

PLAN DE INVESTIGACIÓN

DATOS DEL DOCTORANDO	
Datos personales	
Apellidos	Nombre
Lucas Zaragoza	Rubén
DNI, NIE, Pasaporte	Email
06282454G	ruben.lucas.zaragoza@gmail.com
Datos Académicos	
Programa de Doctorado en	Línea de investigación
Tecnologías de la Información y las Comunicaciones	Inteligencia artificial, aprendizaje por refuerzo, Aplicación práctica de Deep reinforcement learning en robots autónomos
Título de la tesis	

DATOS DE DIRECCIÓN DE TESIS	
Director de Tesis	
Apellidos	Cañas Plaza
Nombre	José María
DNI	
Email	josemaria.plaza@urjc.es
Departamento/Universidad/Centro	Teoría de la Señal y Comunicaciones y Sistemas Telemáticos/Universidad Rey Juan Carlos
Codirector de Tesis	
Apellidos	Calvo Palomino
Nombre	Roberto
DNI	46850094Z
Email	roberto.calvo@urjc.es
Departamento/Universidad/ Centro	Teoría de la Señal y Comunicaciones y Sistemas Telemáticos/Universidad Rey Juan Carlos
Tutor (si difiere del Director de Tesis)	
Apellidos	
Nombre	
DNI	
Email	
Departamento/Universidad/ Centro	

Los datos personales recogidos serán incorporados y tratados en el fichero *Tesis Doctorales*, cuya finalidad es la gestión de las tesis doctorales, inscrito en el Registro de Ficheros de Datos Personales de la Agencia Española de Protección de Datos (<https://www.agpd.es/portalwebAGPD/index-ides-idphp.php>). El órgano responsable del fichero es el Vicerrectorado de Títulos Propios, Formación Continua y Posgrado, y la dirección donde el interesado podrá ejercer los derechos de acceso, rectificación, cancelación y oposición ante el mismo es, Vicerrectorado de Títulos Propios, Formación Continua y Posgrado, C/Tulipán s/n, 28933-Móstoles, todo lo cual se informa en cumplimiento del artículo 5 de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.

1. Resumen del proyecto de tesis.

Durante la elaboración de esta tesis doctoral se pretende explorar diversas técnicas de inteligencia artificial y aprendizaje automático en el ámbito de la conducción autónoma.

El objetivo es conseguir que un vehículo autónomo realice maniobras de diversa índole de manera satisfactoria, centrándose la investigación en la implementación de algoritmos de reinforcement learning, Deep reinforcement learning y Deep learning como parte de su sistema de decisión.

En una primera instancia se resolverán una serie de problemas básicos en un simulador que nos permitirá la realización de una evaluación objetiva en un entorno controlado.

Posteriormente, usaremos como tarea de referencia el seguimiento de carril basado en percepción visual en el simulador Carla para diseñar una arquitectura y conjunto de técnicas (Técnicas de procesamiento de imagen, técnicas de entrenamiento como curriculum learning, configuración de las acciones del agente, ajuste de hiperparámetros del algoritmo, función de recompensa base, etc) que aplicaremos en un problema más complejo aún con áreas susceptibles de mejora en el estado del arte.

Finalmente abordaremos la transferibilidad de este conocimiento a un entorno real con un robot físico.

A continuación se mencionan algunas alternativas del problema concreto a abordar:

- Negociación de cruces
- Conducción urbana incluyendo semáforos y peatones
- Ejecución de adelantamientos según las condiciones de circulación
- Conducción multiagente en convoys utilizando comunicación entre vehículos
- Configuración dinámica de prioridades en base a un agente preentrenado (combustible, comodidad, velocidad, seguridad, etc.)
- Seguridad, explicabilidad y predecibilidad de las decisiones del agente ante situaciones críticas potencialmente peligrosas

Mejorar cualquiera de estos aspectos representaría un avance significativo en el estado actual de la conducción autónoma, ya que se abordan algunos de los desafíos más relevantes y urgentes en este campo.

Durante la realización de las pruebas se colaborará con el desarrollo de la herramienta RL-Studio, lo cual permitirá aplicar distintos algoritmos de aprendizaje automático a robots en un entorno simulado. La versión utilizada se compondrá de:

- Carla 0.9.15 para simular el entorno, los robots y las físicas que aplicarán al problema
- Ros noetic para comunicarse con los robots
- Frameworks de deep learning que nos permitirán implementar los algoritmos con los que se va a trabajar en esta tesis, entre los que destacan Tensorflow, PyTorch, Tensorboard, stable baselines y OpenCV.
- Integración con la herramienta BehaviorMetrics para evaluar y comparar el rendimiento de los agentes entrenados

2. Justificación del tema de investigación. Antecedentes y estado actual del tema.

La aplicación de inteligencia artificial (IA) para abordar desafíos concretos ha experimentado un crecimiento significativo tanto en el sector privado como en el público, abarcando una amplia gama de industrias, como la medicina, las finanzas, la industria y la automoción. Esta tendencia se debe a la capacidad demostrada de la IA para mejorar la calidad de los productos y servicios, al tiempo que reduce los costos asociados, generando beneficios tanto para las empresas como para la calidad de vida de las personas.

Dentro de este contexto, destaca el aprendizaje por refuerzo, un enfoque que permite a las máquinas aprender de manera similar a los seres humanos, explorando su entorno y obteniendo resultados óptimos. Esto se aplica también a los robots, que pueden aprender mediante ensayo y error, desarrollando políticas óptimas para lograr objetivos específicos. Este avance posibilita la automatización de tareas que anteriormente requerían un esfuerzo significativo, mejorando la eficiencia y precisión en su ejecución.

Una de estas tareas en las que la comunidad lleva tiempo aunando esfuerzos es la conducción autónoma, ámbito en el cual se desarrollará la tesis.

La conducción autónoma es un campo de investigación en constante evolución que promete revolucionar la industria automovilística y mejorar la movilidad en todo el mundo. Esto es así dado su gran potencial para abordar desafíos significativos y beneficiar a la sociedad en términos de seguridad vial, eficiencia energética y comodidad.

A continuación se mencionan algunas de las áreas en las que se han producido progresos significativos en los últimos años

- **Sistemas de percepción basados en aprendizaje profundo:**

Los sistemas de percepción actuales emplean redes neuronales convolucionales (CNNs) y arquitecturas avanzadas como transformers para interpretar datos provenientes de sensores como cámaras, LiDAR y radar. Estos sistemas permiten detectar y clasificar objetos, estimar distancias y realizar seguimiento en tiempo real con alta precisión, superando los métodos tradicionales y constituyendo un pilar fundamental en la conducción autónoma.

- **Sistemas de planificación y control de rutas:**

Los avances en algoritmos de planificación permiten generar trayectorias óptimas y seguras en entornos complejos y dinámicos, considerando restricciones del vehículo y condiciones del tráfico. Técnicas como planificación basada en grafos, métodos de optimización y planificación predictiva modelan escenarios futuros para garantizar un comportamiento fluido y seguro.

- **Comunicación vehicular e Internet de las Cosas (IoT):**

La integración de vehículos con infraestructuras inteligentes y otros vehículos mediante tecnologías V2X (Vehicle-to-Everything) habilita la comunicación en tiempo real que mejora la seguridad y eficiencia del tráfico. Estos sistemas utilizan protocolos de comunicación avanzados y técnicas de IoT para intercambiar información crítica sobre el entorno, anticipando eventos y coordinando maniobras.

- **Sistemas de decisión basados en algoritmos de aprendizaje por refuerzo profundo (DRL):**

Los algoritmos DRL, como DQN, SAC, DDPG [1] y PPO [2], han logrado avances significativos al utilizar redes neuronales profundas para aprender políticas de conducción autónoma interactuando con un entorno dinámico en tiempo real. Estos algoritmos pueden gestionar la toma de decisiones complejas, como la aceleración, el frenado y la dirección.

Recientemente, también se ha comenzado a explorar el uso de grandes modelos de lenguaje (LLMs) para asistir en la toma de decisiones en vehículos autónomos, integrando capacidades de razonamiento y planificación a partir de datos multimodales y contextuales, lo que representa un avance prometedor en la interpretación y gestión de escenarios complejos.

- **Transferencia de Conocimiento de Agentes Simulados a Vehículos en el Mundo Real:**

Este enfoque novedoso implica entrenar modelos de conducción autónoma en simulaciones computacionales realistas y luego transferir este conocimiento a vehículos en el mundo real, lo cual resulta de vital importancia en el momento de sacar provecho de los prometedores resultados obtenidos en simuladores como Gazebo, Carla o Sumo. Esto se logra a través de técnicas de adaptación de dominio y validación cruzada.

A continuación se enumeran una serie de **desafíos** que aún podrían ser abordados:

- **Negociación de cruce:**

Las intersecciones se consideran uno de los escenarios más complejos en un marco de conducción autónoma debido a la incertidumbre en los comportamientos de los vehículos circundantes y los diferentes tipos de situaciones que se pueden encontrar.

Las estadísticas muestran que el 60% de las lesiones graves relacionadas con el tráfico en Europa están vinculadas a las intersecciones. En los EE. UU., la Administración Nacional de Seguridad del Tráfico en Carreteras (NHTSA) publica un informe anual sobre seguridad vial [4]. Estos datos revelan que el 29% de todos los accidentes automovilísticos y el 18% de las muertes de peatones ocurren en las intersecciones.

- **Resolución de circuitos predefinidos con tráfico y peatones:**

Los vehículos autónomos deben reconocer y seguir trayectorias específicas en circuitos complejos que incluyen tráfico dinámico y peatones. Esto requiere un control preciso de la velocidad y la dirección. Dada la importancia de proteger al peatón, que resulta el actor más vulnerable en cualquier escenario, podemos encontrar multitud de investigaciones al respecto que han avanzado recientemente el estado en cuestión.

- **Maximización del nivel de seguridad en la conducción autónoma:**

Uno de los problemas principales que aún no se han resuelto por completo es lograr un alto nivel de seguridad y confiabilidad. Si bien los vehículos autónomos han mostrado resultados prometedores en entornos controlados, navegar por escenarios del mundo real complejos sigue siendo un desafío considerable. El comportamiento humano impredecible, las condiciones climáticas adversas y las carreteras mal señalizadas pueden representar obstáculos significativos para los vehículos autónomos hasta el punto de que empresas como Waymo, Tesla y Cruise aún se encuentran en fase de desarrollo para poder garantizar una solución segura. Lo que ha llevado al lanzamiento de algunos modelos con ciertos niveles de

automatización.

Dicho esto, el daño humano, el daño material y la mala reputación que implican los accidentes [7] de estos vehículos autónomos hace de vital importancia centrar esfuerzos en este punto.

- **Comparación objetiva de rendimiento de distintos algoritmos y técnicas de aprendizaje:**

El hecho de que haya tantas variables implicadas en la conducción, la variedad de escenarios que se pueden abordar, los múltiples objetivos que se pueden definir y las distintas técnicas de entrenamiento, hace complicado aseverar con certeza cuál es la mejor técnica y arquitectura del sistema para alcanzar un agente óptimo. En la literatura encontramos multitud de posibilidades y multitud de maneras diferentes de evaluar el rendimiento del agente. Es por ello que para facilitar el progreso de la conducción autónoma con las mayores garantías posibles, una línea de investigación puede ser comparar metódica y profundamente de cualquiera de los factores implicados en esta (ya sea parte perceptiva, planificación de ruta, ejecución de acciones o explicabilidad de estas para reducir accidentes)

- **Optimización de la conducción de modo coordinado:**

Otra línea de investigación en ciernes es la posibilidad de intercomunicar los diferentes vehículos para, prediciendo la secuencia de acciones que el resto de agentes puede tomar, reducir atascos y evitar ineficiencias en la conducción. A esto se podrían añadir distintas entradas como el estado de la vía o el consumo de combustible para también contribuir a una reducción de accidentes y contaminación.

3. Hipótesis de trabajo y Objetivos

Hipótesis de Trabajo:

- Dando al robot las reglas que rigen el entorno simulado y proporcionándole una serie de estados y acciones, el robot será capaz de aprender y ejecutar acciones óptimas basadas en los estados captados a través de sensores de distinta naturaleza.
- Es posible transferir el conocimiento adquirido por el robot en un simulador realista a un entorno real, lo que implica que las habilidades y comportamientos aprendidos en el simulador serán aplicables y efectivos en situaciones reales.
- La utilización de algoritmos de aprendizaje por refuerzo resultará en una mejora significativa en la eficiencia, seguridad o calidad de la conducción.
- El agente transferido al vehículo real será capaz de generalizar, adaptándose a diferentes entornos que no son exactamente iguales a los utilizados para entrenar en el simulador.

Objetivos:

1. Abordar y resolver problemas específicos de conducción autónoma utilizando en un vehículo simulado técnicas de aprendizaje por refuerzo profundo (PPO, DDPG, DQN, SAC, A3C...).
2. Contribuir al desarrollo de una herramienta (RL-Studio) que permita la adaptación sencilla de algoritmos de aprendizaje por refuerzo a diferentes problemas. Esto incluye la configuración de sensores, acciones, objetivos y la naturaleza del problema, de manera que los usuarios futuros puedan resolver problemas de manera rápida y efectiva sin necesidad de implementar algoritmos desde cero.
3. Abordar un problema práctico que aún no ha sido completamente resuelto por la comunidad investigadora o que tiene margen de mejora significativo, lo que implica enfrentarse a desafíos reales y aplicar soluciones innovadoras. A continuación se mencionan algunas posibilidades:
 - Seguimiento de carril basado en percepción visual
 - Negociación de cruces
 - Conducción urbana incluyendo semáforos y peatones
 - Ejecución de adelantamientos según las condiciones de circulación
 - Conducción multiagente en convoys utilizando comunicación entre vehículos
 - Configuración dinámica de prioridades en base a un agente preentrenado (combustible

- Seguridad, explicabilidad y predecibilidad de las decisiones del agente ante situaciones críticas potencialmente peligrosas

4. Contribuir al avance teórico en el campo del aprendizaje por refuerzo a través de artículos propios.

5. Contribuir al avance práctico en el campo del aprendizaje por refuerzo demostrando la efectividad de las soluciones desarrolladas transfiriendo el modelo al entorno real.

4. Metodología

La metodología que se emplea a lo largo de esta investigación se basa en la realización de experimentos que arrojen resultados objetivos evaluando las diferentes propuestas utilizadas.

Las métricas utilizadas diferirán en función de la naturaleza del problema a resolver. En principio, se prevé la utilización en mayor medida de métricas relacionadas con el aprendizaje profundo por refuerzo, pero es posible que exista la necesidad de utilizar otras métricas que permitan la resolución de problemas intermedios e incluso un análisis más minucioso de los resultados obtenidos como velocidad media del vehículo o distancia recorrida.

Entre las posibles métricas de rendimiento específicas de la resolución del problema mediante aprendizaje por refuerzo, serán utilizadas:

1. Métricas intrínsecas durante el entrenamiento (herramienta RLStudio):

- Recompensa acumulada media durante un episodio
- Sumatorio de la recompensa durante un episodio
- Tiempo que tarda el algoritmo en converger a una solución óptima durante el entrenamiento
- Estabilidad y fiabilidad del entrenamiento en función del algoritmo y técnicas utilizadas
- Porcentaje de episodios completados de manera satisfactoria
- Tiempo medio por episodio en alcanzar el objetivo

2. Métricas extrínsecas específicas durante inferencia (herramienta BehaviorMetrics)

- Capacidad del agente a generalizar (comportarse bien ante situaciones o escenarios no observados durante el entrenamiento)
- Solidez del agente ante perturbaciones externas (lluvia, imprevistos, sensores averiados, etc.)
- Desviación del centro del carril
- Distancia recorrida en un circuito concreto
- Velocidad media del vehículo
- Número y gravedad de infracciones cometidas
- Frecuencia de iteración o frames por segundo

Una vez alcanzado un agente óptimo que resuelva el problema en cuestión en un simulador lo más fiel a la realidad posible, se abordará el reto de transferir el conocimiento adquirido a un entorno real con robots que compartan sensores (inputs) y actuadores (outputs) con el entorno simulado durante el entrenamiento.

5. Revistas científicas en las que tiene previsto de forma tentativa publicar los resultados de la tesis doctoral de acuerdo con lo establecido en el artículo 23.3 de la normativa reguladora de los estudios de doctorado de la URJC.

El tipo de investigación desarrollada, teniendo en cuenta su temática, podría publicarse de forma tentativa en las siguientes revistas:

- Neural Computing and Applications, Springer
- Image and Vision Computing, Elsevier
- Journal of Artificial Intelligence Research
- IEEE Transactions on Robotics
- Journal of Machine Learning Research (JMLR)
- Frontiers in Robotics and AI
- Journal of Intelligent & Robotic Systems

También podría ser expuesta en conferencias como las mencionadas a continuación:

- ICRA, IEEE International Conference on Robotics and Automation
- ROBOT Iberian Robotics Conference
- European Conference on Connected and Automated Driving
- IROS IEEE/RSJ International Conference on Intelligent Robots and Systems

6. Plan de trabajo. Describa brevemente las tareas previstas y los resultados esperados en cada bloqu temporal.

1º Año:

1. Estudio de estado de la cuestión para el problema del comportamiento automático de robots.
 2. Estudio en profundidad de técnicas de deep learning y aprendizaje por refuerzo actuales para tener una base clara sobre la que desarrollar la investigación. Dentro de este estudio incluiremos la realización del curso “machine learning” impartido por Andrew NG en coursera.
 3. Implementación y resolución de un problema básico de resolución de un laberinto primero en un simulador propio y luego en Gazebo utilizando técnicas de aprendizaje por refuerzo clásico como Qlearning y Sarsa.
 4. Primeros experimentos en Gazebo y RLStudio con circuitos con restricciones y soluciones basadas en aprendizaje por refuerzo clásico para un robot terrestre móvil.
-

2º Año:

1. Desarrollo de una herramienta (RL-Studio) que permita la aplicación de algoritmos de aprendizaje por refuerzo clásico y Deep learning para resolver problemas básicos con los que se pretende recabar técnicas que doten al doctorando de las herramientas necesarias para abordar problemas complejos.
 2. Además, este desarrollo dotará a RL-Studio de la potencia necesaria para facilitar la implementación de problemas más complejos que se abordarán en los próximos años. El plan de iteración es el siguiente:
 - Generalización de componentes para facilitar la integración de distintos problemas, algoritmos y simuladores de manera rápida y homogénea.
 - Implementación de un modo inferencia para ejecutar los algoritmos entrenados sin la sobrecarga relativa a la fase de entrenamiento.
 - Entrenamiento y ejecución de agentes en problemas canónicos.
 3. Realización del curso “Algorithms, Part 1” impartido por la universidad de Princeton en coursera.
-

3º Año:

1. Extensión de RL-Studio para probar con distintos simuladores (Gazebo, openAI y CARLA) y algoritmos (entre ellos DQN, DDPG, PPO y A3C).
2. Resolución de problemas básicos propuestos por openAI (*mountain car*, *cartpole* y *pendulum*)
3. Extensión de problemas básicos propuestos por openAI (*cartpole* y *mountain car*) para poder extraer de ellos comparativas y lecciones sólidas acerca de las diferentes técnicas y algoritmos a

aplicar en problemas más complejos. Entre la extensión de estos problemas se incluirá:

- la posibilidad de moldear los entornos ofrecidos por openAI con el objetivo de complicar el problema básico (e.g configuración el mundo simulado a resolver, mejora de las físicas aplicadas, inclusion un ciclo de reloj en el simulador de openAI que obligue al algoritmo a decidir la acción con rapidez, modificación de la naturaleza del problema de acciones discretas acciones continuas)
- Trasladar los problemas a un simulador más realista con físicas más sofisticadas (Gazebo)
- Adición de perturbaciones y condiciones adversas

4. Desarrollo de soluciones a problemas más complejos combinando deep learning y reinforcement learning (entre ellos vehículo con autoaparcamiento, sigue-lineas y sigue-carril)

5. Formalización de los problemas resueltos y los bancos de pruebas realizados con objetivo de servir de base para la resolución de futuros problemas y para ilustrar de manera clara el trabajo realizado de manera que se pueda utilizar con objetivos docentes y en publicaciones de diversa índole. Dichos bancos de pruebas consistirán en:

- Aplicación en RL-Studio de distintos algoritmos (solución programática, QLearning, DQN, DDPG, PPO con acciones discretas y PPO con acciones continuas) a la extensión implementada para el problema cartpole (acciones continuas, físicas mejoradas, aplicación de perturbaciones y condiciones iniciales adversas).
- Aplicación en RL-Studio de distintos algoritmos al problema sigue lineas y sigue carril con el objetivo de comparar su rendimiento haciendo uso de las métricas externas evaluadas en BehaviorMetrics (velocidad efectiva, desviación con respecto al centro y avance espacial)

6. Realización del curso “Deep reinforcement learning class” impartido por la la compañía Hugging Face.

7. Colaboración en la redacción del artículo [26] en el que se describe RL-Studio

4º Año:

1. Refinamiento de los agentes entrenados en años anteriores con objetivo de tener una base lo más sólida posible sobre la que pasar a un problema más complejo en un simulador más realista

2. Aplicación de lo previamente aprendido en un simulador fotorealista y con físicas lo más parecido posible al entorno real (CARLA) con el objetivo de facilitar la transferencia de conocimiento al mundo real

3. Resolución del problema sigue-carril con DDPG y PPO con acciones continuas maximizando la suavidad y velocidad de la conducción en un escenario con líneas discontinuas y curvas de distinta índole en ambas direcciones

4. Extensión de RL-Studio para automatizar el análisis de distintas técnicas de percepción y su impacto en el entrenamiento y el comportamiento del agente

- Módulos de percepción basados en redes neuronales
 - MobileV3Small
 - Yolop
- Módulos basados en técnicas clásicas de percepción
 - Thresholding
 - Linear regresion

5. Realización del curso “Deep Learning Specialization” impartido por Andrew NG en Coursera

5º Año:

1. Extensión de BehaviorMetrics para analizar comportamiento específico de agentes en distintas situaciones (curvas, rectas, intersecciones...)
 2. Refinamiento del comportamiento de agentes para incluir suavidad en la conducción (principalmente en curvas) haciendo uso de *curriculum learning*
 3. Comparativa de distintos algoritmos en el problema de seguimiento de carril a velocidades altas en CARLA (DDPG, PPO y SAC)
 4. Extensión de entrenamientos e inferencias a diferentes ciudades de CARLA para justificar la generalización de los agentes
 5. Definición del problema concreto a abordar en la tesis
 6. Redacción de un artículo en el ámbito de la conducción autónoma con la comparativa de los distintos sistemas de decisión abordados durante este año
-

6º Año:

1. Presentación de un artículo en el ámbito de la conducción autónoma.
 2. Resolución definitiva del problema definido tras haber experimentado durante el año anterior con distintos entornos, técnicas y algoritmos
-

7º Año:

1. Experimentar sobre la viabilidad de trasladar la política aprendida a un robot móvil en el mundo real.
 2. Presentación de un artículo con los resultados de la exploración acerca de la transferencia de conocimiento al mundo real.
 3. Elaboración de la memoria de la tesis.
-

7. Bibliografía relevante relacionada con la investigación

Algunas de las referencias más relevantes relacionadas con la investigación incluyen los siguientes trabajos:

1. H. Shao, L. Wang, R. Chen, H. Li, and Y. Liu, "Safety-Enhanced Autonomous Driving Using Interpretable Sensor Fusion Transformer," in Proc. 6th Conf. Robot Learn. (CoRL), PMLR, vol. 205, pp. 726-737, Dec. 2023.
2. J. Zhang et al., "Multi-Modal 3D Object Detection in Autonomous Driving: A Survey," arXiv preprint, arXiv:2106.12735, 2022.
3. S. Kumar and A. Gupta, "Single Camera Object Detection for Self-Driving Vehicle: A Review," Int. J. Vehicle Autonomous Syst., vol. 10, no. 3, pp. 198-210, Jun. 2022.
4. J. Kim and H. Lee, "End-to-End Deep Learning-Based Autonomous Driving Control for High-Speed Environments," J. Autonomous Syst., vol. 7, no. 4, pp. 345-358, Nov. 2021.
5. K. Zhang, R. Li, X. Yu, and P. Wang, "Design of a Robust System Architecture for Tracking Vehicles on Highways Using Front-Facing Mono Camera," Sensors, vol. 22, no. 9, p. 3359, May 2022.
6. D. Dey, S. Bhatia, and M. Kapoor, "Finite State Machine-Based Control Strategy for Autonomous Vehicles," IEEE Trans. Veh. Technol., vol. 69, no. 5, pp. 5149-5158, 2020.
7. B. R. Kiran et al., "Deep Reinforcement Learning for Autonomous Driving: A Survey," IEEE Trans. Intell. Transp. Syst., vol. 23, no. 6, pp. 4909-4926, 2021.
8. Y. Liu, Q. Zhang, and D. Zhao, "A Reinforcement Learning Benchmark for Autonomous Driving in Intersection Scenarios," arXiv preprint, arXiv:2109.10557, 2021.
9. J. W. Lavington et al., "TorchDriveEnv: A Reinforcement Learning Benchmark for Autonomous Driving with Reactive, Realistic, and Diverse Non-Playable Characters," arXiv preprint, arXiv:2405.04491, 2024.
10. A. Kendall et al., "Learning to Drive in a Day," in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), pp. 8248-8254, 2019.
11. J. Liu, Y. Cui, J. Duan, and H. Li, "Reinforcement Learning-Based High-Speed Path Following Control for Autonomous Vehicles," IEEE Trans. Veh. Technol., early access, 2024. DOI: 10.1109/TVT.2024.3352543.
12. B. Ben Elallid, H. El Alaoui, and N. Benamar, "Deep Reinforcement Learning for Autonomous Vehicle Intersection Navigation," arXiv preprint, arXiv:2310.08595, 2023.
13. M. Cederle, M. Fabris, and G. A. Susto, "A Distributed Approach to Autonomous Intersection Management via Multi-Agent Reinforcement Learning," arXiv preprint, arXiv:2405.08655, 2024.
14. A. Singh, "End-to-end Autonomous Driving using Deep Learning: A Systematic Review," arXiv preprint, arXiv:2308.14358, 2023.

15. J. Valente, C. Ramalho, P. Vinha, C. Mora, and S. Jardim, "Using Machine Learning to Understand Driving Behavior Patterns," *Procedia Comput. Sci.*, vol. 231, pp. 3416-3425, 2024. DOI: 10.1016/j.procs.2024.06.363.
16. K. Zhao et al., "Learning Residual Model of Model Predictive Control via Random Forests for Autonomous Driving," [venue unknown], 2023.
17. F. Codevilla et al., "End-to-End Driving via Conditional Imitation Learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 1314-1322, 2018.
18. J. Cheng, Y. Chen, and Q. Chen, "PLUTO: Pushing the Limit of Imitation Learning-based Planning for Autonomous Driving," *arXiv preprint, arXiv:2310.00960*, 2023.
19. A. Moncalvillo et al., "Imitation Learning for Vision-Based Autonomous Driving with Ackermann Cars," in *Proc. XXIV Workshop Physical Agents*, pp. 206-219, Univ. Alicante, 2024.
20. Y. Bengio et al., "Curriculum Learning," in *Proc. 26th Int. Conf. Mach. Learn. (ICML)*, pp. 41-48, 2009.
21. A. Graves et al., "Automated Curriculum Learning for Neural Networks," in *Proc. 34th Int. Conf. Mach. Learn. (ICML)*, pp. 1311-1320, 2017.
22. B. Uppuluri et al., "CuRLA: Curriculum Learning Based Deep Reinforcement Learning for Autonomous Driving," *arXiv preprint, arXiv:2501.03576*, 2025.
23. Z. Sheng et al., "CurricuVLM: Towards Safe Autonomous Driving via Personalized Safety-Critical Curriculum Learning with Vision-Language Models," *arXiv preprint, arXiv:2404.09991*, 2024.
24. Z. Gu et al., "Safe-state enhancement method for autonomous driving via direct hierarchical reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, 2023. DOI: 10.1109/TITS.2023.3272040.
25. L. M. Schmidt et al., "Can you trust your autonomous car? Interpretable and verifiably safe reinforcement learning," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, pp. 171-178, 2021.
26. B. Gangopadhyay et al., "Hierarchical program-triggered reinforcement learning agents for automated driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10902-10911, 2021.
27. Y. Sun et al., "YOLOP: You Only Look Once for Panoptic Driving Perception," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022.
28. A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv preprint, arXiv:1704.04861*, 2017.
29. W. Liu et al., "SSD: Single Shot MultiBox Detector," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016.
30. J. Wang, Q. Zhang, and D. Zhao, "Benchmarking Lane-changing Decision-making for Deep Reinforcement Learning," [venue unknown], 2024.
31. T. Liu et al., "A Comparative Analysis of Deep Reinforcement Learning-Enabled Freeway Decision-Making for Automated Vehicles," *IEEE Access*, 2024.
32. D. Chen et al., "Deep Multi-Agent Reinforcement Learning for Highway On-Ramp Merging in Mixed Traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, pp. 11623-11638, 2023.
33. L. Cui et al., "Multi-input autonomous driving based on deep reinforcement learning with double-

bias experience replay," IEEE Sens. J., vol. 23, no. 8, pp. 1234-1243, 2023.

34. W. Zhou et al., "Multi-agent reinforcement learning for cooperative lane changing of connected and autonomous vehicles in mixed traffic," IEEE Trans. Veh. Technol., vol. 73, no. 2, pp. 956-967, 2024.

35. X. Li et al., "A Deep Reinforcement Learning-Based Approach for Autonomous Overtaking," IEEE Trans. Veh. Technol., vol. 69, no. 12, pp. 14182-14193, 2020.

36. J. Pan, H. Liu, and Q. Sun, "TORCS-Based Evaluation of Continuous Control RL Algorithms," Int. J. Adv. Robot. Syst., 2023.

37. W. Li et al., "Reinforcement Learning Algorithms for Lane Changing Decision-Making in Autonomous Vehicles," arXiv preprint, arXiv:2407.02521, 2024.

38. L. Zhang and Y. Chen, "DQN vs DDPG: A Performance Comparison in Carla Using Waypoints and Visual Input," Multimedia Tools Appl., 2021.

39. J. Wang and F. Liu, "High-Speed Autonomous Driving via Latency-Aware Deep Reinforcement Learning," Sensors, 2024.

40. T. Nguyen and R. Han, "Delay-Aware Cooperative Adaptive Cruise Control with Multi-Agent RL," arXiv preprint, arXiv:2404.15696, 2024.

41. H. Chen and Z. Sun, "Eco-Cooperative Adaptive Cruise Control Using Reinforcement Learning," Transp. Res. Part D, 2025.

42. J. Fritsch, T. Kuehnl, and A. Geiger, "A New Performance Measure and Evaluation Benchmark for Road Detection Algorithms," in Proc. IEEE Int. Conf. Intell. Transp. Syst. (ITSC), 2013.

43. M. Bansal, A. Krizhevsky, and A. Ogale, "ChauffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst," arXiv preprint, arXiv:1812.03079, 2018.

44. A. Geiger, P. Lenz, and R. Urtasun, "Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2012.

45. J. Phillion and S. Fidler, "Improved Single Camera BEV Perception Using Multi-Camera Training," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2023.

46. D. Unger et al., "Multi-camera Bird's Eye View Perception for Autonomous Driving," arXiv preprint, arXiv:2309.09080, 2023.

47. D. Feng, L. Rosenbaum, and K. Dietmayer, "Towards Safe and Efficient Rule-Based Decision Making in Autonomous Driving," in IEEE Intell. Vehicles Symp., 2020.

48. Y. Kwon et al., "A Hybrid Decision Support System for Adaptive Trading Strategies," Decis. Support Syst., vol. 177, p. 114100, 2024.

49. D. Sadigh et al., "Formal Methods for Safe Autonomous Driving: Explicit FSM Design," in IEEE Intell. Vehicles Symp., 2016.

50. Z. Li et al., "UniMODE: Unified Monocular 3D Object Detection," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 16561-16570, 2024.

51. M. R. Barın, G. Aydemir, and F. Güney, "Robust Bird's Eye View Segmentation by Adapting

DINOv2," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2nd Workshop on Vision-Centric Autonomous Driving, 2024.

52. S.-H. Bae et al., "Finite State Machine based Vehicle System for Autonomous Driving in Urban Environments," in Proc. Int. Conf. Control, Autom. Syst. (ICCAS), 2020.

53. A. Dosovitskiy et al., "CARLA: An Open Urban Driving Simulator," in Proc. 1st Conf. Robot Learn. (CoRL), 2017.

54. E. Paniego et al., "Behavior Metrics: A Toolkit for Evaluating Decision-Making in Autonomous Driving," IEEE Trans. Intell. Vehicles, 2024.

55. F. Bouchard, S. Sedwards, and K. Czarnecki, "A Rule-Based Behaviour Planner for Autonomous Driving," in IEEE Intell. Vehicles Symp. (IV), pp. 611-616, 2015.

56. X. He and C. Lv, "Towards Safe Autonomous Driving: Decision Making with Observation-Robust Reinforcement Learning," Automot. Innov., vol. 6, no. 3, 2023. DOI: 10.1007/s42154-023-00256-x.

57. S. Wu et al., "Continuous Decision-Making in Lane Changing and Overtaking Maneuvers for Unmanned Vehicles," IEEE Trans. Intell. Vehicles, vol. 9, no. 4, pp. 4657-4674, 2024.

58. J. Funke et al., "Collision Avoidance and Stabilization for Autonomous Vehicles in Emergency Scenarios," IEEE Trans. Control Syst. Technol., [year incomplete].

59. Paniego, S., Lucas, R., Fernández, P., Cañas, J. M., & Arranz, I. (Título: RL-Studio: A tool for reinforcement learning methods in robotic)

Fecha y Firma del
doctorando:

Rubén