

SISTEMAS DE CONDUCCIÓN AUTÓNOMA

ROBOTS MÓVILES

ALEJANDRO GARCIA MOLINA

LUIS NOLASCO RAMIREZ

ÍNDICE

INTRODUCCIÓN

HARDWARE

SENSORES

IMU

GPS

LIDAR

RADAR

ELEMENTOS DE VISIÓN

ALGORITMOS DE LOCALIZACIÓN

SIMULADORES

SIMULADORES DE CONDUCCIÓN

STSIM DRIVE SIMULATOR

NATIONAL ADVANCE DRIVING SIMULATOR

MAPEADO

SISTEMAS DE SLAM BASADOS EN LIDAR 3D

RECONOCIMIENTO DE OBJETOS

DATASET KITTI

CALTECH DATASET

REDES NEURONALES CONVOLUCIONALES

FASTER - RCNN

PLANIFICACIÓN DE MOVIMIENTO

INTRODUCCIÓN

Los sistemas de transporte inteligentes son uno de los temas más desafiantes en el entorno tecnológico actual. Desde hace ya unas décadas, se ha fijado la atención en los sistemas de automatización de vehículos, especialmente en automóviles, en los que actualmente podemos ver implementadas este tipo de tecnologías, también en aviones (UAVs, unmanned aerial vehicle), submarinos (AUVs, autonomous underwater vehicles), barcos, etc.

Este tipo de sistemas prometen traer grandes mejoras en seguridad, reducción de costes y optimización de recursos, renunciando al factor humano. Los constantes progresos en el desarrollo de sistemas robustos de control, de percepción y de comunicaciones, han abierto un sendero de evolución casi imparable hacia dicha autonomía.

Para conseguir el mencionado objetivo se utiliza un gran entramado de sensores, como GPS, LIDAR, radar o cámaras, la complejidad de estos sensores reside en fusionar información y que esta sea procesada por un ordenador de a bordo, capaz de emitir diagnósticos sobre el estado del vehículo y su entorno para después, mediante algoritmos, emitir órdenes a la red de actuadores que permitirán el control del vehículo.

Aunque el planteamiento parezca simple o asumible, la conducción autónoma sigue siendo, hoy en día, un problema abierto para la comunidad científica, puesto que el control combinado de las dinámicas longitudinal y lateral de un vehículo a velocidades elevadas (autopistas) y en entornos urbanos, es un desafío de extraordinaria complejidad.

HARDWARE

En este apartado hablaremos de la gran variedad de sensores utilizados a día de hoy, sus ventajas e inconvenientes. Veremos que hay muchas combinaciones posibles para implementar en nuestro sistema, no habiendo unas claramente mejor que otras.

SENSORES

Los sensores son una herramienta imprescindible para el vehículo autónomo, ya que estos le permiten construir el entorno en el que se encuentra, y así poder moverse con seguridad. A pesar de que los vehículos siempre han tenido sensores, la complejidad de los sensores en los vehículos autónomos reside en fusionar dicha información y que esta sea procesada por un ordenador de a bordo, capaz de emitir diagnósticos sobre el estado del vehículo, su entorno, etc., y actuar en consecuencia.

La combinación de sensores, actuadores, algoritmos sofisticados y procesadores potentes son los elementos que hacen posible los vehículos autónomos. Las actividades de este se pueden dividir en dos bloques funcionales:

- Percepción (sensores)
- Toma de decisiones (algoritmos y procesamiento)

Este primer bloque los dividiremos en dos tipos:

Sensores propioceptivos: responsables de recoger información del estado del vehículo

Sensores exteroceptivos: aquellos que recogen información del ambiente y entorno del vehículo

Sensores propioceptivos:

-IMU:

La Unidad de Medidas Iniciales (IMU), consiste en una plataforma fijada al vehículo, la cual consta de tres giróscopos y tres acelerómetros, donde cada par está orientado según los ejes ortogonales X, Y y Z.

Estos sensores informan sobre el movimiento lineal y rotacional de la plataforma, siendo estos datos utilizados para calcular el movimiento y la posición del sistema de referencia local del vehículo (se suelen situar en el centro de gravedad del vehículo)

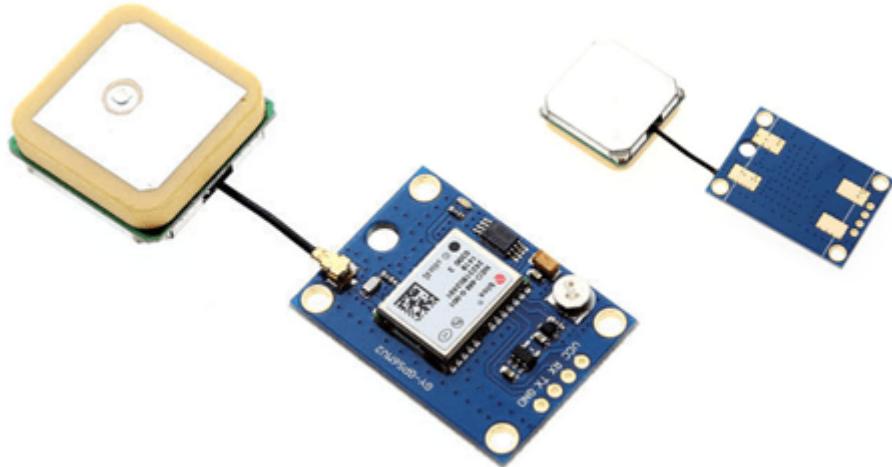


Imagen. Sensor IMU

-GPS:

Es el subsistema que se utiliza para la navegación y guiado del vehículo autónomo. Para determinar las posiciones en el globo, el sistema GPS se sirve de entre 24 y 32 satélites y utiliza el método de la trilateración (método para determinar las posiciones relativas de objetos usando la geometría de los triángulos, de forma análoga a la triangulación.)

Actualmente los sistemas GPS se encuentran en multi-chips, donde solo requieren de alimentación y de una antena. Una vez se computa la localización basada en las señales GPS recibidas, proporciona los datos de salida al procesador. La precisión puede ser hasta de centímetros con los sensores más avanzados, aunque los más usados tendrán una media de unos metros de exactitud.



1 Imagen. Sensor GPS

Sensores exteroceptivos:

-LIDAR:

El sensor LIDAR (Light Detection and Ranging) es un dispositivo que mide la distancia mediante la iluminación de un objetivo con luz láser. Es capaz de mapear de forma muy precisa su entorno mediante haces de luz, calculando incluso velocidades relativas al propio sensor.

La distancia al objeto se determina midiendo el tiempo de retraso entre la emisión del pulso y su detección a través de la señal reflejada.

Empresas como Google y Uber fueron pioneras en la utilización de esta tecnología (más tarde se hablará sobre el vehículo sin conductor de Google), debido a su gran capacidad para generar información, su uso se ha extendido en todos los proyectos relacionados con conducción autónoma, lo que ha generado problemas de abastecimiento, pese al precio de estos dispositivos que alcanza en ocasiones la decena de miles de dólares.

Para obtener la nube de puntos del terreno mediante escáner láser aerotransportado, el módulo debe realizar dos movimientos combinados, una barrida longitudinal y otra transversal, posible mediante un espejo móvil que desvía el haz de luz láser emitido por el escáner. Para obtener las coordenadas de la nube de puntos se necesita la posición del sensor y el ángulo del espejo en cada momento, para esto el sistema se suele apoyar en un GPS diferencial y un sensor inercial de navegación. El resultado son decenas de miles de puntos por segundo.



2 Imagen. Sensor LIDAR

Aunque parece una gran opción, también cuenta con algunas desventajas a tener en cuenta a la hora de decidir en qué sensores se apoyará el sistema autónomo:

- La tecnología aún no está profundamente desarrollada como para ser instalada en millones de coches, son necesarias muchas más pruebas en diferentes condiciones de luz y temperatura, para ofrecer las garantías necesarias.
- Cuenta con un tamaño relativamente grande, aunque Waymo parece estar consiguiendo unidades más compactas en sus prototipos biplaza. Este factor además de afectar a la estética (la mayoría de LIDARS se apoyan en una subestructura independiente en el techo), también afecta a la aerodinámica del vehículo.
- Su alta demanda, los fabricantes no consiguen atenderla, actualmente pueden tardar hasta 6 meses en recibir una unidad. Por este motivo las empresas están diseñando sus propios sensores LIDAR, Waymo, por ejemplo, tiene desarrollados tres sensores diferentes para tres distancias distintas.
- Son muy caros, cada unidad puede costar miles de euros, y un coche debería contar con varios si es de tamaño mediano.

-RADAR:

La tecnología Radar hace uso de ondas de radio para determinar el rango, ángulo o velocidad de objetos, los hay de corto o largo alcance (27-77 GHz).

Los de corto alcance se utilizan para obtener información de las vecindades del coche, en un rango de aproximadamente 30 metros. Se suelen utilizar a bajas velocidades.

Por el contrario, los de largo alcance, cubren distancias de aproximadamente 200 metros, posibles a velocidades altas.

Generalmente, los sensores radar obtienen información de los objetos cercanos, como por ejemplo la distancia a la que se encuentran, su tamaño, o la velocidad que llevan en caso de que el objeto esté en movimiento. La gran precisión de estos sensores, junto a su capacidad de trabajar correctamente en condiciones meteorológicas adversas, hacen de ellos unos sensores de uso permanente en los vehículos autónomos. Funcionan con la misma efectividad en condiciones de luz como de oscuridad, y los sensores de 77GHz recogen información de forma más precisa que los LIDAR en situaciones de niebla, lluvia o nieve. Al igual que estos, no detectan color, contraste ni caracteres.

Los radar son sensores de tamaño reducido y bajo coste, estos ofrecen, como se ha comentado, un buen rango de funcionamiento, pero la resolución que ofrece no es tan buena como la de otros sensores.



3 Imagen. Sensor RADAR

-ELEMENTOS DE VISIÓN:

Los sistemas basados en cámaras se utilizan para detectar caracteres, señales, las marcas de los distintos carriles, pero también, mediante sofisticados algoritmos, detectar objetos, prever sus futuras trayectorias y conocer su velocidad relativa.

Dependiendo de las necesidades, las cámaras pueden ser situadas en diferentes sitios: en los espejos retrovisores, puertas traseras, luna frontal, etc.

Tienen un rango más amplio que el resto de sensores, pero solo en el caso de que haya condiciones adecuadas de luz. De hecho, su rango se deteriora con el nivel de luz, al igual que ocurre con la visión humana.

El hecho de que estos sistemas detecten color, contraste y reconozcan caracteres cubre el hueco del resto de sensores mencionados previamente, ya que no proporcionan ninguna de estas características.

Los sistemas de reconocimiento de imagen mediante cámaras son bastante baratos y de pequeño tamaño, además su resolución es bastante alta.

Algunos problemas principales de este tipo de elementos de visión es el clima (nieblas, lluvias fuertes, nieve, etc) y las sombras proyectadas por otros objetos.



4 Imagen. Cámara trasera de un vehículo

-Reflexión global:

Tras indagar en los sensores que deberá montar un vehículo que pretenda conducir autónomamente, podemos concluir con que no es viable confiar la seguridad del vehículo a un único sensor o incluso dos. La solución residirá en la efectiva combinación de diferentes sensores, y en la correcta fusión de la información generada por estos. Cada sensor tiene sus puntos fuertes y débiles, y la comprobación de que la información recibida por cada uno concuerda con la de los demás evitará tanto falsos negativos como falsos positivos.

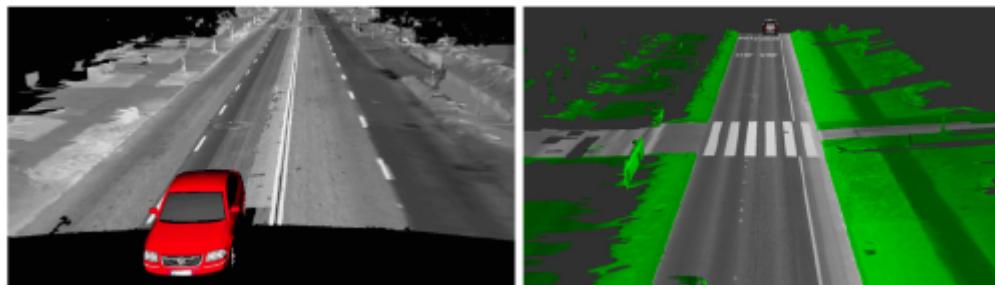
ALGORITMOS DE LOCALIZACIÓN

Hoy día es fácil pensar que los algoritmos de localización en un coche autónomo son sencillos y pueden estar basados, por ejemplo, en el GPS, dado que la mayoría de coches cuentan con uno. Sin embargo, nada más lejos de la realidad dado que únicamente utilizando el GPS del propio coche, este solo sabe qué dirección debe tomar pero no conoce nada sobre su entorno.

Por tanto, en la actualidad uno de los sensores más valorados para vehículos son los escáneres láser 3D o LIDARS. Proporcionan una visión rápida y en 3D del entorno, incluyendo distancias (profundidad) y niveles de reflectancia de cada material para cada dirección del espacio.

Un perfecto ejemplo de coches autónomos es el coche de Google famoso por ser el “coche sin conductor”, cuyo algoritmo de localización funciona de la siguiente manera: Se afronta la localización sólo tras haber realizado un mapa del entorno, que se usará para localizarse asumiendo que éste no va a cambiar mucho con el tiempo. Hay poderosas razones para proceder así: construir un mapa es uno de los mayores problemas de la robótica y la visión por ordenador (el problema se llama SLAM), por tanto es esencial capturar datos controlando los vehículos de manera manual, y entonces emplear el tiempo necesario procesándolos en el laboratorio hasta que se obtenga un mapa aceptable. Sólo entonces, se proporciona este mapa al robot para que pueda localizarse en tiempo real, ya sea de nuevo teleoperado o moviéndose autónomamente.

A la nube de puntos 3D del LIDAR se incorpora la información de «color» que ve el láser, monitorizando el eco de los rayos láser infrarrojos, el sensor determina el nivel de gris (reflectancia) de cada punto escaneado del entorno. Componiendo esta información con los puntos 3D se obtienen imágenes fotorealistas del mundo. Es importante insistir en que el nivel de gris no ha sido obtenido con ningún tipo de cámara, sino directamente a través de los millones de puntos por segundo adquiridos por un LIDAR.



5 Imagen. Mapeado obtenido por sensor LIDAR

A continuación se extrae solamente el plano horizontal de tierra, la única parte del entorno que se empleará de verdad para localización. Según los autores, quedarse con el entorno 3D completo no pareció aportar mucho. Sin embargo, sería fundamental si se emplean otros sensores, como cámaras ópticas. Al no grabar las «imágenes» del suelo con cámaras

normales se evitan algunos problemas típicos de visión. Por ejemplo, nunca se verán sombras proyectadas por luz solar, ya que lo que se mide es directamente la reflectancia de cada punto desde el LIDAR. Además, las imágenes apenas varían con las condiciones meteorológicas.

Después de recopilar toda esta información, debe ser procesada teniendo especialmente en cuenta todos los errores que esta incluye, por ejemplo, los errores de posición de GPS. Para ello, un método efectivo sería la minimización de residuos cuadráticos. Este acercamiento al problema desde la perspectiva de mínimos cuadráticos se llama «**Graph SLAM**».

Al final, el problema se reduce a minimizar esta función de coste:

$$\begin{aligned}
 J = & \sum_t (x_t - g(u_t, x_{t-1}))^T R_t^{-1} (x_t - g(u_t, x_{t-1})) \\
 & + \sum_{t,i} (z_t^i - h_i(m, x_t))^T Q_t^{-1} (z_t^i - h_i(m, x_t)) \\
 & + \sum_t (x_t - (y_t + b_t))^T \Gamma_t^{-1} (x_t - (y_t + b_t)) \\
 & + \sum_t (b_t - \gamma b_{t-1})^T S_t^{-1} (b_t - \gamma b_{t-1})
 \end{aligned} \tag{7}$$

6 IMAGEN: Función de coste original

donde x son posiciones del coche, $g()$ es la función que modela la odometría (incrementos de rotación de las ruedas), $h()$ es un modelo de observación que predice la apariencia que van a tener los puntos LIDAR según se vea el plano del suelo desde una determinada posición x , m es el mapa del suelo propiamente dicho, z las medidas del LIDAR, y las lecturas del GPS y b es una variable utilizada para modelar la lenta variación del sesgo/error en las señales de un GPS.

Se trata de una solución ideal pero inviable debido al grandísimo número de incógnitas. Es por ello, que una solución factible sería eliminar el mapa de la ecuación.

Primero, se afronta un modelo simplificado: solamente usando GPS y datos iniciales, se tiene una primera estimación de la trayectoria del coche. Asumiendo que las orientaciones están más o menos correctas en este modelo, el problema se simplifica enormemente ya que solamente nos queda ajustar posiciones en el plano (2D), mucho más fácil de optimizar debido a que no introducen no linealidades como sí hacen las incógnitas de ángulos y giros. Entonces, los mapas del suelo de puntos sobre los que el coche ha pasado más de una vez se emparejan unos con otros para establecer nuevas restricciones adicionales entre las posiciones del coche. La idea es deshacernos de la función de observación $h()$, y centrarnos en optimizar solamente la trayectoria del vehículo. Así queda esta función de coste mucho más simple:

$$\begin{aligned}
J' = & \sum_t (x_t - g(u_t, x_{t-1}))^T R_t^{-1} (x_t - g(u_t, x_{t-1})) \\
& + \sum_t (x_t - (y_t + b_t))^T \Gamma_t^{-1} (x_t - (y_t + b_t)) \\
& + \sum_t (b_t - \gamma b_{t-1})^T S_t^{-1} (b_t - \gamma b_{t-1}) \\
& + \sum_t (x_t + \delta_{st} - x_s)^T L_{st} (x_t + \delta_{st} - x_s) \quad (9)
\end{aligned}$$

7 IMAGEN: Función de coste simplificada

El resultado es una ecuación la cual un ordenador normal podría resolver en cuestión de segundos con errores típicos de 5cm.

Una vez corregida la trayectoria del coche, construir el mapa del suelo completo es una tarea trivial: sólo hay que proyectar geométricamente los datos del LIDAR desde cada una de las posiciones del vehículo, ya conocidas.

Ahora el coche puede ser conducido con el mapa cargado a bordo y que éste sea capaz de localizarse con precisión. Para hacerlo, se emplea un filtro de partículas (PF) que compara las medidas del LIDAR con la predicción del mapa del suelo. Un PF es un caso particular de estimador Bayesiano, similar a un algoritmo genético con reglas sólidamente fundadas en la estadística. La posición del coche se estima mediante unos *cientos de hipótesis* sobre su posición, que evolucionan en el tiempo de manera que encajan (estadísticamente) con las medidas de los sensores del coche. Cada hipótesis predecirá una medida distinta del LIDAR, y solamente aquellas que más se acerquen a la realidad sobreviven; el resto son descartadas.

A pesar de la simplicidad, y limitaciones, de este tipo de filtros, han demostrado ser extremadamente robustos contra ruido y medidas de elementos inesperados en el entorno.

Uno de los grandes problemas a la hora de realizar un proyecto como el de un coche autónomo es el coste tan elevado de los sensores LIDAR los cuales pueden llegar a costar entre 70000 y 100000 euros o más.

SIMULADORES

Cuando se habla de conducción autónoma es necesario hacer mención a una de las principales herramientas que permiten su desarrollo. La conducción autónoma implica muchos riesgos, debido a ello se utilizan en primera instancia simuladores para, más tarde,

poder implementar el sistema en un entorno real. Su objetivo es reproducir tanto las sensaciones físicas como el comportamiento de los elementos del objeto a simular, eliminando el riesgo de accidentes o daños tanto para el equipo como para el usuario.

Simuladores de conducción

Son el tipo de simulador más popular en la actualidad debido a sus distintos tipos de uso. Se utilizan para la investigación y desarrollo de nuevos sistemas de asistencia a la conducción (ADAS). Aunque también se usan en autoescuelas para poder repetir situaciones concretas, por profesionales para entrenar deportes de motor y en la industria de los videojuegos.

Simuladores de conducción para el desarrollo de sistemas de conducción

STISIM Drive Simulator

Simulador programable y totalmente interactivo. Utilizado por más de 500 universidades, hospitales, empresas y centros de entrenamiento para realizar investigaciones, evaluaciones y tratamientos. Posee sobre 90 simulaciones de conducción listas para ser ejecutadas con diversas situaciones de conducción, entornos personalizables y librerías para añadir elementos y objetos de carretera.



8 Imagen. STISIM Drive Simulator

National Advanced Driving Simulator (NADS)

Es un centro sostenido de investigación de seguridad en el transporte en el Parque de investigación de la Universidad de Iowa, Estados Unidos. Hace uso de un conjunto de simuladores de conducción para llevar a cabo estudios de investigación para los sectores público y privado.

Ofrece simuladores con el objetivo de lograr de la mejor manera posible los requisitos de cada proyecto: NADS-1, el simulador de mayor fidelidad del mundo; NADS-2, simulador de base fija; NADS MiniSim, simulador portátil enfocado a computadoras de bajo coste.



9 Imagen. NADS

Mapeado

Los sistemas de localización y mapeo simultáneos (SLAM, simultaneous location and mapping) han sido estudiados durante años y forma parte esencial de los sistemas de conducción autónoma, dotando al vehículo de mapas y localización actualizada al momento. Se utilizan dos tipos de sensores para implementarlo: sensores propioceptivos, captan información interna del vehículo (velocidad, aceleración, orientación), y los exteroceptivos que captan información del exterior. Ambos tipos se combinan para localizar y mapear.

Para poder realizar una conducción autónoma el vehículo debe ser capaz de construir un mapa incremental del entorno y a la vez saber cuál es la posición exacta en la que se encuentra dentro de dicho mapa.

Existen diversas fuentes de incertidumbre y, por lo tanto, de error que podrían ser fatales si sucedieran en un entorno real. Entre ellos:

- Ruido de los sensores
- Entornos donde aparecen simetrías
- Cambios en el entorno.
- Uso de modelos probabilísticos

Sistemas de SLAM basados en Lidar 3D

Light Detection And Ranging es una tecnología que permite medir la distancia entre un sensor y un objeto a través de un rayo de luz láser para obtener medidas de posicionamiento georreferenciadas mediante modelos digitales tridimensionales de alta precisión. Es de los más utilizados en la conducción autónoma.

A continuación algunos ejemplos de grandes empresas y universidades que implementan las técnicas de SLAM:

El proyecto CityMobil2 recorrió más de 26000 kilómetros durante varios meses en diferentes ciudades, para poder obtener las suficientes medidas y generar un mapa de gran resolución.

La Universidad de Stanford participó en el Urban Challenge. Su vehículo generaba un mapa terrestre utilizando GPS, IMU y un Lidar 3D de 64 capas de múltiples escaneos sin conexión. La localización se llevó a cabo dentro del mapa con un filtro de histograma 2D con una precisión de 10 cm.



10 Imagen. Vehículo autónomo de Stanford e histograma 2D en tiempo real.

La Universidad de ULM, creaba un mapa compuesto por características MSER procedentes del sistema de visión del coche y de una rejilla Lidar 2D georreferenciada mediante un GPS RTK. La localización se llevó a cabo a través del método Montecarlo. Las pruebas se realizaron a lo largo de 5 km, todas ellas con una precisión de 10 cm.

BMW, ofrece una descripción general de su experiencia con la conducción autónoma durante miles de kilómetros en la vía pública. Se basan en el método del carril marcado mediante el uso de visión Lidar, así como odometría y GPS. Para la detección de los límites de la carretera necesitan Lidar y sensor de radar. Su mapa integra información semántica y geométrica (carril, modelos, conectividad, etc) así como datos de localización (marcas de carril y posición de los límites de la carretera).

RECONOCIMIENTO DE OBJETOS

Las tendencias recientes en Deep Learning han logrado un rendimiento de detección impresionante en muchos casos distintos, incluida la detección de objetos. Además, hoy en día, varios modelos básicos de DNN se comparten públicamente y se puede utilizar como paso de inicio para un mayor entrenamiento y puesta a punto de diferentes tareas de clasificación de objetos.

El **dataset KITTI** es una de las bases de datos más populares en el contexto de la detección de objetos para vehículos autónomos. Para la detección de peatones en concreto también es muy popular la **base de datos de Caltech**.

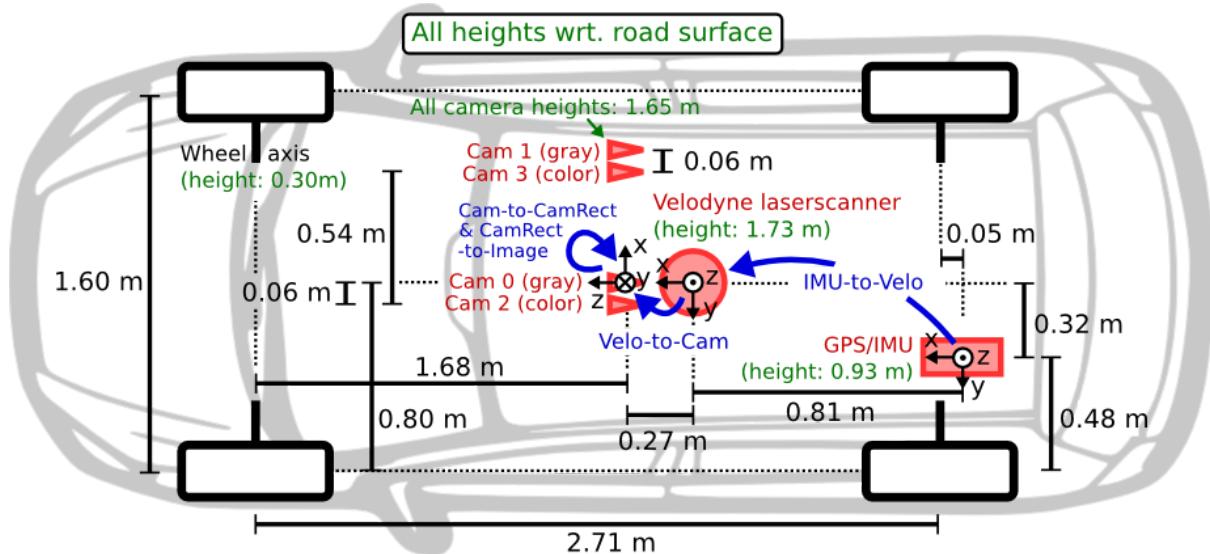
Las redes neuronales convolucionales permitieron una mejora significativa en el rendimiento de los detectores de objetos. Al principio, estas redes utilizaban un enfoque de ventana deslizante. Sin embargo, la localización precisa de los objetos es un desafío debido a los grandes campos receptivos. Más adelante, se proponen R-CNNs para solucionar el problema de la localización con un paradigma de reconocimiento usando regiones. Generan propuestas de región mediante búsqueda selectiva, extraen un vector de características de longitud fija para cada propuesta usando una red convolucional y clasifican cada región con un SVM (support vector machine) lineal. Las redes basadas en regiones son computacionalmente costosas, pero se han propuesto varias mejoras para reducir la carga computacional. Usando la agrupación de pirámides espaciales que permiten calcular un mapa de características convolucionales para toda la imagen con solo una ejecución de la red convolucional en contraste con la R-CNN que debe aplicarse en muchas regiones de imágenes.

Dataset KITTI

KITTI es una base de datos capturada por vehículos que circulan alrededor de la ciudad de Karlsruhe, en áreas rurales y en autopistas donde se captan hasta 15 vehículos y 30 peatones a la vez en imágenes. Estos vehículos están altamente equipados con dos cámaras de alta definición RGB y otras dos en escala de grises, un escáner de velocidad de alta resolución de 360°, un GPS y una IMU.

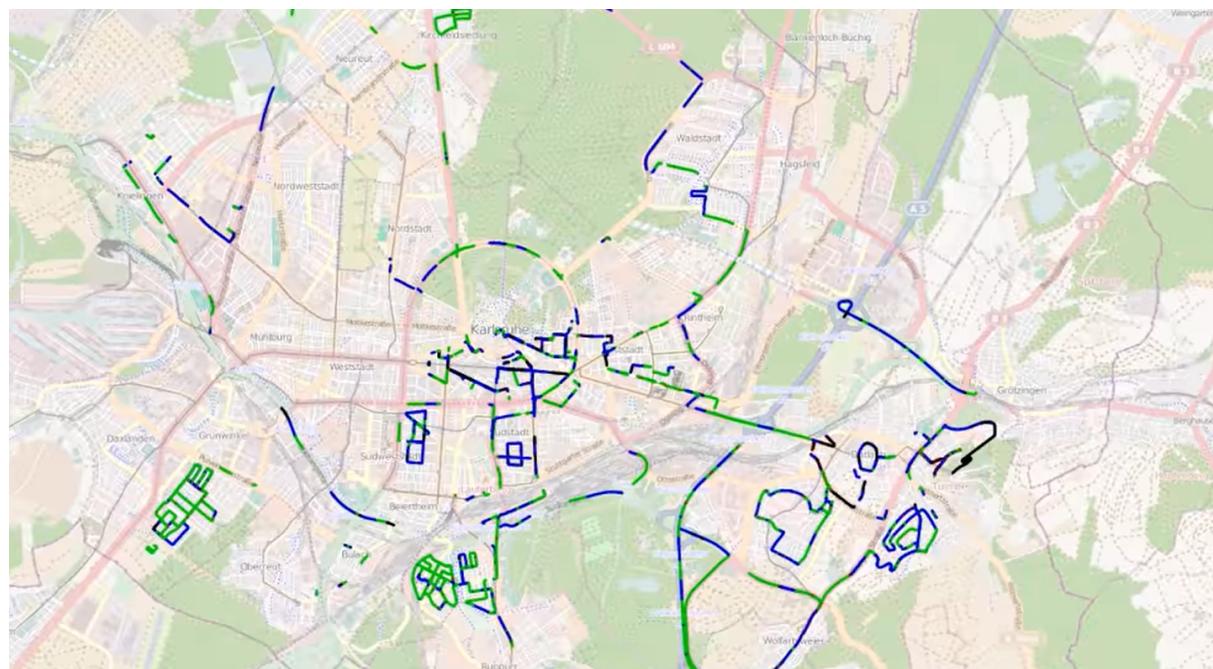


11 Imagen. Vehículo utilizado para la base de datos KITTI



12 Imagen. Medidas del vehículo

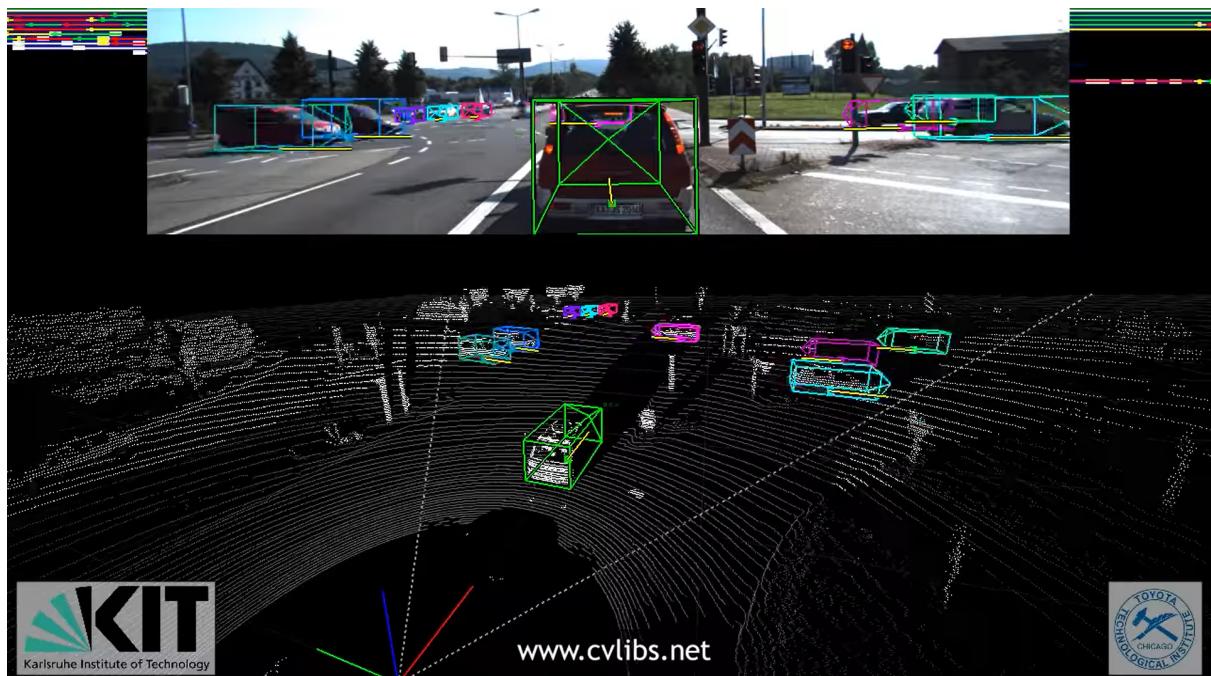
Los diferentes datos captados por los vehículos consisten en 22 vídeos estéreo con una longitud total de 40 kilómetros



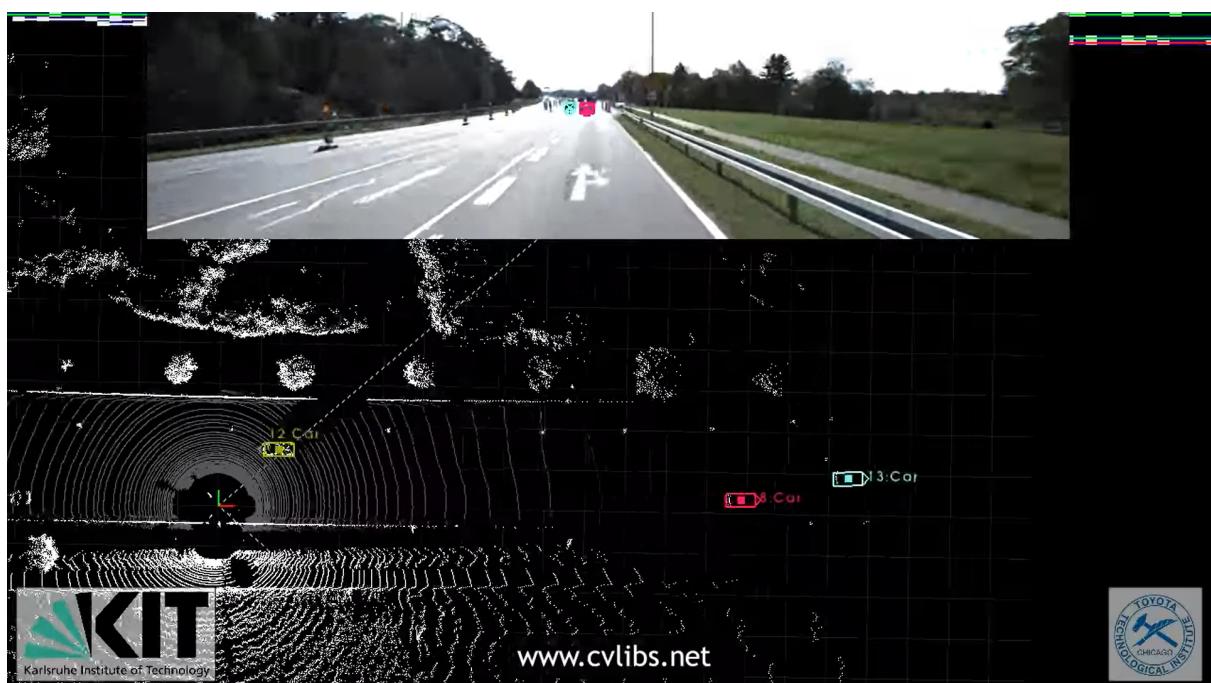
13 Imagen. Trayectorias seguidas por la base de datos KITTI.

Están disponibles todos los datos obtenidos por todos los sensores para todas las trayectorias realizadas pudiendo hacer uso así de sus datos en crudo. Además de eso, la propia empresa ha realizado una serie de benchmarks para cada una de las tareas específicas: estéreo, optical flow, scene flow, profundidad, odometría, detección de objetos

2D y 3D, seguimiento de objetos, carreteras y varios más. Y proveen métricas de evaluación y una web de evaluación online



14 Imagen. Benchmark de reconocimiento de objetos 3D



15 Imagen. Benchmark de reconocimiento de objetos 3D.

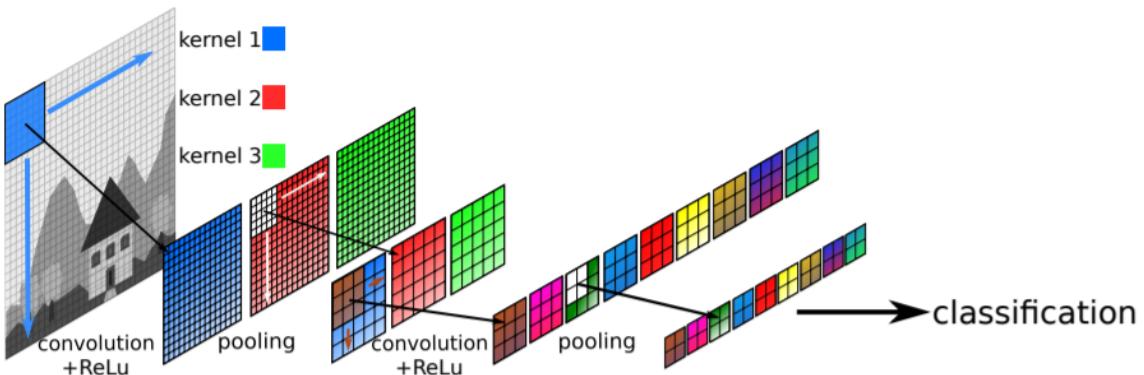
Caltech Dataset

Utilizado principalmente en la detección de peatones, consiste en 6 conjuntos de entrenamiento de 1GB cada uno. Cada uno además contiene entre 6 y 13 archivos de secuencia de un minuto con anotaciones e información, cada archivo de secuencia consta de una serie de frames de video y cada frame contiene una cabecera con información respecto al mismo. Al igual que el dataset KITTI también proporcionan sus propios benchmarks..

REDES NEURONALES CONVOLUCIONALES

El tipo de red más usado a la hora de clasificar y procesar imágenes. Son un tipo de red neuronal especializado en procesar datos ordenados en forma de cuadrícula y reciben su nombre de una operación matemática llamada convolución.

La convolución aplica diferentes filtros a la imagen y obtiene como salida una imagen transformada y de dimensión reducida para cada filtro. Después se aplica la función de activación, entre las que destaca la ReLU y se vuelve a reducir la imagen extrayendo los valores más significativos. Una vez aplicado esto se conecta con una red neuronal convencional para realizar una clasificación de la imagen.

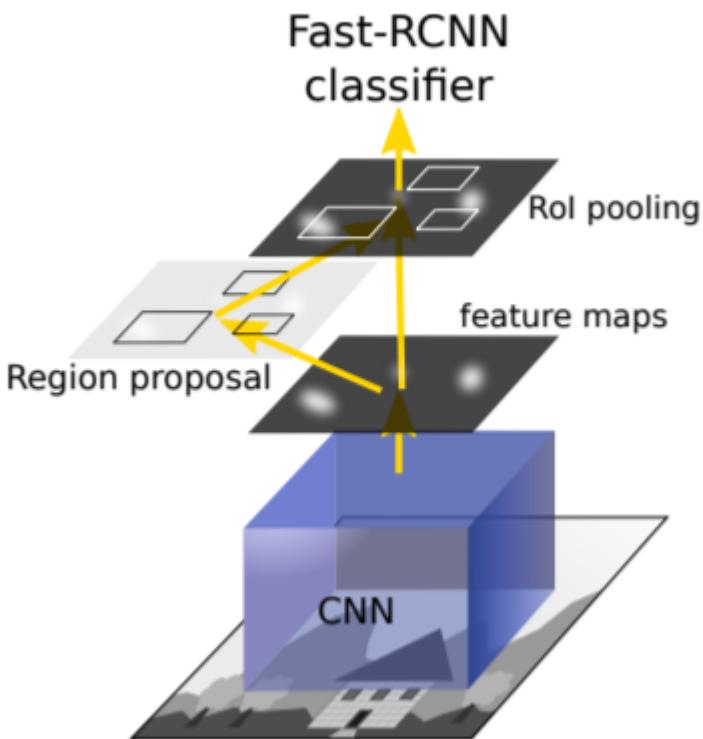


16 Imagen. Esquema de una red convolucional con dos capas de convolución y pooling.

Faster R-CNN

El rápido avance del campo de los vehículos autónomos ha hecho necesario no solo la clasificación de imágenes sino también sistemas capaces de clasificar cada objeto dentro de una imagen, localizarlo y posicionarlo en la misma, convirtiendo la detección de objetos en una tarea mucho más compleja que solo la clasificación de imágenes.

Se utiliza una arquitectura basada en propuestas de región de interés. Se extraen regiones de interés que contienen candidatos a ser algún objeto a detectar y se utiliza el clasificador en estas regiones.



17 Imagen. Arquitectura del modelo Faster R-CNN

PLANIFICACIÓN DE MOVIMIENTO

Waymo (división de vehículos autónomos de Google) y Tesla son las compañías que actualmente se encuentran en la fase más avanzada del desarrollo de coches autónomos. Google comenzó su proyecto de automóvil autónomo en 2009 y completó más de 1,498,000 millas de forma autónoma hasta marzo de 2016 en Mountain View, California, Austin, TX y Kirkland, WA. Diferentes sensores (por ejemplo, cámaras, radares, LiDAR, GPS) permiten detectar peatones, ciclistas, vehículos, obras en carreteras, etc. en todas las direcciones. De acuerdo con sus informes de accidentes, los automóviles autónomos de Google estuvieron involucrados sólo en 14 accidentes, y 13 de los 14 incidentes fueron causados por terceros. En 2016, el proyecto continuó con el nombre de Waymo, que es una compañía independiente de tecnología de autoconducción.

Por otro lado, Tesla Autopilot es un sistema avanzado de asistencia a la conducción desarrollado por Tesla que se lanzó por primera vez en 2015 con la versión 7 de su software. El nivel de automatización del sistema permite la automatización total, pero requiere la plena atención del conductor para tomar el control si es necesario. A partir de octubre de 2016, todos los vehículos producidos por Tesla estaban equipados con ocho cámaras, doce sensores ultrasónicos y un radar orientado hacia adelante para permitir la capacidad de conducción automática. Un automóvil condujo de forma autónoma el 98 % del tiempo desde Holmdel hasta Atlantic Highlands en el condado de Monmouth NJ, así como

otras 10 millas en Garden State Parkway sin intervención. Para cumplir este objetivo, se utilizó una red neuronal convolucional que predice el control del vehículo directamente desde las imágenes que obtiene del exterior, utilizando la computadora Nvidia Drive PX2.

A la hora de planificar la ruta a seguir por un vehículo autónomo se siguen, de forma habitual, los siguientes pasos:

- A. **Planificación del camino.** Se decide la ruta a seguir por el vehículo desde su ubicación actual hasta el destino solicitado a través de la red de carreteras. La red de carreteras se suele representar como un grafo dirigido donde se pueden usar algoritmos como el Dijkstra o A* para calcular la ruta más corta.
- B. **Decisión de comportamiento.** El vehículo debe poder navegar por la ruta seleccionada e interactuar con otros participantes del tráfico de acuerdo a las normas de circulación y tránsito. Es el paso responsable de seleccionar la acción a realizar en función de la situación frente a la que se encuentre, como por ejemplo detenerse en un semáforo. Un problema ampliamente estudiado en esta capa es el de predecir y estimar las trayectorias futuras de otros usuarios de la vía pues, el mundo real se caracteriza por las intenciones inciertas del resto del entorno.
- C. **Planificación de ejercicios.** Una vez seleccionado el comportamiento en la capa anterior es necesario convertirlo en una ruta o trayectoria. Esta trayectoria debe cumplir con la dinámica del vehículo, ser cómoda para los pasajeros y evitar colisiones.
- D. **Control de vehículos.** Se usa un controlador de retroalimentación para seleccionar la entrada del conductor apropiada y corregir el error resultante.

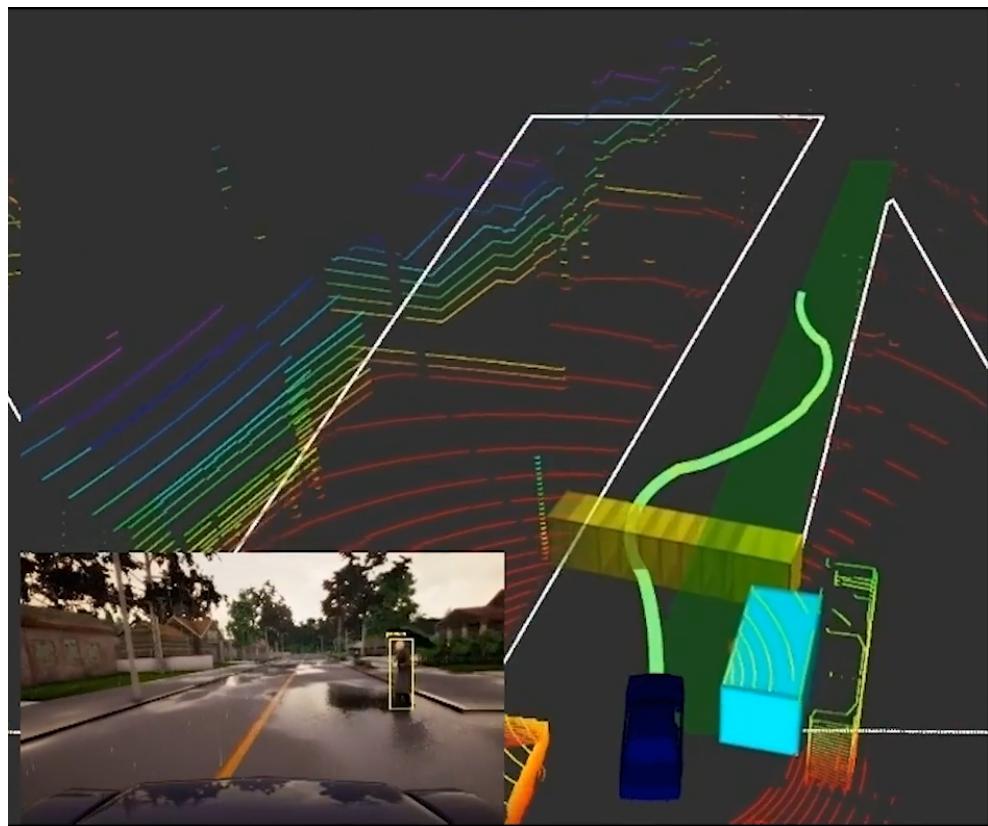
La razón por la que es tan importante la planificación de movimientos rápidos para vehículos autónomos es porque la planificación de movimientos abarca el proceso mediante el cual el vehículo decide qué hacer a continuación. Hacer este proceso más rápido no solo significa que el vehículo puede tomar decisiones más rápidamente, sino que también puede tomar decisiones mucho mejores: mantener al pasajero y a todos los que lo rodean lo más seguros posible.

Mientras que la mayoría de los coches autónomos se basaban en sistemas de planificación de movimiento que funcionan a una velocidad de entre 3 Hz y 10 Hz, Realtime Robotics implementó un nuevo sistema que funcionaba a 1.000 Hz, pasando por una red masiva de datos de posibles trayectorias en paralelo.

La mayoría de sistemas calculaban una trayectoria capaz de esquivar obstáculos dinámicos y estáticos pero utilizan mucha potencia de cálculo y son generalmente lentos tomando unos 100 ms para su cálculo. Este tiempo era adecuado para autopistas y autovías pero no es suficiente para entornos urbanos densos y poco estructurados.

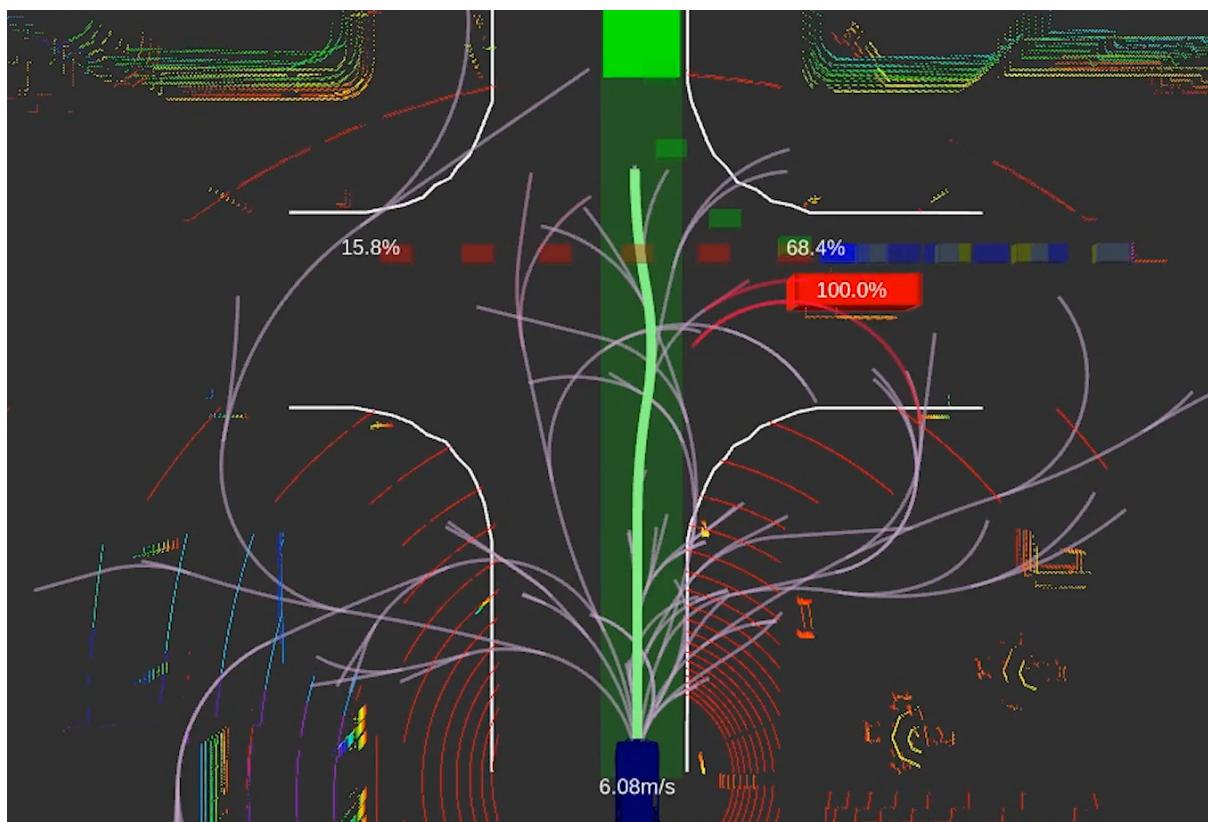
El nuevo procesador de Realtime Robotics reduce su peso y consumo además de introducir dos nuevas capacidades que no estaban presentes en otros competidores: reactividad instantánea y conciencia del riesgo.

- Reactividad instantánea. La enorme velocidad de procesamiento permite reaccionar a tiempo ante situaciones peligrosas y ser capaz de reaccionar en tiempo real permitiendo así a los vehículos autónomos alcanzar mayor velocidad sin riesgo.



18 Imagen. Reactividad instantánea.

- Conciencia del riesgo. Con la información obtenida calcula la posibilidad de que sucedan distintas situaciones de riesgo en el entorno en que se encuentra y prepara, de forma preventiva, trayectorias capaces de hacerse cargo de esas situaciones.



19 Imagen. Conciencia del riesgo

A día de hoy, las empresas punteras en este campo utilizan esta tecnología con incluso más potencia permitiendo así la completa navegación autónoma en autopista como se puede comprobar con los coches Tesla. Por entornos urbanos aún hay complicaciones debido a que su seguridad y precisión no es tanta por ello la legislación aún no ha aceptado su uso en muchos países.

Bibliografía:

<https://eprints.ucm.es/id/eprint/15311/1/T33773.pdf>

<https://zaguan.unizar.es/record/69868/files/TAZ-TFG-2017-4623.pdf>

<https://idus.us.es/bitstream/handle/11441/80759/TFG-1847-GONZALEZ.pdf?sequence=1&isAllowed=y>

https://ebuah.uah.es/dspace/bitstream/handle/10017/45087/TFM_Ribalda_Fernandez_2020.pdf?sequence=1&isAllowed=y

<https://idus.us.es/bitstream/handle/11441/94482/TFG-2732-PINILLA%20ALONSO.pdf?sequence=1&isAllowed=y>

https://idus.us.es/bitstream/handle/11441/54393/TFM_PatriciaL%C3%B3pezTorres.pdf?sequence=1&isAllowed=y

<http://www.cvlibs.net/datasets/kitti/>

http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/

<https://programmerclick.com/article/25201146870/>

<https://blog.bricogeek.com/noticias/tecnologia/sistema-de-planificacion-de-movimiento-ultra-rapido-para-coches-autonomos/>

Ignacio Arriola, Detección de objetos basada en Deep Learning y aplicada a vehículos autónomos.