

Visión artificial - Proyecto

Definición de problema:

Con el pasar del tiempo, la sociedad ha evolucionado, han cambiado tanto las personas como nuestro entorno, llevando al humano a un estilo de vida diferente.

Un ejemplo de esto se puede ver en la manera de alimentarse. El hecho de ir a conseguir alimentos para cocinarlos en casa ha cambiado. Ya sea por falta de tiempo, por incapacidad de cocinar o por simple deseo, los humanos han creado establecimientos que satisfagan tales aspectos y el uso de los mismos se ha vuelto de uso cotidiano para los usuarios.

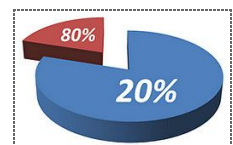
Se sabe que, en Colombia, el 38% de la población come una o más vez a la semana fuera de casa [1]. Con esta cifra podemos notar la gran demanda que se tiene para el servicio de alimentos.

Ahora bien, esto no es solo viable para servicios de alimentación, también se puede tomar la venta de tecnología, de muebles, servicios (peluquería, lavado de carros, entre otros), etc. Podemos ver el hecho de comprar y/o pagar por un servicio es algo cotidiano, que todos llegamos a hacer.

El humano se ha vuelto un ser consumista [2], por lo que siempre está en búsqueda de comprar más cada vez. Si bien puede sonar bien, no quita el hecho de que contribuye a que surjan nuevos establecimientos con el fin de aprovechar esto.

Toda empresa, con ánimo de lucro, siempre está en búsqueda de ganancias, de poder obtener alguna utilidad del trabajo que esta realiza. Esto no es tarea fácil, debido al gran número de competencia que puede surgir con el tiempo, por lo que se debe buscar una manera de lograr captar la atención de los (nuevos o actuales) consumidores [3]

Si bien las estrategias parecen apuntar a captar al consumidor nuevo, no se debe perder de vista al actual. Como señala el principio de Pareto, el 80% de los ingresos de una empresa provienen del 20% de los clientes totales [4]. Con esto podemos mirar que el mantener a los clientes llega a ser bastante importante.



Por otro lado los clientes actuales, en su mayoría, llegan a ser clientes frecuentes, por lo que el hecho de tenerlos asegura compras con regularidad. Aspecto que los clientes nuevos pueden no tener. Otro aspecto a tener en cuenta es el hecho de que el tener clientes frecuentes, los cuales están fidelizados con la marca, puedan llamar a otros, aumentando de esta manera la cantidad de consumidores gracias a la publicidad voz a voz que puedan hacer estos.

Una manera para conservar los consumidores no puede limitarse al hecho de ejecutar publicidad recordatoria, también deberá evaluar otros métodos. Entre ellos puede estar la personalización del servicio [5], exclusividad, beneficios, etc.[6]

En cuanto al aspecto de la personalización, uno de los métodos puede ser el recomendar productos al cliente los cuales pueda estar interesado. O también puede ser el tener precios especiales para estos usuarios.

Si bien, en este momento una aproximación a un “trato especial” puede ser vista en aplicaciones como “leal”, donde un usuario puede acumular puntos al hacer compras en algún establecimiento afiliado para así poder canjear los mismos por productos. Esto solo se puede realizar por medio de una tarjeta con código o teniendo la aplicación en el celular.



Cómo muestra “leal”, el hecho de dar un trato especial al cliente llega a ser más efectivo que el marketing tradicional, siendo hasta 15 veces mejor [7]

Ahora, ¿No sería bueno encontrar un método diferente, el cual no necesite un registro previo y que funcione automáticamente, que trate diferente a los clientes que van con frecuencia al establecimiento dando así un valor agregado?

Objetivo del proyecto:

El objetivo de este proyecto es lograr el reconocimiento de personas en relación a un local. Para esto, se generará información sobre el flujo de clientes que ingresen al local, clasificar los tipos de clientes (frecuentes o nuevos) y lograr reconocer a los clientes frecuentes cuando ingresen de nuevo al establecimiento todo con el fin de poderles dar un trato especial.

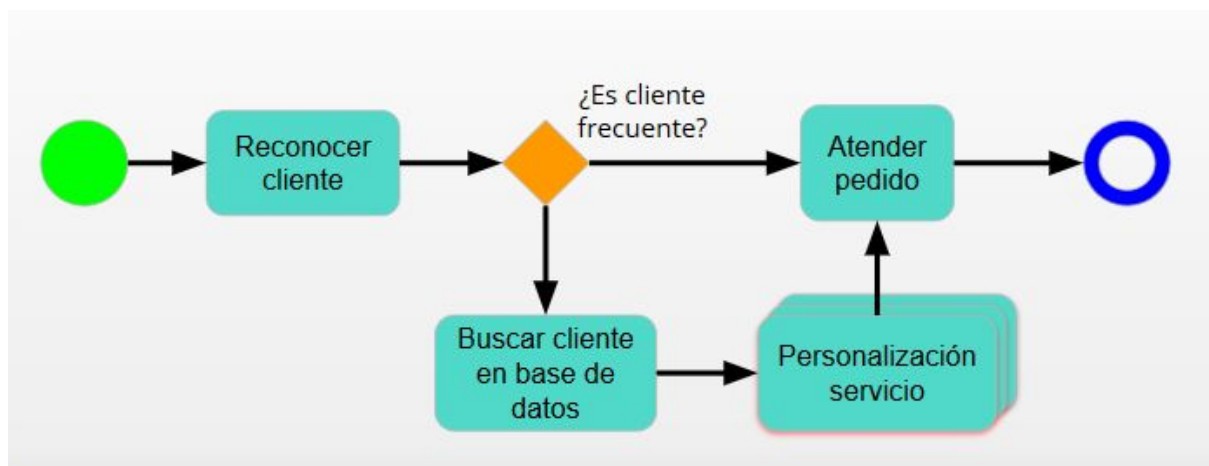
Actividades tentativas:

Se pondrá una cámara en la entrada de un local la cual pueda identificar el rostro del usuario. Una vez obtenido el rostro se procederá a buscar el rostro en la base de datos, con el fin de mirar si el cliente a estado con anterioridad en el establecimiento.

Una vez encontrado el cliente, se procederá a personalizar el servicio que se le será prestado, aplicando descuentos a los productos o recomendando el menú habitual del cliente . En el caso de que no se identifique el rostro por ser un cliente nuevo se prestará el servicio regular.

Por otro lado, se pueden llegar a sacar más datos tomando en cuenta la cantidad de personas que entran. Se podrá tomar el flujo de clientes del establecimiento, como también las horas en las que ingresan más clientes, haciendo que se pueda aprovechar de mejor manera los recursos del establecimiento (empleados, productos disponibles, etc.)

A grandes rasgos, el siguiente modelo representa lo que se realizaría e



En cuanto a los aspectos tecnológicos, se usarán cámaras que, dependiendo las condiciones, pueden ser normales, infrarrojas o con reconocimiento 3D (dependiendo de los recursos que se tengan).

Por otra parte, se usará una base de datos donde estarían almacenado los rostros de los clientes, esto para poder identificar quien ya a estado en el establecimiento con anterioridad y quien no.

Restricciones:

- Las imágenes deben estar a una distancia entre 15 y 30 cm respecto al cliente
- El fondo debe ser plano, con el fin de no afectar la toma del rostro

- Se debe tener una luz blanca para lograr tomar la mayor cantidad de detalles posible sin afectar el tono de la persona.
- La persona debe estar ubicada por el centro de la foto, con el fin de que se logre tomar los detalles del rostro.

Métodos pensados:

- **Bordes:** Con el fin de mirar la máxima cantidad de detalles del usuario, con lo que se podrá comprar después. Se utiliza el filtro de canny, ya que se evidencio que es el filtro que logra mostrar más detalles.
- **Apertura / Cierre:** Se hace con el fin de apreciar de mejor manera los bordes de las imagen. Por otro lado, con el fin de disminuir el ruido que pueda existir en una imagen.
- **Binarización:** Reducción de una imagen a colores blanco y negro.
- **Número de forma de cada área:**

Proceso:

- **Selección de rostros.**

Para la resolución de este problema se probaron dos métodos distintos, con el fin de obtener tener diferentes comparaciones para mejorar el resultado, aquellos serán descritos a continuación:

A. Método 1:

Primeramente se debe capturar una imagen con la cual se puede comparar o guardar en el computador. En nuestro caso, no logramos utilizar la cámara de los computadores por medio de la funcionalidad VideoCapture() de opencv, por lo que se optó por utilizar directamente fotos que hayan sido tomadas desde celular.

Una vez teniendo estas imágenes, se ingresa la imagen al programa donde primero se suaviza un poco la imagen gracias a la función GaussianBlur(), esto con el fin de reducir la presencia de ruido en la imagen, se tuvo la precaución de que no se aplicará de manera agresiva este filtro, ya que podría quitar detalles que lograron ser útiles para la identificación del rostro.

Seguido de esto, se procede a ubicar la cara de la persona y recortarla, con el fin de guardar en memoria lo que es importante para el reconocimiento. Para hacer este reconocimiento primeramente se binariza la imagen gracias a la función threshold(), con lo que se obtiene una imagen en blanco y negro.



Teniendo la imagen binarizada se ubica un óvalo en la mitad de la imagen, donde debería de estar la cara de la persona. Se empieza a recorrer cada fila de píxeles con el fin de contar cuántos cambios de color hay, esto quiere decir que el fondo que es plano tendría un cambio de color (o un número bastante pequeño, ya que se puede obtener ruido de la imagen). La cantidad de cambios de color se guarda, siendo esto el estado inicial de la imagen.



Cambio de color = 1



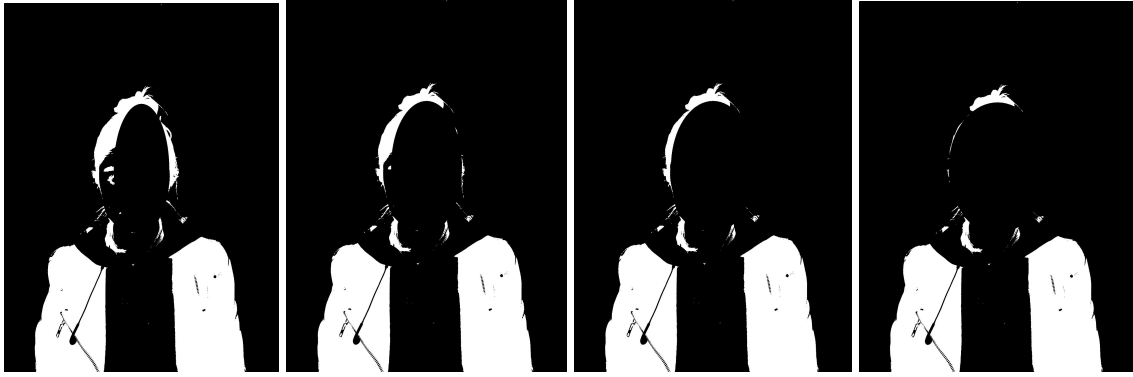
Cambio de color = 5



Cambio de color = 3

Se pudo observar que cuando se tiene la cara descubierta y se cuenta el número de cambios de color es bastante alto, esto debido a los detalles que tiene la cara. Por lo que si se unifica la cara cómo si fuese uno se reduciría esa cantidad, de manera que si se encuentra el óvalo que cubre más correctamente la cara y reduzca la cantidad de cambios de color será el óvalo que está sobre la cara.

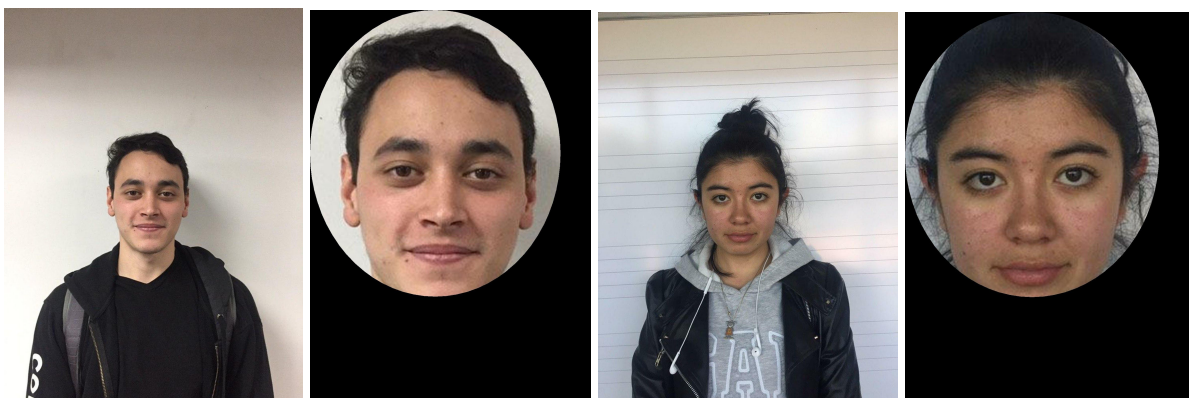




Se itera varias veces cambiando primeramente el ancho del óvalo, aumentando hasta que el número de cambios de color reduzca. Seguido de esto se hace lo mismo, solo que cambiando la altura del ovalo. Por último se prueba moviendo el óvalo desde la izquierda hasta la derecha.

Al terminar este proceso, se obtiene las coordenadas y el tamaño del óvalo por lo que se resta de la imagen original. Logrando aproximaciones bastante cercanas a la cara de la persona.

Por último, se recorta la imagen para que se guarde principalmente la cara de las personas.





Imágenes originales y resultado final.

B. Método 2:

Por otro lado, se consideró conseguir todos los bordes de una imagen por medio de Canny. Donde, teniendo esta imagen se mira aquellos bordes que estén cercanos, con el fin de mirar a estos cómo si fuera una sola área.

Seguido de esto, se binariza la imagen con el objetivo de que todos los valores que no sean de color negro, se vuelven completamente blancos. Gracias a esto último se obtendrán áreas, las cuales representarán las facciones de la cara.

Para finalizar, se consigue el número de forma de cada una de las áreas identificadas anteriormente en la imagen.

Filtrado:

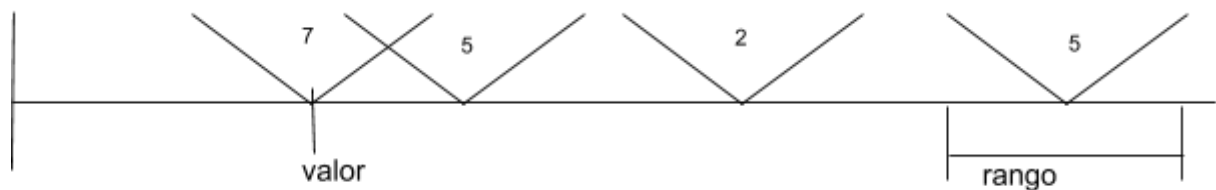
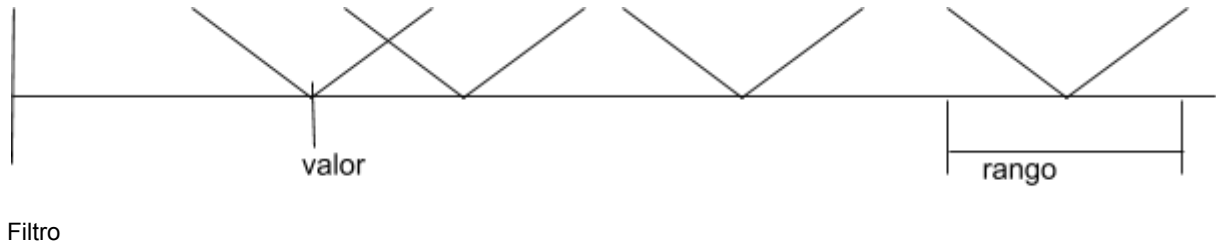
Se utilizan filtros con el fin de buscar imágenes más similares, según las áreas vistas. Esta etapa tiene dos partes:

1. Entrenamiento:

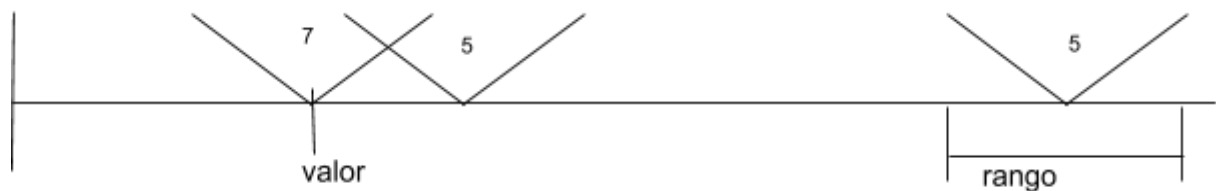
Para el entrenamiento del filtro, a partir de los números de forma previamente conseguidos se ubican en un espacio de números reales positivos donde se le asignan dos cualidades: Rango y Aparición

Rango: Distancia a la que un valor nuevo es visto como igual a un valor ya existente. Todos los filtros tienen el mismo rango.

Aparición: Cantidad de valores que han caído en un mismo filtro. En caso de que un nuevo valor caiga en uno de los filtros, a este filtro se le suma 1 en aparición, de lo contrario se crea un nuevo filtro en la posición del número de forma.



Este proceso se repite sobre un conjunto de imágenes.



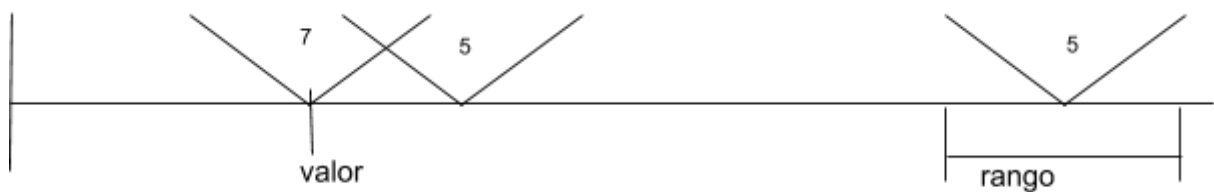
Al final se eliminan los filtros con menor cantidad de apariciones.

Se hace esto con la idea de que los filtros, al haber sido entrenados con un mismo objeto (en este caso caras), las áreas que habrían sido mayormente prevalentes en las imágenes serán las mismas o por lo menos muy similares. Además, al hacer la limpieza se eliminan las áreas con menor cantidad de apariciones.

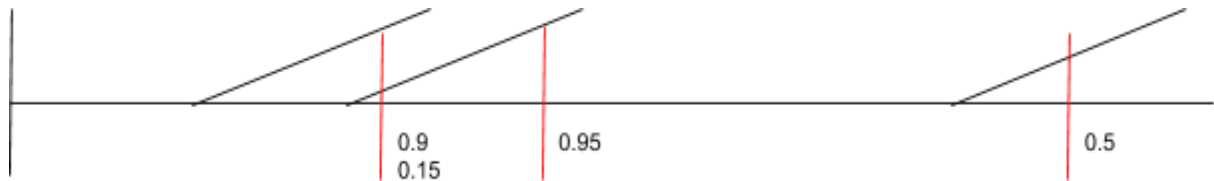
2. Utilización e identificación:

Con los filtros ya entrenados y obtenidos anteriormente, se consiguen los números de forma de cada una de las áreas para luego ser comparados con los filtros. En caso de ser válido, el área se evalúa respecto al filtro y se retorna un valor.

Teniendo como referencia los siguientes filtros:



Se evalúa como:



Con el fin de que se puedan diferenciar los dos lados de cada filtro .

En este caso sería:

Filtro 1 = 0.9

Filtro 2 = 0.15 + 0.95

Filtro 3 = 0.5

Estos valores se ingresan a un vector $V = [0.9, 1.1, 0.5]$

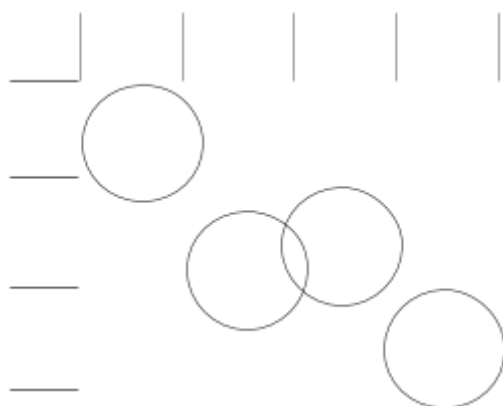
V luego se compara utilizando la distancia euclidiana a los demás vectores conseguidos de imágenes distintas.

En caso de que la distancia sea menor a un límite previamente establecido, se busca el vector con el cual la distancia sea menor.

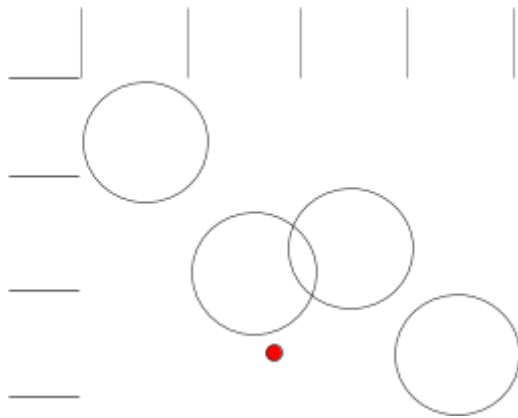
Ejemplo:

Para este ejemplo se utilizaran vectores únicamente con 2 filtros, por facilidad al momento de hacer los dibujos.

En este ejemplo, cada uno de los círculos representa una imagen ya clasificada.

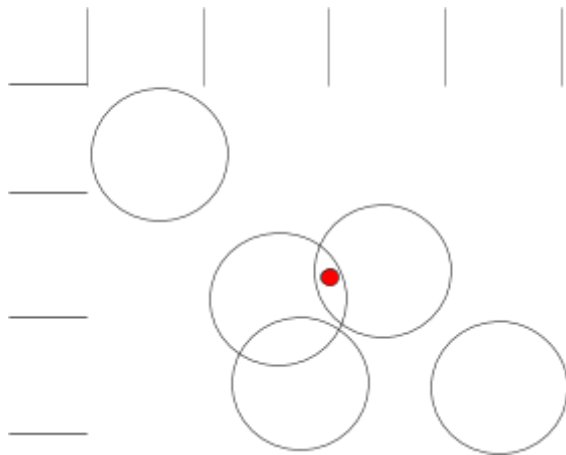


En este caso de un vector $[2.5, 1.7]$, este sería ingresado como :

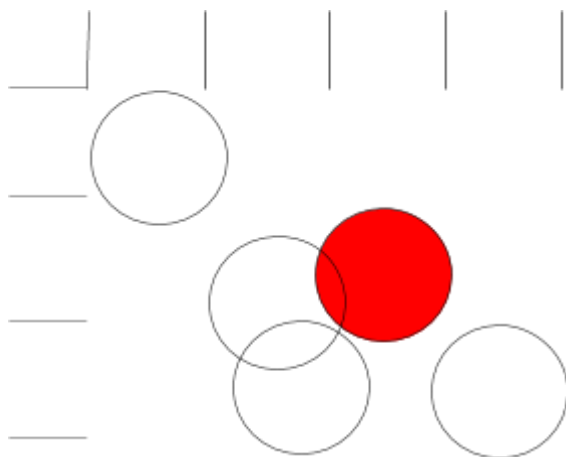


creando una nueva área.

en el caso del vector $[1.7, 2]$



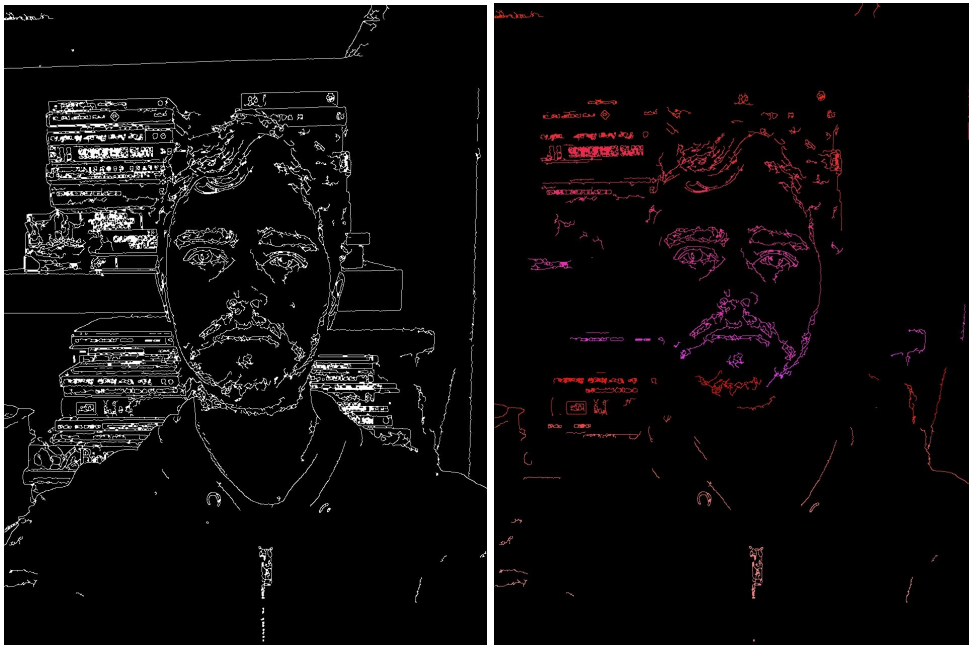
se decidirá el área la cual tenga el centro más cercano al nuevo dato

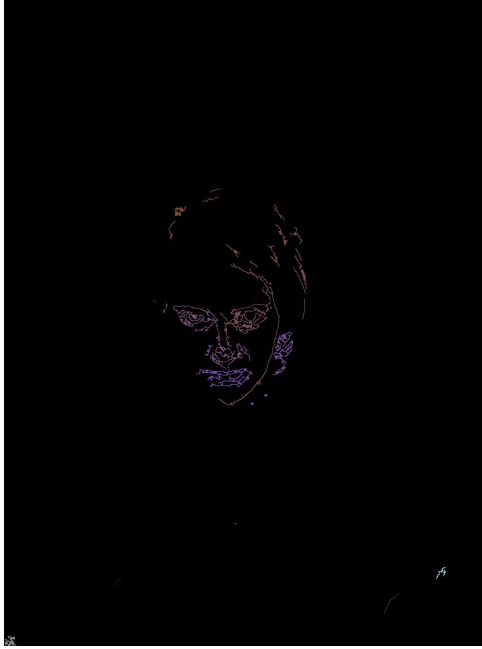


Haciendo que a esta área se le sume 1 en la cantidad de apariciones.

Esta funcionalidad no logró ser completamente implementada debido a no haber encontrado el radio ideal para cada una de las áreas. Además de esto, hubo problemas encontrados con la obtención de los vectores, en los que dos imágenes de personas no relacionadas llegaban a tener una gran cercanía.

Resultados del filtrado:







Conclusiones y problemas encontrados:

- En las imágenes se puede notar que además de salir facciones faciales, también sale bastante ruido del fondo, esto puede deberse a que las imágenes utilizadas para el entrenamiento del filtro no fueron suficientes (18 imágenes).
- Otro problema de este método, es que el método para conseguir el número de forma realizado por nosotros es particularmente lento, haciendo que la identificación de las caras tome bastante tiempo.
- No se logró encontrar los hiper-parámetros adecuados para el funcionamiento correcto del sistema knn, o la forma correcta de representar los datos al momento de compararlos por cercanía. Haciendo que no se haya podido verificar el funcionamiento correcto para la identificación de personas.