



DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN CARRERA DE SOFTWARE

ARTÍCULO ACADÉMICO PREVIO A LA OBTENCIÓN DEL TÍTULO
DE INGENIERO EN SOFTWARE

TEMA: "Re-identificación de personas mediante características biométricas blandas: silueta corporal y textura de la ropa en un entorno de videovigilancia multicámara."

AUTORES:


**CHANGO CAISABANDA, DAVID BERNARDO
CUYO CHILUISA, VICTOR HUGO**

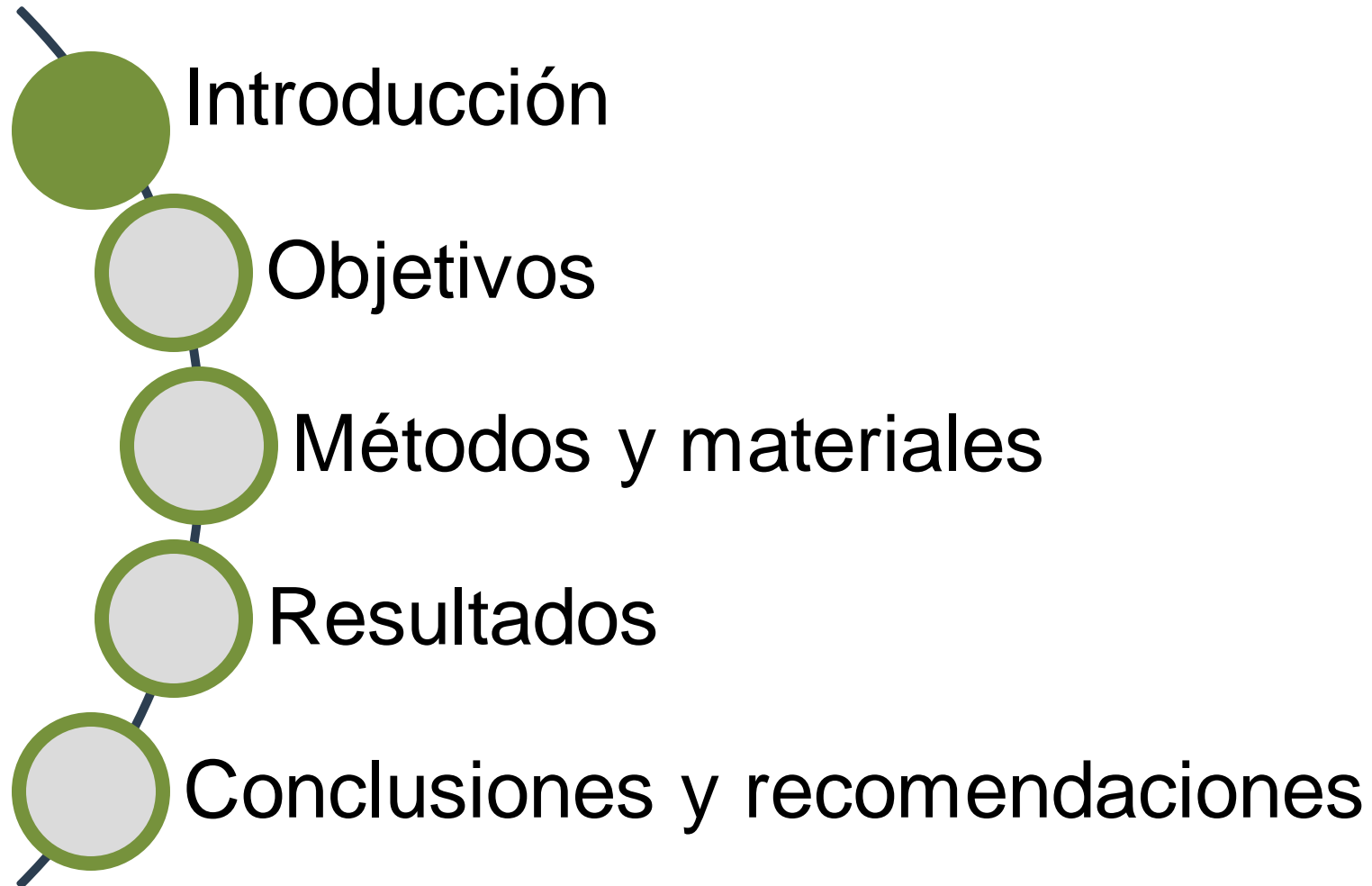
DIRECTOR:

DR. CARRILLO MEDINA, JOSE LUIS

**LATACUNGA
AGOSTO, 2022**



- 
- Introducción
 - Objetivos
 - Métodos y materiales
 - Resultados
 - Conclusiones y recomendaciones



Problemática



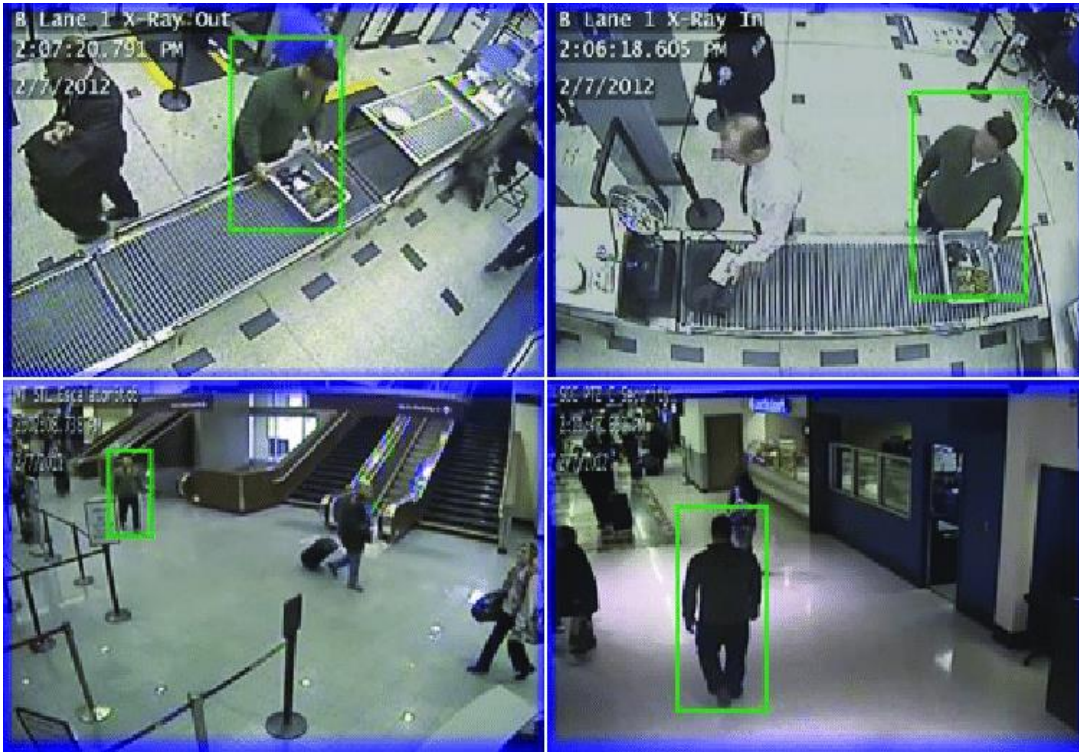
Un incremento notable de la **inseguridad** de las personas dentro de instituciones públicas y privadas.

La video vigilancia **tradicional** no está exenta de **errores humanos**, por parte del vigilante.

La extracción de características biométricas duras (huella dactilar, rostro, etc) **es compleja**.

Introducción

Re-Identificación

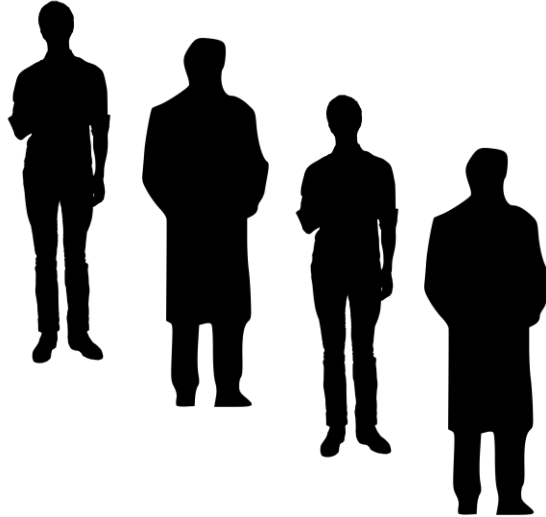


La Re-Identificación de personas consiste en la **recuperación de instancias** entre cámaras no superpuestas, que se encarga del seguimiento **de personas**.



Está atrayendo mucha atención debido a su demanda y al crecimiento de las redes de cámaras instaladas dentro de aeropuertos, campus universitarios, edificios, etc., **cuyo propósito es la seguridad**.

Patrones y características (Silueta humana y textura de vestimenta)



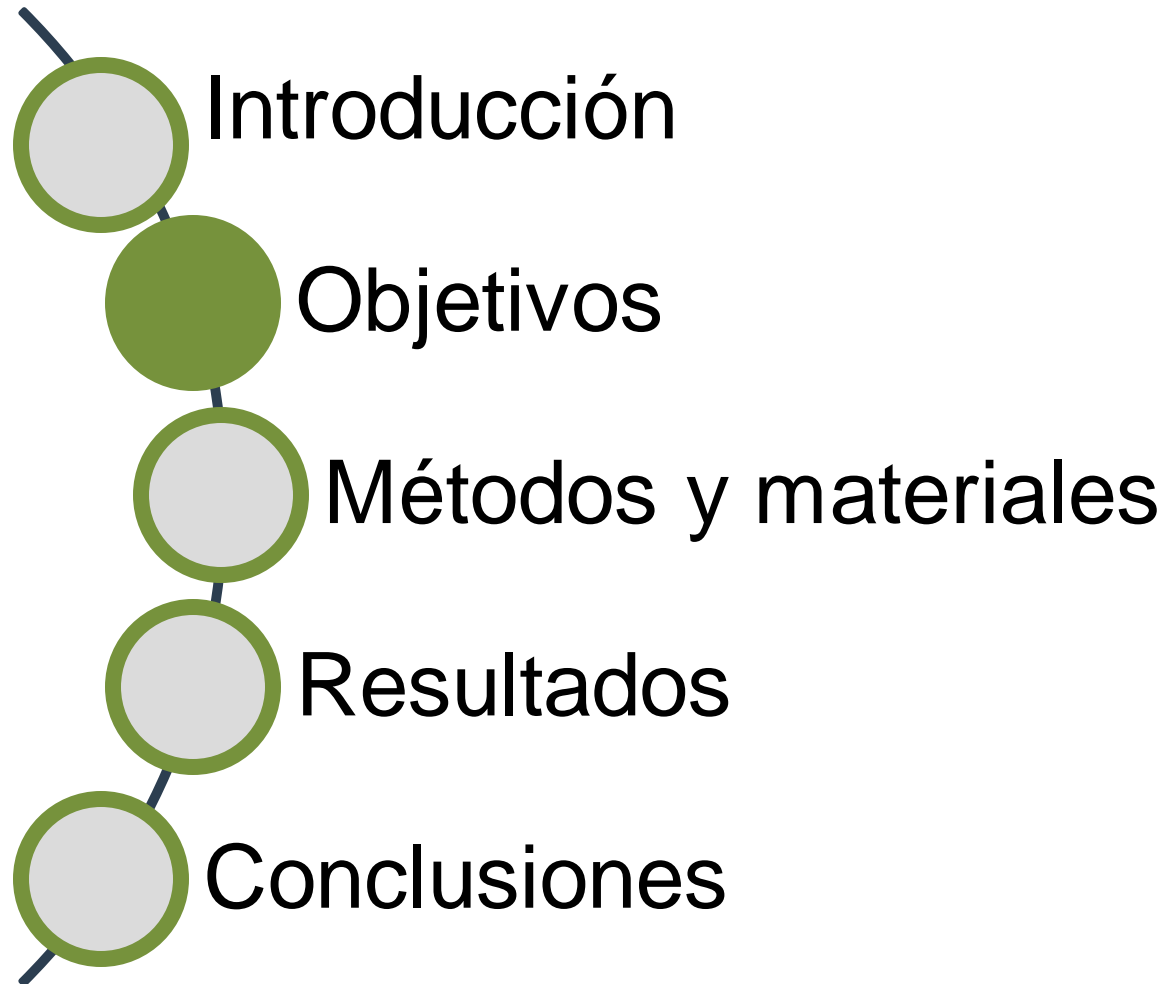
Silueta corporal: la precisión del 77,40%, con HOG y SVM.

No tienen un gran **poder discriminatorio** como las biométricas duras, sin embargo, son de gran ayuda en la identificación de personas.



Textura: el porcentaje de precisión es del 48.12%, con K-means y SVM.

Gracias a resultados de diversos **trabajos realizados**, se **evidencia** que estas características son **viabiles** para una correcta Re-Identificación





OBJETIVO GENERAL

Desarrollar un sistema de Re-Identificación y seguimiento de personas usando sus características soft-biométricas: silueta del cuerpo y textura de la vestimenta, con algoritmos de Visión por Computadora y Machine Learning.

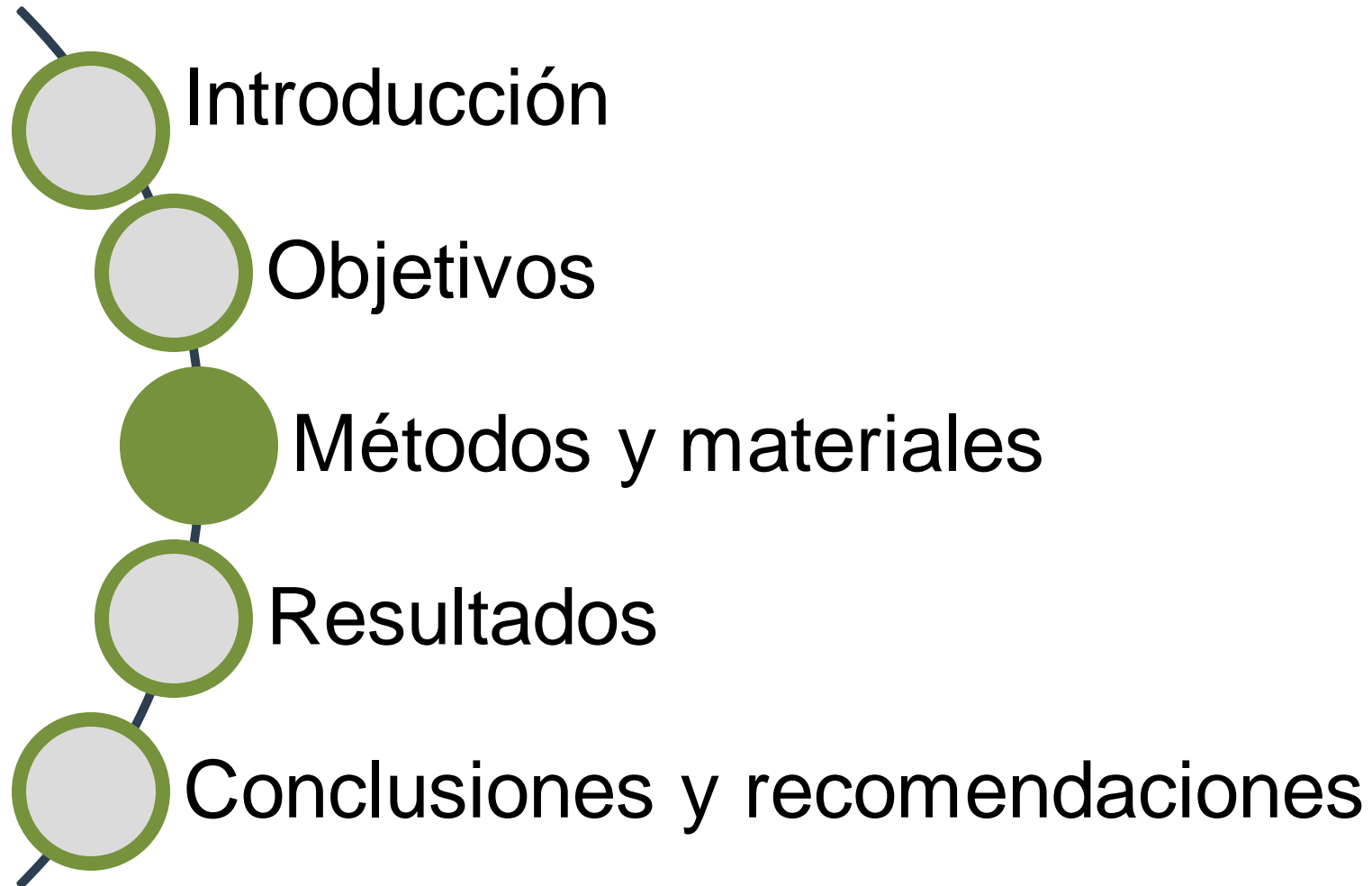
Investigar el estado del arte sobre métodos y técnicas para la Re-ID de personas utilizando características soft-biométricas: silueta del cuerpo y textura de la vestimenta, con algoritmos de Visión por Computadora y Machine Learning.

Implementar un sistema de Re-ID a través del uso de características soft-biométricas: silueta del cuerpo y textura de la vestimenta de personas, con algoritmos de Visión por Computadora y Machine Learning.

Evaluar el sistema y analizar los resultados obtenidos durante su ejecución.



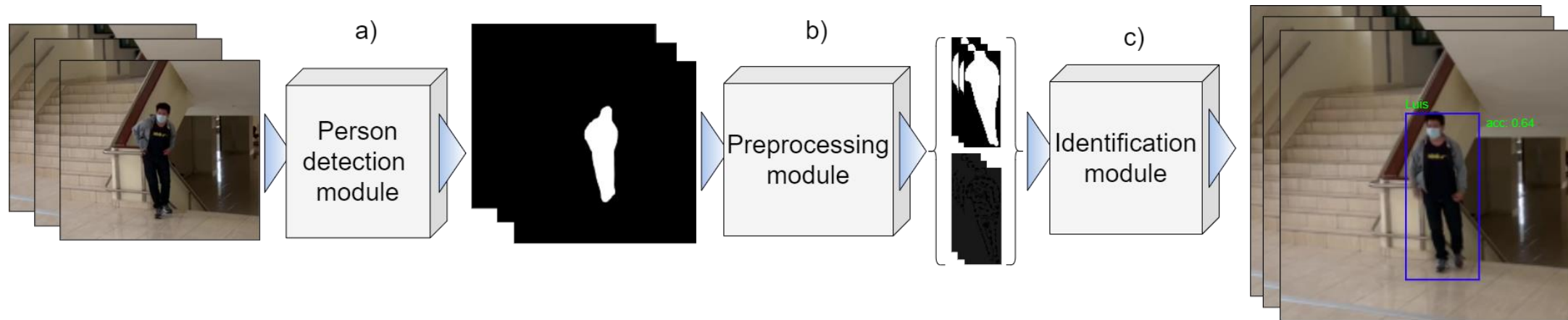
OBJETIVOS ESPECÍFICOS



Construcción del sistema



La construcción del sistema se hizo utilizando el lenguaje de programación Python.

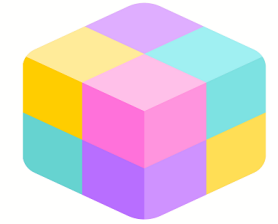


Esquema funcional del sistema de Re-Identificación de persona se conforma por tres módulos:

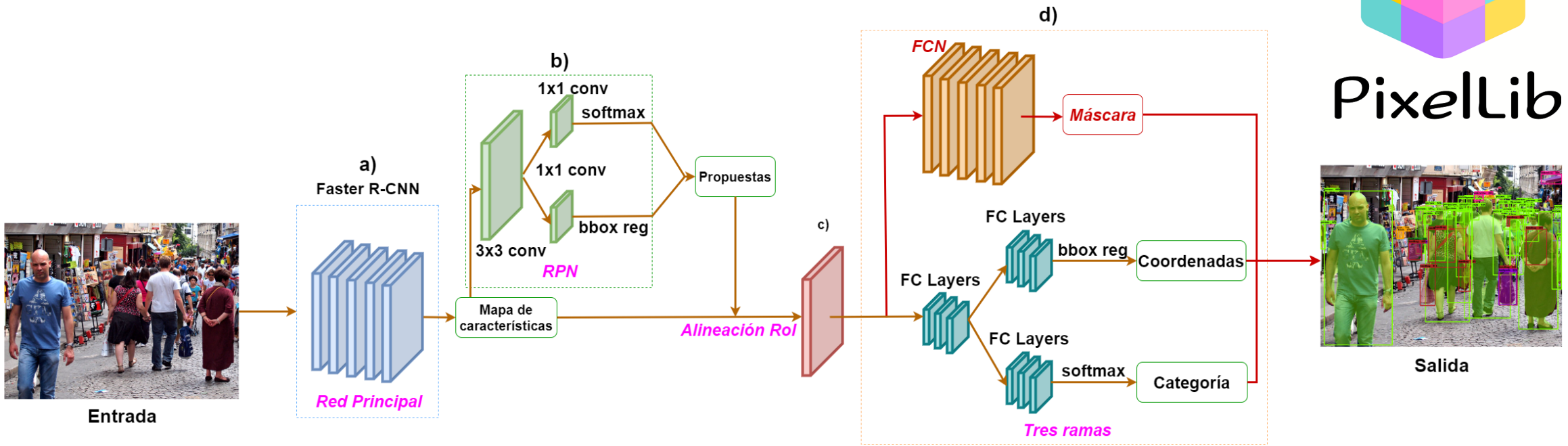
- a) Detección y segmentación de personas.
- b) Preprocesamiento.
- c) Identificación.

Detección y segmentación de personas

Se utilizó un modelo pre-entrenado de la librería PixelLib.



PixelLib

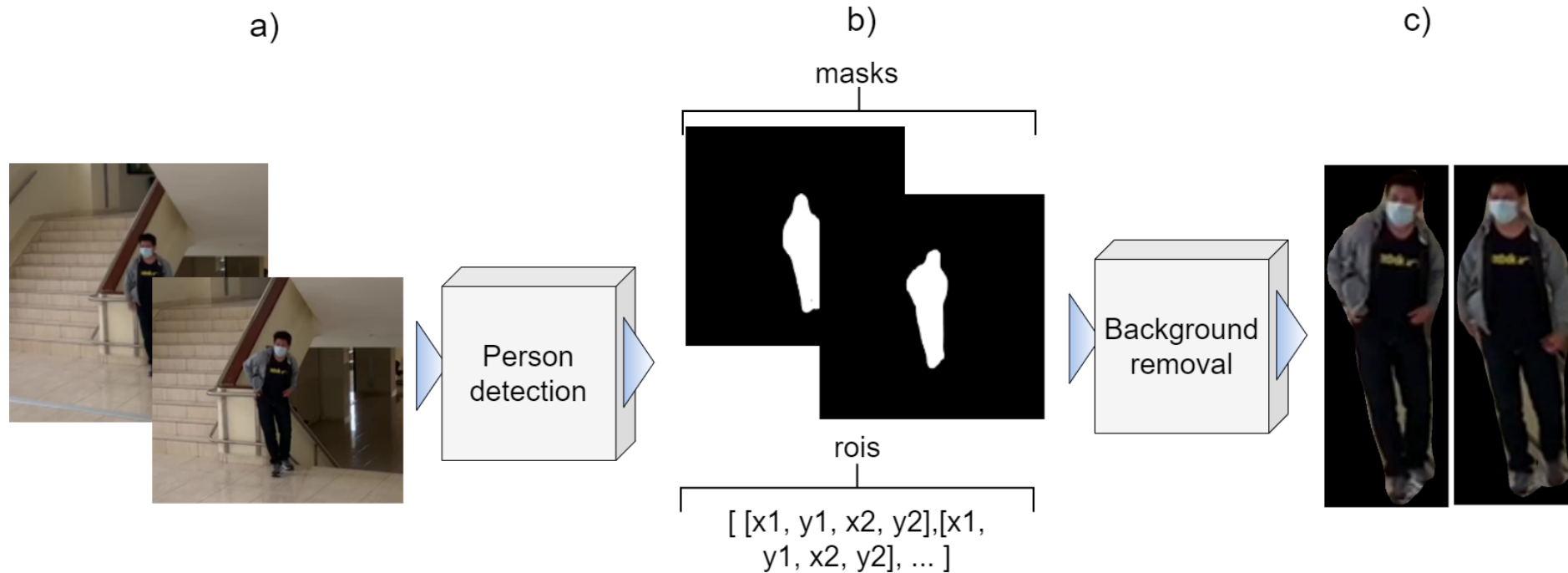


El modelo Mask R-CNN para la detección y segmentación de personas. Consta de:

- Una red troncal basada en Faster R-CNN.
- Una red de propuesta de regiones.
- Mapeo de regiones de interés generadas.
- Generación de máscara binaria, cuadros delimitadores y precisión de la clasificación.

Preprocesamiento

Eliminación de fondo



Método de sustracción de fondo, que consiste en:

- a) Lectura de la imagen.
- b) Generación de máscaras binarias y coordenadas de cajas delimitadoras.
- c) Recorte de la imagen y eliminación del fondo.

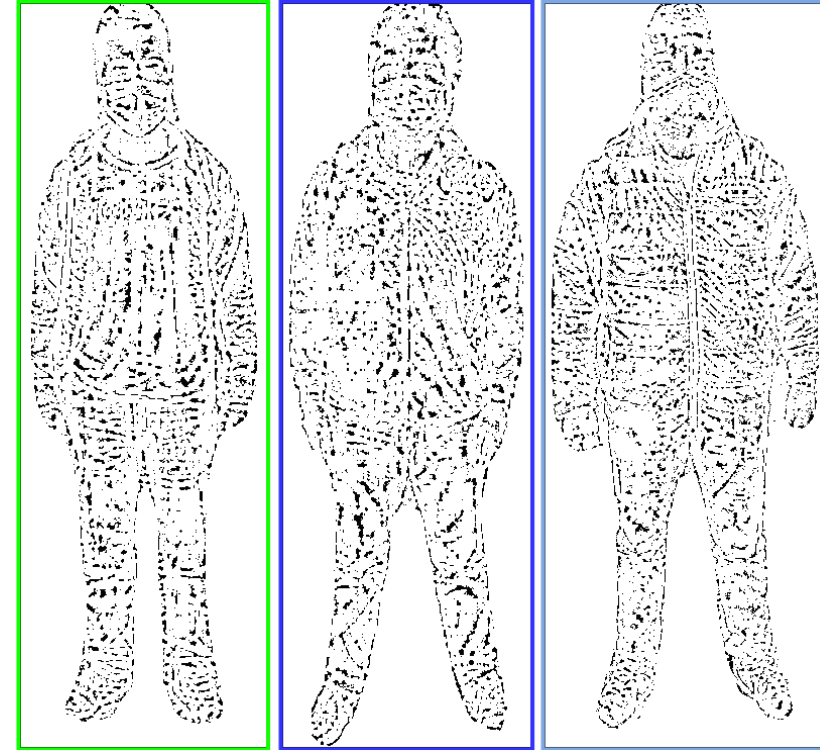
La eliminación del fondo se utiliza ampliamente en escenarios en los que es necesario detectar objetos en primer plano en secuencias de imágenes.

Preprocesamiento

Extracción de las texturas de la ropa Patrón Binario Local Unitario (LBPU)

El filtro **LBP** etiqueta cada pixel de una imagen al analizar su vecindario, estudia el nivel de gris de cada píxel y codifica esta comparación utilizando **números binarios**.

El filtro **LBPU** genera un descriptor de texturas uniforme más **robusto** que otras de sus variaciones, por su **invariancia ante rotaciones**, **simplicidad computacional**, y **tolerancia a cambios de iluminación**.



Imágenes filtradas mediante LBPU.

Preprocesamiento

Extracción de la silueta corporal de la persona
Máscara binaria (Silueta)

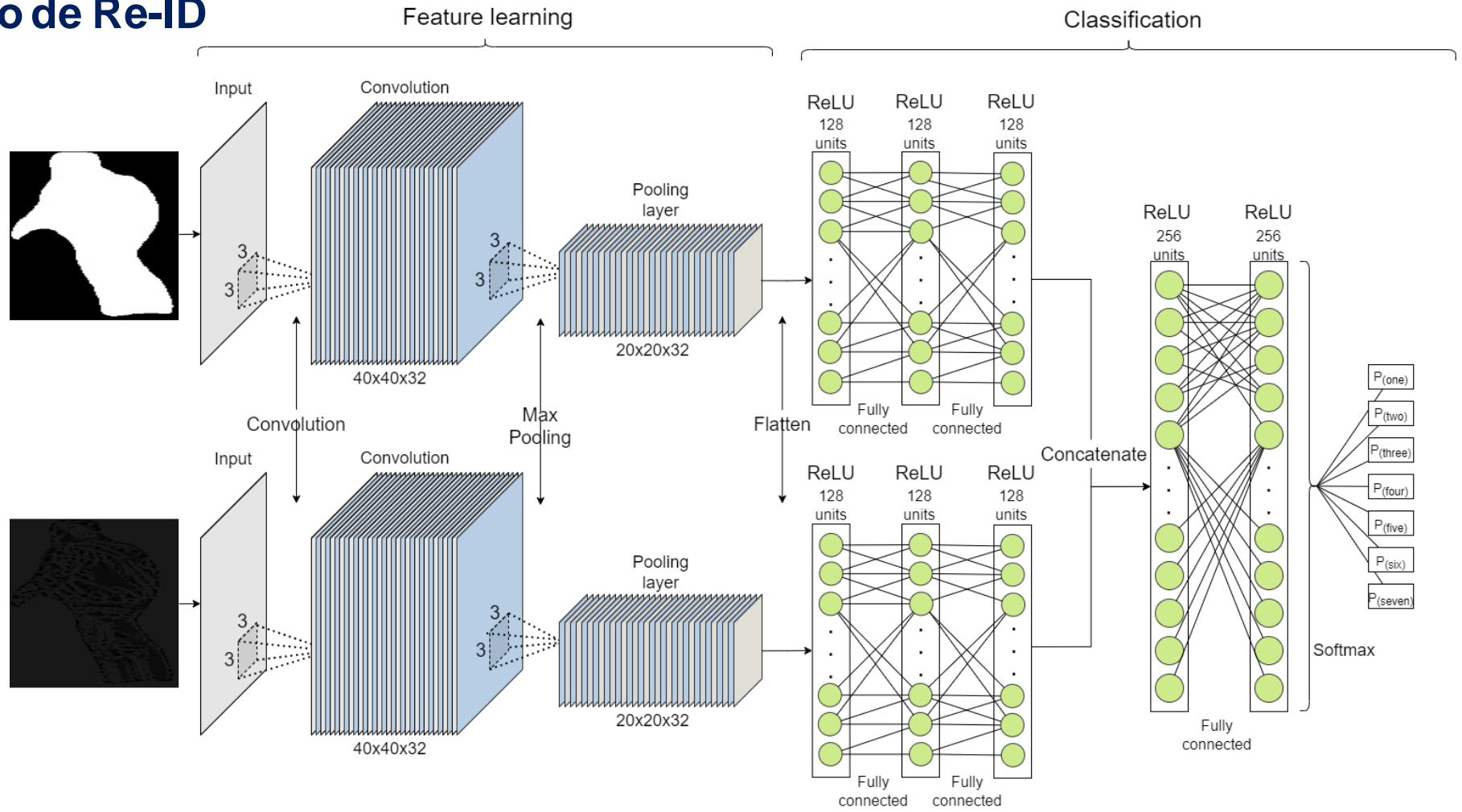


Máscaras binarias de personas detectadas.

Se reutilizó la máscara generada por el modelo Mask R-CNN como el descriptor de la silueta de la persona, la misma que se ha sido recortada utilizando las Regiones de Interés.

Diseño del modelo de Re-ID

Se propone un modelo multi-entrada con dos ramas implementado en Keras. Este recibe una máscara binaria y una imagen filtrada utilizando LBPU. La salida de este modelo es un vector de probabilidades.



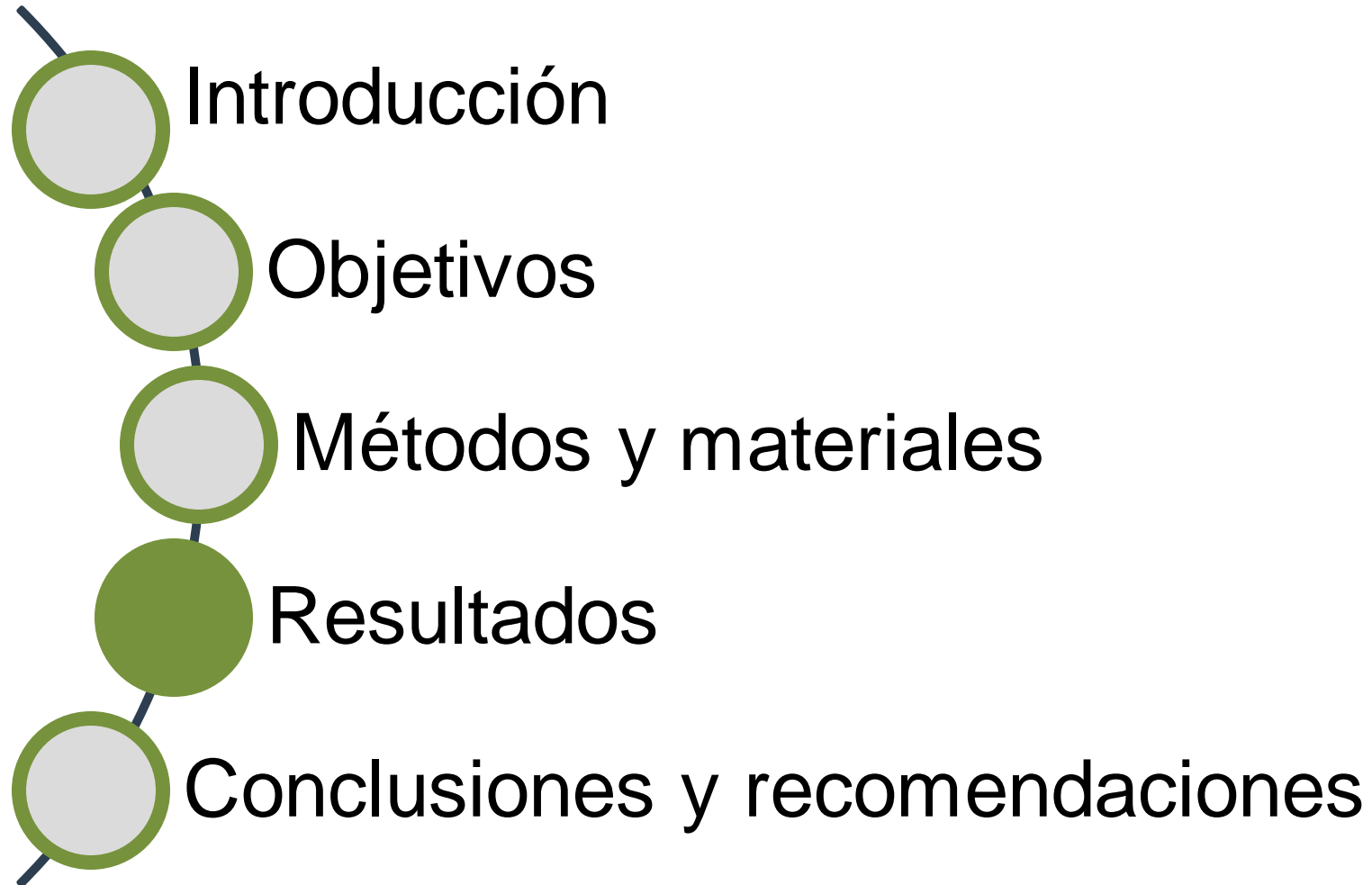
Arquitectura del modelo de Re-Identificación

Conjunto de datos



Conjunto de datos de muestras

- Dos escenarios con **iluminación natural**.
- Diferentes capturas en distintas perspectivas.
- Se utilizó la técnica de **aumentación de datos**.
- Se generaron **10 variaciones** de cada captura.
- A partir de **85 imágenes de entrenamiento** y **20 de prueba**, tenemos un total de **1862 imágenes de entrenamiento** y **481 de prueba**, por persona (12,295 imágenes de 7 identidades).



Entorno de pruebas

Características

MSI GL62M 7RDX

- Windows 10
- Nvidia GTX 1050 – 4GB GDDR5
- RAM – 16 GB
- Procesador Intel core i7-7700HQ

Herramientas

- Python 3.7.13
- CUDA 10.0
- CUDNN 7.6.3
- Tensorflow 1.15.2



Entrenamiento del modelo de Re-Identificación

$$exactitud = \frac{VP + VN}{VP + VN + FP + FN}$$

Modelo	Descriptores	Exactitud	Tiempo de entrenamiento
1	LBP	68.54%	59 seg
2	Máscara binaria	42.99%	50.8 seg
3	Máscara binaria + LBP	71.58%	2 min 20 seg

Tabla I. Resultados de los modelos durante el entrenamiento

Se entrenaron tres modelos de Re-ID con diferentes entradas y sus resultados de exactitud fueron capturados para su análisis.

Análisis e interpretación:
La combinación de dos características soft-biométricas logró los mejores resultados.

Capacidad de Re-Id

Modelo	Descriptores	Precisión	Exhaustividad	F1	Exactitud
1	LBPU	0.5095	0.4035	0.4241	40.35%
2	Máscara binaria	0.2895	0.3158	0.2745	31.58%
3	Máscara binaria + LBPU	0.6245	0.4561	0.4555	45.61%

Se planteó una prueba, en la que se juntaron 57 imágenes con varias imágenes de cada identidad, para determinar la capacidad de Re-ID del sistema.

Análisis e interpretación:

El modelo 3 logró los mejores resultados con respecto al resto, con una precisión del 45,61% y una puntuación F1 de 0,4555, lo que refleja una mejora de la precisión del 5,26% y 14,03%, y una mejora del puntaje F1 de 0,0314 y 0,1810 respecto a los modelos 1 y 2.

$$exactitud = \frac{VP + VN}{VP + VN + FP + FN}$$

$$precisión = \frac{VP}{VP + FP}$$

$$exhaustividad = \frac{VP}{VP + FN}$$

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

Rendimiento del sistema

Modelo	Descriptores	Tiempo (seg)			Exactitud
		Total	Pre-procesamiento	Identificación	
1	LBP	2.2061	0.0079	0.0107	33.13%
2	Máscara binaria	2.0559	0.0063	0.0095	30.31%
3	Máscara binaria + LBP	2.7787	0.0058	0.0101	38.57%

Tabla II. Resultados del rendimiento del sistema.

Se preparó un experimento en un entorno controlado, en el que se tomaron las grabaciones de dos cámaras disjuntas.

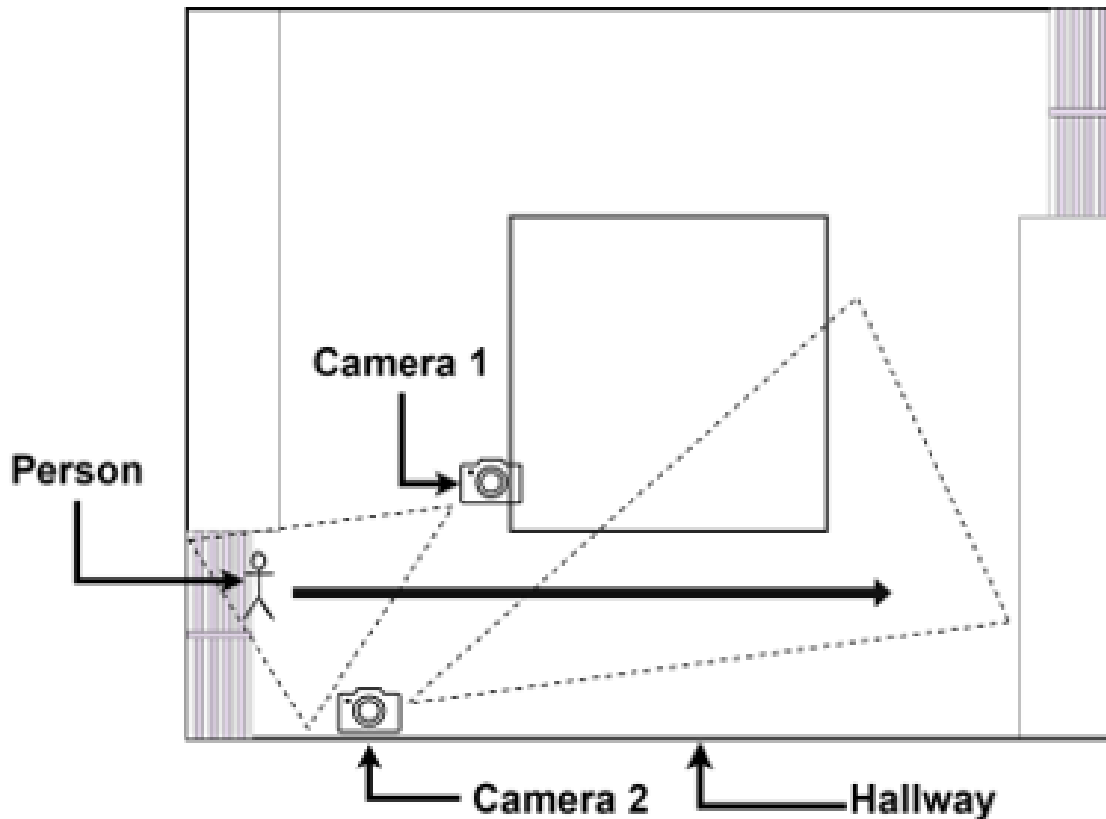
Análisis e interpretación:

Los resultados denotan que el modelo combinado tiene un tiempo de ejecución total mayor al resto, debido a la cantidad de descriptores a procesarse. Sin embargo, los resultados de exactitud son mejores.

Resultados

Simulación

Distribución de las cámaras



Dos cámaras disjuntas de **12 y 13 megapíxeles** situadas en un pasillo con iluminación natural.

Las **dos cámaras** se colocaron una frente a la otra

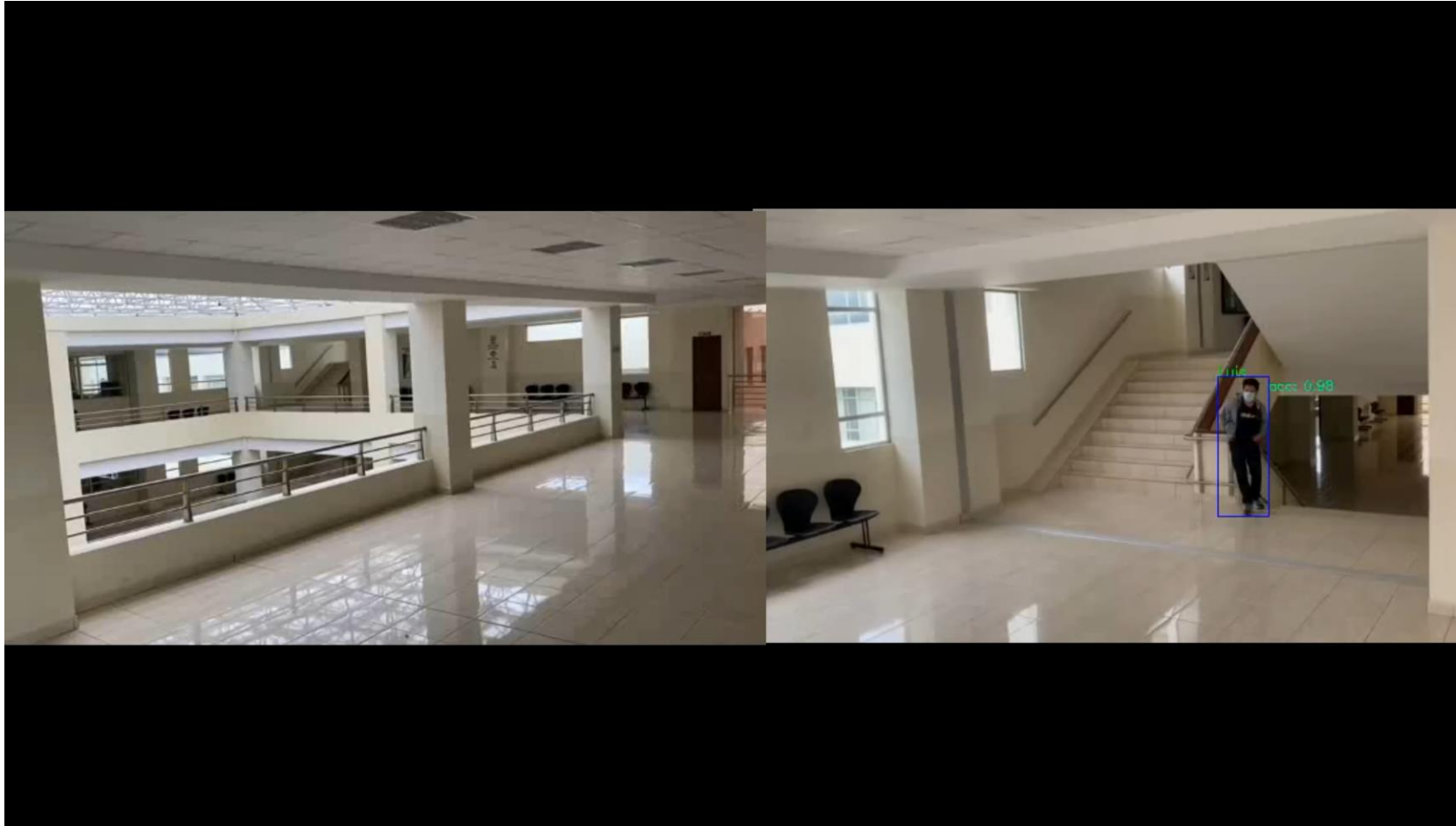
La **cámara 1** apunta hacia las escaleras del pasillo (empezando por punto de la trayectoria de las personas filmadas) y la **cámara 2** apunta al lado opuesto.

La duración de los vídeos fue de **21 segundos**, leídos con una latencia de **25 fotogramas** por segundo y una resolución de **3840x2160 píxeles**.

Resultados

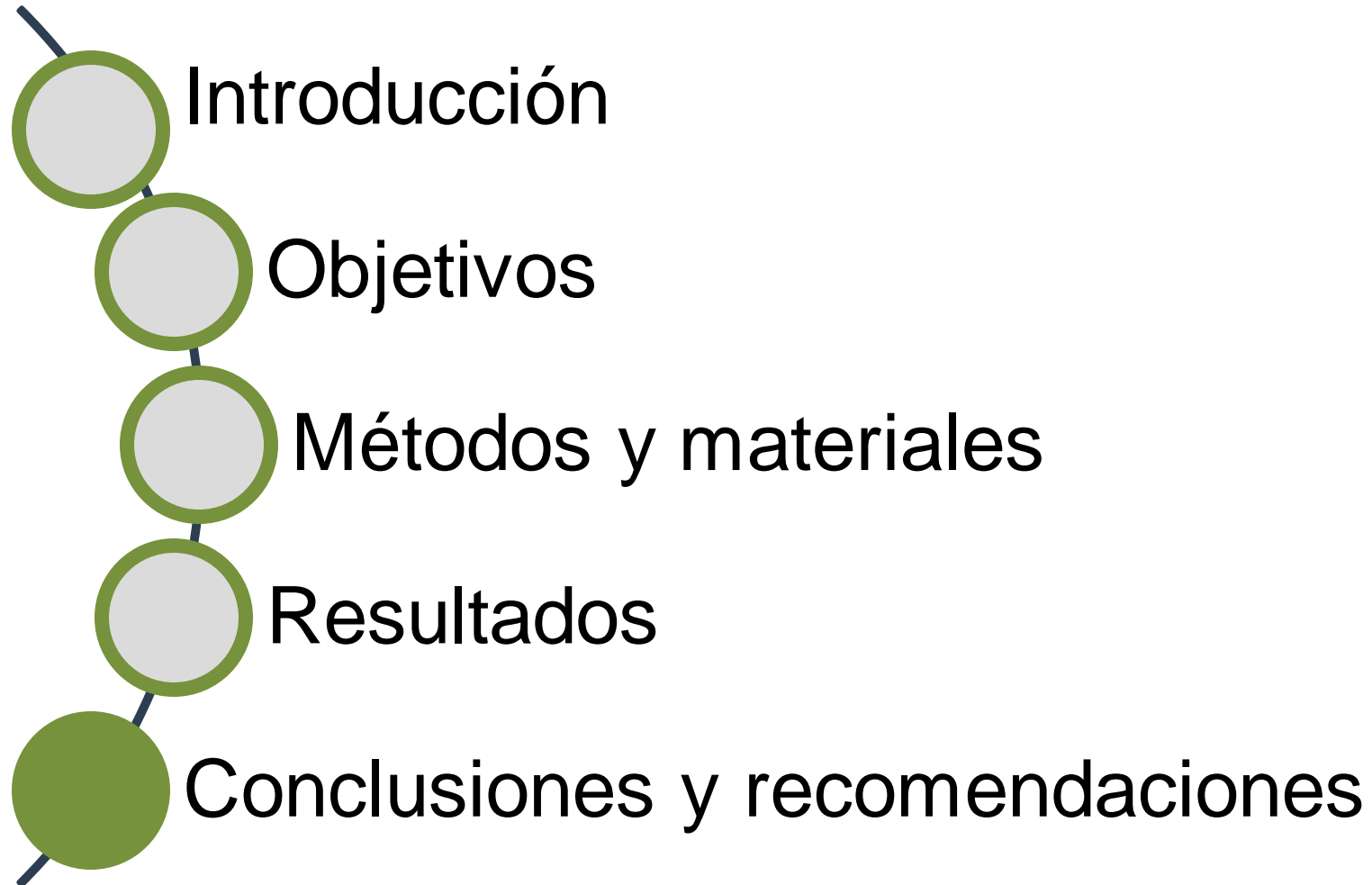
Simulación

Cámara 2



Cámara 1

URL: <https://drive.google.com/drive/folders/1JbavJbY6SHP5LrB-ZwK1ZanP3NsREv4A?usp=sharing>



Conclusiones y recomendaciones

- El uso de características soft-biométricas en la literatura ha demostrado ser una **solución viable** en **ambientes complejos** de videovigilancia, donde la extracción de características biométricas fuertes no es posible.
- La combinación propuesta de características soft-biométricas para la Re-Identificación de personas demostró ser una solución más **eficiente** frente a su uso individual.
- La aplicación de la aumentación de datos es una técnica muy útil para enfrentar el sesgo que suele existir en los conjuntos de datos, **aumentando** la variación de los mismos y mejorando la capacidad de **generalización** de los modelos.
- Los resultados indicaron que el **rendimiento** del modelo de Re-Identificación propuesto disminuye en un ambiente de **producción**, debido a que la extracción de las características de las imágenes tomadas para la validación de los modelos es complicada, esto se puede mejorar **aumentando** el **número de características** que se tomaron en cuenta.
- Se espera **mejorar** la solución propuesta utilizando algoritmos **más eficientes** y la **combinación** de más características soft-biométricas para mejorar la **robustez** de la **re-identificación** de las personas.

¡Gracias!