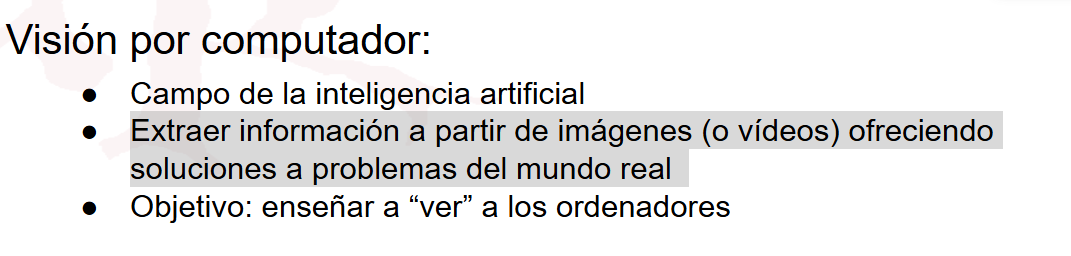
Ejercicios visión por computador

**Ejercicio 1.**

Resume en qué consiste el proceso de visión.

* Extraer información a partir de imágenes (o vídeos) ofreciendo soluciones a problemas del mundo real



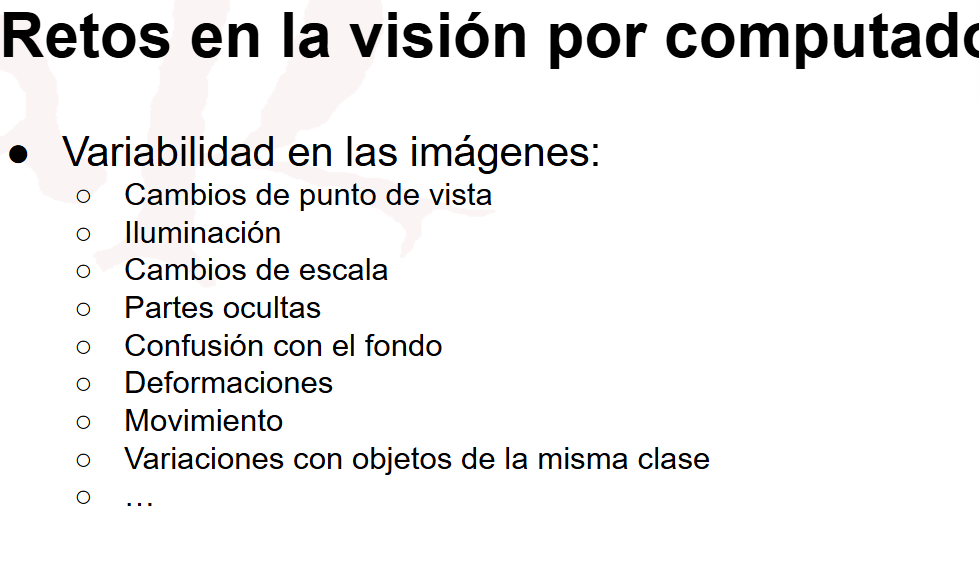
**Ejercicio 2.**

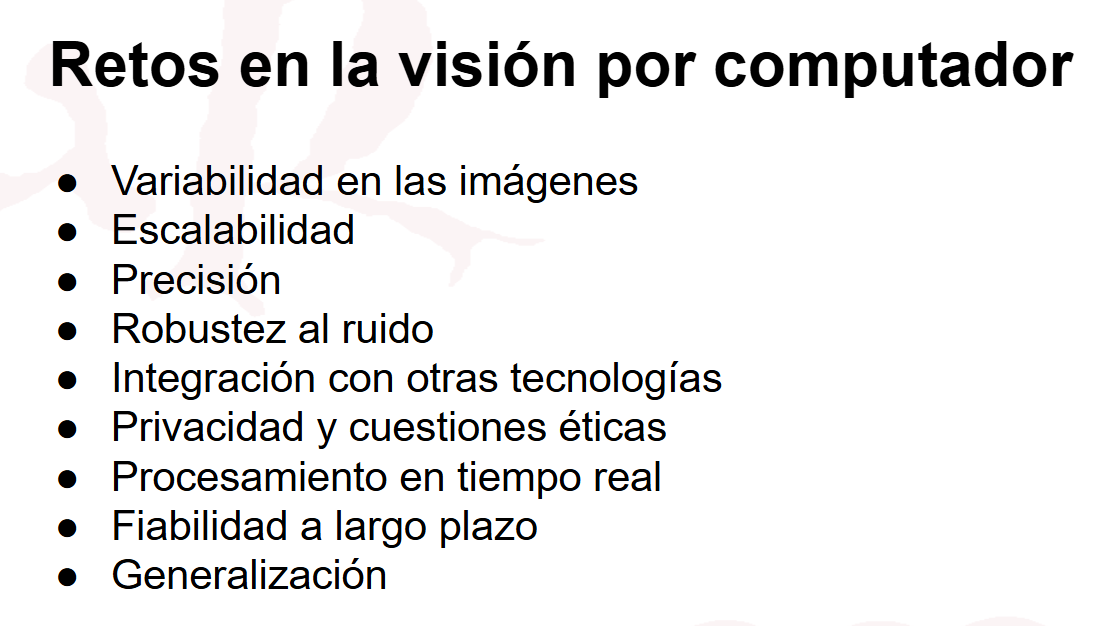
¿Qué es el salto semántico?

En **visión por computador**, el **salto semántico** (**semantic gap** en inglés) se refiere a la diferencia entre la información que puede extraerse de los datos visuales (como píxeles o características de bajo nivel) y el significado que un humano asocia a esos mismos datos.

**Ejercicio 3.**

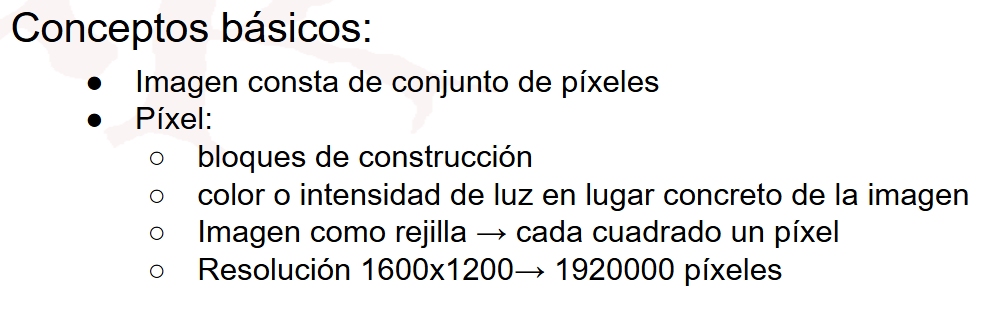
Nombra los retos de la visión por computador.





**Ejercicio 4.**

¿Qué representa un píxel dentro de una imagen?



**Ejercicio 5.**

¿Cómo se representa una imagen dentro del ordenador?

En un ordenador, una **imagen digital** se representa como una estructura de datos organizada que almacena información sobre cada punto (píxel) que la compone. La representación depende del tipo de imagen (binaria, en escala de grises, en color o multicanal) y del formato de almacenamiento (JPEG, PNG, RAW, etc.).

**Ejercicio 6.**

¿Qué rango de valores toman los píxeles de una imagen en escala de grises?

De 0 a 255

**Ejercicio 7.**

En una imagen en escala de grises, ¿qué píxel será más oscuro, uno con valor 10 o uno con valor 200?

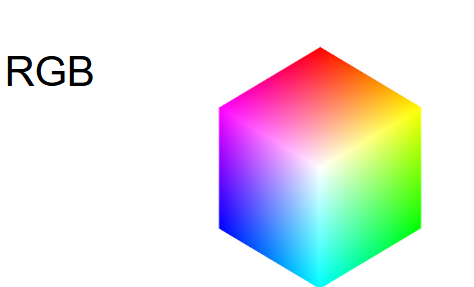
Uno de 10



**Ejercicio 8.**

En una imagen RGB, un píxel se representa mediante una tupla con tres valores que indican tres colores ¿cuáles son?

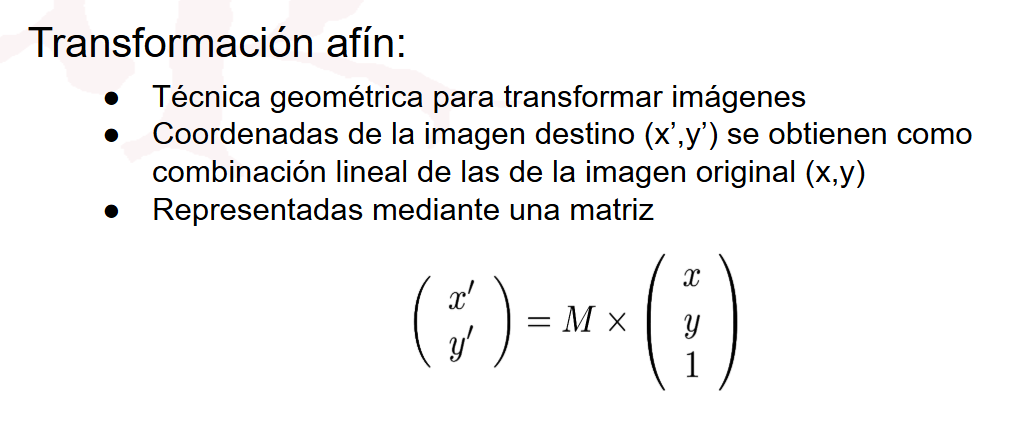
Azul, verde y rojo



**Ejercicio 9.**

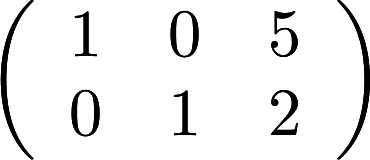
¿Qué es una transformación afín?

Una **transformación afín** es un tipo de transformación geométrica en el espacio euclidiano que preserva el paralelismo entre líneas (pero no necesariamente longitudes o ángulos). Es ampliamente utilizada en **visión por computador** y **procesamiento de imágenes** para realizar operaciones como rotación, traslación, escalado, sesgado (shear) o combinaciones de estas.

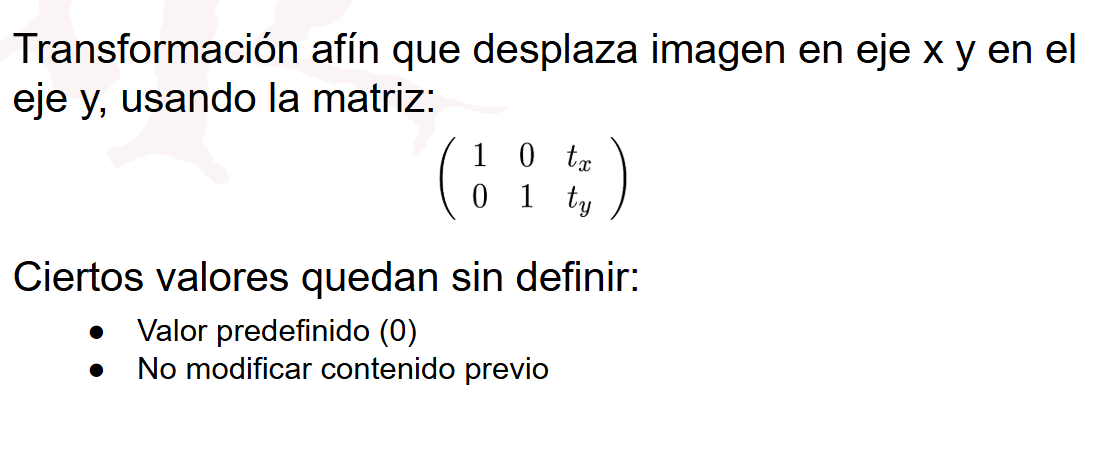


**Ejercicio 10.**

Dada la siguiente matriz de traslación

[](http://api.gmath.guru/cgi-bin/gmath?%5Cdpi%7B480%7D%5Cleft(%5Cbegin%7Barray%7D%7Bccc%7D1%260%265%5C%5C%0A0%261%262%5Cend%7Barray%7D%5Cright))

¿Cuál será el desplazamiento vertical de la imagen? ¿Y el horizontal?



Vertical 2 y horizontal 5

**Ejercicio 11.**

¿Por qué en el proceso de traslación algunos píxeles de la imagen destino quedan sin definir? ¿Cómo se definen?

La respuesta no está tal cual en las diapositivas. Algunos píxeles quedan sin definir porque no tienen ningún pixel al que correspondan en la imagen original. Por ejemplo si trasladas la imagen un pixel hacia abajo, los píxeles  de la primera fila de la imagen destino no corresponden con ningún pixel de la imagen origen. En esos casos se da un valor por defecto que suele ser 0.

Preservar la **relación de aspecto** (proporción ancho/alto) al escalar una imagen es crucial para evitar distorsiones como estiramientos o aplastamientos. Por ejemplo, si cambias solo el ancho sin ajustar el alto proporcionalmente, un círculo puede volverse un óvalo. En código, esto se soluciona calculando un factor de escala común para ambas dimensiones. Ignorar esta relación solo es válido en casos específicos, como redes neuronales que requieren imágenes cuadradas (usando padding o crop). La frase que mencionaste sobre "valores sin definir" parece ser un error técnico ajeno al concepto.

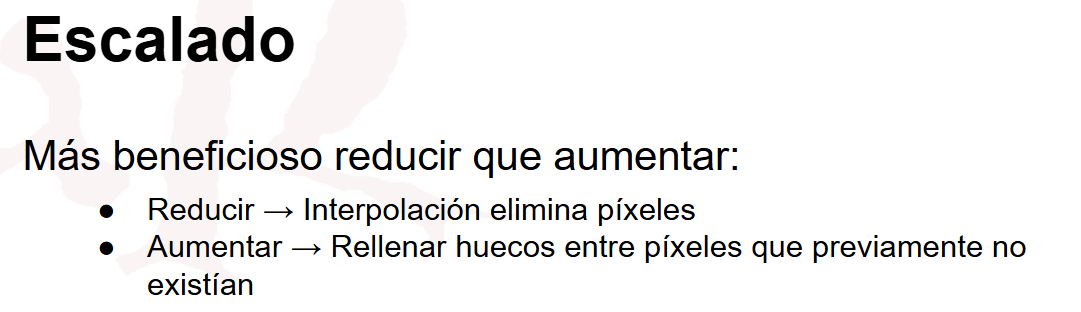
**Ejercicio 12.**

¿Qué dos valores son necesarios a la hora de rotar una imagen?

El angulo y el punto de rotación

**Ejercicio 13.**

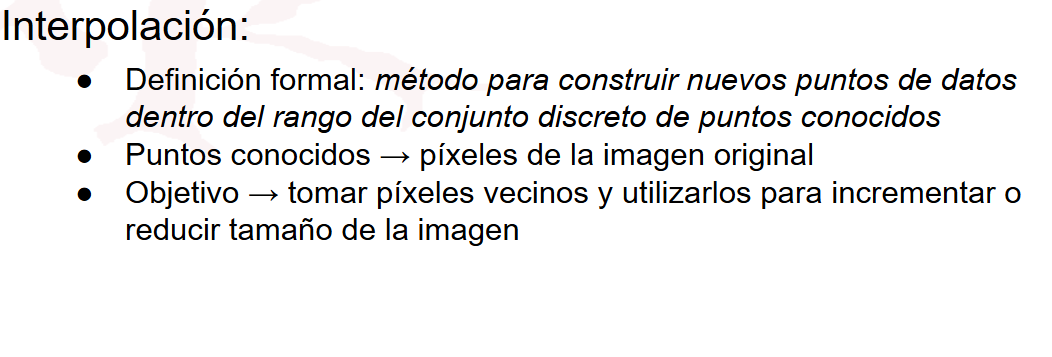
A la hora de escalar una imagen, ¿por qué es importante tener en cuenta la relación de aspecto?



**Ejercicio 14.**

¿Qué es la interpolación?

**Interpolación** en visión por computador y procesamiento de imágenes es el **método matemático para estimar valores desconocidos de píxeles** cuando se redimensiona, rota o transforma una imagen. Como las imágenes son matrices discretas de píxeles, al deformarlas o cambiar su tamaño, los nuevos píxeles no siempre coinciden con las posiciones originales. La interpolación determina cómo rellenar esos huecos.



**Ejercicio 15.**

¿Por qué es en general más beneficioso reducir el tamaño de una imagen que incrementarlo?

Porque al reducirlo fusionas pixeles con información que ya se tiene y al aumentar tienes que crear nuevos pixeles en correlación con los de su alrededor, lo que hacer perder calidad a la imagen.

**Ejercicio 16.**

¿Qué es una región de interés?

na **Región de Interés (ROI, por sus siglas en inglés: Region of Interest)** es una porción específica dentro de una imagen o conjunto de datos que se selecciona para su análisis, procesamiento o manipulación debido a su relevancia en una tarea concreta. En visión por computador y procesamiento de imágenes, trabajar con una ROI es el "foco de atención" dentro de una imagen, esencial para optimizar y precisar análisis en visión por computador

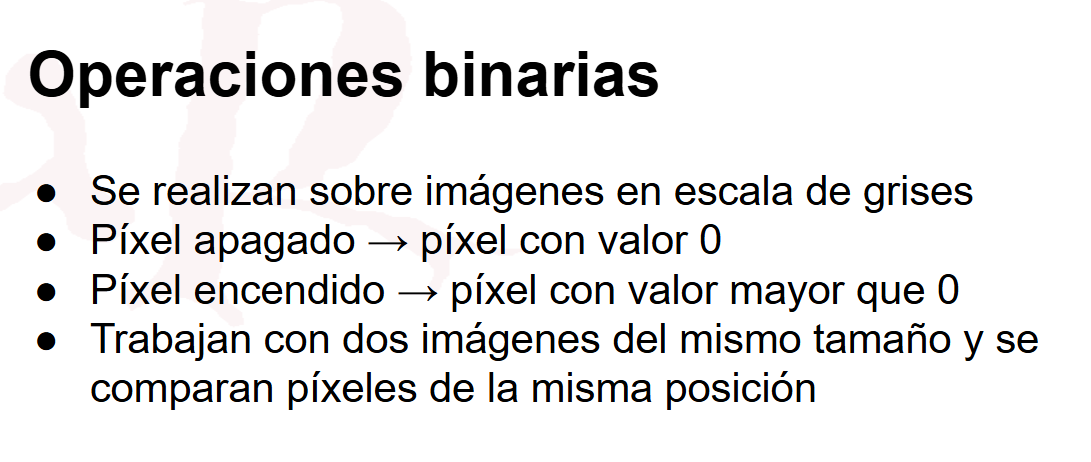
**Ejercicio 17.**

¿Cómo se utilizan las máscaras para recortar zonas que no son rectangulares?

Mediante mascaras binarias

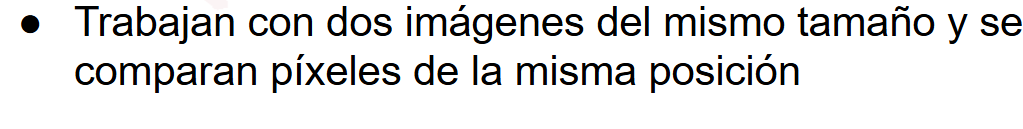
**Ejercicio 18.**

¿Sobre qué tipo de imágenes se realizan las operaciones binarias?



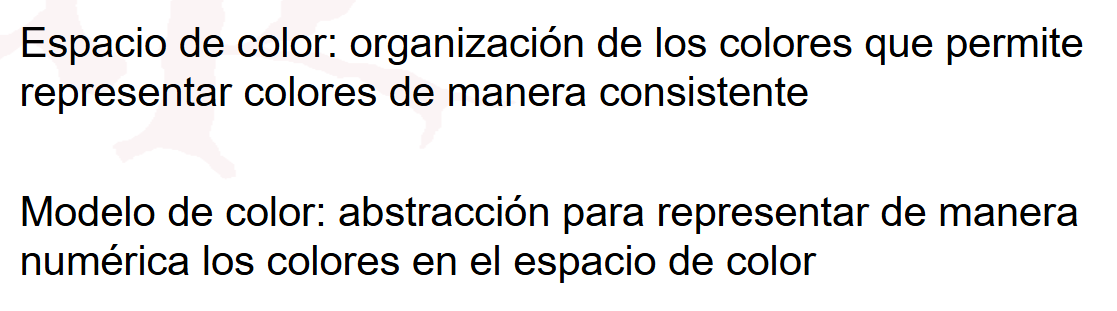
**Ejercicio 19.**

¿Con cuántas imágenes suelen trabajar las operaciones binarias?



**Ejercicio 20.**

¿Qué es un espacio de color? ¿y un modelo de color?



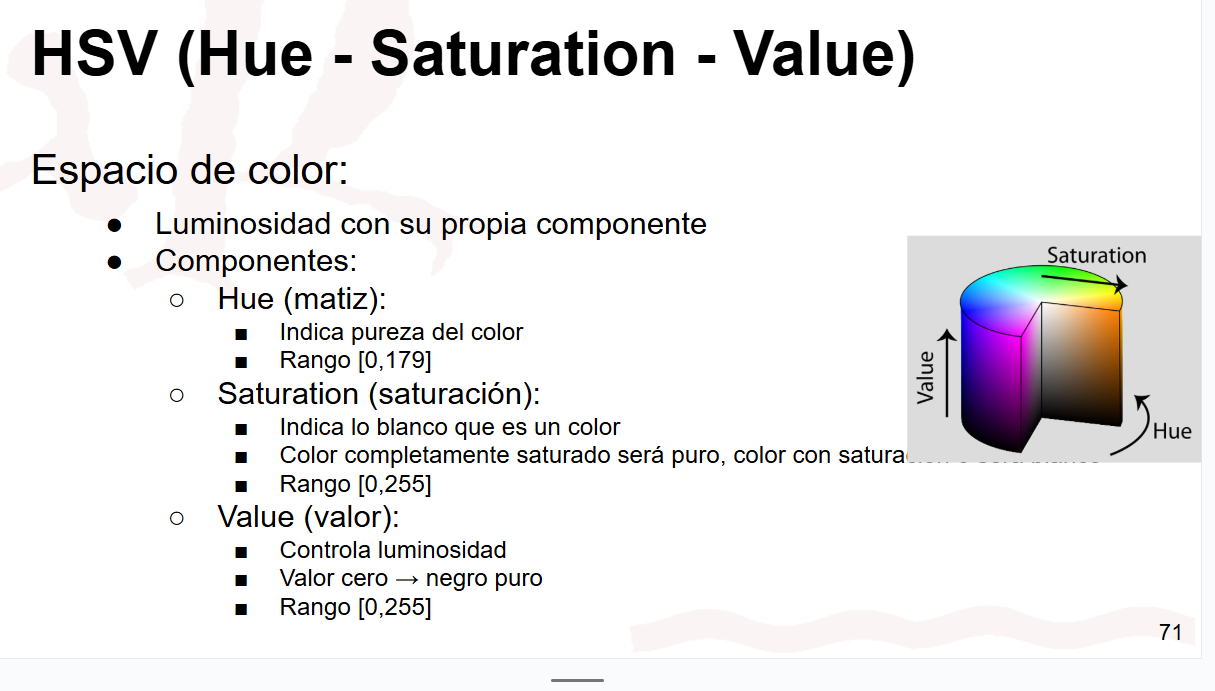
**Ejercicio 21.**

¿Por qué decimos que el espacio de color RGB es aditivo?

Porque el color obtenido es resultado de la mezcla en diferentes intensisdades (de 0 a 255) de los colores azul, rojo y verde.

**Ejercicio 22.**

¿Qué representan las componentes del espacio de color HSV?



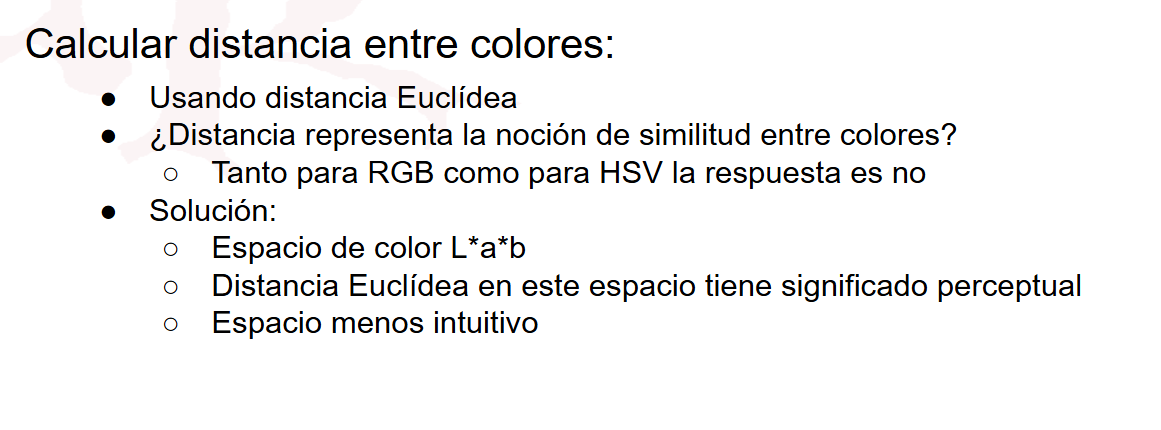
**Ejercicio 23.**

¿Por qué es preferible usar el espacio de color HSV cuando se quiere seguir el color de un objeto en una imagen?

* Más fácil definir rango de colores válidos

**Ejercicio 24.**

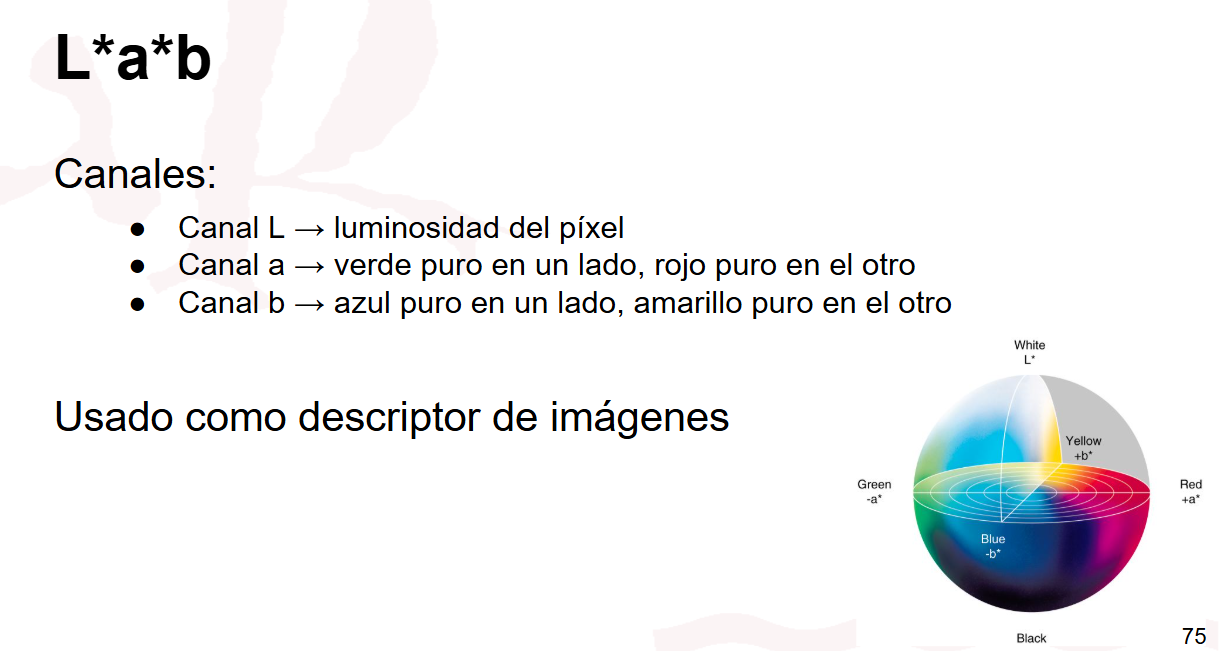
¿Qué problema soluciona el espacio de color L\*a\*b?



**Ejercicio 25.**

¿Por qué se suele utilizar el espacio de color L\*a\*b como descriptor de imágenes?

Porque puedes calcular distancia entre colores y te   
debido a su capacidad de representar la percepción visual del color de manera más uniforme que otros espacios de color.



**Ejercicio 26.**

¿Por qué no es correcto decir que una imagen en escala de grises es una imagen en blanco y negro? ¿Qué sería en realidad una imagen en blanco y negro?

Una imagen en blanco y negro sería una que solo tiene el valor 1 y el valor 0 para diferenciar entre negro y blanco.

**Ejercicio 27.**

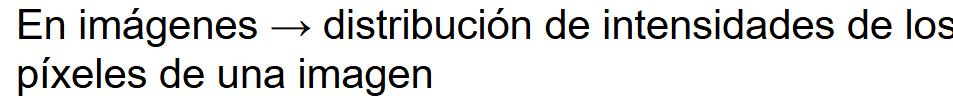
¿Por qué la transformación de RGB a escala de grises no es uniforme?

Porque hay muichos más valores en rgb que en escala de grises.

La **transformación de RGB a escala de grises** no es uniforme porque los canales de color (rojo, verde y azul) no contribuyen por igual a la percepción humana del brillo. El ojo humano es más sensible al verde que al rojo o al azul, por lo que se utilizan **pesos específicos** para cada canal al calcular el valor de gris.

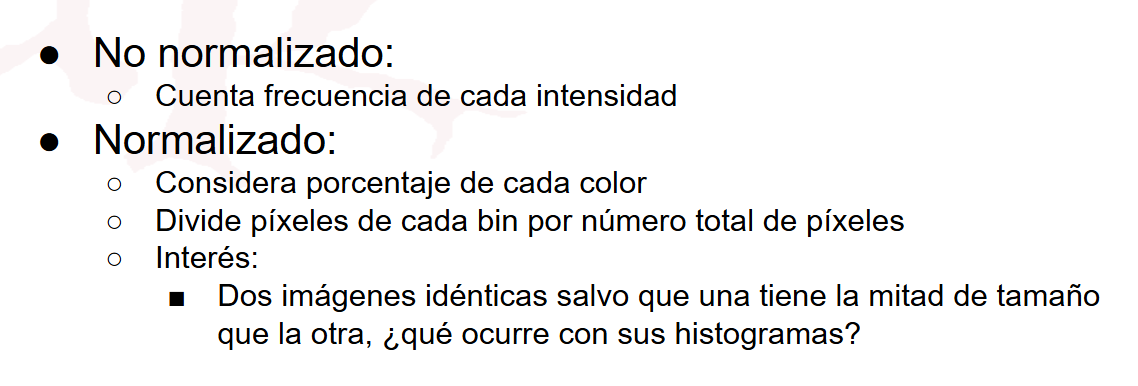
**Ejercicio 28.**

¿Por qué tienen tanta utilidad los histogramas?

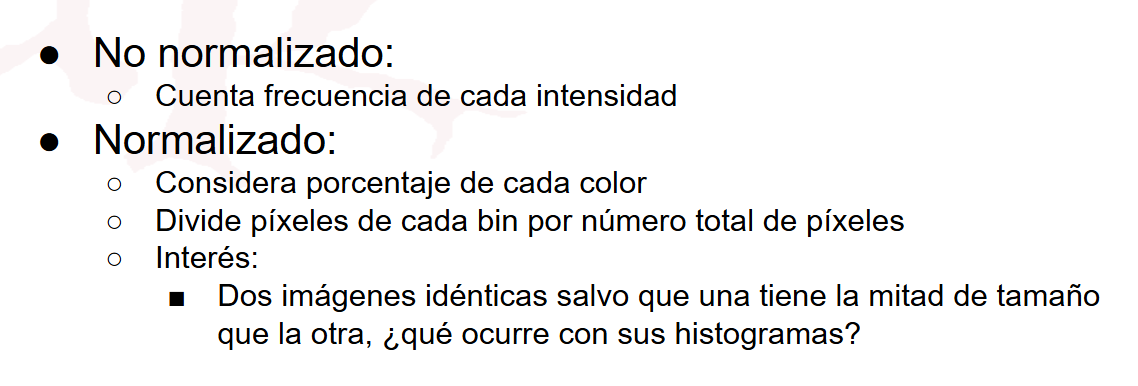


**Ejercicio 29.**

¿Cuál es la diferencia entre un histograma normalizado y uno no normalizado?

 **Ejercicio 30.**

¿Qué es un histograma normalizado? ¿Por qué es interesante considerar histogramas normalizados?



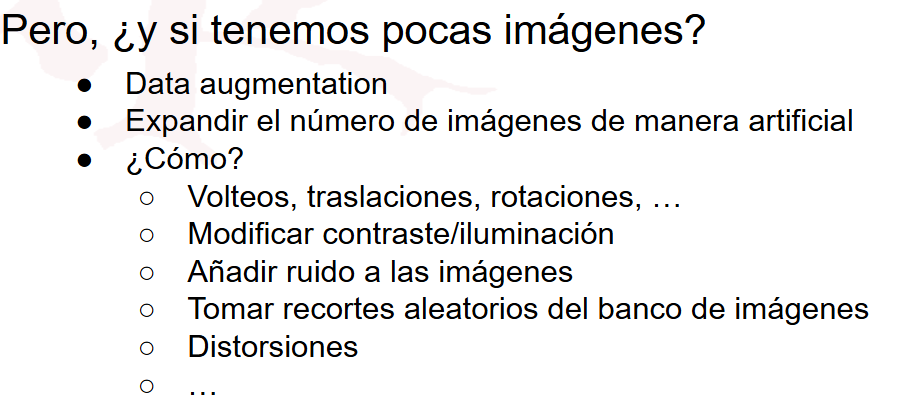
**Ejercicio 31.**

¿Cómo se representan los histogramas de color?

Con 3 líneas, de los colores primarios, con mayor o menos tamaño dependiendo de su participación en el pixel correspondiente.

**Ejercicio 32.**

Explica en qué consiste la técnica de data augmentation y cómo se logra.



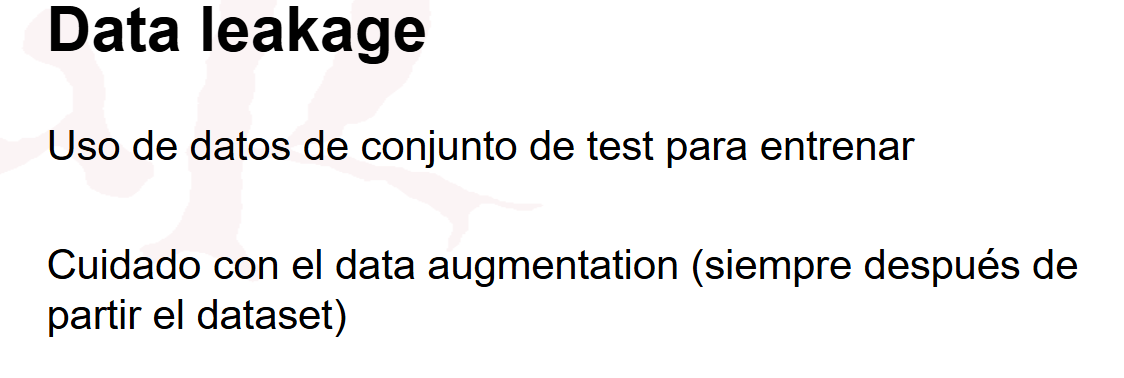
**Ejercicio 33.**

¿Por qué el aumento de datos ha de realizarse después de haber partido el dataset en los conjuntos de entrenamiento y test?

Porque sin oserian dependientes los resultados y no se valoraría si la inteligencia artificial es capaz de responder a imágenes nuevas sino a imágenes que ya conoce.

**Ejercicio 34.**

¿Qué es el data leakage?



**Ejercicio 35.**

¿Qué problemas tiene usar la propia imagen como conjunto de descriptores?



**Ejercicio 36.**

Dada la siguiente imagen representada mediante una matriz de números, ¿cuál sería su vector de descriptores utilizando las proyecciones? ¿y su vector de descriptores utilizando un histograma no normalizado con las bins [0,35], (35,75], y (75, 255]?

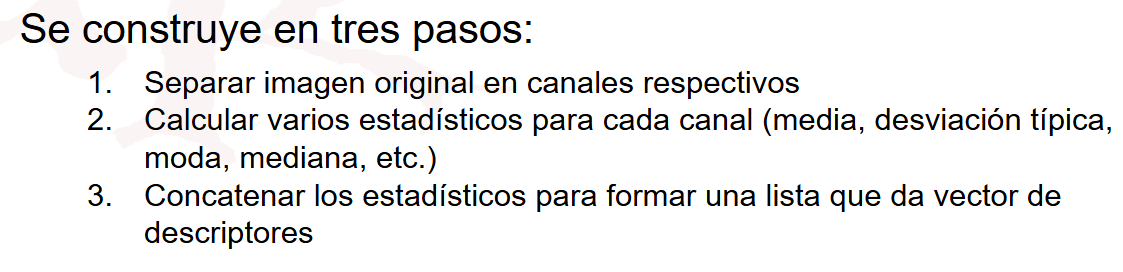
|  |  |  |  |
| --- | --- | --- | --- |
| 10 | 20 | 30 | 60 |
| 40 | 50 | 60 | 150 |
| 70 | 80 | 90 | 240 |
| 120 | 150 | 180 |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 3 |
| 1 | 2 | 2 | 5 |
| 2 | 3 | 3 |  |

[60,150,240,120,150,180] La matriz es lo rojo y lo otro, es su suma para calcularlo, 3,4,2

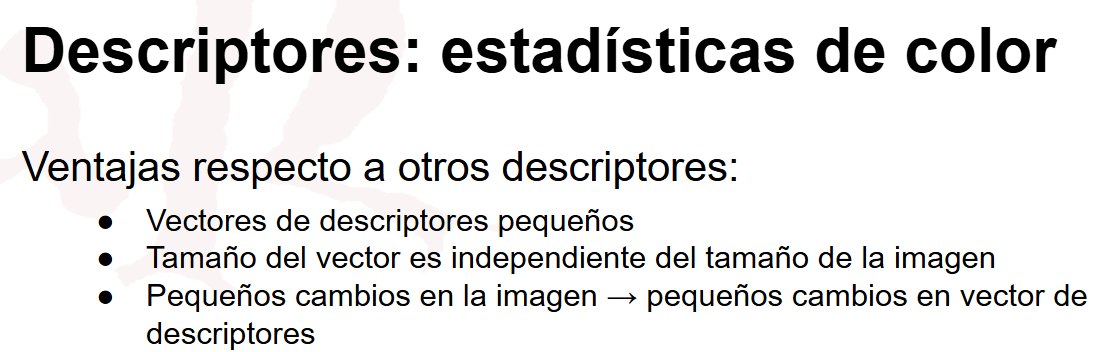
**Ejercicio 37.**

Explica los pasos a seguir para construir un descriptor basado en estadísticas de color.



**Ejercicio 38.**

Explica las ventajas de utilizar un descriptor basado en estadísticas de color en lugar de usar descriptores basados en la imagen completa o en proyecciones.



**Ejercicio 39.**

Tanto los histogramas como las estadísticas de color caracterizan la distribución de color de una imagen, pero ¿cuál es la diferencia entre ellos?

Los **histogramas** muestran la distribución detallada de frecuencias de los valores de píxeles en una imagen (ej: cuántos píxeles tienen intensidad 150 en el canal rojo), proporcionando una visualización completa pero sin información espacial. Las **estadísticas de color** (como media, desviación estándar o asimetría) resumen numéricamente esas distribuciones, ofreciendo métricas globales para análisis rápidos o comparaciones. Mientras los histogramas son ideales para ajustes de contraste o corrección de color, las estadísticas son útiles en machine learning o evaluación automatizada de propiedades como el brillo promedio. En esencia, el histograma revela el "qué" (distribución exacta) y las estadísticas el "cómo" (características agregadas).

**Ejercicio 40.**

¿Por qué es conveniente normalizar los histogramas cuando se trabaja con distintos tamaños de imagen?

Para que no haya diferencias de proporcionalidad.

**Ejercicio 41.**

¿Cuándo se utiliza el top-5 error rate como métrica para evaluar un clasificador?

Cuando hay muchas categorias

**Ejercicio 42.**

¿En qué consiste la técnica de la ventana deslizante y para qué se utiliza?

La **técnica de la ventana deslizante** (sliding window) es un método utilizado en visión por computador y procesamiento de imágenes que consiste en **analizar una imagen mediante una ventana rectangular que se desplaza sistemáticamente** (píxel a píxel o con saltos) sobre la imagen, extrayendo información o aplicando operaciones en cada posición.

**Ejercicio 43.**

Detalla los pasos de la aproximación clásica para construir un clasificador de imágenes.

Recordad los 5 pasos de un clasificador de imágenes:

Preparar el dataset.

Separar en conjuntos de entrenamiento

, test y validación.

Extraer descriptores.

Entrenar el modelo.

Evaluar el modelo.

En muchas ocasiones los pasos 2 y 3 son intercambiables

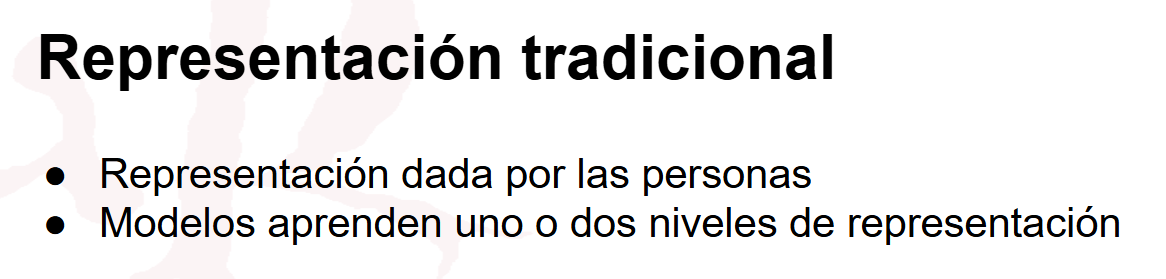
**Ejercicio 44.**

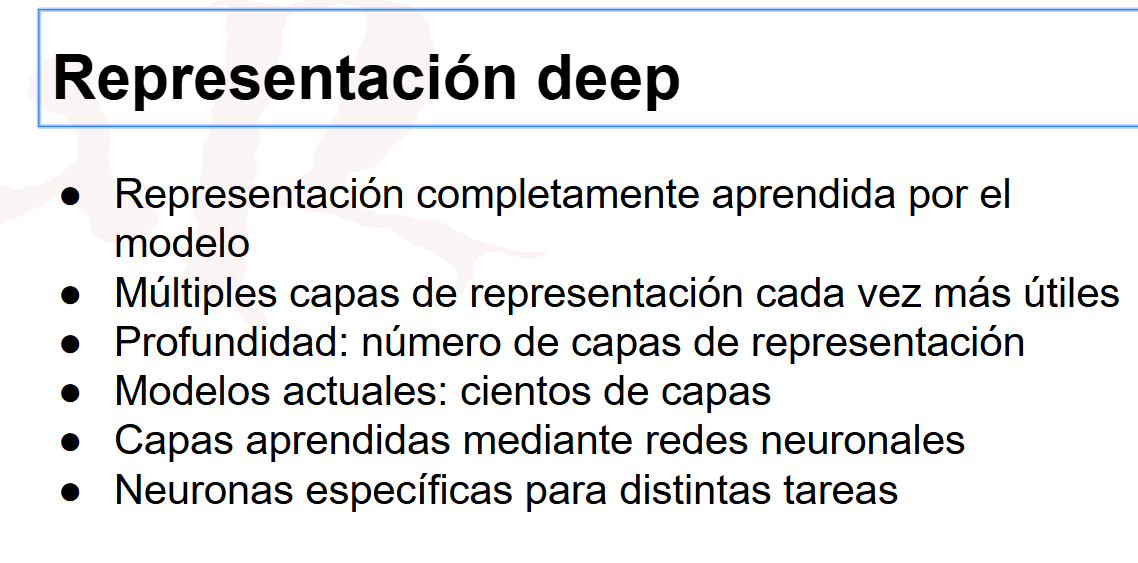
Explica cómo se puede utilizar un clasificador de imágenes para detectar objetos. Detalla también los problemas que tiene dicha aproximación y cómo se pueden resolver.

Un **clasificador de imágenes** puede usarse para detectar objetos aplicándolo a múltiples ventanas desplazadas sobre la imagen (sliding window), pero este enfoque clásico (ej: HOG + SVM) tiene problemas como **ineficiencia computacional, baja precisión con objetos pequeños o ocluidos, y falsos positivos**. Para resolverlo, los métodos modernos como **YOLO** o **Faster R-CNN** reemplazan la ventana deslizante con redes neuronales que predicen bounding boxes y clases en una sola pasada, logrando mayor velocidad y precisión gracias al aprendizaje automático de características y el uso de region proposals. Mientras los clasificadores tradicionales son útiles en entornos con pocos datos, las CNN dominan en aplicaciones reales que exigen robustez.

**Ejercicio 45.**

¿Cuál es la diferencia principal entre el modo tradicional y el modo deep para clasificar imágenes?

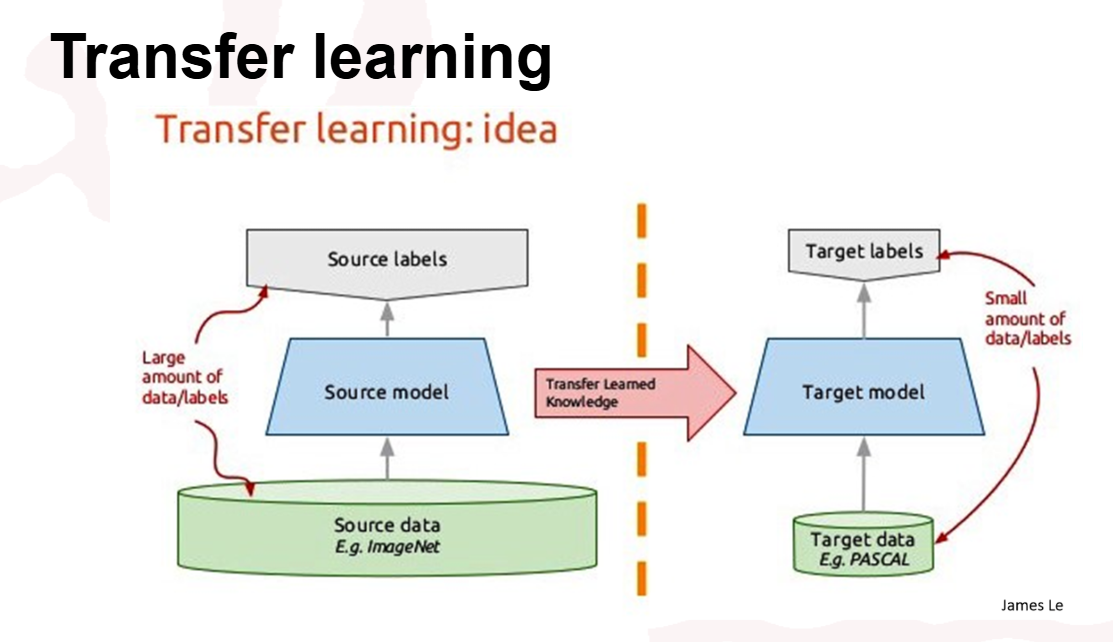




**Ejercicio 46.**

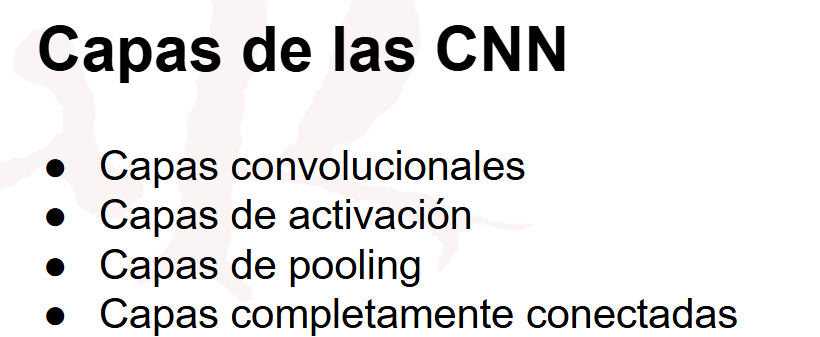
¿En qué consiste la técnica de transfer learning y para qué se utiliza? Explica los distintos modos de transfer learning que existen.

El **Transfer Learning** (aprendizaje por transferencia) es una técnica en la que un modelo preentrenado en una tarea grande (ej: clasificación de ImageNet) se **reutiliza** o **adapta** para una nueva tarea relacionada (ej: detectar tumores en radiografías), aprovechando el conocimiento ya aprendido (como bordes, texturas o formas).



**Ejercicio 47.**

¿Qué tipos de capas aparecen en una red convolucional?



**Ejercicio 48.**

Dada la siguiente matriz

Dada la siguiente matriz

|  |  |  |
| --- | --- | --- |
| 1 | 2 | 3 |
| 6 | 5 | 4 |
| 7 | 9 | 8 |

Y el filtro

|  |  |
| --- | --- |
| 1 | 0 |
| 0 | -1 |

Indica el resultado de convolucionar el filtro con la matriz sin usar padding y con un desplazamiento de 1.

|  |  |
| --- | --- |
| -4 | -2 |
| -3 | -3 |

**Ejercicio 49.**

Dada la siguiente matriz

|  |  |  |
| --- | --- | --- |
| 1 | 2 | 3 |
| 6 | 5 | 4 |
| 7 | 9 | 8 |

Indica el resultado de aplicar una capa de max pooling con filtro 3x3 y un desplazamiento de 1. Son 9

.

**Ejercicio 50.**

Explica qué son los ataques adversarios.

Un Ataque Adversario corresponde a una entrada al modelo que, de manera intencional, ha sido ligeramente modificada y que es capaz hacer que este modelo genere una salida incorrecta

**Ejercicio 51.**

Un algoritmo de detección ha producido las siguientes predicciones. Sabiendo que había 5 objetos en la imagen. Calcula los valores de precision y recall asociados.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rank | Confianza | ¿Correcta? | Precision | Recall |
| 1 | 0.95 | True | 1/1 | 1/5 |
| 2 | 0.90 | False | 1/2 | 1/5 |
| 3 | 0.88 | False | 1/3 | 1/5 |
| 4 | 0.83 | False | 1/4 | 1/5 |
| 5 | 0.72 | True | 2/5 | 2/5 |
| 6 | 0.64 | True | 3/6 | 3/5 |
| 7 | 0.57 | False | 3/7 | 3/5 |
| 8 | 0.34 | False | 3/8 | 3/5 |
| 9 | 0.22 | True | 4/9 | 4/5 |
| 10 | 0.20 | True | 5/10 | 5/5 |