- Diseño de redes neuronales convolucionales



Segmento 5: Tareas clásicas de VC

- Face ID y Face verification
- Preparando un modelo para producción TFlite y TensorRT
- End-to-end Computer Vision problem: Ejemplo caso práctico

Segmento 6: Despliegue de modelos en sistemas embebidos

- Introducción a sistemas embebidos
- Configuración de Jetson Nano y Raspberry Pi
- Preparando un modelo para producción TFlite y TensorRT
- End-to-end Computer Vision Problem: Ejemplo caso práctico



Teórico

- Probabilidad
- Estadística
- Álgebra Lineal
- Conocimientos básicos de aprendizaje automático y profundo



Prerrequisitos

Teórico

- Probabilidad
- Estadística
- Álgebra Lineal
- Conocimientos básicos de aprendizaje automático y profundo

Práctico

- Programación estructurada y orientada a objetos
- Python



Actividades	Porcentaje (%)
Entrega 1	10 %
Entrega 2	20 %
Entrega 3	40 %
Entrega 4	30 %



Leer las aplicaciones adicionales en la dispositiva 7.



Tabla de Contenido

1 Visión por Computador: ¿Qué y para qué?

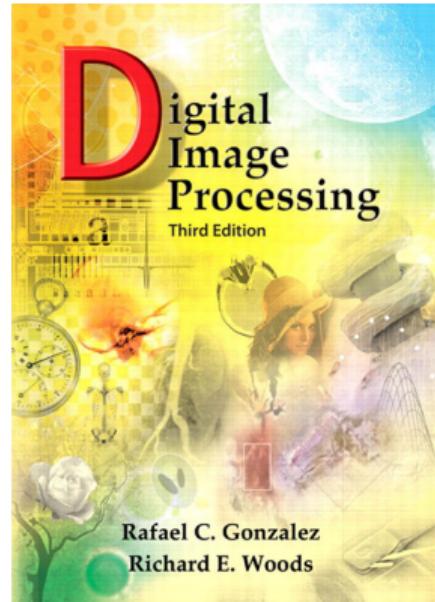
2 Enfoque del Curso

3 Estructura del Curso

4 Recursos y Referencias



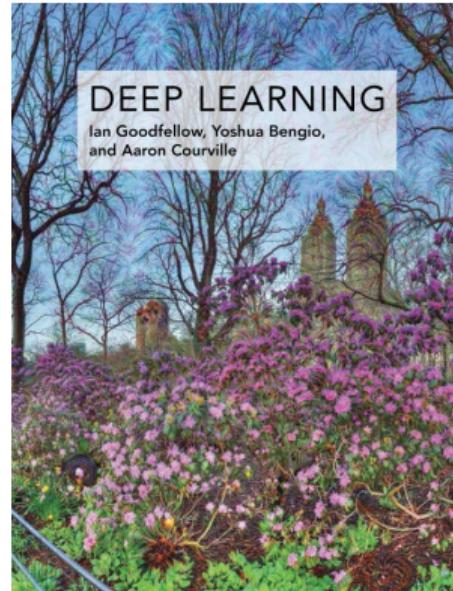
Recursos I



(a) Link

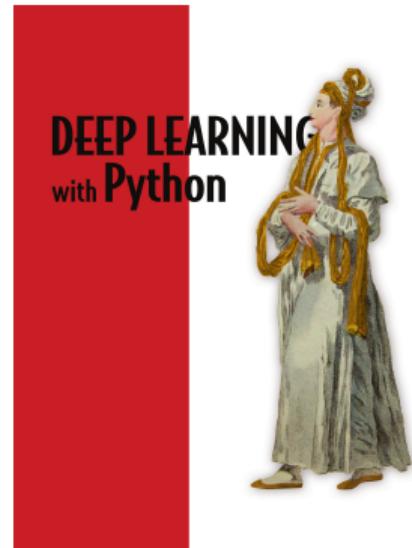


Recursos II

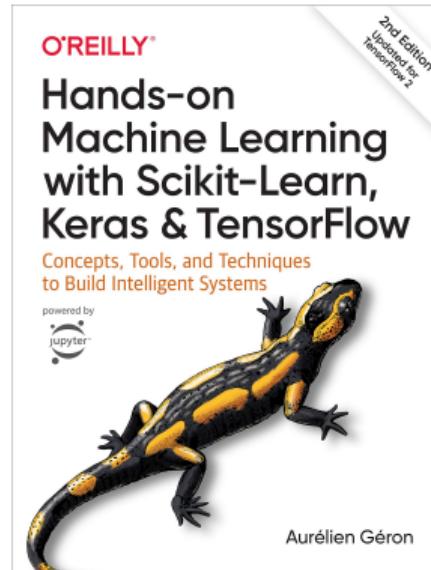


(b) Link

Recursos III



(c) Link



(d) Link



(e) NPTEL



Recursos VI



(f) José Pablo Alvarado Moya



Bibliografía I

-  AI, C. (2021).
Drone AI: Counting Animals With Computer Vision.
YouTube.
-  Ballard, D. H. and Brown, C. M. (1982).
Computer Vision.
Prentice Hall, New Jersey, USA.
-  Intel (2022).
Computer Vision in Healthcare.
-  Langhee, A. (2020).
Football Analytics — Deep Learning — Computer Vision — Artificial Intelligence.
YouTube.



Bibliografía II

-  Stockman, G. and Shapiro, L. G. (2001).
Computer vision.
Prentice Hall PTR.
-  Szeliski, R. (2022).
Computer Vision - Algorithms and Applications, Second Edition.
Texts in Computer Science. Springer.
-  Tesla (2016).
Autopilot full self-driving hardware (neighborhood long).
-  Trucco, E. and Verri, A. (1998).
Introductory techniques for 3-D computer vision, volume 201.
Prentice hall Englewood Cliffs.

