

Sistemas de Percepción para Vehículos Autónomos

Arturo de la Escalera Hueso

Laboratorio de Sistemas Inteligentes

www.uc3m.es/islab

Dpto. Ingeniería de Sistemas y Automática

Universidad Carlos III de Madrid



UNED

Ciclo de Conferencias del Máster y
del Programa de Doctorado en
Ingeniería de Sistemas y de Control



- ◆ **Presentación del LSI**
- ◆ **Entendimiento del entorno viario**
- ◆ **Percepción del entorno**
- ◆ **Conclusiones**

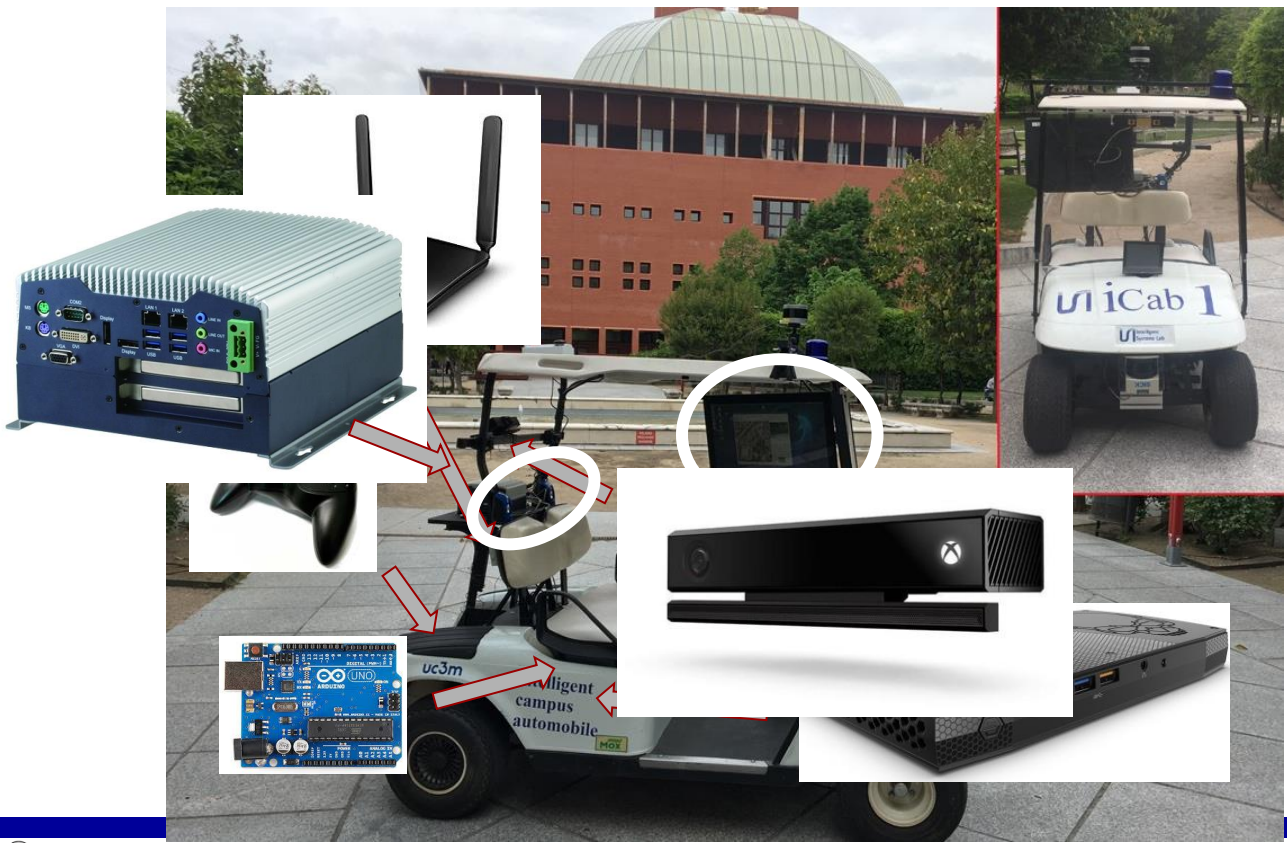
◆ Equipo

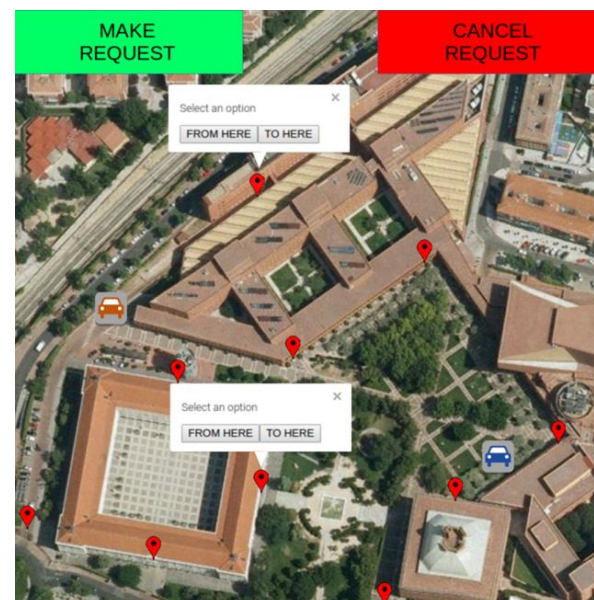
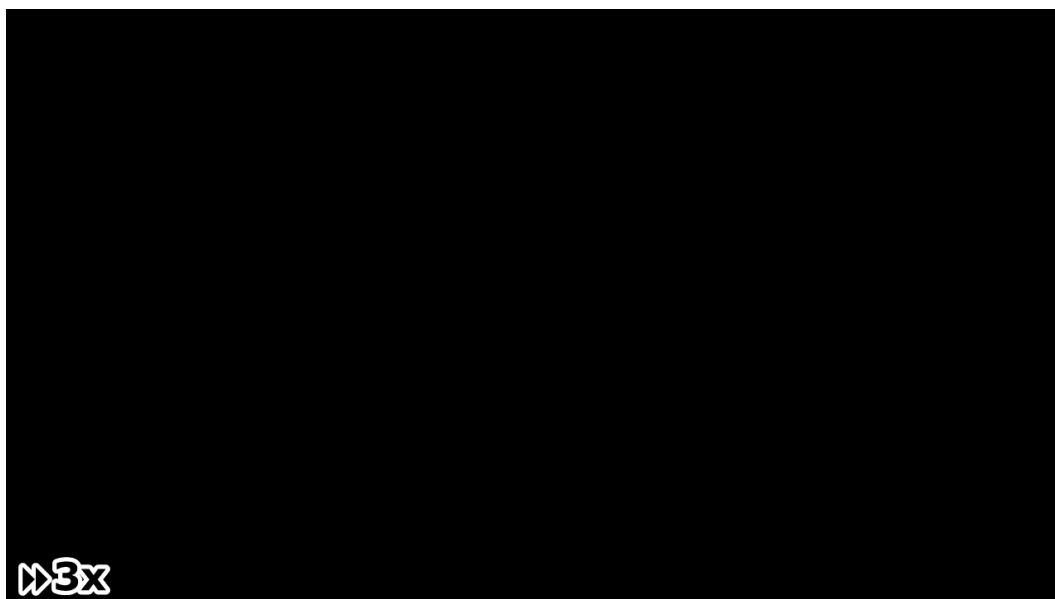
- ❖ 2 Catedráticos
- ❖ 3 Profesores titulares
- ❖ 4 Profesores doctores
- ❖ 9 Estudiantes de doctorado
- ❖ 12 Ingenieros de proyectos
- ❖ 1 Profesor asociado





iCab (Intelligent Campus Automobile)





Autonomous Driverless Automobile (Ada)



Ciclo de Conferencias del Master y del Programa de Doctorado en Ingeniería de Sistemas y de Control

Arturo de la Escalera
Sistemas de Percepción para Vehículos Autónomos
28 de Octubre de 2020

◆ Centros de Arte, Cultura y Turismo (Cabildo de Lanzarote)

❖ Ruta de los Volcanes

- ✓ 11Km, fuertes pendientes, curvas cerradas, poca cobertura
- ✓ Imposibilidad de actuar en el PN
- ✓ Nivel 4
- ✓ Falta de procedimiento de homologación
- ✓ Lograr la confianza del usuario

<https://citiestimanfaya.com>



CITIES Timanfaya

- ◆ Centros de Arte, Cultura y Turismo (Cabildo de Lanzarote)
- ◆ Albufera
- ◆ Asociación Española de la Carretera
- ◆ Cevimap
- ◆ 2RK
- ◆ GMV
- ◆ Instituto de Seguridad del Vehículo Automóvil, UC3M
- ◆ Mapfre
- ◆ SGS-GMR
- ◆ VTI

Promotores	 <p>Centros de Arte, Cultura y Turismo Cabildo de Lanzarote</p>	 <p>FONDO DE DESARROLLO DE CANARIAS</p>	 <p>CABILDO DE LANZAROTE</p>	 <p>Gobierno de Canarias</p>
Coordinación y gestión del proyecto	 <p>Asociación Española de la Carretera</p>			
Dirección Técnica				
Coordinación tecnológica	 <p>ISVA «DUQUE DE SANTOMAURO»</p>	 <p>uc3m</p>	 <p>Universidad Carlos III de Madrid</p>	
	 <p>vti Vázquez y Torres Ingeniería SL.</p>	 <p>SGS</p>	 <p>gmv INNOVATING SOLUTIONS</p>	 <p>GMR Homologación</p>
	 <p>MAPFRE</p>	 <p>CESVIMAP</p>	 <p>albufera Energy Storage</p>	

CITIES Timanfaya



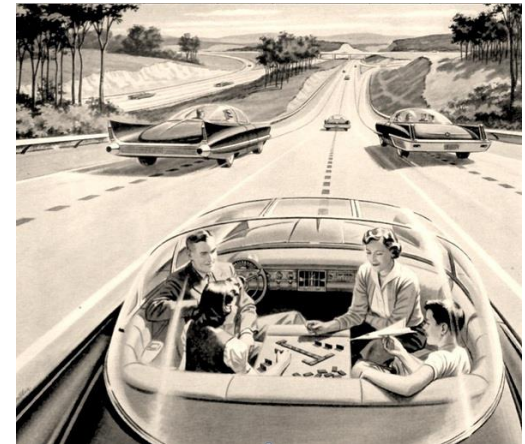
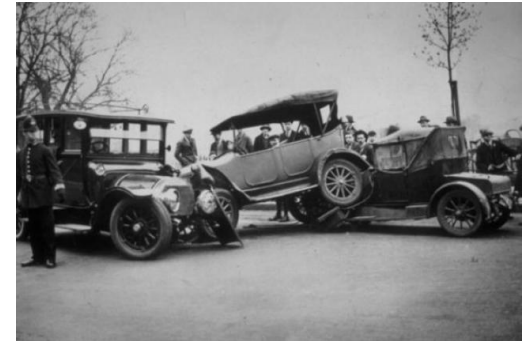
- ◆ **Presentación del LSI**
- ◆ **Entendimiento del entorno viario**
 - ❖ Necesidades
 - ❖ Dificultades
 - ❖ Sensores
- ◆ **Percepción del entorno**
- ◆ **Conclusiones**

- ◆ **Vehículo capaz, por sus propios medios, de realizar todas las tareas necesarias para ir desde un punto a su destino, de forma segura y confortable, sin ninguna intervención humana**

- ❖ ¿Eléctrico?
- ❖ ¿De uso compartido?
- ❖ ¿Conectado?



- ❖ Menos accidentes
 - ✓ 1.1 fallecimiento por 10^8 de millas recorridas (2012)
¿Objetivo 10^9 ?
- ❖ Mejor uso de la infraestructuras existentes
 - ✓ Más coches pueden usar la misma infraestructura
- ❖ Reducción de la infraestructuras
 - ✓ Más espacio dedicada a jardines
- ❖ Mejor aprovechamiento del tiempo
- ❖ Vehículos autónomos compartidos
 - ✓ Reducción de costes
- ❖ Inclusión social: disponibles para todos sin restricción de:
 - ✓ Edad (demasiado joven ... o mayor)
 - ✓ Enfermedad
- ❖ Accesibilidad
 - ✓ Facilitar el acceso al centro de las ciudades





Intelligent Systems

Vehículos autónomos

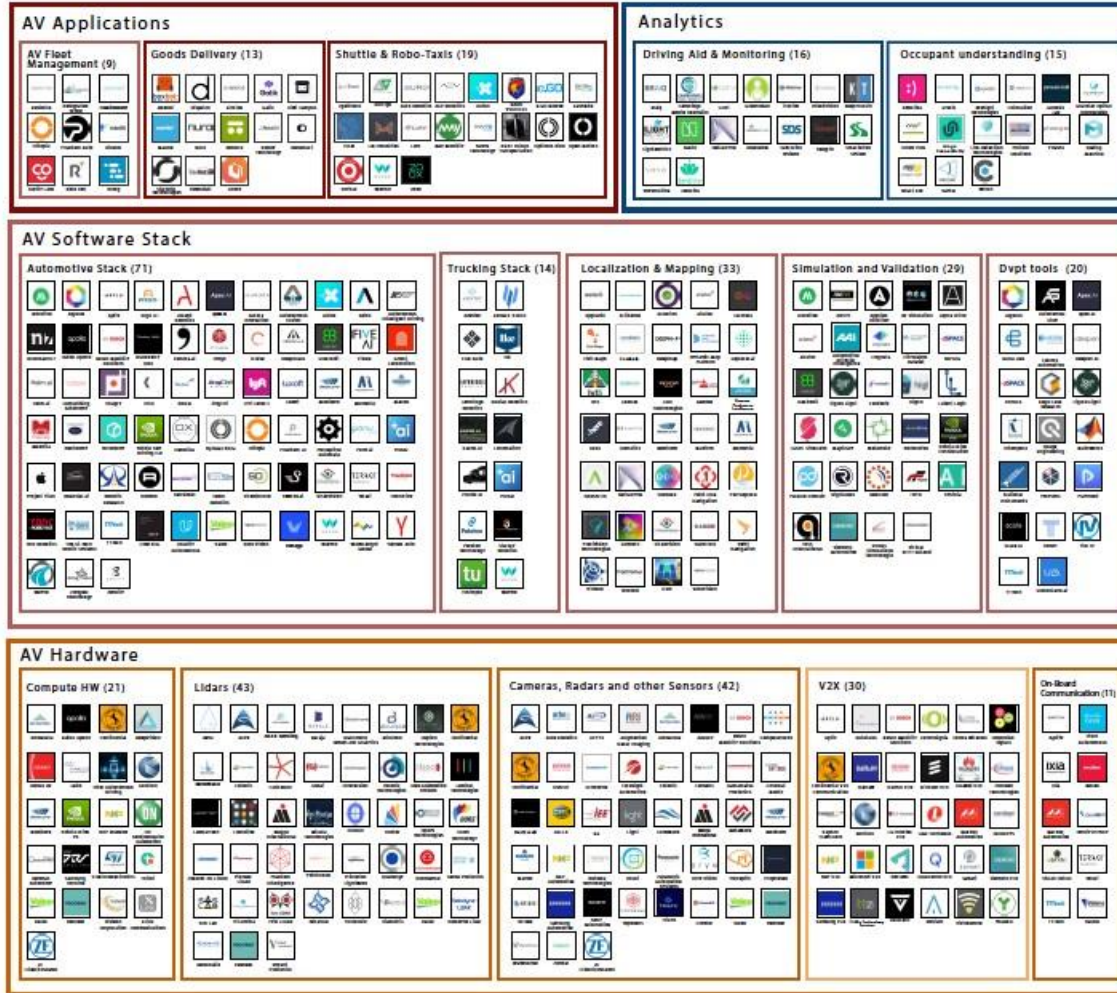
Created By:
Marc Amblard

The Autonomous Vehicles landscape
Last Update August 2019

Designed By: **rwLABS**
Powered By: **spoke**

◆ La
au

le los vehículos



35

¡386!

UNED



Ciclo de Conferencias del Máster y del
Programa de Doctorado en Ingeniería de
Sistemas y de Control

Arturo de la Escalera
Sistemas de Percepción para Vehículos Autónomos
28 de Octubre de 2020

◆ La percepción del entorno

- ❖ Tarea clave
- ❖ Tarea más compleja



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



◆ Los entornos de tráfico son:

- ❖ Muy estructurados desde el punto de vista de la infraestructura
- ❖ Poco estructurados desde el de los participantes y situaciones



- ◆ **Detección**
- ◆ **Clasificación**
- ◆ **Seguimiento**
- ◆ **Segmentación semántica**



◆ Sensores

❖ Exteroceptivos

✓ Captan el entorno alrededor del vehículo

- Ultrasonidos
- Radar
- Cámaras
- LiDAR

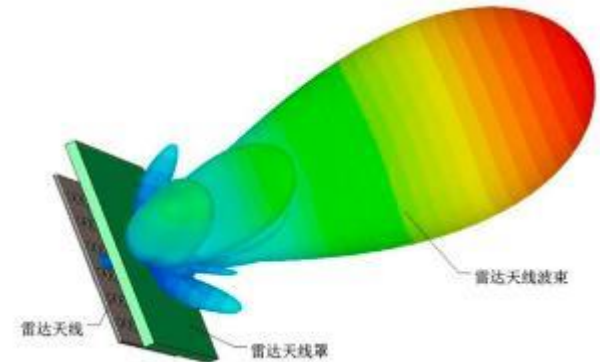
❖ Propioceptivos

✓ Miden variables internas

- GNSS
- IMU

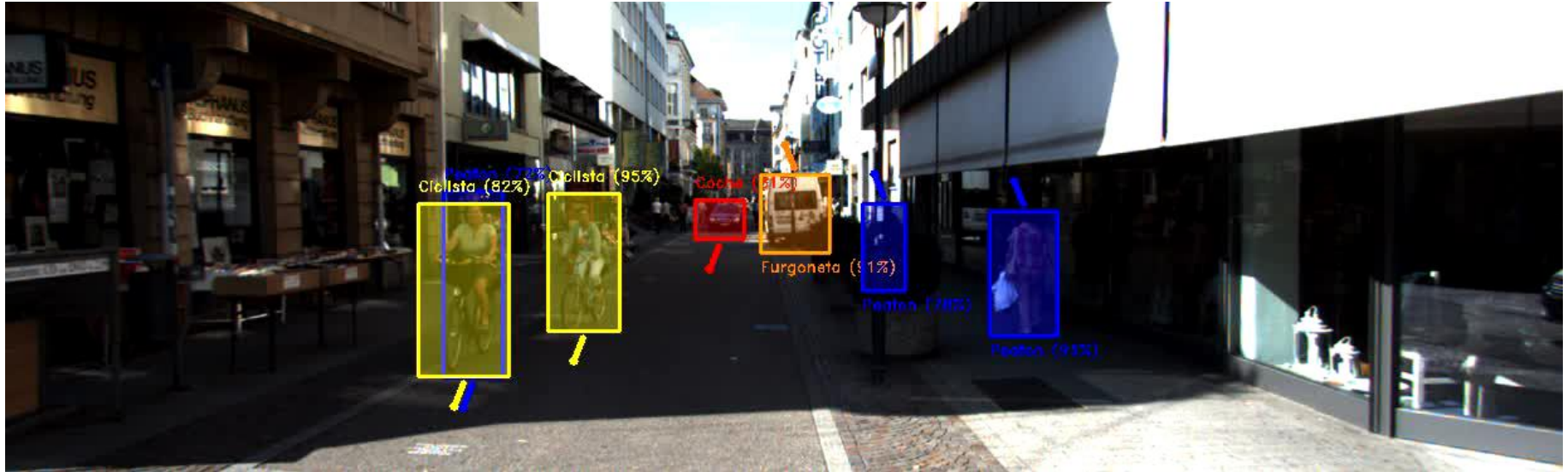


- ◆ **Radio detection and ranging**
- ◆ **Usan ondas de radio**
 - ❖ Frecuencias 24/77 GHz
 - ❖ millimetre wave radar (MMW)
- ◆ **Tiempo de vuelo**
- ✓ **Detección de múltiples objetivos**
- ✓ **Inmunes a cambios de iluminación y lluvia**
- ✓ **Distancia, velocidad y dirección de movimiento**
- ☒ **Escasa información sensorial**
- ☒ **Poca resolución lateral**



- ◆ Percibir la luz reflejada por los objetos (espectro visible) o el calor (infrarrojo lejano o térmicas)
- ◆ Información sensorial muy rica
- ◆ Gran influencia de las condiciones atmosféricas
- ◆ Resolución muy variable
- ◆ Gran diversidad de campo de vista
- ◆ Importancia del rango dinámico





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.

IR Pedestrian Detection
Cold Weather



www.uc3m.es/islab

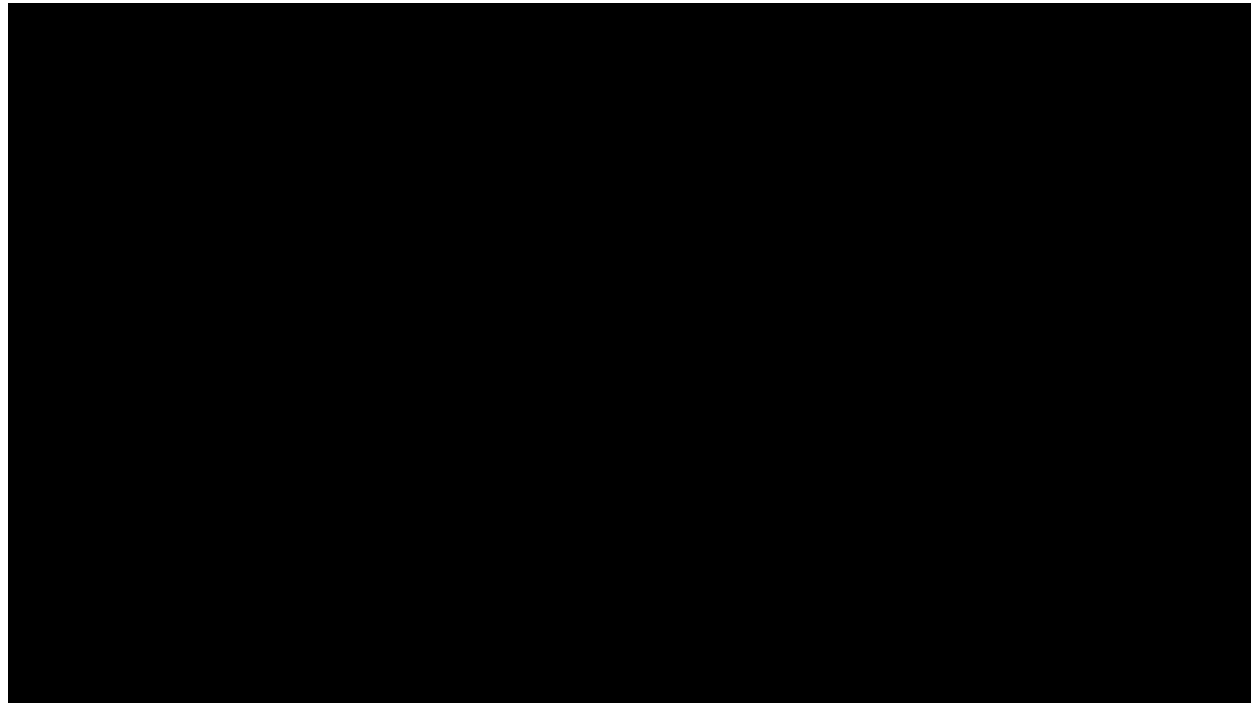
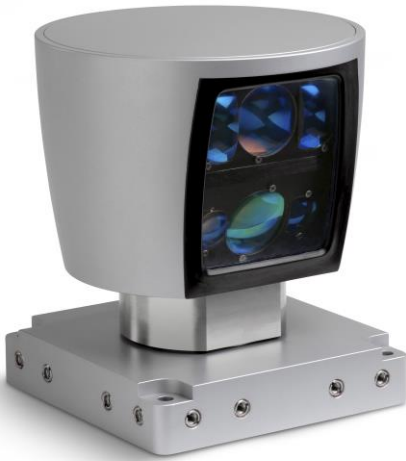
IR Pedestrian Detection
Hot Weather



www.uc3m.es/islab

D. Olmeda, C. Premebida, U. Nunes, J.M. Armingol and A. de la Escalera.
Pedestrian Detection in Far Infrared Images. Integrated Computer-Aided
Engineering, Vol: 20 (4), pp: 347-360, 2013

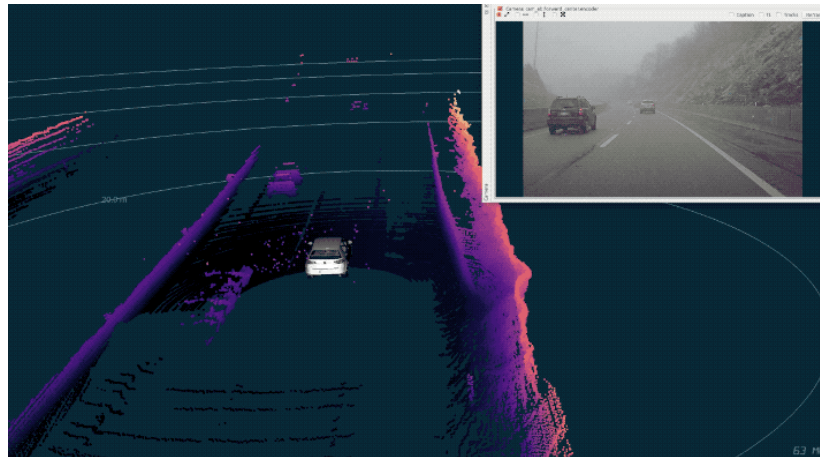
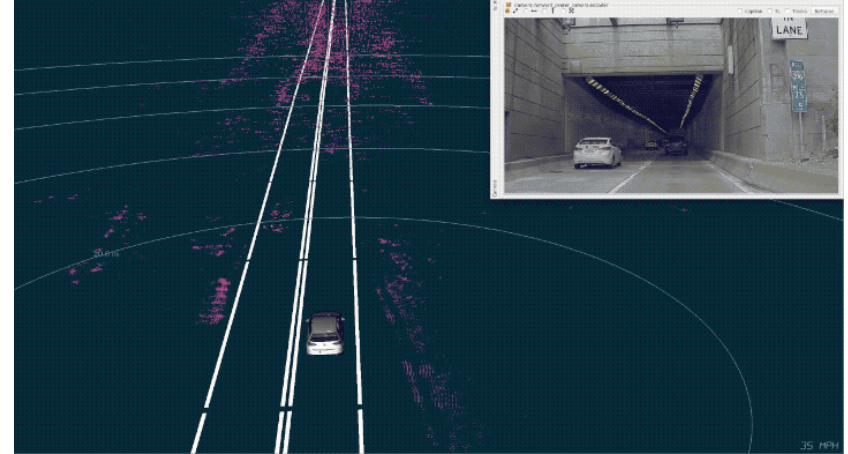
- ◆ Con espejos y motores se obtienen medidas 3D
- ◆ La intensidad del pulso retornado puede ayudar a clasificar el objeto



Dificultades de la percepción

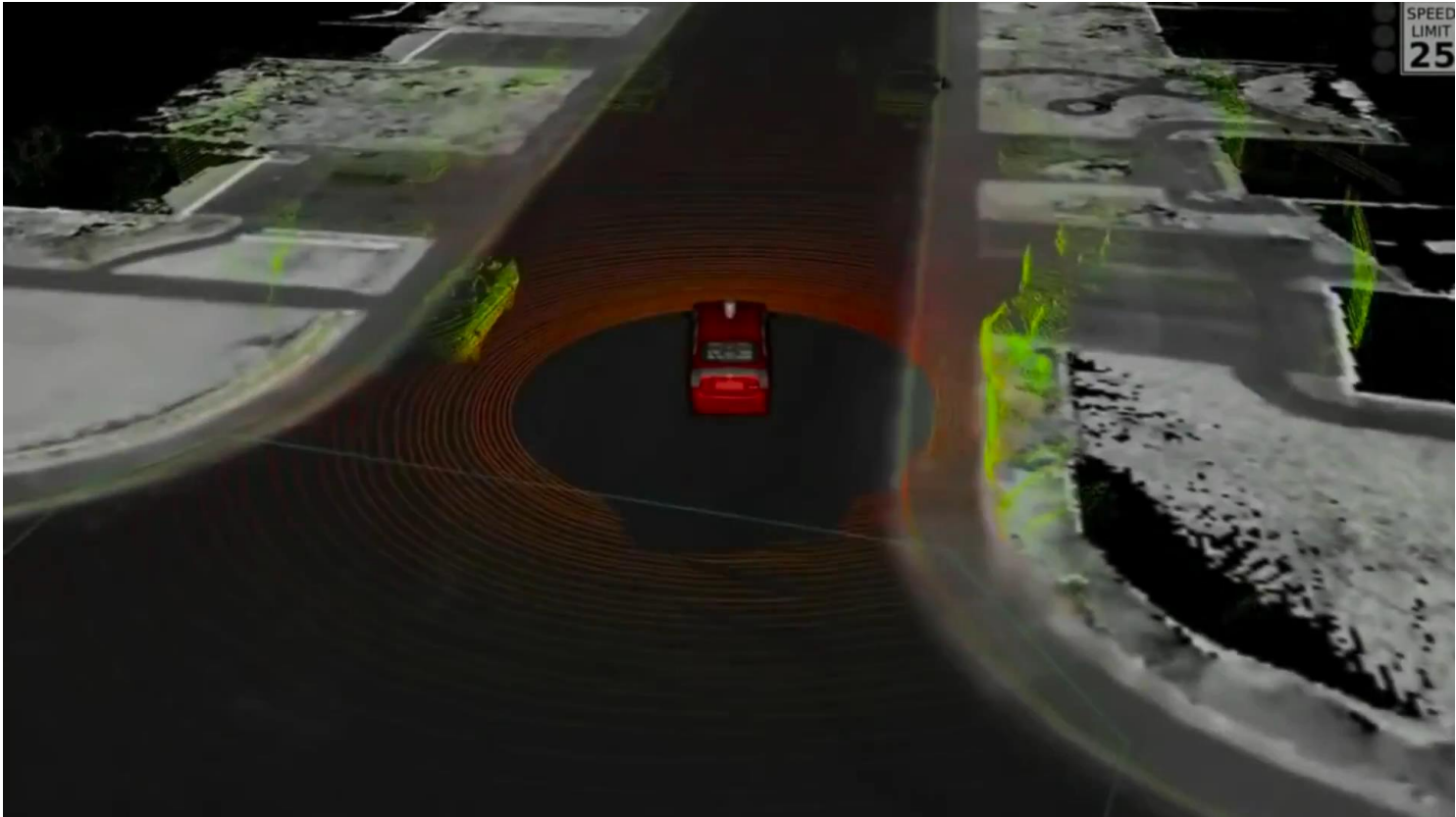


Dificultades de la percepción



Fuente: <https://medium.com/@curmson>

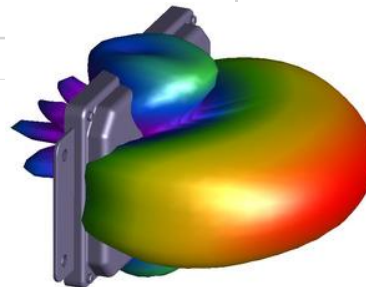
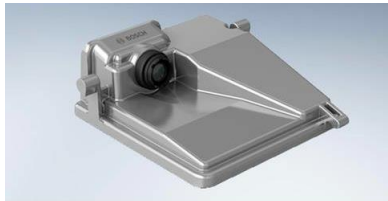


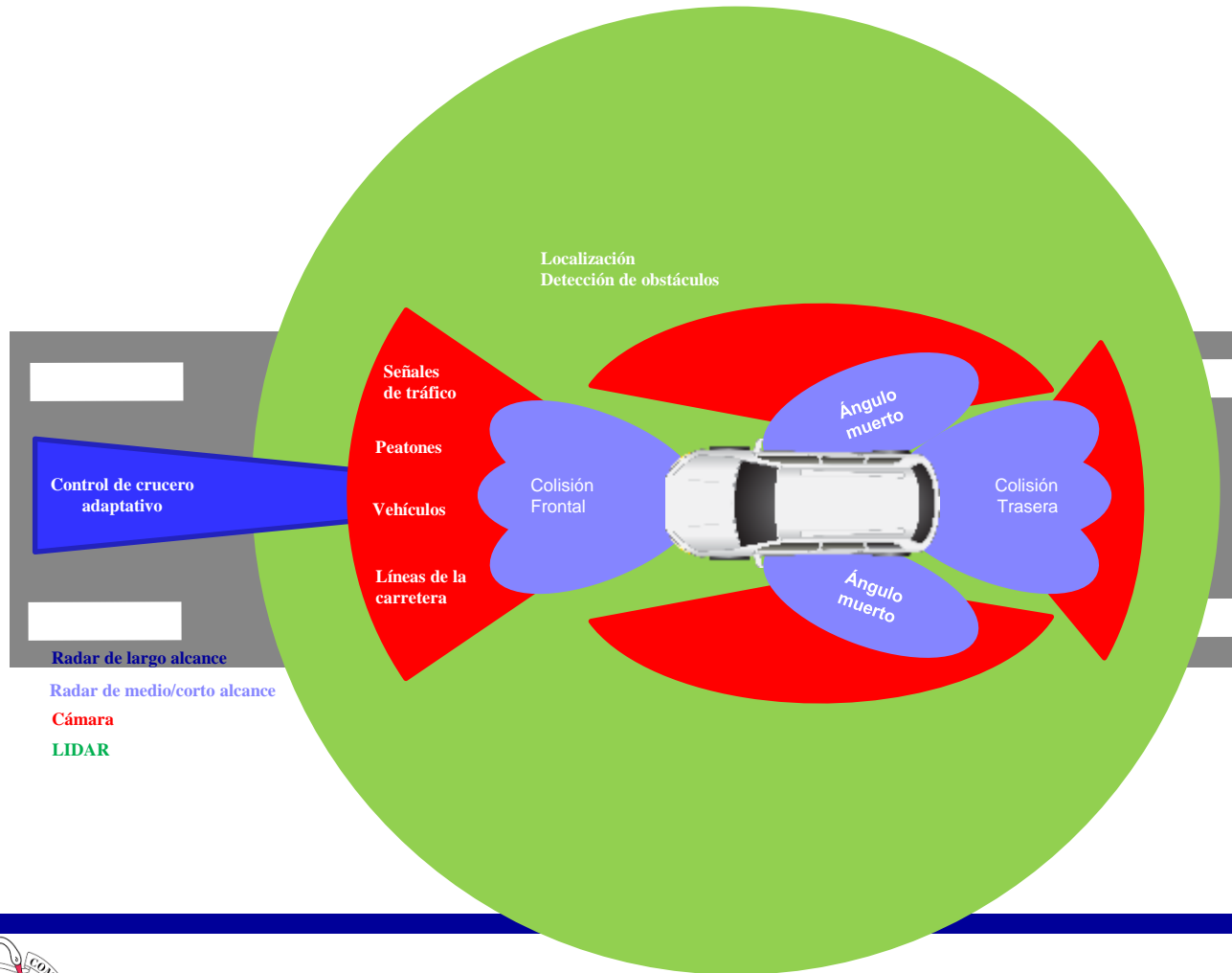


<https://youtu.be/weXDUc5Osto>

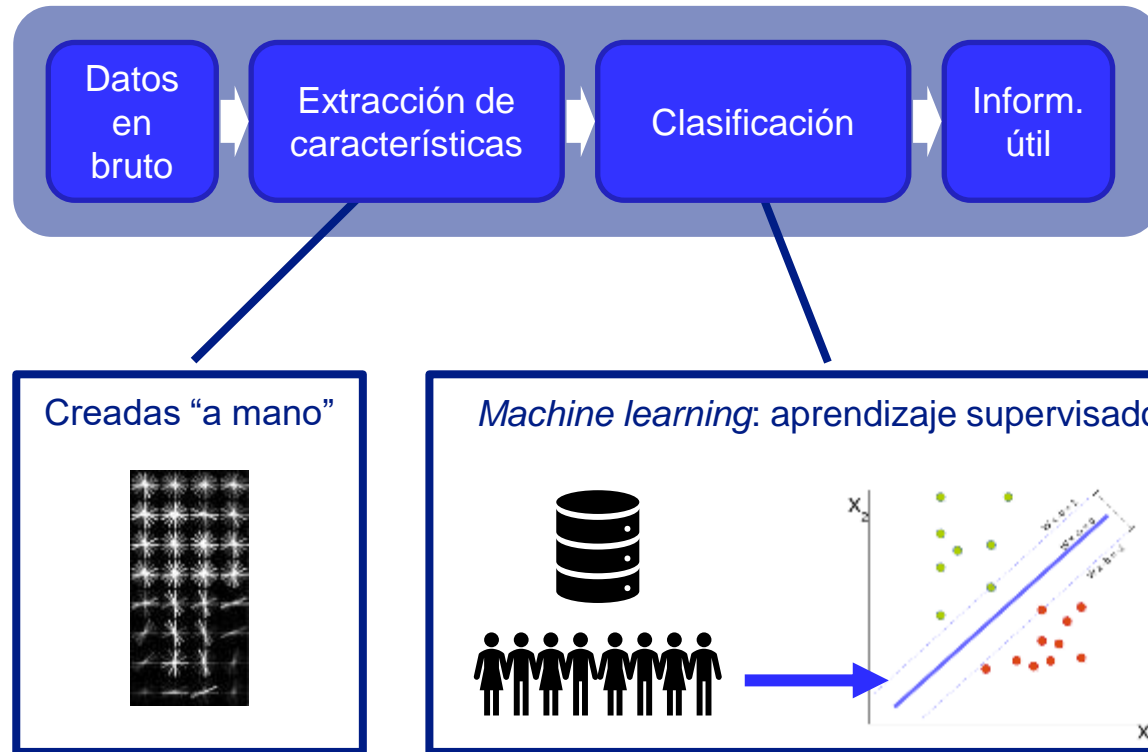
Waymo: encountering an elderly lady chasing a duck in a wheel chair

	Funcionamiento en mal tiempo	Funcionamiento con mala luminosidad	Detección obstáculos	Clasificación de obstáculos
Cámaras	Red	Yellow	Yellow	Green
Radar	Green	Green	Yellow	Red
LiDAR	Yellow	Green	Green	Yellow
Bueno	Green			
Regular	Yellow			
Insuficiente	Red			

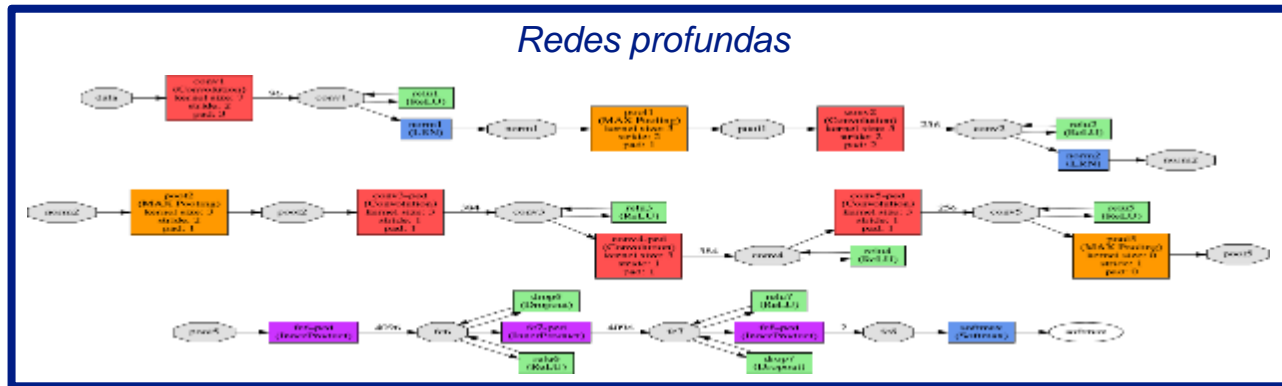
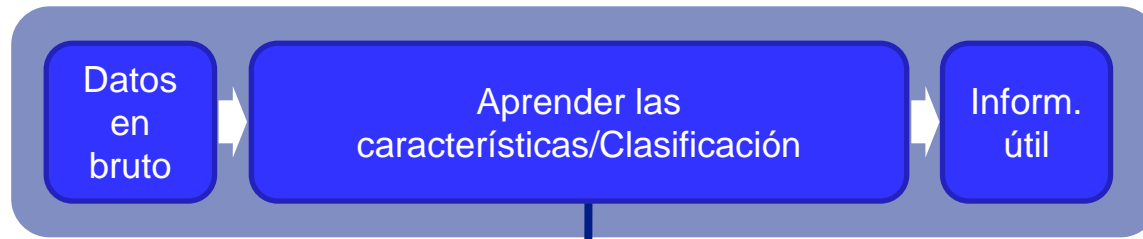




- ◆ **Presentación del LSI**
- ◆ **Entendimiento del entorno viario**
- ◆ **Percepción del entorno**
 - ❖ Detección espacio libre
 - ❖ Odometría visual
 - ❖ Detección de carriles
 - ❖ Reconocimiento de objetos
 - ❖ Segmentación semántica
 - ❖ Obtención información 3D monocular
- ◆ **Conclusiones**



Aprendizaje profundo: *deep learning*



- ◆ Hirschmuller, H.: Stereo Processing by Semiglobal Matching and Mutual Information, IEEE T on PAMI 30, 328-341 (2008)

Cámara
Izquierda



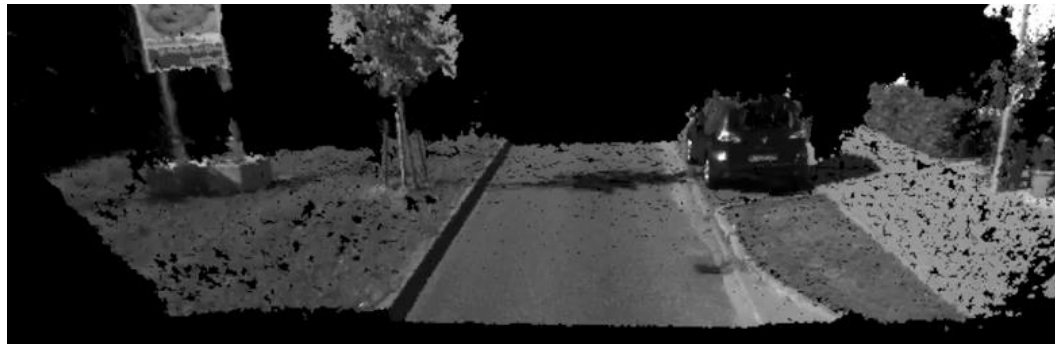
Imagen de
Disparidades



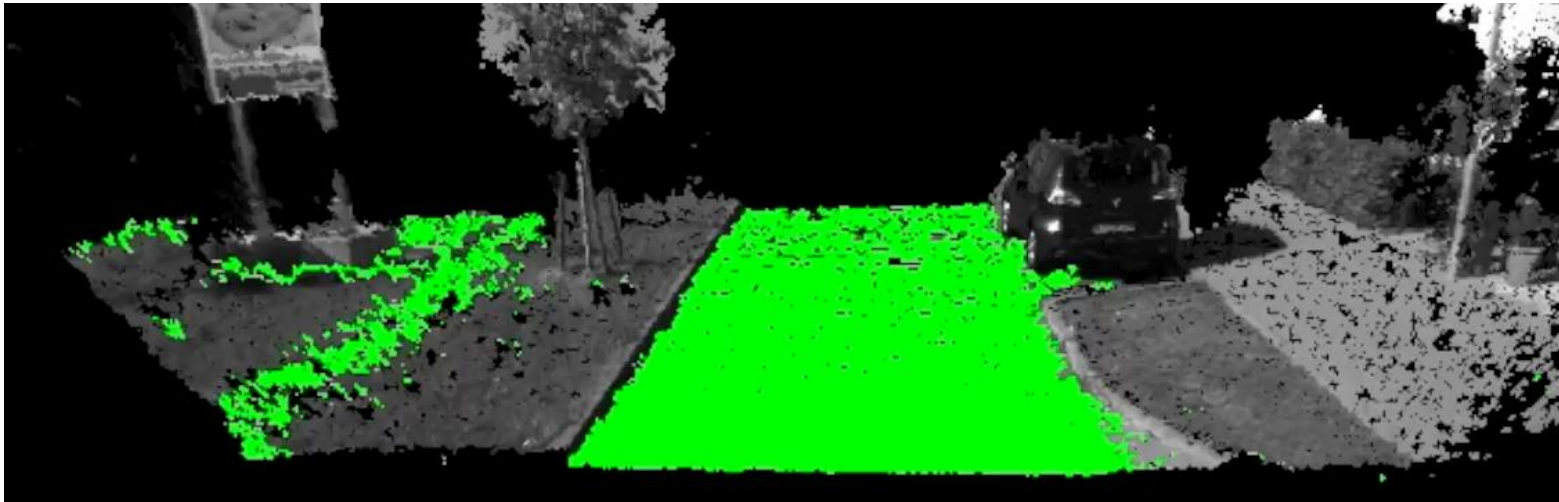
Datos 3D



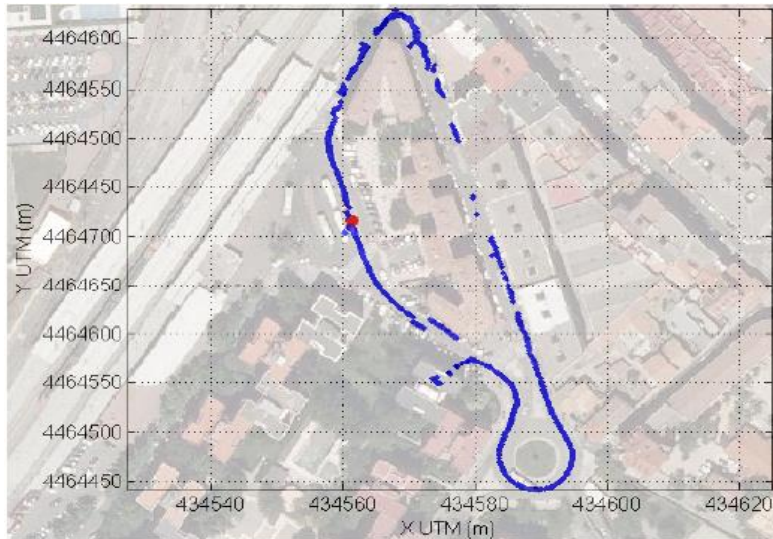
Max Dist 20m
Min Dist 6m
Anchura 2m



- ◆ A partir de la nube de puntos, se obtiene el plano de la carretera
 - ❖ El algoritmo detecta un plano perpendicular a un eje dentro de una desviación angular máxima especificada
 - ✓ Ransac

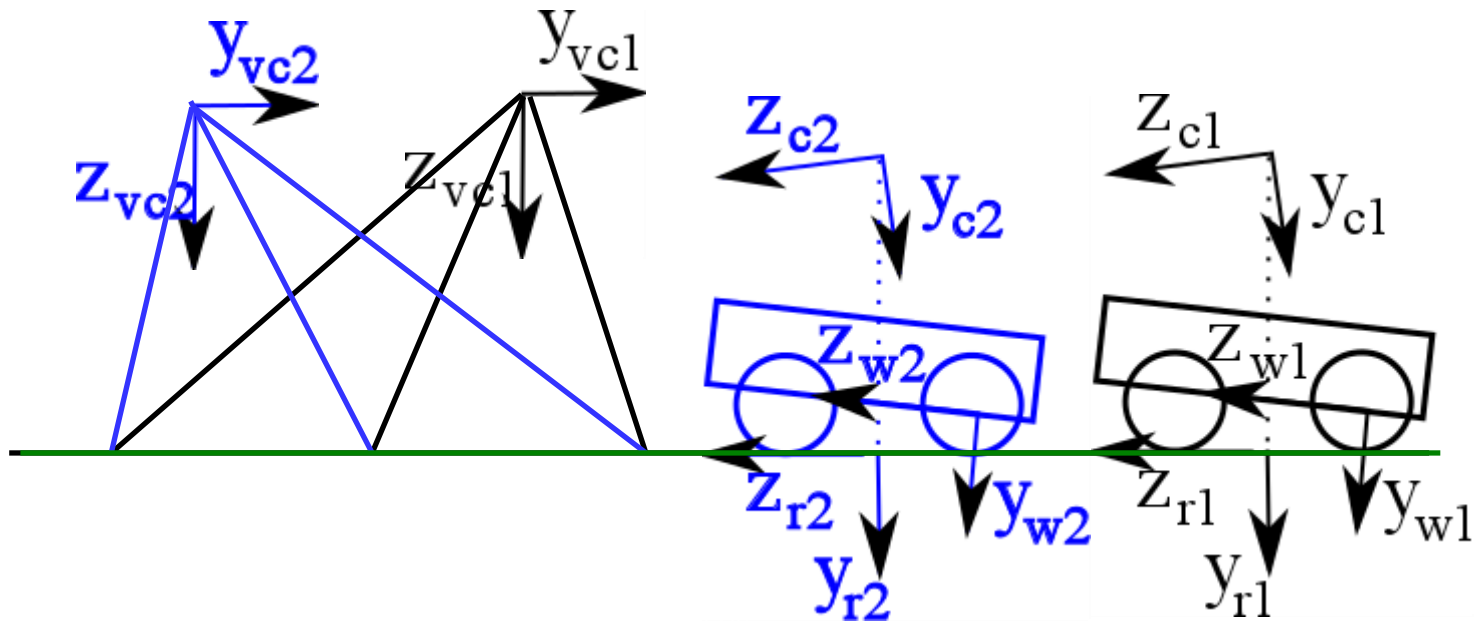


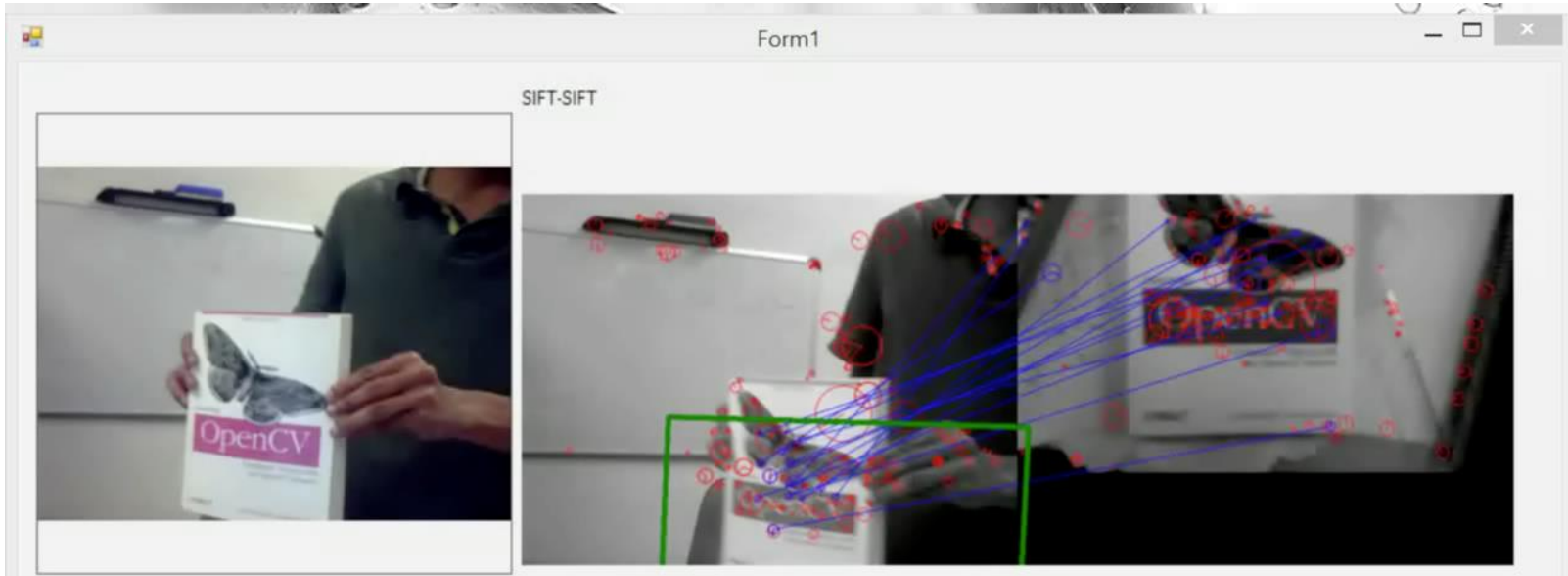
◆ Problemas del GPS en entornos urbanos

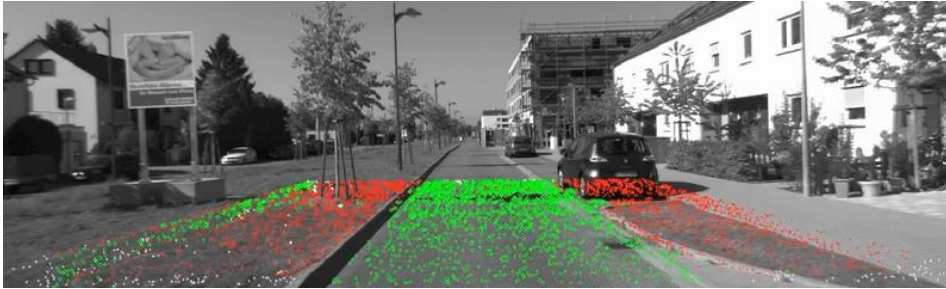


- ◆ **Estimar el movimiento de un vehículo utilizando sólo información visual**
- ◆ **Tres pasos:**
 - ❖ Detectar puntos de interés en cada imagen y emparejarlos entre dos consecutivas
 - ❖ Encontrar y eliminar las coincidencias incorrectas
 - ❖ Estimar el movimiento relativo de la cámara

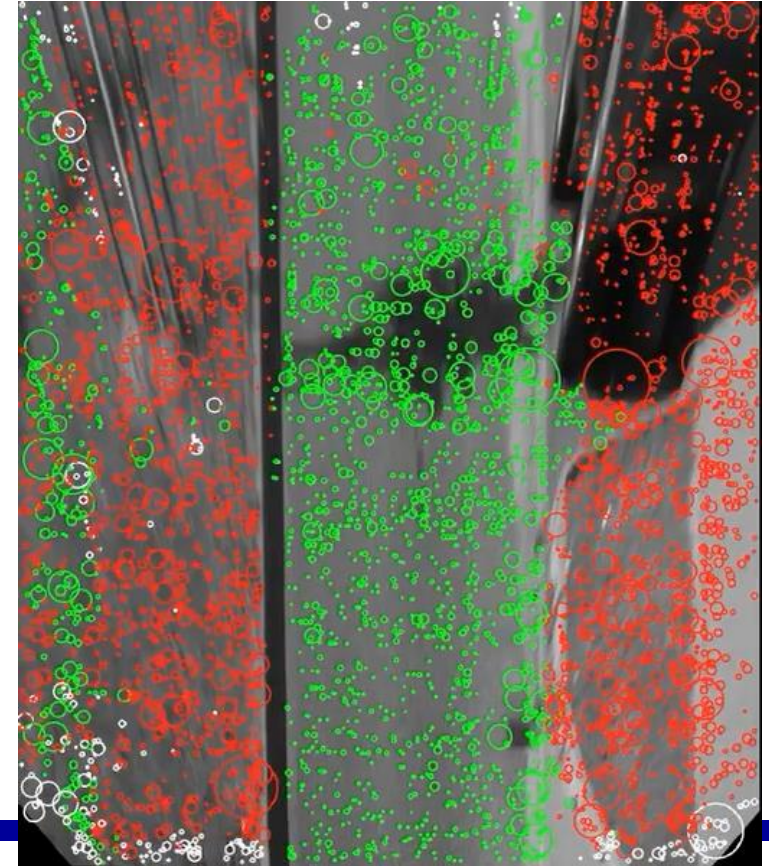
Odometría visual

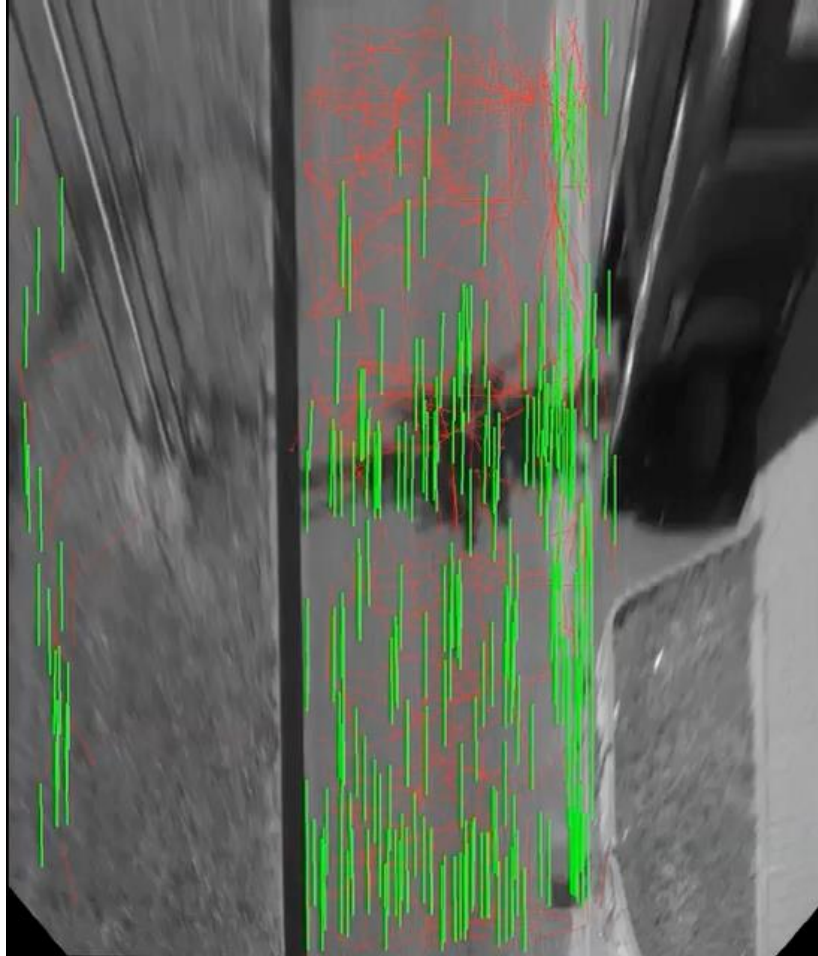






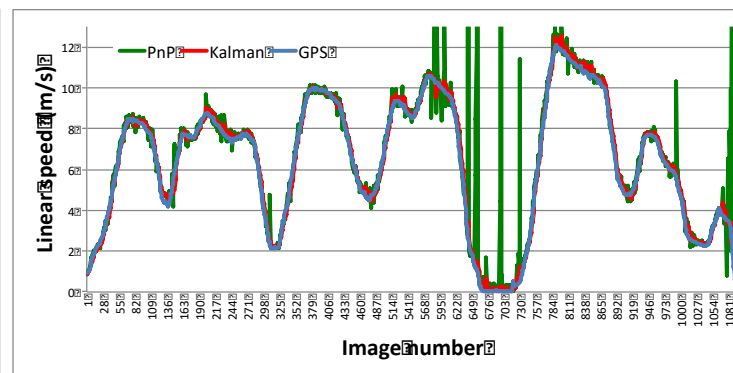
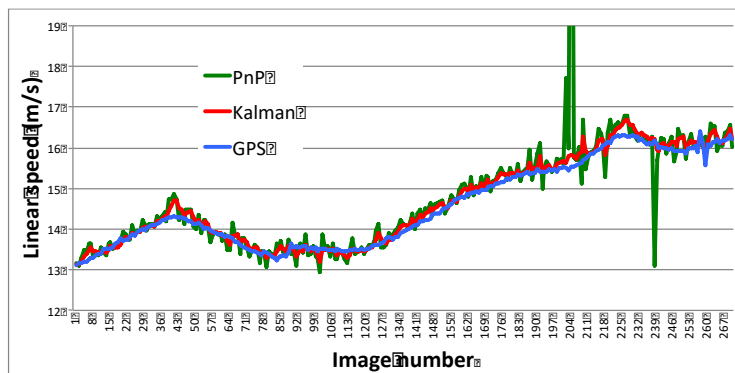
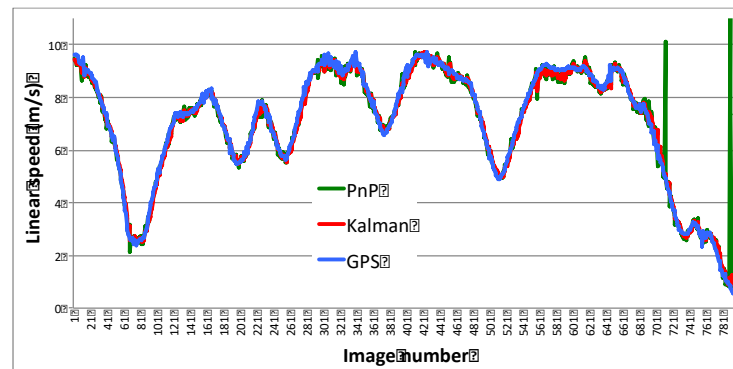
- ◆ Se detectan 6000 características SIFT en la imagen de vista de pájaro
- ◆ Algunas características corresponden a áreas texturizadas de la carretera
- ◆ Otros están relacionados con la proyección de coches, peatones, edificios, etc.
- ◆ Se revisa si los puntos pertenecen al suelo o no

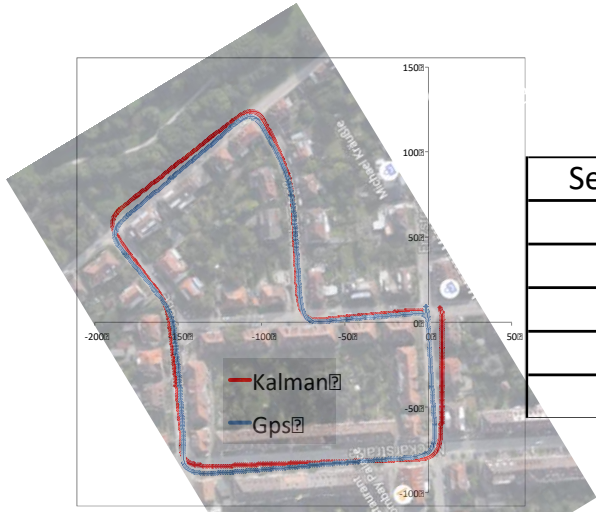




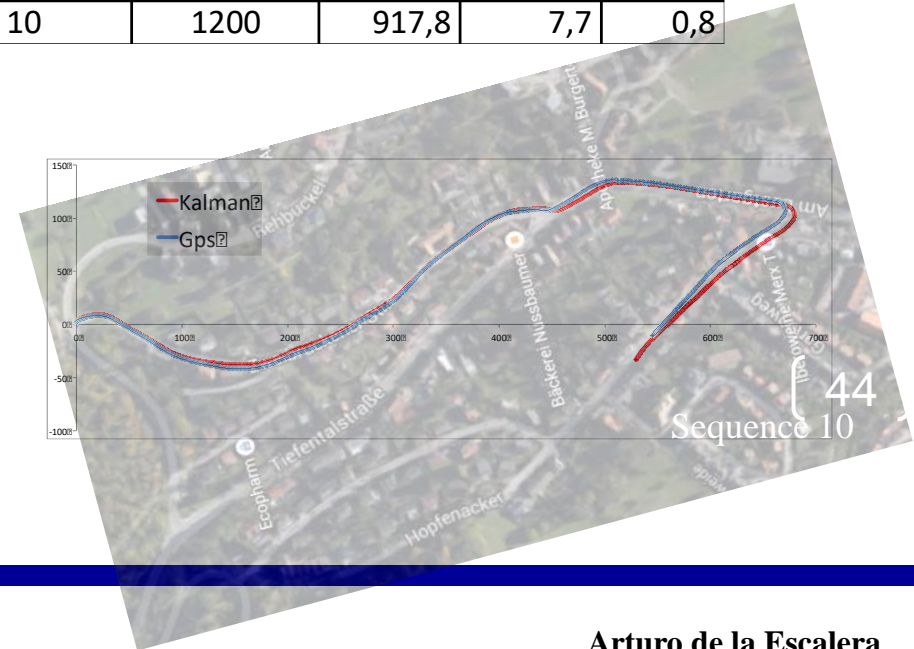
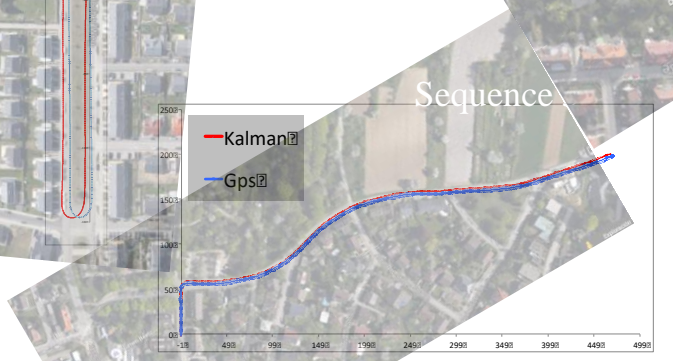
Root Mean Square error

Sequence	Num.ima	linear speed	angular speed
		RMS	RMS
3	800	0,181	0,024
4	270	0,151	0,009
6	1100	0,301	0,027
7	1100	0,337	0,038
10	1200	0,271	0,015

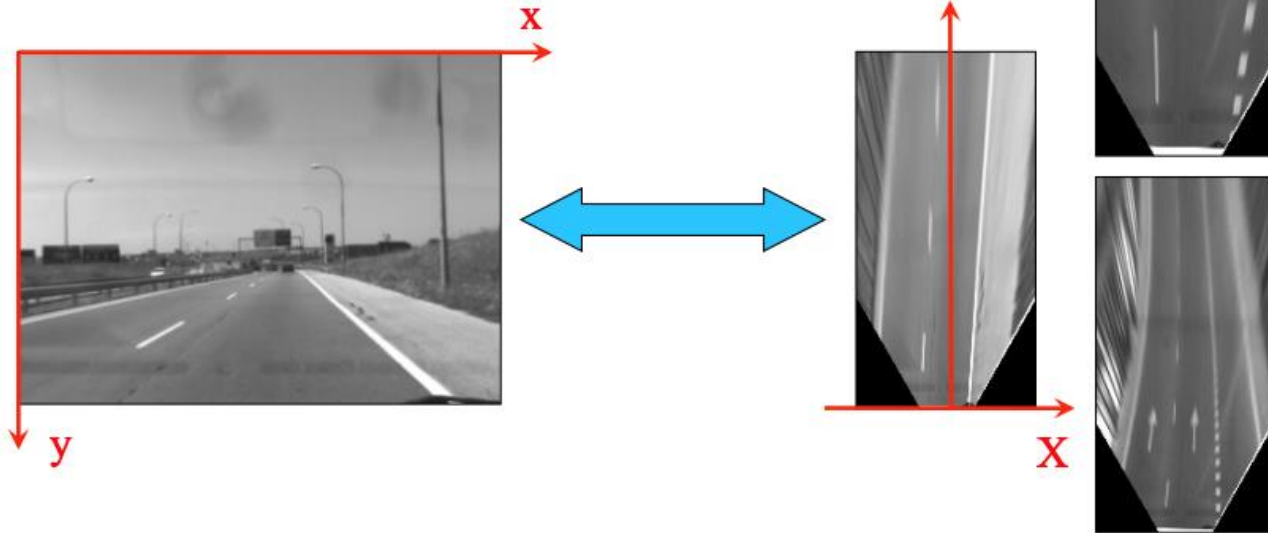




Sequence	Num.ima	Dist. (m)	Error(m)	Error%
3	800	558,9	3,5	0,6
4	270	393,6	3,2	0,8
6	1100	1229,4	8,7	0,7
7	1100	694,4	8,8	1,3
10	1200	917,8	7,7	0,8

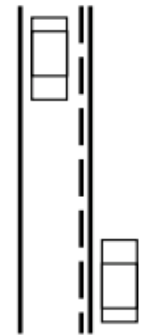
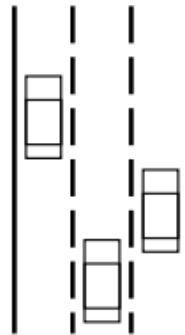
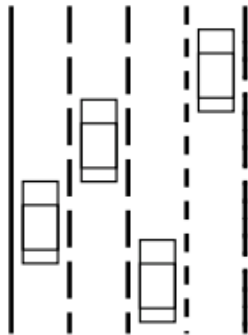
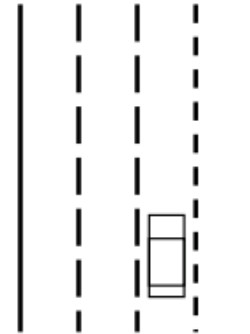


- Carriles
 - Perspectiva inversa

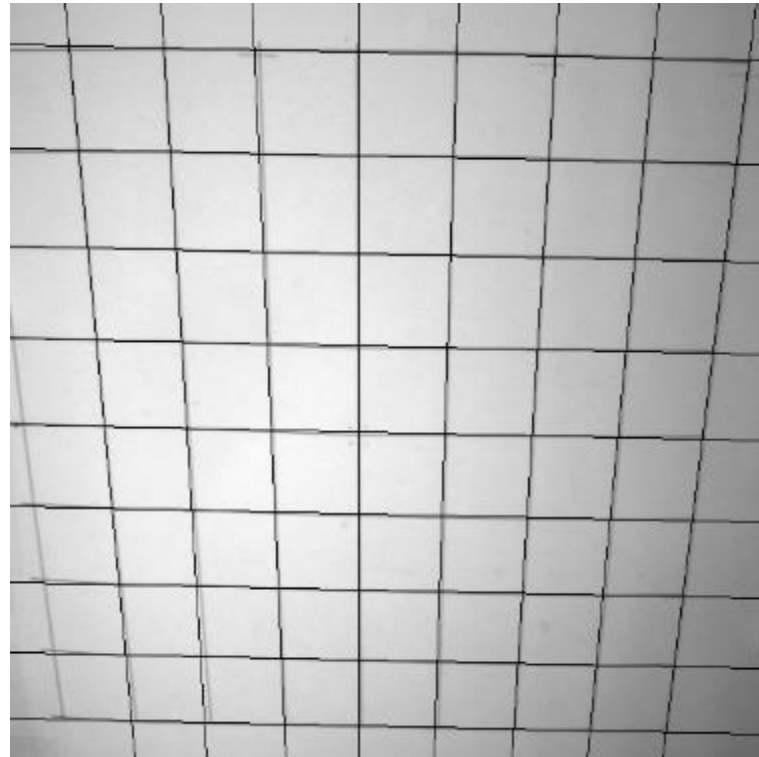
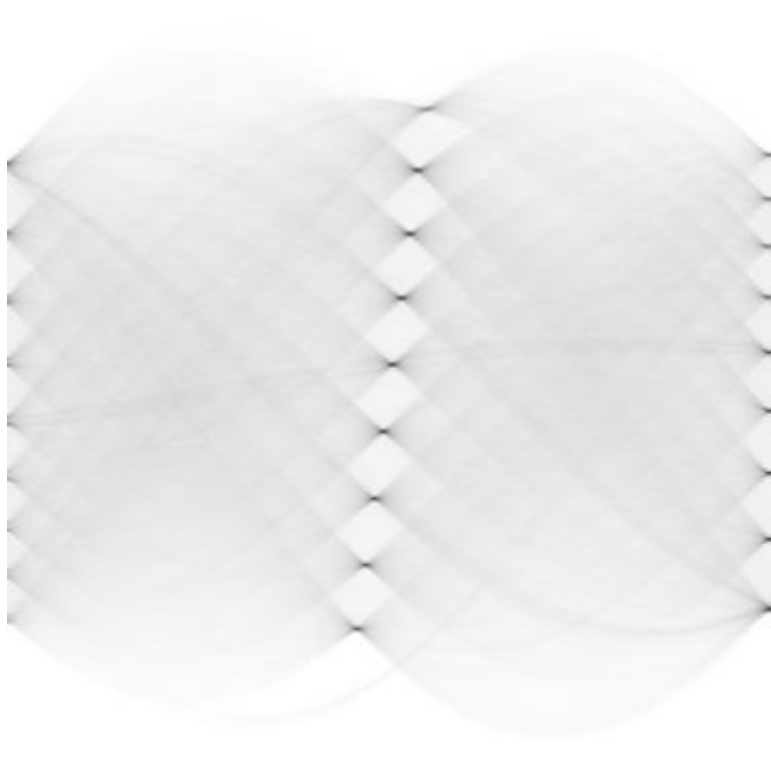


Detección de carriles


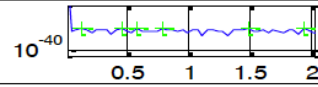
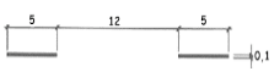
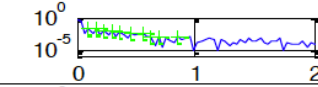

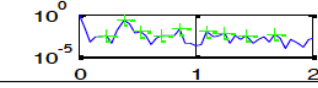
- Reconocer los tipos de líneas
 - Continua, discontinua, incorporación, arcén, ...
- Identificar tipo vía y nº carriles
 - Modelo adaptable al tipo de vía

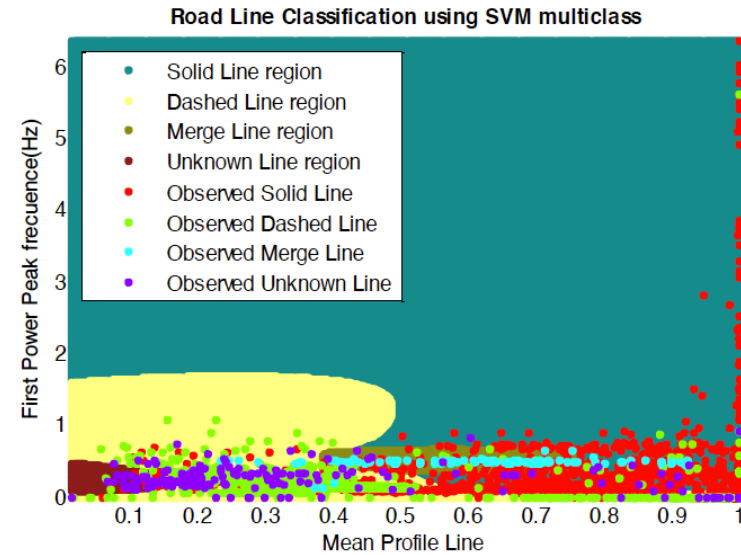
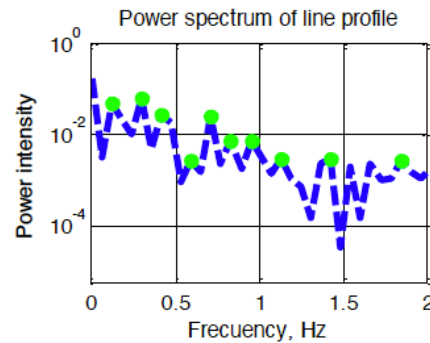
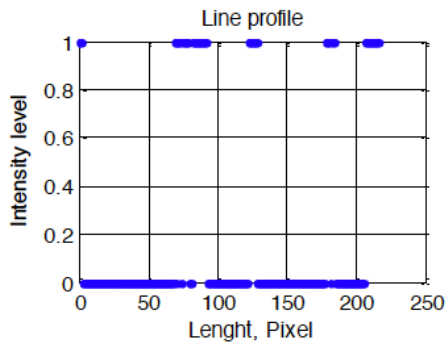


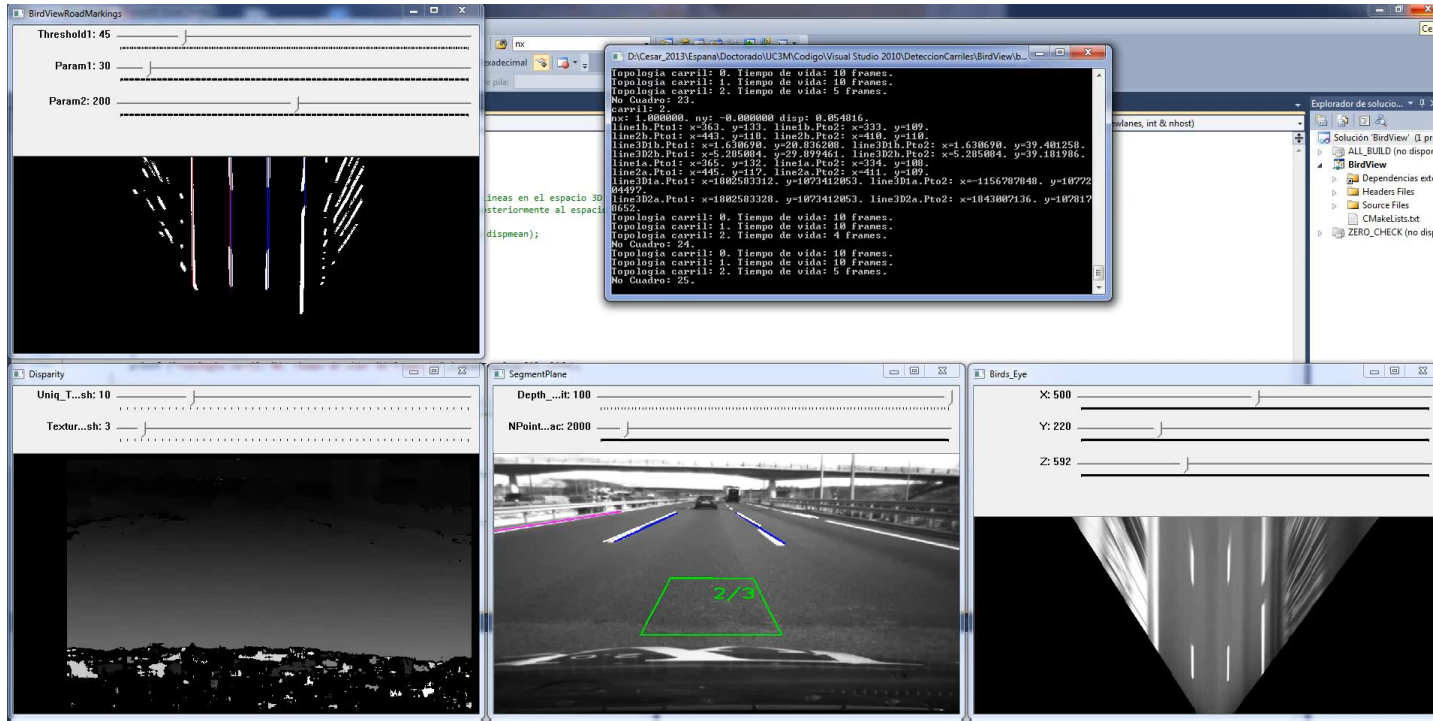
Transformada de Hough



Detección de carriles

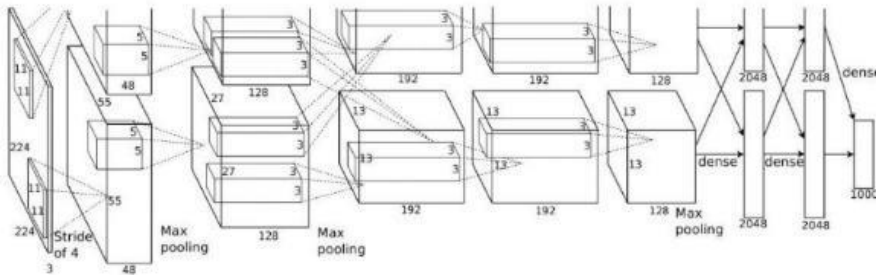
Line type	Spatial representation	Frequency representation
Solid		
Dashed		
Merge		





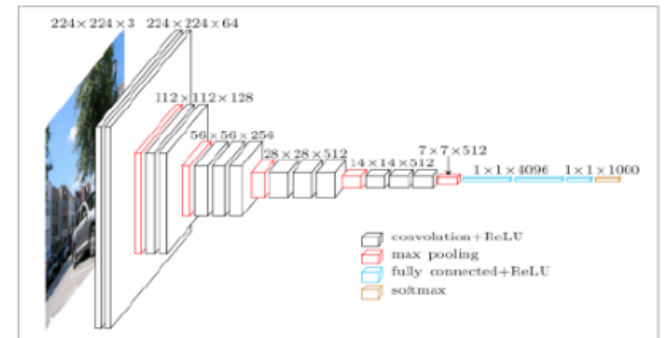
◆ AlexNet (2012)

- ❖ 650,000 neuronas
- ❖ 5 capas de convolución
- ❖ 60 millones de parámetros



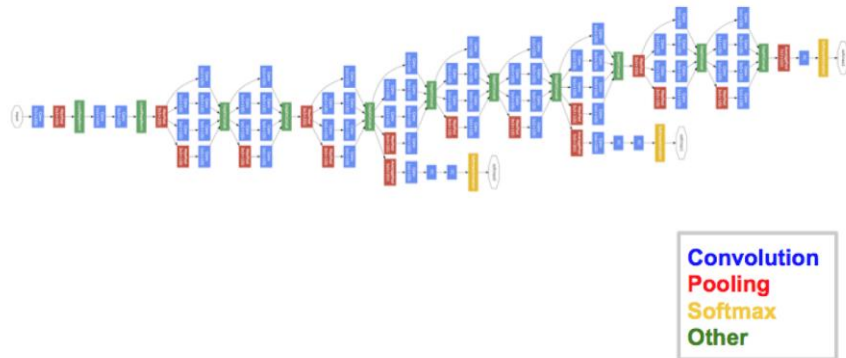
◆ VGGNet (2014)

- ❖ 16 capas convolucionales
- ❖ Entrenada en 4 GPUs durante 2–3 semanas
- ❖ 138.357.544 parámetros



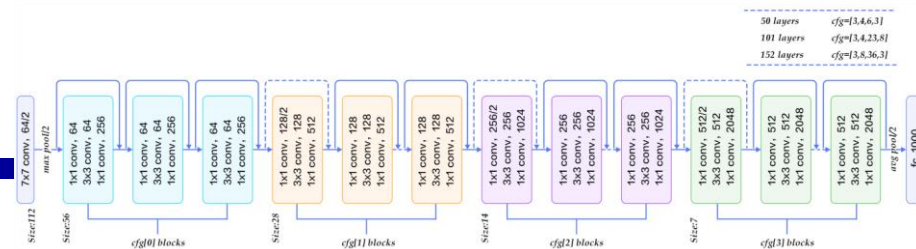
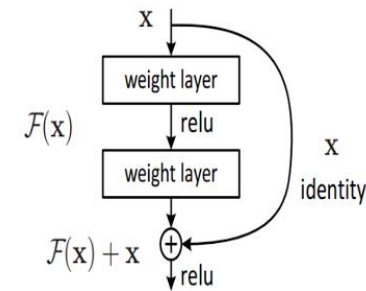
◆ Inception (2014 y 2016)

- ❖ El modelo se compone de una unidad básica denominada "Inception cell"
- ✓ Se realizan una serie de convoluciones a diferentes escalas y, posteriormente, se agregan los resultados



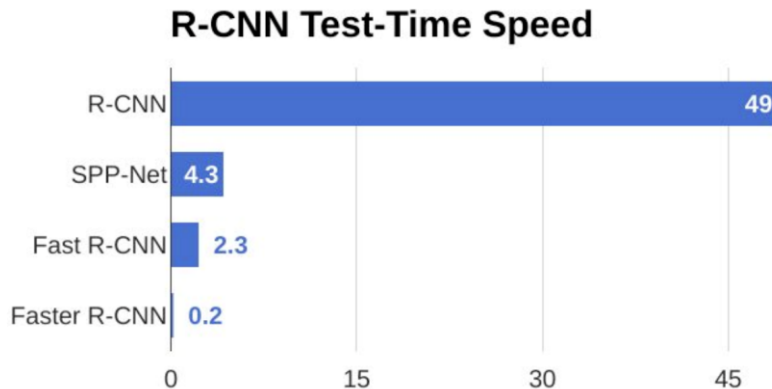
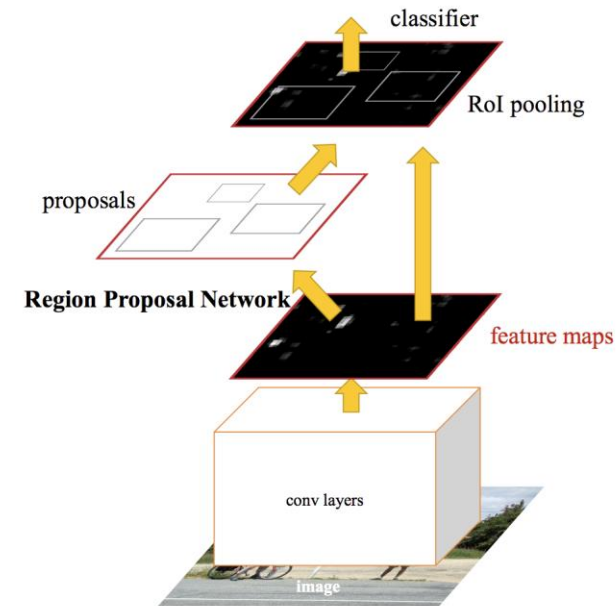
◆ Residual Neural Network (ResNet) (2015)

- ❖ Se aumenta el número de capas introduciendo una conexión residual (con una capa identidad).
- ✓ Esta capa pasa a la siguiente directamente, mejorando el proceso de aprendizaje.
- ✓ 152 capas teniendo menor complejidad que VGGNet
- ✓ ILSVRC 2015, tasa de error del 3,57%
- Supera el rendimiento a nivel humano en este conjunto de datos



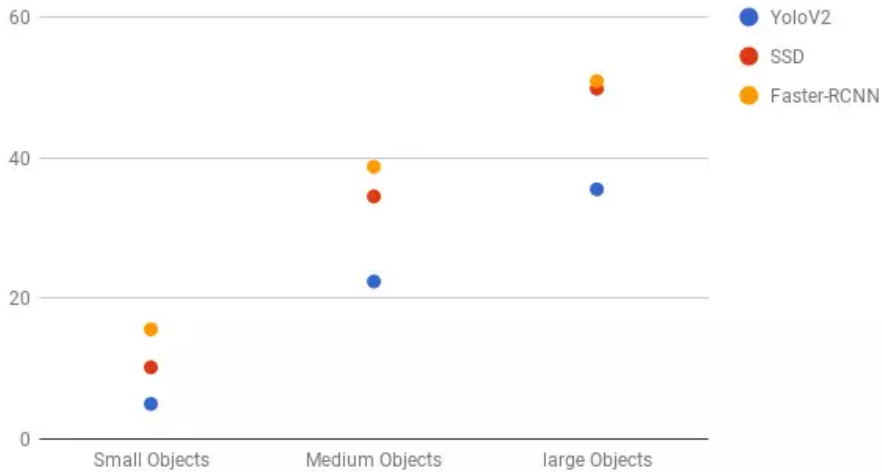
Faster R-CNN

- ◆ Heredera de la R-CNN y Fast R-CNN
- ◆ Se utiliza una red independiente para predecir las propuestas de región.
- ◆ Las propuestas de región pronosticadas se remodelan mediante una capa pooling



Comparativa

Accuracy



<https://cv-tricks.com>

Clasificación: estado del semáforo

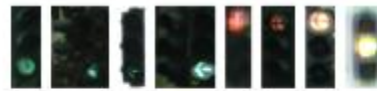


- Mapas digitales
- Posicionamiento usando visión

Regiones precisas



(a) Bosch



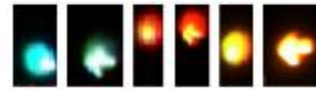
(b) Cityscapes



(c) LaRA



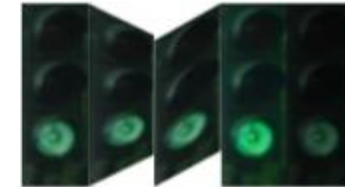
(d) LISA-Day



(e) LISA-Night



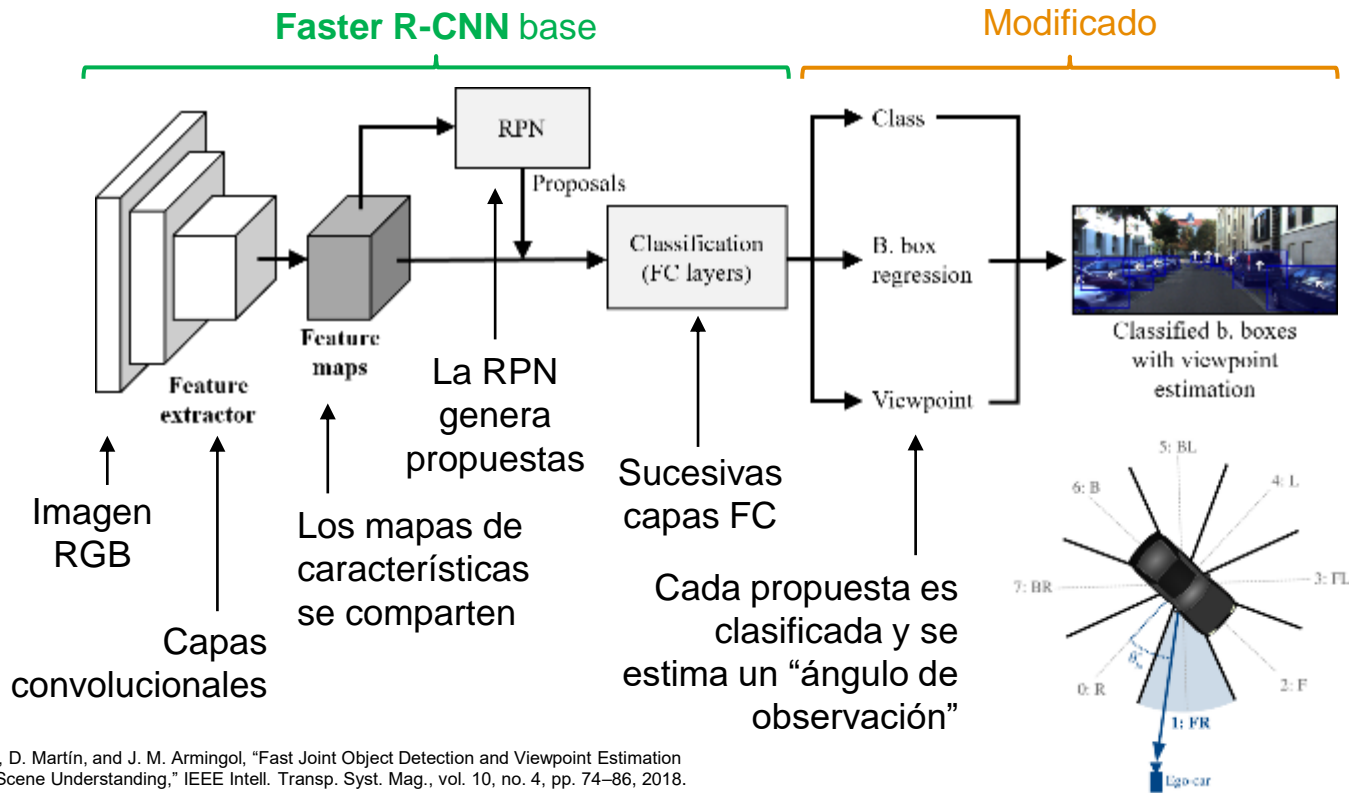
(f) WPI



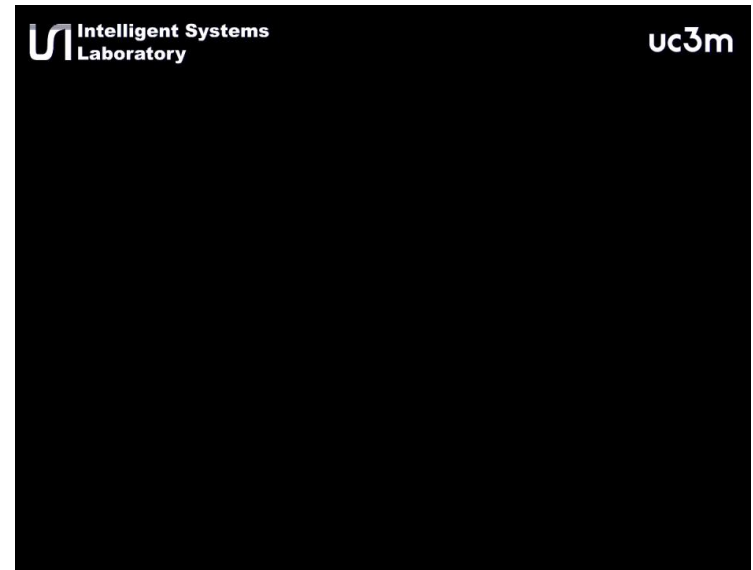
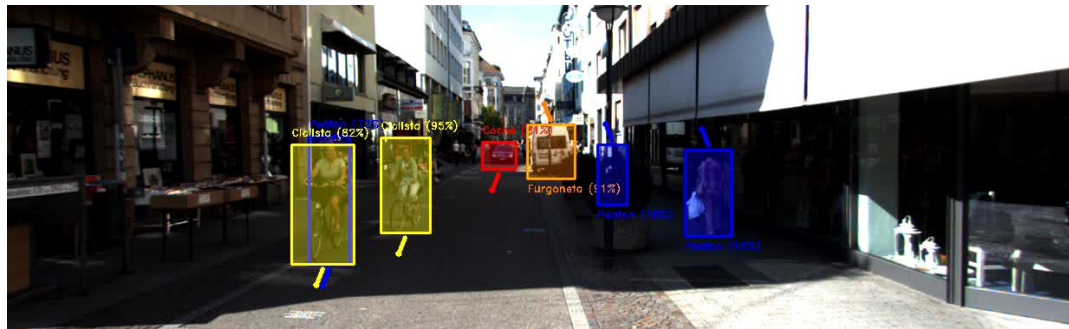
- Combinación de varias bases de datos + aumento de datos
- F1-score: **90,3%** (vs. 67,6% Random Trees)

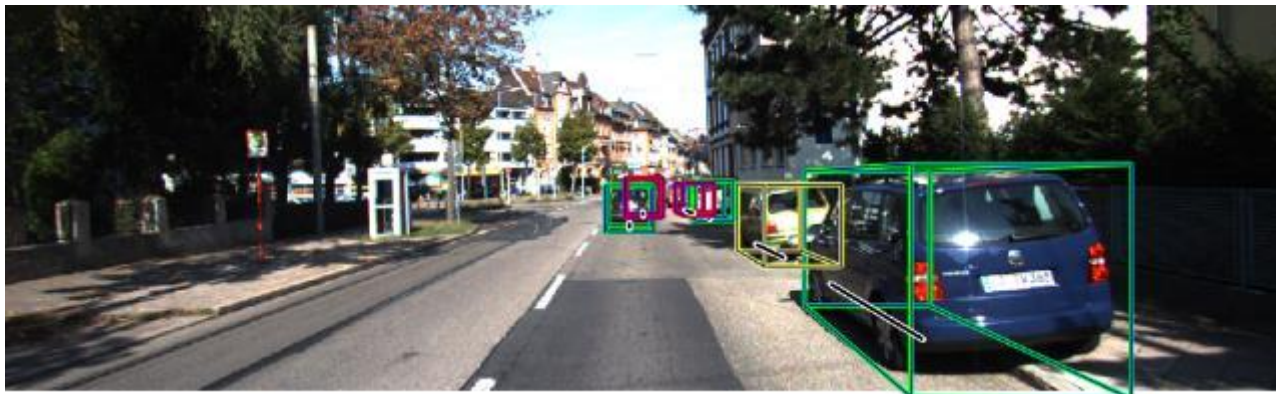
C. Fernández, C. Guindel, N.-O. Salscheider, and C. Stiller, "A Deep Analysis of the Existing Datasets for Traffic Light State Recognition," in Proc. IEEE International Conference on Intelligent Transportation Systems (ITSC), 2018, pp. 248–254

Detección 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.





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

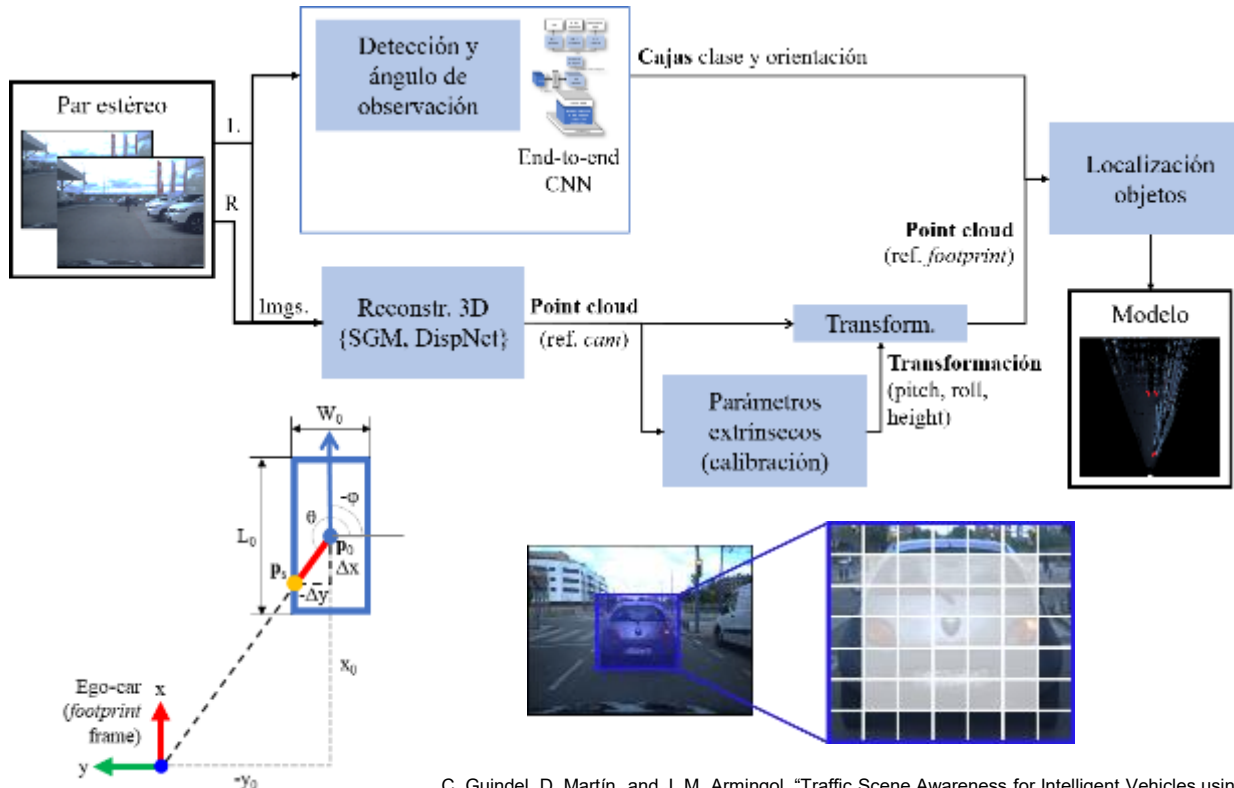


- ◆ ¿Y las dimensiones del objeto?
- ◆ ¿Y su orientación en coordenadas del mundo?

Distancia
¿Punto más cercano?
¿Centroide?

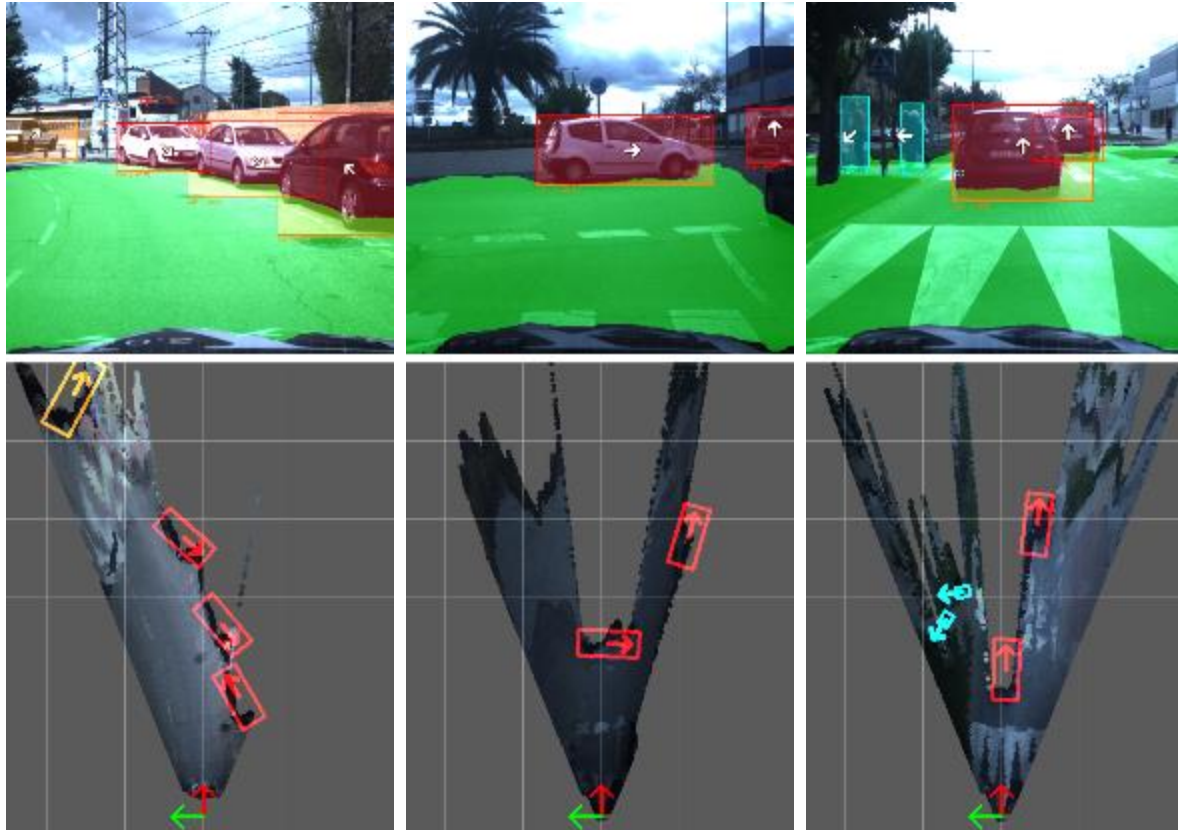
No hay medidas de distancia para las zonas no visibles por el sensor

Localización 3D: estéreo

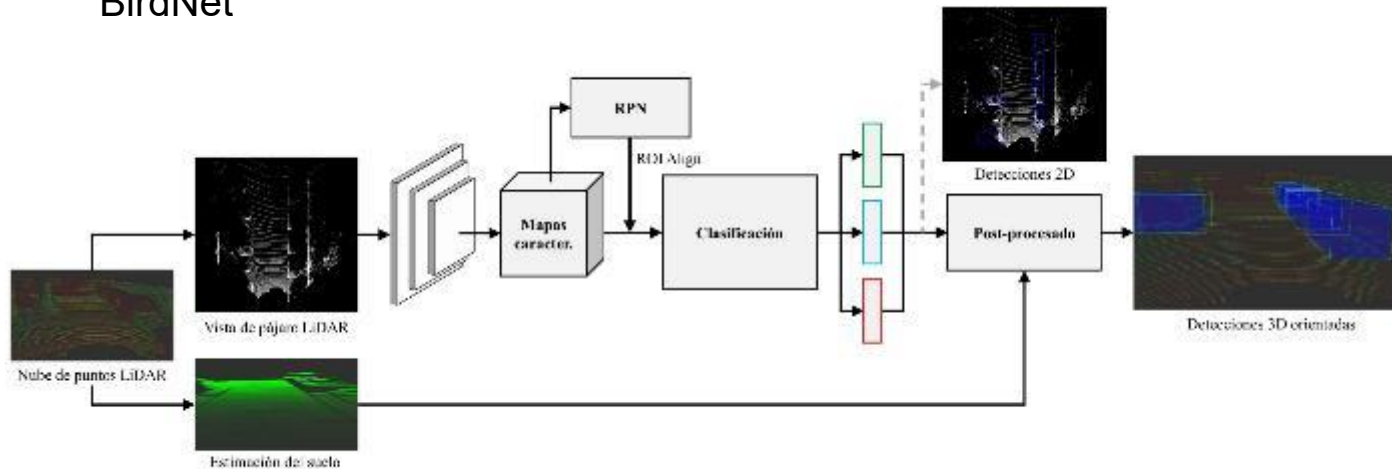


C. Guindel, D. Martín, and J. M. Armingol, "Traffic Scene Awareness for Intelligent Vehicles using ConvNets and Stereo Vision," *Rob. Auton. Syst.*, vol. 112, pp. 109–122, 2019.

Localización 3D: estéreo

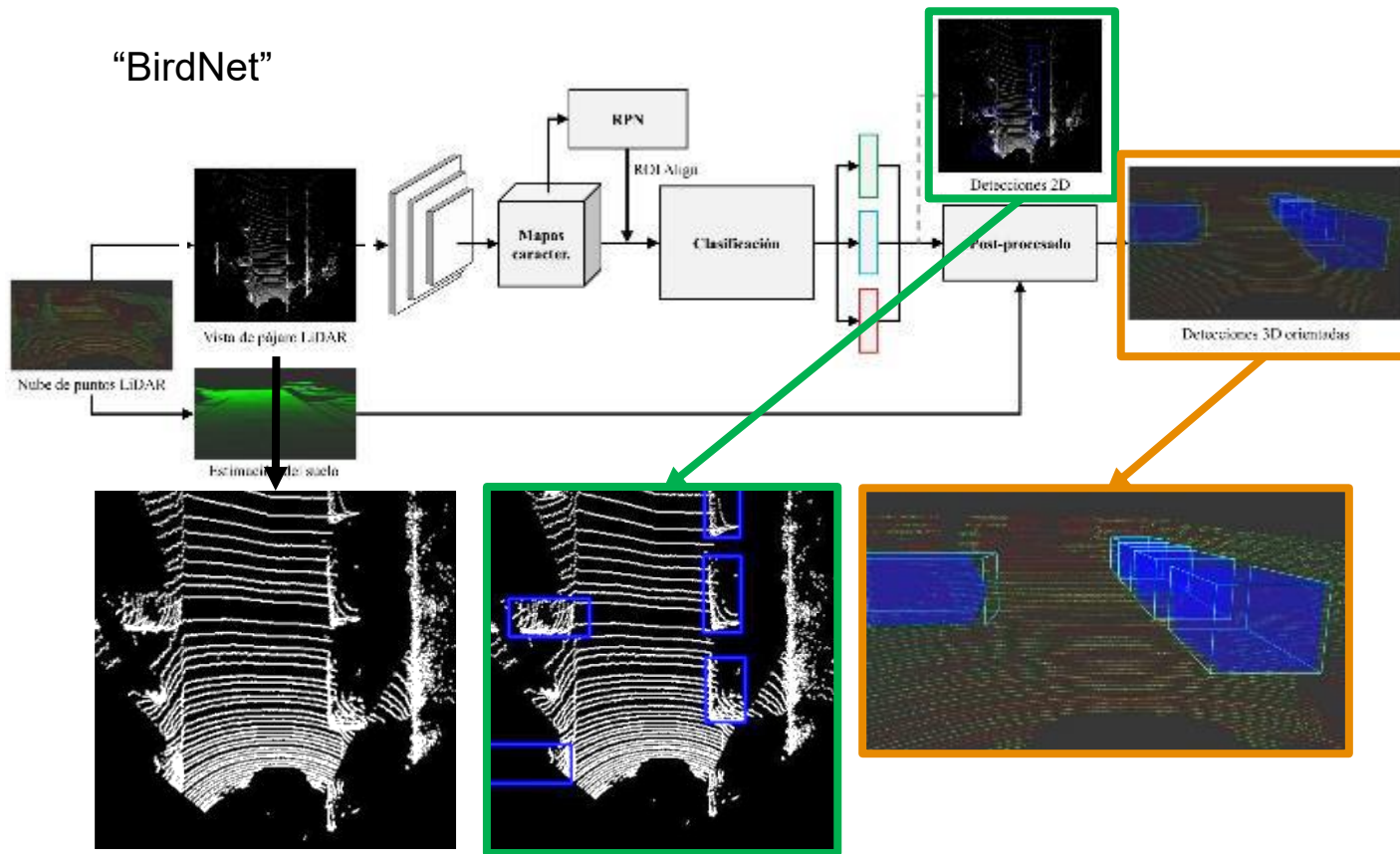


“BirdNet”



- ◆ 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.



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

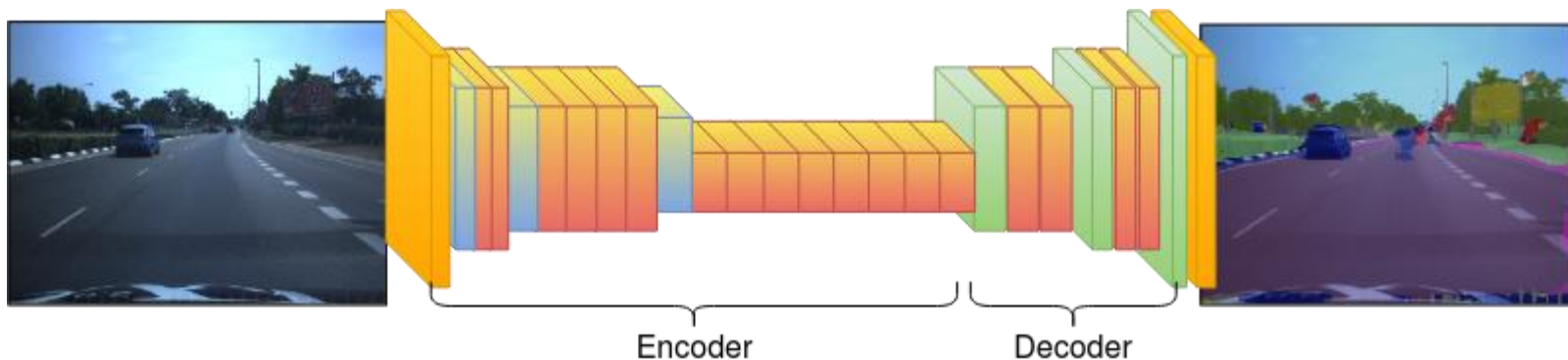


- ◆ **Asignación de una clase a cada píxel**
- ◆ **Reconoce obstáculos pero no los diferencia entre sí**
 - ❖ Los métodos “de instancias” sí (aplicando detección)
- ◆ **Especialmente útil para elementos de la infraestructura:**
 - ❖ Carretera, carriles, etc.



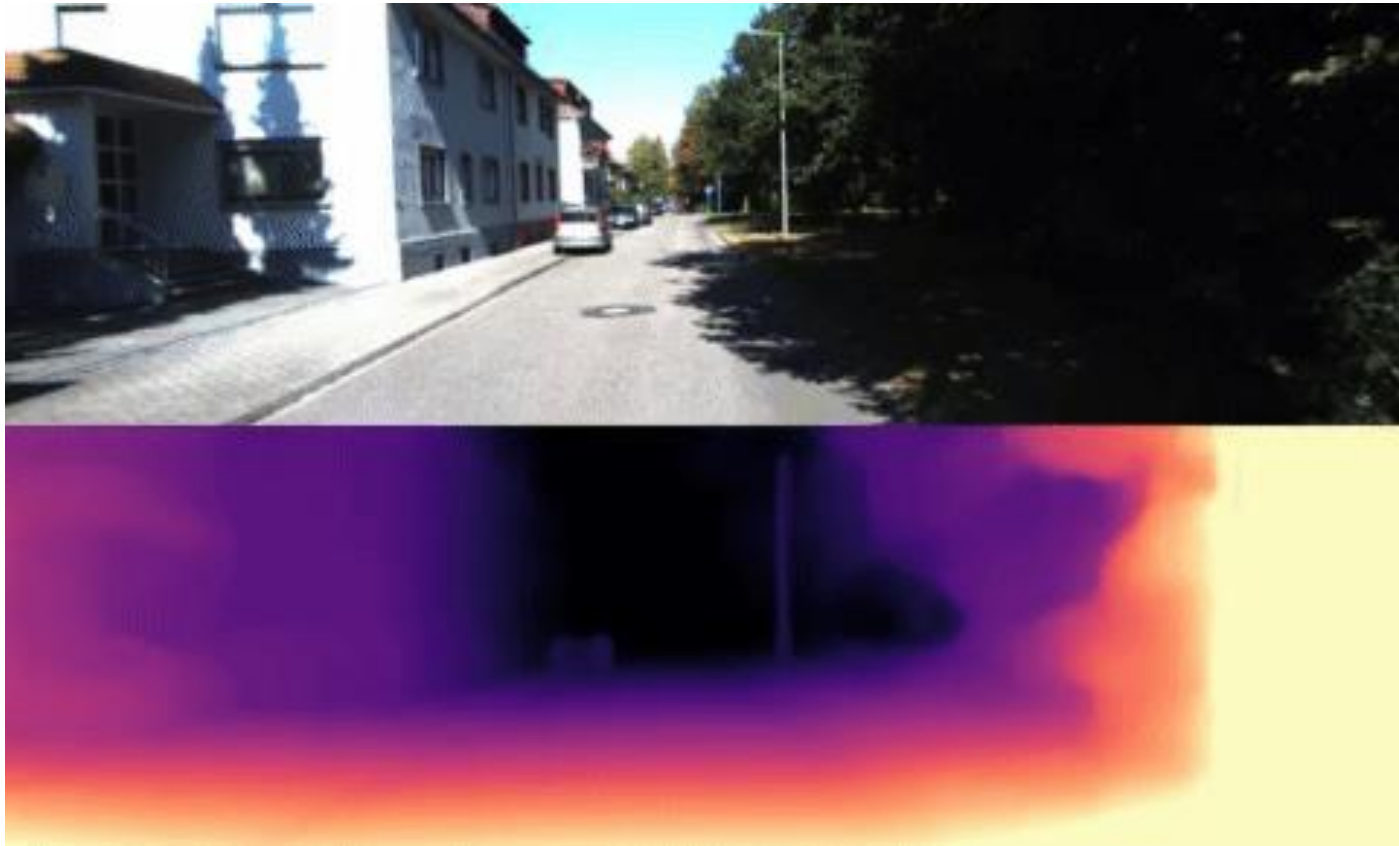
◆ Redes convolucionales

- ❖ Comprensión completa del contexto de conducción por una clasificación a nivel de píxel
- ❖ Clasificación de 19 tipos de objetos
- ❖ 7Hz para una resolución de 1024x512)



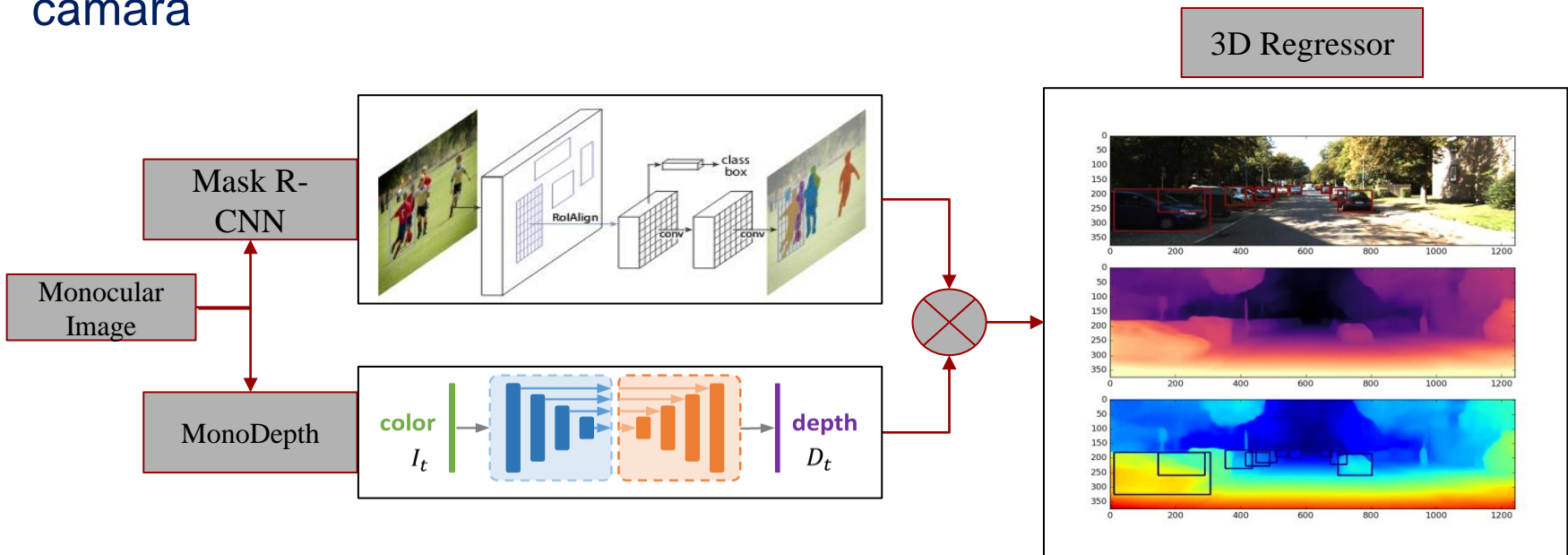


Obtención de distancias a partir de una imagen monocular



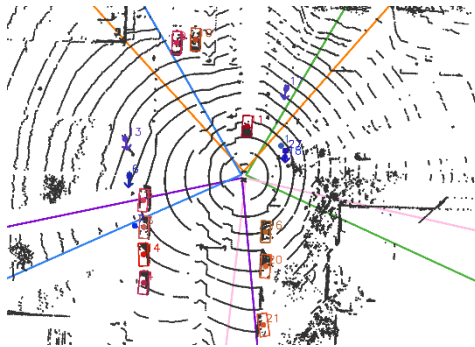
Detección 3D Monocular

- Novedoso enfoque para la detección de objetos en 3D usando sólo imágenes monoculares
- Detección de objetos 2D en paralelo (Máscara R-CNN 100[ms]) y estimación semántica de la profundidad (MonoDepth 70[ms])
- Esto permite estimar en 3D los objetos RGB-D extraídos con una sola cámara

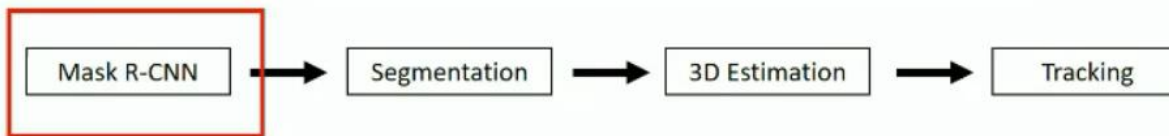
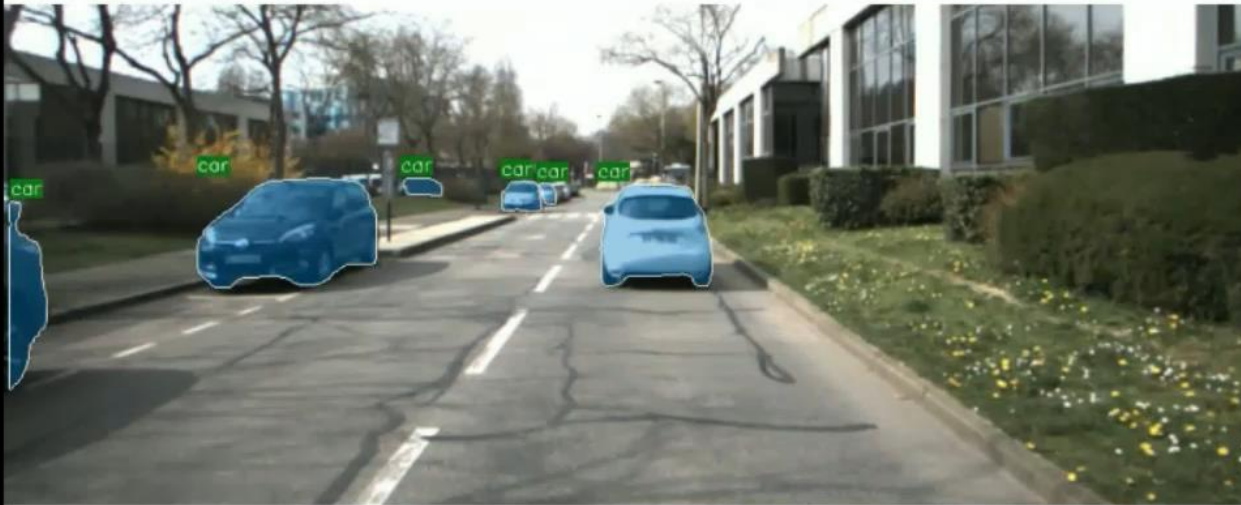


Diseño de Sistema de Percepción para Vehículo Autónomo de Renault

- ◆ Proyecto Tornado
- ◆ Demostración de conducción autónoma por calles interurbanas
- ◆ Tecnología patentada



Diseño de Sistema de Percepción para Vehículo Autónomo de Renault



- ◆ **Presentación del LSI**
- ◆ **Entendimiento del entorno viario**
- ◆ **Percepción del entorno**
- ◆ **Conclusiones**

- ◆ **El desarrollo de los vehículos autónomos depende críticamente de sus algoritmos de percepción**
- ◆ **Gran avance debido a:**
 - ❖ Se dispone de herramientas muy potentes para el análisis de la información sensorial
 - ❖ Existen múltiples bases de datos públicas
 - ❖ Publicación no solo de artículos sino también de código
- ◆ **Líneas de trabajo**
 - ❖ Fusión sensorial: radar/imagen/Lidar
 - ❖ Predicción de actividades



José María Armingol



Fernando García



David Martín



Pablo Marín



Carlos Guindel



Jorge Beltrán



Irene Cortés

Gracias por su atención

arturo.delaescalera@uc3m.es

www.lsi-uc3m.es

https://www.youtube.com/channel/UCfqeOk1EIJ1-VfSYCs_nL4A

