

Procesamiento de imágenes digitales con opencv y redes convolucionales

Edwin Dair Zapata Duque. Author^{1, 2}

1. Instituto de Física, Univesidad de Antioquia, Medellín, Colombia

2. dair.zapata@udea.edu.co

Abstract

El procesamiento de imagen es una gran herramienta a la hora de optimizar, analizar y clasificar, por ejemplo se pueden detectar el cancer y melanomas; en el campo industrial es usado para clasificar productor en mal estado. Este trabajo se toma un documento, en este caso el formulario E14 de la registraduría para votaciones al senado del año 2022 el cual se le extrajo cierta información con el objetivo de manipularla después. El lenguaje de programación usado es python y sus librerías opencv y pytesseract.

Keywords: openCV; pytesseract; python

Introducción

Se entiende por procesamiento digital como el conjunto de técnicas que se aplican a las imágenes digitales con el objetivo de mejorar ya sea calidad o extraer alguna información de interés. Uno de los principales procesos es el de filtrado el cual, como resultado de su aplicación, se obtiene otra imagen con ciertas características que nos permiten efectuar ciertas operaciones sobre ella. Entre los objetivos del filtrado se persigue suavizar la imagen, es decir, reducir la cantidad de intensidad entre píxeles adyacentes; también se busca la eliminación del ruido: eliminar aquellos píxeles cuya intensidad es muy diferente a la de los vecinos y cuyo origen es externo a la imagen; Realzar bordes y detectarlos. Bordes son aquellos píxeles donde se produce un cambio brusco en la función intensidad. En definitiva el filtrado se realiza sobre los píxeles.

El proceso de filtrado puede llevarse a cabo sobre los dominios de frecuencia o espacio. En el dominio se asocian las frecuencias de la transformada de Fourier con patrones de variación de las intensidades de la imagen. La frecuencia más baja ($x = 0, y = 0$) corresponde al promedio de los valores de gris de la imagen (x, y) se hacen más

grandes las frecuencias altas corresponden a cambios rápidos en los tonos de gris, que son bordes o ruido, cambios de tonalidad suaves implican frecuencias bajas y cambios bruscos frecuencias altas.

Existen tres tipos básicos de filtros.

1. Filtro paso bajo: Atenúa las frecuencias altas y mantiene las bajas.
2. Filtro paso alto: Atenúa las frecuencias bajas y mantiene las altas.
3. Filtro paso banda: Deja pasar un rango de frecuencias definido entre dos valores dados.

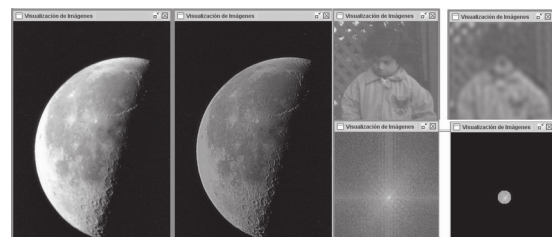


Figure 1: filtrado en el dominio de frecuencia

En el dominio del espacio los filtros tienen como objetivo modificar la contribución de determinados rangos de frecuencias a la formación de la imagen. El término espacial se refiere al hecho de que el filtro se aplica directamente a la imagen y no a una transformada de la misma, es decir, el nivel de gris de un píxel se obtiene directamente en función del valor de sus vecinos aquí las operaciones de filtrado se llevan a cabo directamente sobre los píxeles de la imagen. En este proceso se relaciona, para todos y cada uno de los puntos de la imagen, un conjunto de píxeles próximos al píxel objetivo con la finalidad de obtener una información útil, dependiente del tipo de filtro aplicado, que permita actuar sobre el

píxel concreto en que se está llevando a cabo el proceso de filtrado para, de este modo, obtener mejoras sobre la imagen y/o datos que podrían ser utilizados en futuras acciones o procesos de trabajo sobre ella.

Los filtros en el dominio del espacio pueden clasificarse en:

1. Filtros lineales, asu vez esto se clasifican en filtros paso bajo, paso alto y filtros de banda.
2. Filtros no lineales

La forma de operar de los filtros lineales es por medio de la utilización de máscaras figura 22 que recorren toda la imagen centrando las operaciones sobre los píxeles que se encuadran en la región de la imagen original que coincide con la máscara y el resultado se obtiene mediante una computación (suma de convolución) entre los píxeles originales y los diferentes coeficientes de las máscaras.

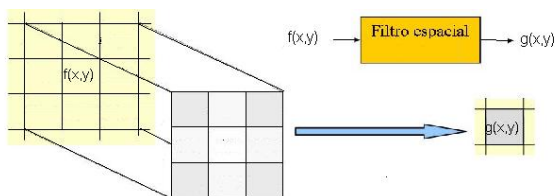


Figure 2: gg



Figure 3: Para realizar un filtrado en el dominio del espacio se realiza una convolución (barrido) del núcleo sobre la imagen

Los filtros espaciales no lineales también operan sobre entornos. Ejemplos de filtros no lineales habituales son los filtros mínimo, máximo y de mediana que son conocidos como filtros de rango. El filtro de mediana tiene un efecto de difuminado de la imagen, y permite realizar una eliminación de ruido, el filtro de máximo se emplea para buscar los puntos más brillantes de una imagen produciendo un efecto de erosión, y el filtro de mínimo se emplea con el objetivo contrario, buscar los puntos más oscuros de una imagen produciendo un efecto de dilatación

Visión por computador

La visión artificial o visión es conjunto de herramientas y métodos que permiten obtener, procesar y analizar imágenes del mundo real con la finalidad de que puedan ser tratadas por un ordenador, se busca producir el mismo efecto hacen los ojos en los humanos: tratar que los maquinas puedan comprender una imagen o secuencia de imágenes y actuar según convenga en una determinada situación. Esta comprensión se consigue gracias a distintos campos como la geometría, la estadística, la física y otras disciplinas. La adquisición de los datos se consigue por varios medios como secuencias de imágenes, vistas desde varias cámaras de video o datos multidimensionales desde un escáner médico. Para esto usa algoritmos matematicos para clasificación de imágenes o toma de decisiones.

Entre otras cosas, la detección de objetos tiene amplio uso cuando se habla de la visión artificial. Aqui aplica cómo detectar la presencia de objetos en una imagen sobre la base de su apariencia visual, bien sea atendiendo al tipo de objeto (una persona, un coche) o a la instancia del objeto (mi coche, el coche del vecino). Generalmente se pueden distinguir dos partes en el proceso de detección: la extracción de características del contenido de una imagen y la búsqueda de objetos basada en dichas características.

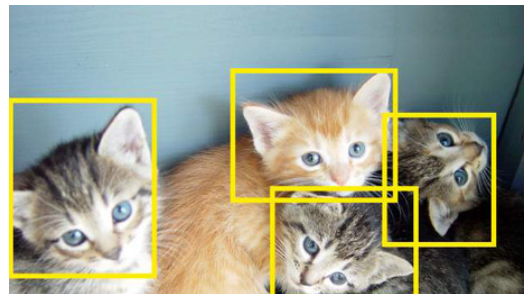


Figure 4: Ejemplo de aplicación de detección de objetos, en este caso, tenemos idenbtificados 4 lindos gatitos.

Para trabajar con imagenes digitales primero que todo se deben binarizar, para esto hay a disposición varios métodos:

1. Simple thresholding(humbralización): es el método más simple de la segmentación de imágenes cuyo objetivo es separar un objeto de interés del fondo de una imagen. Como su nombre lo indica, para el humbral definido lo compara con cada intensidad de los pixeles, si estos son menosres al humbral lo lleva a cero, pero si es mayor, al humbral le otarga un valor de 255. Las siguiembtes figuras ilustran el proceso.
- 2.

Umbral: $T=130$		
130	20	210
$130 > T?$	$20 > T?$	$210 > T?$
35	255	70
$35 > T?$	$255 > T?$	$70 > T?$

Figure 5: Ejemplo de aplicación de detección de objetos, en este caso, tenemos identificados 4 lindos gatitos.

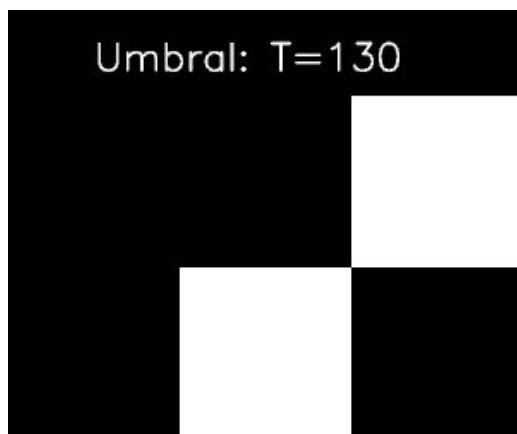


Figure 6: Ejemplo de aplicación de detección de objetos, en este caso, tenemos identificados 4 lindos gatitos.

Python para el procesamiento de imagen

¿Qué es el OCR? proceso dirigido a la digitalización de textos, con el cual se pueden identificar y extraer texto, para luego almacenarlos en forma de datos el cual podremos manipular con algún software. Para llevar a cabo el OCR con Python necesitaremos tesseract, que es la librería que se encarga de todo el trabajo pesado y el procesamiento de imágenes.

Todos los algoritmos OCR siguen 4 etapas en su implementación.

1. Binarización.
2. Fragmentación de la imagen.
3. Adelgazamiento de los componentes.
4. Comparación con patrones.

Binarización Convertir una imagen en blanco y negro, de tal forma que se preserven las propiedades esenciales de la imagen, una manera de usarlo es con un histograma de imagen.

Binarización Detección de los contornos o regiones de la imagen, basándose en la información de intensidad o información espacial.

Adelgazamiento de los componentes Este procedimiento consiste en ir borrando sucesivamente los puntos de los contornos de cada componente de forma que se conserve su tipología.

Comparación con patrones se comparan los caracteres obtenidos anteriormente con unos teóricos (patrones) almacenados en una base de datos.

OpenCV (visión por computadora de código abierto) es una biblioteca de funciones de programación destinadas principalmente a la visión por computadora en tiempo real el cual se utiliza para leer una imagen y realizar ciertas técnicas de procesamiento de imágenes.

Proyecto

La siguiente figura es la primera página del formulario E14 de la registraduría nacional. La intención inicial era crear una red neuronal que fuera capaz de extraer toda la información como la cantidad de votos por partido y por candidato como también el número de mesa y lugar de votación entre otras cosas, sin embargo, aunque sí es posible la implementación requiere más tiempo en investigación, por tal motivo acoté mi proyecto y me propuse extraer lo que tiene que ver con la localización de la mesa de votación como son el departamento, municipio, lugar y zona.

Como dije en la introducción, lo primero que se debe de hacer a una imagen es binarizarla. Mi documento ya está en blanco y negro, sin embargo como ejercicio lo sometí a umbralización el cual entrega la imagen binarizada y los contornos.

Como dije en la introducción, lo primero que se debe de hacer a una imagen es binarizarla. Mi documento ya está

Figure 7: Portada de el formulario E14 de la registraduría para votación al senado



Figure 9: Contornos de la imagen

en blanco y negro, sin embargo como ejercicio lo sometí a humbralización el cual entrega la imagen binarizada y los contornos.

Como dije en la introducción, lo primero que se debe de hacer a una imagen es binarizarla. Mi documento ya está en blanco y negro, sin embargo como ejercicio lo sometí a humbralización el cual entrega la imagen binarizada y los contornos.

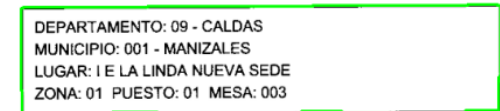


Figure 10: Extracción del contorno de interés al cual se le aplicará OCR para reconocer el texto.

Luego de obtener aislado mi zona de interés aplico OCR el cual me extrae la información dentro del contorno.



Figure 8: Sample thresholding entrega tanto imagen como contornos.

Los contornos se dibujan sobre la imagen original, ya queda aplicar querarquización de contornos y extraer mi zona de interés.

ksdfnskhfksbhfbsfsdfsdf

DEPARTAMENTO: 09 - CALDAS

MUNICIPIO: 001 - MANIZALES

LUGAR: I E LA LINDA NUEVA SEDE

ZONA: 01 PUESTO: 01 MESA: 003

Figure 11: letras

obtengo el data frame

	0	1	2	3	4	5
fila1	DEPARTAMENTO	MUNICIPIO	LUGAR	ZONA	PUESTO	MESA
fila2	09 - CALDAS	001-MANIZALES	E LA LINDA NUEVA SEDE	01	01	003

Figure 12: letras