

# Práctica 2 de Visión Artificial

Grado en Enxeñería Informática – Computación

Curso 23/24

Esta práctica consiste en aplicar los conocimientos adquiridos sobre técnicas de procesado de imagen para desarrollar una metodología que permita abordar uno de los siete problemas de visión artificial planteados.

- Las práctica deberá realizarse de forma **individual**.
- Se utilizará uno los siguientes lenguajes, o entornos de desarrollo, a elección del alumno:
  1. **Python 3 + NumPy/SciPy**. Acompañado de librerías de procesado de imágenes (e.g. scikit-image, OpenCV etc.). Esta es la opción recomendada.
  2. **Matlab, Octave** o herramientas similares. Acompañado de los toolboxes y librerías de procesado de imagen que se consideren convenientes.
- Está permitido el uso de cualquier técnica de visión por computador disponible que permita abordar el problema, siempre que se demuestre el conocimiento de la técnica utilizada y se justifique el motivo por el que se usa.

## Entrega y Evaluación

- Para poder superar esta práctica es necesario **entregar**:
  1. Software de procesado de imágenes desarrollado que permita ver resultados parciales.
  2. Ejemplos de resultados precalculados (solo si el software es muy lento).
  3. Pequeño documento de memoria (1-2 folios), que indique lo siguiente, a modo de guion:
    - Esquema global del método. Incluyendo los pasos (o subproblemas) identificados que permiten llegar a la solución propuesta.
    - Las soluciones aportadas a cada uno de estos subproblemas, incluyendo:
      - Metodologías consideradas, independientemente de que hayan sido descartadas o seleccionadas para formar parte de la solución final.
      - Motivos por los que se han considerado cada una de estas metodologías.
      - Evaluación cuantitativa y/o cualitativa de los resultados de cada solución.
      - Identificación de problemas y posibles mejoras.
- Los alumnos deberán **entregar y defender la práctica antes de la fecha límite, aún por fijar**.
  - Todos los entregables anteriormente mencionados deben subirse al moodle.

- Se concertara una cita de defensa de prácticas con alguno de los profesores de la asignatura en la que se mostrará el software de la práctica funcionando en un portátil, junto con el documento de memoria.
  - Los alumnos que presenten prácticas con indicios de plagio obtendrán una calificación de Suspenso.
- En la evaluación de la práctica se tendrán en cuenta:
- Que el alumno ha sido capaz de identificar un conjunto de funcionalidades parciales (pasos) y aportar metodologías para solucionar estas funcionalidades que, en conjunto, puedan aportar una solución al problema planteado.
  - Que el alumno ha sido capaz de explorar técnicas que puedan tener sentido (a priori), comprender su funcionamiento práctico, y evaluarlas para aportar soluciones a los subproblemas identificados, hasta dar con la solución final. En este sentido, probar varias técnicas alternativas sin justificación, no implica mejor nota.
  - No se busca solucionar el problema al 100 %, sino aplicar los conocimientos de la asignatura a un caso práctico real. Así, conseguir mejor o peor resultado en algunas o todas las imágenes no implica por sí solo ni mejor ni peor nota.

## Trabajos

Uno de los siguientes trabajos será asignado por el profesor de prácticas a cada alumno. Las siguientes secciones contienen una descripción detallada de cada uno de ellos.

- 1.- Segmentación y mejora de documentos
- 2.- Detección de campo, balón y jugadores de fútbol
- 3.- Segmentación de iris, pupila y esclerótica
- 4.- Detección y reconocimiento de señales de tráfico
- 5.- Segmentación y evaluación de crecimiento de plantones
- 6.- Segmentación de pulmones en imagen CT torácica
- 7.- Segmentación de disco óptico y copa para diagnóstico de glaucoma en retinografías

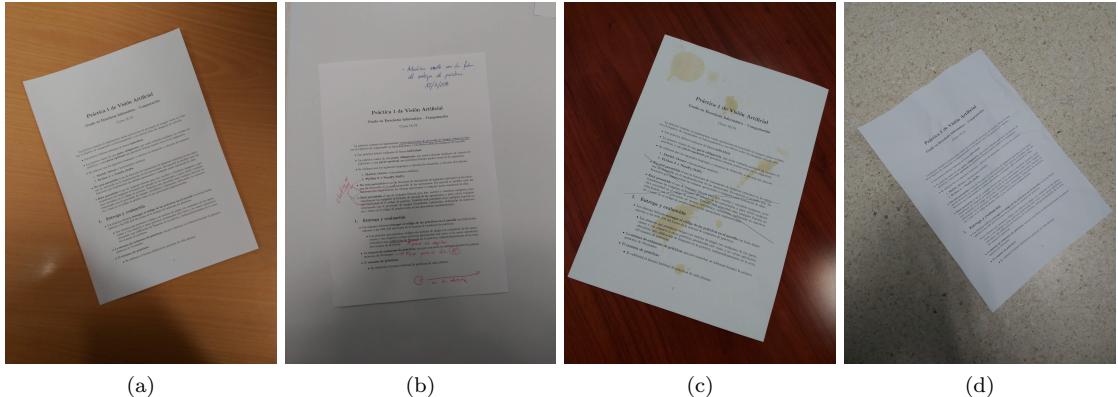
# 1. Segmentación y mejora de documentos

El uso generalizado de teléfonos móviles inteligentes equipados con cámara motiva la creación de herramientas inteligentes de utilidad, que puedan hacer uso de metodologías de análisis de imagen. Un ejemplo de este tipo de herramientas es el uso de la cámara como escáner de documentos.

## 1.1. Materiales

Se proporcionan un conjunto de fotografías en formato JPG tomadas con un teléfono móvil convencional. El alumno puede tomar nuevas fotografías con cualquier móvil si se considera conveniente.

- Cada fotografía contiene una única hoja de documento sobre un fondo uniforme.
- Se proporcionan múltiples fotografías de hojas con marcas de bolígrafo y manchas de café sobre diversos fondos. La siguiente figura muestra algún ejemplo.



## 1.2. Objetivos

Desarrollar un método computacional automático que, a partir de una de estas imágenes, permita:

1. Detectar las esquinas y los contornos del documento. Para ello se asumirá que la hoja tiene forma trapezoidal.
2. Detectar el texto en negro del documento, obviando las manchas y marcas de bolígrafo.
3. Montar un documento reconstruido de fondo blanco y texto en negro uniforme, sin las manchas y marcas de bolígrafo.

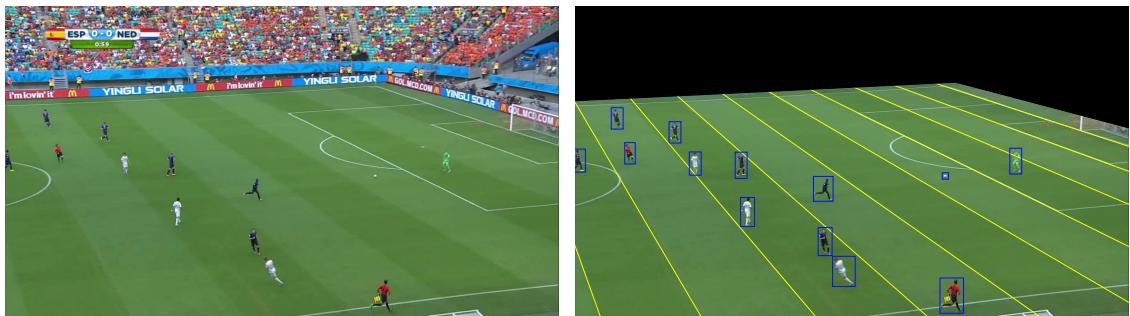
## 2. Detección de campo, balón y jugadores de fútbol

Es posible desarrollar sistemas automáticos de asistencia a la decisión arbitral de diversos deportes que, mediante el análisis de videos deportivos sean capaces de registrar la posición de los jugadores de ambos equipos y el balón, con respecto al terreno de juego, para así proporcionar mediciones e información objetiva sobre el partido. Esta información es relevante, por ejemplo, para extraer estadísticas del partido, o para asistir en decisiones arbitrales. En este trabajo nos centraremos en detectar algunos objetos relevantes en imágenes de partidos de fútbol.

### 2.1. Materiales

Para este trabajo se utilizará un subconjunto de imágenes de partidos de fútbol, de las disponibles en [1]. El subconjunto de imágenes seleccionado está disponible en el moodle, aunque adicionalmente se pueden descargar más imágenes de la web original si se considera conveniente.

- Cada imagen consiste en un fotograma de un partido de fútbol, donde se visualiza parte del terreno de juego, junto con algunos jugadores, algunos árbitros y la pelota.
- Se proporcionan 4 fotogramas de 5 partidos diferentes (20 imágenes). La siguiente figura muestra un ejemplo:



### 2.2. Objetivos

Desarrollar un método computacional que permita, a partir de una imagen dada:

1. Detectar el área ocupada por el terreno de juego (césped)
2. Localizar las líneas de siega del césped, indicativas de la perspectiva.
3. Detectar cada uno de los objetos presentes en el terreno de juego: jugadores, árbitros y balón.  
No es necesario reconocer qué es qué.

### 2.3. Referencias

- [1] Página personal de Namdar Homayounfar.  
<http://www.cs.toronto.edu/~namdar/#publications>

### **3. Segmentación de iris, pupila y esclerótica**

Los patrones biométricos son de especial relevancia en aplicaciones de control de acceso e identificación de personas. Uno de los patrones biométricos más característicos es aspecto visual del iris, ya que puede ser extraído de una imagen del rostro con alta resolución. En este trabajo nos centraremos en la segmentación de la región del iris, así como sus regiones adyacentes (pupila y esclerótica) a partir de imágenes de ojos.

#### **3.1. Materiales**

Para este trabajo se utilizará un subconjunto de imágenes provenientes del Multimedia University Iris Dataset [1]. El subconjunto de imágenes seleccionado está disponible en el moodle, aunque adicionalmente se pueden descargar más imágenes de la web original si se considera conveniente.

- Cada imagen contiene una fotografía en escala de grises del primer plano de un ojo
- Se proporcionan imágenes del ojo izquierdo y derecho para 8 individuos (16 imágenes). En la siguiente figura se muestran algunos ejemplos:



#### **3.2. Objetivos**

Desarrollar un método computacional que permita, a partir de una de estas imágenes:

1. Segmentar la región de la pupila (círculo).
2. Segmentar la región del iris (aro).
3. Segmentar la/s región/es correspondientes a la esclerótica a ambos lados del iris.
4. Identificar las áreas con posibles artefactos de las segmentaciones anteriores (e.g. brillos en la pupila o pestañas en iris y esclerótica).

#### **3.3. Referencias**

[1] MMU Iris Dataset, <http://pesona.mmu.edu.my/~ccteo/>.

Enlaces alternativos en: <http://www.cs.princeton.edu/~andyz/irisrecognition>

## 4. Detección y reconocimiento de señales de tráfico

Si quisieramos construir vehículos autónomos que pudieran utilizar la infraestructura viaria actual, tendríamos que dotar al sistema objetivo, entre otras cosas, de mecanismos de percepción de las señalizaciones presentes en la vía pública. Este trabajo se centrará en la detección y reconocimiento de un subconjunto de las señales de tráfico.

### 4.1. Materiales

Para este trabajo se utilizará un subconjunto de imágenes provenientes del German Traffic Sign Detection Benchmark [1]. El subconjunto de imágenes seleccionado está disponible en el moodle, aunque adicionalmente se pueden descargar más imágenes de la web original si se considera conveniente.

- Cada imagen contiene una escena de tráfico en ciudad o carretera, que contiene una o más señales de tráfico.
- Se proporcionan un total de 12 escenas con señales variadas. La siguiente figura muestra un ejemplo de escena con tres señales objetivo: 2 de peligro, y una de prohibición.



### 4.2. Objetivos

Desarrollar un método computacional que, a partir de una estas imágenes, permita:

1. Localizar todas las apariciones de señales de tráfico en la imagen. Solamente se consideran las señales descritas en el siguiente punto.
2. Identificar la categoría de cada señal de tráfico detectada. Las categorías contempladas son:
  - Prohibición. Rojas y circulares, “Ceda el paso” o “STOP”.
  - Peligro. Rojas y triangulares.
  - Obligación. Azules y circulares.
  - Indicación. Azules y cuadradas.

### 4.3. Referencias

- [1] The German Traffic Sign Detection Benchmark (GTSDB)  
<http://benchmark.ini.rub.de/?section=gtsdb>

## 5. Segmentación y evaluación de crecimiento de plantones

Supongamos que queremos hacer un sistema de visión artificial que ayude a los ingenieros agrónomos a estudiar el crecimiento de las plantas bajo diferentes condiciones de cuidado. La idea es poder cuantificar el tamaño total de la planta, así como el número de hojas, a partir de imágenes cenitales de las mismas.

### 5.1. Materiales

Para este trabajo usaremos un subconjunto de imágenes, disponibles en el moodle, extraídas del Aberystwyth Leaf Evaluation Dataset [1].

- Las imágenes consisten en una visión cenital de una bandeja de  $5 \times 4$  plantones de la especie *Arabis thaliana* tomadas en un instante temporal durante su crecimiento.
- Se han seleccionado un total de 6 imágenes, correspondientes a dos bandejas en tres instantes temporales diferentes.
- Se dispone, además, de las máscaras de segmentación deseadas.



### 5.2. Objetivos

Desarrollar un sistema computacional que permita, dada una de las imágenes anteriormente mencionadas:

1. Separar y segmentar cada una de las plantas, hasta un máximo de  $5 \times 4$  por bandeja.
2. Para cada una de las plantas, segmentar cada una de sus hojas por separado.
3. Proporcionar una evaluación cuantitativa adecuada de las segmentaciones 1 y 2.

### 5.3. Referencias

- Bell, Jonathan, Dee, Hannah M. (2016). Aberystwyth Leaf Evaluation Dataset [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.168158>

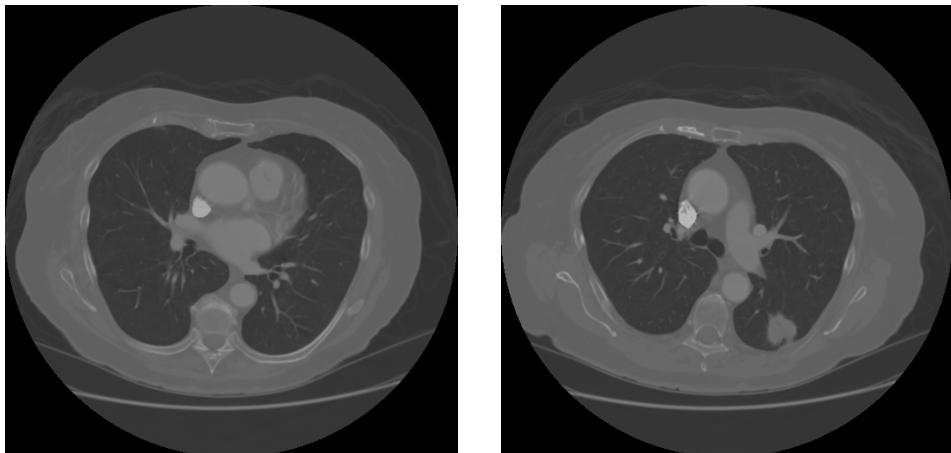
## 6. Segmentación de pulmones en imagen CT torácica

A pesar de su elevado coste, las imágenes CT (Computed Tomography) torácicas son ampliamente utilizadas en diversos procedimientos de análisis clínico, ya que constituyen la mejor manera de inspeccionar los pulmones de un paciente. Así, su uso está ampliamente extendido en el diagnóstico de patologías como el cáncer de pulmón, fibrosis pulmonar, o el ILD (Interstitial Lung Disease).

Para su estudio, el primer paso consiste en la localización y segmentación de la región de interés, que en este caso se corresponde con la región pulmonar. Así, en esta práctica abordaremos la segmentación de las regiones pulmonares en imágenes CT torácicas.

### 6.1. Materiales

Para este trabajo se utilizará un subconjunto de cortes transversales, provenientes de la colección de imágenes CT Lung Image Database Consortium (LIDC) [1], y que están disponibles en el moodle. En la siguiente figura se muestran unos ejemplos:



### 6.2. Objetivos

Desarrollar un método computacional que permita, a partir de un corte transversal de CT, identificar y segmentar las regiones correspondientes a los dos pulmones. Para ello se tiene en cuenta que se pretende extraer el contorno externo de cada pulmón, considerando que:

- Las alteraciones patológicas, como pueden ser los nódulos cancerígenos, forman parte de la región pulmonar.
- La tráquea y los bronquios principales se deben excluir de la región pulmonar.
- Los pulmones no están unidos entre sí.

### 6.3. Referencias

- Armato III, Samuel G., McLennan, Geoffrey, Bidaut, Luc, McNitt-Gray, Michael F., Meyer, Charles R., Reeves, Anthony P., ... Clarke, Laurence P. (2015). Data From LIDC-IDRI. The Cancer Imaging Archive. <http://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX>

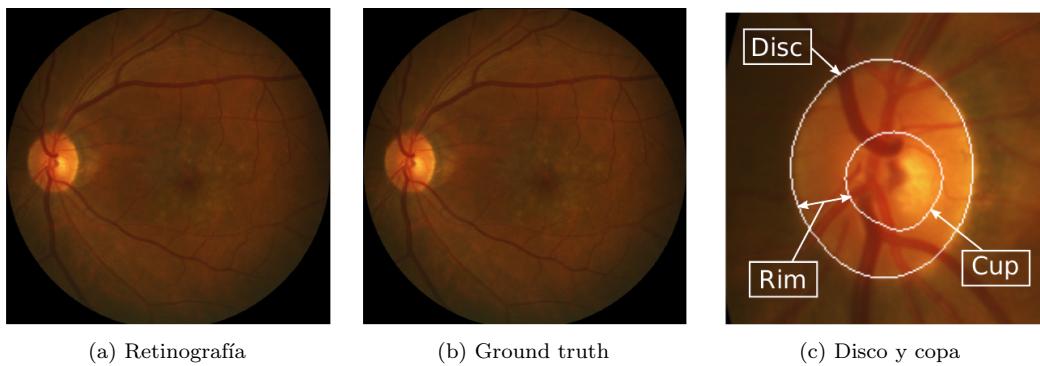
## 7. Segmentación de disco óptico y copa para diagnóstico de glaucoma en retinografías

La retinografía es una modalidad de imagen médica que permite visualizar el fondo de ojo y sus estructuras relevantes. Esto resulta útil para el diagnóstico de enfermedades oculares como el glaucoma. El glaucoma se caracteriza por el aumento de la presión intra-ocular, lo que daña diferentes tejidos de la retina. Una de las lesiones típicas es la deformación del disco óptico, formado por la copa y el anillo neuroretiniano (rim). En los pacientes con glaucoma la copa se agranda, y el rim se reduce. Por lo que la relación entre el tamaño de la copa y del disco es un buen marcador para su diagnóstico.

### 7.1. Materiales

Usaremos un subconjunto de imágenes reescaladas del Retinal Fundus Glaucoma Challenge (REFUGE) [1,2]. Estas imágenes están disponible en el Moodle, aunque adicionalmente se pueden descargar más imágenes de la web original si se considera conveniente.

- Cada imagen consiste en una retinografía a color centrada en mácula
- Se acompañan una imagen de ground truth con el disco y la copa segmentadas.



### 7.2. Objetivos

Desarrollar un método computacional que permita:

1. Detectar la posición aproximada del disco óptico (por ejemplo, el centroide), y recortar una región de interés (ROI) cuadrada que lo contenga.
2. Segmentar el disco óptico y la copa
3. Calcular el ratio copa-disco (CDR) a partir de la segmentación, usando el cociente de las alturas del disco y la copa.
4. Evaluar cuantitativamente: la localización del disco (1), la segmentación del disco y copa (2), y la medición del CDR (3), usando el ground truth.

### 7.3. Referencias

[1] REFUGE: Retinal Fundus Glaucoma Challenge  
<https://refuge.grand-challenge.org/details/>

[2] Orlando et al., “REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs”, Medical Image Analysis 59, 2020.