

Generación sintética de imágenes para segmentación y detección de objetos

 Miguel Sánchez Novo

 miguel.sanchez@iic.uam.es

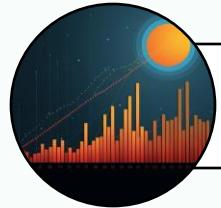
Científico de datos en el Instituto de Ingeniería del
Conocimiento (IIC)
<https://www.iic.uam.es/iic/>

ÍNDICE

1. Utilidad de generación de datos sintéticos
2. Métodos de generación de datos sintéticos
3. Hitos temporales en imágenes
 1. *Modelos de difusión*
 2. *Modelos de generación de vídeo*
4. Tareas en 'computer vision'
5. Hitos temporales en segmentación y detección de objetos
6. Etiquetado de datos
7. Técnica basada en composición de imágenes
8. Conclusiones



1) ¿Son útiles los datos sintéticos?



Aumentar la cantidad de datos



Privacidad y seguridad



Robustez y generalización



Balance de clases



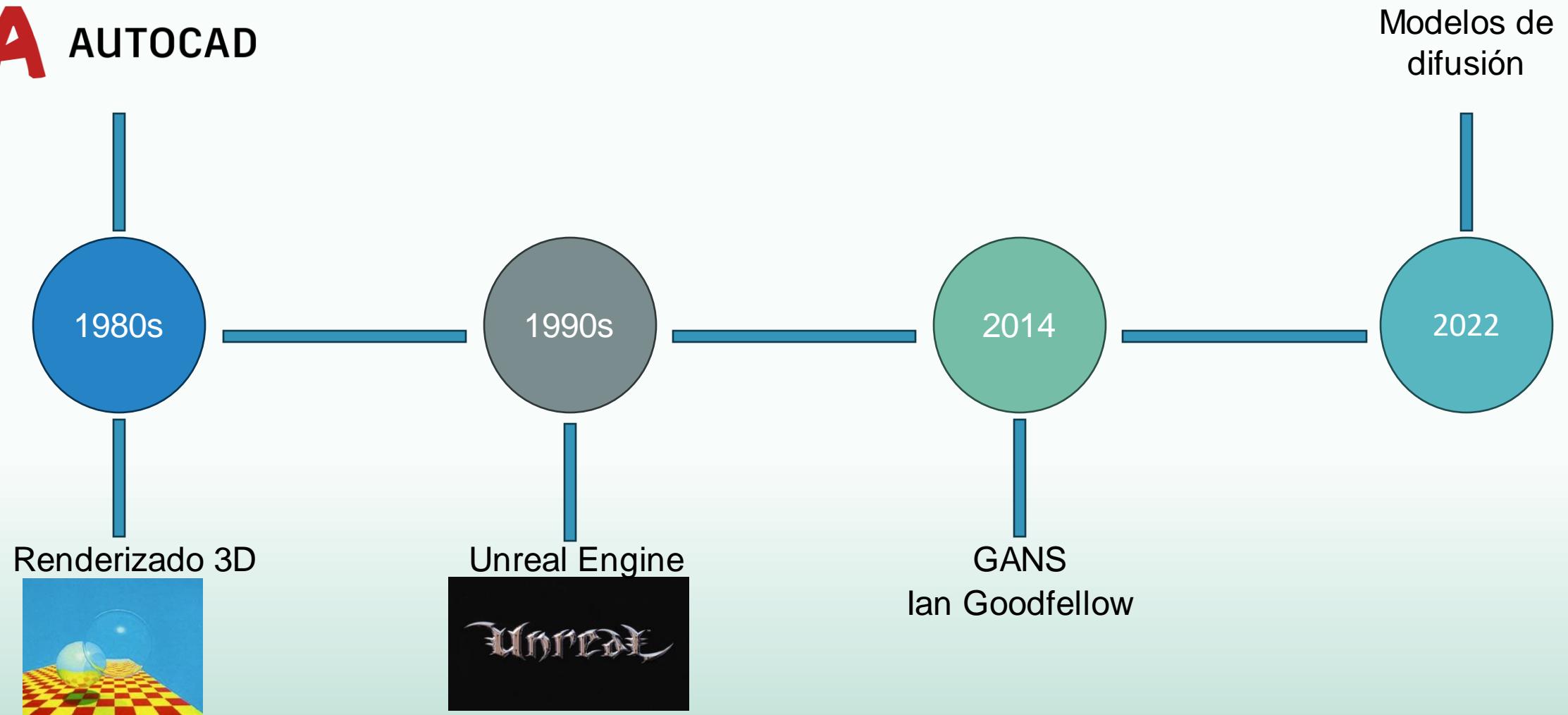
Ahorro de costes

2) Métodos de generación sintética de datos

	Proceso estocástico	Basado en reglas	Modelos generativos
Computación	BAJA 	MEDIA 	ALTA 
Labor humana	BAJA 	ALTA 	MEDIA 
Complejidad del sistema	BAJA 	MEDIA 	ALTA 
Información contenida	BAJA 	MEDIA 	ALTA 

3) Línea temporal generación sintética de imágenes

A AUTOCAD



3.1) Modelos de difusión

The banner features a portrait of a young man with dark hair and brown eyes, looking slightly to the right. He has a small smile and is wearing a dark jacket over a white shirt. To his left, large blue text reads "MODELOS DE DIFUSIÓN" and "¿CÓMO FUNCIONAN?". Above this, smaller text says "SPAIN AI". Below the main title, it says "ÁLVARO BARBERO JIMÉNEZ (CHIEF DATA SCIENTIST IIC)". At the bottom, it says "DALL-E 2, Stable Diffusion... Profundizaremos en el funcionamiento de estos". A circular logo for "IIEVES" is partially visible.



DALL-E 2



Midjourney



Stable Diffusion





3.2) Generación de vídeo

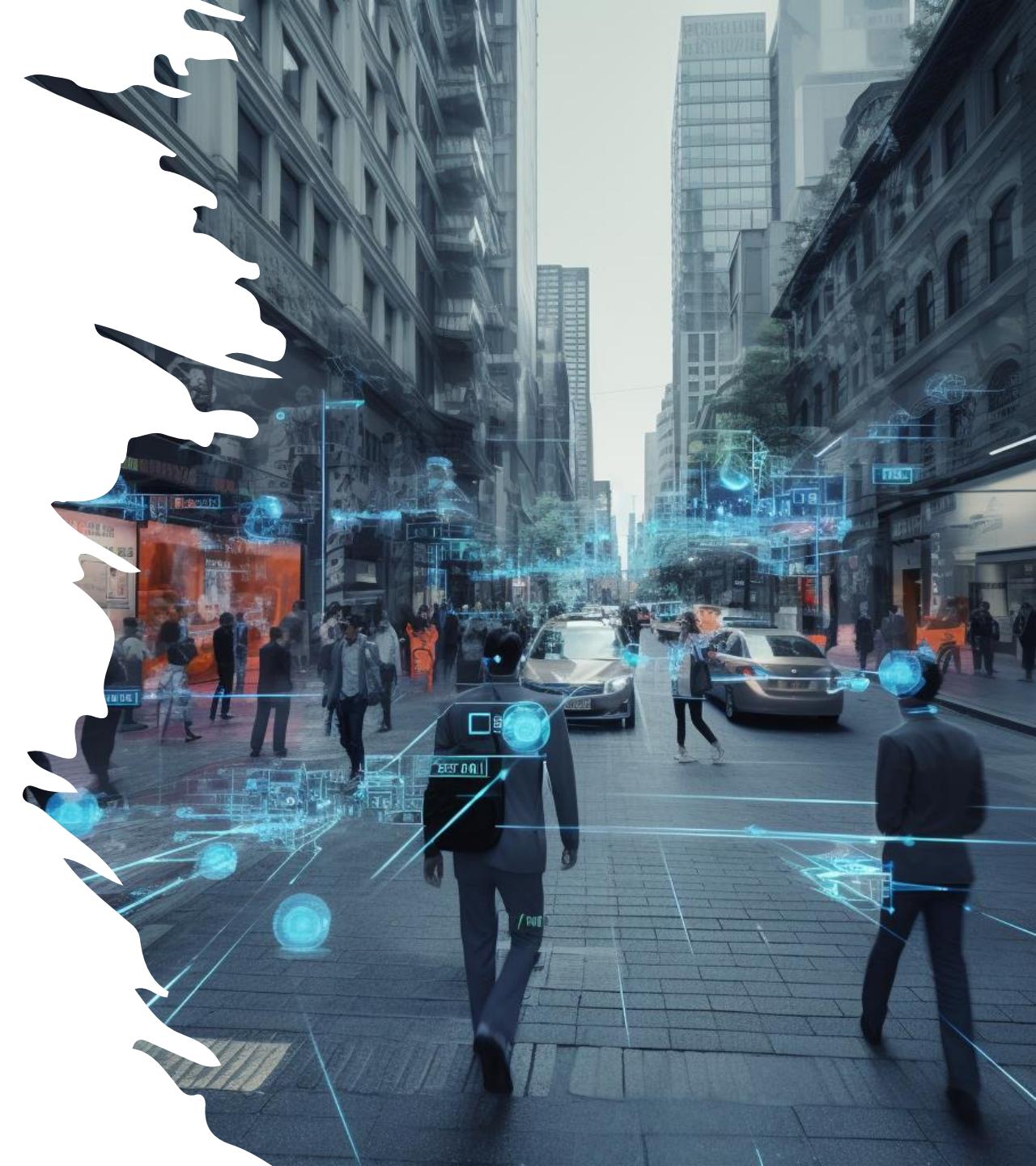
GAIA-1



GAIA-1: A
Generative
World Model
for
Autonomous
Driving

4) Tareas en “Computer Vision”

- Clasificación
- Detección de objetos
- Segmentación
 - Segmentación semántica
 - Segmentación de instancias
 - Segmentación panóptica



5) Línea temporal modelos segmentación y detección

Temin ator (1984)	Fast R- CNN YOLOv1 (2015)	Mask R- CNN (2017)	Cascade Mask R- CNN (2019)	SegFor mer (2021)	SAM YOLO- NAS (2023)
R-CNN (2014)	Faster R-CNN YOLOv2 (2016)	DeepLa b v3 YOLOv3 (2018)	ViT YOLOv4 YOLOv5 (2020)	SWIN v2 YOLOv7 (2022)	



Aplicaciones de segmentación y detección de objetos



Vehículos Autónomos

Detección de Defectos

Vigilancia y Seguridad

Medicina y Salud

Agricultura de Precisión

Mapeo y Cartografía

Robótica



Herramientas que ayudan/facilitan el etiquetado.

Herramientas que guían el etiquetado con modelos pre-entrenados.

Modelos de Zero-Shot

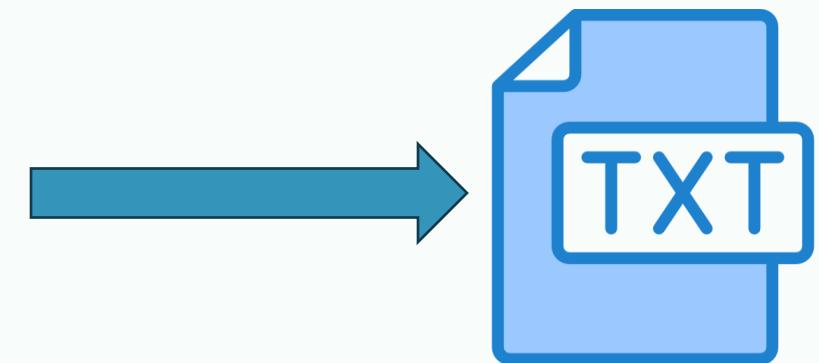
Técnica basada en composición de imágenes

	165	187	209	58	7	
14	125	233	201	98	159	
253	144	120	251	41	147	204
67	100	32	241	23	165	30
209	118	124	27	59	201	79
210	236	105	169	19	218	156
35	178	199	197	4	14	218
115	104	34	111	19	196	
32	69	231	203	74		

Imagen/es de fondo

Imagen/es objetivo

Imagen/es de ruido (optativo)



Ventajas del uso de esta técnica

Entrenar
modelos sin
imágenes

Herramienta de
auto etiquetado

Aumentar el
conjunto de
datos

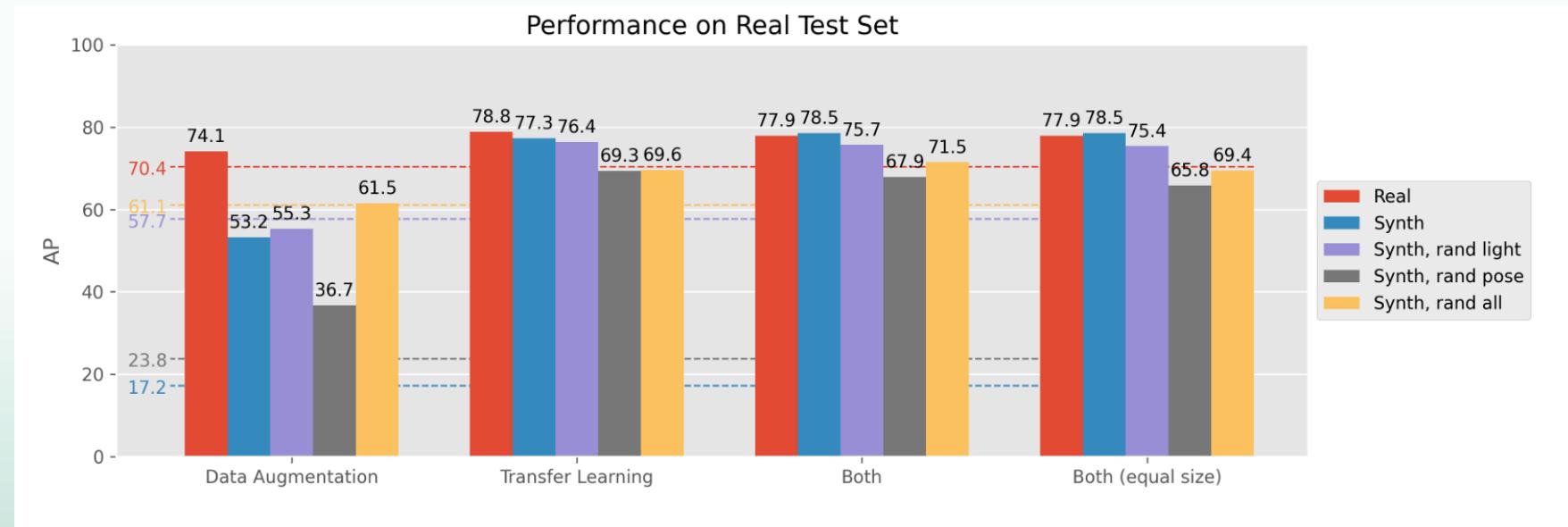
Mejorar la
generalización
de los modelos

Elegir el formato
de etiqueta para
nuestro modelo

¿Funciona?

Analysis of Training Object Detection Models with Synthetic Data

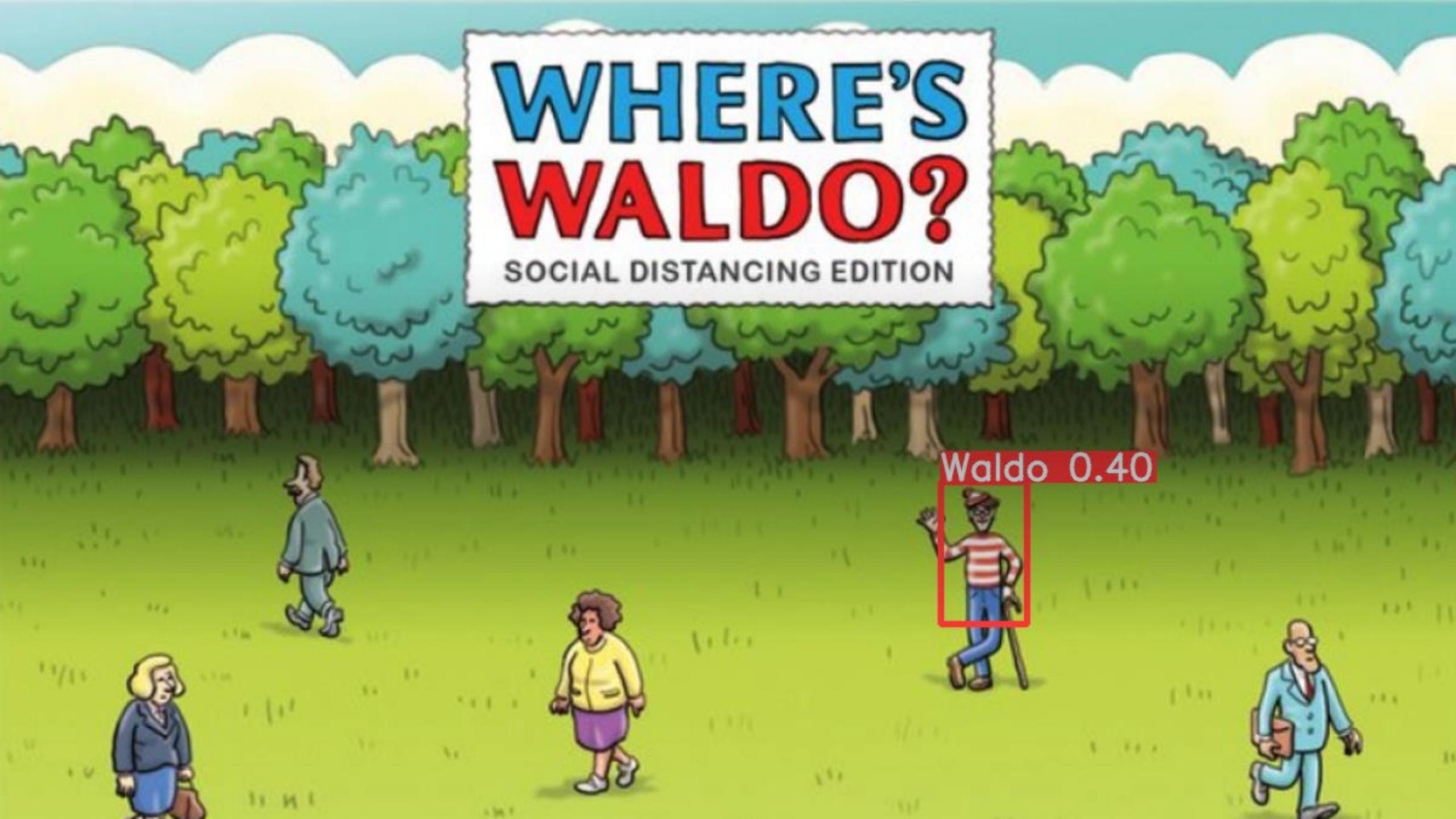
“DIMO dataset”



WHERE'S WALDO?

SOCIAL DISTANCING EDITION

Waldo 0.40



Composición de imágenes

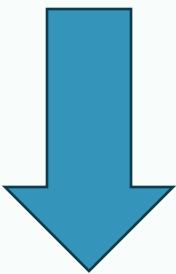
Background



Objetivo



Ruido



Entrenamiento del modelo





¿Jugamos?



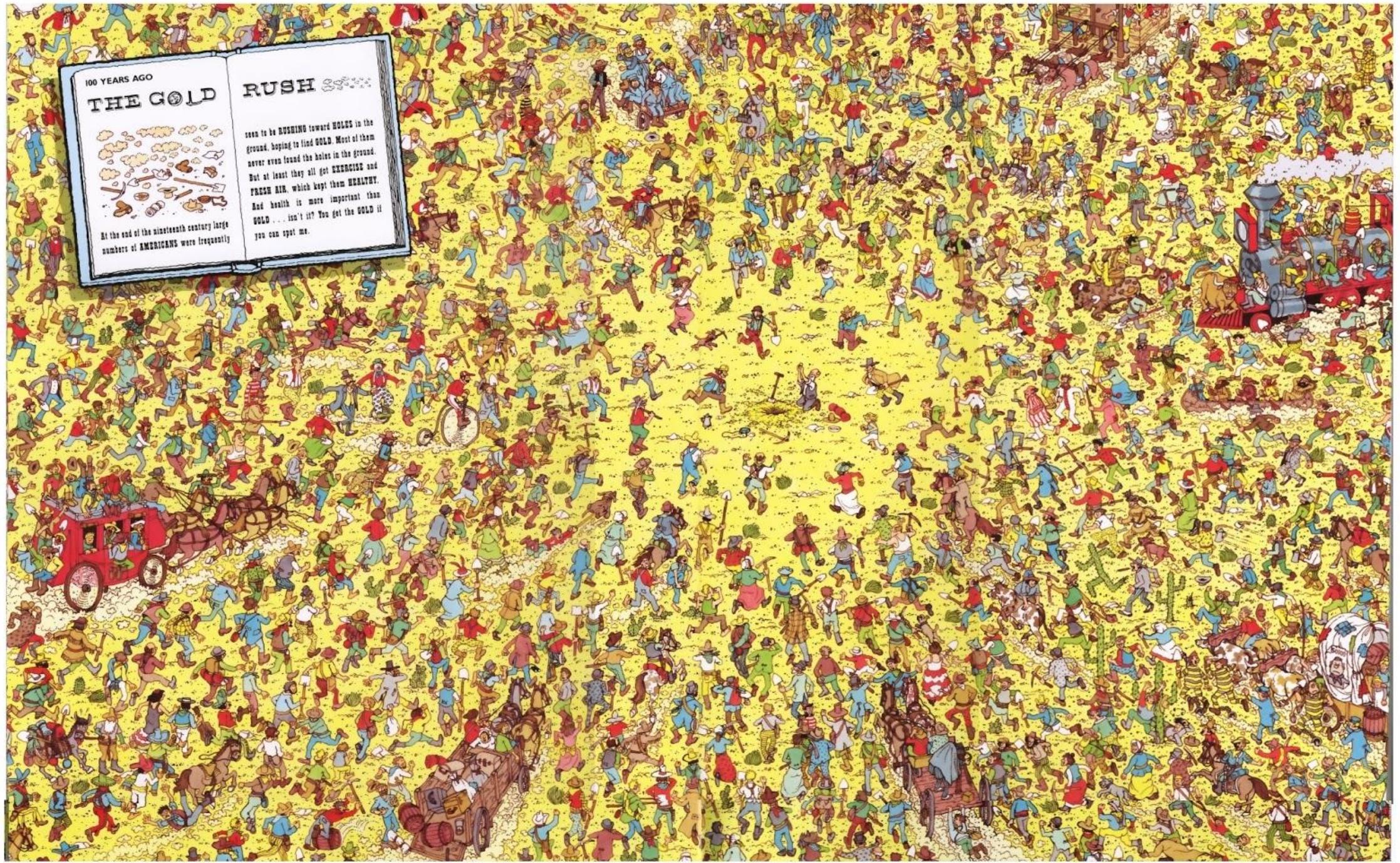


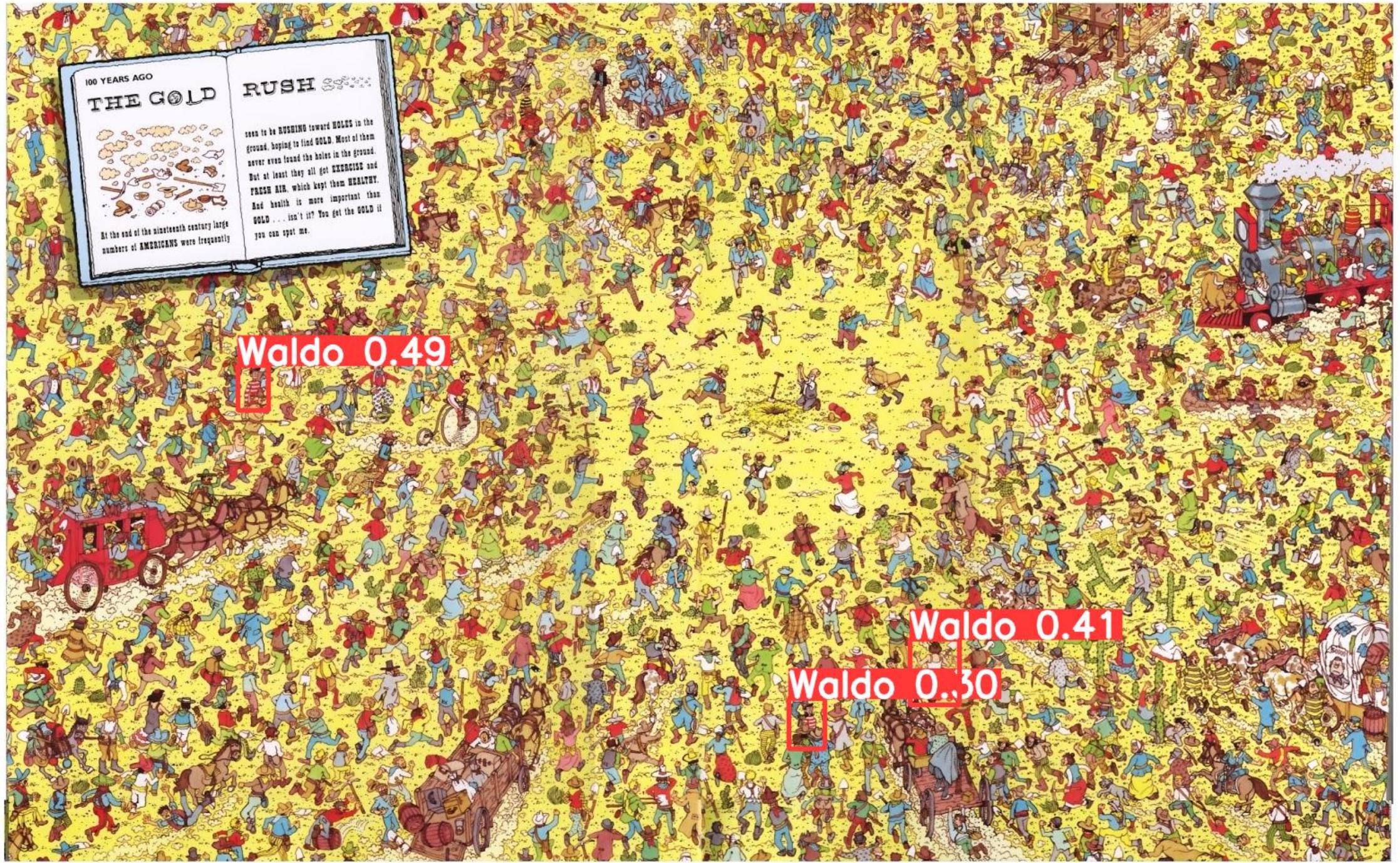
Wal

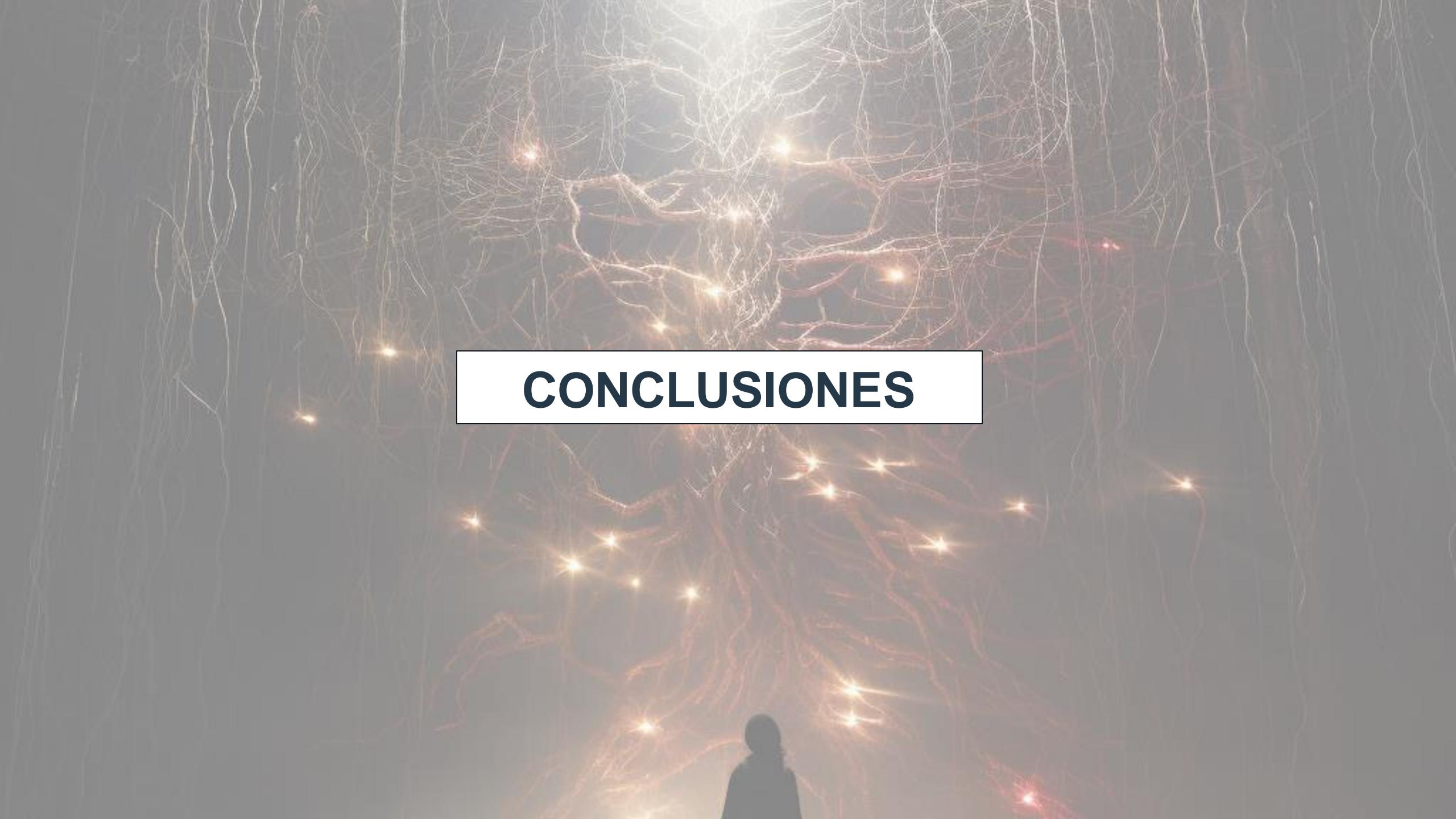




Waldo 0.64







CONCLUSIONES

Generación sintética de imágenes para segmentación y detección de objetos

¡Muchas gracias!



[miguel.sánchez@iic.uam.es](mailto:miguel.sanchez@iic.uam.es)



Miguel Sánchez Novo
Científico de datos en el Instituto de Ingeniería del
Conocimiento (IIC)
Profesor asociado en la Universidad CEU San Pablo