

Sistemas de visión artificial y fusión sensorial

Prof. Dr. David Martín Gómez
Catedrático de Universidad

Laboratorio de Sistemas Inteligentes
Dpto. de Ingeniería Eléctrica
Universidad Carlos III de Madrid

uc3m | Universidad
Carlos III
de Madrid



Curso de especialización
Automatización de vehículos
(Aplicación y oportunidades en el Sector Transporte)

8ª edición

24 de marzo de 2026



Índice

1. Vehículos autónomos

- Motivación y antecedentes
- Taxonomía

2. Sistemas de percepción y fusión de sensores

- Sensores
- *Deep learning*

3. Aplicaciones y fusión de sensores

- Clasificación, detección y localización de objetos
- Segmentación semántica
- Casos específicos de fusión de sensores

Transporte



«Night lights» de dr_zoidberg disponible en Flickr (www.flickr.com/photos/zoidberg72/4081385620/) bajo licencia CC BY-SA 2.0

Transporte



Problemas del transporte



Accidentes

- 1,4 millones de muertes al año en el mundo: más de 9 Boeing 747-400 al día
- Primera causa de muerte en el mundo entre los 5-29 años
- +90% debidos a errores humanos



Contaminación

- Siete millones de muertes por exposición a partículas en el aire.
- 23% de las emisiones de gases de efecto invernadero.

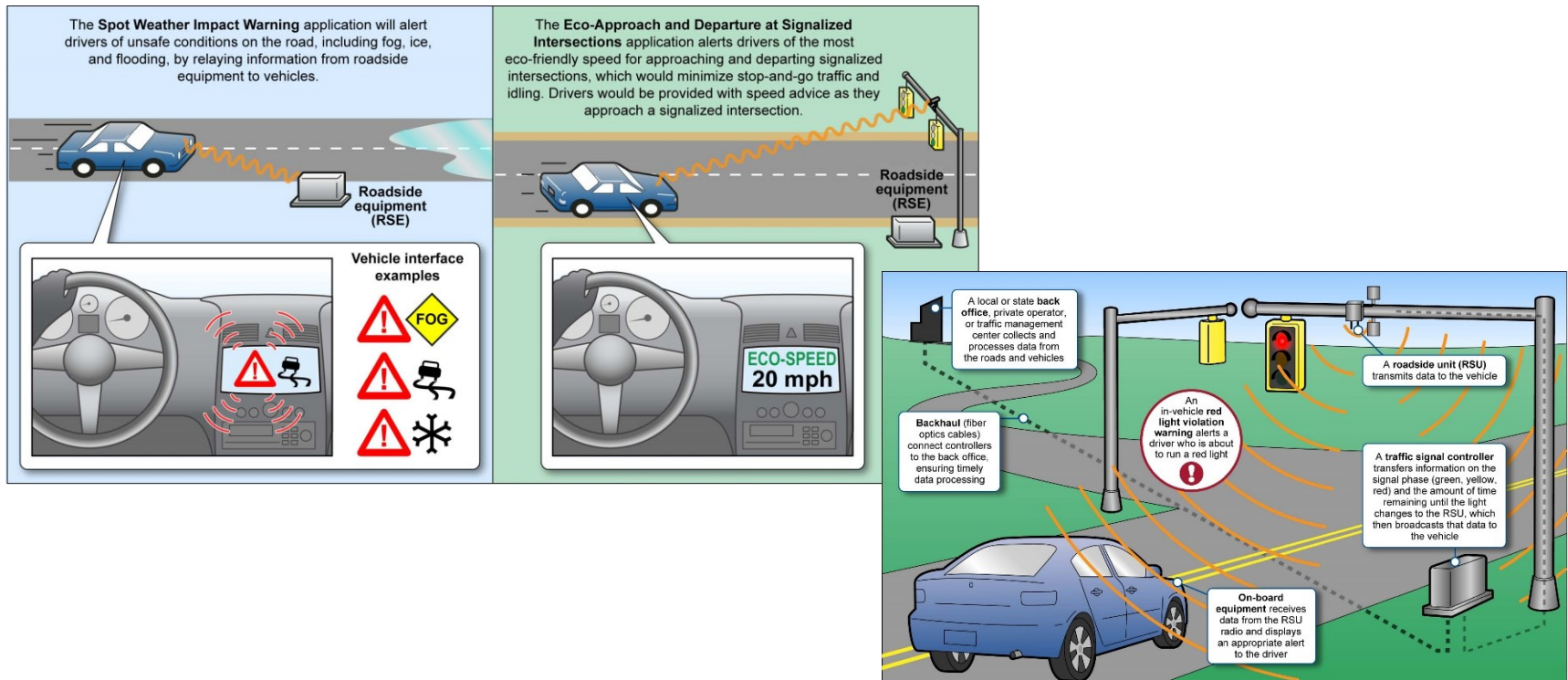


Atascos

- 2% del PIB en Europa.

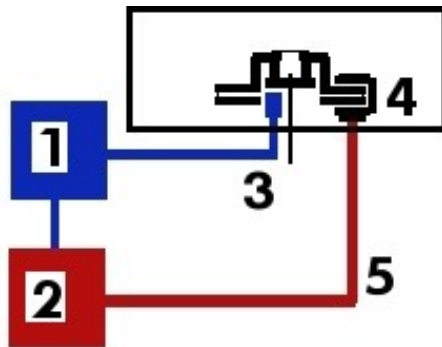
Sistemas Inteligentes de Transporte (ITS)

- Aplicación de la tecnología para mejorar el transporte (carretera):
 - Seguridad
 - Eficiencia
 - Confort



Seguridad activa en los vehículos

- La tecnología se ha utilizado desde los inicios del automóvil para reducir y mitigar los accidentes
- Avance empujado por el desarrollo de la electrónica y los sensores



ABS (1978)



ESP (1990)

Sistemas de asistencia al conductor

Sistemas Avanzados de Asistencia al Conductor (ADAS)

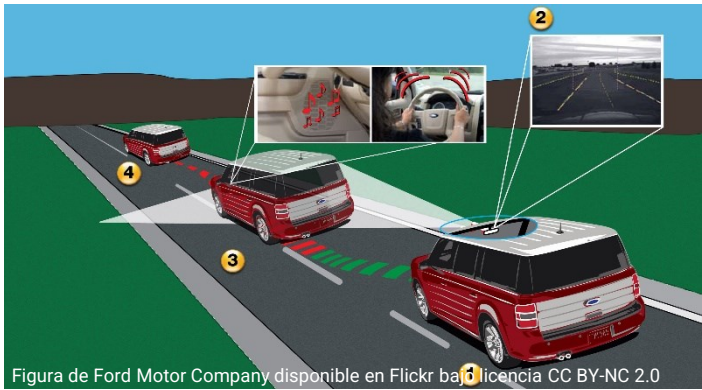
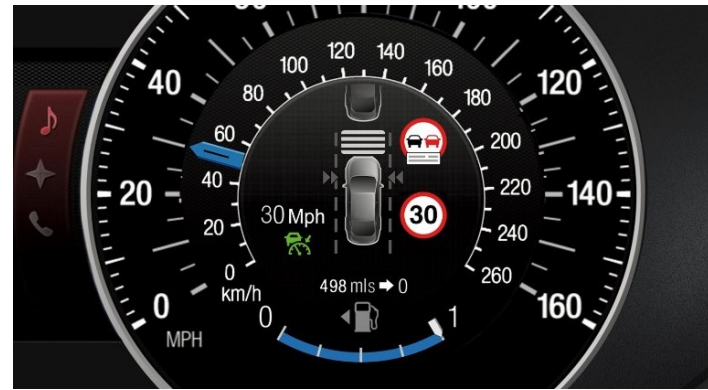


Figura de Ford Motor Company disponible en Flickr bajo licencia CC BY-NC 2.0

Salida de carril (LDW/LKA)



Asistente velocidad (ISA)



Foto de Jaguar MENA disponible en Flickr bajo licencia CC BY 2.0

Freno de emergencia (AEB)



Foto de Emancipator disponible en Wikimedia bajo licencia CC BY-SA 3.0

Detección de punto ciego

Vehículo autónomo

Un vehículo autónomo es aquel capaz de realizar todas las tareas necesarias para moverse de un punto a otro sin necesidad de intervención humana



Vehículo autónomo

- Eliminar al conductor de la ecuación
- Potenciales beneficios:



Reducción de accidentes



Uso más eficiente de las infraestructuras



Uso más eficiente del tiempo



Mejora de la productividad



Acceso al transporte sin restricciones



Cambio de paradigma



Nuevos modelos de negocio

¿Ciencia ficción?

- Fuertes inversiones de fabricantes:
 - Ford: 1.000 millones \$ en Argo AI (2017)
 - Toyota: 2.800 millones \$ en Toyota Research Institute (2018)
 - etc.
- Predicciones de los expertos:
 - Capricho para ricos alrededor de los 2020-30
 - Adopción generalizada a finales de los 2030-inicio de los 2040

Fuente: T. Litman, "Autonomous vehicle implementation predictions. Implications for Transport Planning," Victoria Transport Policy Institute, 2018.



Foto de Dllu disponible en Wikimedia bajo licencia CC BY-SA 4.0



Foto de Metropolitan Area Council disponible en Flickr bajo licencia CC BY 2.0

Antecedentes

- Primeros desarrollos en los 1980-90
- **DARPA Challenges** 2004-2007
 - Promovidos por EE UU para automatizar sus vehículos militares
 - Participación de universidades y empresas
 - Desierto (2004, 2005) y entorno urbano (2007)



Foto de Rob NREC disponible en Wikimedia bajo licencia CC BY-SA 3.0

Antecedentes

Google Self-driving Car (ahora Waymo)

- En 2010, Google desvela sus coches autónomos: han recorrido 140 000 millas sin accidentes.
- 2018: Más de 8 millones de km recorridos en carreteras reales
- Waymo One: servicio de “robotaxis” en Phoenix



Demostraciones públicas alrededor del mundo

- PROUD (VisLab): en tráfico real en Parma (Italia)
- Bertha Benz Memorial Route (Daimler y KIT): 103 km
- BB8 (NVIDIA): 98% de autonomía

A día de hoy...

- *Pilotos automáticos:*
 - Tesla Autopilot
 - Mercedes Drive Pilot
 - ...
- **Funciones:**
 - Mantenimiento de carril
 - Control de crucero adaptativo
 - Frenado de emergencia

Requieren la atención constante del conductor

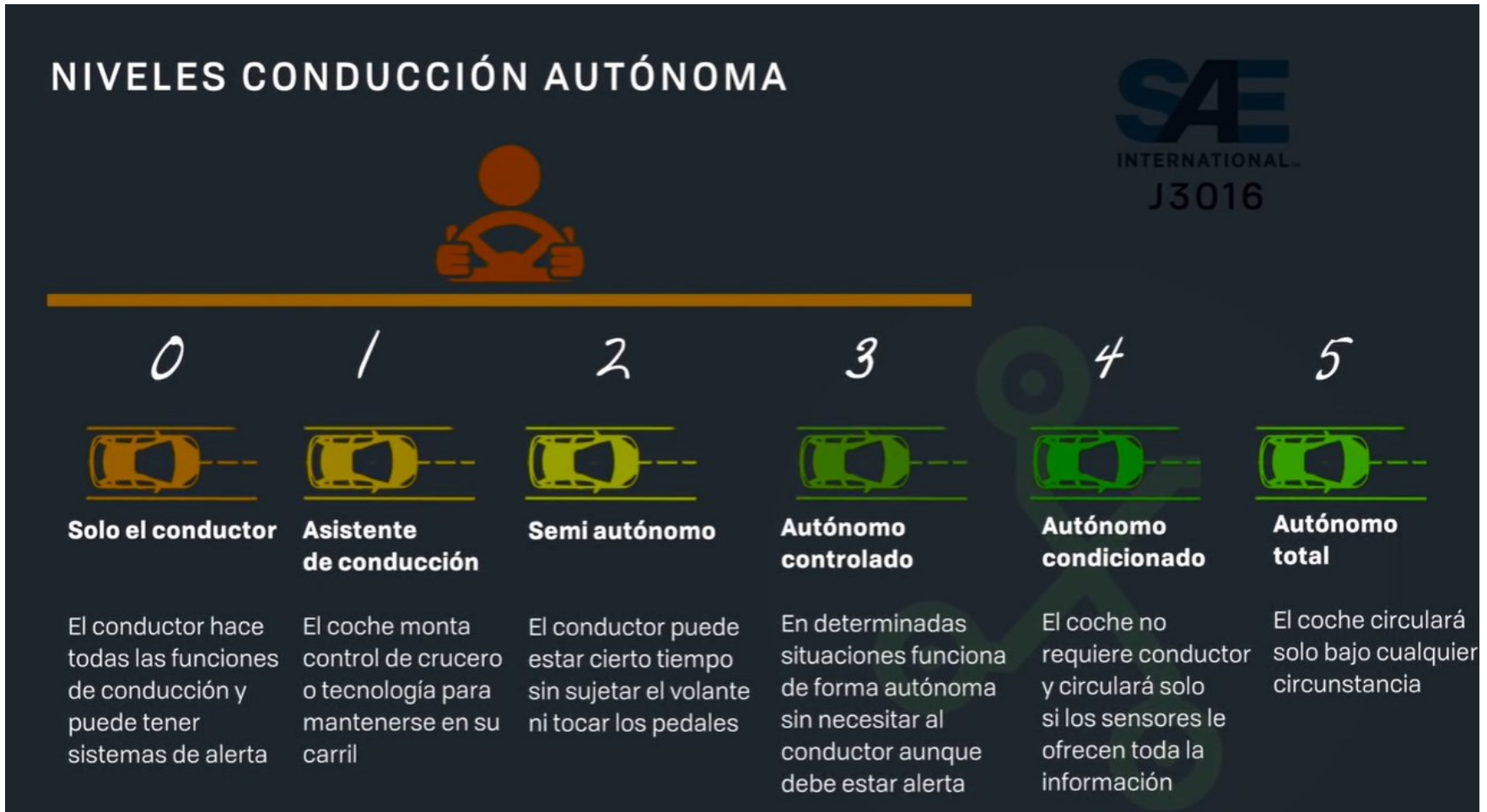


Foto de Marc van der Chijs disponible en Flickr bajo licencia CC BY-ND 2.0



Foto de EurovisionNim disponible en Wikimedia bajo licencia CC BY-SA 4.0

Estado actual de la tecnología: 2010-2026



Niveles

- Definidos por SAE International (desarrolladora de estándares)

Nivel	Nombre	Aceleración, dirección,...	Supervisión del entorno	Respuesta a errores	Modos de conducción
0	Sin automatización	Conductor	Conductor	Conductor	No
1	Asistencia al conductor	Conductor y sistema	Conductor	Conductor	Algunos
2	Automatización parcial	Sistema	Conductor	Conductor	Algunos
3	Automatización condicional	Sistema	Sistema	Conductor	Algunos
4	Automatización alta	Sistema	Sistema	Sistema	Algunos
5	Automatización completa	Sistema	Sistema	Sistema	Todos

Niveles: situación actual (2026)

- **Nivel 0:** coches sin ADAS

Nivel	Nombre	Aceleración, dirección,...	Supervisión del entorno	Respuesta a errores	Modos de conducción
0	Sin automatización	Conductor	Conductor	Conductor	No
1	Asistencia al conductor	Conductor y sistema	Conductor	Conductor	Algunos
2	Automatización parcial	Sistema	Conductor	Conductor	Algunos
3	Automatización condicional	Sistema	Sistema	Conductor	Algunos
4	Automatización alta	Sistema	Sistema	Sistema	Algunos
5	Automatización completa	Sistema	Sistema	Sistema	Todos

Niveles: situación actual (2026)

- **Nivel 1:** Algunos ADAS (mantenimiento de carril, crucero adaptativo, etc.)

Nivel	Nombre	Aceleración, dirección,...	Supervisión del entorno	Respuesta a errores	Modos de conducción
0	Sin automatización	Conductor	Conductor	Conductor	No
1	Asistencia al conductor	Conductor y sistema	Conductor	Conductor	Algunos
2	Automatización parcial	Sistema	Conductor	Conductor	Algunos
3	Automatización condicional	Sistema	Sistema	Conductor	Algunos
4	Automatización alta	Sistema	Sistema	Sistema	Algunos
5	Automatización completa	Sistema	Sistema	Sistema	Todos

Niveles: situación actual (2026)

- **Nivel 2:** “Pilotos automáticos” actuales

Nivel	Nombre	Aceleración, dirección,...	Supervisión del entorno	Respuesta a errores	Modos de conducción
0	Sin automatización	Conductor	Conductor	Conductor	No
1	Asistencia al conductor	Conductor y sistema	Conductor	Conductor	Algunos
2	Automatización parcial	Sistema	Conductor	Conductor	Algunos
3	Automatización condicional	Sistema	Sistema	Conductor	Algunos
4	Automatización alta	Sistema	Sistema	Sistema	Algunos
5	Automatización completa	Sistema	Sistema	Sistema	Todos

Niveles: situación actual (2026)

- **Nivel 3 y siguientes:** Solo prototipos

Nivel	Nombre	Aceleración, dirección,...	Supervisión del entorno	Respuesta a errores	Modos de conducción
0	Sin automatización	Conductor	Conductor	Conductor	No
1	Asistencia al conductor	Conductor y sistema	Conductor	Conductor	Algunos
2	Automatización parcial	Sistema	Conductor	Conductor	Algunos
3	Automatización condicional	Sistema	Sistema	Conductor	Algunos
4	Automatización alta	Sistema	Sistema	Sistema	Algunos
5	Automatización completa	Sistema	Sistema	Sistema	Todos

Niveles: situación actual (2026)

- ¿Dónde está la brecha ? --> **sistemas de percepción y fusión de sensores**

Nivel	Nombre	Aceleración, dirección,...	Supervisión del entorno	Respuesta a errores	Modos de conducción
0	Sin automatización	Conductor	Conductor	Conductor	No
1	Asistencia al conductor	Conductor y sistema	Conductor	Conductor	Algunos
2	Automatización parcial	Sistema	Conductor	Conductor	Algunos
3	Automatización condicional	Sistema	Sistema	Conductor	Algunos
4	Automatización alta	Sistema	Sistema	Sistema	Algunos
5	Automatización completa	Sistema	Sistema	Sistema	Todos

Índice

1. Vehículos autónomos

- Motivación y antecedentes
- Taxonomía

2. Sistemas de percepción y fusión de sensores

- Sensores
- *Deep learning*

3. Aplicaciones y fusión de sensores

- Clasificación, detección y localización de objetos
- Segmentación semántica
- Casos específicos de fusión de sensores

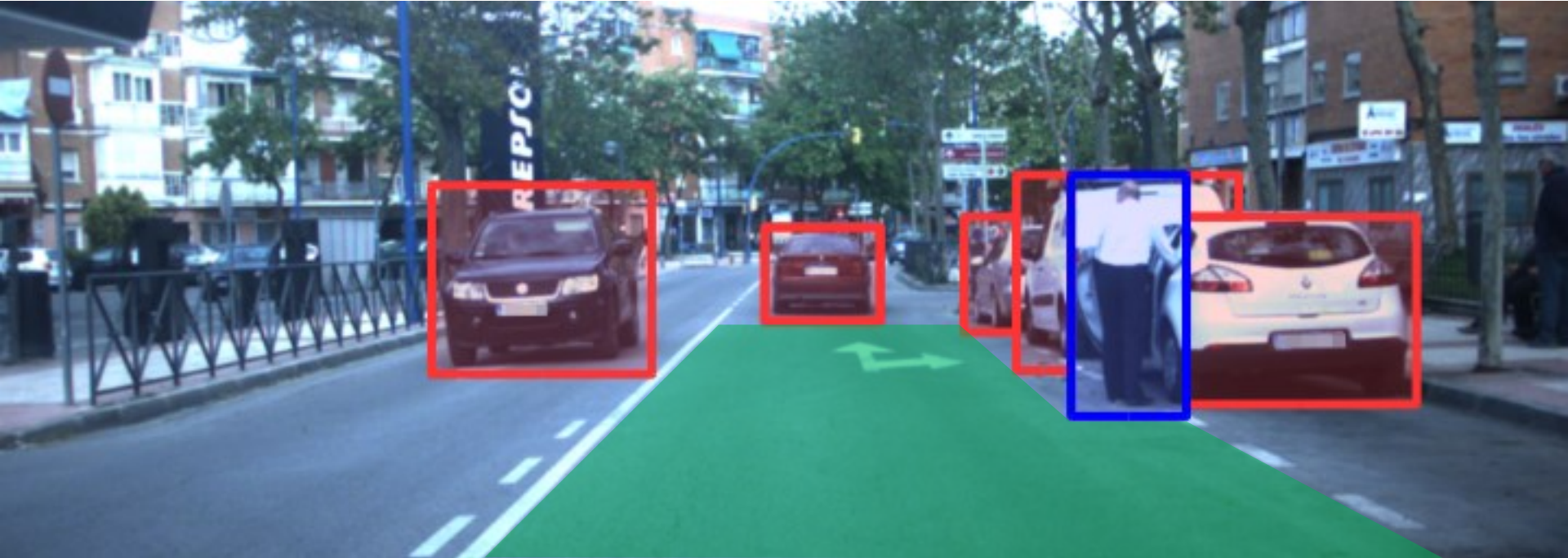
Entendimiento del entorno viario

- La percepción del entorno es clave, y también la tarea más compleja:



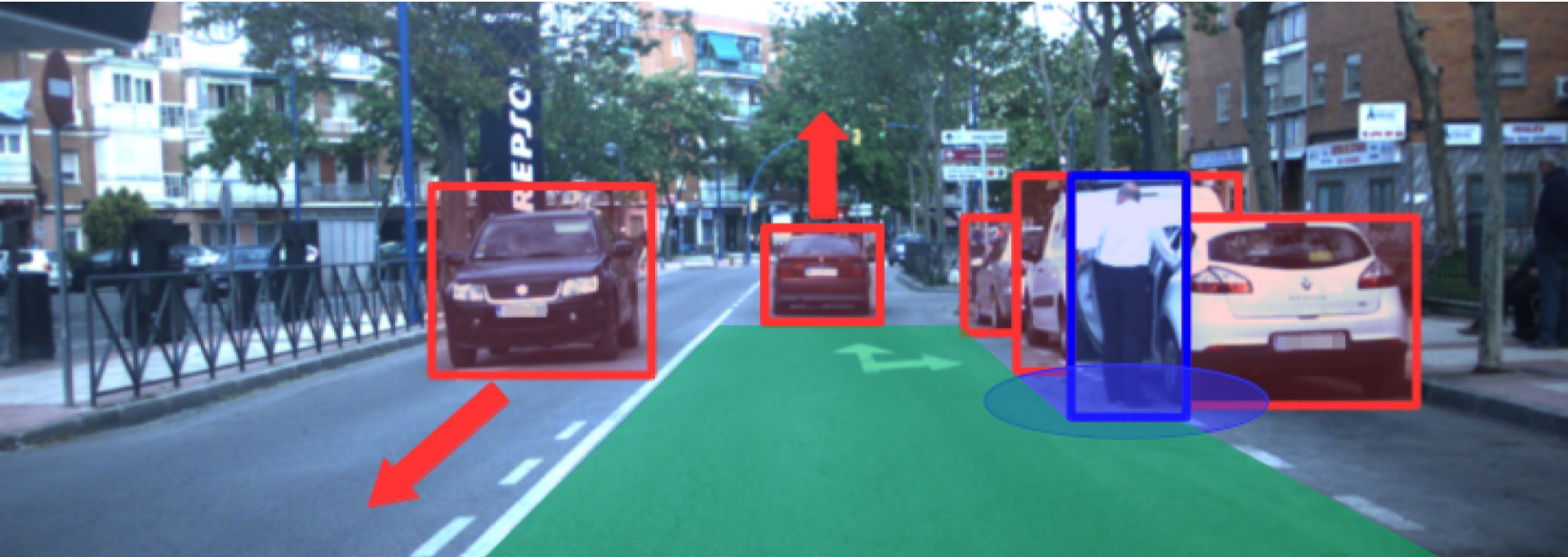
Entendimiento del entorno viario

- La percepción del entorno es clave, y también la tarea más compleja:



Entendimiento del entorno viario

- La percepción del entorno es clave, y también la tarea más compleja:



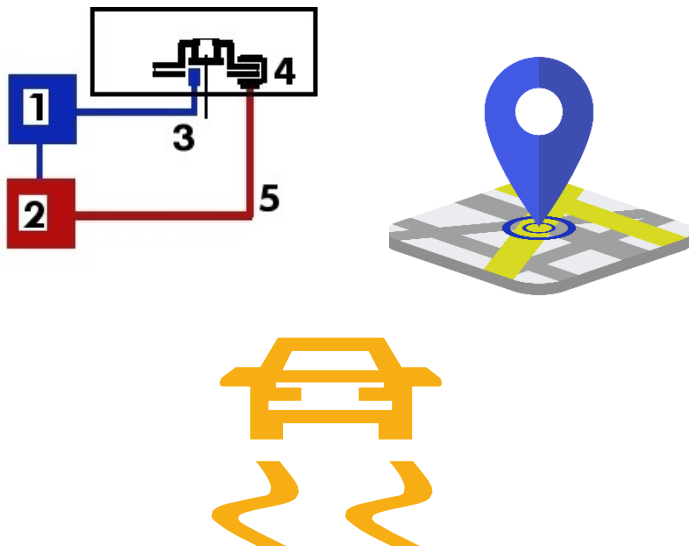
- Variedad de agentes, tamaños y puntos de vista
- Oclusiones
- Condiciones del entorno variables (iluminación, clima, etc.)

Los entornos de tráfico son poco estructurados

Sensores

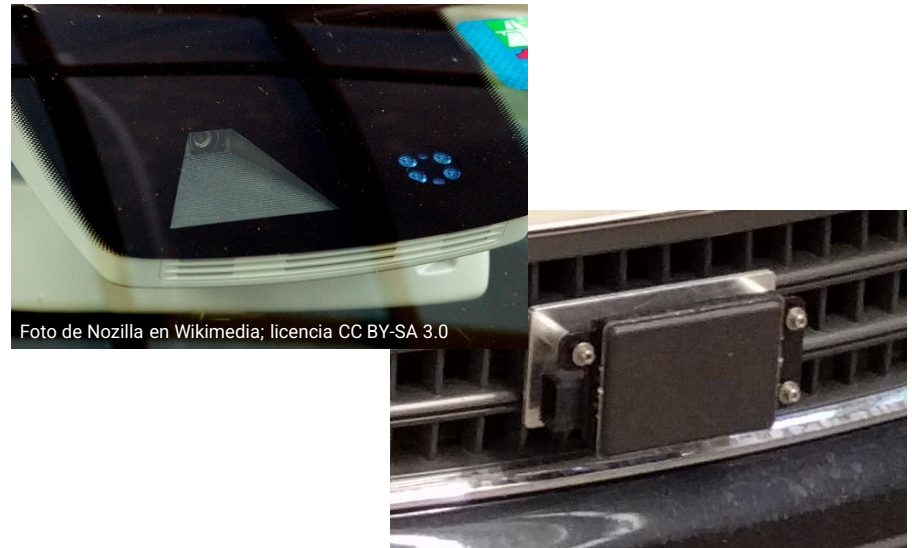
Propioceptivos

- Miden variables internas
- Ej: deslizamiento de las ruedas, localización del propio vehículo



Exteroceptivos

- Captan el entorno alrededor del vehículo
- Sistemas de percepción



Sensores

Exteroceptivos

Activos

- Emiten energía y miden el reflejo
 - Ondas de radio, láser, etc.

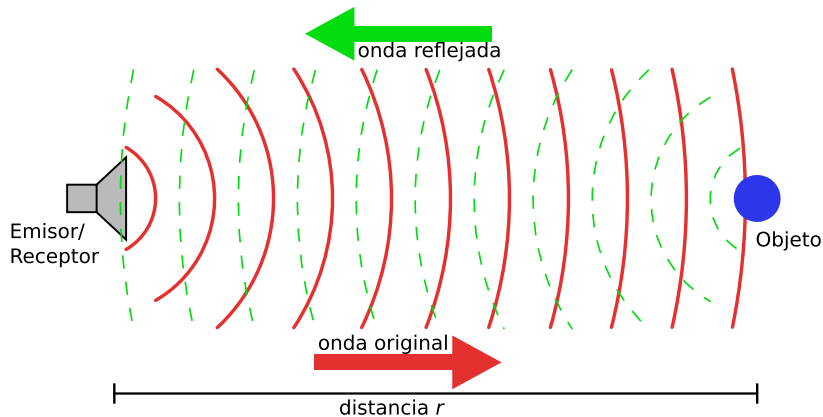
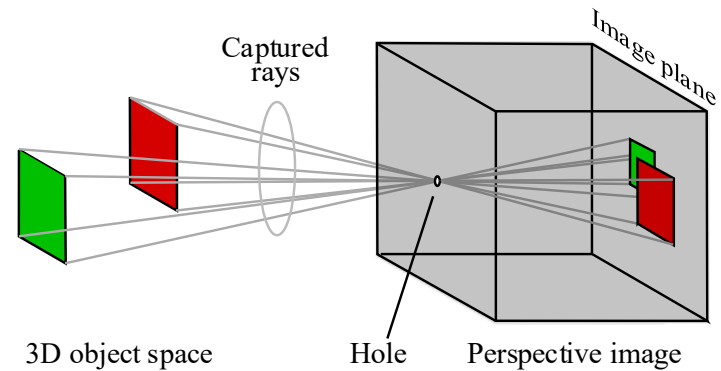


Figura de Georg Wiora en Wikimedia; licencia CC BY-SA 3.0

Ultrasonidos, radar, lidar

Pasivos

- Usan la energía del entorno
 - Luz reflejada en los objetos



Cámaras de vídeo, infrarrojas

Sensores exteroceptivos activos

Ultrasonidos

- ✓ Sencillos y baratos
- ✗ Rango muy limitado

Ej.: Asistente aparcamiento

Radar

- ✓ Mayor rango de detección
- ✗ Información escasa

Ej.: ACC, Frenado automático

Lidar

- ✓ Fiable y robusto
- ✓ Gran rango (hasta 200 m)
- ✗ Elevado coste
- ✗ Información limitada

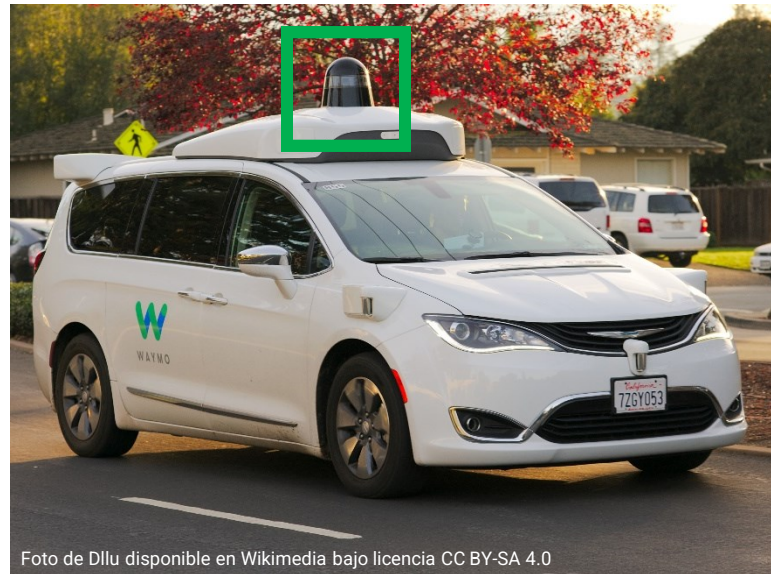


Foto de Dllu disponible en Wikimedia bajo licencia CC BY-SA 4.0

Proporcionan información de distancias

Sensores exteroceptivos activos

Ultrasonidos

- ✓ Sencillos y baratos
- ✗ Rango muy limitado

Ej.: Asistente aparcamiento

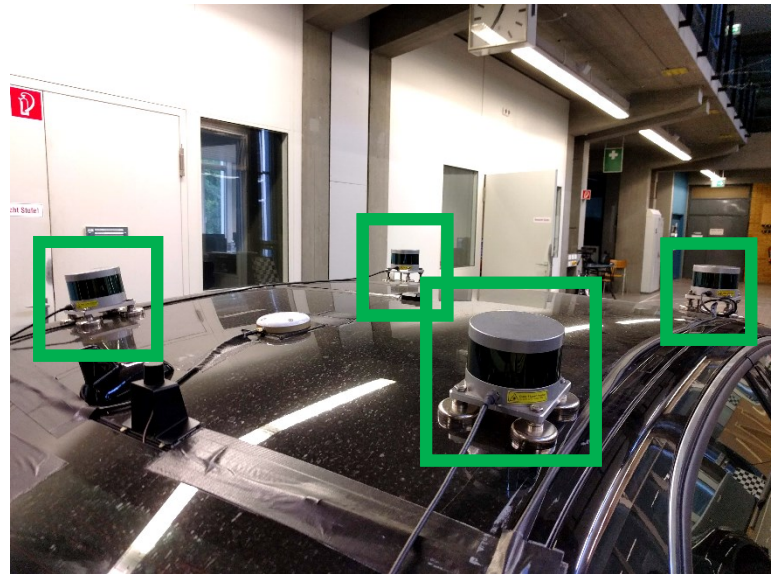
Radar

- ✓ Mayor rango de detección
- ✗ Información escasa

Ej.: ACC, frenado automático

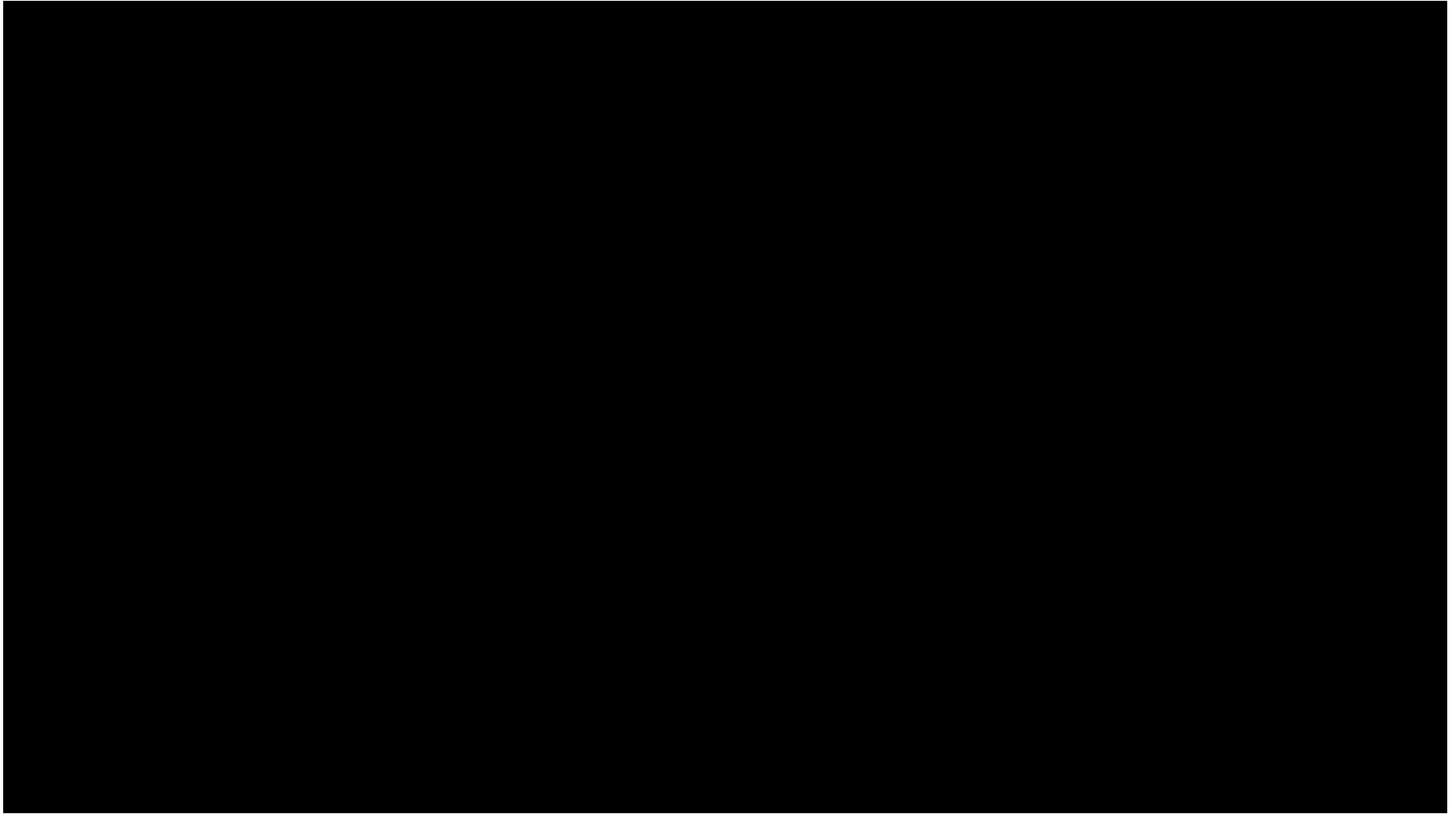
Lidar

- ✓ Fiable y robusto
- ✓ Gran rango (hasta 200 m)
- ✗ Elevado coste
- ✗ Información limitada

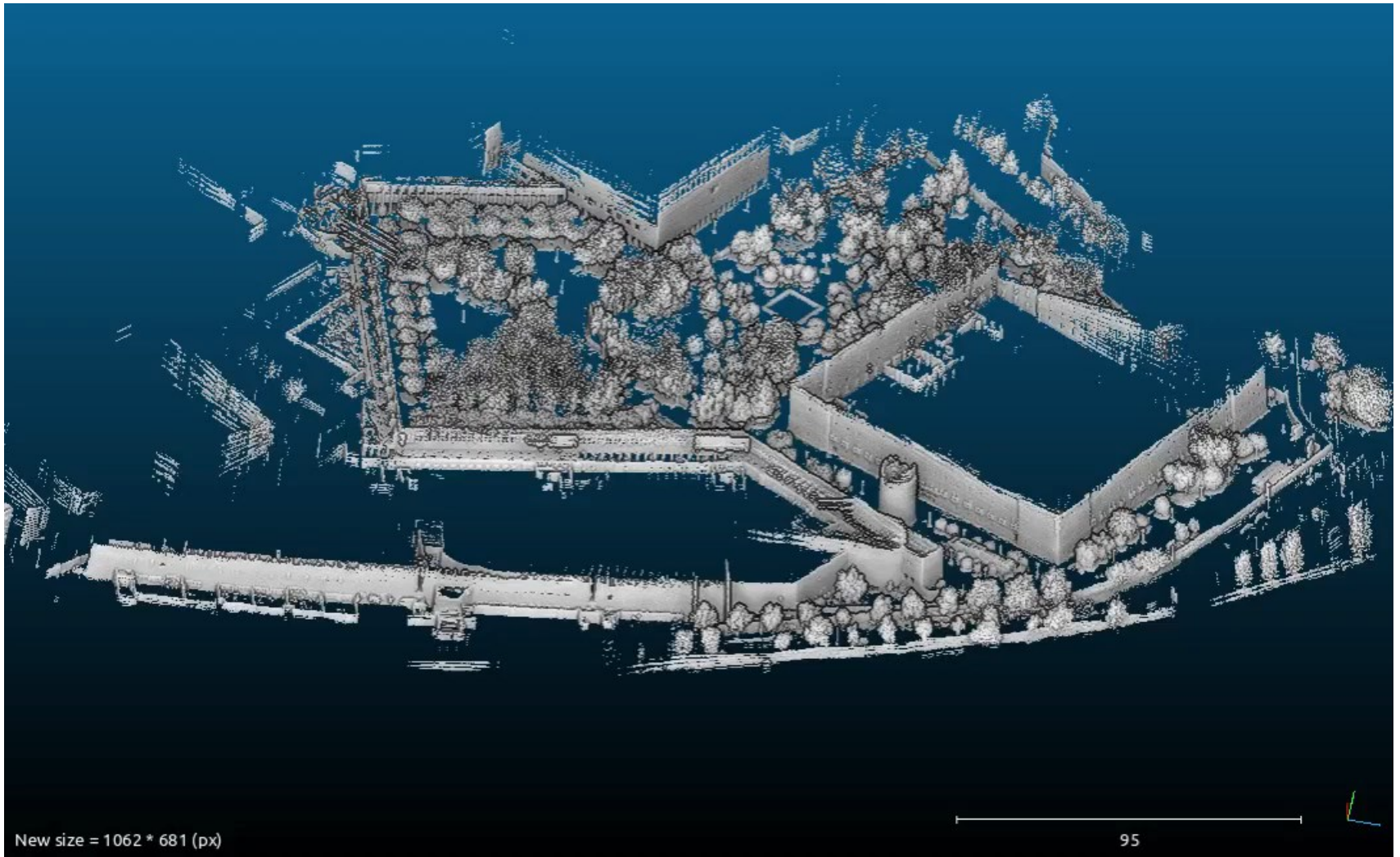


Proporcionan información de distancias

Lidar



Lidar



Lidar

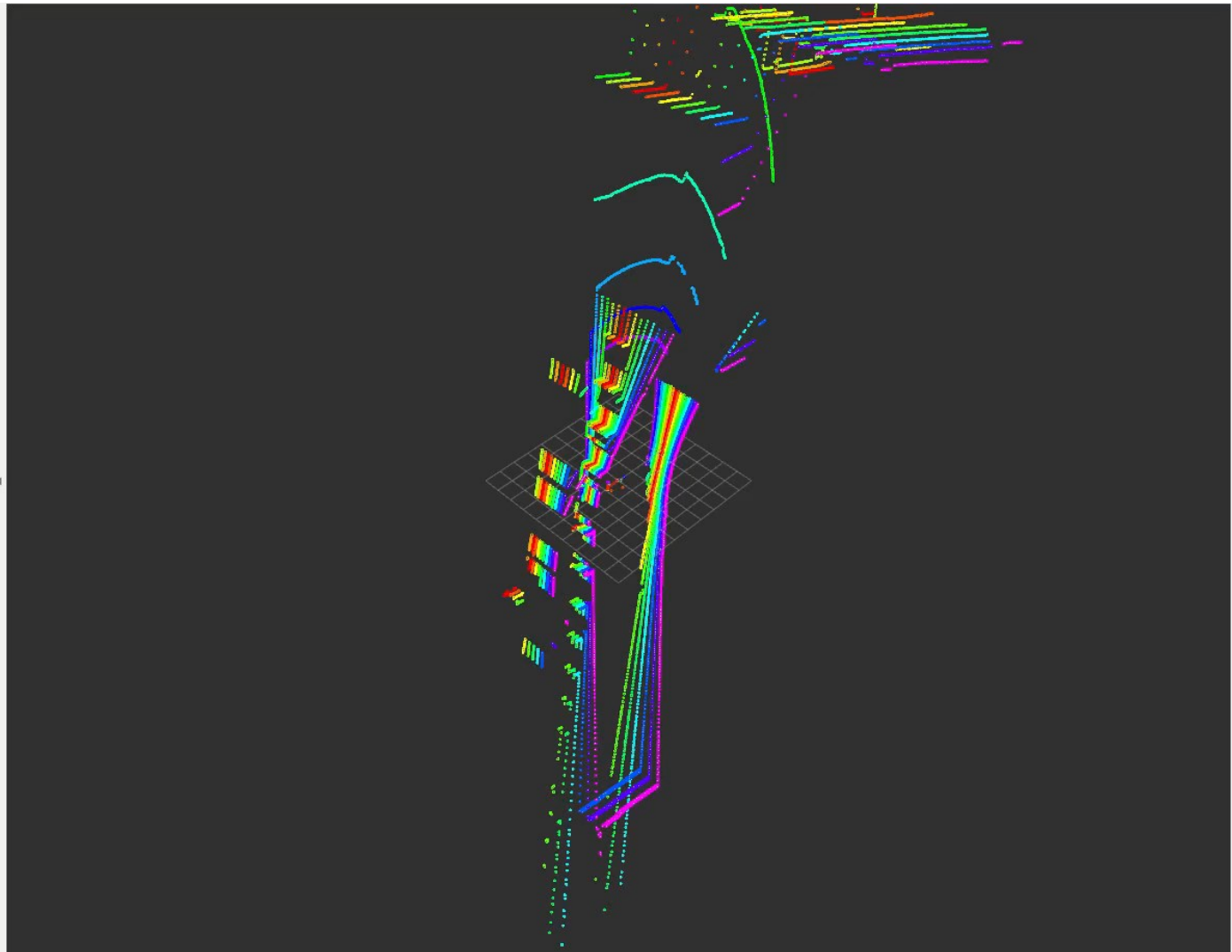
The screenshot shows the ROS Displays window with the following configuration:

- Global Options**
 - Fixed Frame: camera_init
 - Background Color: 48; 48; 48
 - Frame Rate: 30
- Global Status: W...**
 - Fixed Frame: No tf data. Actual error...
- Grid**:
- TF**:
- Odometry**:
- PointCloud2**:
- Odometry**:
- PointCloud2**:
- Image**:
 - Status: Ok
 - Image Topic: /stereo_camera/right/i...
 - Transport Hint: raw
 - Queue Size: 2
 - Unreliable:

Decay Time
Duration, in seconds, to keep the incoming points. 0 means only show the latest points.

Buttons: Add, Duplicate, Remove, Rename

Image window showing a camera view of a brick-paved area with a building in the background.



Sensores exteroceptivos pasivos

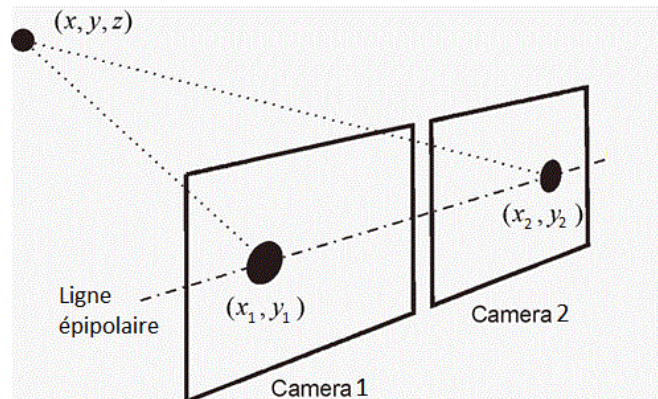
Cámaras de vídeo

- ✓ Información rica
- ✗ Poco robusto (iluminación, clima)
- ✗ Dificultad de procesamiento



Proporcionan información de apariencia

Sistemas estéreo



Sensores exteroceptivos pasivos

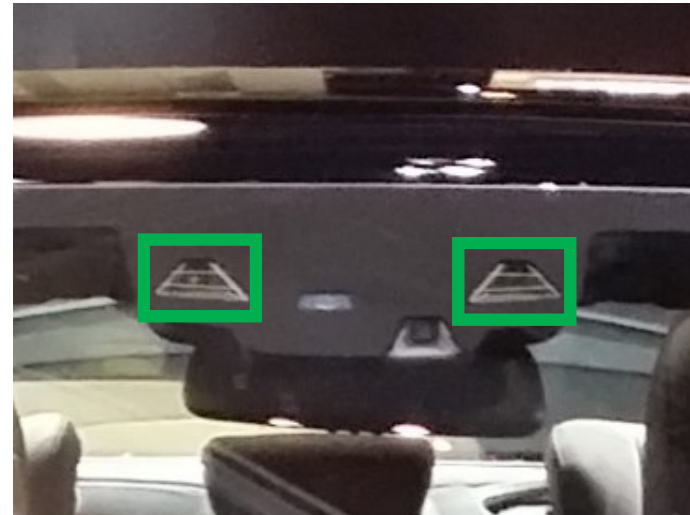
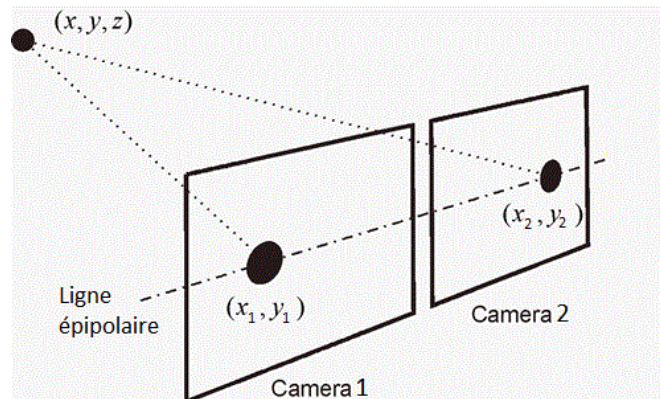
Cámaras de vídeo

- ✓ Información rica
- ✗ Poco robusto (iluminación, clima)
- ✗ Dificultad de procesamiento

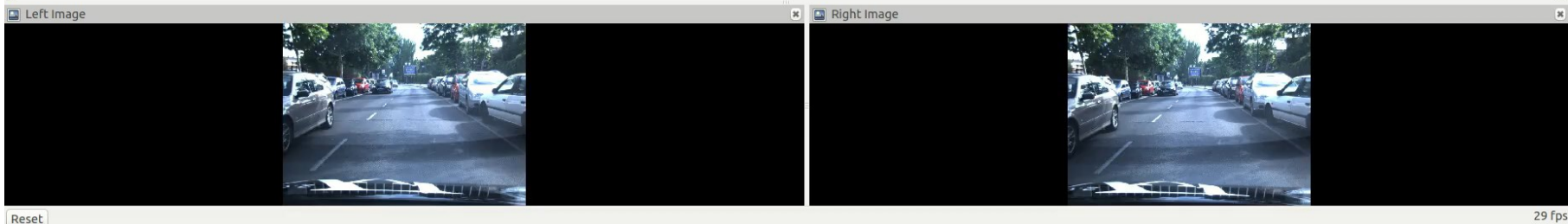


Proporcionan información de apariencia

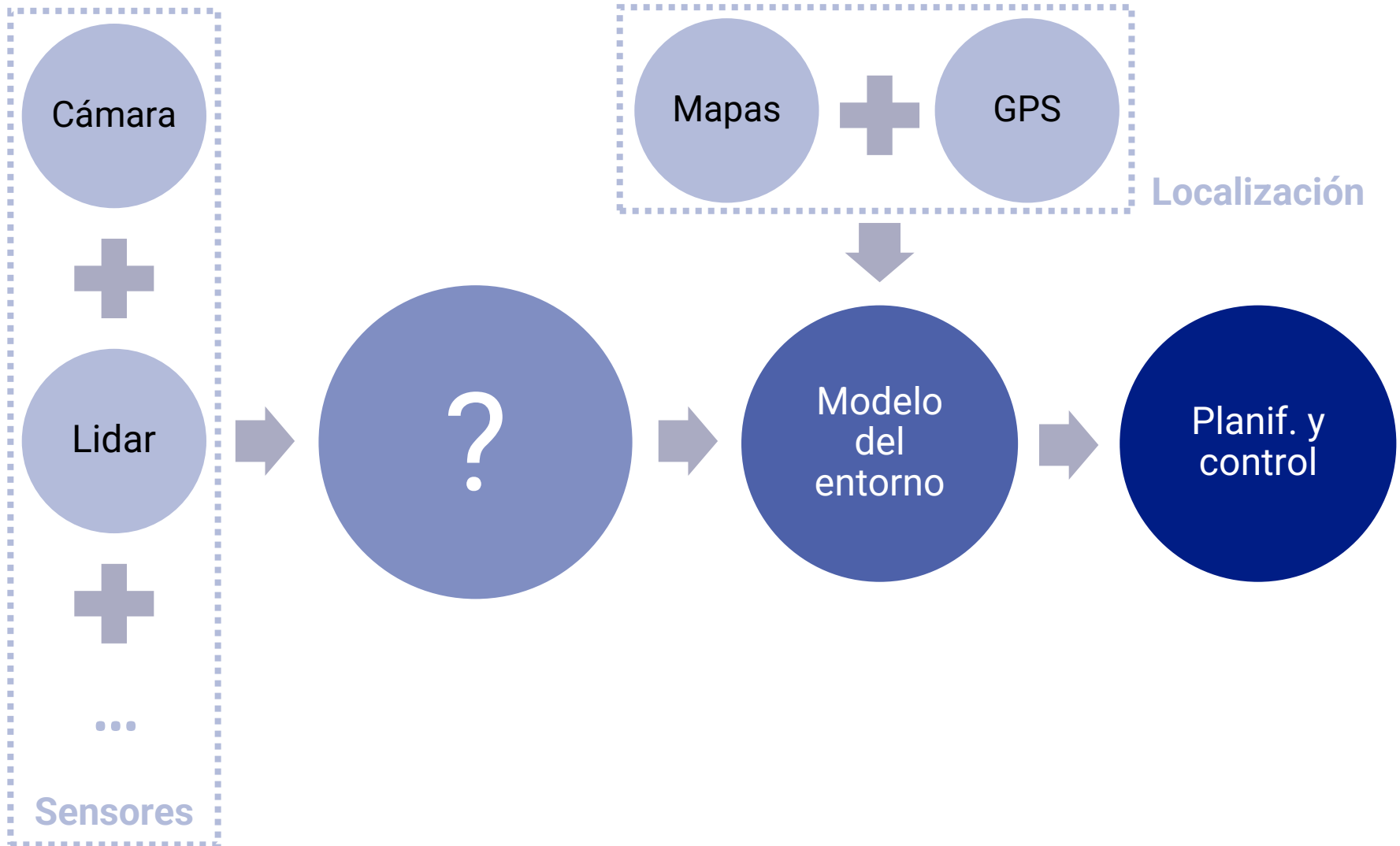
Sistemas estéreo



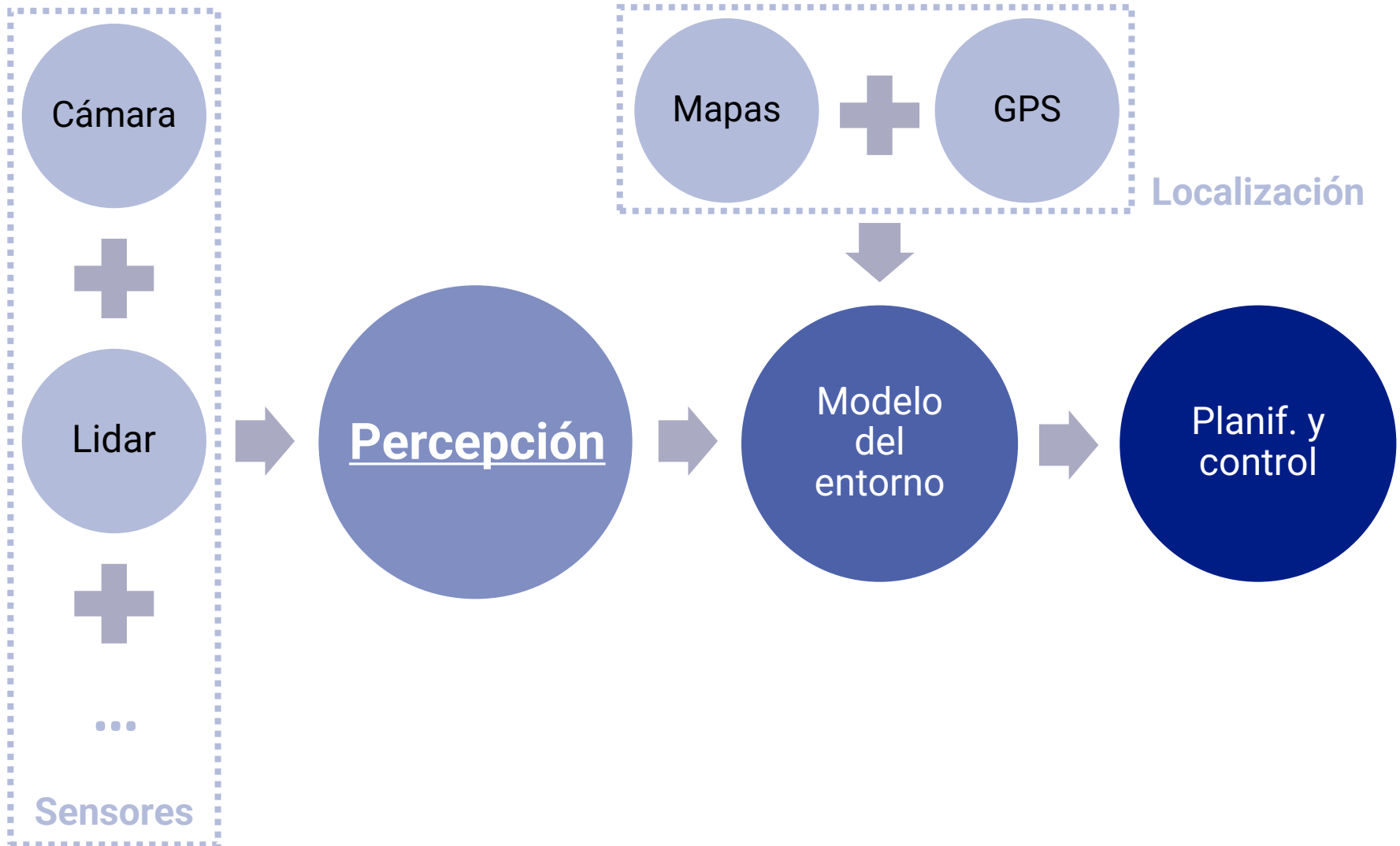
Visión estéreo



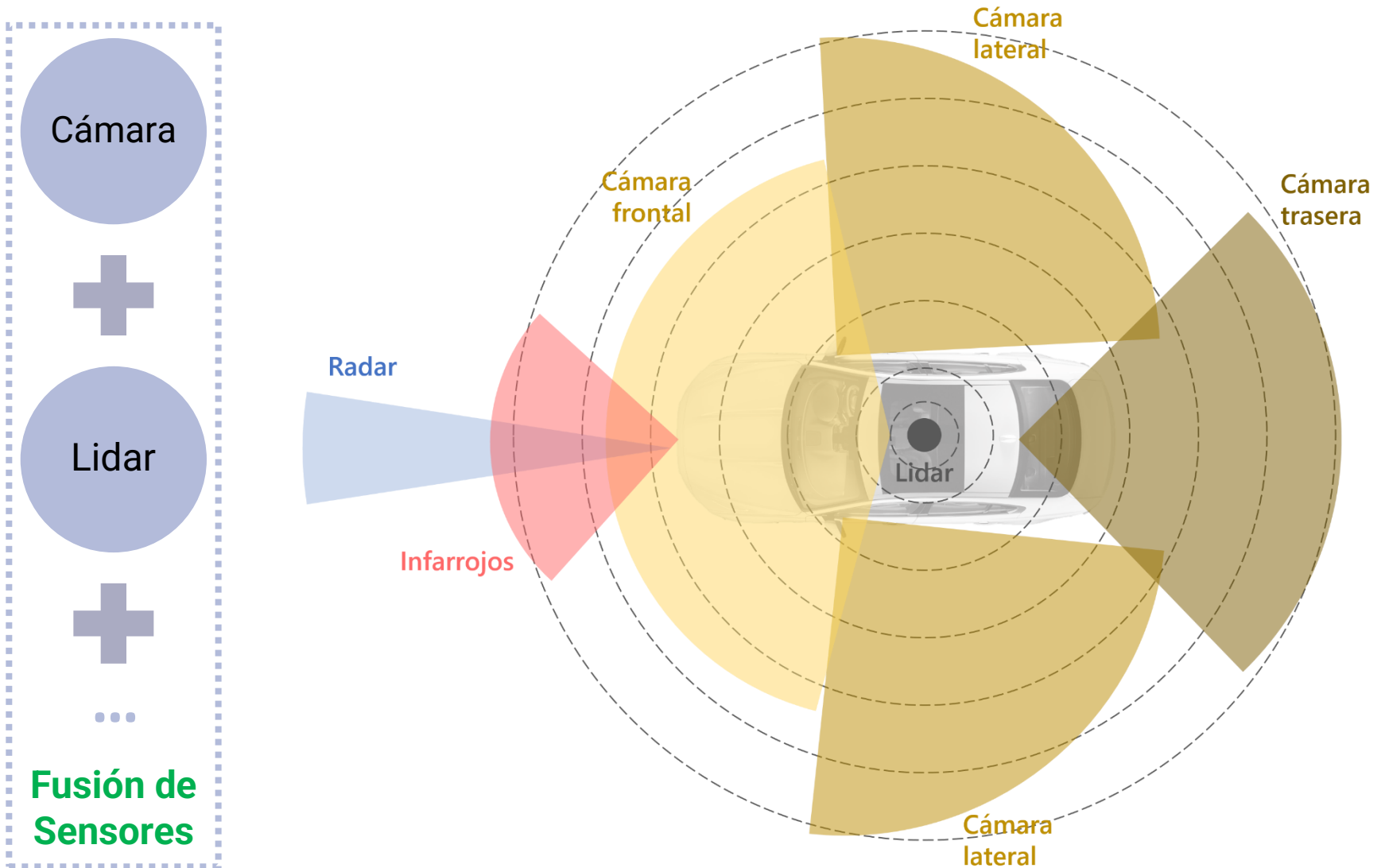
Conducción autónoma



Conducción autónoma



Fusión de datos

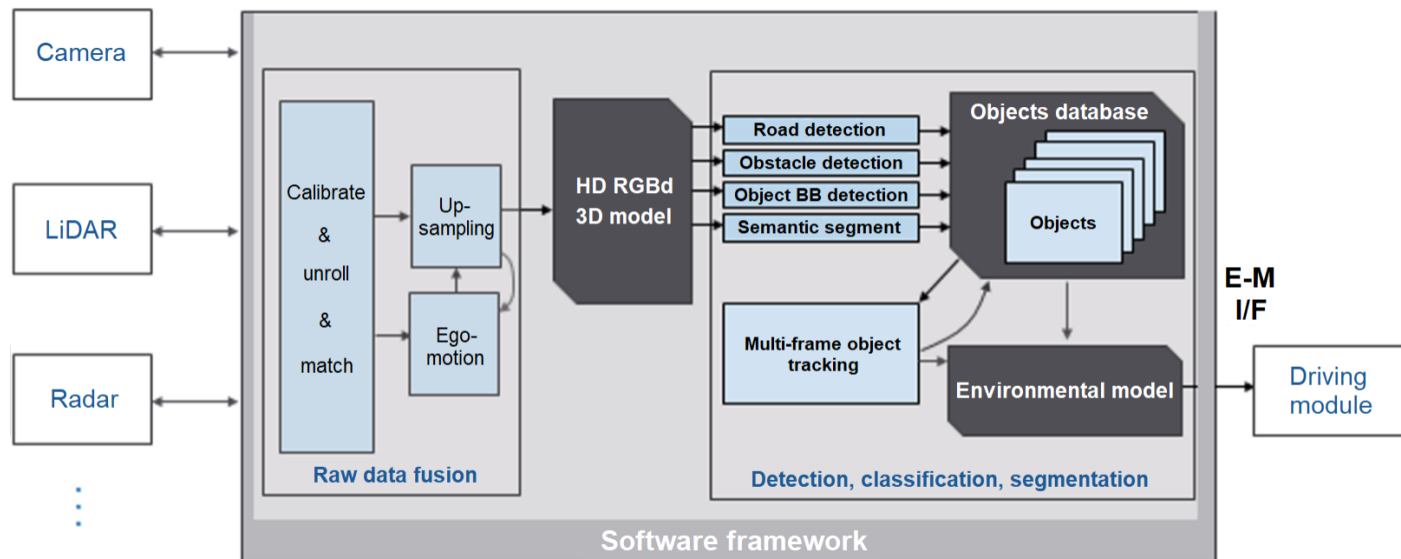


Clasificación de fusión de sensores: tres niveles de abstracción

- 1. Fusión de bajo nivel: Los datos en **bruto** son proporcionados como **entradas** al proceso de fusión, el cual obtiene datos más precisos (reducción del ruido) que las entradas individuales. También es conocido como signal-level fusión.
- 2. Fusión de nivel medio: Los atributos o características de una entidad (forma, textura, posición) son fusionados para obtener **características** que pueden ser usadas para otras tareas. Este tipo de fusión es conocido como fusión a nivel de atributos o **características**.
- 3. Fusión de alto nivel: También conocida como **fusión de decisión**. Obtiene decisiones o representaciones simbólicas como entrada y las combina para obtener una **decisión** más precisa. Las aplicaciones de fusión a este nivel suelen utilizar enfoques Bayesianos.

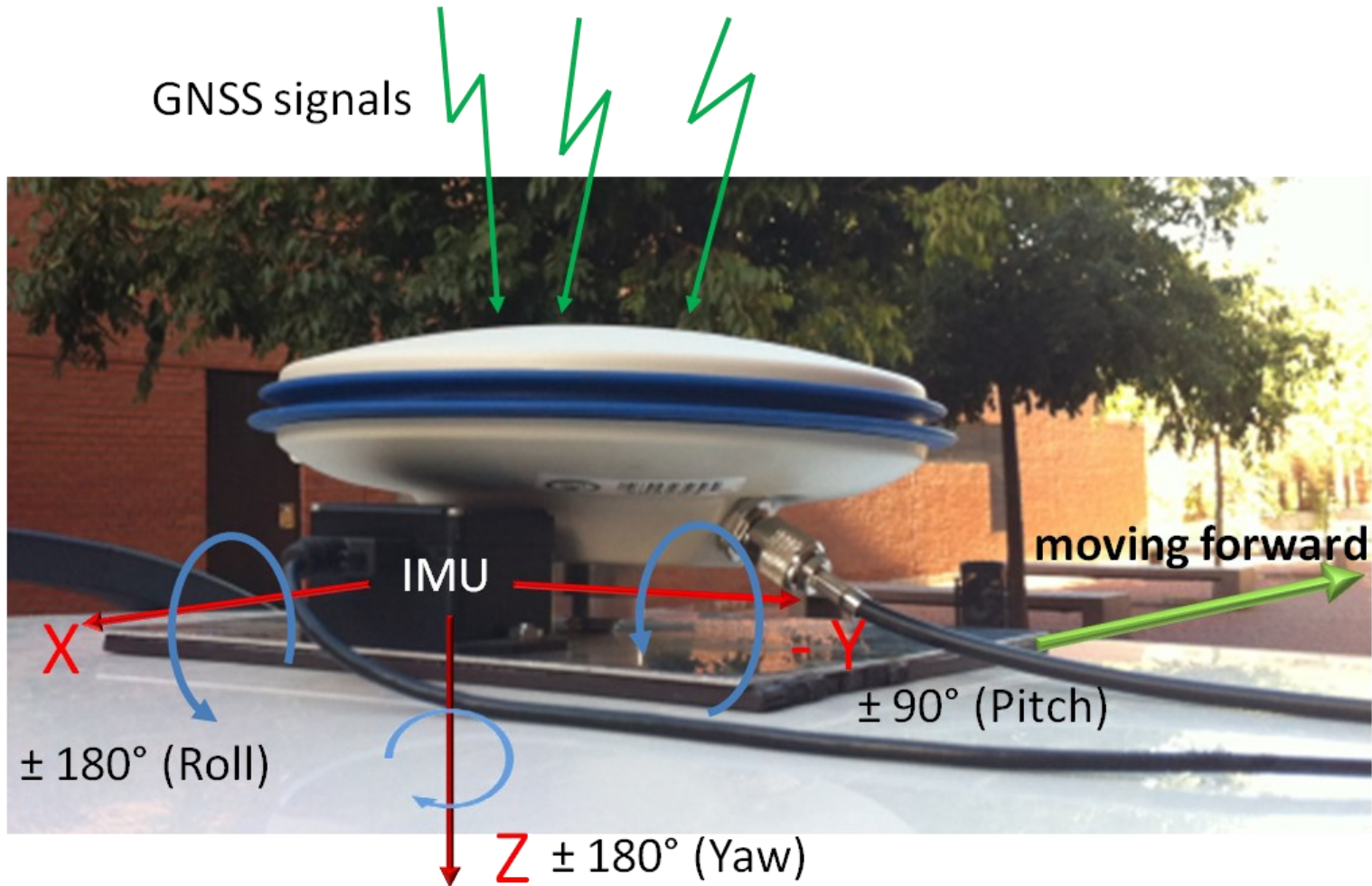
Ejemplo 1 de fusión de sensores de bajo nivel:

[LeddarVision](#) integra datos sin procesar de sensores para generar un modelo 3D RGBD. Los datos de los sensores se fusionan de manera inteligente, agregando información temporal (información de múltiples capturas), para obtener representaciones más precisas que con una sola medición (múltiples mediciones de un solo objeto permiten reducir el error de medición).

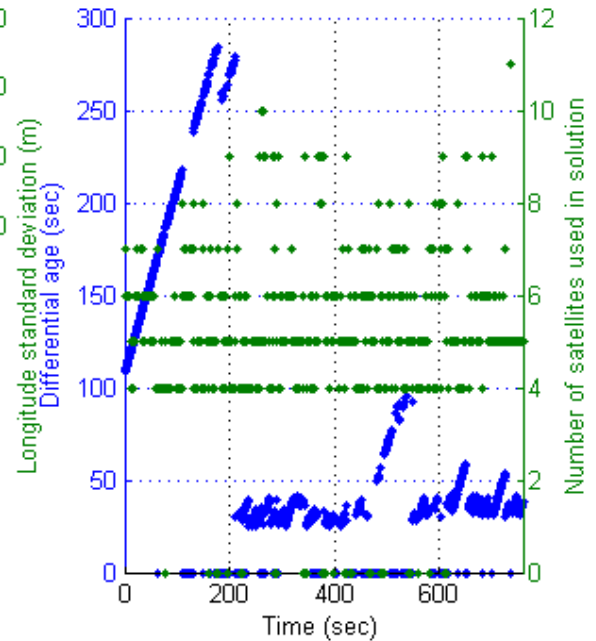
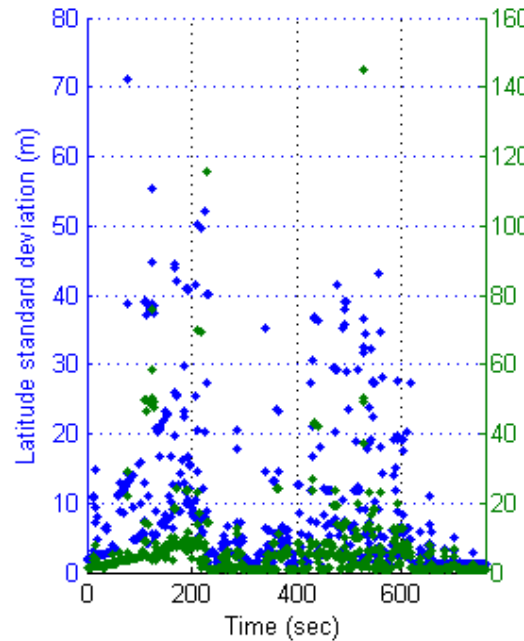
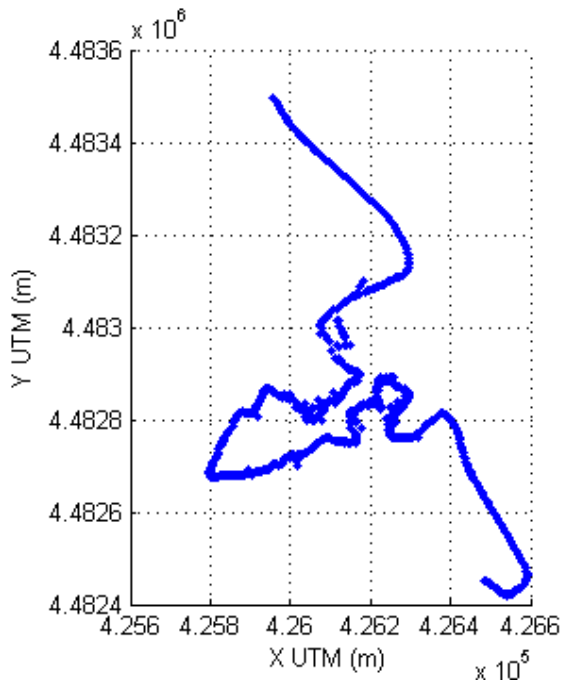


<https://leddartech.com/>

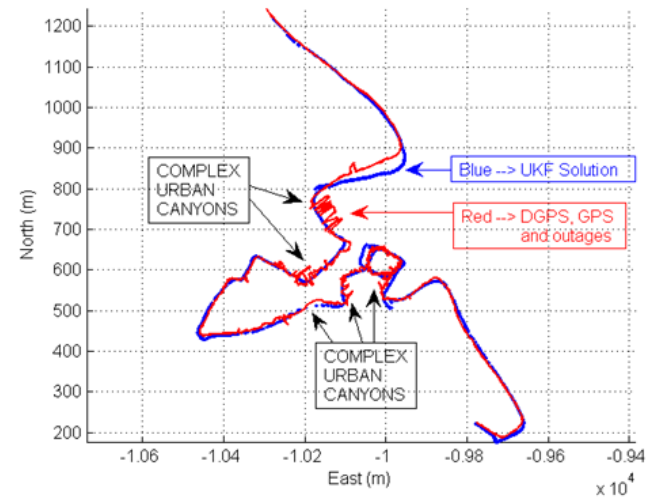
Ejemplo 2. fusión de bajo nivel: receptor GPS + sensor inercial



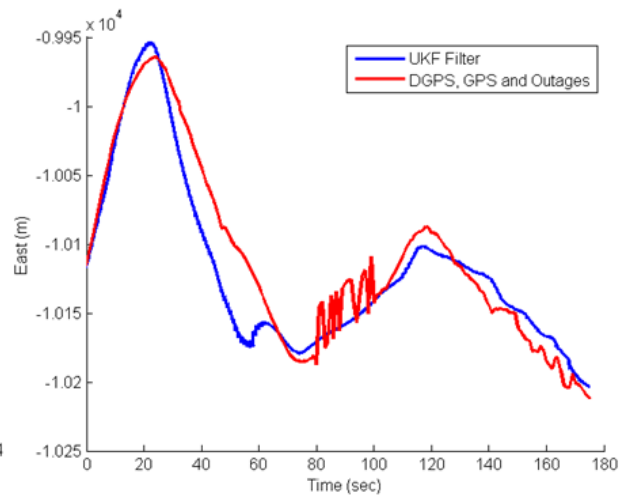
Ejemplo 2. fusión de bajo nivel: receptor GPS + sensor inercial



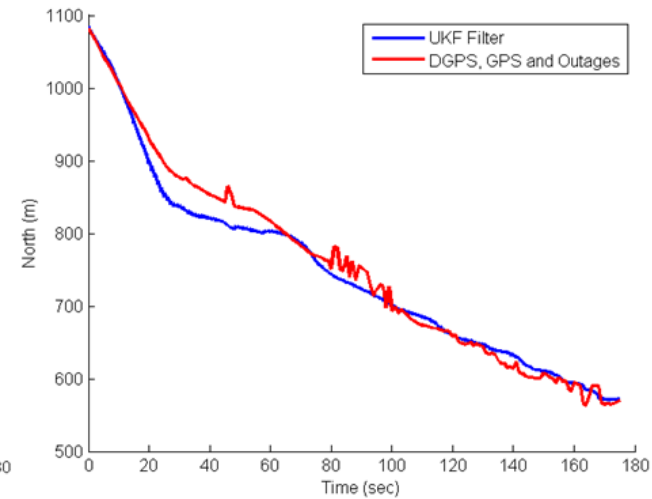
Ejemplo 2. fusión de bajo nivel: receptor GPS + sensor inercial



(a)



(b)

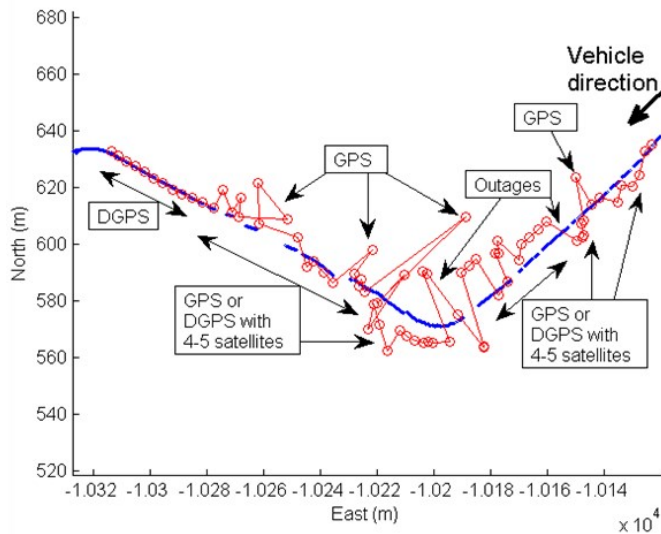


(c)

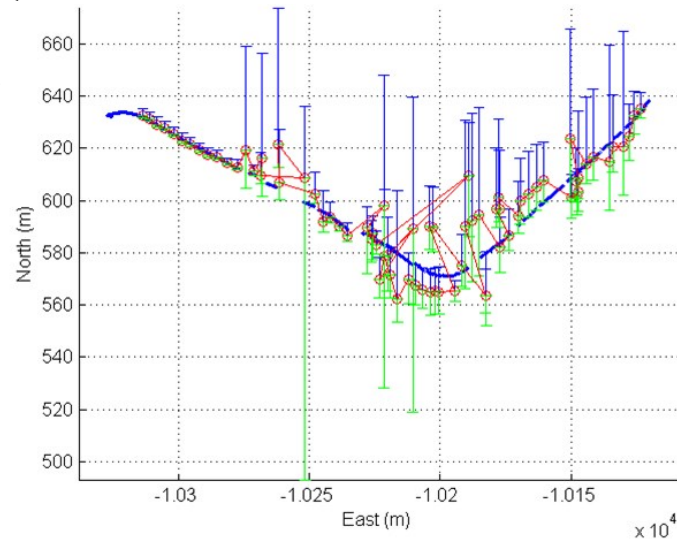
Ejemplo 2. fusión de bajo nivel: receptor GPS + sensor inercial



(a)



(b)

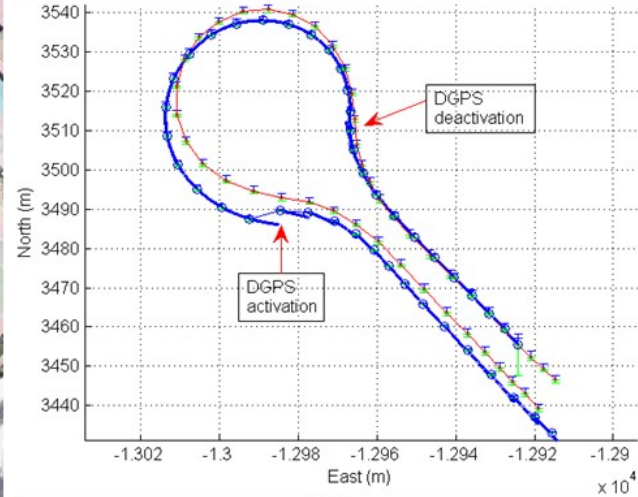


(c)

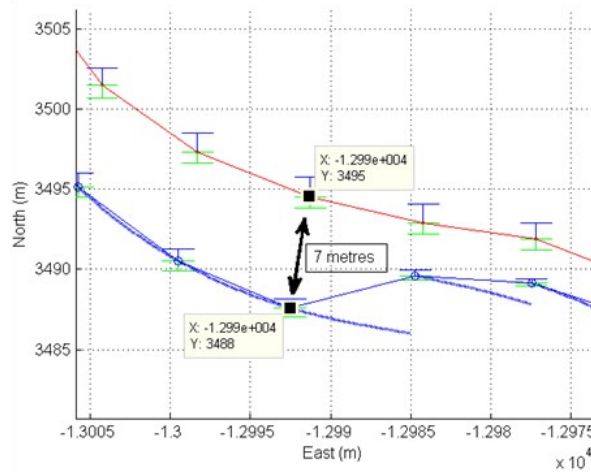
Ejemplo 2. fusión de bajo nivel: receptor GPS + sensor inercial



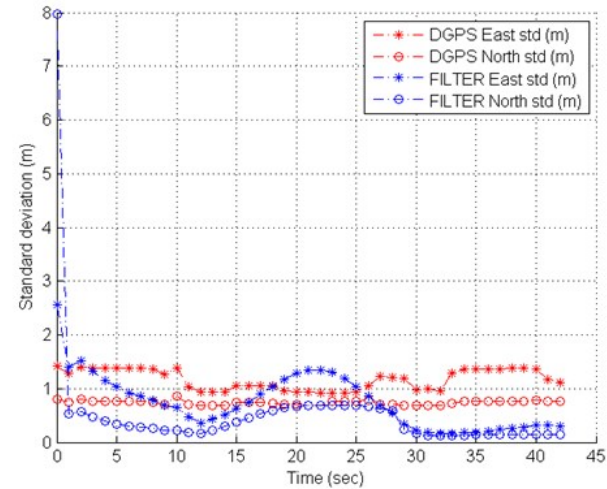
(a)



(b)



(c)

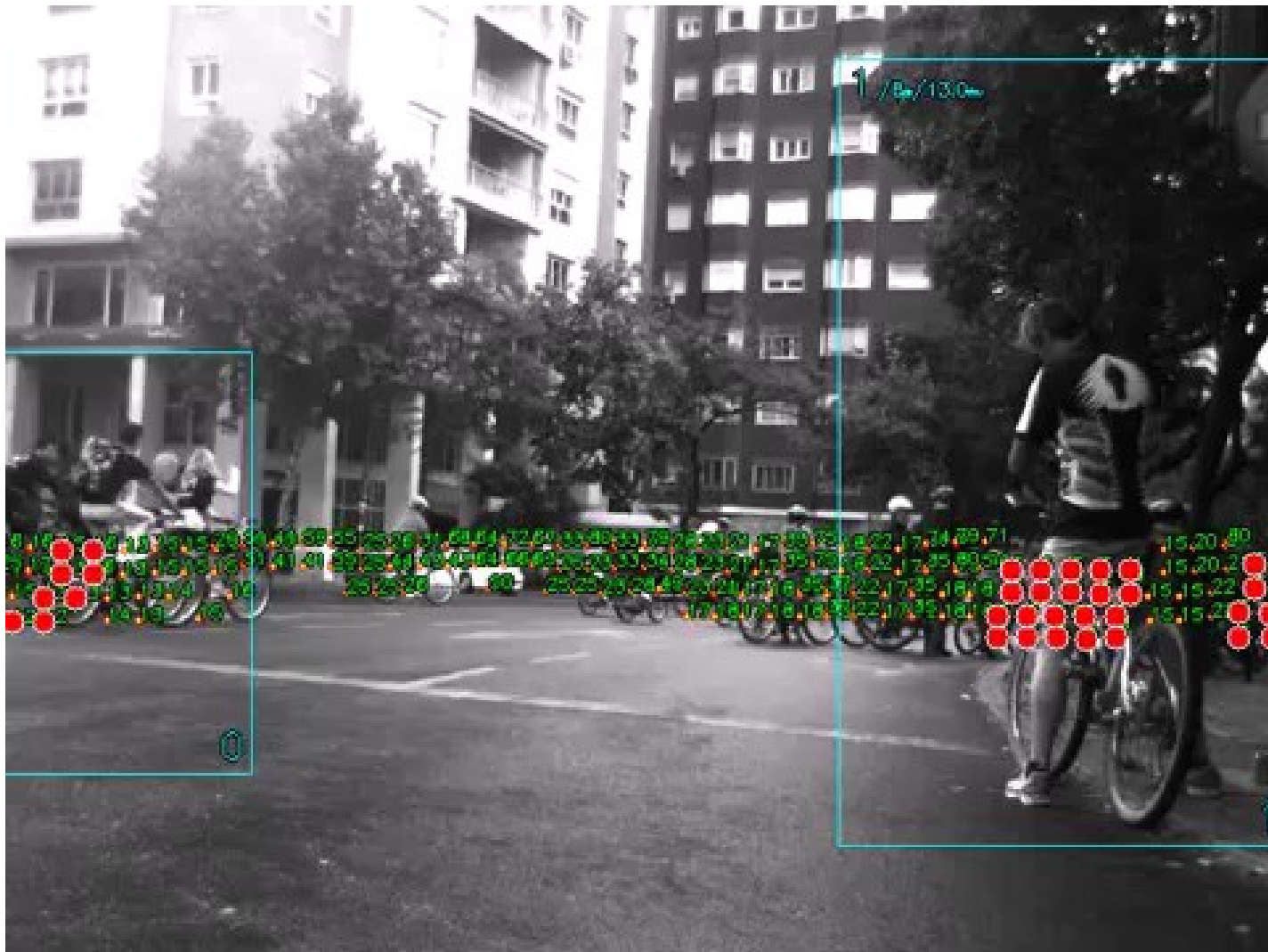


(d)

Ejemplo 3. fusión de bajo nivel: reconocimiento de bicicletas y motos



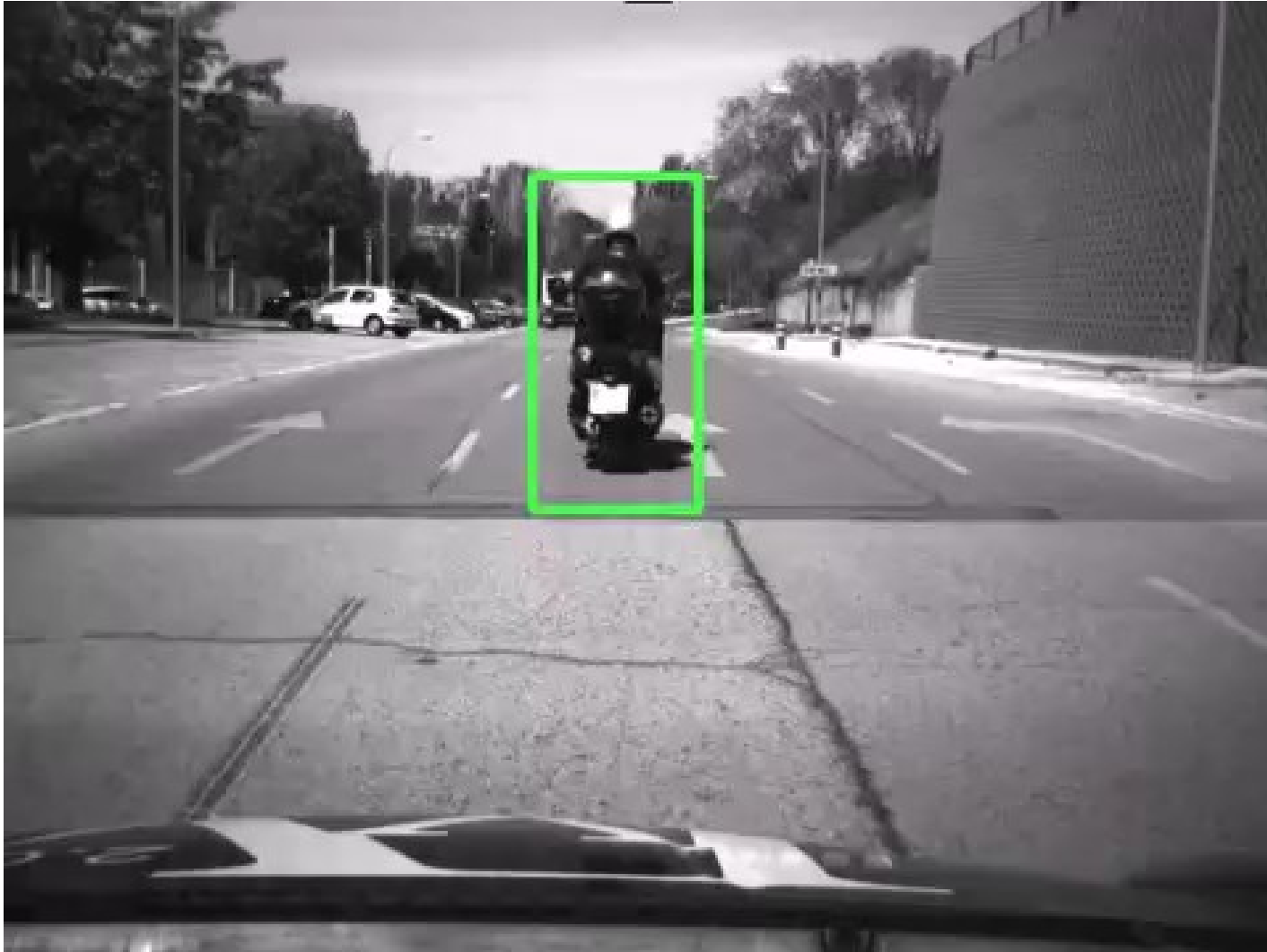
Ejemplo 3. fusión de bajo nivel: reconocimiento de bicicletas y motos



Ejemplo 3. fusión de bajo nivel: reconocimiento de bicicletas y motos



Ejemplo 3. fusión de bajo nivel: reconocimiento de bicicletas y motos



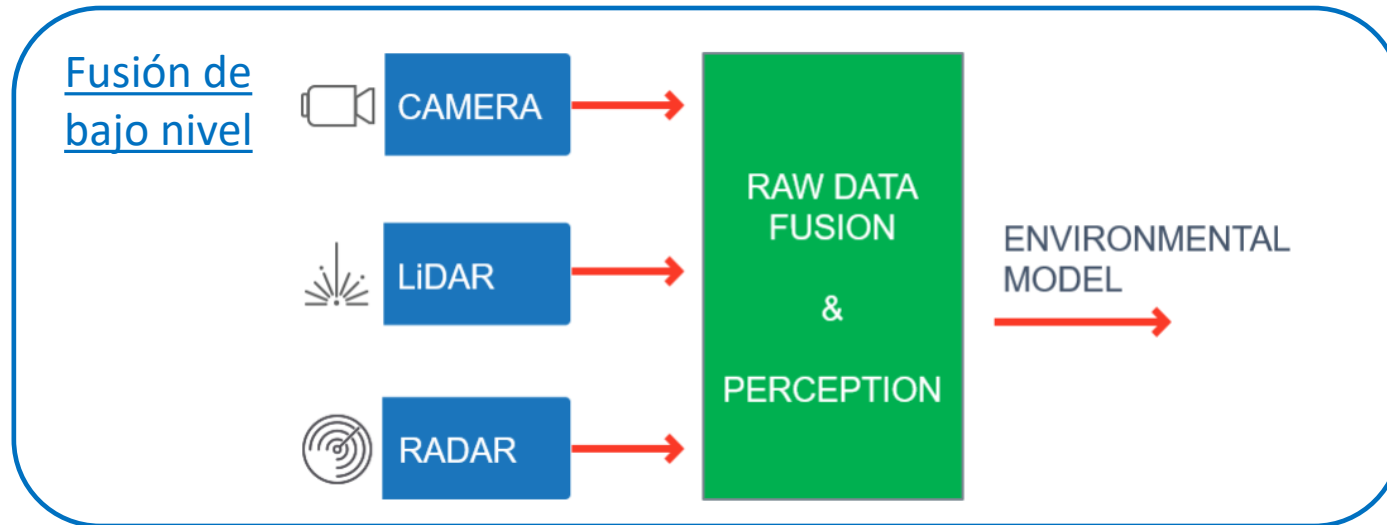
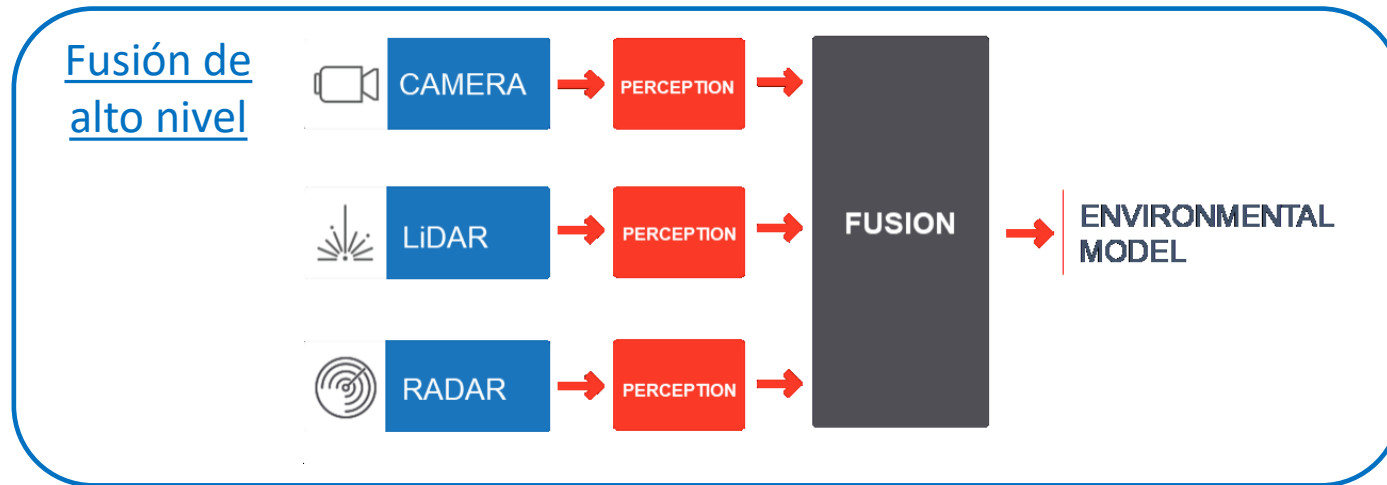
Ejemplo 3. fusión de bajo nivel: reconocimiento de bicicletas y motos



Ejemplo 3. fusión de bajo nivel: reconocimiento de bicicletas y motos



Fusión de bajo nivel contra fusión de alto nivel



2022 LeddarTech Inc. e-book // Sensor Fusion and Perception Technology Overview

Fusión de datos y calibración

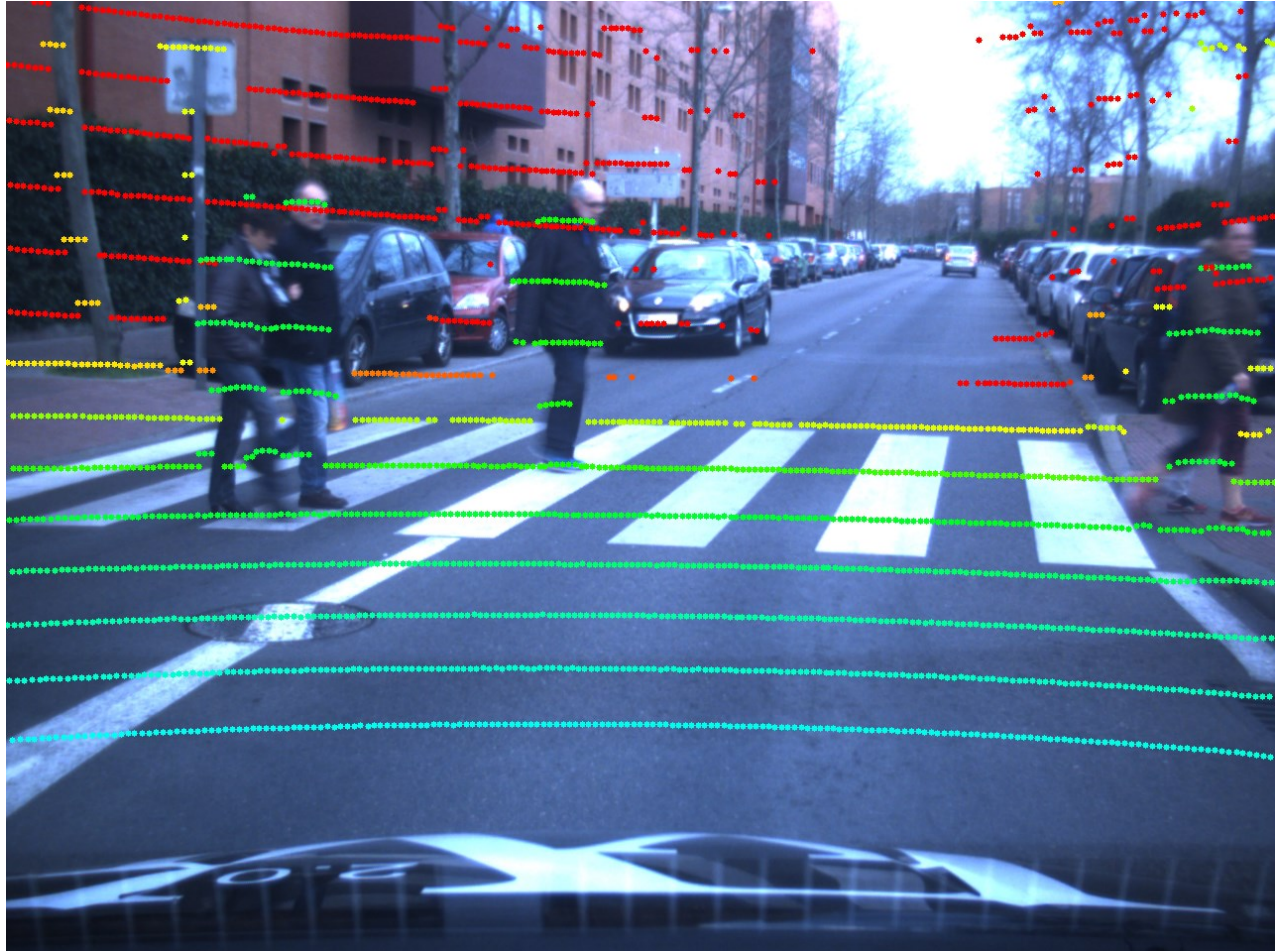
Cámara



Lidar

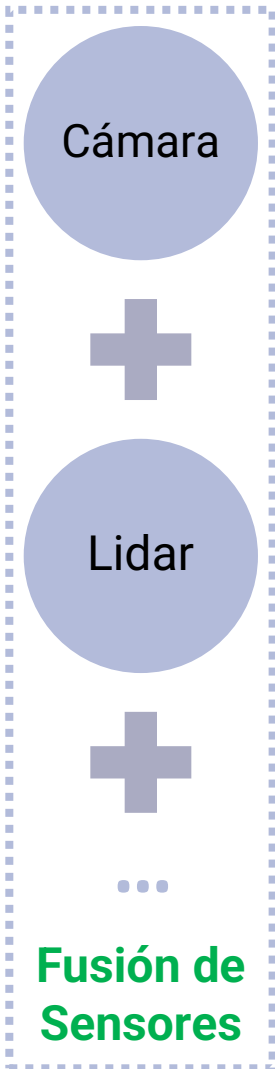


**Fusión de
Sensores**

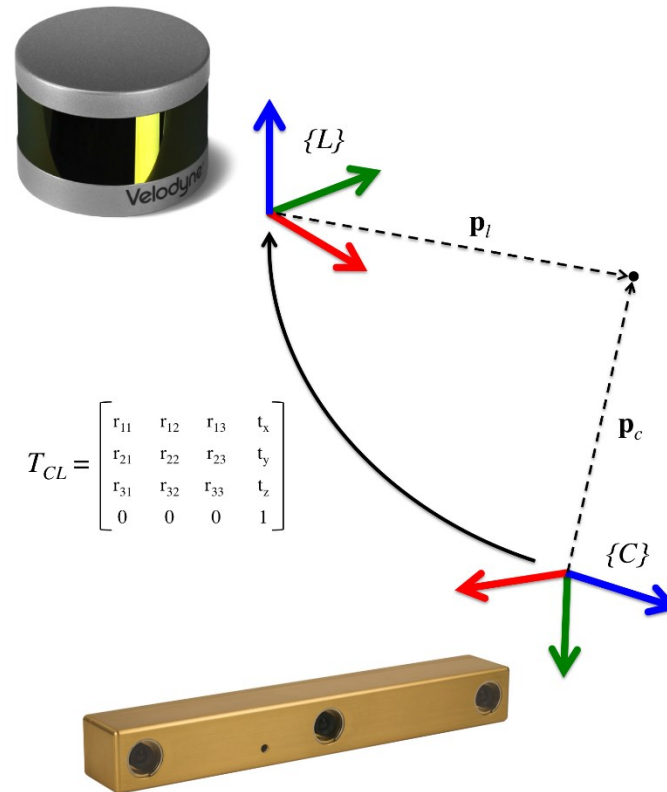


C. Guindel, J. Beltrán, D. Martín, and F. García, "Automatic Extrinsic Calibration for Lidar-Stereo Vehicle Sensor Setups," in IEEE International Conference on Intelligent Transportation Systems (ITSC), 2017, pp. 674–679.

Fusión de datos y calibración

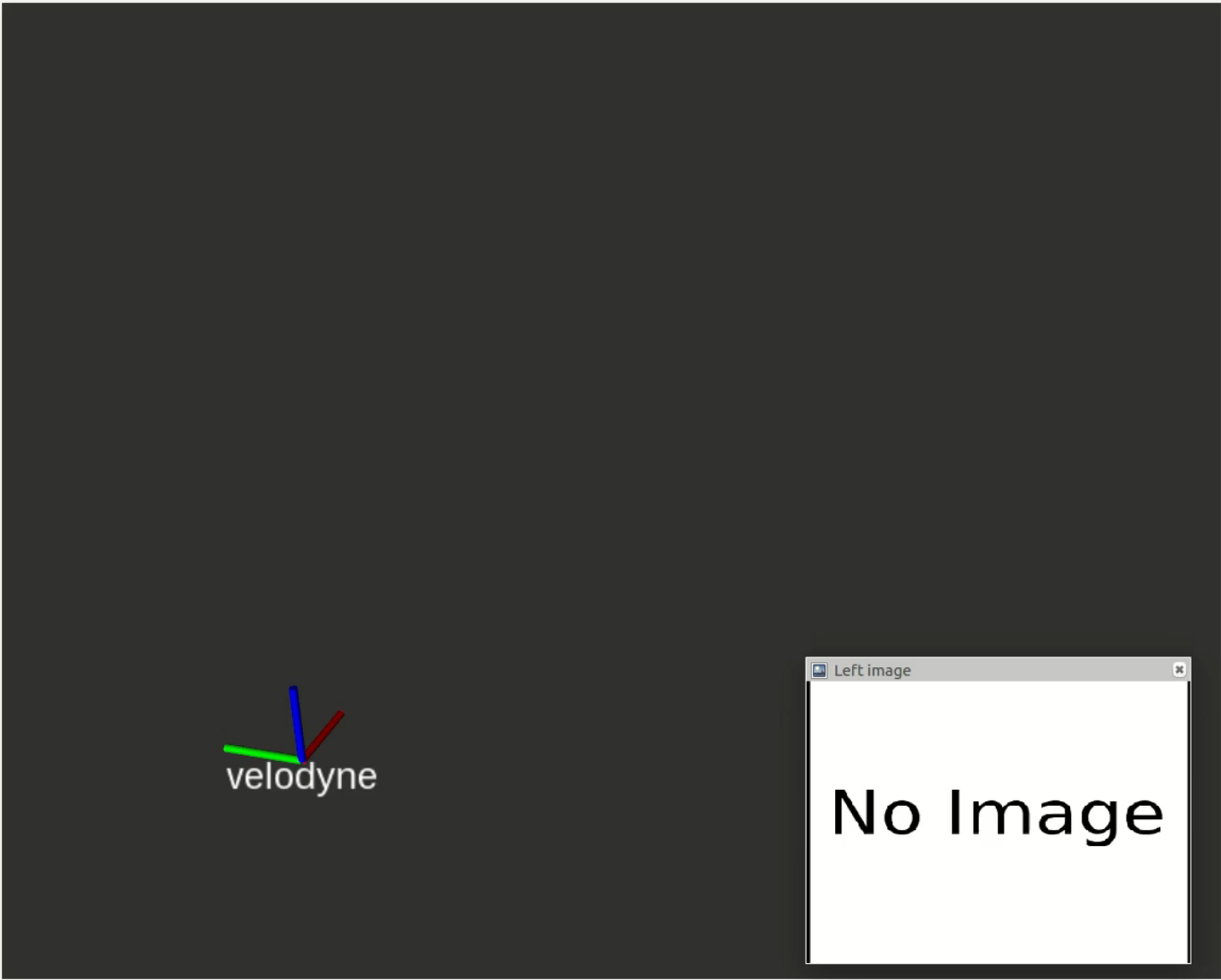


Parámetros extrínsecos

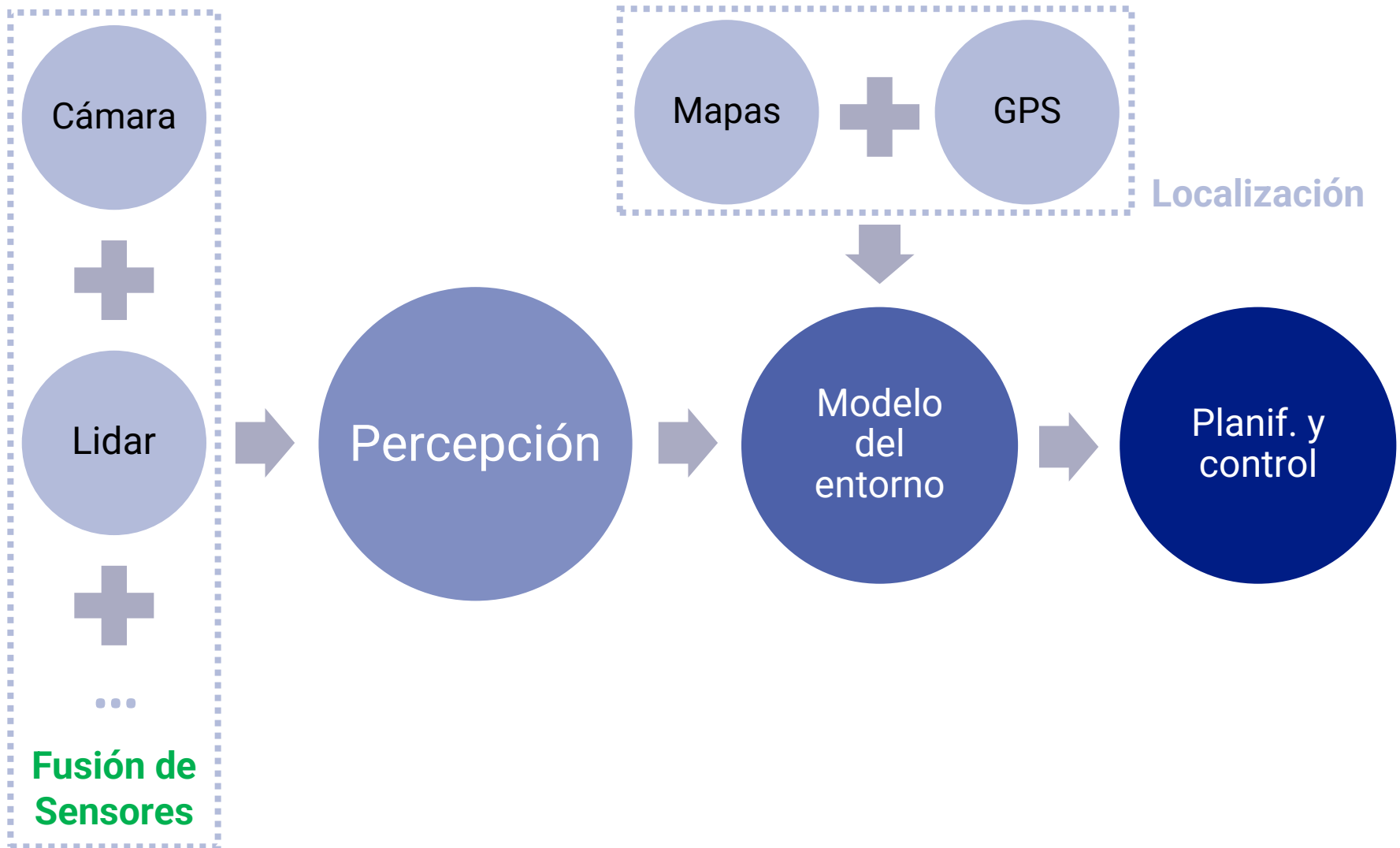


C. Guindel, J. Beltrán, D. Martín, and F. García, "Automatic Extrinsic Calibration for Lidar-Stereo Vehicle Sensor Setups," in IEEE International Conference on Intelligent Transportation Systems (ITSC), 2017, pp. 674–679.

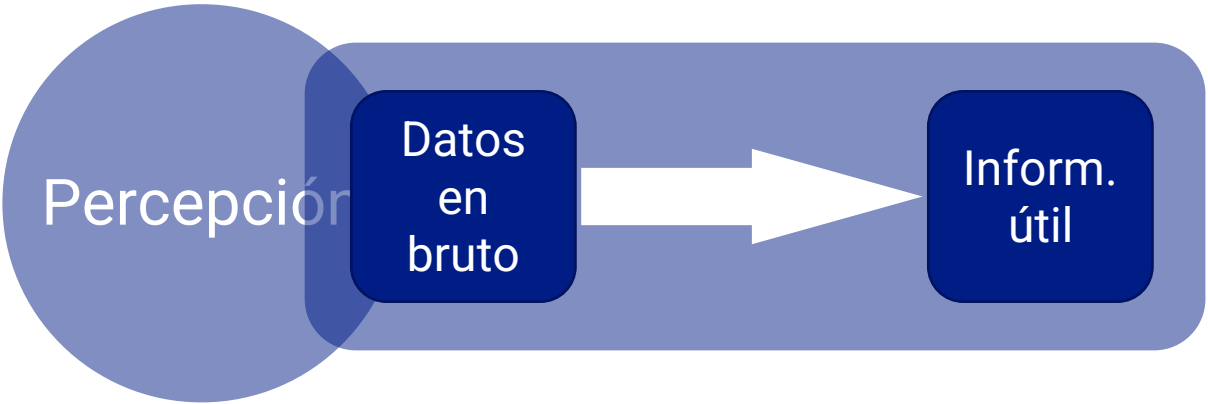
Calibración estéreo-lidar



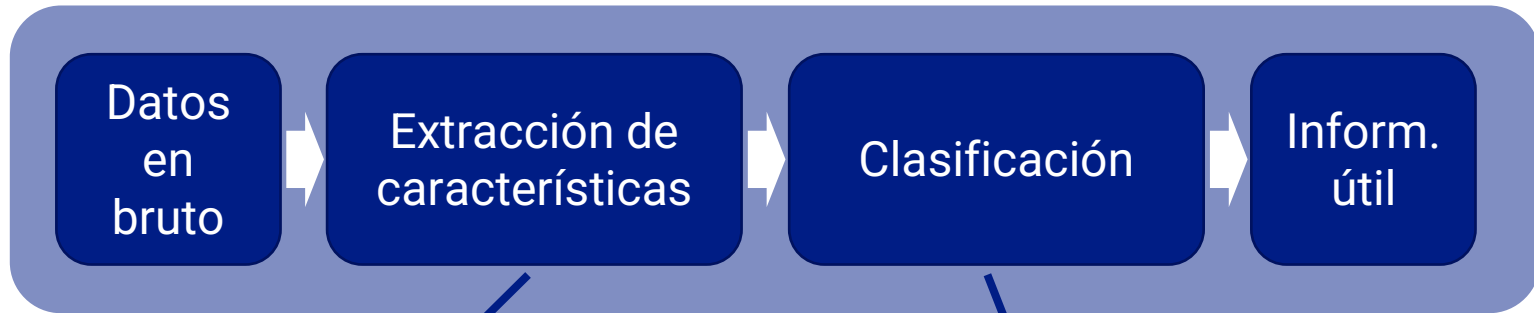
Conducción autónoma



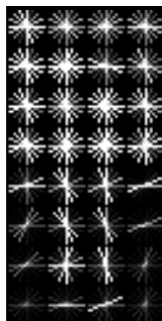
Conducción autónoma



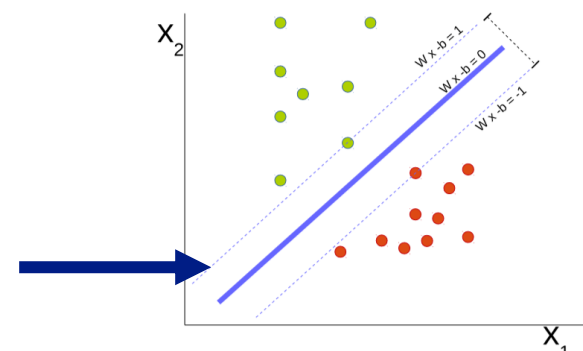
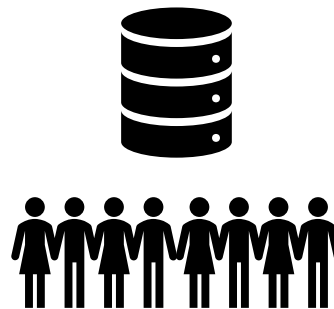
Enfoque clásico



Creadas "a mano"



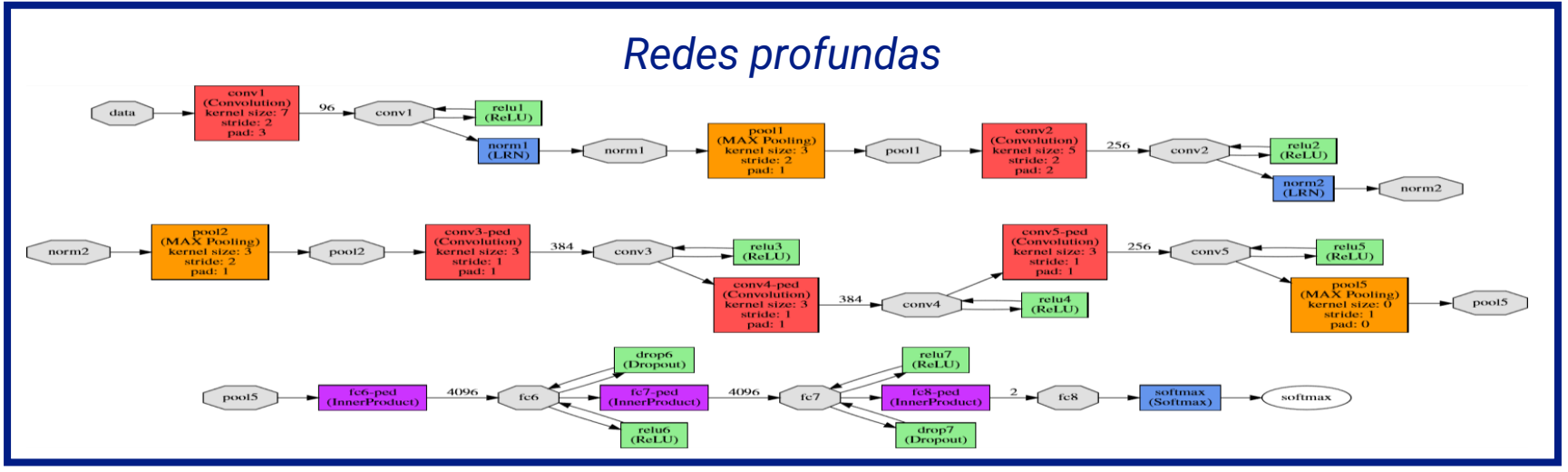
Machine learning: aprendizaje supervisado



Aprendizaje de características: *deep learning*

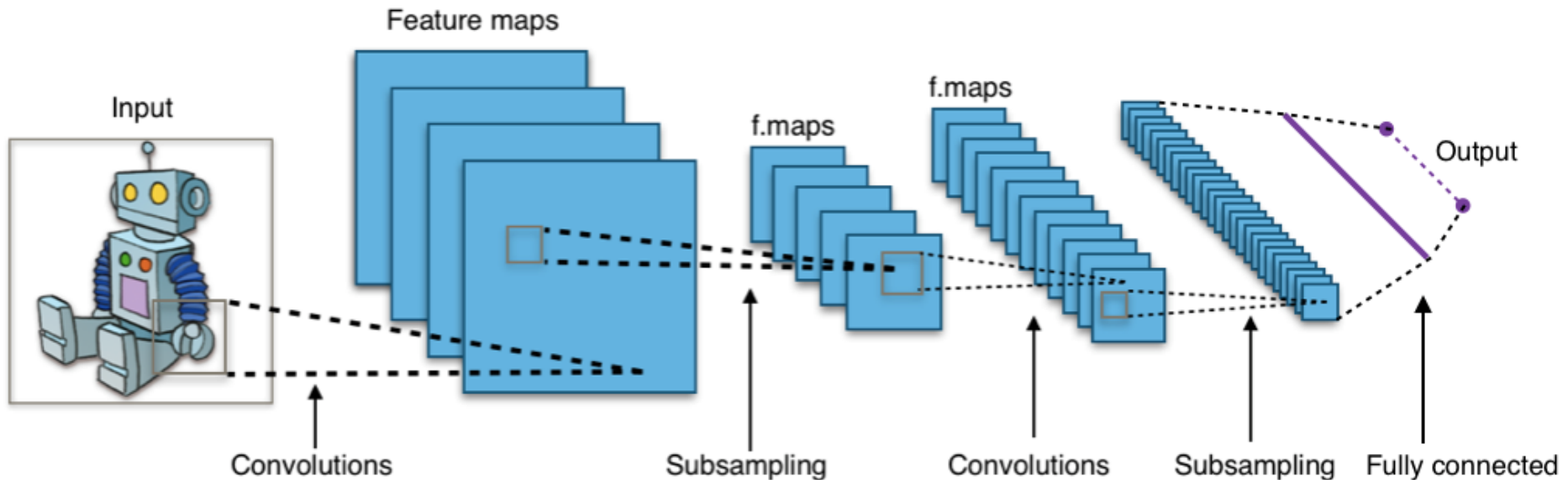


Redes profundas



Redes neuronales convolucionales (CNN)

Para datos con estructura bidimensional (p.ej. imágenes)

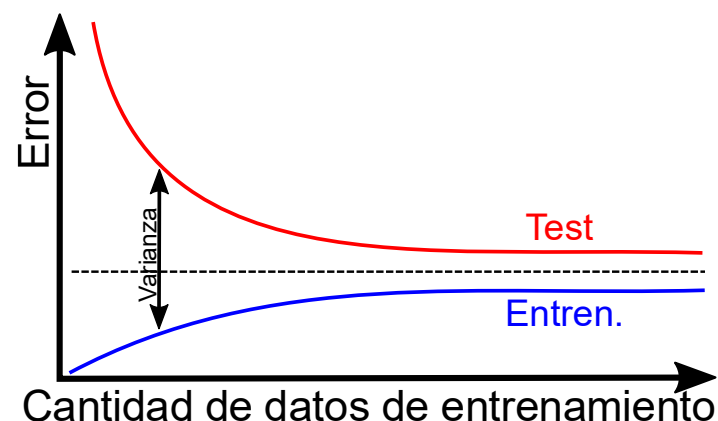
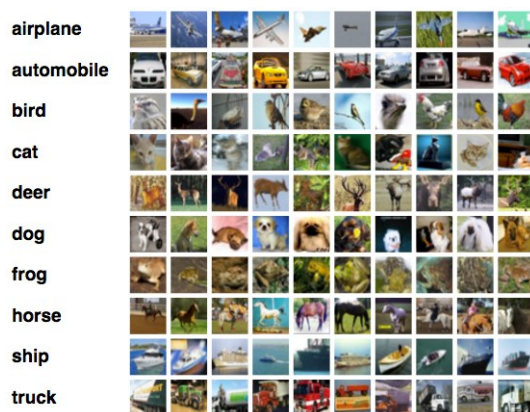


«Typical CNN» de Apex, disponible en Wikimedia (https://commons.wikimedia.org/wiki/File:Typical_cnn.png) bajo licencia CC BY-SA 4.0

- Sucesión de **convoluciones** con filtros aprendidos, **no linealidades** y **reducciones de muestreo**
- La etapa final está compuesta por redes neuronales clásicas (totalmente conectadas)

Aprendizaje

- Minimización de una fcn. de coste (error) en el entrenamiento
- Son necesarias muchas muestras etiquetadas



- Bases de datos públicas:
 - Generalistas: ImageNet, COCO,...
 - Conducción: KITTI, Cityscapes, BDD 100k, Apolloscape,...
 - Permiten el entrenamiento y la comparación de algoritmos



www.image-net.org



cocodataset.org



www.cvlibs.net/datasets/kitti/



www.cityscapes-dataset.com

Índice

1. Vehículos autónomos
 - Motivación y antecedentes
 - Taxonomía
2. Sistemas de percepción y fusión de sensores
 - Sensores
 - *Deep learning*
3. Aplicaciones y fusión de sensores
 - Clasificación, detección y localización de objetos
 - Segmentación semántica
 - Casos específicos de fusión de sensores

Aplicaciones: Laboratorio de Sistemas Inteligentes (LSI)



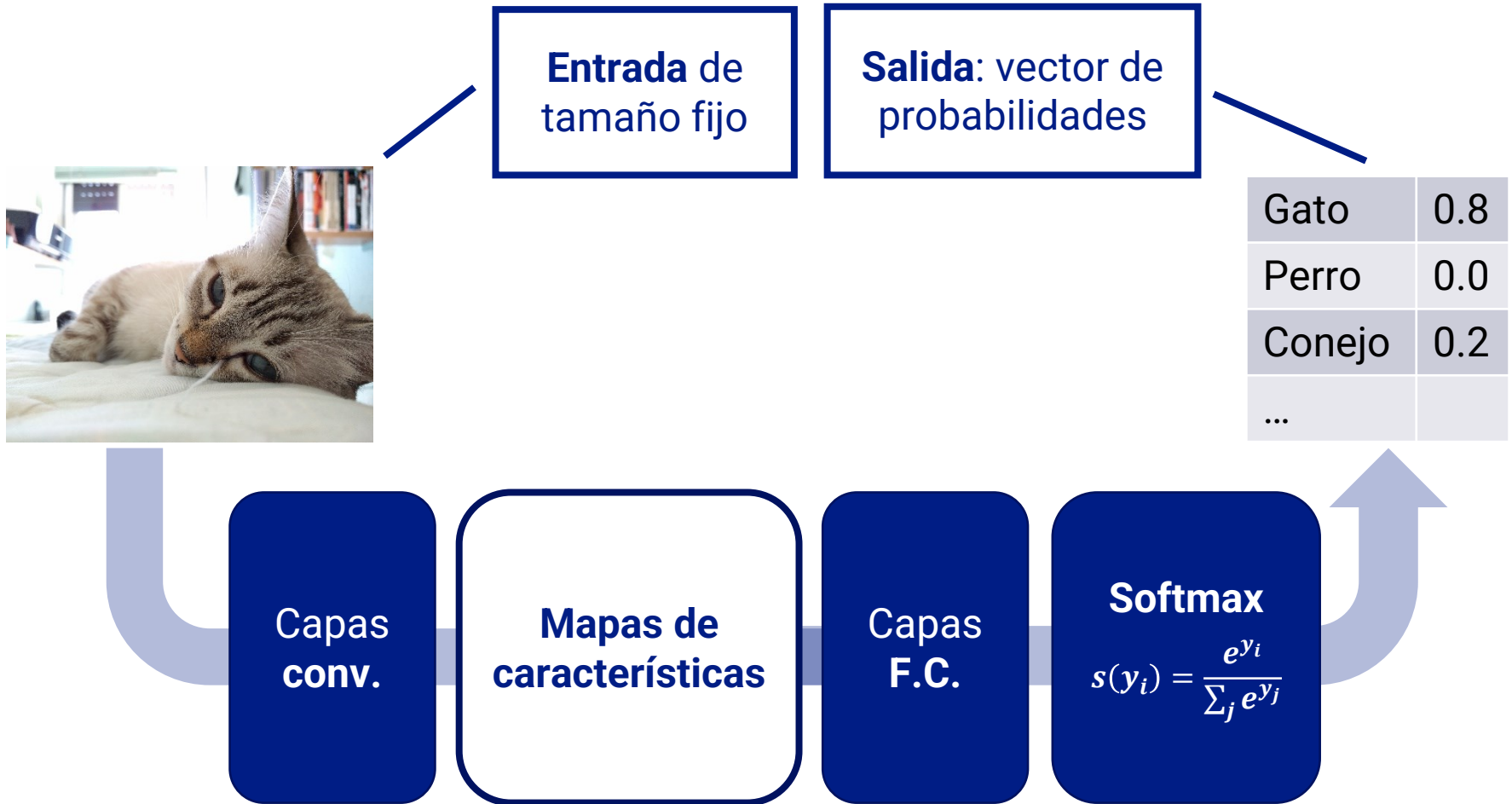
Lab. Sistemas
Inteligentes

uc3m

Universidad
Carlos III
de Madrid

Clasificación

- Aplicación directa de las CNN



Clasificación: señales de tráfico

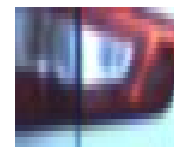
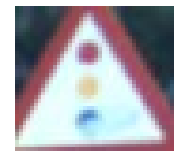
Segmentación

- Basada en el color
- Genera propuestas susceptibles de ser señal



Clasificación

- Aplicación de una CNN por cada propuesta
- Clase “fondo” para errores de segmentación



“Fondo”

Precisión en la clasificación: **99,05%**

Clasificación: estado del semáforo



- Mapas digitales
- Posicionamiento usando visión

Regiones precisas



(a) Bosch



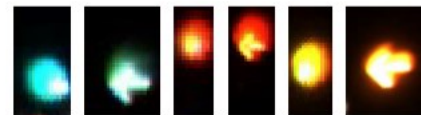
(b) Cityscapes



(c) LaRA



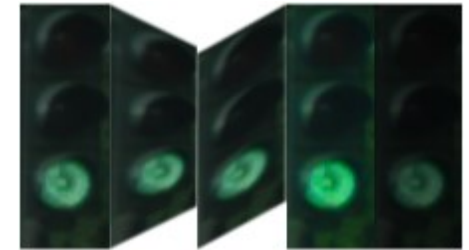
(d) LISA-Day



(e) LISA-Night



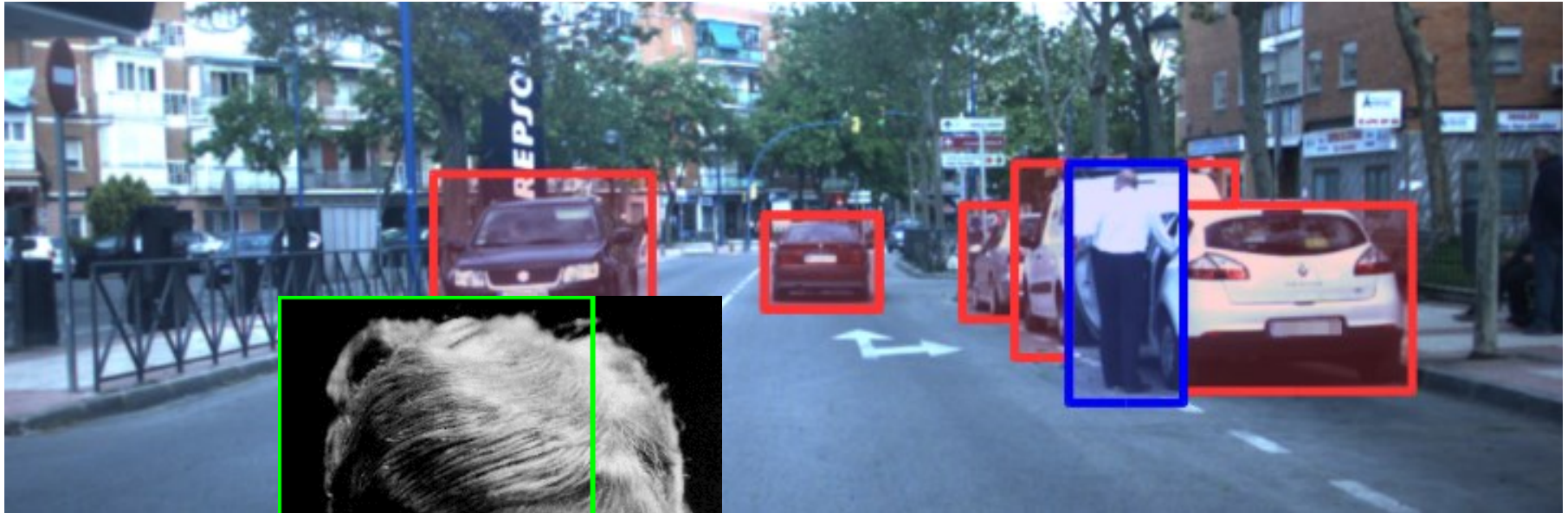
(f) WPI



- Combinación de varias bases de datos + aumentado de datos
- CNN ligera (7 ms)
- F1-score: **90,3%** (vs. 67,6% Random Trees)

C. Fernández, C. Guindel, N.-O. Salscheider, and C. Stiller, "A Deep Analysis of the Existing Datasets for Traffic Light State Recognition," in Proc. IEEE International Conference on Intelligent Transportation Systems (ITSC), 2018, pp. 248–254

Detección



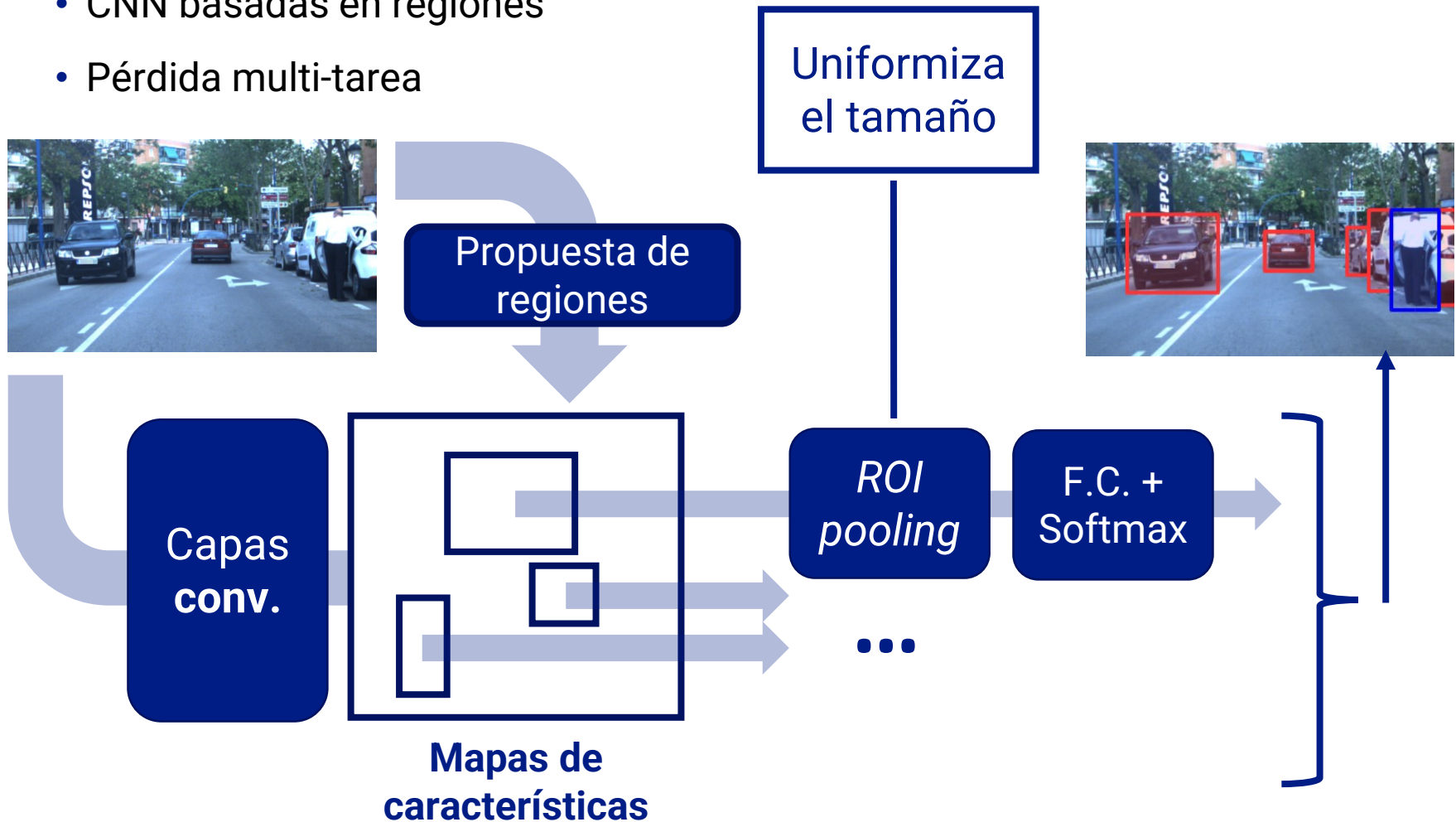
- En las imágenes se detectan los elementos
- Salida de la información relevante y su posición en la imagen



Fuente: www.pyimagesearch.com, Adrian Rosebrock

Detección: paradigma (Fast) R-CNN

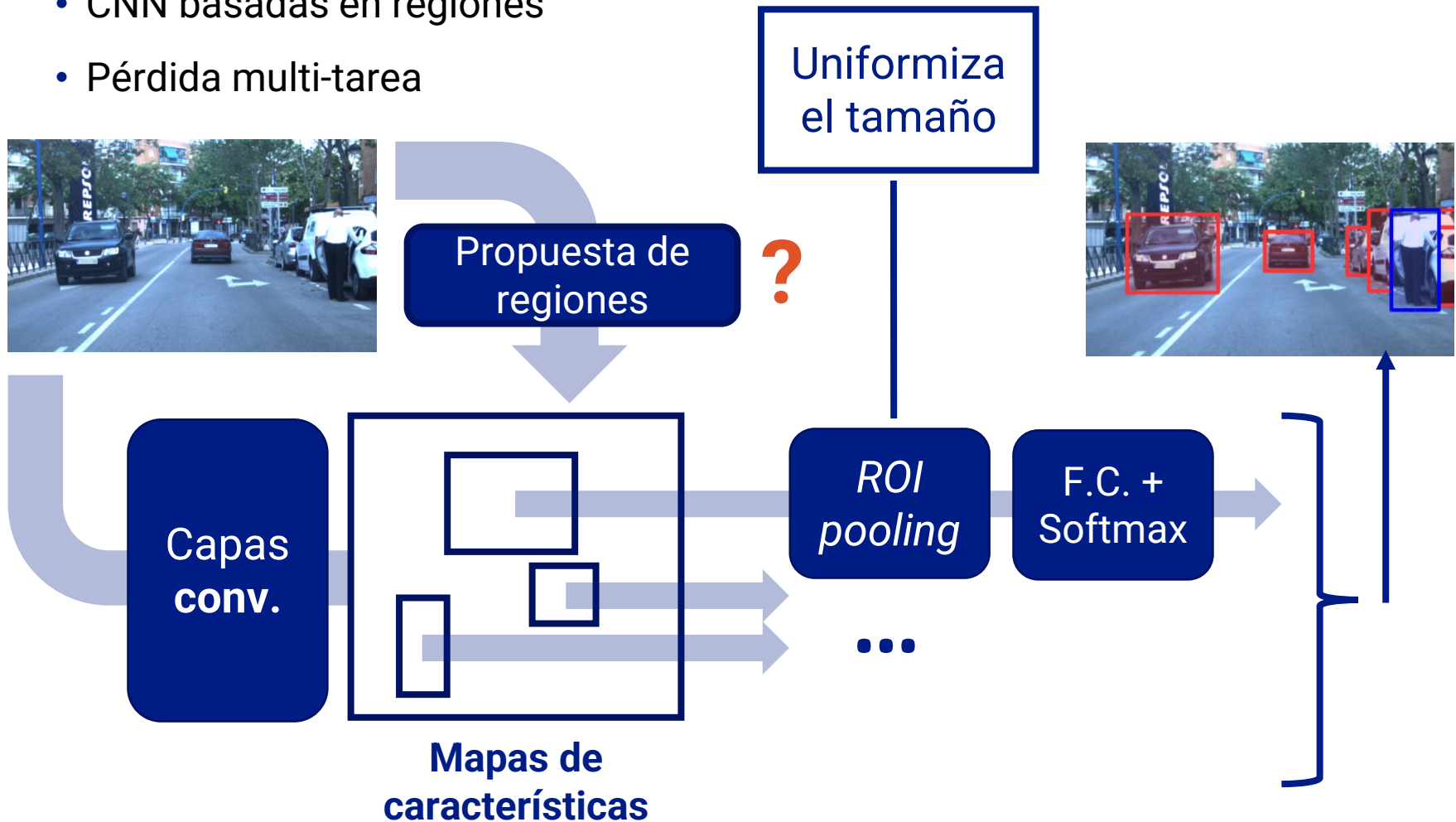
- CNN basadas en regiones
- Pérdida multi-tarea



R. Girshick, "Fast R-CNN," in Proc. IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440–1448.

Detección: paradigma (Fast) R-CNN

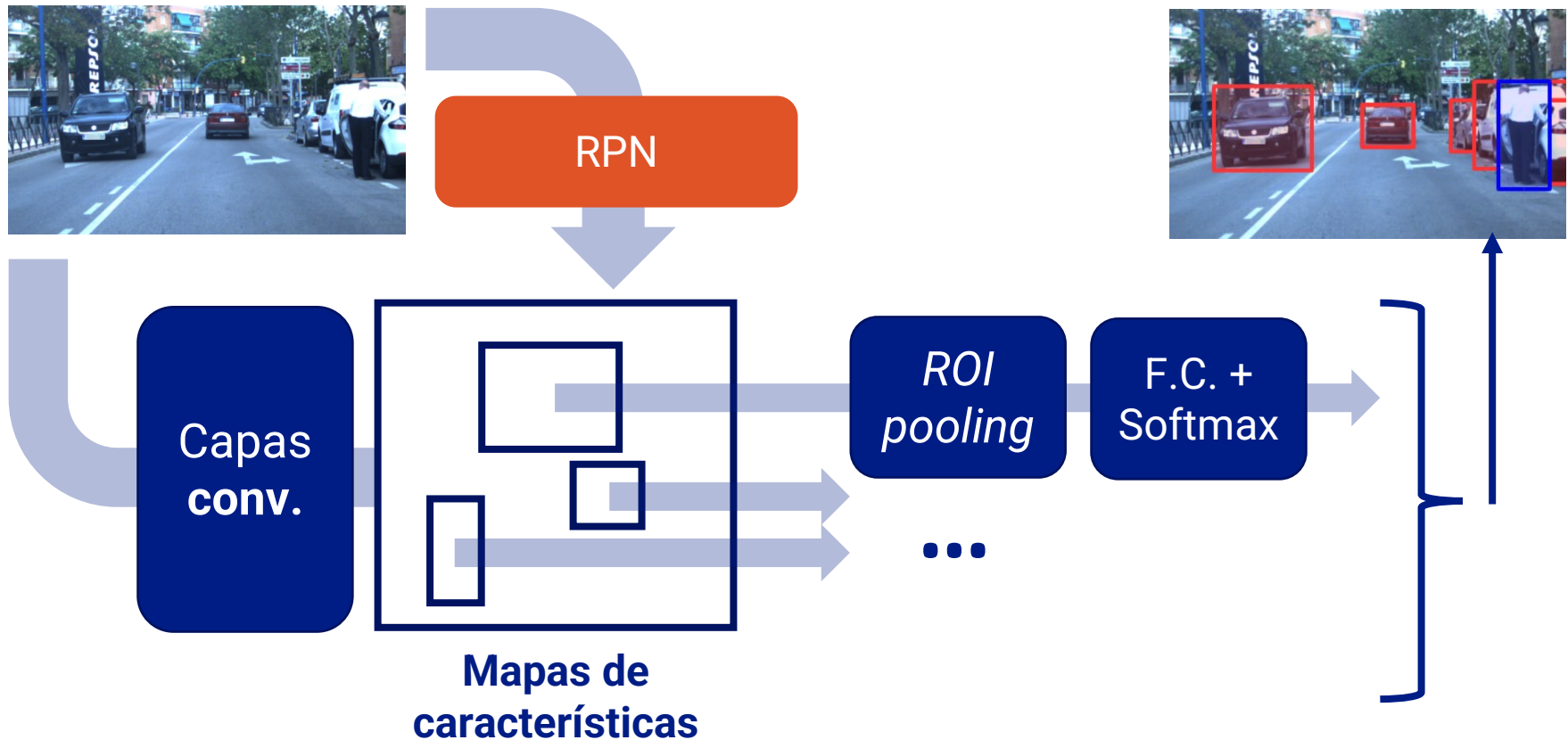
- CNN basadas en regiones
- Pérdida multi-tarea



R. Girshick, "Fast R-CNN," in Proc. IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440–1448.

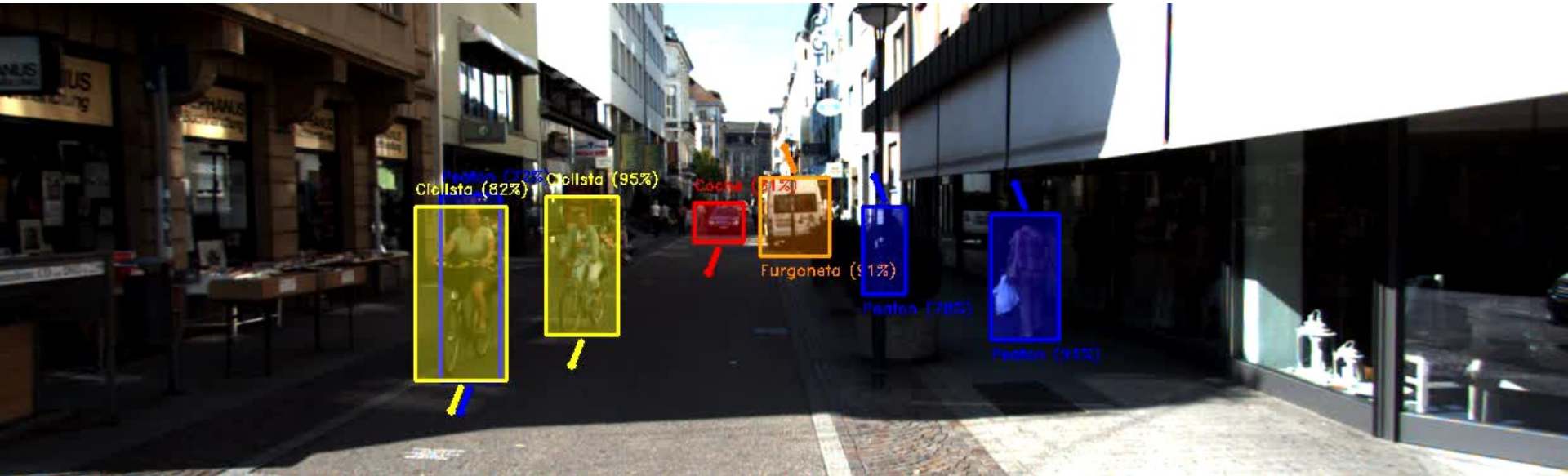
Detección: paradigma **Faster R-CNN**

- “Red de propuesta de regiones” (**RPN**)
- Pequeña CNN que clasifica “cajas ancla” en objeto/no objeto



S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” in Advances in Neural Information Processing Systems (NIPS), 2015, pp. 91–99

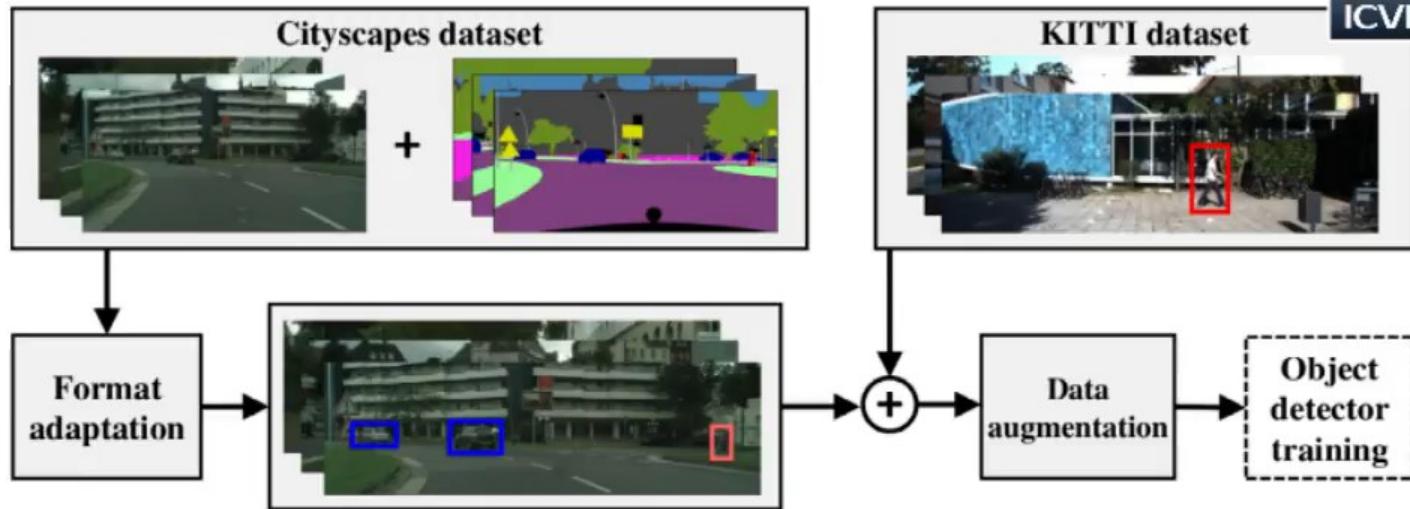
Entendimiento del entorno urbano y de carretera



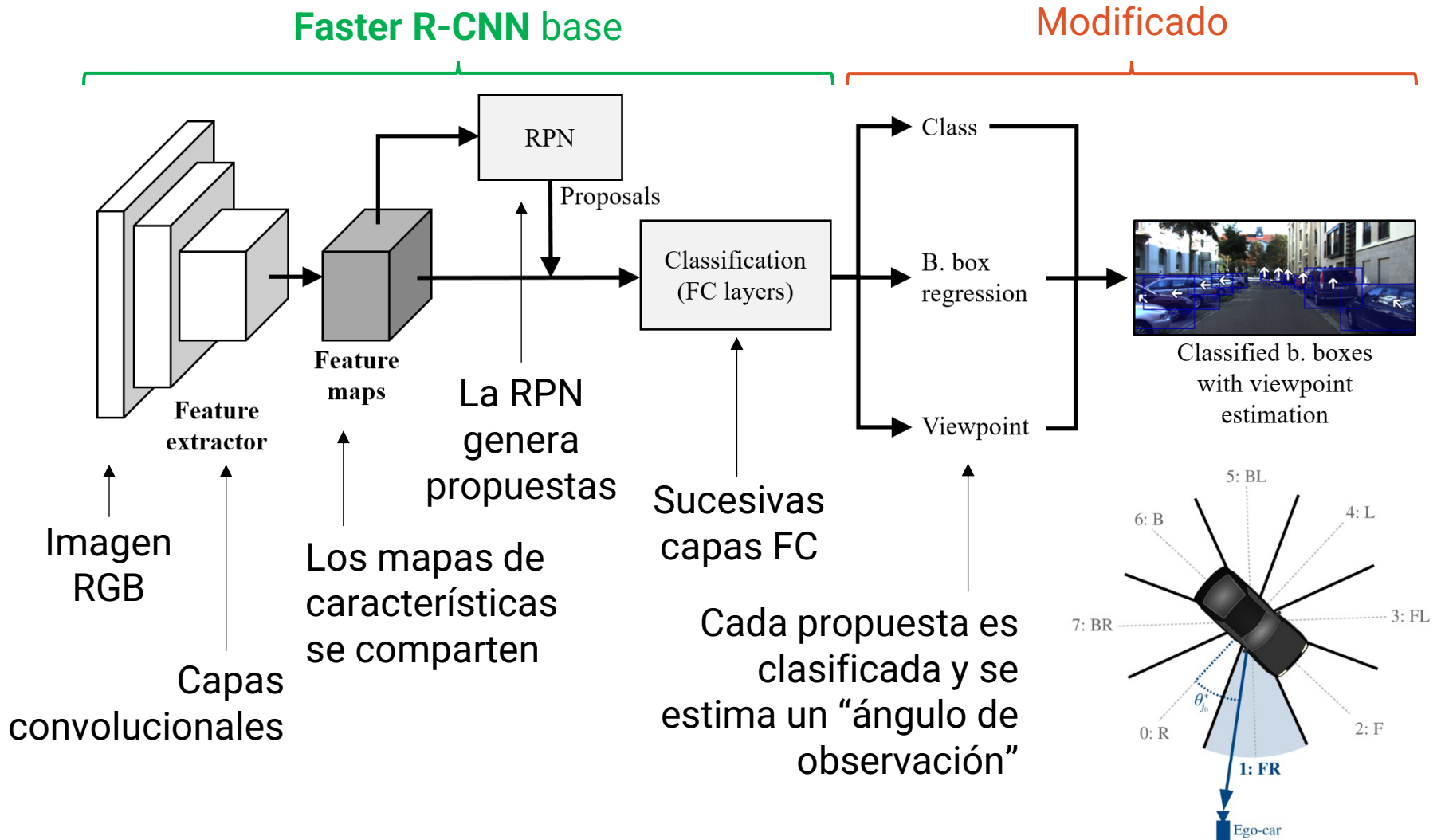
Entendimiento del entorno urbano y de carretera

Analysis of the Influence of Training Data on Object Detectors aimed at Road Environments

C. Guindel, D. Martín and J. M. Armingol



Detección y punto de vista (orientación)



C. Guindel, D. Martín, and J. M. Armingol, "Fast Joint Object Detection and Viewpoint Estimation for Traffic Scene Understanding," IEEE Intell. Transp. Syst. Mag., vol. 10, no. 4, pp. 74–86, 2018.

Detección y punto de vista (orientación)



(a)



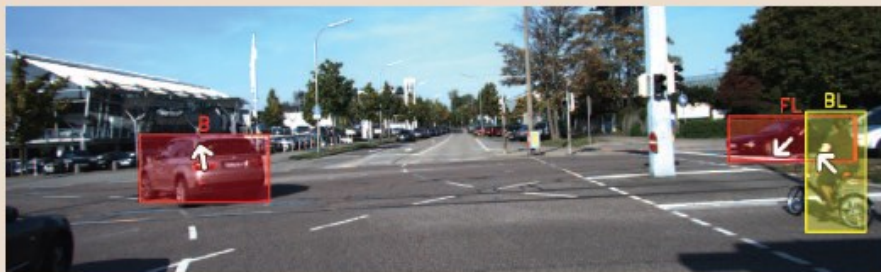
(b)



(c)



(d)

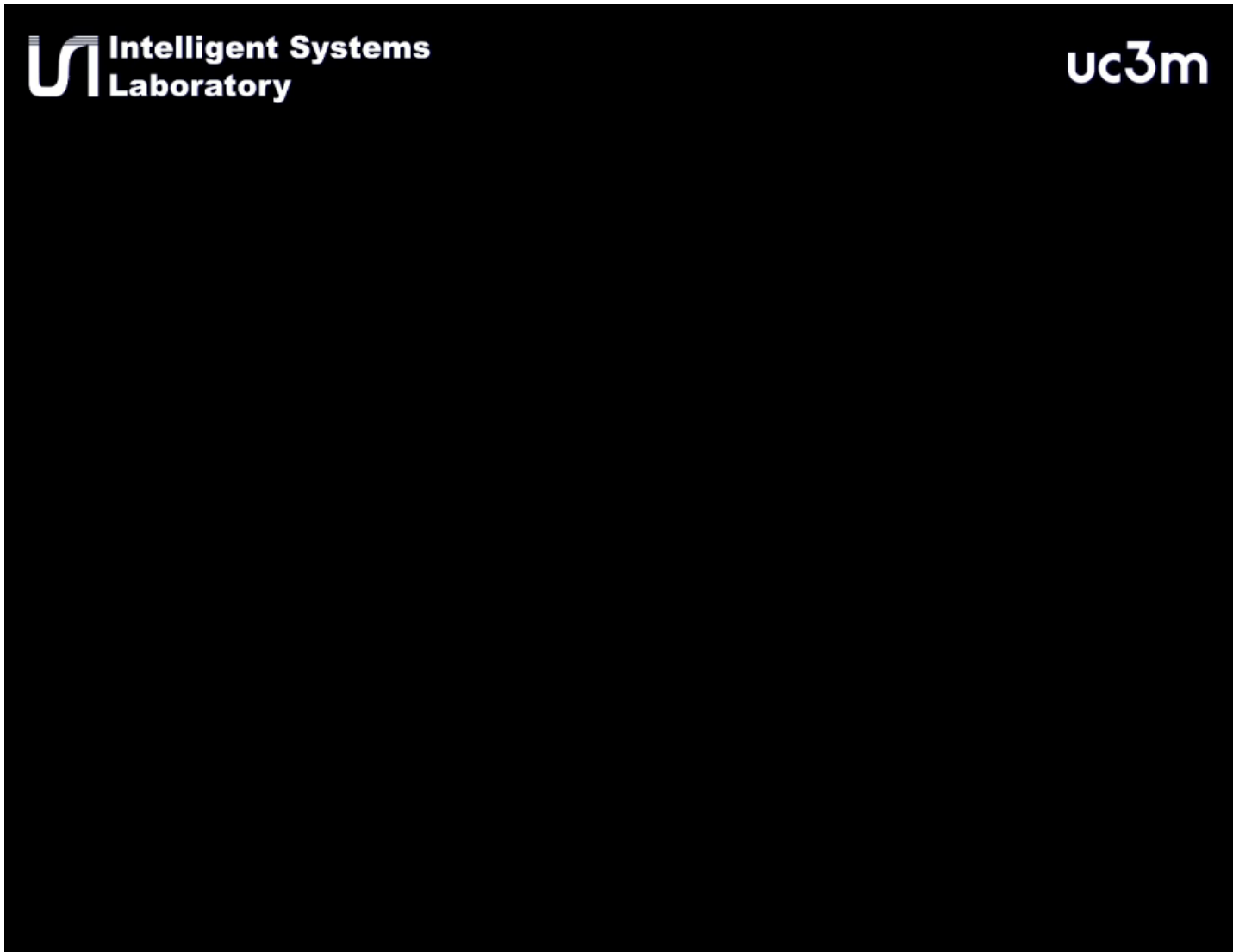


(e)



(f)

Detección y punto de vista (orientación)



IS Intelligent Systems
Laboratory

uc3m

Detección y punto de vista (orientación)

¿Cómo mejorar aún más el rendimiento?

1. Introducir información estéreo

Mapa de disparidades como un “cuarto canal”



↖
Par
estéreo

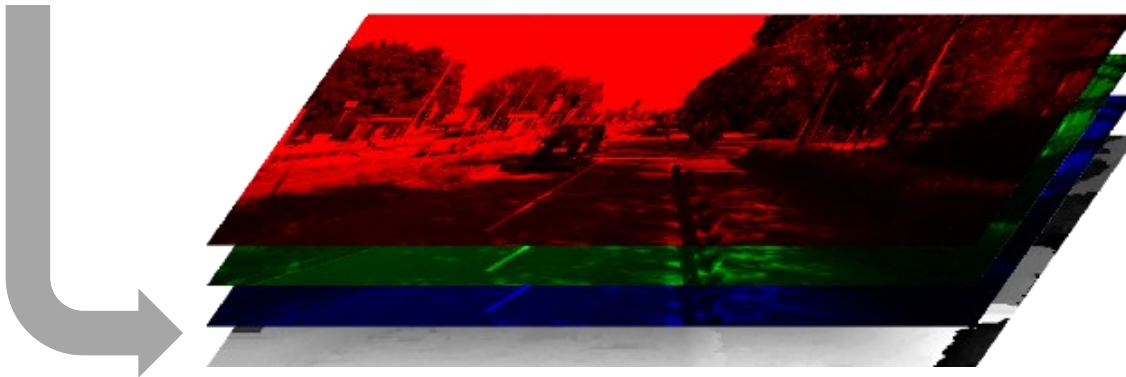
C. Guindel, D. Martín, and J. M. Armingol, “Stereo Vision-Based Convolutional Networks for Object Detection in Driving Environments,” in Computer Aided Systems Theory - EUROCAST 2017, pp. 427–434.

Detección y punto de vista (orientación)

¿Cómo mejorar aún más el rendimiento?

1. Introducir información estéreo

Mapa de disparidades como un “cuarto canal”



Aprox.
+1 mAP

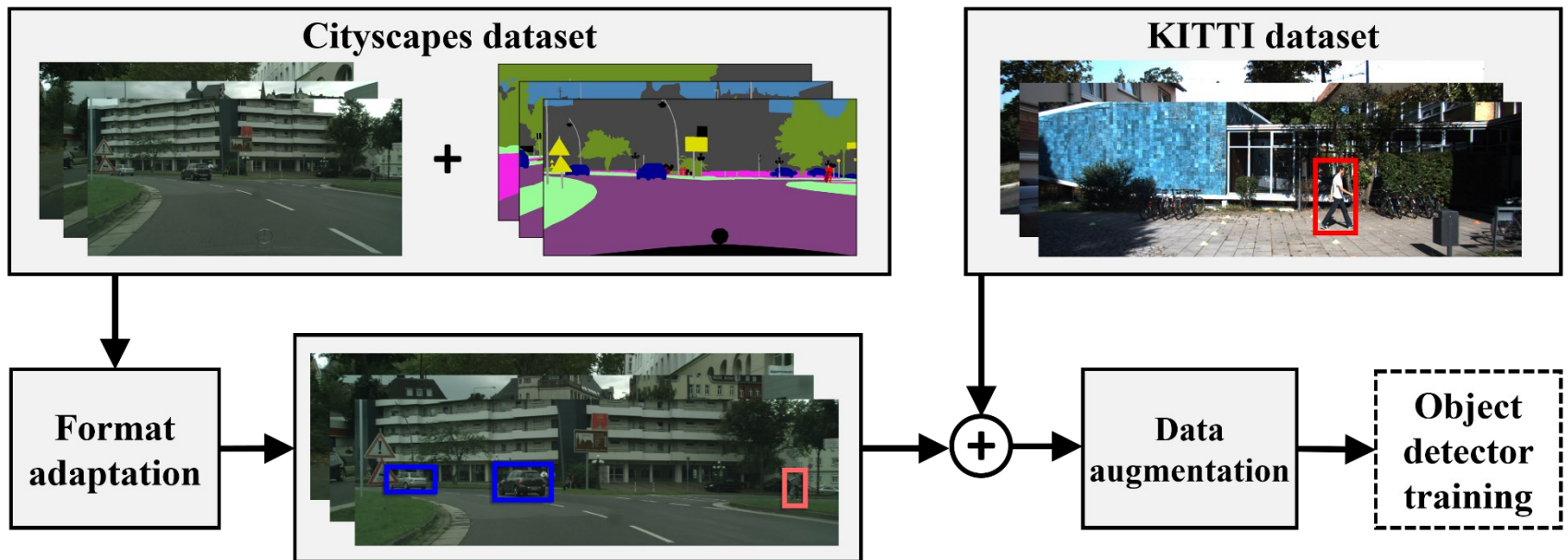
C. Guindel, D. Martín, and J. M. Armingol, “Stereo Vision-Based Convolutional Networks for Object Detection in Driving Environments,” in *Computer Aided Systems Theory - EUROCAST 2017*, pp. 427–434.

Detección y punto de vista (orientación)

¿Cómo mejorar aún más el rendimiento?

2. Combinar diferentes bases de datos

Mejora sustancial al introducir muestras adicionales



- BB DD de 3.712 imágenes: 65,98 % mAP
- Combinación 6.687 imágenes: **73,69%** mAP

C. Guindel, D. Martín, J. M. Armingol, and C. Stiller, "Analysis of the Influence of Training Data on Road User Detection," in Proc. IEEE International Conference on Vehicular Electronics and Safety (ICVES), 2018, pp. 21–26.

Detección y punto de vista (orientación)

¿Cómo mejorar aún más el rendimiento?

3. Aumentado de datos

- Distorsiones aleatorias en las imágenes
- Ayuda a generalizar el modelo (otras situaciones, cámaras, etc.)



C. Guindel, D. Martín, J. M. Armingol, and C. Stiller, "Analysis of the Influence of Training Data on Road User Detection," in Proc. IEEE International Conference on Vehicular Electronics and Safety (ICVES), 2018, pp. 21–26.

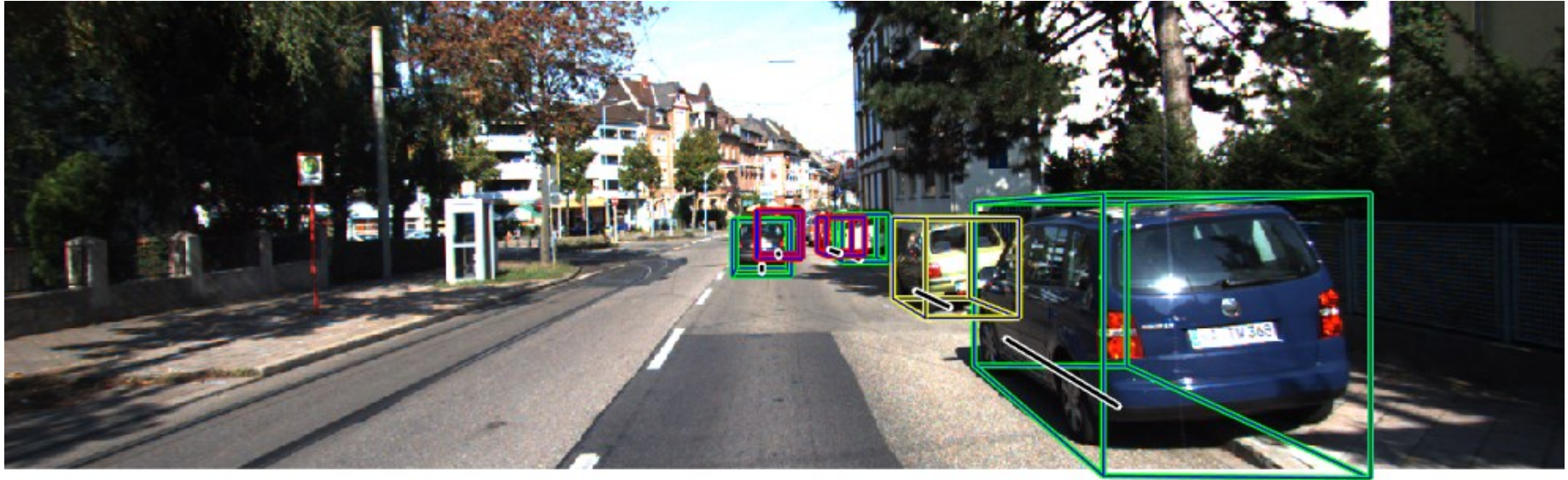
Otros sistemas de visión artificial: detección de líneas de la carretera

The image displays a software interface for road line detection, consisting of several windows and a console window.

- BirdViewRoadMarkings:** A window with three sliders: Threshold1: 45, Param1: 30, and Param2: 200. Below the sliders is a grayscale image of a road with white lines.
- Disparity:** A window with two sliders: Uniq_T...sh: 10 and Textur...sh: 3. Below the sliders is a grayscale image of a road with a dark, textured foreground.
- SegmentPlane:** A window with two sliders: Depth...it: 100 and NPoint...ac: 2000. Below the sliders is a grayscale image of a road with colored lines (magenta, blue, green) and a green trapezoid labeled "2/3".
- Birds_Eye:** A window with three sliders: X: 500, Y: 220, and Z: 592. Below the sliders is a grayscale image of a road with white lines, viewed from a high angle.
- Console Window:** A window displaying the following text:

```
Topologia carril: 0. Tiempo de vida: 10 frames.
Topologia carril: 1. Tiempo de vida: 10 frames.
Topologia carril: 2. Tiempo de vida: 5 frames.
No Cuadro: 23.
carril: 2.
nx: 1.000000. ny: -0.000000 disp: 0.054816.
line1b.Pto1: x=363. y=133. line1b.Pto2: x=333. y=109.
line2b.Pto1: x=443. y=118. line2b.Pto2: x=410. y=110.
line3D1b.Pto1: x=1.630690. y=20.836208. line3D1b.Pto2: x=1.630690. y=39.401258.
line3D2b.Pto1: x=5.285084. y=29.899461. line3D2b.Pto2: x=5.285084. y=39.181986.
line1a.Pto1: x=365. y=132. line1a.Pto2: x=334. y=100.
line2a.Pto1: x=445. y=117. line2a.Pto2: x=411. y=109.
line3D1a.Pto1: x=1802583312. y=1073412053. line3D1a.Pto2: x=-1156787848. y=1077204497.
line3D2a.Pto1: x=1802583328. y=1073412053. line3D2a.Pto2: x=1843807136. y=1078178652.
Topologia carril: 0. Tiempo de vida: 10 frames.
Topologia carril: 1. Tiempo de vida: 10 frames.
Topologia carril: 2. Tiempo de vida: 4 frames.
No Cuadro: 24.
Topologia carril: 0. Tiempo de vida: 10 frames.
Topologia carril: 1. Tiempo de vida: 10 frames.
Topologia carril: 2. Tiempo de vida: 5 frames.
No Cuadro: 25.
```

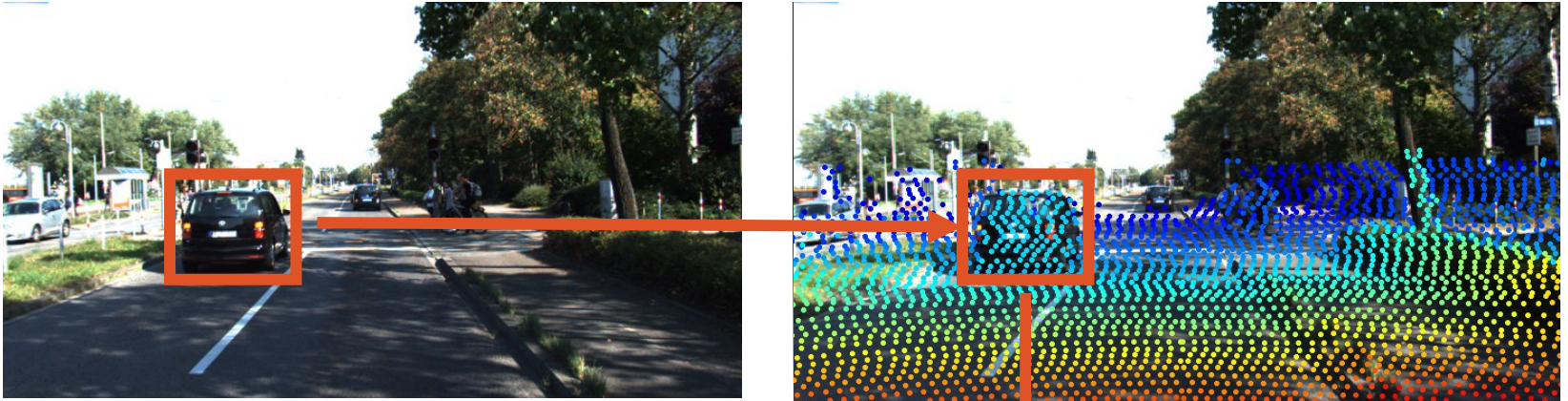
Localización 3D



Fuente: KITTI dataset

- Coordenadas en la imagen: (u,v) ✓
 - ¿Coordenadas en el espacio (X,Y,Z) ?
 - ¿Tamaño de los objetos (l, w, h) ?
 - ¿Yaw (θ) ?
- Se necesita información geométrica:
 - Estéreo: una sola modalidad, poco preciso
 - Lidar: alta precisión, menos información

Localización 3D



Distancia

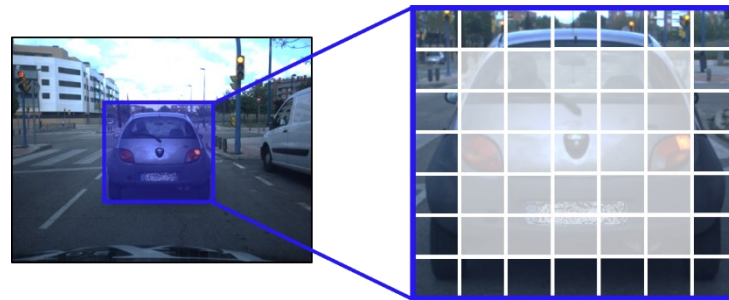
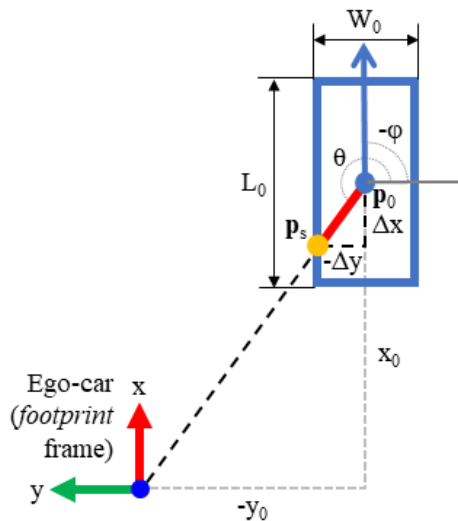
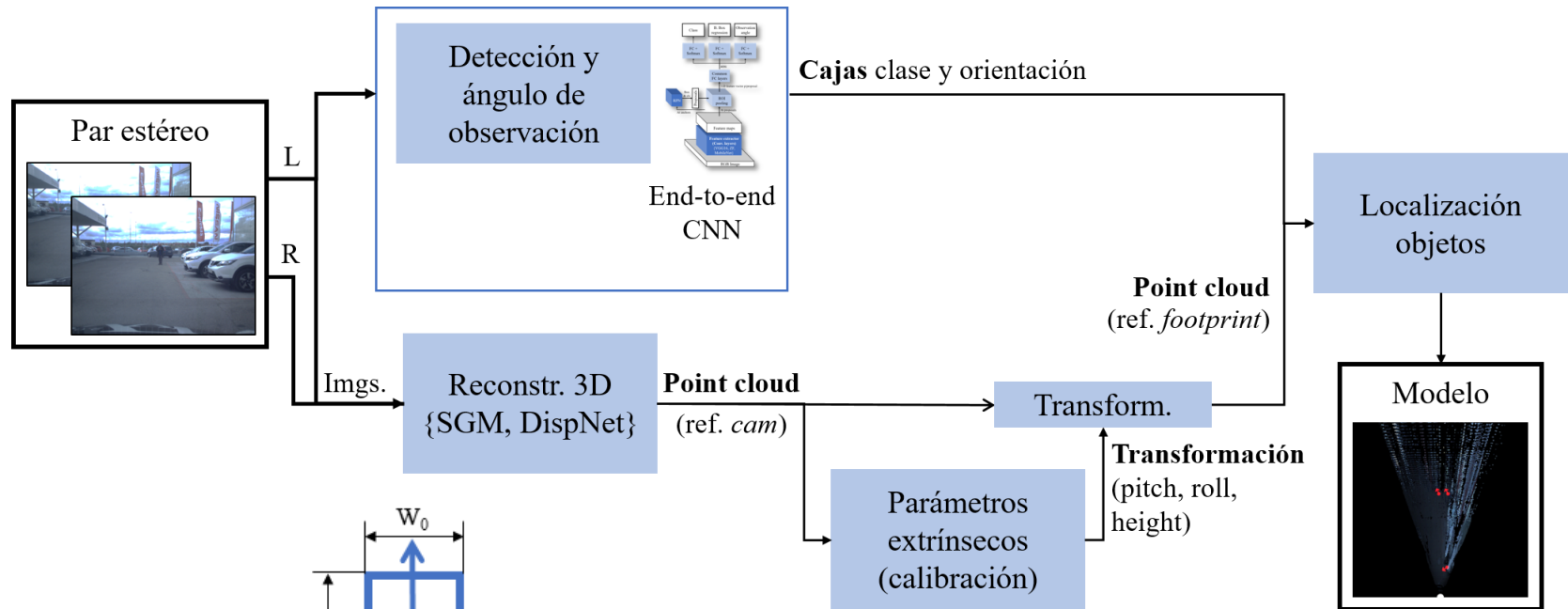
¿Punto más cercano?

¿Centroide?

- ¿Y las dimensiones del objeto?
- ¿Y su orientación en coordenadas del mundo?

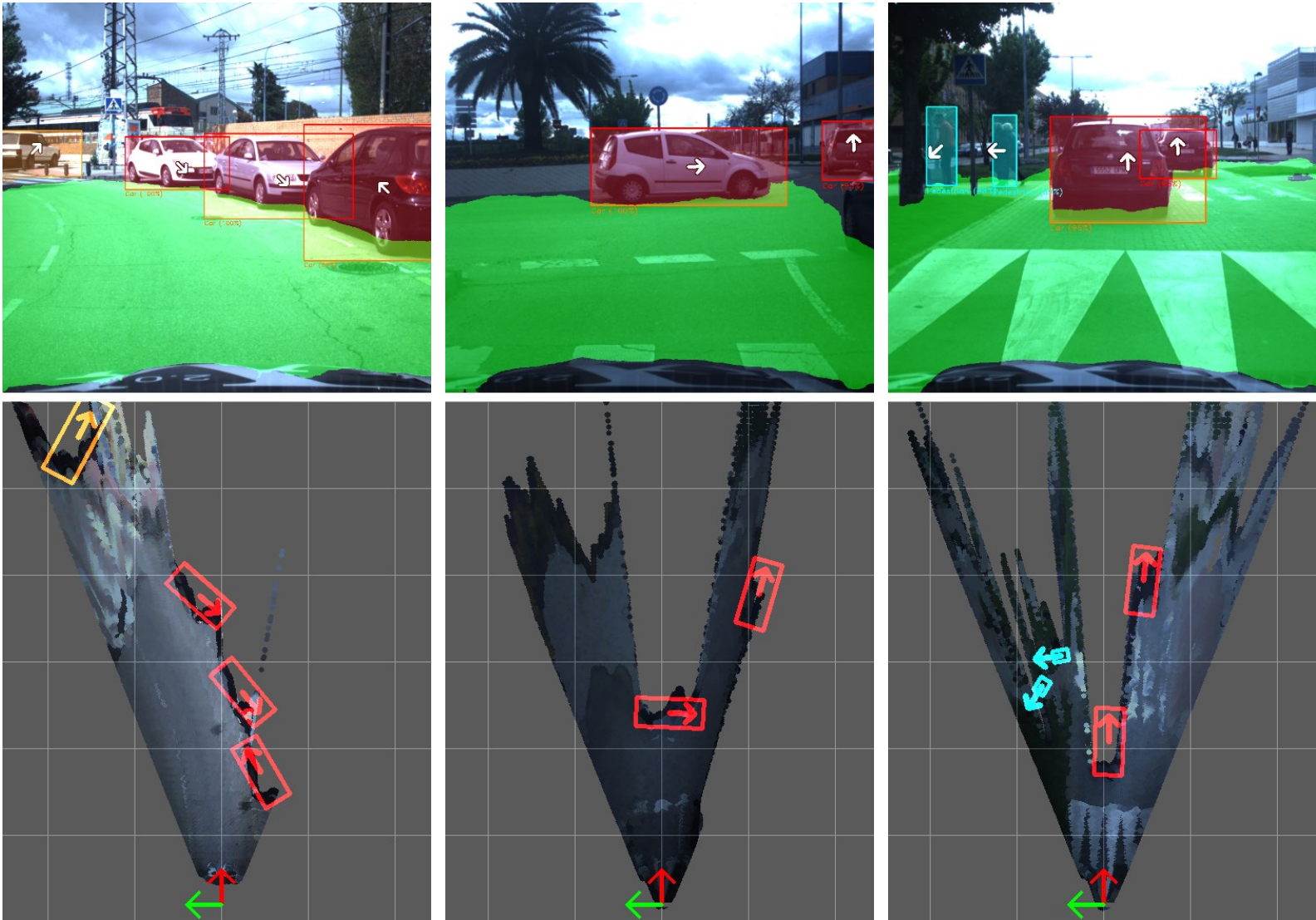
No hay medidas de distancia para las zonas no visibles por el sensor

Localización 3D: estéreo

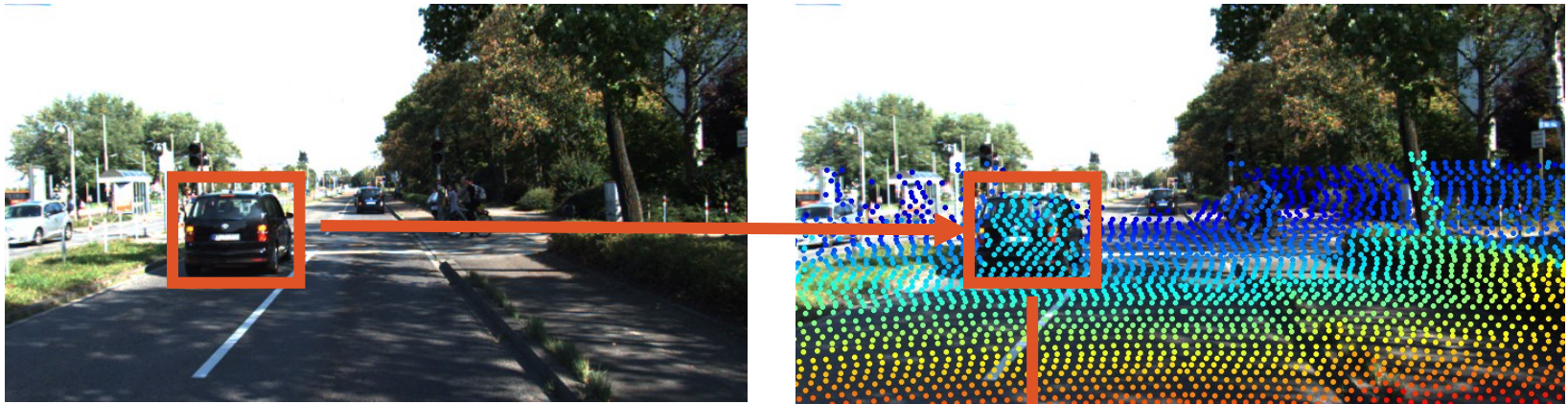


C. Guindel, D. Martín, and J. M. Armingol, "Traffic Scene Awareness for Intelligent Vehicles using ConvNets and Stereo Vision," *Rob. Auton. Syst.*, vol. 112, pp. 109–122, 2019.

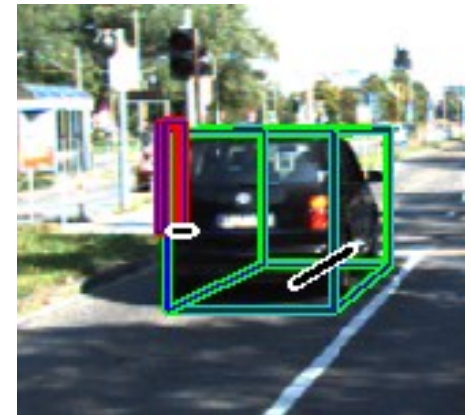
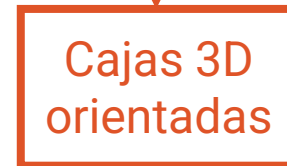
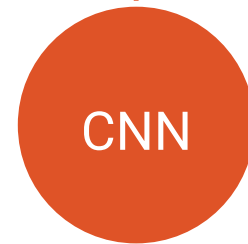
Localización 3D: estéreo



Localización 3D: alternativas

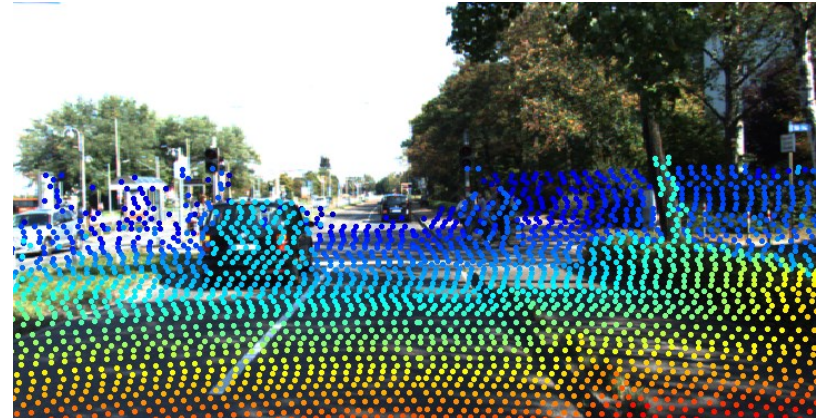


- Algunos métodos utilizan una red neuronal para estimar:
 - Puntos que pertenecen verdaderamente al objeto
 - Centro del objeto
 - Dimensiones del objeto
- Ej: Frustum PointNets



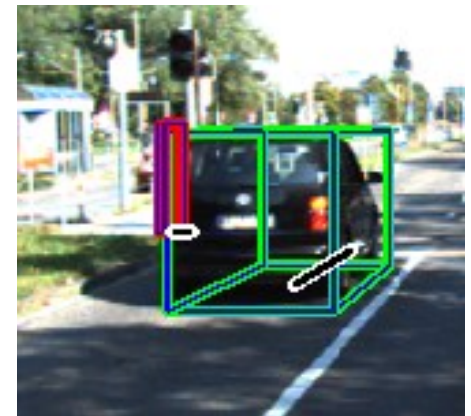
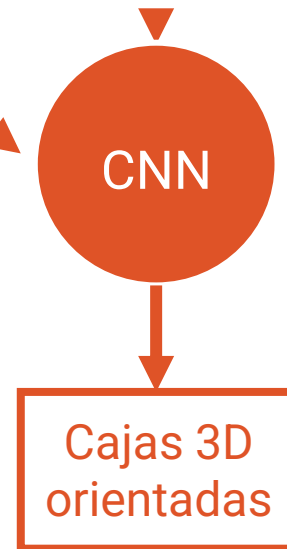
C. R. Qi, W. Liu, C. Wu, H. Su, and L. J. Guibas, "Frustum PointNets for 3D Object Detection from RGB-D Data," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 918–927.

Localización 3D: alternativas



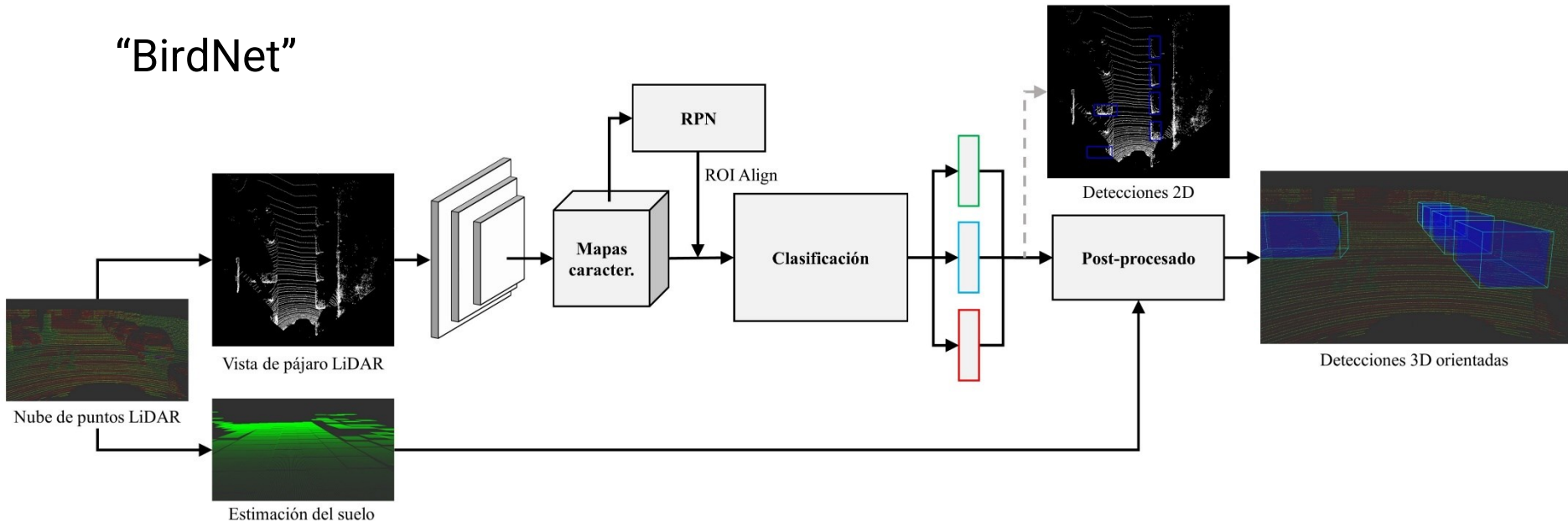
- Otros métodos estiman las detecciones directamente de la fusión de cámara y lidar
- El reto es la representación de la información
- Ej: Multi-View 3D networks (MV3D)

X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, "Multi-View 3D Object Detection Network for Autonomous Driving," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 6526–6534.



Detección y localización 3D: lidar

“BirdNet”

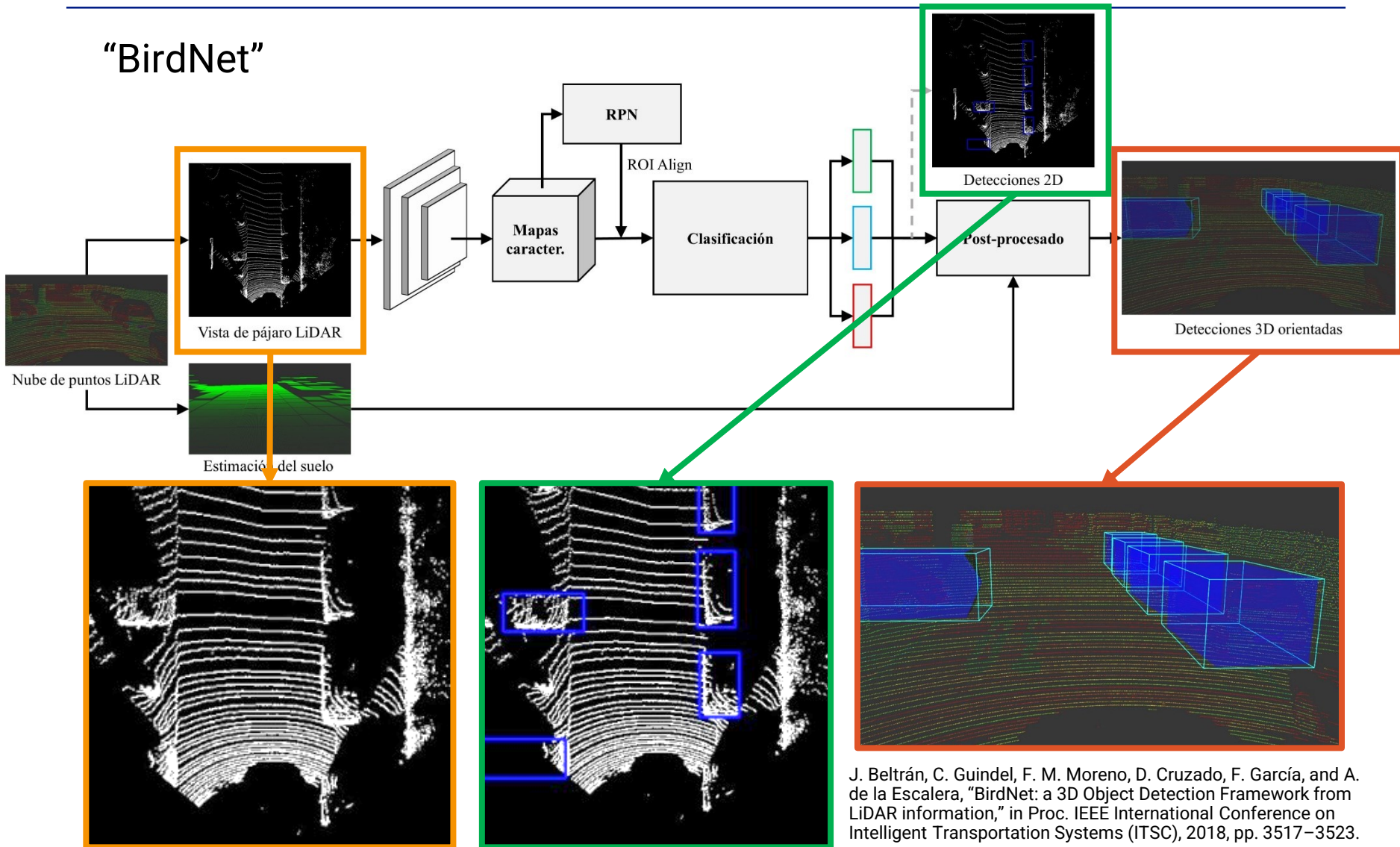


- Usa únicamente información del lidar (distancias)
- Esta información se representa de forma análoga a una imagen
- Detección y localización 3D simultáneas en 360°

J. Beltrán, C. Guindel, F. M. Moreno, D. Cruzado, F. García, and A. de la Escalera, “BirdNet: a 3D Object Detection Framework from LiDAR information,” in Proc. IEEE International Conference on Intelligent Transportation Systems (ITSC), 2018, pp. 3517–3523.

Detección y localización 3D: lidar

“BirdNet”



J. Beltrán, C. Guindel, F. M. Moreno, D. Cruzado, F. García, and A. de la Escalera, “BirdNet: a 3D Object Detection Framework from LiDAR information,” in Proc. IEEE International Conference on Intelligent Transportation Systems (ITSC), 2018, pp. 3517–3523.

Detección y localización 3D: lidar

BirdNet: a 3D Object Detection Framework from LiDAR Information

Jorge Beltrán, Carlos Guindel, Francisco Miguel Moreno,
Daniel Cruzado, Fernando García, and Arturo de la Escalera

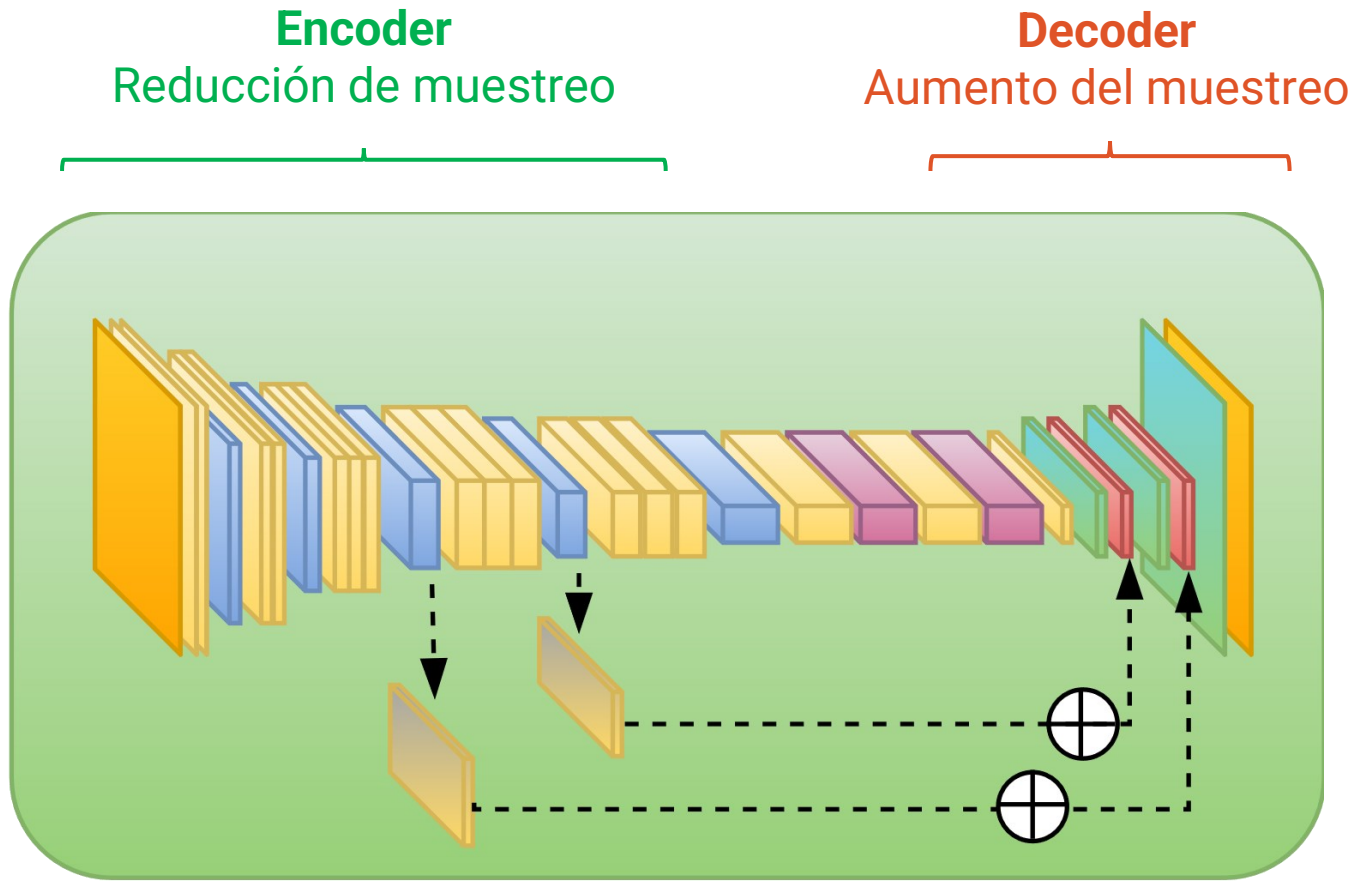
Segmentación semántica



- Asignación de una clase a cada píxel
- Reconoce obstáculos pero no los diferencia entre sí
 - Los métodos “de instancias” sí (aplicando detección)
- Especialmente útil para elementos de la infraestructura:
 - Carretera, carriles, etc.

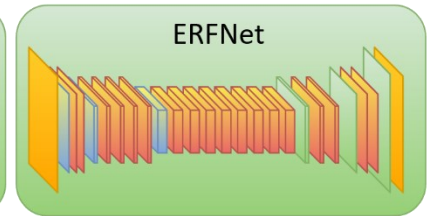
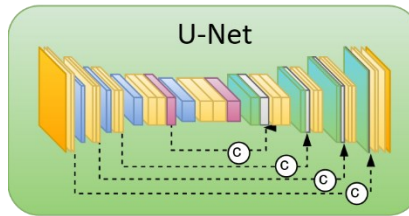
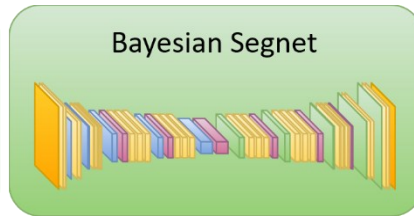
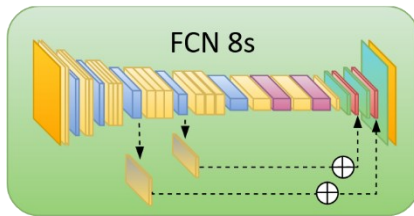
Segmentación semántica

- Redes totalmente convolucionales



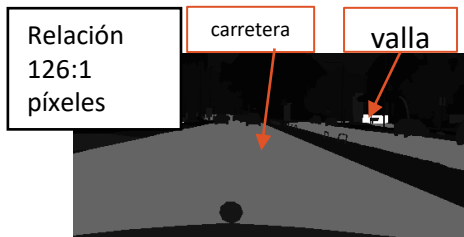
Segmentación semántica

Arquitecturas

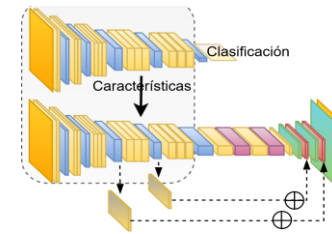


Técnicas de entrenamiento

Equilibrado de clases



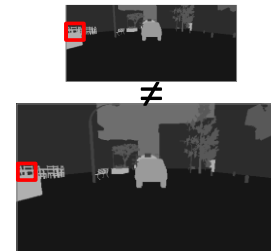
Transferencia de aprendizaje



Aumentado de datos



Reescalado de la imagen



A. Barrera, C. Guindel, F. García, and D. Martín, "Análisis, evaluación e implementación de algoritmos de segmentación semántica para su aplicación en vehículos inteligentes," en XXXIX Jornadas de Automática, 2018, pp. 983–990.

Segmentación semántica



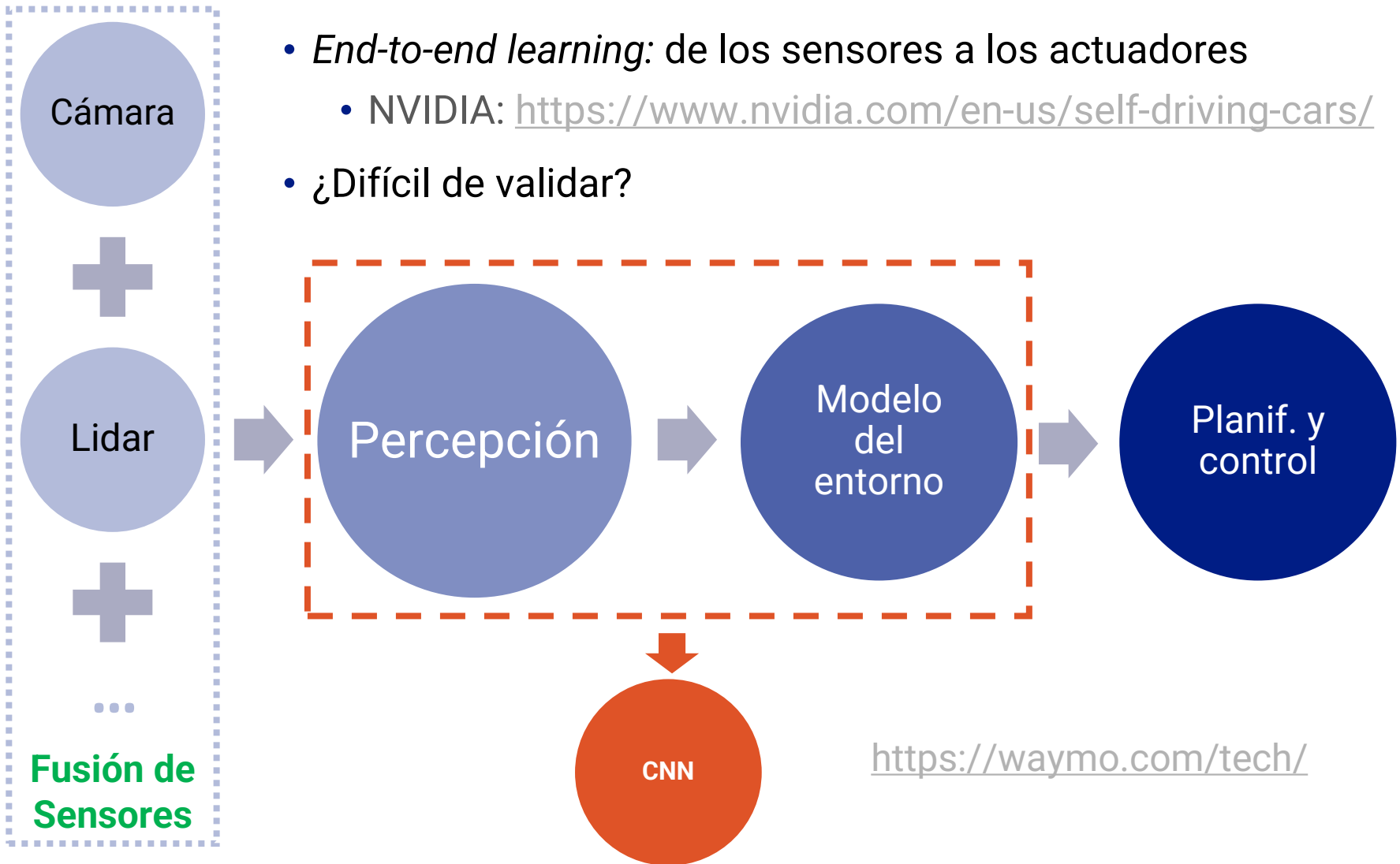
Segmentación semántica: carriles



Segmentación semántica: resultados



¿Futuro?



- *End-to-end learning*: de los sensores a los actuadores
 - NVIDIA: <https://www.nvidia.com/en-us/self-driving-cars/>
- ¿Difícil de validar?

Arquitectura de navegación autónoma nivel 3-4

Icabs



Arquitectura de navegación autónoma nivel 3-4

Cámara Trinocular: odometría visual, mapa de percepción local en tiempo real, navegación autónoma, etc

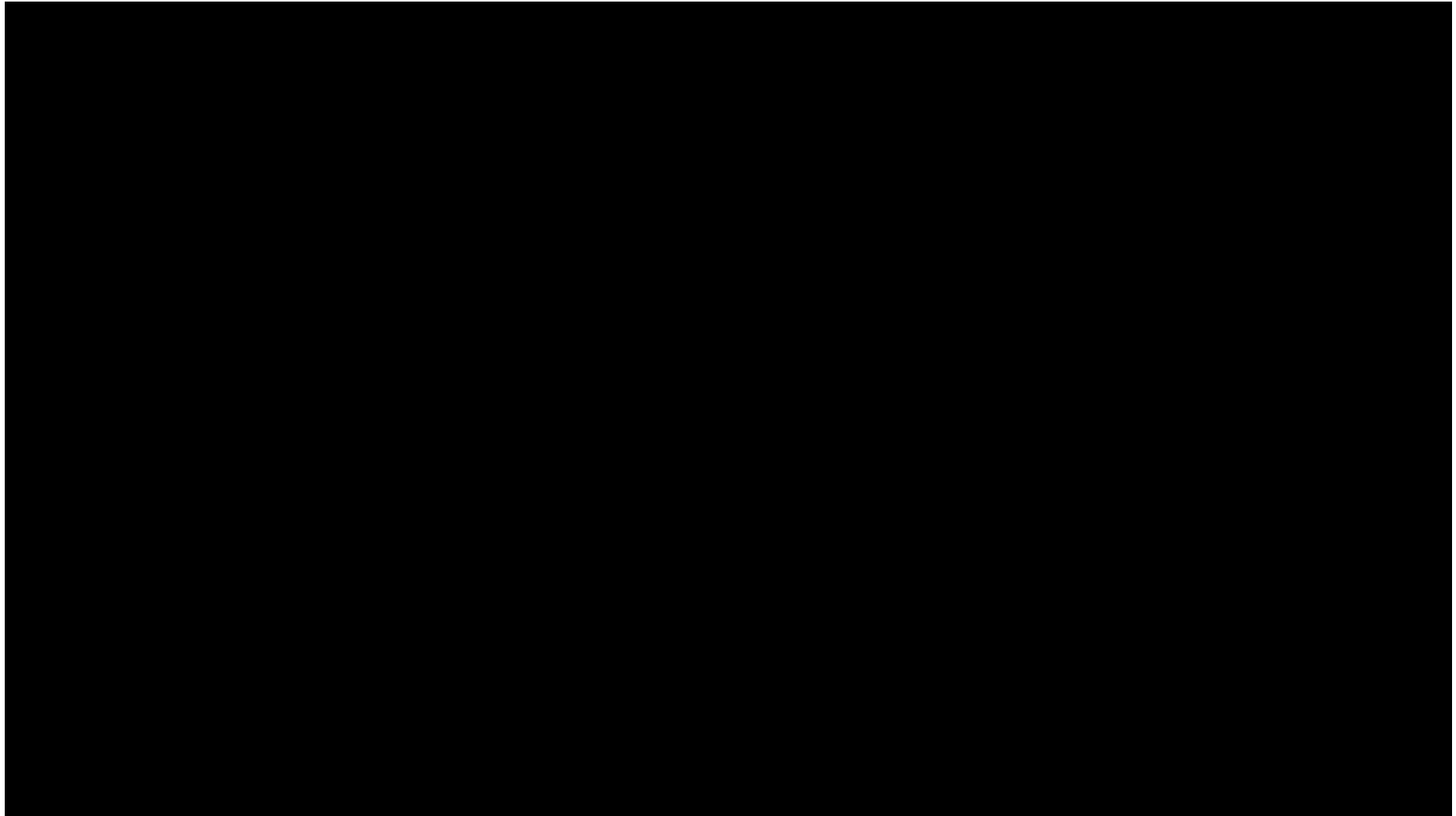


Arquitectura embebida basada en ROS

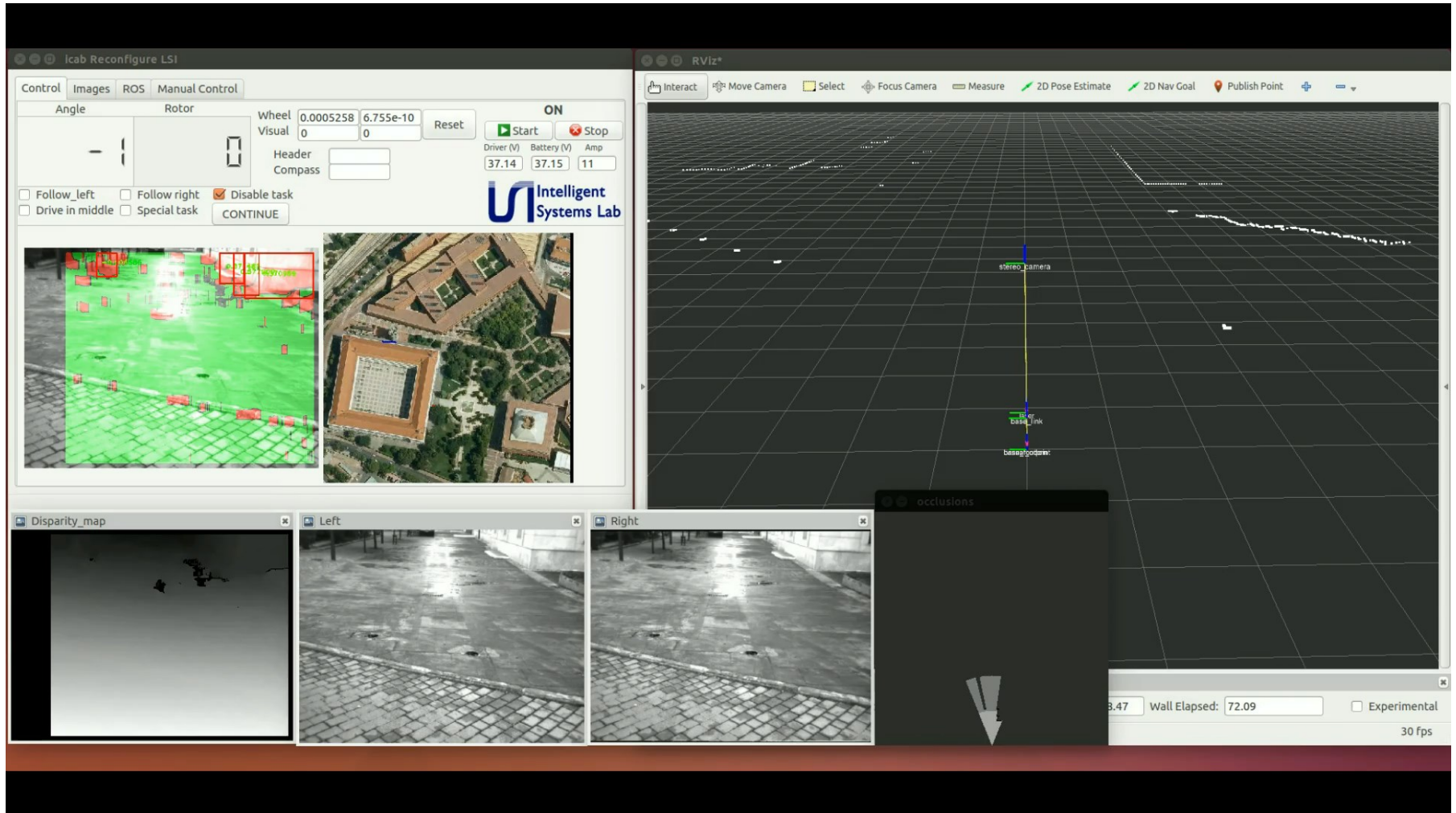
Cooperación entre vehículos autónomos

Láser: detección y evitación de obstáculos estáticos y dinámicos en tiempo real

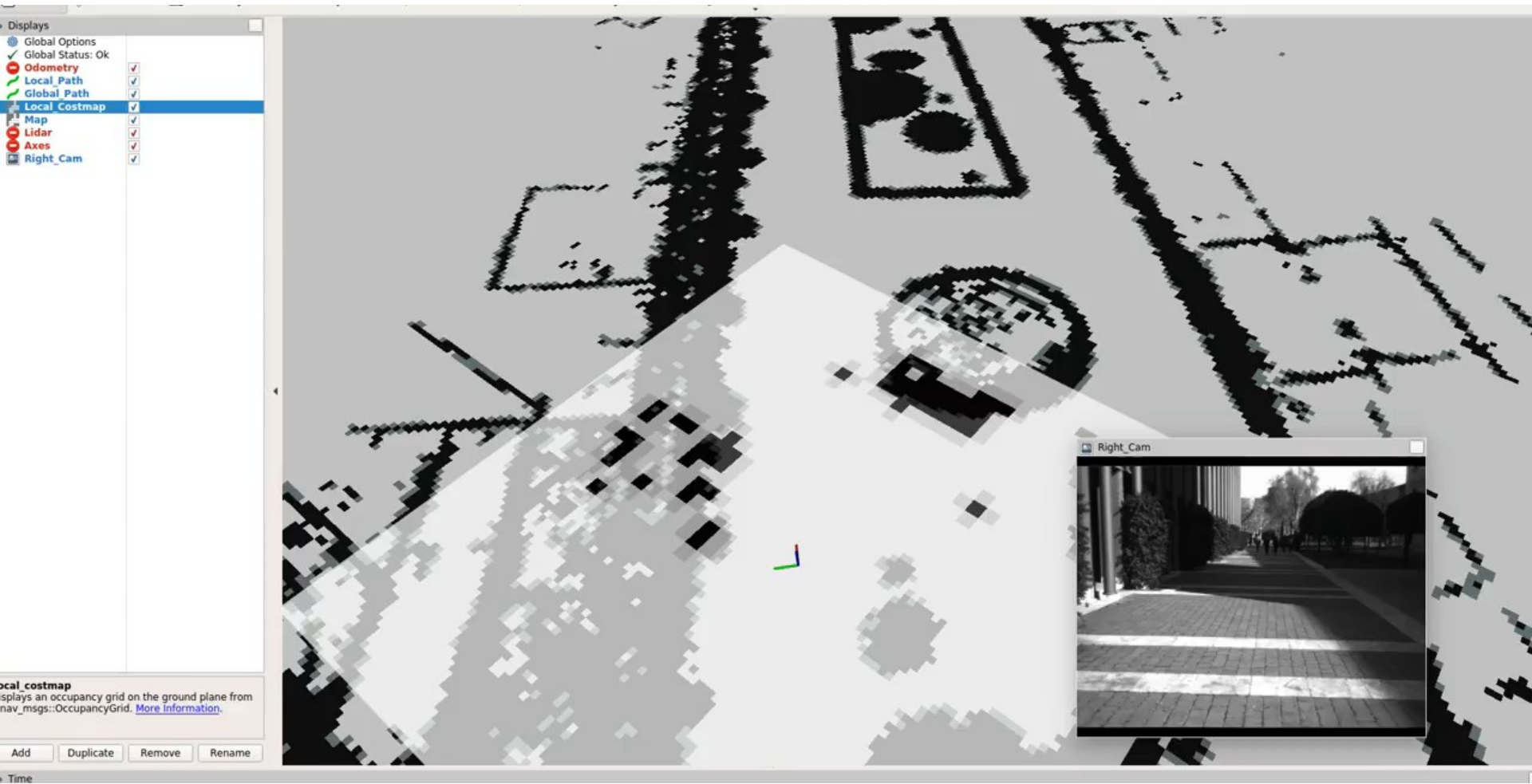
Arquitectura de navegación autónoma nivel 3-4



Arquitectura de navegación autónoma nivel 3-4



Arquitectura de navegación autónoma nivel 3-4



Arquitectura de navegación autónoma nivel 4

- ¿Dónde está la frontera del conocimiento en toma de decisión?

Nivel	Nombre	Aceleración, dirección,...	Supervisión del entorno	Respuesta a errores	Modos de conducción
0	Sin automatización	Conductor	Conductor	Conductor	No
1	Asistencia al conductor	Conductor y sistema	Conductor	Conductor	Algunos
2	Automatización parcial	Sistema	Conductor	Conductor	Algunos
3	Automatización condicional	Sistema	Sistema	Conductor	Algunos
4	Automatización alta	Sistema	Sistema	Sistema	Algunos
5	Automatización completa	Sistema	Sistema	Sistema	Todos

Módulo avanzado de autoconocimiento para la toma de decisión en los iCabs

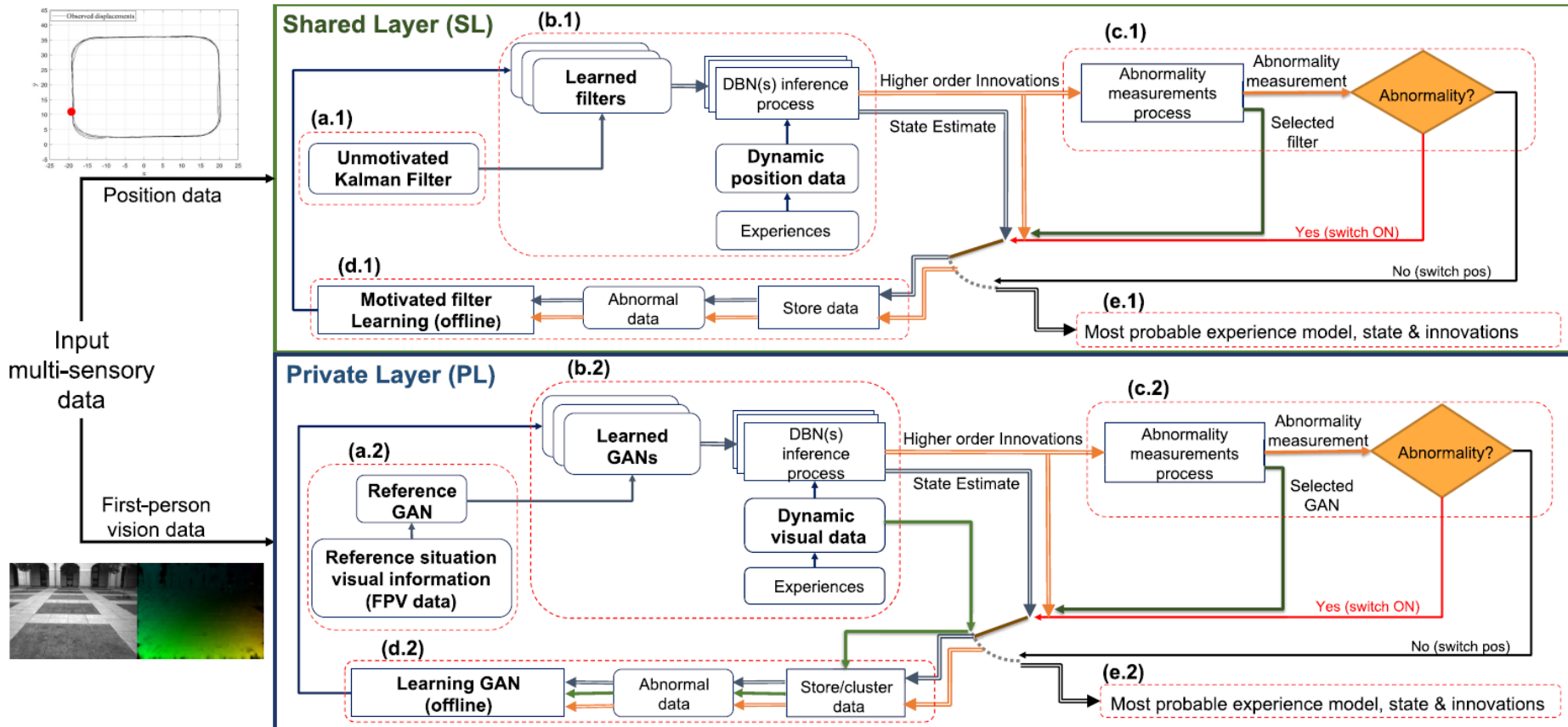
Autoconocimiento y sistemas avanzados de toma de decisión:

1. Capacidad de detectar situaciones **novedosas** que no han sido **modeladas previamente**
2. Anticipar estados situacionales/contextuales y mejorar la efectividad de los módulos de **toma de decisiones**
3. Inyectar comportamiento humano de **toma de decisiones** para adaptar un vehículo autónomo a una gran variedad de entornos de tráfico

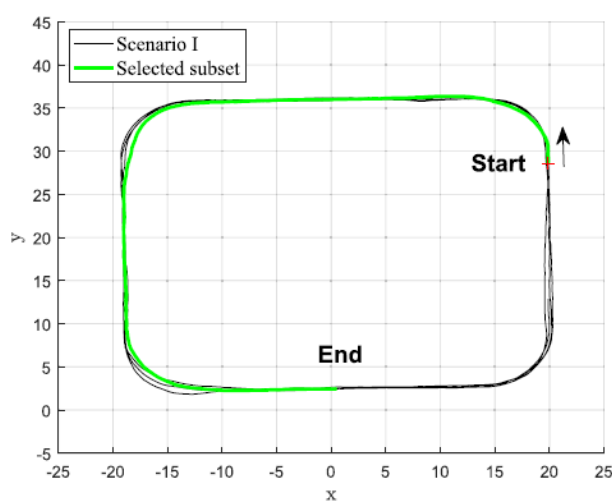
Módulo avanzado de autoconocimiento para la toma de decisión en los iCabs



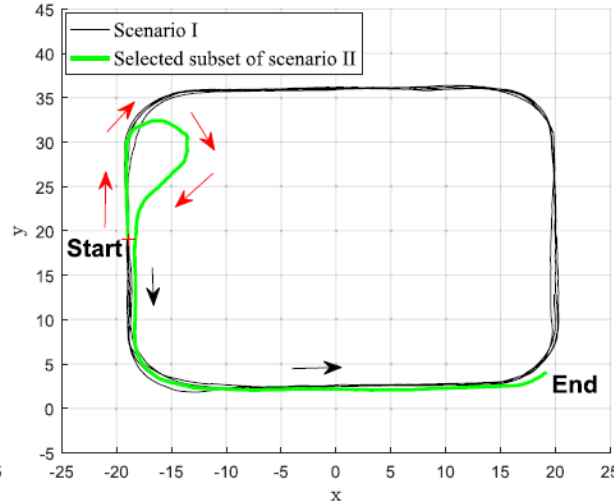
Módulo avanzado de autoconocimiento para la toma de decisión en los iCabs



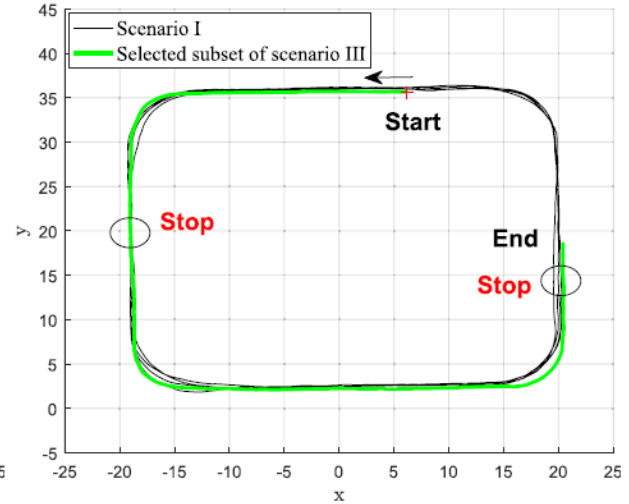
Módulo avanzado de autoconocimiento para la toma de decisión en los iCabs



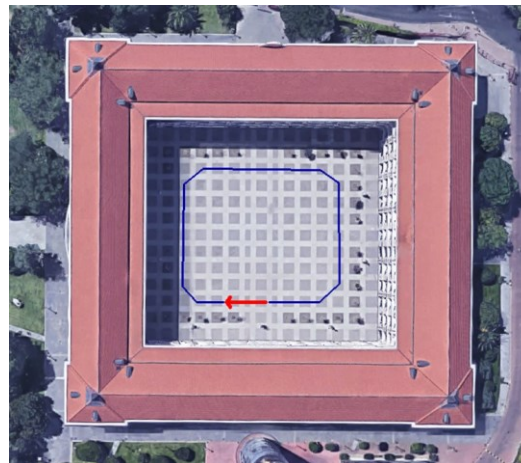
(a) perimeter monitoring



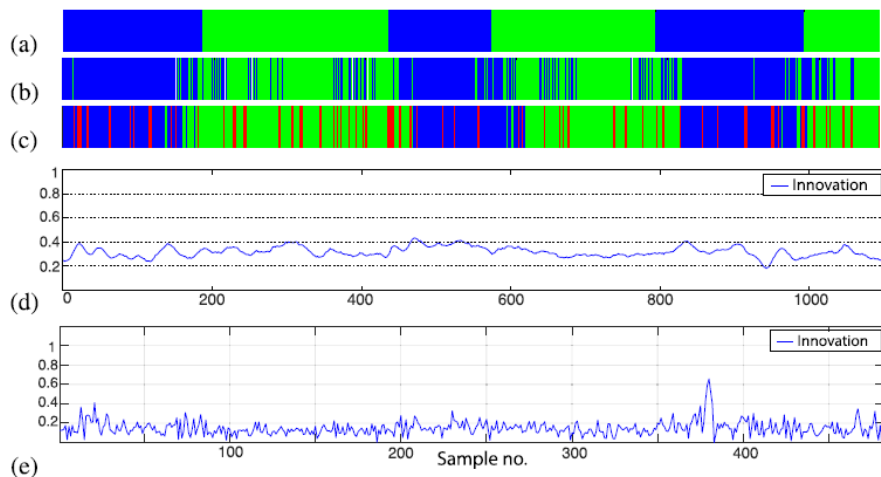
(b) U-turn



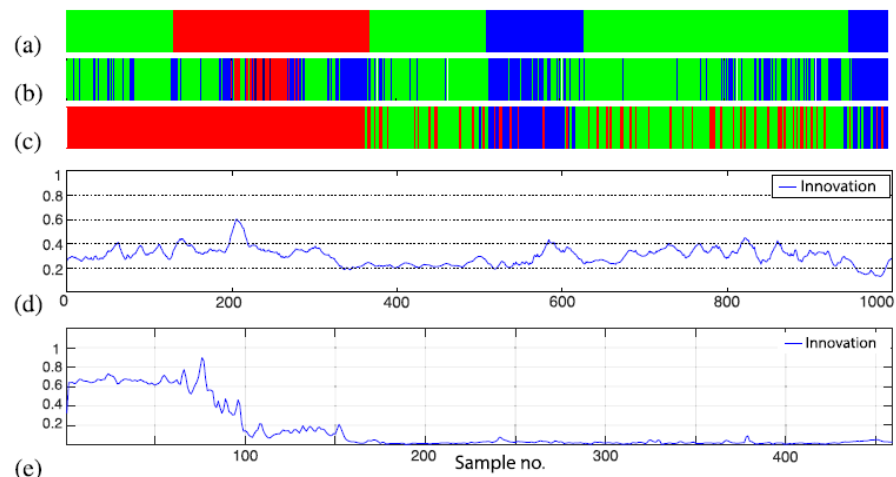
(c) emergency stop



Módulo avanzado de autoconocimiento para la toma de decisión en los iCabs

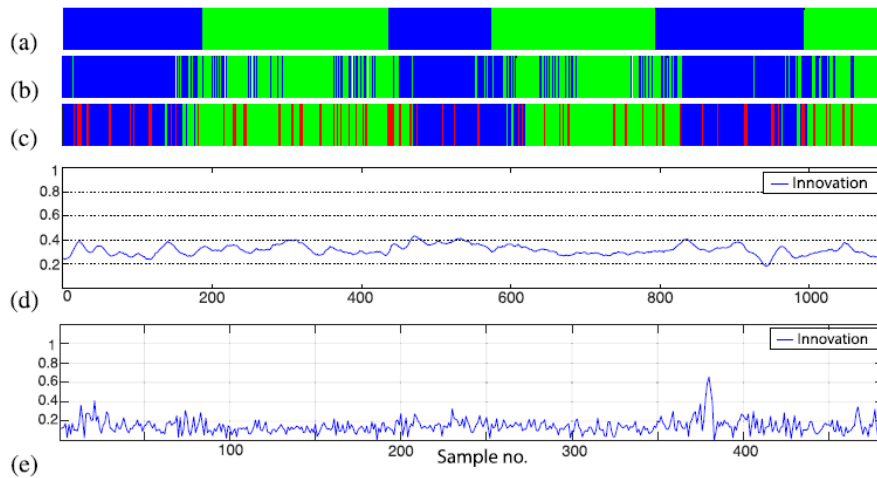


**Comportamiento normal en
ambas capas PL y SL**

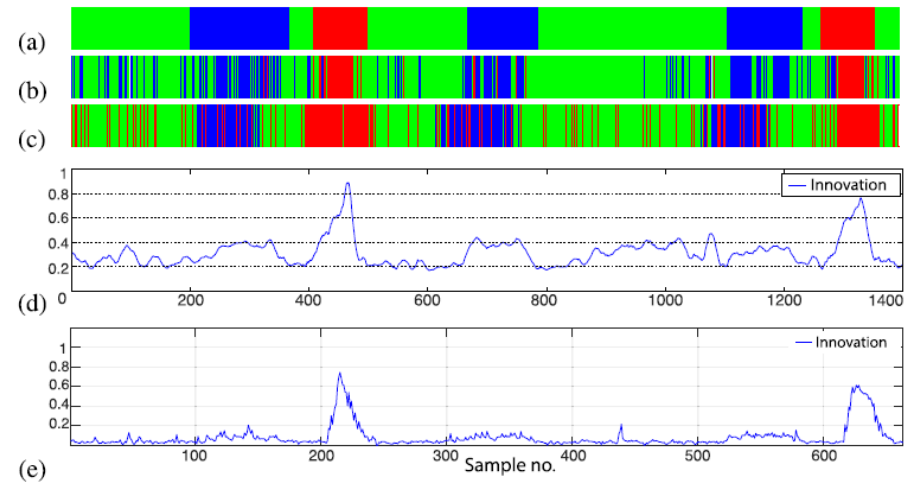


**Autoconocimiento de
anomalía en el escenario de
evitación de peatón mediante
un giro en U**

Módulo avanzado de autoconocimiento para la toma de decisión en los iCabs



**Comportamiento normal
en ambas capas PL y SL**



**Autoconocimiento de
anomalía en el escenario
de paradas de emergencia**

Muchas gracias por
vuestra atención