

# Diseño, implementación y evaluación de una estrategia de detección de objetos abandonados en aplicaciones de videovigilancia

Jesús Mudarra Luján

Departamento de Electrónica

28 de julio de 2021

# Índice

## 1. Introducción

- Objetivos

## 2. Estudio teórico

- Técnicas utilizadas
- Algoritmos de detección y seguimiento utilizados

## 3. Desarrollo

- Timeline desarrollo del proyecto
- Desarrollo algoritmo detección de objetos abandonados

## 4. Resultados

- Métricas de calidad
- Bases de datos
- Resultados entrenamientos
- Resultados detección, seguimiento y detección de objetos abandonados

## 5. Conclusiones y líneas futuras

## 6. Proyectos derivados

## Introducción

## Sistemas de videovigilancia automatizados

Han despertado un gran interés en los últimos años en la monitorización de lugares públicos y privados.

- En los últimos tiempos la detección de objetos se ha investigado para detectar eventos de gran interés como **objetos abandonados** y **vehículos estacionados** ilegalmente.
  - Los desarrollos de técnicas de detección de objetos abandonados se encuentran constantemente enfrentados contra diferentes **desafíos** durante su implementación.
  - Los sistemas de detección de objetos abandonados actuales se centran principalmente en dos etapas principales: **detección estacionaria** y **clasificación de objetos**.

## Introducción

Los objetos abandonados se pueden determinar mediante dos reglas: el objeto aspirante se encuentra estático o desatendido.

- El primer enfoque corresponde a una **regla espacial**, en la que un objeto se considera desatendido si el propietario del objeto se encuentra apartado del objeto.

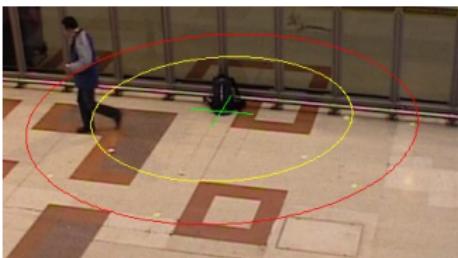


Figura 1: Persona sobre pasando la zona de alarma (marcado en amarillo)

# Introducción

- El segundo enfoque define una **regla temporal** en la que un objeto se considera estacionario si se encuentra inmóvil durante un cierto período de tiempo, dependiendo de la aplicación, siendo típicamente 30 o 60 segundos.

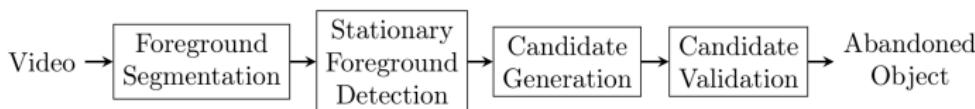


Figura 2: Marco de referencia en detección de objetos abandonados

# Introducción

En los últimos años se ha mostrado un **gran progreso** en la detección de personas debido a la aparición de métodos de aprendizaje profundo o Deep Learning.

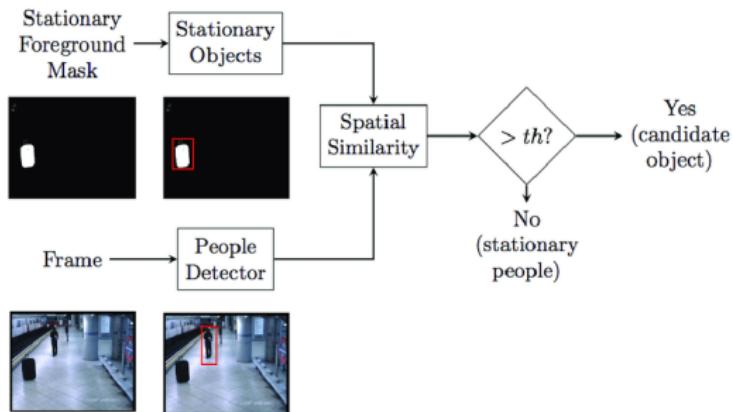


Figura 3: Diagrama de bloques del módulo de generación de candidatos

## Introducción

## Objetivos

Una meta

El objetivo que se persigue es el desarrollo de una estrategia de detección de objetos abandonados mediante el uso de CNN en aplicaciones de videovigilancia en espacios como aeropuertos, estaciones de metro, edificios o cualquier tipo de infraestructura que disponga de una o varias cámaras de videovigilancia.

Para alcanzar dicho objetivo, se han planteado los siguientes puntos:

- Revisión del Estado de Arte.
  - Evaluación de algoritmos de detección de objetos más relevantes.
  - Evaluación de algoritmos de seguimiento o de objetos más relevantes.
  - Implementación y evaluación de un algoritmo de detección de objetos abandonados.

# Estudio teórico

## Técnicas utilizadas - Segmentación de objetos en primer plano

- Obtener una imagen que se encuentre en el **fondo** sin ningún objeto en movimiento.
- El modelo de representación debe ser **robusto** ante posibles alteraciones en la iluminación u objetos que han sido introducidos y/o retirados.
- Dos principales **problemas**: detección de cambios y la detección de movimientos.

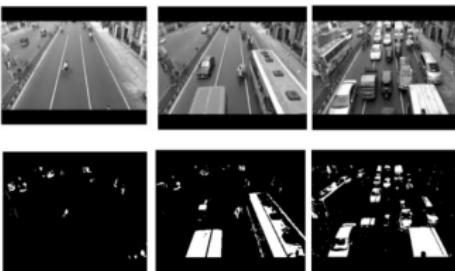


Figura 4: Ejemplo sustracción del fondo

# Estudio teórico

## Técnicas utilizadas - Detección de objetos estacionarios

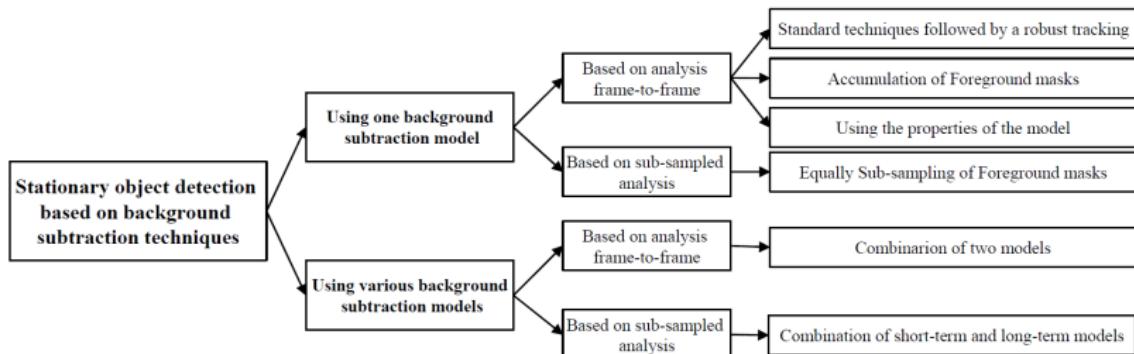


Figura 5: Clasificación de sustracción del fondo basados en métodos de detección de objetos estacionarios

# Estudio teórico

## Técnicas utilizadas - Reconocimiento del comportamiento

- La detección de comportamientos **anómalos** en la videovigilancia es un reto en la visión por computadora y últimamente ha experimentado importantes avances.
- Las etapas de procesamiento no permiten comprender el tipo de **acción** que realiza el objeto en movimiento ni determinan si su comportamiento es normal o no.
- Densidad de escenas e interacción de objetos en movimiento.
- Métodos de clasificación: supervisado, semi-supervisado y no supervisado.



Figura 6: Comportamientos anómalos en una persona



Figura 7: Comportamientos anómalos en múltiples personas

# Estudio teórico

## Técnicas utilizadas - Detección de personas y objetos

- Los algoritmos de detección de personas y objetos se clasifican según el enfoque utilizado para generar o extraer los objetos iniciales y en base al modelo de persona.
- La **segmentación** genera directamente los objetos candidatos a ser persona u objetos y fácilmente se rechaza las áreas irrelevantes de la imagen, es decir, sin objetos de interés.
- La **búsqueda exhaustiva** consiste en escanear la imagen completa buscando similitudes con el modelo de persona elegido en múltiples escalas y ubicaciones.

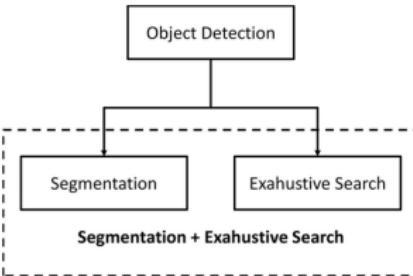


Figura 8: Clasificación en la detección de personas

# Estudio teórico

## Algoritmo de detección de objetos YOLOv4

En vista a las métricas se puede concluir que, de todos los detectores de objetos en tiempo real actuales, YOLOv4 es el mejor en términos de **velocidad y precisión**.

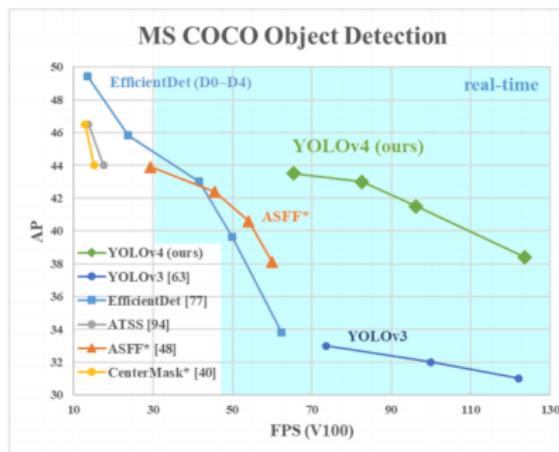


Figura 9: YOLOv4 frente a otros detectores

# Estudio teórico

## Algoritmo de detección de objetos YOLOv4

- La detección primero divide la imagen en una cuadrícula de SxS.
- En cada una de las celdas se **predice** N posibles cuadros delimitadores y se calcula el nivel de **certidumbre**.
- Se eliminan los cuadros que estén por debajo de un límite.
- A los cuadros restantes se les aplica un paso de “non-max suppression”.

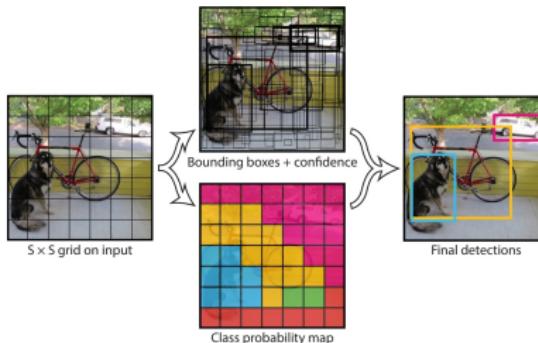


Figura 10: Funcionamiento del algoritmo YOLO

# Estudio teórico

## Algoritmo de seguimiento de objetos Deep SORT

- Utiliza los pesos pre-entrenados o entrenados de YOLOv4 para extraer los cuadros delimitadores de las detecciones.
- Los cuadros delimitadores sirven de entrada al algoritmo SORT para realizar el seguimiento.
- Mediante una DNN entrenada el algoritmo asocia los tracks activos y los que se han perdido.

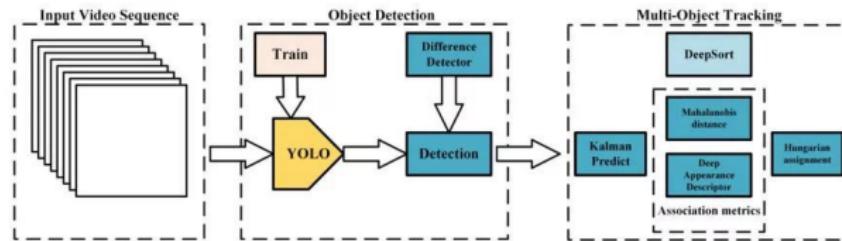


Figura 11: Esquema etapas detección y seguimiento con YOLOv4 y Deep SORT

## Desarrollo

## Timeline desarrollo del proyecto

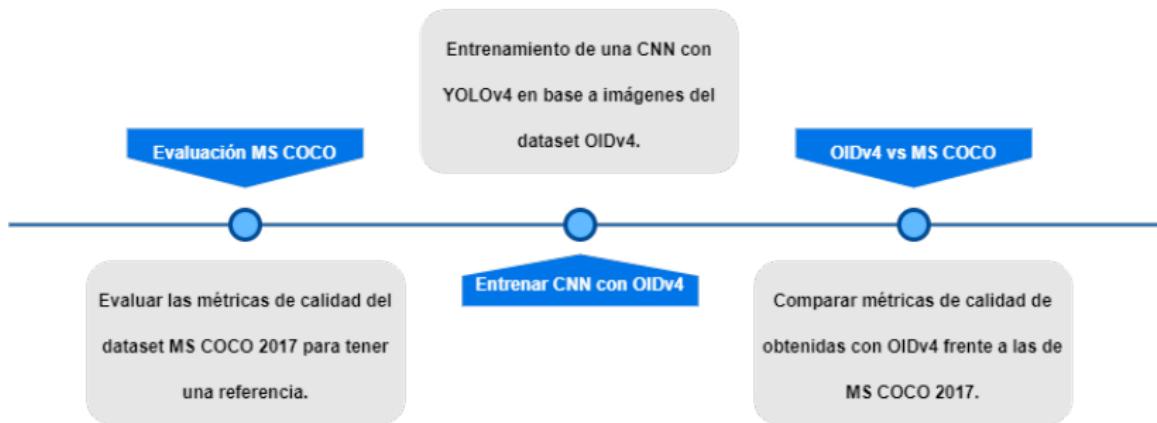


Figura 12: Timeline de la evaluación de los datasets MS COCO y OIDv4

# Desarrollo

## Timeline desarrollo del proyecto

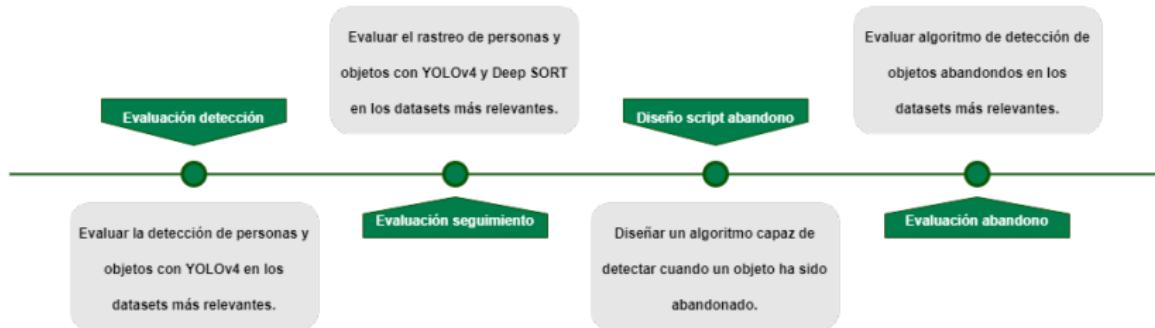


Figura 13: Timeline de la evaluación de los distintos algoritmos

# Desarrollo

## Desarrollo algoritmo detección de objetos abandonados

- Asociación **persona-objeto** al principio de la secuencia de vídeo.
- Determinar si un objeto no tiene **propietario**.
- Medir **distancia** en fotogramas posteriores.
- Aviso si el objeto se encuentra en **zona de alarma**.
- Aviso si el objeto ha sido **abandonado**.



(a)



(b)



(c)

Figura 14: Imágenes extraídas del dataset AVSSAB2007

# Desarrollo

## Diagrama algoritmo detección de objetos abandonados

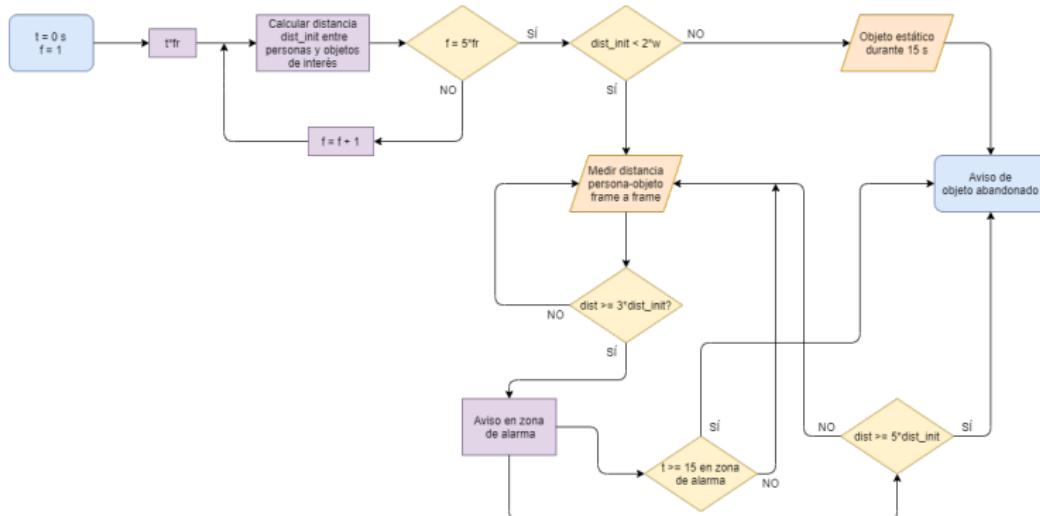


Figura 15: Diagrama algoritmo detección de objetos abandonados

## Resultados

## Métricas de calidad

- La IoU es una medida basada en el índice Jaccard que evalúa la superposición entre dos cuadros delimitadores.

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{blue area}}{\text{red + green + blue areas}}$$


Figura 16: Área de superposición IoU entre los cuadros delimitadores

- True Positive (**TP**), True Negative (**TN**), False Positive (**FP**) y False Negative (**FN**).
  - La **precisión** es la capacidad de un modelo para identificar solo los objetos relevantes.

## Resultados

## Métricas de calidad

- El **Recall** es la capacidad de un modelo para encontrar todos los casos relevantes (todos los cuadros delimitadores de ground truth).
  - El **F-Score** se trata de una medida estadística de precisión muy utilizada en las pruebas test de algoritmos.
  - La precisión media (**AP**) es el valor medio de 11 puntos en la curva **P-R** para cada posible umbral (cada probabilidad de detección) para la misma clase (precisión-Recall). Por otro lado, el **mAP** es la media de los **AP** de todas las categorías de objetos.

# Resultados

## Bases de datos - PETS2007 dataset

- Resolución 768x576 a 25 fps.
- Cámaras 1 y 3: Canon MV-1 1xCCD w/progressive scan.
- Cámaras 2 y 4: Sony DCR-PC1000E 3xCMOS.
- Secuencias que contienen tres tipos de escenarios con una complejidad ascendente: personas merodeando, robo de equipaje y equipaje desatendido/abandonado.



(a)



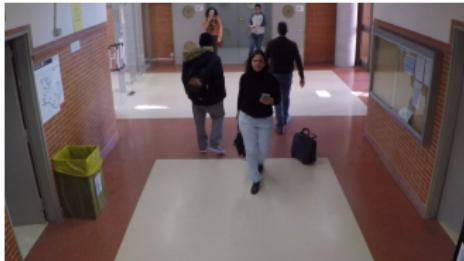
(b)

Figura 17: Imágenes extraídas del dataset PETS2007

# Resultados

## Bases de datos - GBA2018 dataset

- Resolución 1920x1080 a 60 fps.
- Cámara GoPro HERO4.
- Contiene 8 secuencias en 2 escenarios donde tienen lugar eventos como abandono de maletas, mochilas y bolsas de mano.



(a)



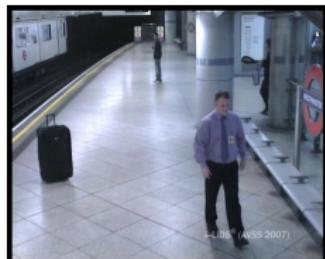
(b)

Figura 18: Imágenes extraídas del dataset GBA2018

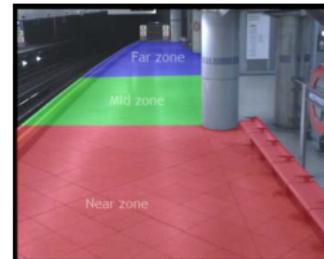
# Resultados

## Bases de datos - AVSSAB2007 dataset

- Resolución 720x576 a 25 fps.
- Modelo de la cámara no especificado.
- Subconjunto de datos del dataset AVSS 2007 creado para el i-LIDS bag and vehicle detection challenge.
- La región de interés está dividida en tres zonas: cercana, media y lejana.



(a)



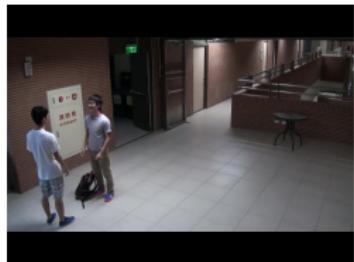
(b)

Figura 19: Imágenes extraídas del dataset AVSSAB2007

# Resultados

## Bases de datos - ABODA dataset

- Resoluciones de 720x480 a 30 fps y 640x480 a 30 fps.
- Modelo de la cámara no especificado.
- 11 secuencias etiquetadas con varios escenarios de aplicaciones reales.
- Escenas de aglomeraciones de personas que se encuentran en aeropuertos o estaciones de tren, iluminación variable de día y grabaciones nocturnas.



(a)



(b)

Figura 20: Imágenes extraídas del dataset ABODA

# Resultados

## Resultados de los entrenamientos vs MS COCO

Dataset	TP	FP	FN
MS COCO	22.730	10.889	13.027
OIDv4 test 1	430	527	273
OIDv4 test 2	6.600	7.366	5.226

Tabla 1: Comparativa métricas de calidad entre los test en OIDv4 y MS COCO [1]

Dataset	Precision (%)	Recall (%)	F-score (%)	average IoU (%)	mAP @ 0.5 (%)
MS COCO	67,61	63,57	65,53	56,04	64,16
OIDv4 test 1	44,93	61,17	51,81	34,13	66,25
OIDv4 test 2	47,26	55,81	51,18	37,83	60,97

Tabla 2: Comparativa métricas de calidad entre los dos test en OIDv4 y MS COCO [2]

# Resultados

## Resultados detección de personas y objetos

- Mejora en las predicciones respecto a la versión previa YOLOv3.
- Confusión en la detección de las clases `handbag` y `backpack`.
- Gran influencia de los cambios de iluminación, la distancia, el ángulo y la resolución de la cámara.



Figura 21: Detección de personas y objetos en el dataset AVSSAB2007

# Resultados

## Resultados seguimiento de personas y objetos

- Totalmente dependiente de la etapa detección.
- Elevados cambios de ID en grandes aglomeraciones de personas.
- Deficiencia en el seguimiento a través de occlusiones.



Figura 22: Seguimiento de personas y objetos en el dataset PETS2007

# Resultados

## Resultados detección objetos abandonados



Figura 23: Bolsa de mano abandonada en el pasillo de cafetería de la EPS (1)

# Resultados

## Resultados detección objetos abandonados

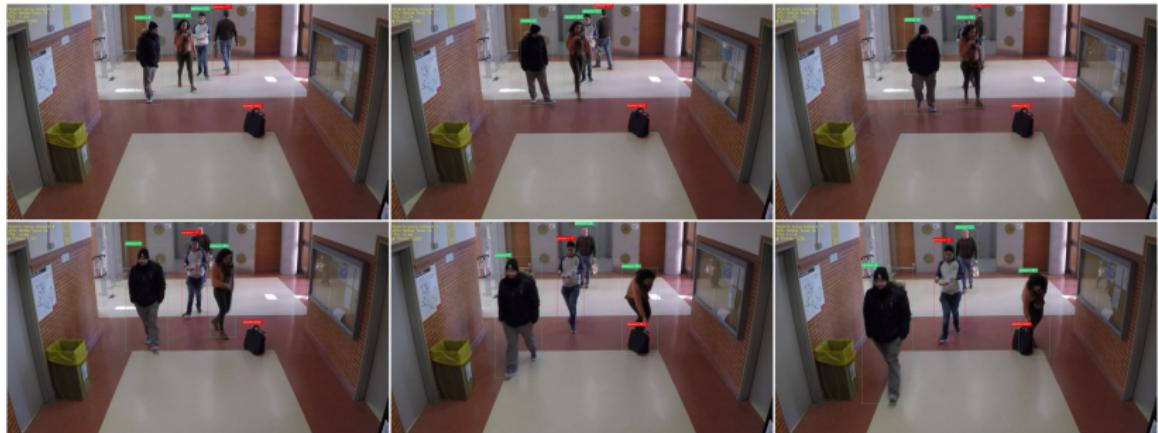


Figura 24: Bolsa de mano abandonada en el pasillo de cafetería de la EPS (2)

## Conclusiones y líneas futuras

## Conclusiones

- YOLOv4 presenta **mejores** métricas en el dataset MS COCO frente a otros detectores.
  - Resultados desfavorables en los dos entrenamientos realizados con el dataset Open Images Dataset v4.
  - Deep SORT trabajaba **excelentemente** junto a YOLOv4.
  - El algoritmo de detección de objetos abandonados trabaja correctamente.
  - Problemas en el seguimiento de personas y objetos debido a **cambios** de ID y **occlusiones**.

# Conclusiones y líneas futuras

## Líneas futuras

- Implementar como **detector Scaled-YOLOv4, YOLOv5, YOLOR o YOLOX.**
- Implementación en tarjetas **NVIDIA Jetson.**
- Diferenciar cuando un objeto ha sido **abandonado o robado.**
- Entrenamiento de la red neuronal con otros **datasets.**

# Proyectos derivados



Figura 25: Mobile World Congress 2021



Figura 26: Exposición del Ninot 2021

- Posibilidad de crear un startup de **inteligencia artificial** enfocada a la seguridad por **videovigilancia**.
- **Colaboraciones** con empresas de gran impacto social.
- Oportunidad de estar presente en la mayor **feria tecnológica** del mundo.

**¡Muchas gracias por su atención!**

**¿Tienen alguna pregunta?**