



MACHINE LEARNING CON NVIDIA JETSON Y MATLAB

Introducción

Diego Banda R.
Universidad Católica San Pablo
diego.banda@ucsp.edu.pe





AGENDA

- Colaboración entre NVIDIA y MATLAB.
- Desafíos del mundo real.
- ¿Qué es la Jetson Nano?
- Flujo de trabajo: MATLAB → NVIDIA Jetson.
- Ejemplo: Detección y Conteo de Vehículos.



Colaboración entre MathWorks y NVIDIA

Integrar tarjetas gráficas de NVIDIA con MATLAB

Cómputo de Propósito General

Soporte de tarjetas NVIDIA en más de 1000 funciones:

- MATLAB
- Deep Learning Toolbox
- Image Processing Toolbox
- Signal Processing Toolbox
- ¡Y muchas más!



On-Premise y Cloud

MATLAB y NVIDIA GPU Cloud

Acceso instantáneo a GPU locales (On-Premise) y en la nube con MATLAB y NGC (NVIDIA GPU Cloud).



Sistemas Embebidos

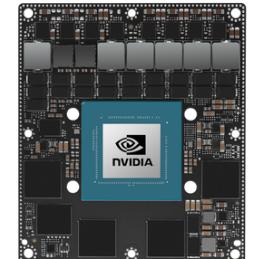
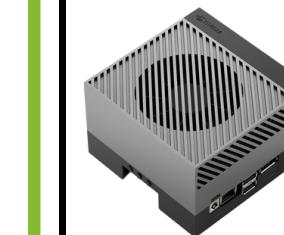
Generación CUDA optimizada desde MATLAB

MATLAB

GPU Coder

CUDA

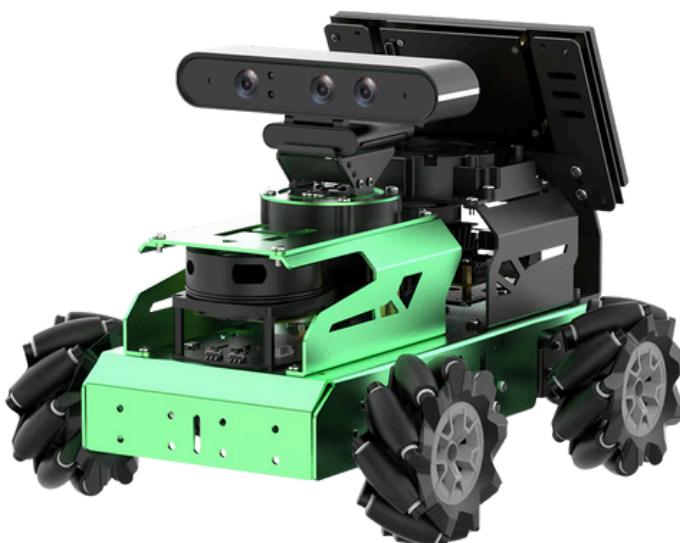
cuBLAS, cuSOLVER,
cuDNN, TensorRT





Desafíos al crear una aplicación de IA para dispositivos de borde (edge)

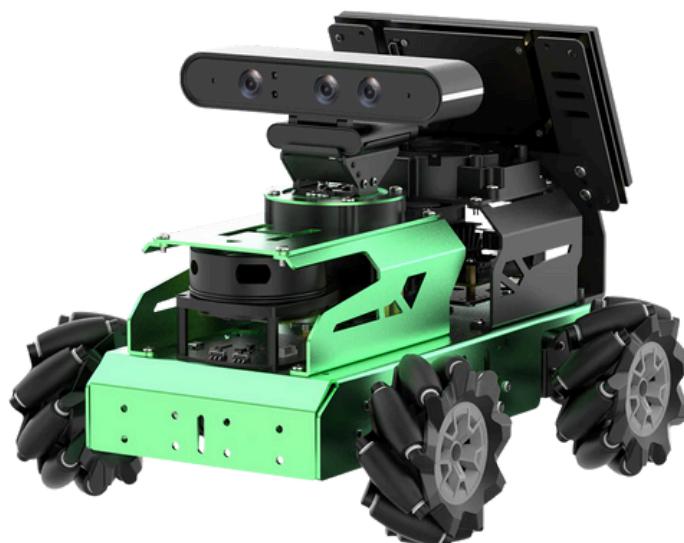
- Elegir el hardware adecuado para la aplicación de IA.
- Simular y probar la aplicación de IA.
- Implementar la aplicación de IA en el dispositivo de borde (edge).



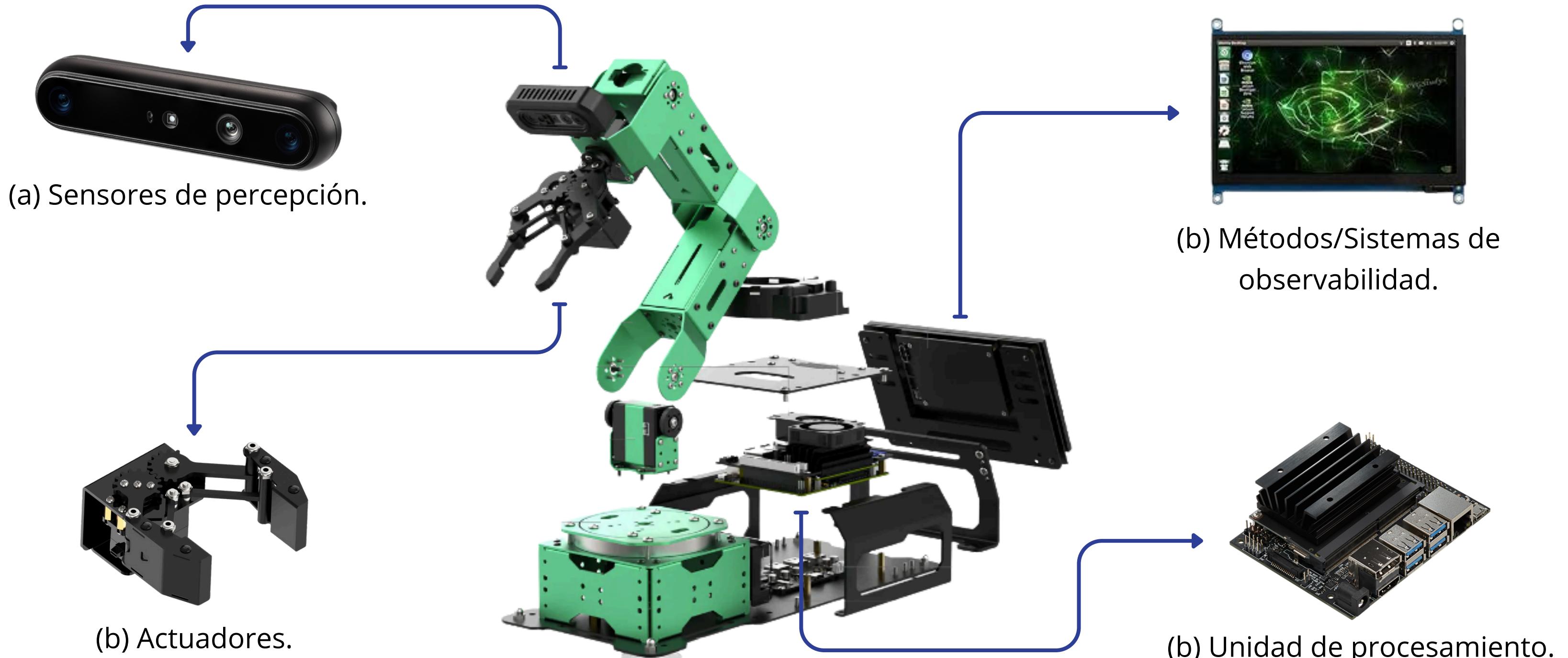


Desafíos al crear una aplicación de IA para dispositivos de borde (edge)

- Elegir el hardware adecuado para la aplicación de IA.
- Simular y probar la aplicación de IA.
- Implementar la aplicación de IA en el dispositivo de borde (edge).



¿Qué es realmente importante?





¿Qué es realmente importante?

Sensores de Percepción

Ver, extraer y recolectar información

- Capturan el entorno real (visión, profundidad, distancias).
- Aportan datos críticos para decisiones en tiempo real.
- La calidad del sensor determina la precisión de la IA.



Actuadores

Actuar y ejecutar

- Ejecutan las decisiones de la IA en el mundo físico.
- Permiten interacción (mover, manipular, transportar).
- Su precisión y velocidad impactan en el rendimiento.



Sistemas de Observabilidad

Monitorear y controlar

- Permiten monitorear y depurar la aplicación.
- Facilitan la interacción humano-máquina.
- Aseguran trazabilidad y confianza en la IA.



ROS



La Familia de Dispositivos Jetson

Desde aplicaciones básicas hasta grandes proyectos...

Múltiples kits de desarrollo – El mismo software



Educadores, Estudiantes, Creadores (Makers)

Desarrolladores de productos comerciales

JETSON NANO 2GB

5W | 10W

0.5 TFLOPS (FP16)

JETSON NANO 4GB

5W | 10W

0.5 TFLOPS (FP16)

JETSON XAVIER NX

10W | 15W

7 TFLOPS (FP16),
21 TOPS (INT8)

JETSON AGX XAVIER

10W | 15W | 30W

11 TFLOPS (FP16),
32 TOPS (INT8)

JETSON AGX ORIN

15W | 30W | 50W

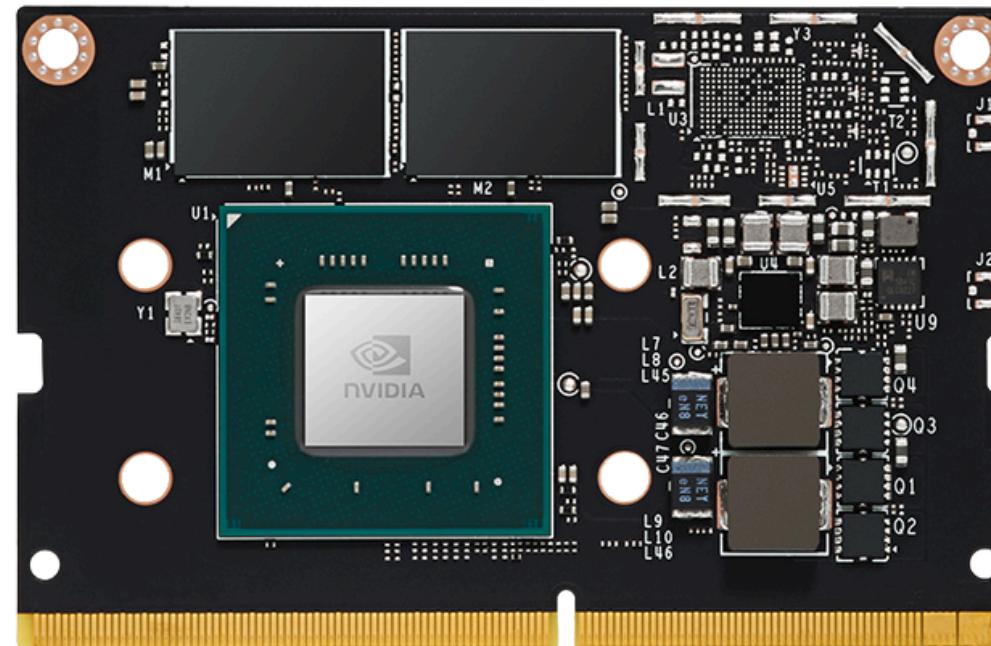
Hasta 60W

270 TOPS (INT8)



Jetson Nano Developer Kit

Categoría	Jetson Nano 4GB
AI Performance	Hasta 472 GFLOPS FP16
GPU	NVIDIA Maxwell GPU, 128 núcleos CUDA
DL Accelerator	No incluye NVDLA
Vision Accelerator	ISP con soporte para sensores RAW hasta 24 MP
CPU	Quad-core ARM Cortex-A57 @ 1.43 GHz, 48KB I-cache + 32KB D-cache por núcleo, 2MB L2 compartida
Memory	4GB LPDDR4, bus 64-bit, 1600 MHz, 25.6 GB/s
Storage	eMMC 5.1 de 16GB integrado
Video Encode	H.265/H.264/VP8 hasta 2160p30 o 1080p120
Video Decode	H.265/H.264/VP9/VP8/MPEG-2/VC-1 hasta 2160p60 o 1080p240
Camera	12 carriles MIPI CSI-2 (soporte hasta 4 streams) / hasta 1400 Mpix/s
PCI Express	1x4, 1x2 + 2x1 PCIe Gen2 (4 carriles máx.)
Ethernet	Gigabit Ethernet (10/100/1000)
Mechanical	69.6 mm × 45 mm, conector SO-DIMM 260 pines
Power	5–10 W, entrada 5V



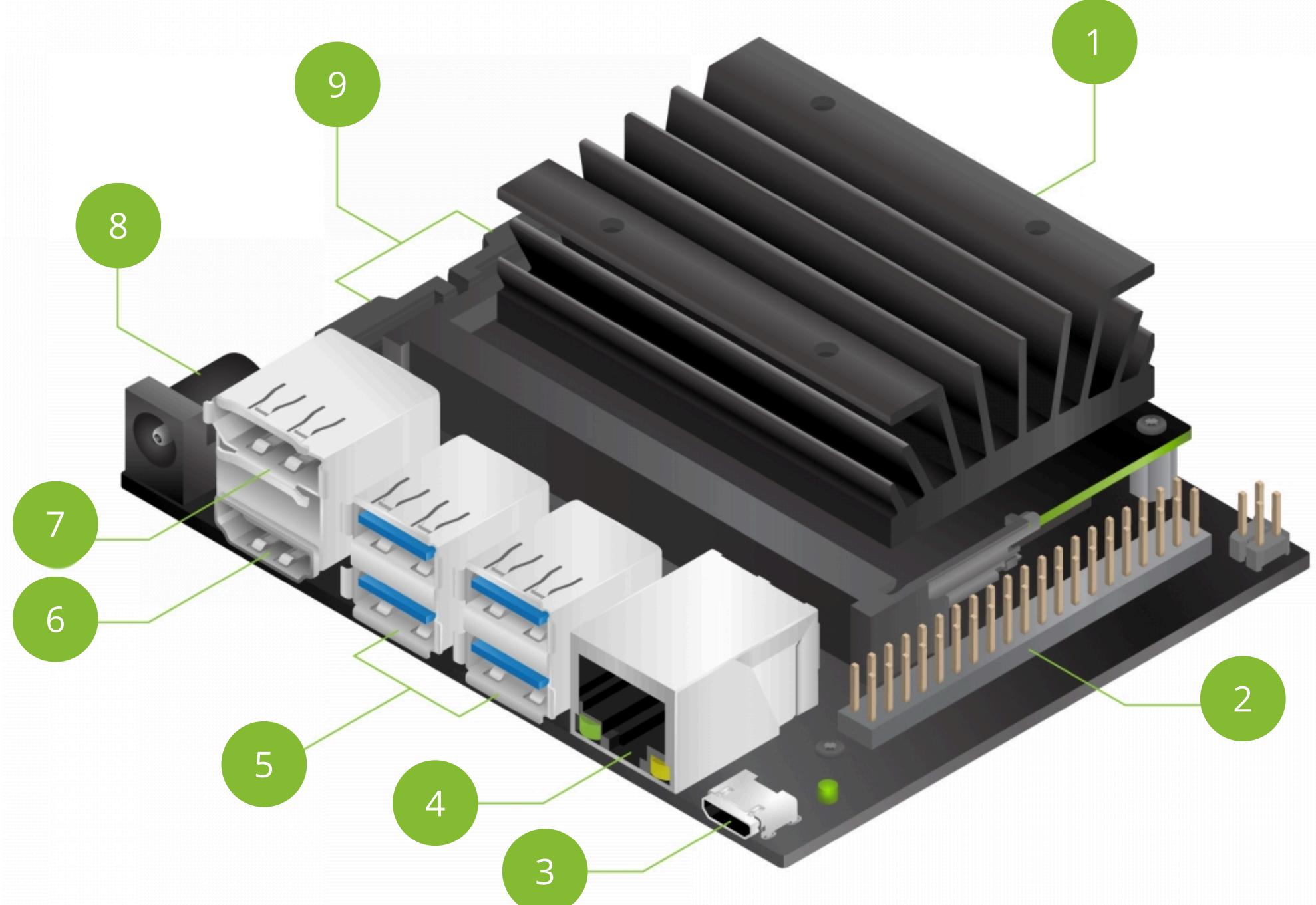
(a) Tarjeta Jetson Nano.



(b) Jetson Nano Developer Kit 4GB.



Jetson Nano Developer Kit

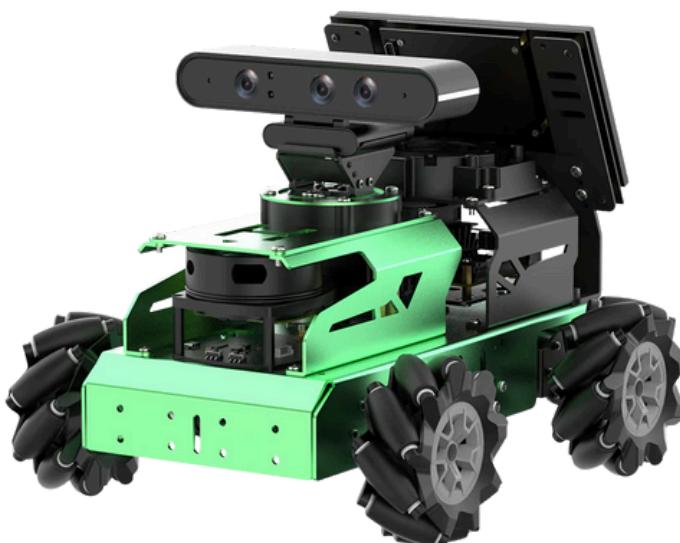


- 1 Ranura para tarjeta microSD para almacenamiento principal.
- 2 Cabezal de expansión de 40 pines.
- 3 Puerto micro-USB para entrada de alimentación de 5 V o para modo de dispositivo.
- 4 Puerto Gigabit Ethernet.
- 5 Puertos USB 3.0 (x4).
- 6 Puerto de salida HDMI
- 7 Conector DisplayPort
- 8 Conector de barril de CC para entrada de alimentación de 5 V
- 9 Conectores de cámara MIPI CSI-2



Desafíos al crear una aplicación de IA para dispositivos de borde (edge)

- Elegir el hardware adecuado para la aplicación de IA.
- Simular y probar la aplicación de IA.
- Implementar la aplicación de IA en el dispositivo de borde (edge).





Pipeline de Desarrollo

Preparar los Datos

- Extraer/descargar datos.
- Preprocesamiento de los datos.
- Análisis y perspectiva humana.

Entrenar Modelos

- Diseño de modelos y entrenamiento.
- Entrenamiento acelerado por Hardware.
- Interoperabilidad

Simular y Probar

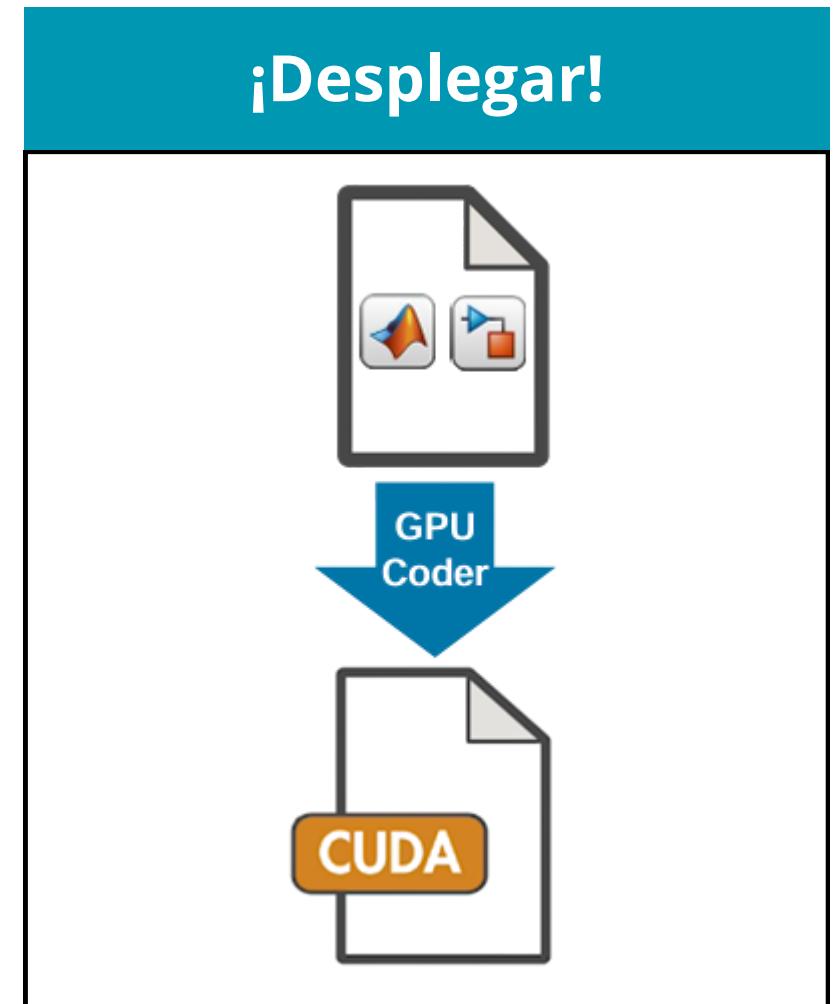
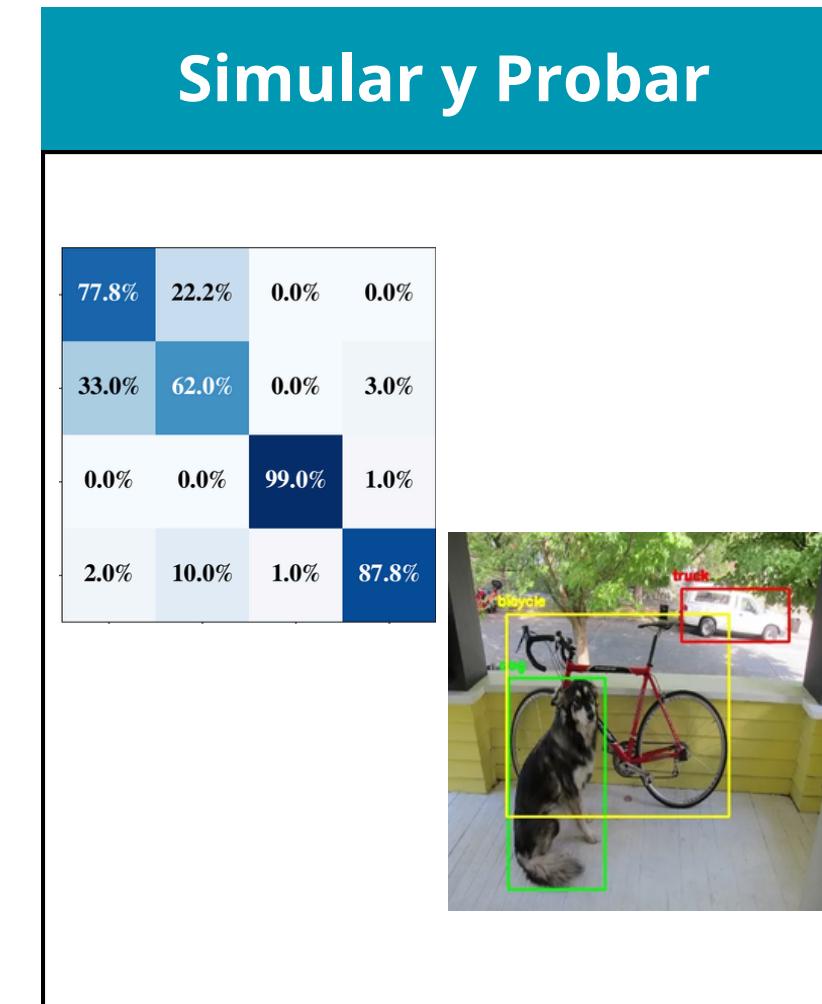
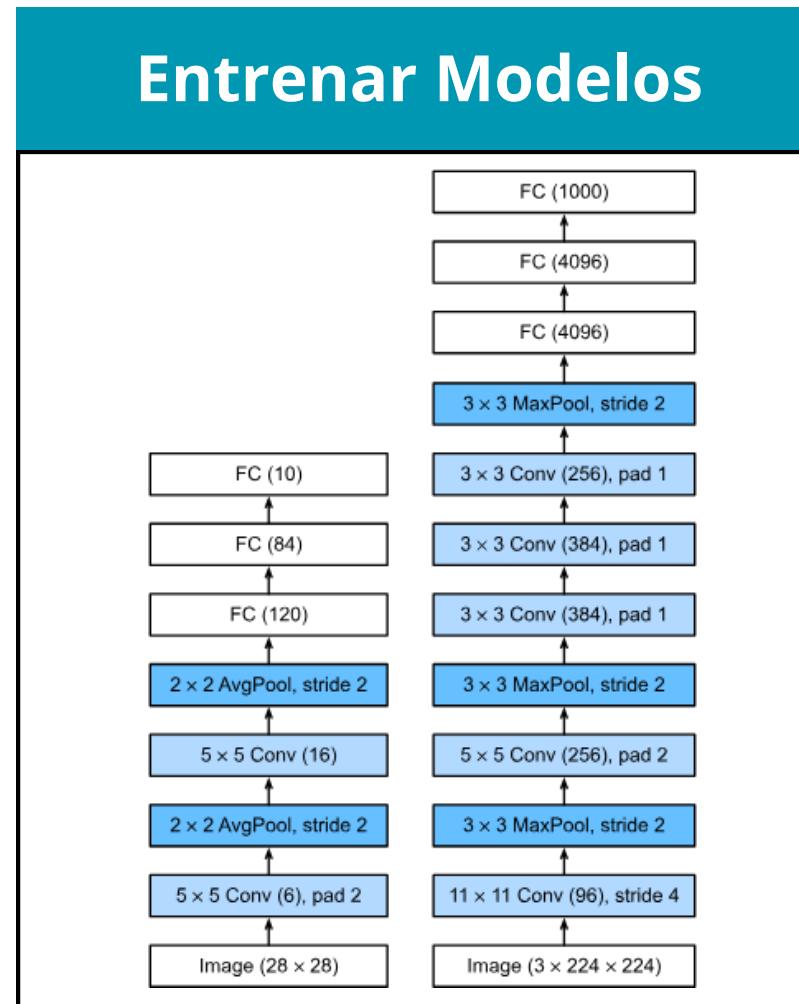
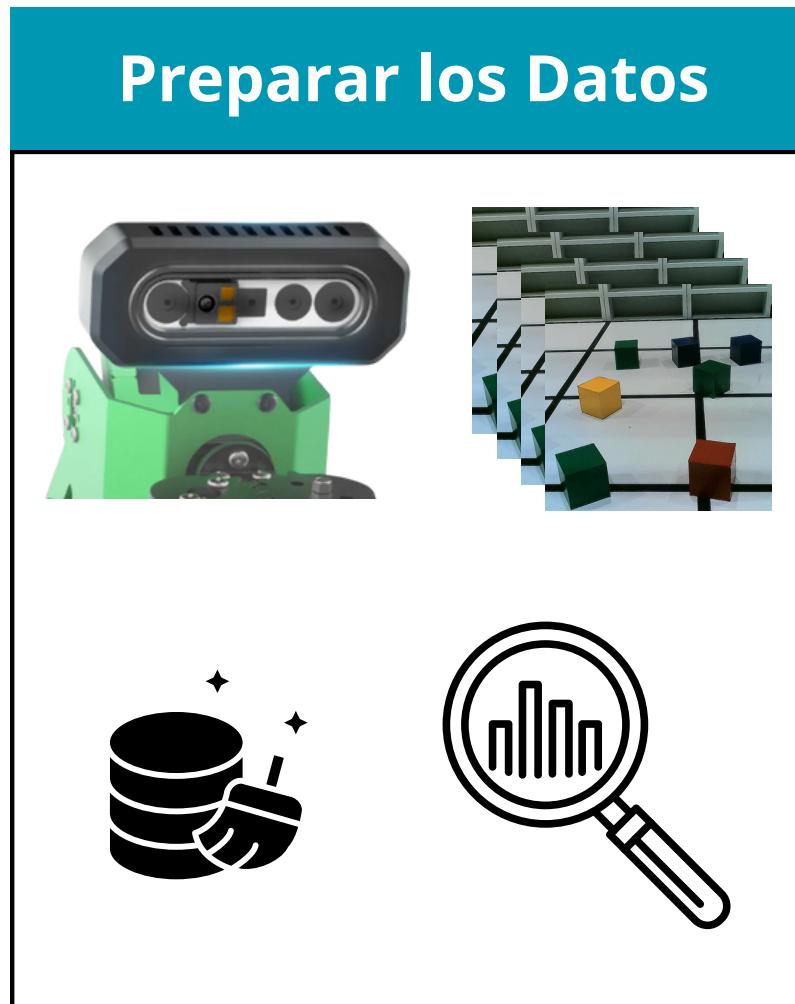
- Fine-tuning de modelos.
- Simulación del sistema.
- Verificación y validación.

¡Desplegar!

- Sistemas embebidos.
- Clusters.
- Cloud, On-Premise.



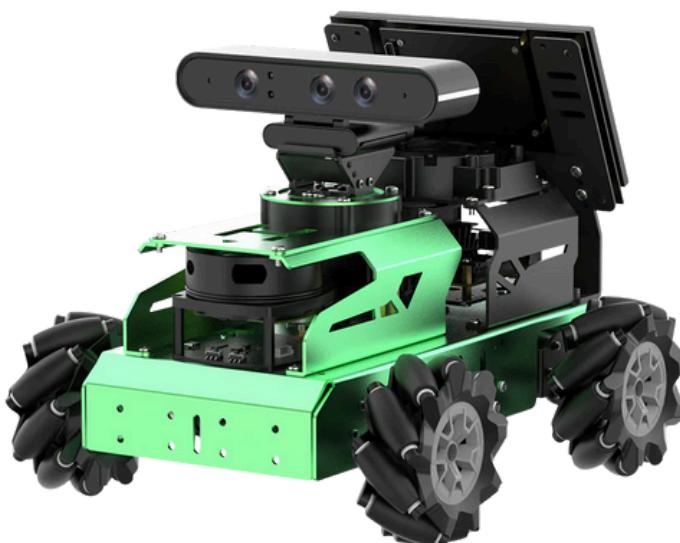
Pipeline de Desarrollo





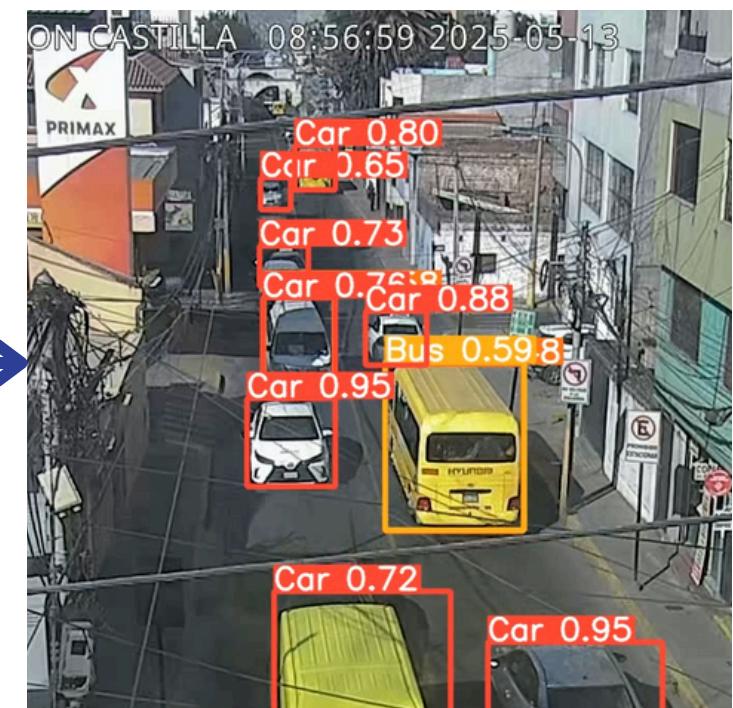
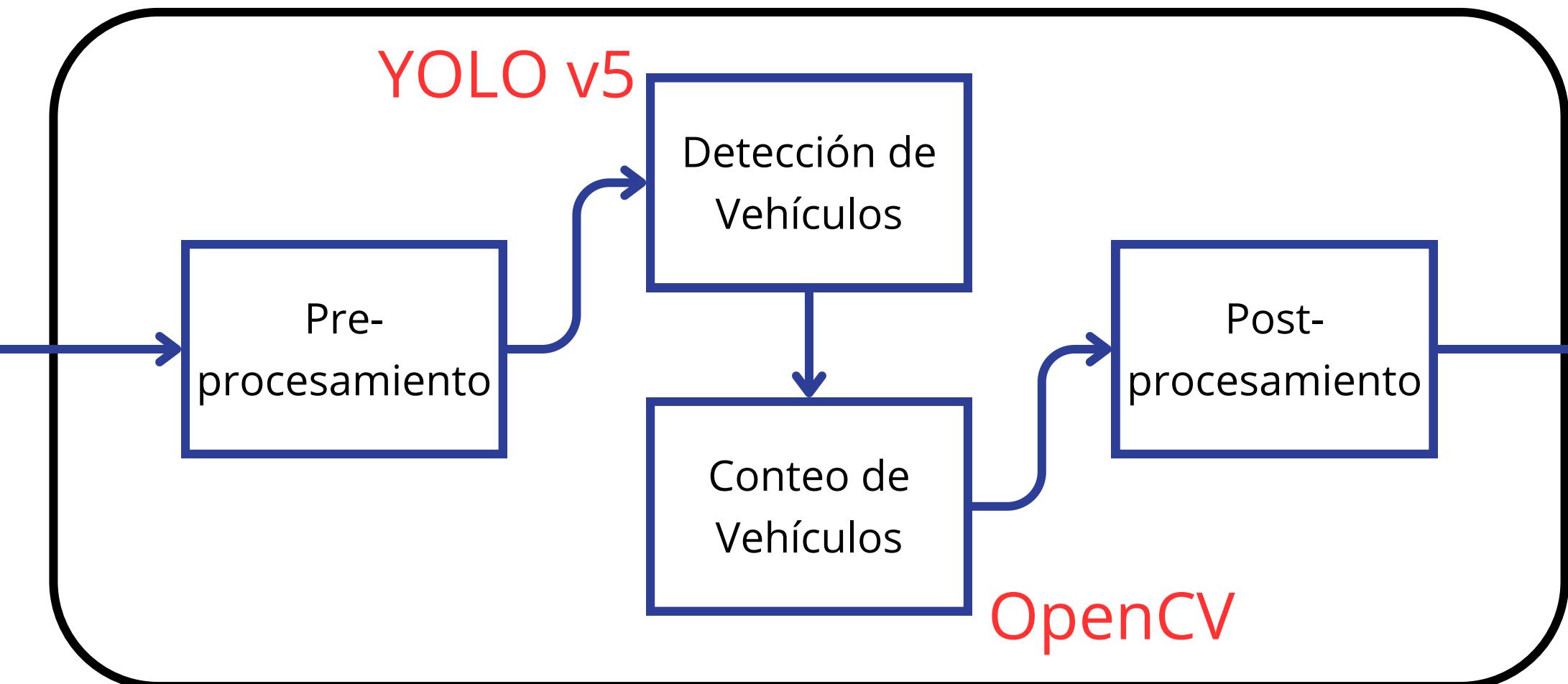
Desafíos al crear una aplicación de IA para dispositivos de borde (edge)

- Elegir el hardware adecuado para la aplicación de IA.
- Simular y probar la aplicación de IA.
- Implementar la aplicación de IA en el dispositivo de borde (edge).



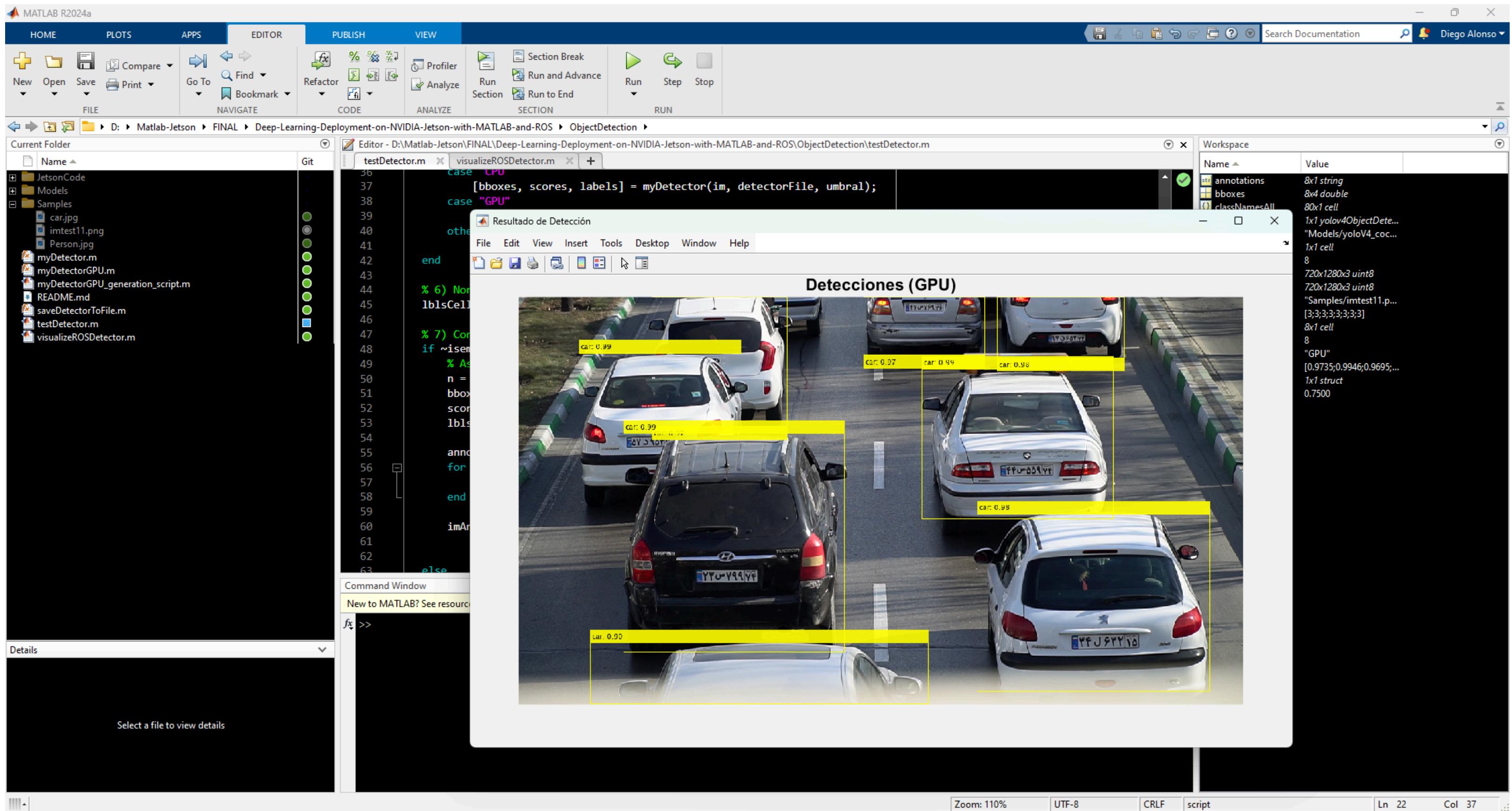


Ejemplo: Detección y Conteo de Vehículos



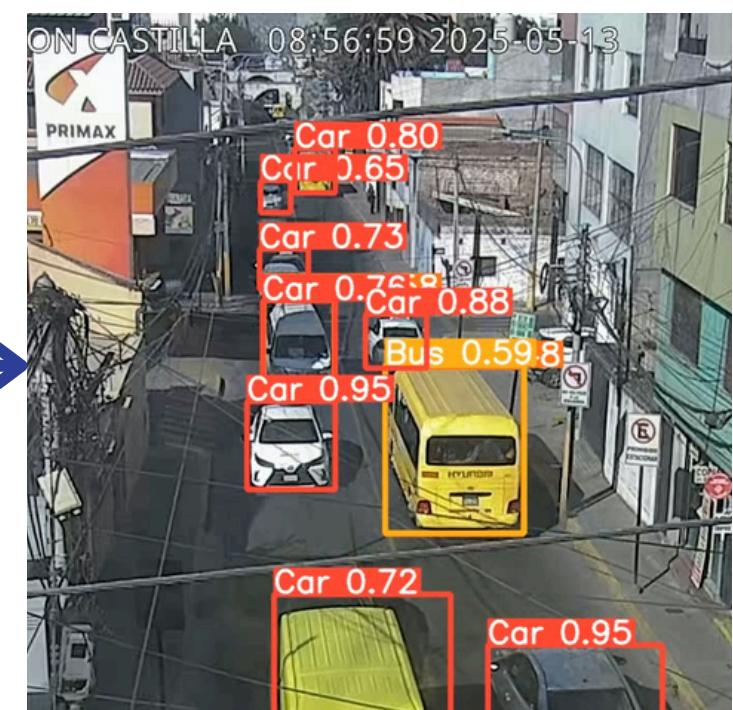
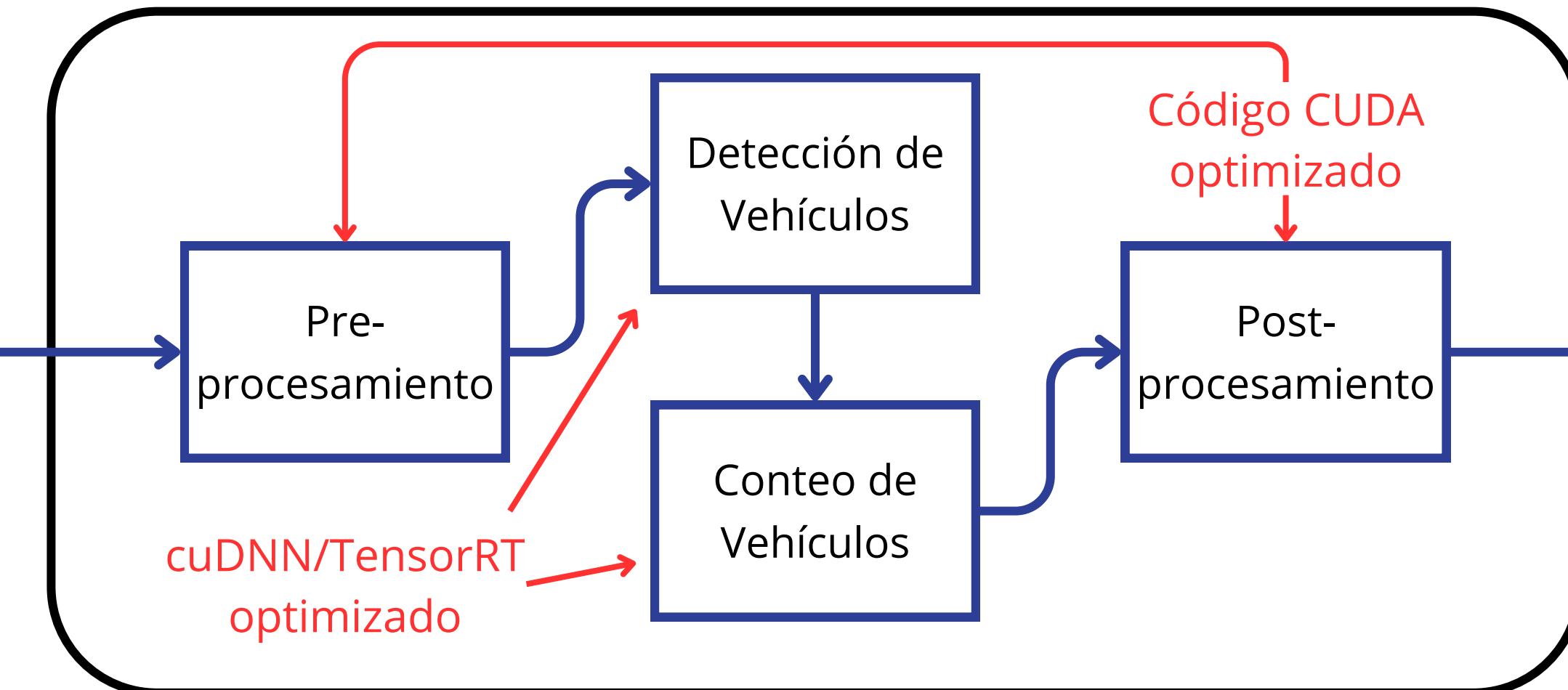


Ejecutar simulación en la GPU local (on-premise)



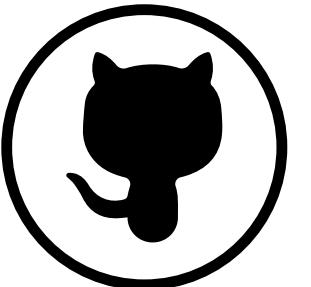


Usar GPU Coder para generar código CUDA y desplegarlo en Jetson Nano Dev. Kit





Página de recursos



DiegoABR07/Deep-Learning-Deployment...

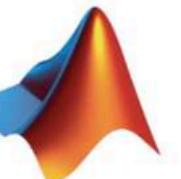


Implementation of object detection and classification models on NVIDIA Jetson Nano Developer Kit, using MATLAB R2024a, GPU Coder, cuDNN, TensorRT...

1 Contributor 0 Issues 0 Stars 0 Forks



matlab-deep-learning/MATLAB-Deep-Learning...



Discover pretrained models for deep learning in MATLAB

4 Contributors 12 Issues 533 Stars 117 Forks



¿PREGUNTAS?



Diego Banda R.
Universidad Católica San Pablo
diego.banda@ucsp.edu.pe

