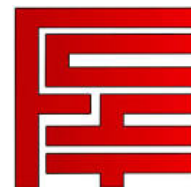




UNIVERSIDAD MAYOR DE SAN SIMÓN  
FACULTAD DE CIENCIAS Y TECNOLOGÍA  
ING. INFORMÁTICA



## Visión Artificial

**Estudiantes:** Palacios Galarza Osmar Ivan  
Rodriguez Blanco Josue Brandon  
Tonconi Copa Jesus Ernesto

**Carrera:** Ing Informática

**Materia:** Inteligencia Artificial.II

**Docente:** Lic. Carmen Rosa Garcia Perez

**Fecha:** 2/05/2023

I-2023  
Cochabamba - Bolivia

### **Introducción.**

La visión artificial es una rama de la inteligencia artificial que se centra en el desarrollo de algoritmos y sistemas que permiten a las computadoras "ver" y entender imágenes y videos, de manera similar a como lo hacen los seres humanos. La visión artificial utiliza la óptica, la electrónica, la matemática y la informática para procesar y analizar imágenes digitales.

La visión artificial se basa en la adquisición de datos a través de cámaras u otros dispositivos de imagen, y en la extracción de información de esos datos mediante algoritmos de procesamiento de imagen y aprendizaje automático. Estos algoritmos pueden ser entrenados con conjuntos de datos de imágenes etiquetados y clasificados, lo que permite a las computadoras reconocer patrones y características en nuevas imágenes.

En general, la visión artificial permite a las computadoras realizar tareas que antes eran exclusivas de los seres humanos, como el reconocimiento de objetos y la interpretación de escenas. A medida que la tecnología continúa evolucionando, es probable que veamos aún más aplicaciones innovadoras de la visión artificial en el futuro.

### **Historia**

La historia de la visión artificial se remonta a la década de 1950, cuando los investigadores comenzaron a explorar el uso de computadoras para analizar imágenes y extraer información de ellas. En ese momento, el procesamiento de imágenes se realizaba de manera manual, lo que limitaba la capacidad de analizar grandes conjuntos de datos.

En la década de 1960, se desarrollaron los primeros algoritmos para el procesamiento de imágenes, que incluían técnicas de segmentación y detección de bordes. A medida que se desarrollaban nuevas técnicas y algoritmos, la visión artificial se utilizó en una variedad de aplicaciones, incluyendo la clasificación de imágenes, la identificación de objetos y la detección de patrones.

En la década de 1970, se introdujeron las técnicas de reconocimiento de patrones y se comenzaron a utilizar técnicas de inteligencia artificial en el procesamiento de imágenes. Esto permitió el desarrollo de algoritmos más avanzados para la clasificación y el análisis de imágenes.

En la década de 1980, se desarrollaron los primeros sistemas comerciales de visión artificial, que se utilizaron principalmente en aplicaciones industriales, como la inspección de calidad. En ese momento, la visión artificial se había convertido en una herramienta práctica y eficiente para una variedad de aplicaciones.

En la década de 1990, se introdujeron técnicas más avanzadas de procesamiento de imágenes, como la visión estéreo y la visión 3D. También se desarrollaron los primeros

sistemas de reconocimiento facial, que se utilizan hoy en día en aplicaciones de seguridad y vigilancia.

En la década de 2000, se produjeron importantes avances en el campo del aprendizaje automático, lo que permitió el desarrollo de algoritmos de clasificación más precisos y eficientes. Esto llevó a una mayor adopción de la visión artificial en una variedad de campos, incluyendo la medicina, la biología y la robótica.

En la actualidad, la visión artificial se utiliza cada vez más en aplicaciones en tiempo real, como la conducción autónoma y los asistentes virtuales. Además, los avances en la tecnología de hardware y software están permitiendo el desarrollo de sistemas de visión artificial cada vez más avanzados y eficientes.

### **Aprendizaje Automático.**

El aprendizaje automático es útil en la visión artificial porque puede ayudar a las computadoras a reconocer patrones en grandes conjuntos de datos. En lugar de programar manualmente reglas para reconocer objetos o patrones específicos, el aprendizaje automático permite que la computadora "aprenda" automáticamente a reconocer patrones y objetos mediante la exposición a grandes conjuntos de datos de entrenamiento.

### **Procesamiento de imágenes.**

1. **Adquisición de imágenes:** La adquisición de imágenes es el primer paso en el procesamiento de imágenes. Este proceso implica la captura de imágenes digitales utilizando una cámara digital, un escáner o cualquier otro dispositivo de captura de imágenes.
2. **Preprocesamiento:** El preprocesamiento implica la corrección de la imagen para eliminar ruido, mejorar la calidad de la imagen y normalizar la iluminación y el contraste. Las técnicas de preprocesamiento comunes incluyen la corrección de la exposición, la eliminación de ruido y la normalización de la iluminación.
3. **Segmentación:** La segmentación implica la identificación de regiones de interés dentro de la imagen. Esto se logra mediante técnicas de umbralización, detección de bordes, agrupamiento de píxeles y clasificación de píxeles.

### **TRANSFORMACIONES MORFOLÓGICAS:**

Las transformaciones morfológicas se encargan de cambiar la forma y estructura de los objetos. Estas herramientas matemáticas permiten obtener componentes que dan una idea de la forma y estructura de los objetos que forman la imagen. Además, permiten modificar estas formas para separar los objetos unos de otros, obtener contornos primarios, descomponer formas complejas en formas más simples, obtener contornos dentro de entornos ruidosos, reconstruir elementos distorsionados, etc..

Generalmente esta serie de transformaciones se realizan en imágenes previamente binarizadas (en blanco y negro).

### DILATACIÓN BINARIA

Dados dos conjuntos A y B de  $Z^2$ , la dilatación, denotada por  $A \oplus B$ , se define como:  $A \oplus B = \{c \in E^N \mid c = a + b \text{ para todo } a \in A \text{ y } b \in B\}$

Es decir, dada una máscara B (formada por unos y ceros), la dilatación de A por B es el conjunto de todos los desplazamientos de x tales que B y A se solapen en al menos un elemento distinto de cero.

Por ejemplo, dada una máscara B formada por una matriz

0	1	0
1	1	1
0	1	0

y una figura A formada por una matriz:

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	1	1	1	0	0
0	0	1	0	1	0	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

$A \oplus B$ :

0	0	0	0	0	0	0
0	0	1	1	1	0	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

La dilatación, también llamada “crecimiento”, “llenado”, “expansión”, etc., produce un efecto de engrosamiento de los bordes del objeto. Este efecto es muy

valorado para aumentar el contorno de los objetos y unir líneas discontinuas de estos, producidas por algún filtrado, etc.

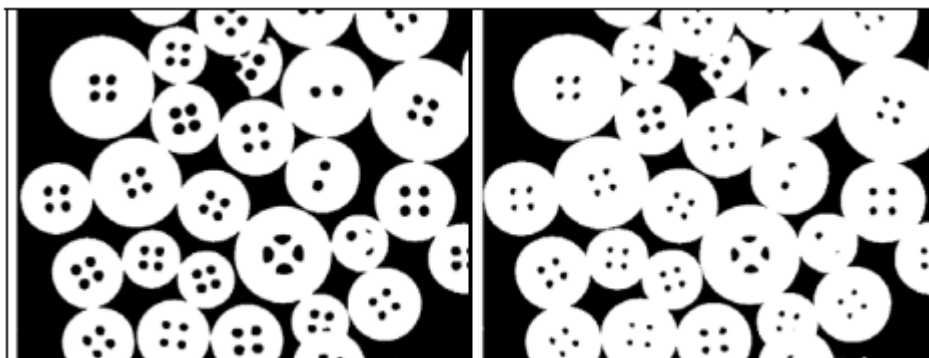


Imagen binarizada (izquierda) y dilatación de la misma(derecha)

### EROSIÓN BINARIA

Dados dos conjuntos A y B de  $\mathbb{Z}^2$  la erosión, denotada por  $A \ominus B$ , se define como:

$$A \ominus B = \{x \in E \mid x + b \in A \text{ para todo } b \in B\}$$

Dada una imagen A se erosiona por B cuando para todos los puntos x tales que B, trasladado por x, está contenido en A. Es decir, la erosión pone a cero todos los pixels de la imagen que no contengan completamente al elemento estructurante en su entorno.

Por ejemplo, usando la máscara del caso anterior:

0	1	0
1	1	1
0	1	0

y una figura A formada por una matriz:

0	0	0	0	0	0	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0

0	0	0	0	0	0	0
---	---	---	---	---	---	---

$A \ominus B$ :

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	1	1	1	0	0
0	0	1	1	1	0	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Si la dilatación expandía los bordes y contornos de los objetos, la erosión reduce los contornos de los objetos. Se utiliza para separar objetos que están unidos por una pequeña parte de sus contornos.

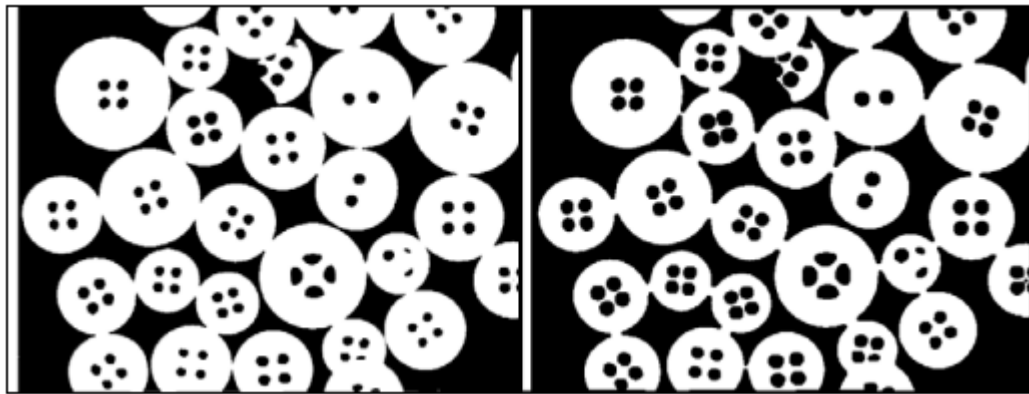


Imagen binarizada (izquierda) y erosionada (derecha)

## APERTURA

La apertura consiste en realizar en una imagen una erosión y después una dilatación. Aunque a simple vista puede parecer que la imagen va a quedar como está, esto no es así debido a que la dilatación no es una operación inversa de la erosión.

La apertura es muy utilizada para:

Segmentación de objetos, separando unas formas de otras.

Descomposición de objetos en elementos más simples.

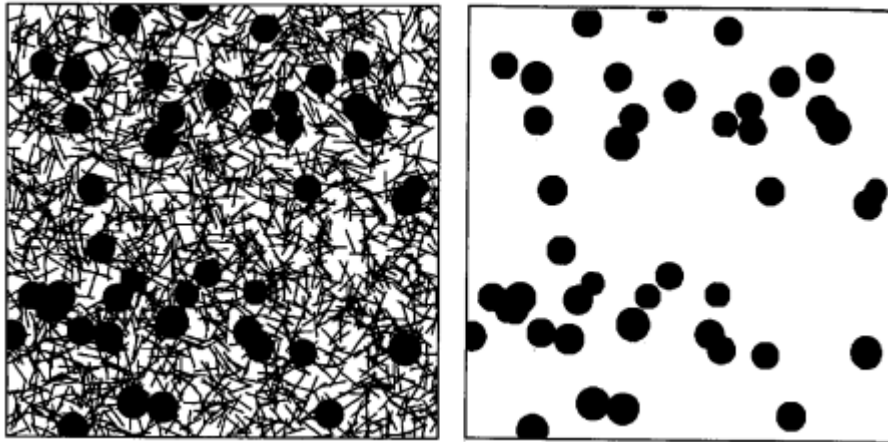
Extracción de formas determinadas en un entorno con ruido.

Eliminar salientes estrechos.

Separar objetos que no están demasiado pegados.

Aumentar los agujeros que están dentro de los objetos.

Las aperturas pueden estar formadas por un número de erosiones y dilataciones diferentes, produciendo resultados muy dispares.



## CIERRE

El cierre hace exactamente lo contrario que la apertura, primero se realiza la dilatación y después la erosión. La combinación de las operaciones de apertura y cierre pueden servir para múltiples operaciones: filtrado, segmentación, etc.; simplemente con jugar con el tipo de máscaras y tamaños.

- a) imagen original
- b) apertura con disco de radio 1
- c) cierre con disco de radio 4
- d) apertura con disco de radio 3
- e) apertura con rectángulo de 21x20
- f) residuo de la apertura anterior
- g) apertura con máscara de estructura vertical.



(a)



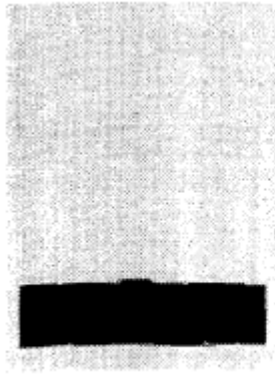
(b)



(c)



(d)



(e)



(f)



(g)

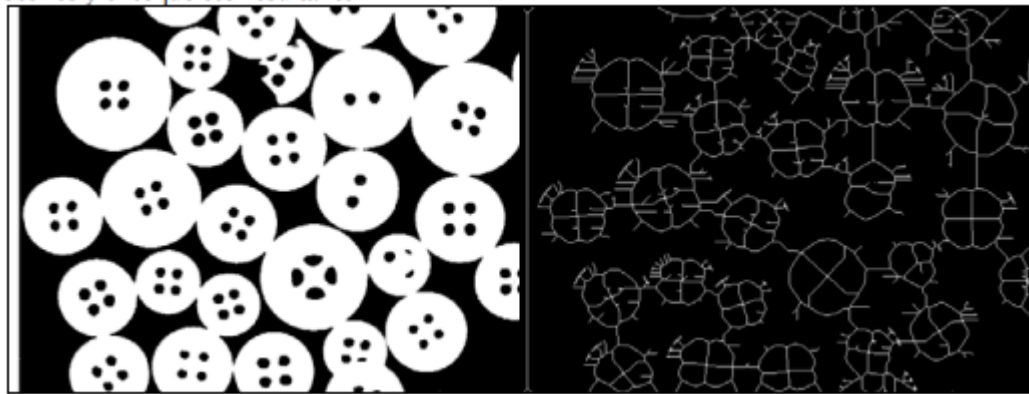
## ESQUELETOS

Este tipo de operación reduce los objetos de una imagen, a un esqueleto de grosor de un pixel. Es decir, dado un objeto cualquiera va reduciendo su grosor hasta dejarlo con un grosor de un pixel.

El esqueleto conseguido da una idea de la forma inicial del objeto.



El problema de este tipo de transformación, es el tiempo prohibitivo que supone su cómputo. Sólo pudiéndose implementar con hardware específico que acelere el proceso de cálculo.



### OTRAS TRANSFORMACIONES

(bottom hat): Que consiste en realizar un cierre de la imagen y después restar el resultado a la imagen original.

‘bridge’: Une pixels sin conectar.

‘clean’: Elimina pixels sueltos.

‘fill’: Rellena pixels que están a 0 y rodeados de 1.

‘diag’: Hace un relleno diagonal para eliminar la conectividad 8.

‘hbreak’: Elimina pixels que están conectados en forma de H.

‘majority’: Pone un pixel a 1, si 5 o más vecinos (una máscara de 3x3) están a 1.

‘dilate’: Realiza la dilatación (3x3).

‘erode’: Realiza la erosión (3x3).

‘open’: Realiza una apertura (3x3).

‘close’: Realiza un cierre (3x3).

‘remove’: Elimina pixels interiores. Pone un pixel a cero si los cuatro vecinos están a 1 (con conectividad 4). ‘shrink’: Reduce objetos a puntos. Si estos tienen agujeros, los convierte a puntos formando un anillo que los une al contorno exterior.

‘skel’: Realiza la esqueletización.

‘spur’: Elimina pixels que aparezcan sueltos y solo unidos por una esquina de los mismos. ‘thicken’: Engorda objetos añadiendo pixels al exterior de los mismos.

‘thin’: Adelgaza los objetos, quitando pixels de sus contorno.

‘tophat’: Realiza la apertura de la imagen y se la resta a la imagen original.

## TÉCNICAS DE SEGMENTACIÓN

### SEGMENTACIÓN BASADA EN UMBRALIZACIÓN

El proceso de umbralización basado en el histograma, consiste en agrupar los píxeles según sus niveles de intensidad luminosa. Por ejemplo, si se analiza la imagen de una manzana roja en un fondo blanco, el histograma estará formado por dos montañas correspondientes a los píxeles del nivel de gris (si la imagen obtenida es en niveles grises) y al fondo blanco.

Si en vez de tener un objeto, se tiene varios con diferentes niveles de grises, aparecen varios montículos en el histograma correspondientes a cada objeto

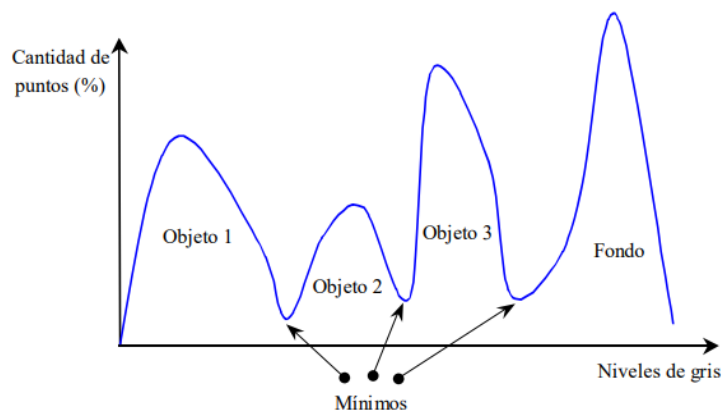
### MÉTODO P-CUANTIL

Este método usa el conocimiento sobre el área o el tamaño del objeto que se desea extraer para segmentar la imagen. Si se supone que en una aplicación determinada los objetos ocupan alrededor del p por ciento del tamaño del área de la imagen. Utilizando este conocimiento es posible particionar la imagen de modo que asigne el p por ciento de los pixels al objeto

Es claro que el uso de este método es muy limitado. Sin embargo, en aplicaciones como preprocesamiento para reconocimiento de caracteres es muy útil.

### UMBRALIZACIÓN BASADA EN LA BÚSQUEDA DE MÍNIMOS

Dado que un objeto con unos niveles de grises homogéneos va a tener en su histograma un montículo determinado, el procedimiento consiste en obtener los mínimos de esos valles y asignarlos como valores iniciales de los umbrales buscados.



## *TÉCNICAS DE SEGMENTACIÓN BASADAS EN BORDES*

Las técnicas de segmentación basadas en bordes se basan en los bordes encontrados en la imagen por los detectores.

### EXTRACCIÓN DE LA FRONTERA

Las fronteras son bordes unidos que caracterizan la forma de un objeto. Son, por tanto, útiles para calcular rasgos geométricos como tamaño u orientación

encontrar fronteras o bordes puede ser complicado debido a las definiciones de conectividad utilizadas. Por ejemplo, si se utiliza la 4-conectividad, segmentos que aparentemente forman parte de una figura pueden ser clasificados como disjuntos. Por otro lado, si se utiliza la 8-conectividad, los segmentos pueden estar conectados pero también puede haber conexiones no deseadas entre el interior y el exterior de una figura. En general, hay problemas asociados con la definición de conectividad que deben ser resueltos de manera cuidadosa.

### SEGUIMIENTO DE CONTORNO

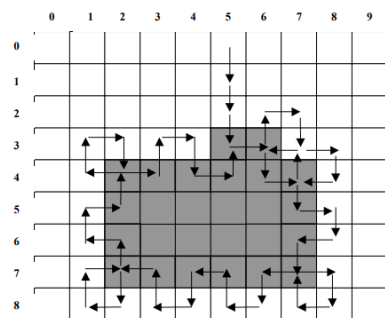
Como el nombre sugiere, los algoritmos de seguimiento de contorno trazan las fronteras ordenando los puntos de tipo borde sucesivos. Un algoritmo muy utilizado para determinar los puntos del contorno de un objeto es el algoritmo de la tortuga.

Se empieza desde un punto cercano al objeto.

Se avanza pixel a pixel hasta encontrar un pixel negro correspondiente al contorno a determinar.

Empieza el algoritmo de la tortuga que consiste en avanzar de la siguiente manera: si está en un punto negro gira  $90^\circ$  en SENTIDO CONTRARIO A LAS AGUJAS DEL RELOJ y avanza un pixel, si en cambio está en un punto blanco gira  $90^\circ$  en SENTIDO DE LAS AGUJAS DE RELOJ y avanza un pixel. Se van almacenando las coordenadas de todos los pixels negros que se van encontrando. El algoritmo termina cuando llega de nuevo al primer punto negro encontrado.

Como en el vector del contorno aparecen algunos puntos repetidos (algunas esquinas), se eliminan. (siempre que sea un objeto cerrado y no se salga de la imagen).

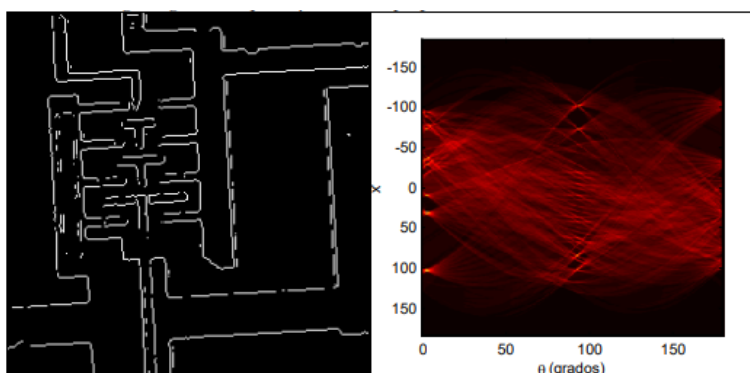


### TRANSFORMADA HOUGH

La transformada Hough, sirve para vectorizar una imagen de líneas y contornos. Es decir, dada una imagen de la que se han extraído sus contornos se procede a convertir estos en una serie de vectores. Este tipo de transformada no obtiene una serie de líneas que definen los objetos. El problema es que el cálculo de la transformada de Hough es muy lento, debido a que ésta tiene que desplazarse por todos los contornos ajustando las líneas buscadas a éstos.

### TRANSFORMADA RADON

Otro tipo de transformada es la transformada Radon. Este tipo de transformada consiste en proyectar todos los puntos de la imagen en una línea diagonal que va girando de 0 a 180 grados. Es decir, para cada grado se tiene un vector que representa a la recta con esa inclinación y el número de puntos que han incidido en ella. Donde mayores valores se obtengan significará que hay una recta perpendicular a esa inclinación.



### TRANSFORMADA DEL RELLENADO CRECIENTE

Consiste en asignar un valor a cada pixel según la distancia a la que esté del borde. Esto permite distinguir fácilmente los objetos de una imagen.

0	0	0	0	0	0	0	0	0
0	1	1	1	1	1	0	0	0
0	1	2	2	2	1	0	0	0

0	1	2	3	2	1	0	0	0
0	1	2	3	2	1	1	1	0
0	1	2	3	2	2	2	1	0
0	1	2	3	3	3	2	1	0
0	1	2	2	2	2	2	1	0
0	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0	0

### SEGMENTACIÓN ORIENTADA A REGIONES

Las regiones resultantes de estos procesos de segmentación deberán ser homogéneas y maximales, donde maximal significa que los criterios de homogeneidad no se cumplirán tras la unión de una región con alguna de sus adyacentes.

El criterio más simple de homogeneidad utiliza el nivel de gris medio de la región, algunas propiedades simples o un vector de medias para imágenes multibanda.

### UNIÓN DE REGIONES

El método más natural para el crecimiento de las regiones es comenzar considerando cada pixel de la imagen como una región, obviamente en este caso cada región es homogénea, pero no necesariamente cumple el criterio de ser las regiones maximales. El proceso habrá de repetirse hasta que las regiones sean maximales. En algoritmo, el método sería:

Definir una segmentación inicial que cumpla el criterio de homogeneidad.

Definir un criterio para unir regiones adyacentes.

Unir las regiones adyacentes si cumplen el criterio de unión. Parar cuando no puedan unirse dos regiones sin romper el criterio de homogeneidad.

En general, podría decirse que lo más importante es seleccionar el criterio para realizar la unión. Algunos criterios de tipo heurístico son: pixels que a ambos lados tienen una diferencia menor que un cierto umbral

### CRECIMIENTO DE REGIONES POR AGREGACIÓN DE PIXELS

el crecimiento de regiones es un procedimiento que agrupa pixels o subregiones en regiones mayores. La forma más sencilla de este proceso es la llamada agregación de pixels que comienza con un conjunto de pixels semilla y a partir de ellos hace crecer la región añadiendo a dichos pixels semilla aquellos vecinos que tienen propiedades similares (nivel de gris, color, textura)

5	6	5	7	8
6	7	5	8	8
5	5	6	9	9
5	6	6	7	7
6	5	7	8	9

píxeles semilla = [(0, 0), (4, 4)]

podríamos definir que los píxeles vecinos que tienen un nivel de gris a lo sumo 1 punto de diferencia del nivel de gris del píxel actual se agreguen a la región.

1	1	1	0	0
1	1	1	0	0
1	1	1	0	0
1	1	1	1	1
0	1	1	1	1

### *SEGMENTACIÓN MEDIANTE EXTRACCIÓN Y ANÁLISIS DE TEXTURAS*

Dentro de las técnicas de segmentación actuales, cobran gran importancia todas aquellas que tratan las texturas que se encuentran en la imagen. Realmente, no hay una única y universal definición de textura, aunque se suele determinar que la medida de texturas es la medición en una imagen de la rugosidad, suavidad y regularidad dentro de ella.

### *TÉCNICAS DE CLUSTERIZADO*

Las técnicas de clusterizado o agrupación de datos es un método muy utilizado en la clasificación automática de objetos.

En este caso, se trata de definir un conjunto de características asociadas a cada píxel generando un vector multidimensional por cada punto de la pantalla. La técnica de clusterizado se encarga de agrupar dichos vectores según su grado de semejanza entre ellos y así, finalmente se obtienen N grupos de píxeles con características semejantes entre ellos

### *SEGMENTACIÓN POR MOVIMIENTO*

segmentación de objetos en movimiento. En este caso, la segmentación se aplica a varias imágenes obtenidas en el tiempo, de forma que por comparación se detecta el objeto en movimiento y se extraen sus características

4. **Extracción de características:** La extracción de características es una etapa fundamental en la visión artificial, ya que permite identificar y analizar características relevantes de una imagen que pueden ser útiles para la identificación de objetos o patrones en la misma. La extracción de características se utiliza en una amplia variedad de aplicaciones, incluyendo el reconocimiento de objetos, la detección de anomalías, la clasificación de imágenes, la segmentación de imágenes y la detección de bordes.

Existen diferentes tipos de características que se pueden extraer de una imagen, incluyendo formas, texturas, colores, características de movimiento, entre otras. Cada tipo de característica puede ser útil en diferentes aplicaciones, dependiendo del problema que se esté tratando de resolver.

A continuación, se describen algunos de los métodos más comunes utilizados para la extracción de características en la visión artificial:

**Características basadas en bordes:** Los bordes son cambios abruptos en los niveles de intensidad de una imagen y pueden ser utilizados para identificar formas y contornos en una imagen. Los métodos basados en bordes pueden ser utilizados para detectar bordes en una imagen y luego utilizar estos bordes para extraer características como la forma y el tamaño de los objetos en la imagen.

**Características basadas en texturas:** La textura se refiere a los patrones repetitivos de las características de la imagen, como la rugosidad, la suavidad, la granularidad, entre otros. Los métodos basados en texturas pueden ser utilizados para extraer características que son útiles en la clasificación de objetos, como la presencia de ciertas texturas en la imagen.

**Características basadas en formas:** Las formas son características importantes de los objetos en una imagen y pueden ser utilizadas para la identificación y clasificación de objetos. Los métodos basados en formas pueden ser utilizados para extraer características como el área, la relación de aspecto, la circularidad y la simetría de los objetos en la imagen.

**Características basadas en colores:** El color es una característica importante de las imágenes y puede ser utilizado para la identificación y clasificación de objetos. Los métodos basados en colores pueden ser utilizados para extraer características como la distribución de colores en la imagen, la presencia de ciertos colores en la imagen y la relación entre los colores en la imagen.

Además de los métodos mencionados anteriormente, existen muchos otros métodos para la extracción de características en la visión artificial, incluyendo características basadas en la forma de onda, características basadas en la transformada de Fourier, características basadas en la geometría y características basadas en la textura fractal, entre otros.

5. **Clasificación:** La clasificación implica la asignación de etiquetas a las regiones de interés identificadas en la imagen. Esto se logra mediante técnicas de aprendizaje automático, como la clasificación supervisada y no supervisada.

Ejemplos de clasificación:

**El reconocimiento facial:** Es una tecnología de visión artificial que se utiliza para identificar y verificar la identidad de una persona mediante el análisis de características faciales únicas, como la forma de los ojos, la nariz, los labios y la estructura de la cara. Esta tecnología es capaz de comparar la imagen de una cara capturada por una cámara con una base de datos de imágenes previamente registradas para determinar si hay una coincidencia.

El reconocimiento facial se ha utilizado en una variedad de aplicaciones, como la seguridad de acceso a edificios, el control de fronteras, la identificación de sospechosos criminales y la gestión de la asistencia de los empleados. También se ha utilizado en aplicaciones comerciales, como el marketing personalizado y la autenticación en dispositivos móviles.

El reconocimiento facial utiliza técnicas de aprendizaje automático y redes neuronales para identificar patrones en las imágenes faciales y crear modelos de reconocimiento facial precisos. Sin embargo, el uso de la tecnología de reconocimiento facial ha generado controversias en cuanto a la privacidad y el uso indebido de los datos personales. Por lo tanto, su uso y regulación son objeto de debate en muchos países y comunidades.

**Eigenfaces:** Es una técnica de análisis de componentes principales (PCA) que se utiliza comúnmente en el reconocimiento facial para identificar y verificar la identidad de una persona.

El método de Eigenfaces se basa en el concepto de que las imágenes faciales pueden ser representadas como una combinación lineal de componentes básicos o "eigenfaces", que son vectores de características que se derivan de un conjunto de imágenes faciales. Estos eigenfaces se utilizan para encontrar patrones y características distintivas de la cara humana que se pueden utilizar para identificar y reconocer rostros.

**La detección de rostros en tiempo real:** Es una técnica de visión artificial que se utiliza para detectar la presencia y ubicación de rostros humanos en una imagen en tiempo real, generalmente a través de una cámara en vivo. Esta técnica se basa en algoritmos de detección de objetos y aprendizaje automático para analizar la imagen de la cámara y detectar la presencia de rostros humanos.



La detección de rostros en tiempo real se utiliza en una variedad de aplicaciones, como la videovigilancia, la interacción humano-computadora, el reconocimiento facial y la autenticación de usuario. La técnica se ha vuelto cada vez más común en los últimos años gracias a la mejora en la tecnología de hardware y software, como las redes neuronales convolucionales (CNN) y las unidades de procesamiento gráfico (GPU) de alta velocidad.

La detección de rostros en tiempo real se realiza en tiempo real, lo que significa que los resultados se generan en tiempo real y se actualizan continuamente a medida que cambia la imagen de la cámara. Esto requiere un procesamiento rápido y eficiente de la imagen, lo que se logra mediante el uso de técnicas de optimización de algoritmos y hardware especializado.

En general, la detección de rostros en tiempo real es una técnica avanzada de visión artificial que tiene un gran potencial en una variedad de aplicaciones, pero también plantea desafíos técnicos y éticos, como la privacidad y la precisión en la detección.

**Instancie.-** se refiere a la identificación de objetos o regiones específicas en una imagen o video, de manera que cada objeto o región se pueda analizar y comprender de forma independiente.

La instancia es una técnica clave en la visión artificial, y se logra mediante técnicas de segmentación de imagen y clasificación de objetos, que se basan en algoritmos de aprendizaje automático y redes neuronales. Una vez que se han identificado las instancias, se pueden realizar una variedad de análisis y tareas de visión artificial en cada objeto o región, como la detección de características, la identificación de patrones y la clasificación.

En resumen, la instancia en visión artificial se refiere a la identificación de objetos o regiones específicas en una imagen o video, lo que permite un análisis detallado y una variedad de tareas de visión artificial en cada objeto o región.

**El reconocimiento basado en características** es una técnica común en la visión artificial para identificar y reconocer objetos o patrones en imágenes o videos. Esta técnica se basa en la extracción y comparación de características únicas de cada objeto o patrón, en lugar de comparar la imagen completa.

La técnica de reconocimiento basado en características se utiliza en una variedad de aplicaciones de visión artificial, como el reconocimiento de objetos en imágenes, la identificación de huellas dactilares, la identificación de patrones de textura y la detección de anomalías en imágenes médicas.

En resumen, el reconocimiento basado en características es una técnica de visión artificial que se basa en la extracción y comparación de características únicas de objetos o patrones en lugar de comparar la imagen completa, y se utiliza en una variedad de aplicaciones de visión artificial

**La localización de objetos.-** es una técnica clave en la visión artificial que implica detectar la presencia y ubicación de objetos o regiones de interés en una imagen o video.

La localización de objetos puede ser realizada mediante técnicas de segmentación de imagen y análisis de características, y se puede utilizar para una variedad de aplicaciones, como la detección de objetos en imágenes médicas, la detección de peatones en imágenes de vigilancia, la detección de vehículos en imágenes de tráfico, entre otros.

En resumen, la localización de objetos y el reconocimiento de ubicación son técnicas importantes en la visión artificial que permiten detectar la presencia y ubicación de objetos o regiones de interés en una imagen o video, lo que tiene diversas aplicaciones en campos como la medicina, la vigilancia y la automatización industrial

**El uso del contexto.-** es una técnica en la visión artificial que implica utilizar información adicional sobre el entorno en el que se encuentra un objeto o región de interés para mejorar la precisión de su detección o identificación.

Por ejemplo, si se está tratando de detectar un objeto en una imagen, el contexto puede ser información sobre la escena en la que se encuentra el objeto, como el fondo, la iluminación, la presencia de otros objetos, entre otros. Al usar esta información contextual, se pueden crear algoritmos más precisos que pueden detectar y localizar objetos con mayor precisión.

El uso del contexto puede mejorar la precisión de una amplia variedad de aplicaciones de visión artificial, como la detección de objetos en imágenes de vigilancia, la identificación de rostros en imágenes, la detección de peatones en imágenes de tráfico, entre otras.

En resumen, el uso del contexto es una técnica en la visión artificial que implica utilizar información adicional sobre el entorno en el que se encuentra un objeto o región de interés para mejorar la precisión de su detección o identificación. Esta técnica puede mejorar la precisión de una amplia variedad de aplicaciones de visión artificial.

6. **Post Procesamiento:** El post procesamiento implica la mejora de la precisión y la calidad de la imagen después de la clasificación. Las técnicas de postprocesamiento comunes incluyen la eliminación de errores de clasificación y la mejora de la resolución de la imagen.

El postprocesamiento en visión artificial es una etapa que se realiza después de la detección o clasificación de objetos o regiones de interés en una imagen o video. Esta etapa tiene como objetivo refinar los resultados obtenidos a través de la detección o clasificación para obtener una salida más precisa y coherente.

Por ejemplo, después de detectar los objetos en una imagen, se puede realizar un post procesamiento para eliminar los falsos positivos, es decir, objetos que fueron detectados pero no están presentes en la imagen. Para esto, se pueden utilizar técnicas como la eliminación de bordes falsos o la eliminación de objetos pequeños o ruidosos.

Otro ejemplo de post procesamiento en visión artificial es la segmentación semántica, que implica asignar una etiqueta semántica a cada píxel en una imagen. Después de realizar la segmentación semántica, se puede realizar un post procesamiento para mejorar la coherencia de los resultados, como la eliminación de píxeles no deseados o la fusión de regiones pequeñas.

El post procesamiento es importante porque puede mejorar significativamente la precisión y la calidad de los resultados de la detección o clasificación de objetos o regiones de interés. Además, puede ser utilizado en una amplia variedad de aplicaciones de visión artificial, como la detección de objetos en imágenes médicas, la detección de peatones en imágenes de tráfico, la segmentación semántica en imágenes aéreas, entre otras.

### **Aplicaciones.**

La visión artificial es una disciplina que se enfoca en la creación de sistemas que puedan percibir, interpretar y comprender el mundo visual de manera similar a como lo hace un ser humano. A continuación se presentan algunas de las aplicaciones más comunes de la visión artificial:

- **Inspección de calidad:** La visión artificial se utiliza para inspeccionar la calidad de productos y detectar defectos. Por ejemplo, en la industria alimentaria se utiliza para detectar contaminantes y verificar el tamaño y la forma de los productos.
- **Reconocimiento de objetos:** La visión artificial se utiliza para reconocer objetos en imágenes y videos. Por ejemplo, se puede utilizar para reconocer rostros en fotografías y videos de vigilancia.

- **Automatización de procesos:** La visión artificial se utiliza para automatizar procesos en la industria, como el ensamblaje de piezas en líneas de producción.
- **Realidad aumentada:** La visión artificial se utiliza para crear experiencias de realidad aumentada, donde se superponen elementos digitales sobre el mundo real.
- **Vehículos autónomos:** La visión artificial se utiliza para crear sistemas de conducción autónoma, que puedan percibir el entorno y tomar decisiones en tiempo real.
- **Medicina:** La visión artificial se utiliza en medicina para el diagnóstico y tratamiento de enfermedades, como el cáncer y las enfermedades cardiovasculares.
- **Agricultura de precisión:** La visión artificial se utiliza en la agricultura de precisión para el monitoreo y la gestión de cultivos, incluyendo el conteo de plantas, la detección de malezas y la estimación de rendimientos.

Estas son solo algunas de las aplicaciones de la visión artificial. Con el avance de la tecnología y la creciente disponibilidad de datos, es probable que surjan muchas más aplicaciones en el futuro.

### **Desafíos.**

La visión artificial es un campo muy emocionante de la inteligencia artificial, pero también presenta algunos desafíos importantes. Algunos de los principales desafíos son los siguientes:

- **Variabilidad en los datos:** Las imágenes pueden variar mucho en términos de iluminación, ángulo, escala, oclusión, deformación y otros factores. Esto puede hacer que sea difícil para los algoritmos de visión artificial reconocer patrones consistentes.
- **Reconocimiento de objetos en tiempo real:** Los algoritmos de visión artificial a menudo requieren una gran cantidad de cálculos para procesar una sola imagen. Esto puede ser un desafío en situaciones en las que se necesita un reconocimiento de objetos en tiempo real, como en los sistemas de conducción autónoma.
- **Sesgo en los datos:** Los datos utilizados para entrenar los algoritmos de visión artificial pueden estar sesgados en función de factores como la demografía de los sujetos o el entorno en el que se recopilaron los datos. Esto puede llevar a errores en la detección y clasificación de objetos.

- **Interpretación subjetiva:** La interpretación de una imagen a menudo es subjetiva y puede depender de factores como la cultura, la experiencia y la perspectiva individual. Esto puede hacer que sea difícil para los algoritmos de visión artificial dar resultados precisos y consistentes.
- **Privacidad y seguridad:** Las tecnologías de visión artificial a menudo se utilizan para la vigilancia y el reconocimiento facial, lo que plantea preocupaciones sobre la privacidad y la seguridad. Existe el riesgo de que los sistemas de visión artificial puedan ser utilizados de manera inapropiada o incluso malintencionada.
- **Aprendizaje supervisado:** El aprendizaje supervisado es un método común para entrenar algoritmos de visión artificial. Sin embargo, esto puede requerir grandes conjuntos de datos etiquetados, que pueden ser costosos y requerir mucho tiempo para recopilar y anotar.
- **Escasez de datos:** En algunos casos, puede haber una escasez de datos disponibles para entrenar algoritmos de visión artificial. Esto puede ser un desafío en áreas como la medicina, donde los datos pueden ser confidenciales y difíciles de obtener.
- **Adaptación a nuevos entornos:** Los algoritmos de visión artificial pueden ser entrenados en un conjunto de datos específico, pero pueden tener dificultades para adaptarse a nuevos entornos o situaciones. Por ejemplo, un algoritmo entrenado para reconocer objetos en interiores puede tener dificultades para hacerlo en exteriores.

### Herramientas.

Existen muchas herramientas y bibliotecas de software para ayudar a desarrollar soluciones de visión artificial. Aquí hay algunas de las herramientas más populares:

- **OpenCV:** Es una biblioteca de software libre que ofrece un amplio conjunto de algoritmos para procesamiento de imágenes y visión artificial. OpenCV es compatible con varios lenguajes de programación, incluyendo C++, Python y Java.
- **TensorFlow:** Es una biblioteca de software libre para el aprendizaje automático desarrollado por Google. TensorFlow incluye herramientas para el desarrollo de redes neuronales y el procesamiento de imágenes.
- **PyTorch:** Es una biblioteca de software libre para el aprendizaje automático desarrollado por Facebook. PyTorch es una herramienta popular para el desarrollo de modelos de visión artificial y procesamiento de imágenes.
- **Keras:** Es una biblioteca de software libre para el aprendizaje automático desarrollado por la Fundación TensorFlow. Keras se centra en la simplicidad y la facilidad de uso, lo que lo hace ideal para el desarrollo rápido de prototipos de soluciones de visión artificial.

- Caffe: Es una biblioteca de software libre desarrollada por el equipo de investigación de inteligencia artificial de Berkeley. Caffe es una herramienta popular para el desarrollo de soluciones de visión artificial en tiempo real.
- Scikit-learn: Es una biblioteca de software libre para el aprendizaje automático en Python. Scikit-learn incluye herramientas para la clasificación de imágenes y el procesamiento de imágenes.
- MXNet: Es una biblioteca de software libre desarrollada por la empresa de tecnología Amazon. MXNet es una herramienta popular para el desarrollo de soluciones de visión artificial a gran escala.

### Referencias.

- Técnicas y algoritmos básicos de visión artificial / Ana González Marcos... [etal.] (integrantes del Grupo de Investigación EDMANS). – [Logroño] : Universidad de La Rioja, Servicio de Publicaciones, 2006
- Szeliski, R. (2010). Computer Vision: Algorithms and Applications. Springer Science & Business Media.
- ¿Qué es la visión artificial? IBM [¿Qué es la visión artificial? | IBM](#)