

27-marzo-2025:

-En el enunciado de la pregunta 01.c se agrega una frase para clarificar que depende del resultado de 01.b.

-Se cambia la numeración de la pregunta 04 para ser consistente con las preguntas anteriores.

-En la sección Entrega se deja claro que no se puede usar ChatGPT ni similares.

# Control 1

CC5213 – Recuperación de Información Multimedia

Departamento de Ciencias de la Computación

Universidad de Chile

Profesor: Juan Manuel Barrios

Fecha entrega: lunes 21 de abril de 2025

## Semana 01

Considere la siguiente imagen de 7 x 7 en escala de gris 8 bits:

	1	2	3	4	5	6	7
1	0	0	100	0	0	0	100
2	0	0	0	100	0	100	0
3	100	0	0	0	100	0	0
4	0	100	0	100	0	100	0
5	0	0	100	0	0	0	100
6	0	100	0	100	0	0	0
7	100	0	0	0	100	0	0

- a) (0.5 puntos) Escriba la imagen resultante de aplicar un filtro de mediana de 3x3 sobre la imagen **A**. Realice la operación únicamente donde el bloque de 3x3 esté completamente contenido en la imagen.

- b) (0.5 puntos) Escriba la imagen resultante de hacer convolución entre la imagen **A** y el siguiente kernel **K** de 3 x 3:

$$\mathbf{K} = \begin{array}{|c|c|c|} \hline 1 & -2 & 1 \\ \hline -2 & 4 & -2 \\ \hline 1 & -2 & 1 \\ \hline \end{array}$$

Realice la convolución únicamente donde el kernel esté completamente contenido en la imagen. Asuma que la imagen resultante tiene profundidad 32f, es decir, puede contener floats.

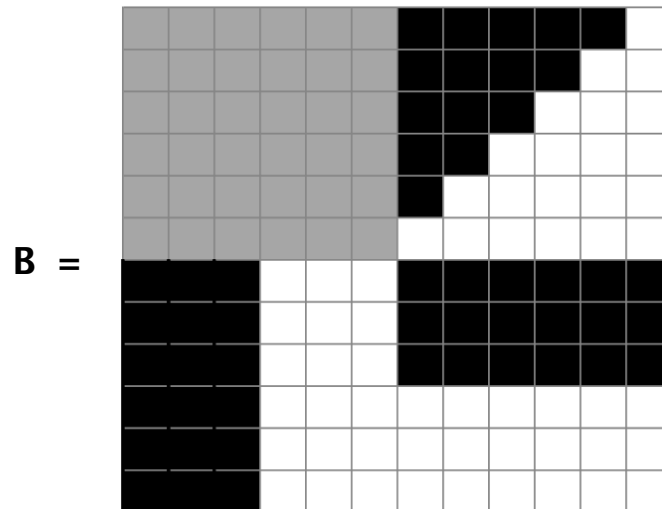
- c) (0.25 puntos) Considere la función umbral  $U_t$  que asigna blanco (255) cuando el valor absoluto de un pixel supera el valor  $t$  y negro (0) si no, es decir:

$$U_t(x) = \begin{cases} 255 & \text{si } |x| > t \\ 0 & \text{si no} \end{cases}$$

Escriba la imagen resultante de aplicar la función umbral  $t=600$  al resultado de la convolución entre la imagen **A** y el kernel **K** (que fue calculado en la parte b).

## Semana 02

Considere la siguiente imagen de 12 x 12 pixeles en escala de gris 8 bits:



Nota: los pixeles blancos tienen valor=255, los pixeles negros tienen valor=0 y los pixeles grises tienen valor=128

- a) (0.25 puntos) Dibuje el histograma de grises normalizado de 256 bins de la imagen **B**. Sea preciso en los ejes del histograma (qué valor representa cada bin y cuál es su altura).
- b) (0.25 puntos) Dibuje el histograma de grises normalizado de 256 bins de la imagen **B** usando una división de 2x2 zonas. Sea preciso en los ejes del histograma (qué valor representa cada bin y cuál es su altura).

- c) (1 punto) Dibuje el histograma de orientaciones de gradiente (HOG) de 180 bins para el rango  $-90^\circ$  a  $90^\circ$  de la imagen **B** usando una división de  $2 \times 2$  zonas. Calcule los gradientes mediante convolución con los filtros de Prewitt, esto es:

$$I_x = \frac{\partial I}{\partial x} = I * \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad I_y = \frac{\partial I}{\partial y} = I * \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Realice una división estricta de zonas, es decir, no considere los pixeles donde el cálculo del gradiente incluya pixeles fuera de la zona.

Sea preciso en los ejes del histograma (qué valor representa cada bin y cuál es su altura).

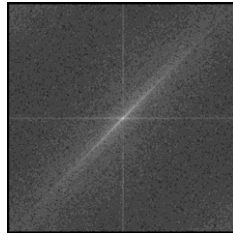
- d) (0.5 puntos) Dibuje el histograma de bordes normalizado (EHD) usando una división de  $2 \times 2$  zonas (división estricta). Considere los siguientes cinco filtros de orientación:

1	-1	1	1	$\sqrt{2}$	0	0	$\sqrt{2}$	2	-2
1	-1	-1	-1	0	$-\sqrt{2}$	$-\sqrt{2}$	0	-2	2
1		2		3		4		5	

Sea preciso en los ejes del histograma (qué valor representa cada bin y cuál es su altura).

## Semana 03

- a) (0.25 puntos) ¿En qué caso es conveniente implementar el filtro de convolución por medio de una DFT en vez de usar la definición original? Explique en máximo 2 líneas.
- b) (0.25 puntos) Suponga que encontró en su computador un archivo llamado `foto_dft.jpg` que al abrirlo ve la siguiente imagen en tonos de gris:



Cuando ve la imagen, recuerda que es la visualización de la DFT del archivo `foto.jpg` que lamentablemente borró hace algún tiempo. ¿Es posible usar `foto_dft.jpg` para obtener la imagen original `foto.jpg`? Explique en máximo 2 líneas.

- c) Para la siguiente imagen de 5 x 5 en escala de gris 8 bits:

	1	2	3	4	5
1	0	100	100	100	0
2	0	100	100	100	0
3	0	100	100	100	0
4	0	100	100	100	0
5	0	100	100	100	0

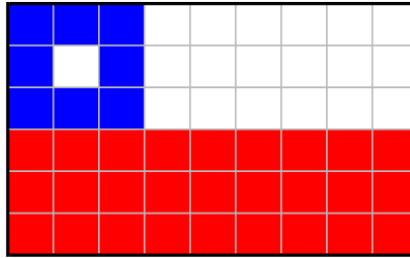
- i. (0.5 puntos) Escriba la imagen resultante de calcular el Local Binary Pattern (LBP)<sup>1</sup> en cada pixel de **C**. Realice la operación únicamente en los pixeles donde el cálculo esté completamente contenido en la imagen.
- ii. (0.25 puntos) Dibuje el histograma normalizado del descriptor LBP de **C**. Sea preciso en los ejes del histograma (qué valor representa cada bin y cuál es su altura).

---

<sup>1</sup> Debe seguir los pasos descritos en las slides "03.2-Descriptores globales gris (Texturas).pdf"

## Semana 04




- a) (0.25 puntos) Dada la siguiente imagen de 9 x 6 pixeles en color RGB 24 bits, dibuje el histograma global normalizado por canal de la imagen, esto es, la unión de tres histogramas de grises, uno para cada canal de la imagen. Sea preciso en las alturas de los bins y el valor que representa cada bin de cada histograma.



Nota: los colores azul y rojo son colores puros

- b) Dadas las siguientes dos imágenes y sus histogramas de colores:






**Histograma 1= { ( 0.3 ) , ( 0.2 ) , ( 0.5 ) }**

-Bin 1 representa colores cercanos a **R=172 G=137 B=150**

-Bin 2 representa colores cercanos a **R=218 G= 38 B=128**

-Bin 3 representa colores cercanos a **R=129 G=175 B=103**



**Histograma 2= { ( 0.3 ) , ( 0.4 ) , ( 0.3 ) }**

-Bin 1 representa colores cercanos a **R= 94 G=202 B=103**

-Bin 2 representa colores cercanos a **R=126 G= 87 B= 70**

-Bin 3 representa colores cercanos a **R=241 G=195 B= 11**

Se desea calcular la **Earth Mover's Distance (EMD)** entre los histogramas de ambas imágenes. Para esto siga los siguientes pasos:

- i. (0.75 puntos) Calcule la **Matriz de Costos** usando como ground-distance entre colores la distancia euclidiana en el espacio CIE LAB.

Para convertir un color RGB a CIE LAB puede usar la función `cvtColor` de OpenCV (ver ejemplo en “Anexo 4.1-Espacios de color RGB y CIE.ipynb”) o también puede usar el convertidor publicado en esta página: <https://colormine.org/convert/rgb-to-lab>

- ii. (0.25 puntos) Señale una posible **Matriz de Flujos**. No necesariamente debe ser la matriz óptima pero debe ser una matriz válida.
- iii. (0.25 puntos) Usando la Matriz de Costos y la Matriz de Flujos anteriores **calcule la EMD** entre los histogramas 1 y 2.

## Entrega

- Puede desarrollarlo en papel y enviar una foto (.jpg, .png), o puede desarrollarlo en formato digital en una planilla (.xlsx .ods), un documento (.docx) u otro formato exportado a .pdf.
- El plazo máximo de entrega es el **lunes 21 de abril de 2025** hasta las 23:59. Existirá una segunda fecha (por definir) para entregar su respuesta sin descuentos en la nota.

**El control es \*individual\* y debe ser de su autoría, es decir, no pueden ser resueltos por otro estudiante, no se pueden copiar respuestas de Internet, no se permite usar ChatGPT ni similares. En caso de detectar copia o plagio se asignará nota 1.0 a las o los estudiantes involucrados.**