

Segmentación de manos basada en redes neuronales.

Su uso en un sistema de control de TV.

Carla Luna Gennari

Directora: Prof. Laura Lanzarini

Co-Director : Prof. César Estrebou

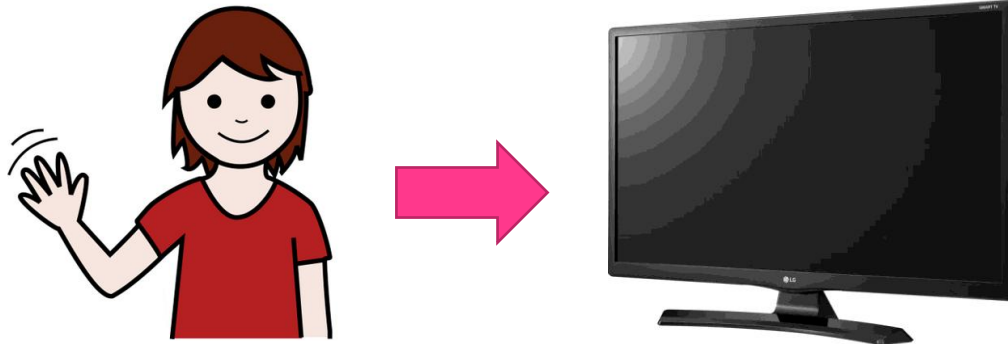


UNIVERSIDAD
NACIONAL
DE LA PLATA

Objetivo

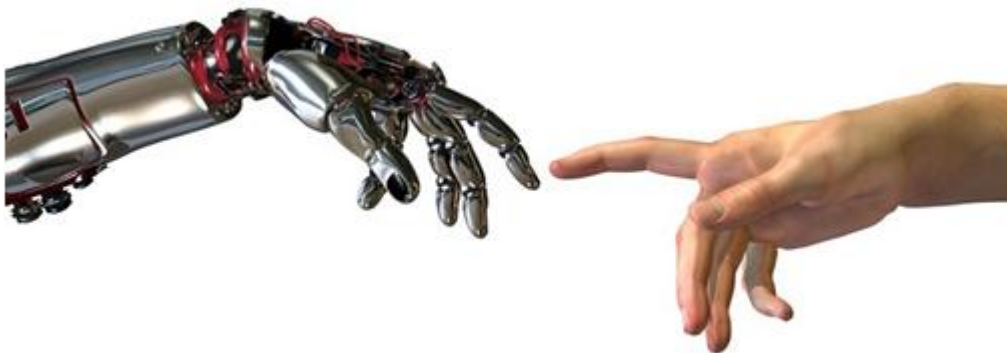


- Diseñar un reconocedor de gestos hechos con las manos sin usar marcadores
 - Segmentación de manos
 - Reconocimiento de gestos
- Construir un prototipo capaz de controlar un TV mediante envío de señales infrarrojas como resultado del reconocimiento de gestos hechos con las manos.



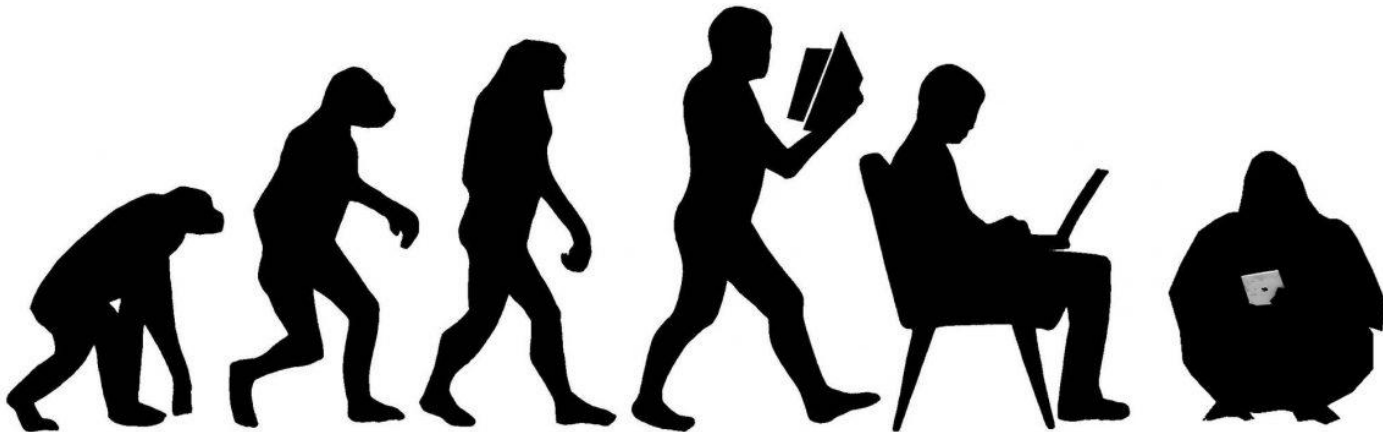
Introducción

- El reconocimiento de gestos permite a los seres humanos comunicarse con las máquinas e interactuar naturalmente sin ningún dispositivo mecánico.
- Traducir los gestos a instrucciones ejecutables por máquinas resulta útil para utilizar computadores sin contacto.



Comunicación Hombre-Máquina

- HCI es el área de la informática que estudia la interacción entre las personas y las computadoras.
- Ventajas
 - Mejorar el uso de dispositivos electrónicos.
 - Las personas con motricidad reducida podrían utilizar dispositivos con facilidad.
 - Independiente del idioma





Agenda

- ❖ Marco teórico
- ❖ Aplicación de control de TV
- ❖ Pruebas y Resultados
- ❖ Conclusiones y trabajos futuros



Reconocimiento de gestos

- Gesto: manifestación de una expresión corporal la cual tiene un significado conocido por lo que puede ser interpretado por las distintas personas.
- Los gestos fueron los primeros medios de comunicación en la antigüedad.



Clasificación de gestos

- Los gestos tienen diferentes definiciones según el contexto. Las clasificaciones pueden variar según el área de estudio.
- Clasificaciones
 - Efrón (1941)
 - Kendon (1981)
 - Mc Neill (1992)
 - Wexelbat (1998)
 - Karam (2002)



Taxonomía de los gestos

- Clasificación en base a las características observables:
 - Estáticos
 - **Dinámicos**
- Clasificación en base a la parte del cuerpo
 - **Hechos con las manos**
 - Hechos con la cabeza y la cara
 - Hechos con el cuerpo
- Clasificación en base al tipo de gesto
 - **En Línea**
 - Fuera de Línea

Dispositivos de captura



- Sensores infrarrojos



- Guantes con cable



- Cámara de profundidad



- Cámara web

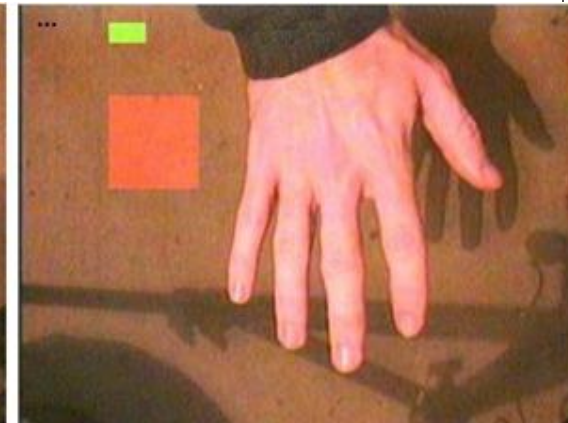
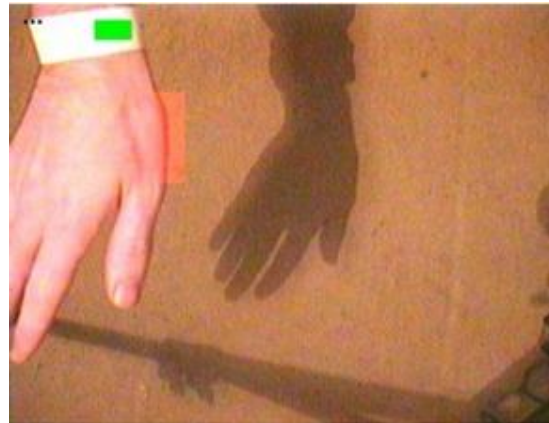
Segmentación de manos

- Segmentación de manos con marcadores



- Clasificaciones para lenguaje de señas

- Clasificación del alfabeto



Visión computacional

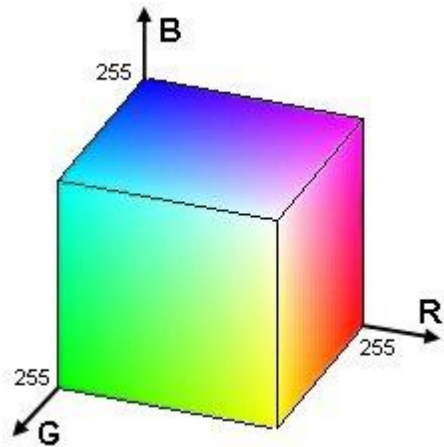
Imagen Digital



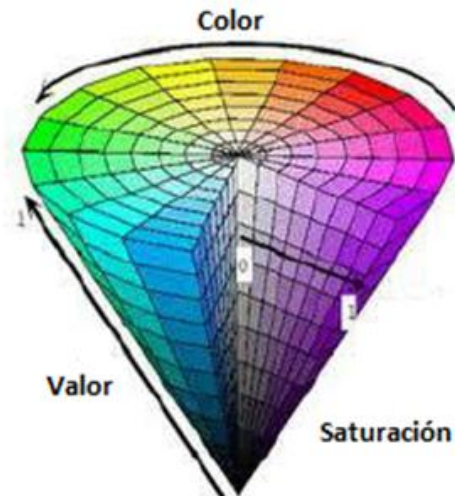
Representación

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

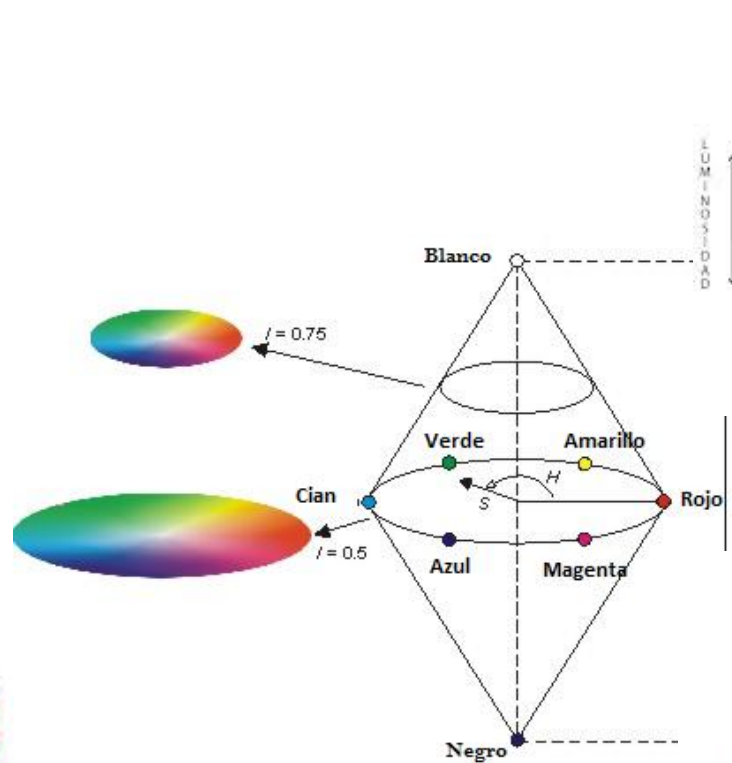
Sistemas de color



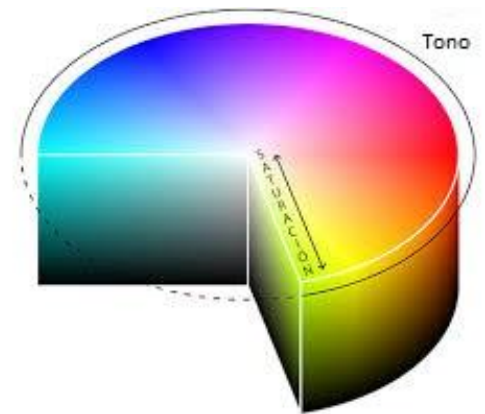
• RGB



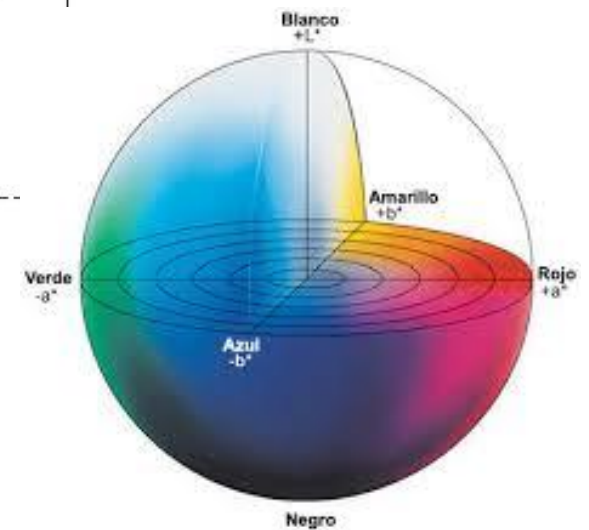
• YCbCr



• HSI



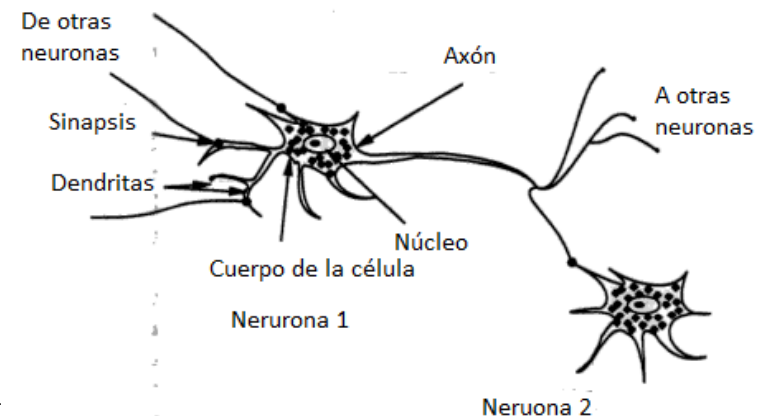
• TSL



• Cie-Lab

Redes neuronales artificiales

- Es un paradigma de procesamiento de información basado en el funcionamiento de los sistemas nerviosos biológicos.
- Características
 - Robustez y tolerancia a fallos
 - Flexibilidad
 - Capacidad de adaptarse a diversas situaciones
 - Gran capacidad de procesamiento



Red neuronal de Energía Coulombica Restringida

- Muchos autores utilizan esta red neuronal
- Se puede utilizar cualquier sistema de color
- El algoritmo es muy sencillo
- Es una forma simple de detectar cuando es color de piel o no.



Entrenamiento de la red

- La red neuronal es entrenada con imágenes de color de piel de distintas personas



- Se calcula la distancia de cada uno de los píxeles con las neuronas existentes.
- Si el píxel no es clasificado se crea una nueva neurona

Entrenamiento de la red

- Luego de recorrer el conjunto de entrenamiento, se evalúa la densidad de las neuronas.
- Las neuronas que no superan el umbral son descartadas.
- Se recorren los píxeles no clasificados hasta que dejan de crearse neuronas nuevas.



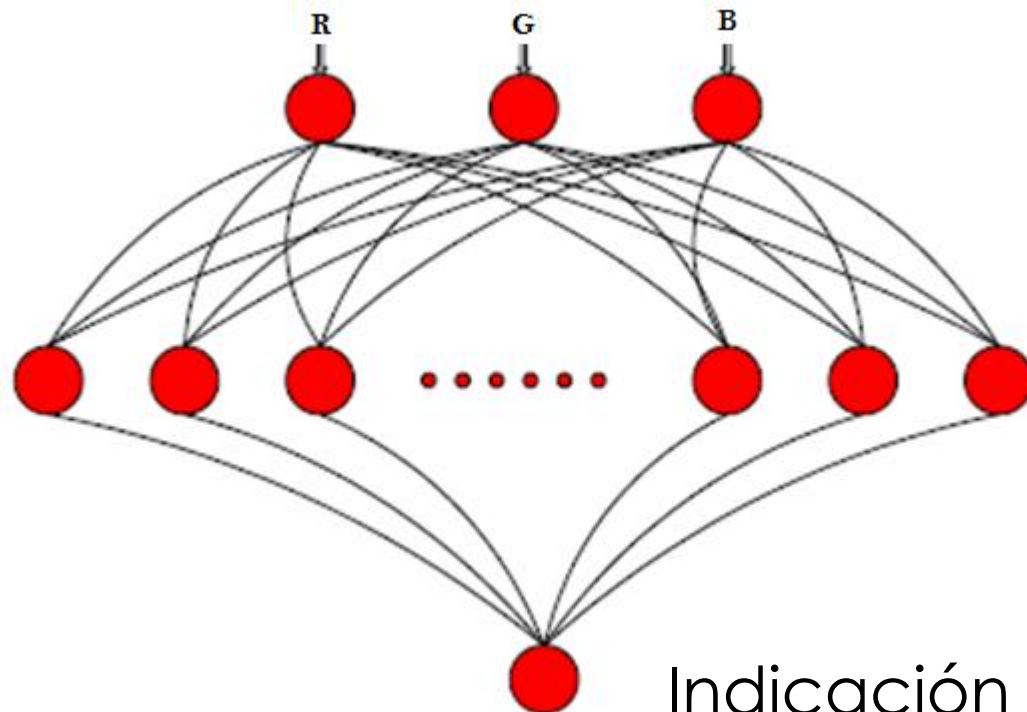
Red neuronal de Energía Coulombica Restringida

- **Arquitectura**

Píxeles de la imagen



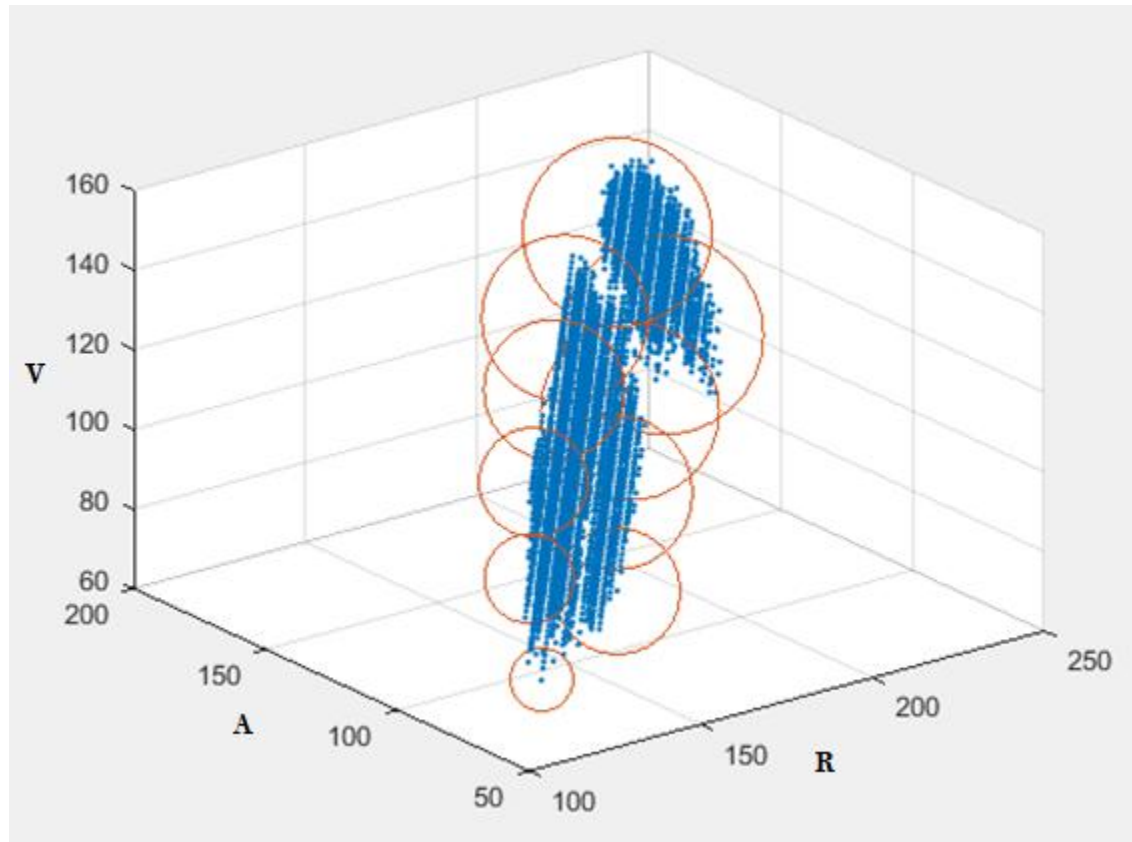
Neuronas Entrenadas



Indicación si es piel o no

Entrenamiento de la red

- Neuronas entrenadas en RGB que conforman los espacios de color de la piel entrenada.





Agenda

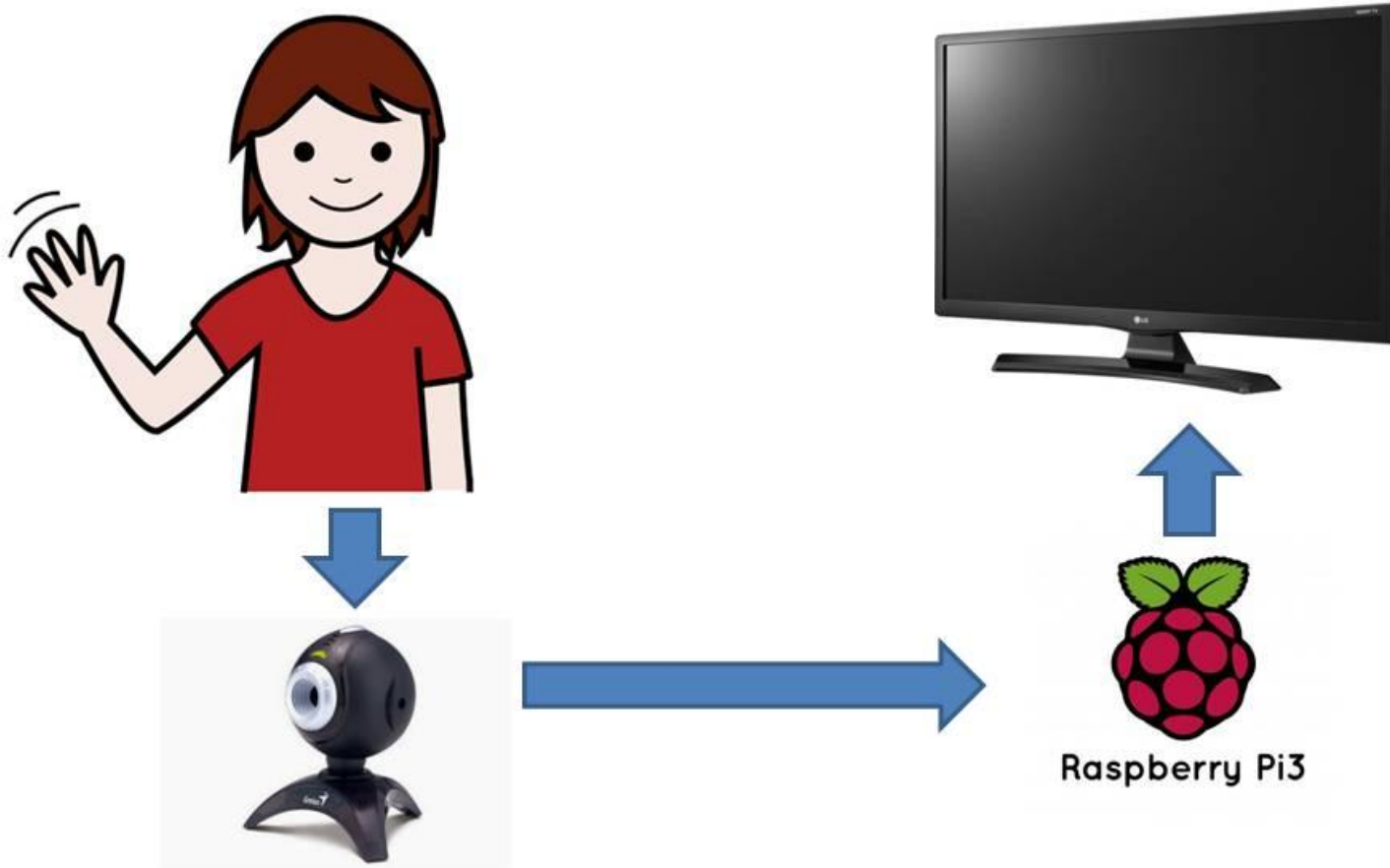
- ❖ Marco teórico
- ❖ Aplicación de control de TV
- ❖ Pruebas y Resultados
- ❖ Conclusiones y trabajos futuros



Aplicación de control de TV desarrollada

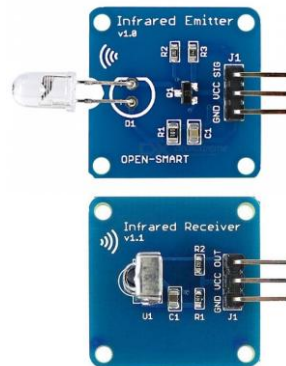
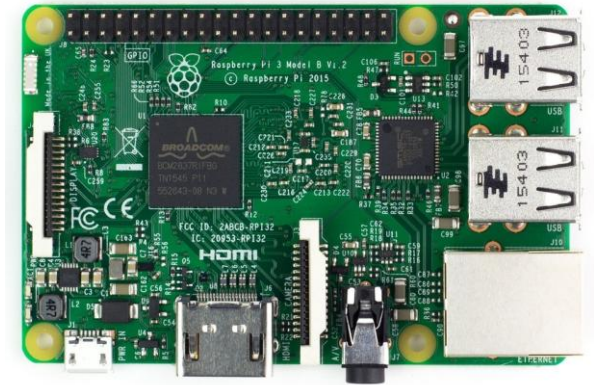
- El aporte central de esta tesina es la construcción de un prototipo que permite controlar un TV mediante envío de señales infrarrojas como resultado del reconocimiento de gestos hechos con las manos.

Aplicación de control de TV desarrollada



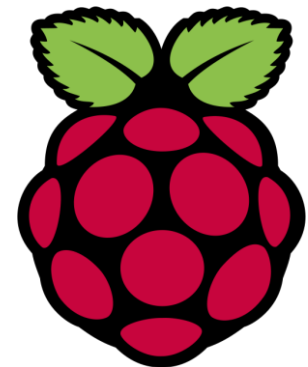
Hardware de captura utilizado

- Raspberry PI3
- Cámara Web
- Sensor de Luminosidad
- Emisor y receptor Infrarrojo



Hardware de captura utilizado

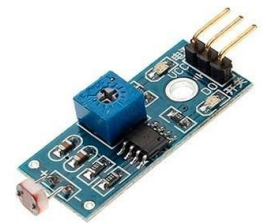
- Ventajas de Raspberry
 - Bajo costo
 - Su tamaño reducido
 - Capacidad de cómputo para procesamiento de imágenes
 - Permite expansión (Wifi, Bluetooth, Entrada/Salida de bajo nivel, HDMI, Audio)
 - Código libre
 - Soporte activo de su comunidad



Hardware de captura utilizado

- Sensor de iluminación

- La intensidad de luz afecta los valores almacenados en los pixeles
- Se utiliza para ajustar la red neuronal RCE
- El sensor devuelve Lux (Unidad de medida de la iluminancia, nivel de iluminación)



Software de reconocimiento desarrollado

- Para realizar el reconocimiento primero se recortan las manos que se encuentran dentro de una imagen
 - Diversos colores de piel
 - Posibles sombras entre objetos
 - Cambios en la intensidad de Luz
 - Fondos complejos



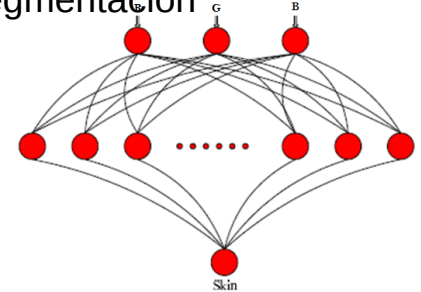
Sistema de reconocimiento desarrollado



Adquisición



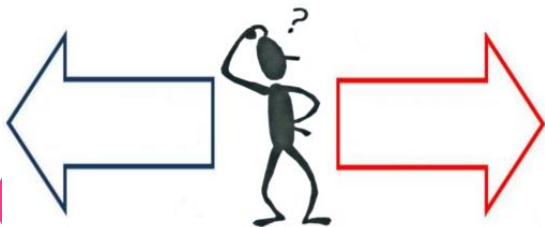
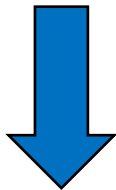
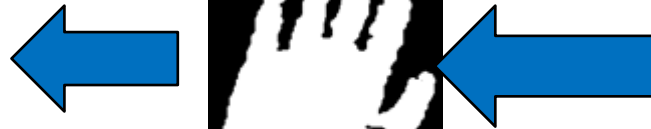
Segmentación



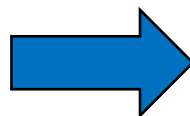
Detección del gesto



Extracción de características

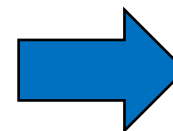


Detección del movimiento



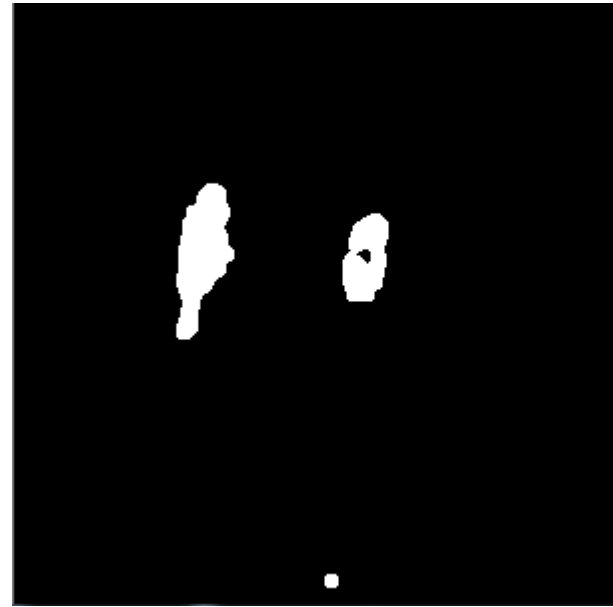
LIRC

Linux Infrared Remote Control
Ejecución de la acción



Software de reconocimiento desarrollado

- Procesamiento de la imagen con la red neuronal RCE



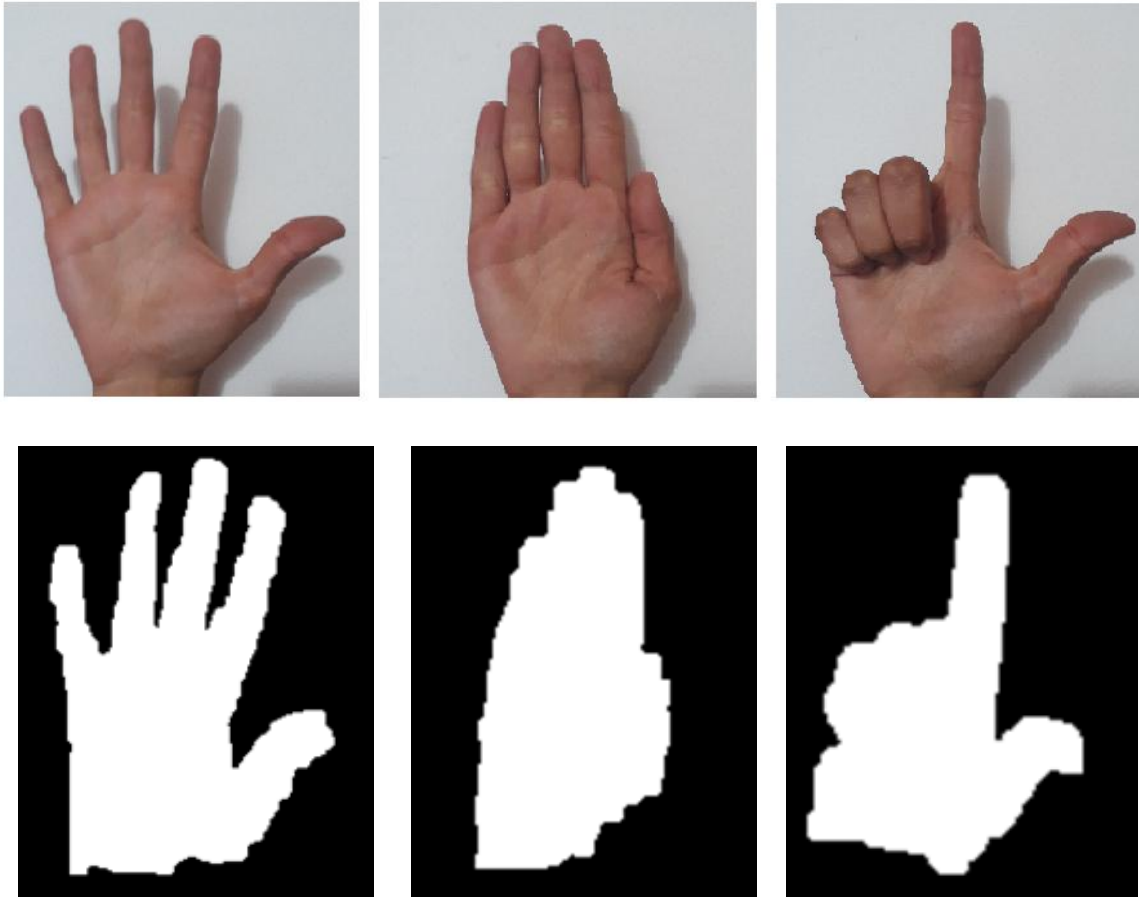
Software de reconocimiento desarrollado

- Mejoramiento de contornos
 - Operación de clausura
- Extracción de regiones
- Cálculos de tamaño



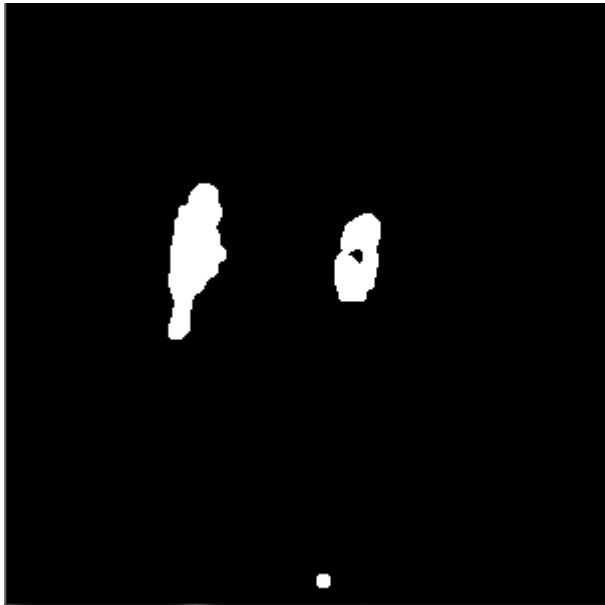
Software de reconocimiento desarrollado

- Template matching



Software de reconocimiento desarrollado

- Detección de la mano en la imagen



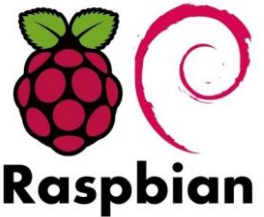
Software de reconocimiento desarrollado

- Movimiento
 - Con la posición del centroide de la mano detectada se evalúa la dirección en que se realiza el movimiento



Software de reconocimiento desarrollado

- Luego se identifica el comando a ejecutar
- El comando es enviado mediante la biblioteca LIRC (Linux Infrared Remote Control)
- Esta biblioteca permite decodificar y enviar señales de infrarrojos de muchos controles remotos comúnmente utilizados
- El comando enviado es interpretado por el TV





Agenda

- ❖ Marco teórico
- ❖ Aplicación de control de TV
- ❖ Pruebas y Resultados
- ❖ Conclusiones y trabajos futuros



Pruebas y resultados

- BBDD para medir el desempeño de la red

Para
medir
aciertos



BBDD MOHI
250 Imágenes

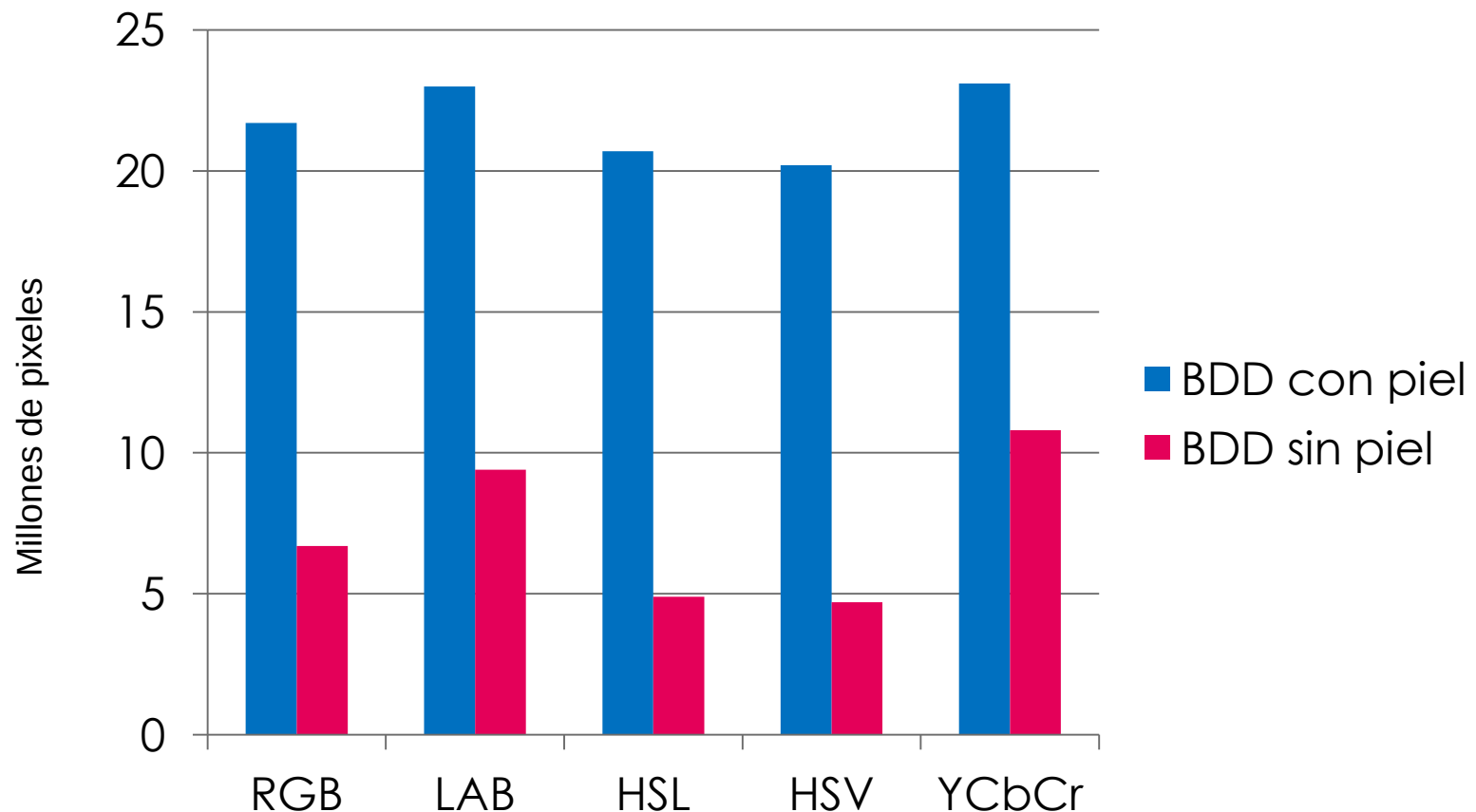


BBDD Ad-Hoc
250 Imágenes

Para
medir
errores

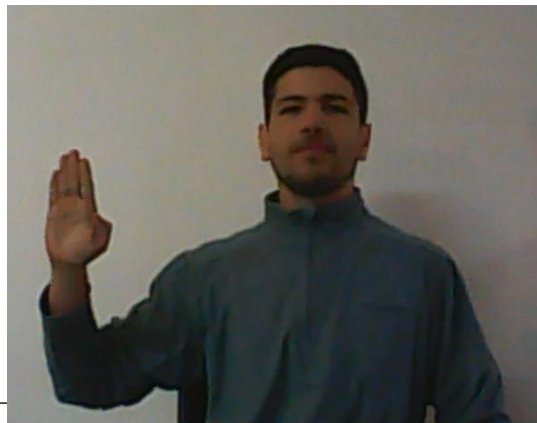
Pruebas y resultados

- **Red Neuronal RCE aplicada a los distintos sistemas de color**



Pruebas y resultados

- Características de la BBDD
 - 11 sujetos de prueba
 - 2 condiciones de iluminación (30 y 600 lux)
 - 2 distancias (1,80m y 3m)
 - 3 formas de manos
 - 4 movimientos: arriba, abajo, izquierda, derecha
 - 12 gestos en total (462 videos)



Pruebas y resultados



- Subir volumen



- Aumentar canal



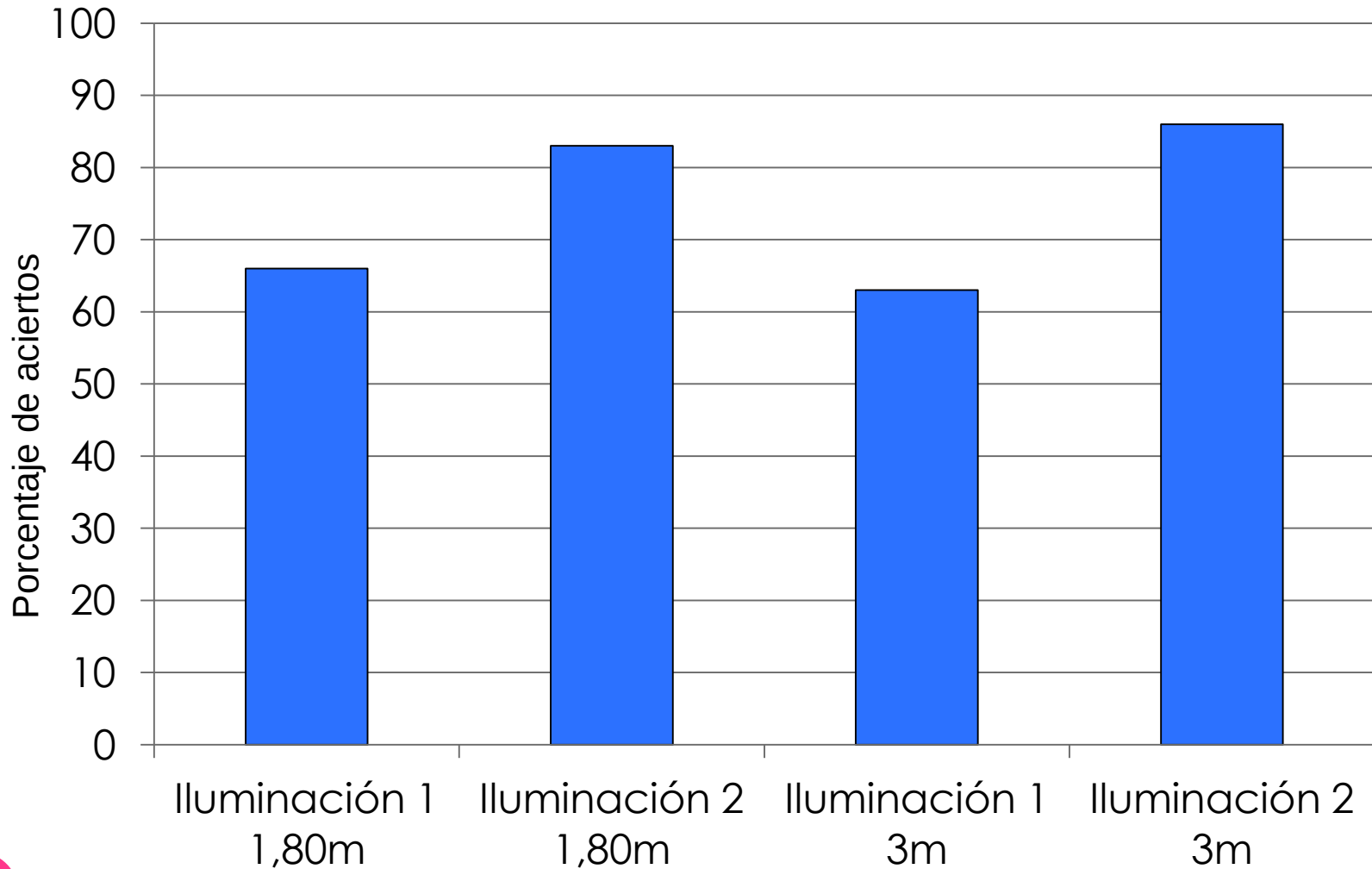
- Bajar volumen



- Disminuir canal

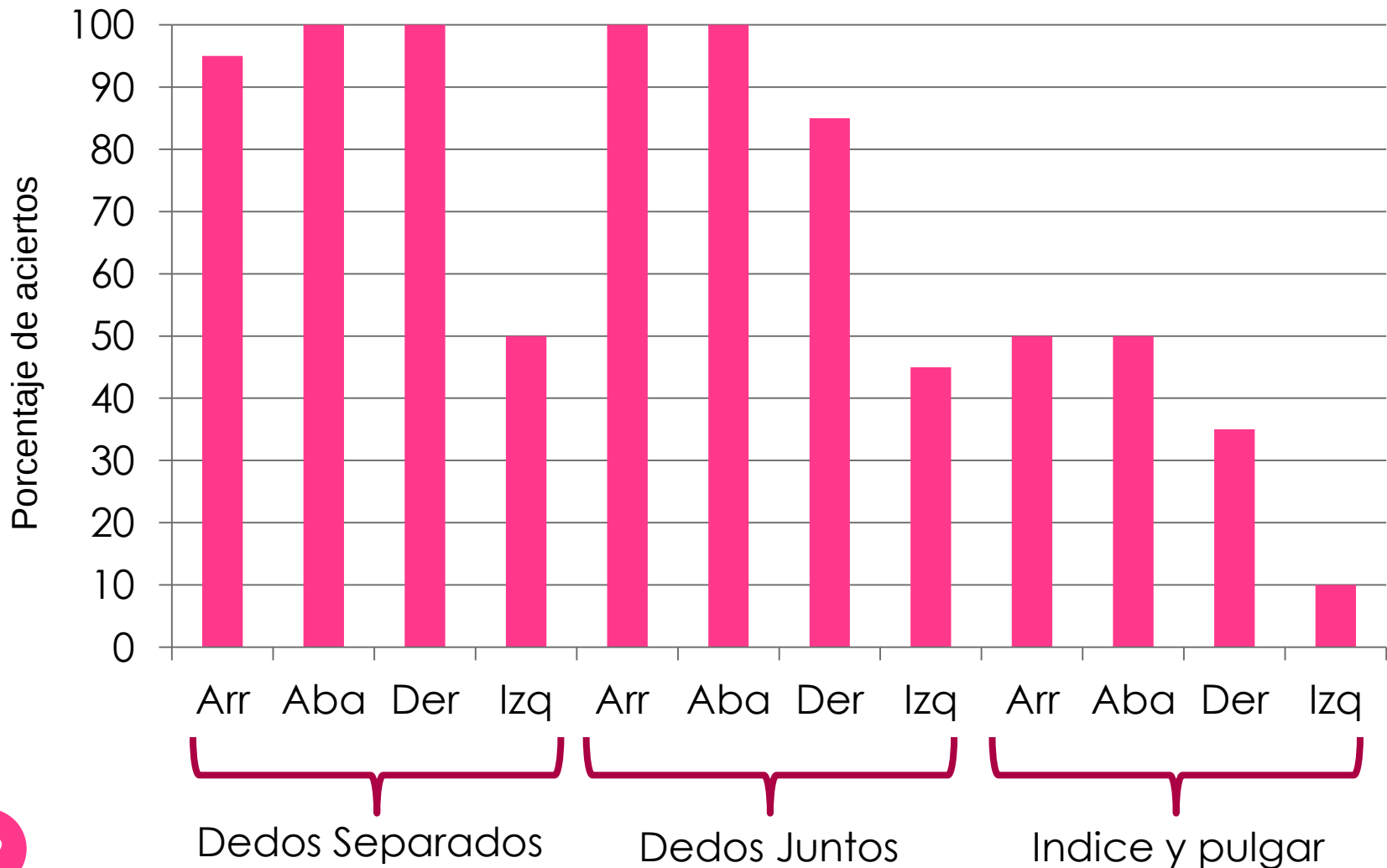
Pruebas y resultados

- **Tasa de acierto promedio para cada escenario**



Pruebas y resultados

- Tasa de acierto promedio para cada gesto



Pruebas y resultados

- Fallas en el reconocimiento

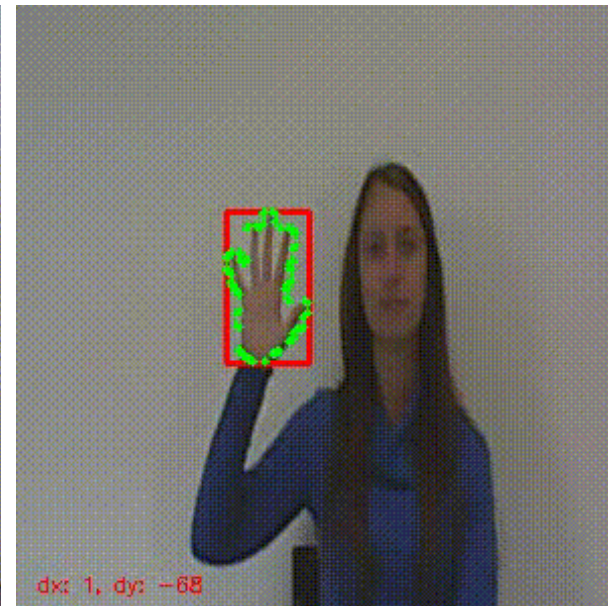
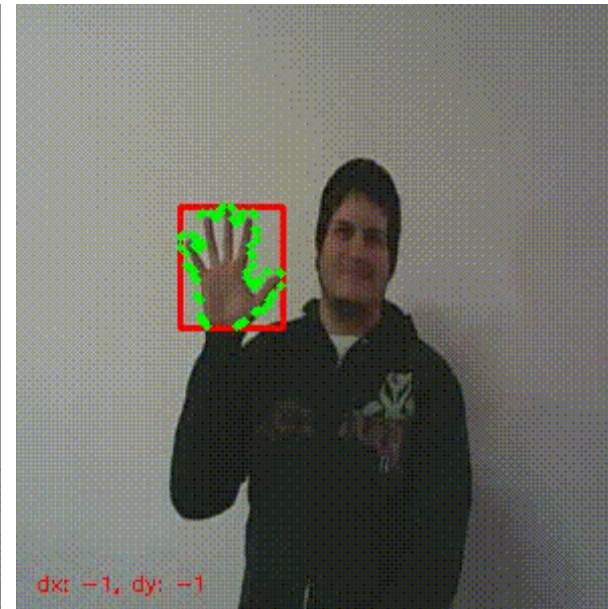
- Manos sobre la cara



- Sombras entre los dedos



Pruebas y resultados





Agenda

- ❖ Marco teórico
- ❖ Aplicación de control de TV
- ❖ Pruebas y Resultados
- ❖ Conclusiones y trabajos futuros ←

Conclusiones

- Los resultados de la segmentación de la piel con la red neuronal RCE fueron satisfactorios.
- La red neuronal se puede utilizar con cualquier sistema de color.

Conclusiones

- El dispositivo construido es capaz de reconocer gestos simples.
- Cuando la intensidad de luz es baja, se dificulta la segmentación. La corrección de la intensidad de luz no es suficiente para resolver el problema

Conclusiones

- Las distancias evaluadas dieron resultados satisfactorios, no sería necesario un sensor de profundidad al utilizarse con distancia menor a 3m.
- Los resultados fueron publicados y expuestos en el Workshop Innovación en Sistemas de Software (WISS) de CACIC 2017.

<http://sedici.unlp.edu.ar/handle/10915/63867>

Trabajos futuros

- Reemplazar Raspberry PI para reducir el costo.
- Mejorar la segmentación incorporando al entrenamiento de la red neuronal, además del color de la piel, información de la sombra.
- Adaptar una cámara de luz infrarroja para ambientes con baja intensidad de luz.

¿Preguntas?

