

Anteproyecto de Trabajo Fin de Grado

Esteban Gómez García

4 de octubre de 2024

Autor: Esteban Gómez García

Tutores: Luis Miguel Bergasa Pascual y nombreCoTutor

Titulación: Grado en Ingeniería Informática

Título: Detección de objetos 3D con métodos de deep learning y fusión temporal para cámara y LiDAR

Departamento: Departamento de Electrónica

1 Introducción

En los últimos años, la detección de objetos en tres dimensiones (3D) ha ganado un rol central en el desarrollo de vehículos autónomos, convirtiéndose en un desafío crucial para mejorar la precisión y seguridad en la conducción automatizada. La combinación de sensores LiDAR y cámaras ha permitido capturar tanto información de profundidad como de color y textura, aspectos esenciales para la interpretación completa del entorno en el que se desenvuelven los sistemas autónomos. Este Trabajo de Fin de Grado (TFG) se enfoca en la detección de objetos 3D mediante la integración de métodos de deep learning y fusión temporal aplicados a imágenes obtenidas de cámaras y datos de sensores LiDAR.

El proyecto utiliza el modelo YOLO (You Only Look Once) para la detección de objetos, principalmente vehículos, en un entorno dinámico captado por un coche autónomo. Mediante el uso de PyTorch, se entrena y optimiza la estimación tridimensional de YOLO, permitiendo una comprensión espacial avanzada de los elementos presentes en el entorno de conducción. Además, para alcanzar una representación temporal coherente y precisa, se propone una fusión de datos basada en la fusión temporal de distintas secuencias de imágenes de cámara y datos LiDAR, capturadas en diferentes momentos y posiciones. Esta técnica de fusión temporal tiene como objetivo lograr una detección más robusta y confiable, especialmente en condiciones de tráfico y situaciones complejas.

Para el desarrollo y la evaluación del modelo, se recurrirá a datasets ampliamente utilizados en el campo de la conducción autónoma, como KITTI. La implementación de estas técnicas permitirá obtener una detección más precisa de los objetos 3D alrededor de un vehículo, contribuyendo al desarrollo de sistemas de conducción autónoma más avanzados y seguros.

2 Objetivos

El objetivo fundamental de este proyecto es el desarrollo e implementación de un sistema de detección y estimación de objetos en 3D utilizando redes neuronales de deep learning, específicamente el modelo YOLO, en combinación con un esquema de fusión temporal de datos provenientes de sensores de cámara y LiDAR, montados en un vehículo autónomo. Este sistema busca mejorar la precisión en la detección y localización tridimensional de objetos en el entorno, en especial vehículos, contribuyendo a la eficiencia y seguridad en la conducción autónoma.

Los objetivos específicos de este proyecto son los siguientes:

- Estudiar la arquitectura de YOLO y su adaptación a la detección 3D:
 1. Realizar una revisión de la arquitectura YOLO y analizar su capacidad para detectar y estimar objetos en 3D a partir de imágenes 2D.
 2. Implementar la versión de YOLO optimizada en PyTorch para la estimación de profundidad, ajustando sus características para mejorar el reconocimiento tridimensional de objetos.
- Diseñar e implementar un esquema de fusión temporal de datos de cámaras y LiDAR:
 1. Desarrollar un modelo de fusión de datos que integre las estimaciones en tiempo real de los sensores de cámara y LiDAR.
 2. Explorar técnicas de fusión temporal para mejorar la precisión y consistencia en la detección de objetos en el entorno, utilizando información recopilada en intervalos de tiempo sucesivos.
- Entrenar y optimizar el modelo YOLO en un entorno de conducción autónoma:
 1. Experimentar con diferentes técnicas de entrenamiento para optimizar el rendimiento del modelo en la detección y clasificación de objetos 3D, tomando en cuenta las limitaciones de la inferencia en tiempo real.
- Evaluar el rendimiento del sistema de detección y fusión temporal:
 1. Medir la precisión y rapidez de detección del sistema evaluándolo en condiciones de simulación y con datos de entornos reales.
- Implementar mejoras de precisión y reducir los tiempos de inferencia del modelo:
 1. Proponer ajustes en el modelo de YOLO y en la fusión de datos para optimizar la precisión y minimizar el tiempo de respuesta.
 2. Desarrollar herramientas de análisis de error y realizar pruebas con datasets específicos de conducción autónoma para lograr mejoras consistentes en el rendimiento.

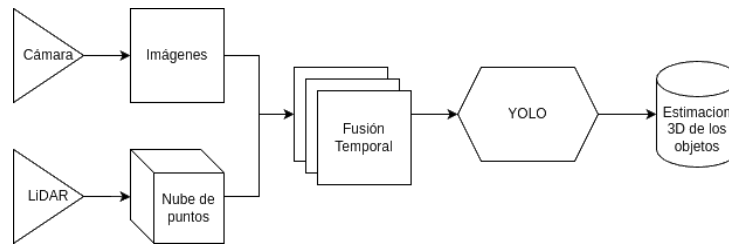


Figura 2: Flujo del sistema de detección y estimación 3D.

3 Metodología y plan de trabajo

Aquí se incluirá una descripción (puede ser incluso una enumeración) clara de las etapas que se van a seguir, y si es posible se deberá incluir un diagrama de Gantt. Por ejemplo algo del estilo a:

Estas son las fases de desarrollo que se van a seguir para la consecución de los objetivos del proyecto descritos en la sección 2:

1. Formación inicial (1 mes)

- Formación en ...
- Consulta de la API ...
- Consulta bibliográfica ...
- Profundización en herramientas de soporte ...
- ...

2. Diseño del entorno ... (0,5 meses)

3. Diseño, implementación y evaluación del ... (2 meses)

- Definición del ...
- Implementación del ...
- Evaluación del ...

4. Implementación y evaluación de ... (1 mes)

5. Diseño, implementación y evaluación del módulo ... (1 mes):

- Definición de ...
- Definición de ...
- Implementación y evaluación ...

6. Integración y ... (1 mes)

7. Documentación ... (0,5 meses)

4 Medios

Aquí se describen de los medios necesarios para realizar el TFG. Por ejemplo:

Las herramientas que van a ser necesarias para desarrollar este proyecto son las siguientes:

- PC compatible
- Sensor ...
- Sistema operativo GNU/Linux **gnulinux**
- Entorno de desarrollo Emacs/Vim **emacs**
- Procesador de textos L^AT_EX **lamport94**
- Control de versiones CVS **cvs**
- Compilador C/C++ gcc **gcc**
- Gestor de compilaciones make **make**
- Robot con movilidad.
- ...

Otros recursos necesarios para la elaboración del proyecto son:

- Herramientas ...
- Sistema de desarrollo ...
- ...