|  |
| --- |
| **Universidad Internacional de La Rioja (UNIR)**  **ESIT**    **Máster Universitario en Industria 4.0** |
|  |
| Detección de contaminación marina mediante la recogida, análisis y procesado de imágenes |

Logotipo

Descripción generada automáticamente

**Trabajo Fin de Máster**

**presentado por:** Álvarez Martín, Rubén

**Director/a:** Sánchez Lugilde, Nuria

Resumen

Abstract

Índice de contenidos

[1. Introducción 9](#_Toc111329320)

[1.1. Motivación 9](#_Toc111329321)

[1.2. Planteamiento del proyecto 10](#_Toc111329322)

[1.3. Estructura de capítulos 11](#_Toc111329323)

[2. Contexto y estado del arte 12](#_Toc111329324)

[2.1. Descripción general del contexto del proyecto 12](#_Toc111329325)

[2.2. Proyectos relacionados con el tema del TFM 15](#_Toc111329326)

[2.2.1. Sensor espacial para detectar residuos 15](#_Toc111329327)

[2.2.2. Drones para detectar contaminación marina 15](#_Toc111329328)

[2.2.3. Drones para localizar vertidos en el mar 15](#_Toc111329329)

[2.3. Tecnologías relacionadas con el tema del TFM 16](#_Toc111329330)

[*2.3.1.* *Machine learning* 16](#_Toc111329331)

[*2.3.2.* *Big data* 16](#_Toc111329332)

[2.3.3. Robótica 18](#_Toc111329333)

[2.3.4. MQTT 22](#_Toc111329334)

[2.4. Conclusiones sobre el estado del arte 24](#_Toc111329335)

[3. Descripción general de la contribución del TFM 25](#_Toc111329336)

[3.1. Objetivos 25](#_Toc111329337)

[3.2. Metodología del trabajo 25](#_Toc111329338)

[3.3. Descripción general de las partes o componentes de la propuesta 26](#_Toc111329339)

[3.3.1. Alcance y limitaciones 26](#_Toc111329340)

[3.3.2. Tecnologías implicadas 26](#_Toc111329341)

[3.3.3. Arquitectura, componentes e integración de tecnologías. 27](#_Toc111329342)

[3.3.4. Resultados esperados 34](#_Toc111329343)

[3.3.5. Presupuesto 35](#_Toc111329344)

[3.3.6. Planificación general 37](#_Toc111329345)

[4. Desarrollo específico de la contribución 38](#_Toc111329346)

[4.1. MARLIT 38](#_Toc111329347)

[4.1.1. Organización del proyecto 38](#_Toc111329348)

[4.1.2. Ejecución de algoritmo 39](#_Toc111329349)

[4.1.3. Conclusiones del proyecto 40](#_Toc111329350)

[4.2. Métodos de prevención 42](#_Toc111329351)

[4.2.1. Actuaciones en caso de contaminación baja 42](#_Toc111329352)

[4.2.2. Actuaciones en caso de contaminación media 42](#_Toc111329353)

[4.2.3. Actuaciones en caso de contaminación alta 43](#_Toc111329354)

[4.3. Algoritmo de detección de plásticos 43](#_Toc111329355)

[4.3.1. Descripción general del algoritmo 43](#_Toc111329356)

[4.3.2. Desarrollo del algoritmo 46](#_Toc111329357)

[4.3.3. Conclusiones y resultados 46](#_Toc111329358)

[5. Conclusiones y trabajos futuros 47](#_Toc111329359)

[5.1. Conclusiones 47](#_Toc111329360)

[5.2. Líneas de trabajo futuras 47](#_Toc111329361)

[Referencias bibliográficas 48](#_Toc111329362)

Índice de figuras

[Figura 1. Evolución de la producción de plástico a nivel mundial. 14](#_Toc111481780)

[Figura 2. Plásticos en el mar actualmente y predicción 14](#_Toc111481781)

[Figura 3. Relación entre tecnologías de IA 17](#_Toc111481782)

[Figura 4. Plano elaborado por un aspirador Roomba 19](https://alumnosunir-my.sharepoint.com/personal/ruben_alvarez470_comunidadunir_net/Documents/TFM-industria.docx#_Toc111481783)

[Figura 5. Robot Da Vinci en un quirófano 20](#_Toc111481784)

[Figura 6. Vehículo Tesla con imágenes de sus sensores 20](#_Toc111481785)

[Figura 7. Robot de logística en un almacén 21](#_Toc111481786)

[Figura 8. Robot Curiosity en la superficie de Marte 21](#_Toc111481787)

[Figura 9. Protocolo MQTT 23](#_Toc111481788)

[Figura 10. Diagrama de Gantt de los *sprints* planificados 26](#_Toc111481789)

[Figura 11. Arquitectura de tecnologías 27](#_Toc111481790)

[Figura 12. Drone aéreo para captura de imágenes aéreas y captura de datos 28](#_Toc111481791)

[Figura 13. Drone submarino para captura de imágenes submarinas 29](#_Toc111481792)

[Figura 14. Drone de recogida de residuos 29](#_Toc111481793)

[Figura 15. Imagen de polución capturada por cámara infrarrojos 30](#_Toc111481794)

[Figura 16. Imagen capturada por una cámara a bordo de un drone 31](#_Toc111481795)

[Figura 17. Representación en Grafana de temperatura y humedad 32](#_Toc111481796)

[Figura 18. Arquitectura del sistema planteado 33](#_Toc111481797)

[Figura 19. Organización de proyecto MARLIT 39](#_Toc111481798)

[Figura 20. Modelos y funciones cargadas en RStudio 41](#_Toc111481799)

[Figura 21. Entrenamiento con un dispositivo 41](#_Toc111481800)

[Figura 22. Diferentes clasificaciones de Deep Learning 44](#_Toc111481801)

[Figura 23. Búsqueda manual de imágenes en Google 45](#_Toc111481802)

[Figura 24. Descarga de dependencias al ejecutar el código Python 46](#_Toc111481803)

[Figura 25. Descarga de imágenes realizando una búsqueda en Google 46](#_Toc111481804)

[Figura 26. Eliminación de imágenes duplicadas 47](#_Toc111481805)

[Figura 27. Fotografía descartada para el algoritmo de detección 47](#_Toc111481806)

[Figura 28. Fotografía empleada para la detección de plásticos 48](#_Toc111481807)

Índice de tablas

[Tabla 1. Diferencias de Big Data y machine learning 18](#_Toc107526856)

[Tabla 2. Presupuesto de hardware 31](#_Toc107526857)

[Tabla 3. Presupuesto de software 32](#_Toc107526858)

[Tabla 4. Presupuesto total 33](#_Toc107526859)

# Introducción

En este capítulo, se introducen los conceptos básicos en los que se apoya este proyecto. Se explican y se ponen en contexto las tecnologías que permiten llevar a cabo la detección de basuras y residuos, el estado actual de la robótica (con especial mención a los drones), los sensores usados en este tipo de sistemas (particularmente cámaras y sensores) y cómo se ha llegado al punto en el que se encuentran.

## Motivación

En la actualidad, la contaminación del planeta se ha convertido en un problema no sólo atmosférico, sino que también afecta al suelo y a los océanos, llegando a todos los seres vivos que habitan en él. La emergencia climática se extiende a todo el ecosistema y biosfera, generando consecuentemente, gases de efecto invernadero, temperaturas extremas, escasez de agua, degradación del suelo, incremento de los incendios o fenómenos meteorológicos como tifones o huracanes. A esto se suma el deshielo de los polos, la deforestación, la pérdida de biodiversidad debido a la desaparición de sus hábitats naturales que, causado por la actividad humana, provoca una subida de temperaturas a nivel global.

Es por este deterioro del planeta, causado por la actividad y mala gestión del ser humano, que resulta esencial materializar acciones para cambiar el rumbo de la catástrofe medioambiental.

La contaminación de mares y océanos es un problema que afecta a todo el hábitat oceánico. Contaminación marina (Euroinnova Business School, 2022) se refiere a derrames de petróleo, vertidos ilegales, de plásticos, de residuos tóxicos, redes fantasmas[[1]](#footnote-2), aguas residuales, etc.

Este proceso comenzó en la década de los 70, cuando debido a la gran cantidad de agua en los océanos, se creía que tenía la capacidad suficiente de poder diluir todos los contaminantes sin generar ningún tipo de consecuencias. Es en esta época cuando se vertieron una gran cantidad de químicos y productos tóxicos y se han ido acumulando en las aguas y cadenas tróficas.

A eso se debe la urgencia en investigar la manera de detectar el nivel de plásticos y contaminación marina, generada sobre todo por plásticos y por la polución de los barcos y buques que navegan los océanos. Gran parte de esta contaminación es generada por los barcos de pesca, que generan de 600 a 1500 millones de toneladas de CO2 debido a las grandes redes que arrastran por el fondo del mar. Estas redes son también las responsables de gran parte de los plásticos que hay en el océano, generando 580000 toneladas de plástico vertidas en el mar. El 46% de plásticos que forman parte de la Isla de Plástico del Pacífico (de la que se hablará en el próximo capítulo) tiene su origen en las redes de pesca. Más adelante se profundizará en este tema.

En conclusión, la finalidad de este proyecto es analizar la contaminación marina y los residuos plásticos en los océanos. Consistirá en utilizar drones que capturarán imágenes aéreas, procesándolas y tratándolas para su posterior análisis. Con los datos obtenidos tendremos la posibilidad de actuar inmediatamente sobre las zonas más perjudicadas, ya sea con drones de limpieza o protegiendo la zona para evitar la navegación de buques y/o navíos pesqueros.

Esto podría convertirse en una valiosa herramienta para trabajar contra el cambio climático y sus consecuencias medioambientales.

## Planteamiento del proyecto

Como se ha dicho, se parte de la necesidad y la urgencia de reducir la contaminación en el océano, particularmente el nivel de plásticos y micro plásticos. En este proyecto se van a delimitar áreas oceánicas concretas, tomando imágenes de estas. Esto va a ser posible gracias a drones con cámaras y sensores de infrarrojos para así poder detectar la contaminación provocada por los navíos que frecuenten las zonas. Posteriormente, se procederá a categorizar las zonas según la contaminación que tengan, tanto ambiental como residual, lo que posibilite hacer efectiva una respuesta adecuada sobre aquellas que presenten mayor nivel de contaminación.

En primera instancia, se obtendrán las imágenes para después investigar y probar distintos algoritmos de *machine learning*, tanto en las imágenes capturadas con cámaras normales como los datos obtenidos con los sensores de infrarrojos para, en última instancia, emplear el algoritmo que mejor se adapte a los datos obtenidos. También se tendrán en cuenta muestras de agua para incluirlas en los algoritmos y estudiar si mejoran la predicción y clasificación de muestras.

## Estructura de capítulos

A continuación, se muestra un breve resumen del contenido de cada capítulo:

**Capítulo 1: Introducción y objetivos.** Es un capítulo introductorio que se divide en tres partes. La primera parte justifica y pone en contexto el tema que se trata en el trabajo. La segunda parte presenta el planteamiento del trabajo y la metodología usada para su consecución. La última parte resume la estructura del trabajo por capítulo.

**Capítulo 2: Contexto y estado del arte.** Se pone en contexto el tema tratado, tanto con trabajos y artículos relacionados como describiendo los conceptos de *big data*, *machine learning* y robótica, así como distintos algoritmos para tratar las imágenes (como pueda ser *monoSLAM* o *PTAM*) y poder obtener información útil para su posterior análisis de ellas.

**Capítulo 3: Descripción general de la contribución del proyecto.** Diseño y creación de las pruebas de algoritmos y estudio de la posibilidad de realizar análisis de muestras marinas para mejorar la detección de micro plásticos.

**Capítulo 4: Desarrollo específico de la contribución.** Se explica el proceso seguido para realizar el análisis de las imágenes mediante algoritmos de *machine learning* y se muestran los resultados obtenidos comparando con distintos algoritmos empleados y con distintas imágenes que permitan realizar gran número de pruebas.

**Capítulo 5: Conclusiones y líneas futuras.** Se desarrollan las conclusiones que se derivan del trabajo y se analizan posibles líneas futuras del trabajo.

**Bibliografía.** Consiste en un listado de referencias consultadas para la realización de este trabajo.

# Contexto y estado del arte

En este capítulo se introducen los conceptos básicos en los que se apoya el proyecto. Se explican y se pone en contexto el estado actual de las herramientas *machine learning* y *big data*, así como el estado actual de la robótica y su expansión hasta llegar al punto en el que se encuentra, haciendo especial mención a drones y a captura y tratamiento de imágenes.

## Descripción general del contexto del proyecto

Según la Real Academia Española (RAE) la contaminación se define como “*alterar nocivamente la pureza o las condiciones normales de una cosa o un medio por agentes químicos o físicos*”. Esta alteración del planeta se ha ido fraguando desde los inicios del ser humano tal y como lo conocemos hoy. En la historia ha habido varios sucesos que implicaron desastres medioambientales, se describen en la referencia (EROSKI Consumer, 2014), y algunos de los más importantes son los siguientes:

* Accidente nuclear de Chernóbil. El accidente de la central nuclear Vladímir Ilich Lenin en 1986 es considerado el accidente más grave en la Escala Internacional de Accidentes Nucleares. Millones de personas sufrieron sus efectos en diversos grados.
* Contaminación en el delta del Río Níger. Derrames de la industria petrolera están provocando la destrucción de la riqueza del humedal más grande de África.
* Derrame de petróleo en en Golfo de México. La plataforma petrolífera Deepwater Horizon explotó en 2010 y derramó aproximadamente la cantidad de casi 5 millones de barriles, siendo considerado uno de los peores vertidos petroleros de la historia.
* Desaparición del mar de Aral. Los sistemas de riego provocaron que la superficie del lago se haya reducido un 90%, con la pérdida de ecosistemas como consecuencia directa.
* Destrucción del Amazonas. El 20% de la selva amazónica se ha perdido en las últimas décadas por la acción humana y la deforestación y prácticas agrícolas.
* Escape químico en Bhopal. En 1984 una planta de pesticidas de la ciudad india sufrió una fuga de diversos gases y productos químicos que provocó la muerte de unas 3000 personas y enfermedades en 50000.
* Fuga de dioxinas en Seveso. En 1976 una explosión en una planta de fabricación de pesticidas de Lombardía provocó una masa de vapores de una clase de dioxina. Se sacrificaron más de 80000 animales para evitar la contaminación de la cadena alimentaria y más de 37000 personas fueron expuestas a esta sustancia.
* Impactos ambientales múltiples en el lago Victoria. Este lago, situado en la zona centro-oriental de África, es un centro de impactos medioambientales: contaminación química, residuos, sobrepesca, especies invasoras y algas que provocan reducción de agua y disminución de flora y fauna.
* Incendios en los pozos petroleros de Kuwait. En 1991 el ejercito iraní incendió varios pozos petroleros, causando una enorme contaminación del suelo y del aire.
* Mar de basura plástica en el Pacífico. Se trata de una isla de basura que flota frente al archipiélago de Hawái y triplica el tamaño de Francia. Contiene 1,8 billones de trozos de plástico y 80000 toneladas de plástico, matando a miles de animales marinos de las aguas cercanas.
* Marea negra del Exxon Valdez. En 1989 un petrolero encalló en un arrecife de Alaska, provocando una de las peores mareas negras de la historia. Tras este accidente, las medidas para el transporte petrolero fueron endurecidas considerablemente.
* Vertedero electrónico. La localidad de Guiyu en China tiene el mayor vertedero de residuos de aparatos eléctricos y electrónicos del mundo. Además, no cuenta con las condiciones necesarias para su correcto tratamiento, por lo que provoca envenenamiento por plomo y abortos involuntarios a la población de la región.

Estos episodios, entre otros muchos, han ocurrido en nuestro planeta tanto por errores humanos o acciones intencionadas. Cada acción tiene su consecuencia y a medida que avanzamos en conocimiento vemos que las consecuencias de estas acciones alteran el planeta de forma significativa. Resulta obvio que la única manera de revertir la emergencia climática que nos rodea es realizar acciones de forma inmediata, como cambiar hábitos de consumo y aplicar una economía más sostenible.

Hay que añadir que la producción de plástico aumenta año tras año, tal y como se muestra en la Figura 1. Según las estimaciones, el vertido de plásticos a los océanos se triplicará en 2040 si no se actúa. Siendo como muestra la Figura 2: en 2016 se gestionan 129 millones de toneladas y se quedan sin gestionar 91 millones, esta cifra se verá incrementada casi tres veces su valor en 2040.

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Figura 1. Evolución de la producción de plástico a nivel mundial.

Fuente: Statista, 2021

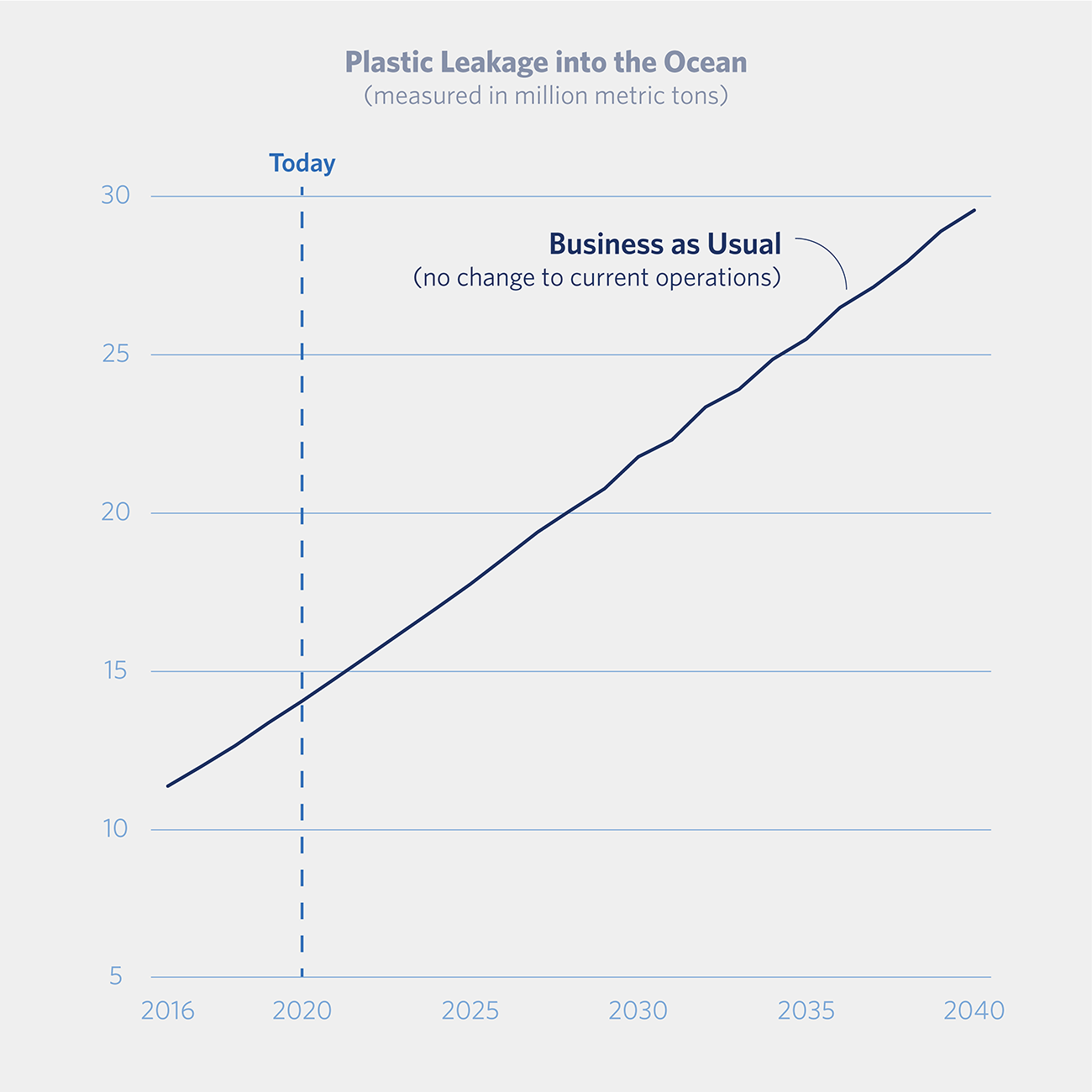


Figura 2. Plásticos en el mar actualmente y predicción

Fuente: Pewtrust, 2020

Además, debido a la exposición del sol, agua y aire el plástico se rompe en microplásticos, que va a ser comido por peces, pájaros o mamíferos. Cada año entre 12000 y 24000 toneladas de plásticos es consumida por peces en el Pacífico Norte, causando malestar estomacal llegando a provocar su muerte y afectando a toda la cadena alimenticia.

## Proyectos relacionados con el tema del TFM

### Sensor espacial para detectar residuos

En este trabajo (Méndez, I., 2020) llevado a cabo por la empresa Orbital EOS, se realiza la detección automática de vertidos contaminantes mediante imágenes de satélites como los de la Agencia Espacial Europea y la Nasa. El sistema permite detectar de forma eficiente y fiable la contaminación en espacios marinos combinando imágenes de satélite, técnicas de Inteligencia Artificial y Big Data. La obtención de imágenes es en tiempo real y con una resolución multiespectral muy amplia, lo cual permite distinguir manchas de petróleo con otras manchas provocadas por fenómenos naturales, como arrecifes o algas.

### Drones para detectar contaminación marina

En este proyecto (Garcia-Garin, O., Monleón-Getino, T., López-Brosa, P., Borrell, A., Aguilar, A., Borja-Robalino, R., Cardona, L., & Vighi, M., 2021) llevado a cabo por la plataforma *Bioinformatics* Barcelona usan técnicas de inteligencia artificial y más de 3700 imágenes de la costa de Cataluña y, a través de *Deep learning*, evalúan la presencia, densidad y distribución de los microplásticos en los mares y océanos del mundo. Han empleado imágenes de la superficie marina obtenidas por drones y avionetas en campañas de monitorización de basura y han desarrollado un algoritmo de *Deep learning* programado en R. Su objetivo es integrar la herramienta que han programado y han bautizado como MARLIT (AZTI, 2021) en sensores como drones o cámaras de alta resolución. Según sus cifras, alcanza un 80% de precisión en detección de macrorresiduos marinos.

### Drones para localizar vertidos en el mar

El proyecto *Brainport* (Tecnología, A. 2017) responde a la necesidad de localizar un vertido en el mar cuando este se produce. Emplea drones y sistemas inteligentes para advertir de la presencia de sustancias contaminantes mediante drones, que mandan datos a un software inteligente para analizarlo y marcar diferentes patrones: barcos, costa, vertidos de crudo, etc.

## Tecnologías relacionadas con el tema del TFM

En este apartado se detallan los habilitadores técnicos y tecnologías relacionadas con este TFM.

### *Machine learning*

*Machine learning* (Iberdrola, 2022) es una disciplina del campo de la Inteligencia Artificial que, mediante algoritmos, otorga a los computadores capacidad de identificar patrones y elaborar predicciones o tomar decisiones. Esto permite a los ordenadores aprender y realizar tareas de forma autónoma sin necesidad de ser programados.

Entre los algoritmos más empleados de *machine learning* están los siguientes:

* Aprendizaje supervisado. Realizan un aprendizaje previo basado en etiquetas asociados a unos datos que permite hacer predicciones o tomar decisiones. Como ejemplo se puede tomar un detector de spam que etiqueta los correos dependiendo del histórico y sus datos asociados como remitente, imágenes, asunto, etc.
* Aprendizaje no supervisado. Estos algoritmos no tienen ningún aprendizaje previo y analizan datos desordenados en búsqueda de patrones que permita organizarlos y ordenarlos. Por ejemplo, para extraer datos de las redes sociales para crear campañas orientadas a un público objetivo.
* Aprendizaje por refuerzo. El algoritmo aprende a partir de la experiencia y es capaz de tomar la mejor decisión ante diferentes situaciones. Este proceso lo lleva a cabo mediante acciones de prueba-error en el que se recompensan las decisiones correctas. Se emplea, por ejemplo, en diagnósticos médicos para predecir la tendencia de los pacientes a ciertas enfermedades.

### *Big data*

*Big data*, según Oracle (Oracle España, 2022), “*son conjuntos de datos de mayor tamaño y más complejos, especialmente procedentes de nuevas fuentes de datos. Estos conjuntos de datos son tan voluminosos que el software de procesamiento de datos convencional sencillamente no puede gestionarlos*”. Tradicionalmente el volumen de datos que se procesaban era ínfimo comparado con lo que permite, tanto el *software* como el *hardware* actualmente*.*

*Big data* se puede resumir en las siguientes características:

* Volumen. La cantidad de datos es importante y tendrán que procesarse grandes volúmenes de datos no estructurados. Pueden ser datos de valor desconocido como clics en una página web, datos obtenidos de sensores o datos de miles de pacientes de un centro sanitario.
* Velocidad. Es el ritmo en el que se reciben los datos y al que se aplican acciones sobre ellos. Algunos productos funcionan a tiempo real y requieren acciones y evaluaciones de manera inmediata.
* Variedad. Los datos se presentan de forma no estructurada o semiestructurados, como texto, audio, video o imagen. Estos datos requieren preprocesamiento adicional para poder tratarlo.

Tanto Big Data como Machine learning forman parte de Inteligencia Artificial. En la Tabla 1 se explican las diferencias entre ellos. En la Figura 3 se representa en un diagrama de Venn que muestra cómo se relacionan las tecnologías que pertenecen a la Inteligencia Artificial.

Diagrama

Descripción generada automáticamente

Figura 3. Relación entre tecnologías de IA

Fuente: Boy, G. (2022)

|  |  |  |
| --- | --- | --- |
| Características | Big Data | Machine learning |
| Tratamiento del dato | Extrae y analiza la información de grandes volúmenes de datos | Utiliza datos de entrada y algoritmos para estimar resultados |
| Tipos | Estructurados, no estructurados y semiestructurados | Aprendizaje supervisado, no supervisado y por refuerzo |
| Interacción humana | Necesaria | No es necesaria |
| Característica principal | Extrae datos sin procesar y busca patrones en ellos. | Aprende de los datos de entrenamiento y actúa como un ser humano para hacer predicciones |
| Usos | Análisis de acciones, análisis de mercado, etc. | Asistencia virtual, recomendaciones de productos, filtrado de spam en el correo, etc |
| Alcance | Manejo de grandes volúmenes de datos y optimización del almacenamiento | Mejorar la calidad de las predicciones, toma de decisiones más rápida y mejora de robots y servicios médicos. |

Tabla 1. Diferencias de Big Data y machine learning

### Robótica

La robótica es una rama tecnológica encargada del diseño y construcción de aparatos que realizan operaciones y trabajos en sustitución de la mano de obra humana. A estos aparatos que se les llama robot y se puede definir como un sistema autónomo programable capaz de realizar tareas de ayuda al ser humano y con aplicaciones en campos diversos como la medicina, el hogar, las fábricas, etc.

Los robots se componen de:

* Sensores. Dispositivos que se encargan de recoger información del entorno. En este grupo se encuentran láseres, cámaras, ultrasonidos u odómetros. Equivalen a los sentidos humanos.
* Controladores. Son losencargados de analizar los datos recogidos por los sensores y elaboran una respuesta que va a ser enviada a los actuadores. En los seres humanos equivale al cerebro.
* Actuadores. Componente electrónico encargado de transformar energía eléctrica, hidráulica o neumática en mecánica. Son los que interactúan con el entorno y equivalen a los músculos humanos.

La robótica es ya una realidad. Los robots no son utilizados únicamente por empresas para labores industriales, si no que se pueden encontrar en la vida cotidiana de las personas. Esto es debido al aumento de eficiencia que generan, reducción de costes o el control de errores que aportan. Algunos ejemplos de robots empleados en la actualidad, tanto en ámbito doméstico como en industrial, son los siguientes:

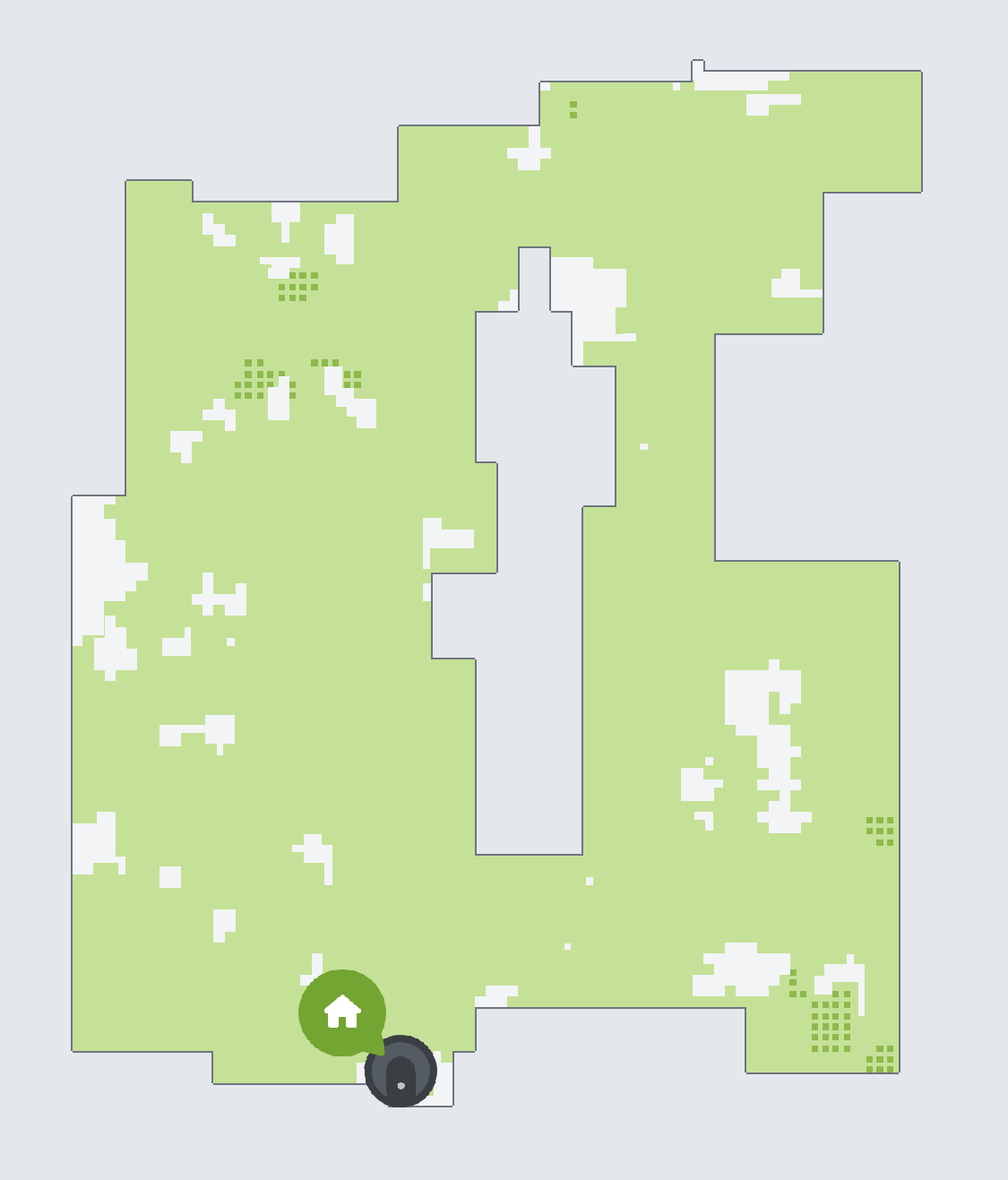
* Aspiradora robótica *Roomba*, robot autónomo que, en los modelos más avanzados, reconocen el entorno, trazan un mapa de la casa (Figura 4) y vacían su depósito automáticamente.

Figura 4. Plano elaborado por un aspirador Roomba

* Robot médico *Da Vinci* (Figura 5), que permite al cirujano operar a través de una consola mejorando así su precisión y reduciendo riesgos en operaciones quirúrgicas.

Imagen que contiene interior, cuarto de hospital, tabla, hombre

Descripción generada automáticamente

Figura 5. Robot Da Vinci en un quirófano

* Vehículos autónomos Tesla (Figura 6), que mediante cámaras y sensores de ultrasonidos analizan el entorno para garantizar una conducción autónoma segura.



Figura 6. Vehículo Tesla con imágenes de sus sensores

* Robots de logística de *Amazon* (Figura 7), que se encargan, de manera autónoma, de localizar la estantería donde se encuentra un paquete solicitado y la desplaza por todo el centro logístico al lugar de destino.

Imagen que contiene edificio, interior, biblioteca, pequeño

Descripción generada automáticamente

Figura 7. Robot de logística en un almacén

* Robot *Curiosity* de la *NASA* (Figura 8), empleado para la exploración de la superficie del planeta Marte.



Figura 8. Robot Curiosity en la superficie de Marte

Para dotar de inteligencia autónoma a los robots se requiere desarrollar sistemas complejos, aplicaciones e infraestructuras. Por ejemplo, hace años, el desarrollo de *software* se realizaba adoptando soluciones *ad-hoc* dotando a cada robot de un diseño específico y con sensores y actuadores concretos. Esto implicaba que había que implementar todo el *software* para un nuevo robot debido a que no se podía aplicar el desarrollado anteriormente. En la actualidad, existen plataformas que permiten el desarrollo de aplicaciones robóticas de forma eficiente y genérica permitiendo así reutilizar aplicaciones creadas en otros robots.

Para proporcionar inteligencia a un robot se desarrolla un *software* que habitualmente se programa con ayuda de herramientas, como los *middleware*s robóticos. El uso de estas herramientas permite introducir una capa de abstracción con los *drivers* y el *hardware* del robot, reduciendo la complejidad y los conocimientos necesarios para realizar desarrollos. Algunos de estos *middlewares* son *Robot Operating System* o ROS (ROS: Home, 2022), *Offshore Robotics for the Certification of Assets* u ORCA (Orca Robotics, 2022) y *Open Robot Control Software* u OROCOS (The Orocos Project, 2022).

### MQTT

MQTT (*MQ Telemetry Transport*) es un protocolo de comunicación máquina a máquina que es muy utilizado para la comunicación entre dispositivos IoT. Está basado en TCP/IP como base y consiste en mensajes de tipo cola. A diferencia de HTTP/1.0, la conexión se mantiene abierta y se reutiliza en cada comunicación con los dispositivos.

Este protocolo funciona como un servicio de mensajería con patrón publicador/suscriptor y conectándose a través de un servidor denominado *broker.* Los mensajes enviados se organizan en tópicos a los cuales se pueden suscribir otros clientes y el bróker le hará llegar los mensajes suscritos. En la Figura 9 se puede observar un esquema de cómo es una comunicación estándar en este protocolo.

Diagrama

Descripción generada automáticamente

Figura 9. Protocolo MQTT

MQTT dispone de un mecanismo de calidad de servicio (QoS) que es empleado para gestionar la robustez de sus mensajes y el control ante los fallos (Llamas L, 2019). Hay 3 niveles posibles:

* QoS 1 unacknowledged (at most one). El mensaje se envía una única vez y en caso de fallo puede que no se entregue.
* QoS 2 acknowledged (at least one). El mensaje se envía hasta que se garantiza la entrega del mensaje. El suscriptor puede recibir mensajes duplicados.
* QoS 3 assured (exactly one). Se garantiza la entrega del mensaje al suscriptor una única vez.

Este protocolo se caracteriza por su ligereza y sencillez, además de las ventajas del sistema publicación/suscripción como son escalabilidad, sincronismo y desacoplamiento entre clientes. Lo hace adecuado para el sistema que se utiliza en este proyecto ya que se van a emplear en dispositivos con baja potencia (sensores, cámaras, drones…) por lo que implica un bajo consumo de energía. También requiere un ancho de banda muy bajo para transmitir los datos, por lo que se adecúa a los requisitos de transmitir los datos en entornos marinos y con posibles problemas de calidad de red.

## Conclusiones sobre el estado del arte

Tras el análisis del estado actual en el que se trata en este proyecto, se puede concluir que se deben efectuar acciones para frenar la emergencia climática que está sufriendo nuestro planeta. Actualmente existe la tecnología adecuada para hacerlo y, en el campo de la contaminación marítima, ya se han llevado a cabo varios proyectos con resultados exitosos.

La base de estos proyectos es el uso de herramientas de análisis de datos y de sistemas robotizados. Por eso en este proyecto se va a implementar un algoritmo de *machine learning* que ayude a detectar qué áreas están más afectadas en cuanto a contaminación marina se refiere. Ya no sólo para este proyecto, si no para todo tipo de contaminación o catástrofes que ya se han producido, evitar que se repitan o minimizar sus consecuencias. Este algoritmo se va a encargar de clasificar las imágenes y dependiendo del tipo de contaminación que detecte se realizarán diferentes acciones. Algunas podrían consistir en utilizar drones acuáticos de recogida de plásticos (Rule, S., 2018) o limitar la navegación de buques marinos en las zonas más afectadas.

# Descripción general de la contribución del TFM

Una vez expuestas las motivaciones y contexto del proyecto, en este capítulo se detallan los objetivos y la metodología empleada.

## Objetivos

El propósito principal de este proyecto es estudiar y mejorar la detección de plásticos y microplásticos en el océano haciendo posible el restablecimiento de la biosfera marina que ha sido dañado por la acción humana. Para cumplir este propósito se han fijado tanto objetivos generales como objetivos específicos:

Objetivo general

Se pretende desarrollar un algoritmo suficientemente autónomo para que procese una colección de imágenes y/o videos capturados mediante drones. Además, debe ser capaz de detectar el nivel de plásticos y contaminación en las zonas en las que se ha capturado esas imágenes.

Objetivos específicos

Como objetivos más específicos se contempla ampliar la información de las bases de datos para realizar el análisis con:

* Imágenes tomadas con sensores infrarrojos.
* Datos obtenidos con sensores de calidad de aire a bordo de los drones.
* Muestras marinas cuyo análisis sirva para detectar el nivel de contaminación marina, tanto de microplásticos como de todo tipo de vertidos.

## Metodología del trabajo

La metodología en la que se ha basado el desarrollo de este proyecto es Scrum (Atlasian, 2022), que es una metodología *agile* en la que se establecen periodos breves de tiempo (o *springs*) para completar una cantidad de trabajo establecida. Se ha establecido un plan de trabajo dividido en *sprints* para afrontar los objetivos previstos:

1. Planteamiento de la propuesta.
2. Estudio del contexto y algoritmos proyectos relacionados.
3. Aprendizaje y uso de la herramienta MARLIT.
4. Elaboración de algoritmos para analizar muestras de agua y datos de los sensores infrarrojos y de calidad del aire.
5. Entrenamiento de los algoritmos y obtención de los resultados.

Imagen que contiene Gráfico

Descripción generada automáticamente

Figura 10. Diagrama de Gantt de los *sprints* planificados

Además, se ha creado un repositorio de *GitHub[[2]](#footnote-3)* en el que se lleva un control de versiones y se almacena el código programado.

## Descripción general de las partes o componentes de la propuesta

Como se ha descrito en apartados anteriores, este proyecto consta de distintas partes y componentes. A continuación, se describen todas ellas de forma breve:

### Alcance y limitaciones

Este proyecto intenta detectar la contaminación de los océanos y ser conscientes de su mal estado. La solución a este problema está en las grandes industrias y en un cambio de mentalidad del ser humano para eliminar hábitos que están perpetuados desde hace mucho, como pueda ser disminuir el nivel de consumo animal para reducir la contaminación que produce la industria pesquera.

### Tecnologías implicadas

Las tecnologías utilizadas y relacionadas con la Industria 4.0 son las siguientes:

* **Robótica**. Empleado para Todo lo relacionado con los drones empleados utiliza este habilitador técnico.
* **Sensores**. Empleados para recoger información del entorno y poder detectar la contaminación del entorno en el que vuelan los drones.
* **Big data**.Encargado de procesar y analizar los datos obtenidos. También se emplea *machine learning* para llevar a cabo predicciones.
* **MQTT.** Protocolo de red muy ligero usado para comunicación de máquina a máquina (M2M), muy empleado en IoT para enlazar dispositivos y mandar información sin interacción humana.

### Arquitectura, componentes e integración de tecnologías.

El esquema de cómo se relacionan las tecnologías y habilitadores técnicos en este proyecto queda representada en un breve esquema en la Figura 11. Como componentes se cuenta con distintos drones, cada uno de ellos con distintos sensores para obtener los datos, con un entorno *cloud* para procesar los datos obtenidos y guardarlos para su visualización y envío a autoridades pertinentes para su posterior toma de decisiones.

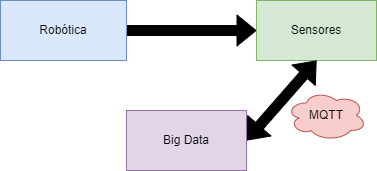


Figura 11. Arquitectura de tecnologías

**Robótica**

Se van a emplear drones para captura de imágenes y datos que se van a analizar posteriormente. Estos drones, tanto aéreos (DJI, 2022) como submarinos (Seasam, 2021) van a ser programados para recorrer un área recogiendo información y/o imágenes que va a ser enviada para su análisis y procesado

Estos drones van a estar programados para que recorran un área de manera autónoma y tomen datos cada minuto. Los drones usados son los siguientes:

* **DJI Matrice 600 Pro** (Figura 12). Este drone se caracteriza por su gran autonomía (38 minutos), por lo que va a permitir capturar más imágenes y obtener más datos en cada vuelo que realice.
  + Su peso es de 9.5 kg y sus dimensiones son 1668 mm x 1518 mm x 727 mm con las hélices, brazos y el soporte del GPS desplegados.
  + Su velocidad máxima es 65 km/h sin viento y tiene una resistencia al viento de 8m/s.
* **Seasam Autonomous ROV** (Figura 13). Este drone puede navegar de manera autónoma y tiene la capacidad de monitorear el medio ambiente.
  + Puede navegar hasta una profundidad de 100 metros.
  + Tiene una batería que le permite navegar durante una hora y media.
  + Sus dimensiones son 55cm x 45cm x 23cm con un peso de 9 kg.
  + La cámara permite capturar imágenes y video en resolución Full HD 1080p.
* **WasteShark** (Figura 14). Es un drone acuático capaz de recoger basura y residuos plásticos en agua. Las características son las siguientes:
  + Tiene una autonomía de 6 horas.
  + Tiene unas dimensiones de 155 cm x 107 cm x 45 cm y pesa 72 kg.
  + Puede recoger 500 kg de escombros al día.
  + Es capaz de escanear y recopilar datos de la calidad del agua.



Figura 12. Drone aéreo para captura de imágenes aéreas y captura de datos

Fuente: DJI, 2022

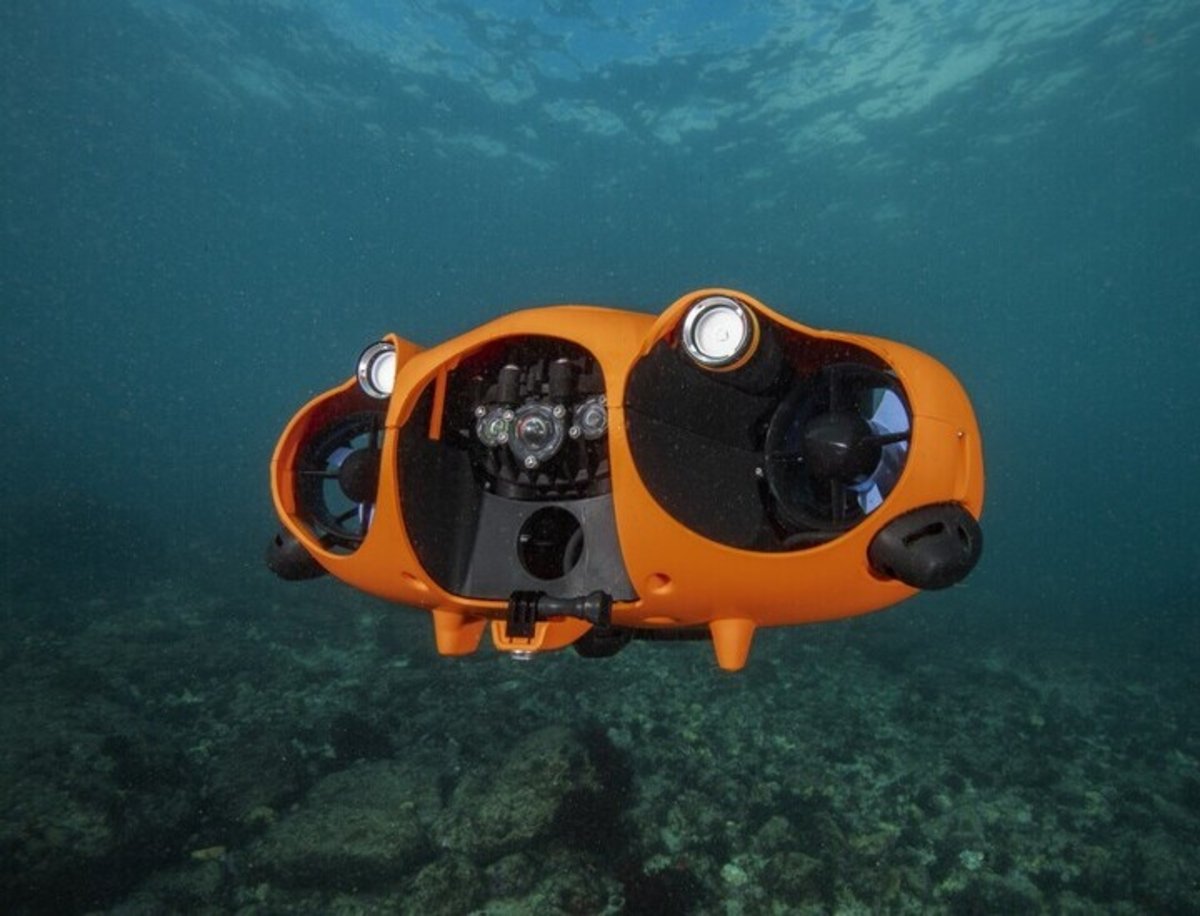


Figura 13. Drone submarino para captura de imágenes submarinas

Fuente: Seasam, 2021

Imagen que contiene agua, barco, viejo, vuelo

Descripción generada automáticamente

Figura 14. Drone de recogida de residuos

Fuente: RanMarine Technology, 2022

**Sensores**

Los sensores empleados van a bordo de los drones para recoger la información necesaria para detectar la contaminación marina. Para ello se utilizan diferentes tipos de sensores:

* **Cámaras infrarrojos**. Permiten medir la radiación que emiten los cuerpos. Esto da la posibilidad de capturar la contaminación que emiten los barcos ya que, de otra manera, no se podrían detectar y valorar cómo afectan a la calidad de aire del entorno. Este tipo de cámara captura la imagen como se ven en la Figura 14 y permite un rápido reconocimiento de la polución generada. La cámara usada a bordo de los drones es Flir duo Pro R[[3]](#footnote-4) que tiene las siguientes características:
  + Resolución 4K
  + Sensores integrados como GPS, barómetro, magnetómetro y unidad de medición inercial
  + Campo de visión de la cámara de 56º x 45º
  + Dimensiones 87 mm x 82mm x 69 mm

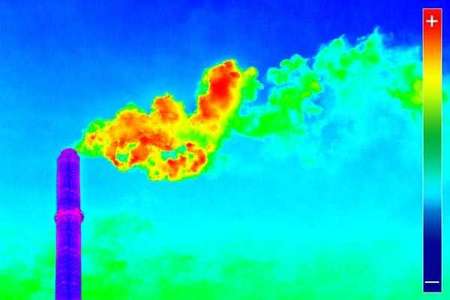


Figura 15. Imagen de polución capturada por cámara infrarrojos

* **Cámaras**. Empleadas para recoger imágenes para su posterior procesamiento y análisis para buscar plásticos en el océano. Sus imágenes capturadas son como las de cualquier cámara, que tengan definición suficiente para distinguir el plástico en ellas y que permita su posterior análisis, como la imagen de la Figura 15.

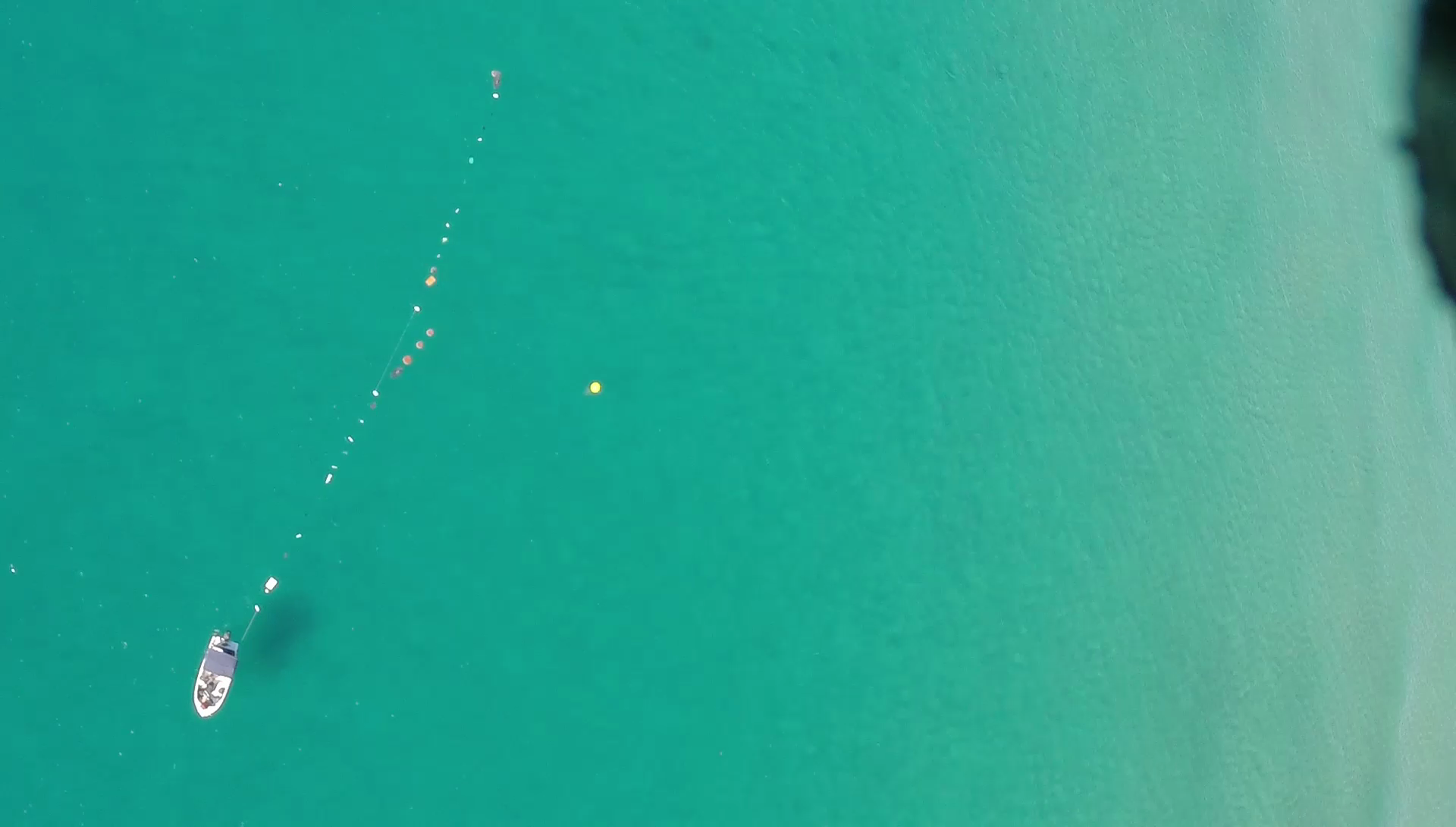


Figura 16. Imagen capturada por una cámara a bordo de un drone

* **De temperatura**. Sensores destinados a recoger la temperatura mediante una señal eléctrica. Se emplean para registrar la variación de temperaturas a lo largo del tiempo y como estas aumentan según la contaminación. Estos sensores registran valores numéricos y después de analizarlos pueden permitir representarlos de forma gráfica en *Grafana*, que se explicará más adelante, para facilitar la lectura e interpretación de imágenes, siendo una gráfica similar a la Figura 16. Los sensores empleados para este proyecto son los Omega OM-90 Series[[4]](#footnote-5), que tienen las siguientes características:
  + Temperatura con rango de medición -30-80ºC con una resolución de 0.01ºC
  + Humedad relativa con rango de precisión de 0-100% y una resolución de 0.01ºC
  + La temperatura y humedad relativa se registran a un intervalo configurable por el usuario
  + Puede almacenar 65520 mediciones de temperatura y humedad.



Figura 17. Representación en Grafana de temperatura y humedad

* **De calidad del aire**. Dispositivos que detectan y miden químicos y contaminantes en el aire. Permiten monitorizar la contaminación aérea y detectar en qué momentos hay más y cómo influyen los buques que navegan por el mar. Su representación es muy similar a la de los sensores de temperatura, ya que los datos son valores numéricos y se van a visualizar en *Grafana*. Se emplea el sensor Gravity PM2.5 Air Quality Sensor[[5]](#footnote-6) que tiene las siguientes características:
  + Rango de funcionamiento de temperatura: -10 – 60ºC
  + Rango efectivo de concentración de masa de partículas: 0~500ug/m³
  + Dimensiones: 67 mm \* 40 mm \* 14 mm

**MQTT**

En este proyecto se emplea el protocolo MQTT para mandar los mensajes de los sensores a los servicios de Amazon que se van a procesar y analizar.

Se va a utilizar el nivel QoS 2 para garantizar la entrega de los mensajes con la información recogida, no teniendo importancia si se reciben duplicados porque se van a procesar posteriormente y la información duplicada no altera en el resultado de la predicción.

**Big data y machine learning**

Para emplear Big Data se va a utilizar el servicio *Kinesis* del entorno *cloud* de *Amazon Web Services* (AWS | Cloud Computing, 2022). Este servicio permite recopilar, procesar y analizar los datos en tiempo real como videos, audios y datos de telemetría de IoT para aprendizaje automático y análisis. Este servicio de AWS se va a comunicar con otros para completar la arquitectura tal y como se muestra en la Figura 17.

****

Figura 18. Arquitectura del sistema planteado

Los servicios que completan esta arquitectura son:

* **IoT Core**. Permite conectar dispositivos a otros dispositivos y a servicios de Amazon Web Services, proteger datos y comunicaciones, procesar y actuar sobre los datos tomados por dispositivos e interactuar entre ellos, aunque carezcan de conexión. Esta arquitectura se emplea para conectar los sensores de los drones con AWS mediante protocolo MQTT y después analizar los datos una vez procesados.
* **Amazon SQS**. Es un servicio de colas de mensajes. Se utiliza para encolar mensajes que no se puedan procesar en el momento y no se pierdan por tener un gran volumen de datos.
* **Grafana**. Empleado para representar los datos de forma gráfica. Gracias a este servicio se puede tener distintas gráficas que representen distintos datos, así como sus predicciones en múltiples paneles de control (Grafana Labs, 2022).

### Resultados esperados

El resultado esperado para este proyecto es mejorar los algoritmos existentes y tener una buena detección de basura en los océanos, así como realizar una predicción de en qué lugares a nivel mundial hay más contaminación analizando solo una pequeña parte del océano.

Para ello hay que tener una buena base de datos con imágenes y datos suficientes para poder analizarlos y tener muestras que permitan obtener un alto porcentaje de acierto.

### Presupuesto

|  |  |  |  |
| --- | --- | --- | --- |
| Producto | Precio/unidad | Cantidad | Precio total |
| Drone | 14000€ | 4 | 56000€ |
| drone acuático | 12000€ | 2 | 24000€ |
| cámara infrarrojos | 4338€ | 2 | 8676€ |
| sensor temperatura/humedad | 75€ | 2 | 150€ |
| sensor calidad aire | 45 € | 2 | 90€ |
| Total |  |  | **168512€** |

Tabla 2. Presupuesto de hardware

Siendo los modelos de este presupuesto los siguientes:

* **Drone**: DJI Matrice 600 Pro
* **Drone acuatico**: Seasam Autonomous ROV
* **Cámara infrarrojos**: Flir duo Pro R
* **Sensor temperatura/humedad**: Omega OM-90 Series
* **Sensor de calidad de aire**: Gravity PM2.5 Air Quality Sensor

Sumando la cantidad necesaria de productos para llevar a cabo el proyecto, la suma total de dinero del hardware necesario es 168440€.

|  |  |  |  |
| --- | --- | --- | --- |
| Producto | Precio/solicitud | cantidad | Precio total |
| aws Iot | 0,0035€ | 250000 | 875€ |
| aws sqs | 0,38€ | 10000 | 3800€ |
| aws kinesis | 0,034€ | 256000 | 8704€ |
| grafana | 9€ | 10 | 90€ |
| total |  |  | **13469€** |

Tabla 3. Presupuesto de software

El precio de AWS IoT y SQS se calcula por el número de peticiones (las del presupuesto son una estimación).

El precio de AWS Kinesis es en relación con la cantidad de ingesta de datos en GB. Como van a ser imágenes de una calidad media se asume que va a ingestar una gran cantidad de datos.

Grafana calcula el precio en relación con las representaciones gráficas que se vayan a emplear. Se estima que van a hacer falta un mínimo de 10 para representar los datos de temperatura, humedad, plásticos, etc.

|  |  |
| --- | --- |
|  | PRECIO |
| Software | 13469€ |
| HARDWARE | 168512€ |
| TOTAL | **181981€** |

Tabla 4. Presupuesto total

### **Planificación** general

Como se ha especificado en anteriores apartados, la metodología llevada a cabo para este proyecto es la metodología *Scrum*. Se divide en cinco *sprints* divididos en planteamiento de la propuesta, estudio del contexto, estudio de MARLIT, elaboración de algoritmo y entrenamiento (Figura 10):

* **Planteamiento de la propuesta**. En esta etapa se elabora una propuesta, incluyendo un plan de trabajo y que frentes va a abarcar todo el proyecto.
* **Estudio del contexto.** Etapa en la que se estudia el estado actual del mar, planteando diferentes catástrofes que han perjudicado seriamente el medio ambiente. También se ha investigado diferentes propuestas existentes que abarcan el mismo problema del que trata este proyecto.
* **Estudio de MARLIT.** Debido a que es la herramienta que más se asemeja a lo que trata el proyecto y además, es de código abierto, se estudia en profundidad esta herramienta descargando el código y ejecutándolo para valorar si es una buena solución del problema.
* **Elaboración de algoritmo.** Programación de un algoritmo que detecte plásticos, así como la elección del lenguaje de programación usado y su correcto desarrollo. La solución se irá actualizando en el repositorio de GitHub expuesto anteriormente.
* **Entrenamiento de algoritmo y corrección de errores.** Entrenamiento del algoritmo elaborado en la anterior fase y corrección de los errores que se puedan encontrar, así como la elaboración de la memoria y últimos ajustes para su presentación.

# Desarrollo específico de la contribución

En este capítulo se desarrolla la contribución aportada al proyecto expuesto

## MARLIT

MARLIT o Medios de evaluación y Atenuación de los Riesgos costeros Locales debidos a Impactos de los Temporales es una aplicación programada en R que se usa para detección de basura marina en imágenes aéreas. Es una aplicación de acceso abierto y permite detectar y cuantificar los plásticos que flotan en el mar a través de análisis de imágenes obtenidas en las costas mediterráneas. Estos análisis emplean técnicas de inteligencia artificial y Deep Learning, lo cual permite evaluar los plásticos y contaminantes, así como su distribución y densidad. El algoritmo desarrollado es una prueba de concepto basada en R Shiny, que es una biblioteca del lenguaje de programación R que permite alojar aplicaciones o levantar páginas web de forma ágil.

Después de obtener estos datos se podría actuar sobre las zonas más afectadas con acciones de limpieza o limitando la navegación en esas zonas para evitar que se genere más contaminación. Las posibles acciones se desarrollarán más adelante.

Para comprender y valorar la aplicación se ha usado el proyecto ubicado en GitHub[[6]](#footnote-7) y se ha ejecutado cada uno de los algoritmos programados para la detección de basura.

### Organización del proyecto

El proyecto está organizado como muestra la Figura 16 y los ficheros están divididos de la siguiente manera:

* Una carpeta la cual contiene modelos entrenados y logos de la empresa y universidad. Además, contiene el código principal para la detección de basura y un fichero RData para almacenar información estadística.
* Ficheros comprimidos en tar.gz que contienen funciones programadas en R para la detección de la basura.
* Algoritmo en R para instalar todos los paquetes necesarios para la correcta ejecución del código.
* Información general del proyecto en formato *markdown*.

Texto

Descripción generada automáticamente

Figura 19. Organización de proyecto MARLIT

### Ejecución de algoritmo

Para ejecutar los algoritmos de este proyecto, primero es necesario ejecutar el código del script *instal\_paquets.R* para instalar las dependencias.

Una vez hecho, ya se puede ejecutar el *script* *app.R*, que es el que realiza la predicción de las imágenes del proyecto y se ubican en la carpeta test IMAGES. En la ejecución se emplean las imágenes que se encuentran en el repositorio del proyecto.

La ejecución del código y el entrenamiento nos devuelve una lista con la precisión de los algoritmos. Los resultados se detallan en el siguiente apartado.

### Conclusiones del proyecto

La ejecución del código principal del proyecto ha llevado varios problemas:

* No se encontraban ciertos paquetes a la ejecución (Figura 19)

Texto

Descripción generada automáticamente

Figura 20. Error en código de proyecto MARLIT

* Los comentarios que explicaban partes del código están en varios idiomas, así como el nombre de los scripts.
* Variables declaradas en distintos idiomas, lo que resta consistencia al código.

Después de solucionar el problema de la instalación, se ejecuta el script de *instal\_paquets.R* para que descargue los necesarios para la ejecución y se han cargado las funciones y modelos (Figura 20), se pueden obtener varias conclusiones del código del proyecto:

* Se pueden usar imágenes .png y .jpg.
* La imagen original es transformada a una matriz de píxeles de 25 x 25
* La precisión del modelo de Deep Learning que se usa por defecto es T.
* Guarda las características del tren de imagen en los ficheros .rds
* El número de rondas para calcular el modelo de Deep Learning es por defecto 150.

Texto

Descripción generada automáticamente con confianza baja

Figura 20. Modelos y funciones cargadas en RStudio

Tras arreglar varios fallos sintácticos del código y ejecutar los distintos algoritmos, obtenemos una lista con las predicciones (Figura 21). En esta ejecución planteada por los desarrolladores del proyecto, el entrenamiento tiene un 85% de precisión, pero después de llevar a cabo la fase de test esta precisión baja a un 75%. Se podría mejorar usando píxeles u otro modelo de color como RGB, pero esto aumentaría el procesamiento y por lo tanto se necesitarían más recursos para los análisis de las imágenes.

Texto

Descripción generada automáticamente

Figura 21. Entrenamiento con un dispositivo

Como conclusión, el código tiene muchos bugs y problemas con su ejecución, además de todos los inconvenientes anteriormente expuestos. Dada su baja precisión (75%), no es considerado un buen proyecto ni algoritmo para llevar a cabo una detección de plástico en el océano.

## Métodos de prevención

En este proyecto se desarrollan métodos de prevención de contaminación marina según los datos capturados por los drones.

Los datos que se emplean para esta prevención son los obtenidos por los sensores de calidad de aire y los datos obtenidos por las cámaras. A partir de los datos obtenidos se pueden llevar a cabo acciones para prevenir el empeoramiento del ecosistema afectado e incluso mejorar las zonas más afectadas.

Para estos métodos de prevención se pueden contemplar varios escenarios según el nivel de contaminación que nos devuelva el análisis de los datos obtenidos de las imágenes.

### Actuaciones en caso de contaminación baja

En caso de baja contaminación del área analizada habría que llevar a cabo alguna acción como limpiar de manera ocasional los plásticos y limitar parcialmente la navegación de barcos en la zona.

Debido a que es un nivel bajo, no hay que llevar a cabo excesivas acciones. Con realizar limpiezas ocasionales para evitar que la contaminación incremente es suficiente en este nivel.

### Actuaciones en caso de contaminación media

En caso de contaminación media hay que efectuar tareas de limpieza de la zona y reducir el plástico encontrado, hay que limitar la navegación de los barcos y prohibir la pesca en un radio de 15km del área afectada.

Estás limitaciones hacen que no vaya a más la contaminación porque los barcos generan muchos residuos y la pesca es lo que más perjudica el medio marino, con las redes de pesca siendo el plástico que más se encuentra y la pesca de arrastre lo que afecta al ecosistema y lo que más daña a la fauna, no sólo por la pesca en sí, sino porque remueve el sedimento marino y genera toneladas de dióxido de carbono.

### Actuaciones en caso de contaminación alta

En caso de contaminación alta hay que actuar inmediatamente en la zona afectada. En caso de llegar a esa situación hay que limpiar de manera inmediata, ya sea mediante drones o mediante barcos/submarinos y vigilar la fauna de la zona para su correcto desarrollo y que no se vea afectada.

Además, se prohíbe de manera inmediata la circulación de barcos por esa zona y se prohíbe la pesca en un radio considerable alrededor de la zona (alrededor de unos 100 km), pues las redes de pesca y la pesca de arrastre es lo que más perjudica el entorno marino.

## Algoritmo de detección de plásticos

Para este proyecto se ha desarrollado un algoritmo básico para la detección de plásticos en un entorno marino. Este algoritmo va a ser capaz de distinguir si hay algún objeto que contenga plástico en una serie de imágenes, por lo que va a servir para detectar plástico en las imágenes capturadas por los drones.

Este algoritmo va a costar de dos partes: uno para obtener las imágenes haciendo web scraping de una búsqueda de Google y otra para detección de plástico y contaminación marina.

### Descripción general del algoritmo

El algoritmo está programado en Python, mediante Google Colab[[7]](#footnote-8), que permite ejecutar el código online y que sea accesible desde cualquier dispositivo con acceso a internet. También se subirá a un repositorio personal de GitHub[[8]](#footnote-9) para poder ejecutarlo también con Jupyter Notebook.

*Computer Vision* es una disciplina de Inteligencia Artificial y esto permite resolver situaciones usando redes neuronales, Deep Learning. Esto permite hacer 3 tipos de clasificaciones (Figura 23):

* Clasificación. Clasifica las imágenes según si en ellas existen el objeto que se desea localizar.
* Detección de objeto. Clasifica y localiza en qué lugar de la imagen se encuentra el objeto deseado.
* Segmentación. Clasifica y localiza el lugar de la imagen y el área exacta que ocupa cada uno de los objetos buscados.

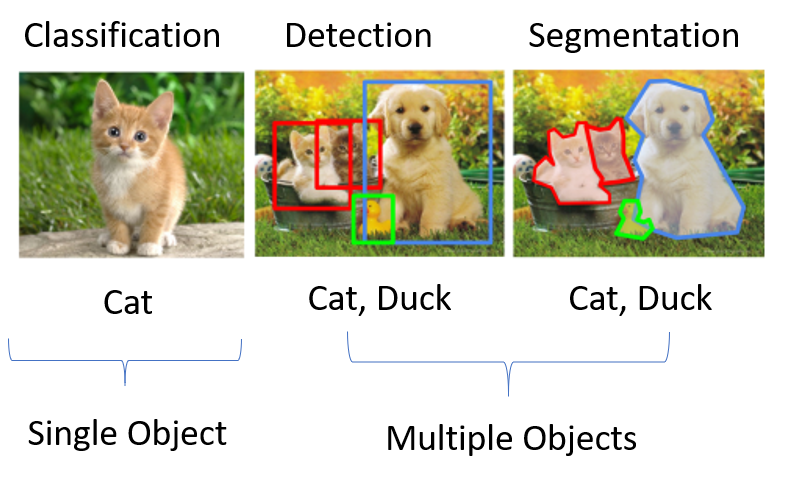


Figura 22. Diferentes clasificaciones de Deep Learning

Fuente: Kolungade, S. (2021, 15 diciembre)

Lo más importante para el correcto desarrollo del algoritmo y un buen nivel de predicción es contar con un buen *dataset* y unas imágenes que faciliten el entrenamiento. El entrenamiento supervisado se entrena con los mismos datos que se esperan como salida.

### Algoritmo para obtener imágenes

Como para este proyecto no se han encontrado grandes *datasets* relacionados con las imágenes deseadas, se va a realizar un *scraping*[[9]](#footnote-10) de Google usando su motor de búsqueda para descargarlas y emplearlas para entrenar nuestro algoritmo. Antes de extraer la información de forma automática, se debe realizar una búsqueda de forma manual para comprobar que nos devuelve imágenes relacionadas con el algoritmo que hemos programado. En este caso “plastic polution” devuelve imágenes muy relacionadas con lo deseado (Figura 23).

Esto permite tener un *dataset* de forma rápida sin tener que hacer una búsqueda exhaustiva y aprovechando el motor de búsqueda de Google, lo que permite entrenar con imágenes de distinto ámbito sólo cambiando las palabras clave en la búsqueda que realiza el código de *scraping*.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Figura 23. Búsqueda manual de imágenes en Google

Ejecutar el algoritmo propuesto descargará las librerías y paquetes necesarios para su correcta ejecución, en caso contrario actualizará los paquetes o no hará nada por tenerlos previamente descargados (Figura 24).

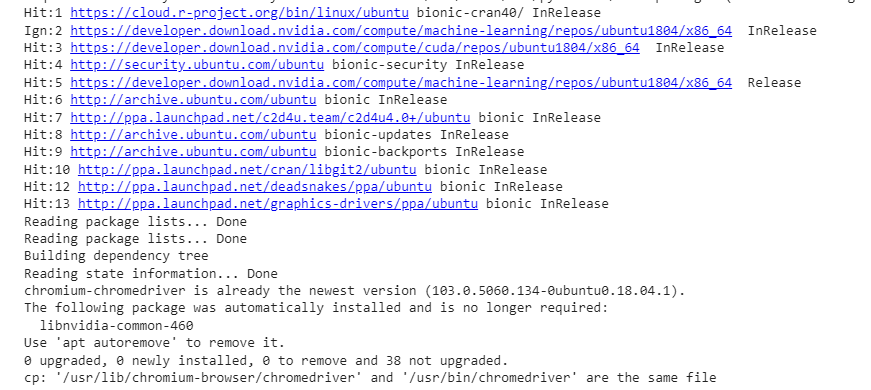


Figura 24. Descarga de dependencias al ejecutar el código Python

A continuación, se han programado una serie de funciones para mejorar la calidad del código. Estas funciones se encargan de diversas tareas como:

* Realizar la búsqueda deseada en Google
* Guardar las imágenes realizando *scraping* de la web
* Convertir la imagen a composición RGB
* Codificar la imagen a formato JPG

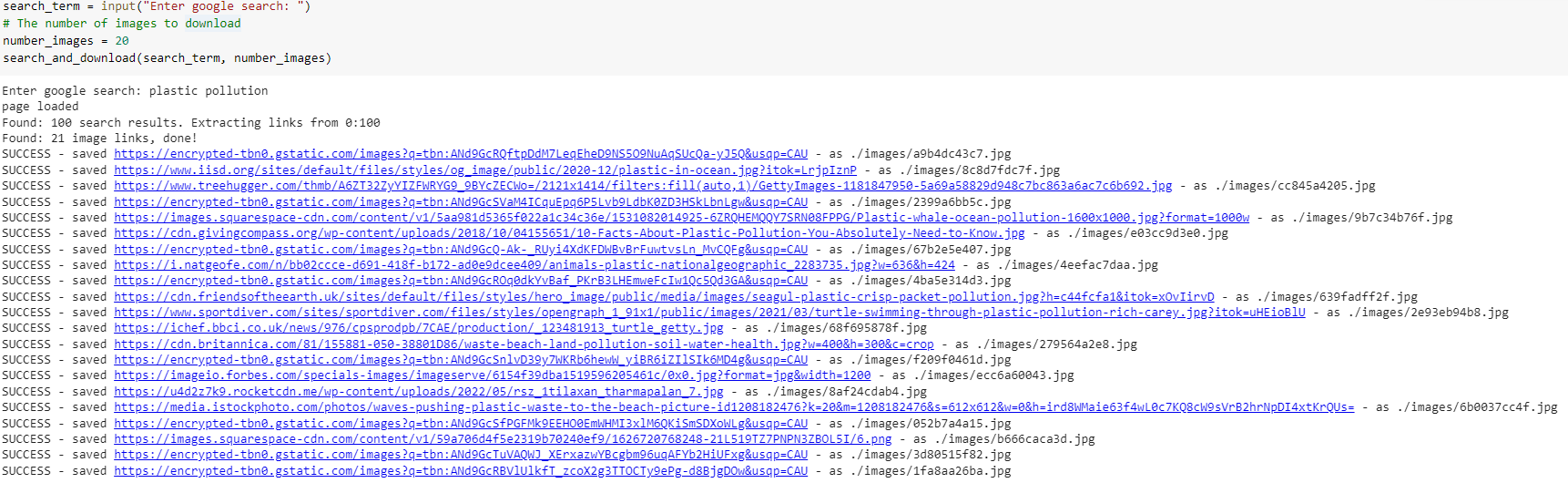
Después de ejecutar la celda en la que se encuentran las funciones para poder guardarlas y que el código pueda usarlas a posteriori, se ejecuta la celda en la que se introduce que palabras clave queremos utilizar para realizar la búsqueda y descargarlas (Figura 25). También se puede elegir en el código cuántas imágenes se tiene que descargar.

Figura 25. Descarga de imágenes realizando una búsqueda en Google

Una vez descargadas las imágenes, se eliminan duplicadas o similares (Figura 26).

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Figura 26. Eliminación de imágenes duplicadas

Como último paso de este algoritmo, seleccionamos foto a foto si contiene plástico u otro material. Si tiene plástico la imagen se guarda, si no tiene plástico o no nos interesa para usar en el algoritmo de detección, se selecciona la opción “Other” y la imagen se descarta y a continuación se selecciona foto por foto si la imagen descargada contiene plástico o no. En la Figura 27 hay un ejemplo de imagen que se descarta porque no interesa para la detección (hay foto en la imagen, pero no se encuentra en agua) y en la Figura 28 una foto que si es válida para el algoritmo de detección.

Un grupo de imágenes de un elefante

Descripción generada automáticamente

Figura 27. Fotografía descartada para el algoritmo de detección

Tortuga nadando en el agua

Descripción generada automáticamente con confianza media

Figura 28. Fotografía empleada para la detección de plásticos

### Algoritmo de detección de residuos

El algoritmo de detección en primera instancia va a cargar las imágenes descargadas en el algoritmo de *web scraping.*

A continuación, se dividen en subconjuntos de entrenamiento, validación y test. En este caso se usa una proporción 70-15-15: 70% para datos de entrenamiento y 15% para validación y test. En caso de querer realizar más pruebas estas proporciones pueden variar a 60-20-20 o 50-25-25, el mejor resultado se obtiene realizando pruebas variando estas proporciones.

Para la solución propuesta se utiliza la librería de Python *imageAI*[[10]](#footnote-11), que es de código abierto y permite crear aplicaciones y sistemas con capacidades autónomas de Deep Learning y Computer Vision usando pocas líneas de código y facilitando la tarea a los desarrolladores. Esta biblioteca contiene el modelo *YoloV3K*, que es uno de los modelos más populares y completos para la detección de imágenes. Para este algoritmo se descarga uno previamente entrenado porque requiere menos datos y menos iteraciones, por lo que su ejecución va a ser mucho más rápida y eficaz.

Al ejecutar el algoritmo, en primera instancia, se instalan los paquetes necesarios y se realizan los *imports* para el correcto funcionamiento del código.

Después se carga el *dataset* obtenido con el algoritmo de *scraping* y se configura la proporción que se ha descrito anteriormente.

### Conclusiones y resultados

# Conclusiones y trabajos futuros

## Conclusiones

## Líneas de trabajo futuras

Referencias bibliográficas

A. (2014, 29 mayo). Los doce peores desastres ecológicos del mundo | EROSKI Consumer. Consumer |. <https://www.consumer.es/medio-ambiente/los-doce-peores-desastres-ecologicos-del-mundo.html>

Atlassian. (2022). Scrum: qué es, cómo funciona y por qué es excelente. <https://www.atlassian.com/es/agile/scrum>

AWS | Cloud Computing (2022). Amazon Web Services, Inc. <https://aws.amazon.com/es/>

Boy, G. (2022, 31 mayo). Las diferencias que hay entre Data Science, Inteligencia Artificial, Machine Learning y Deep Learning. PROGRAMMATIC SPAIN. https://www.programaticaly.com/education/las-diferencias-que-hay-entre-data-science-artificial-intelligence-machine-learning-y-deep-learning

Big Data: ¿En qué consiste? Su importancia, desafíos y gobernabilidad. (2021). PowerData. https://www.powerdata.es/big-data

Breaking the Plastic Wave: Top Findings for Preventing Plastic Pollution #BreakingThePlasticWave. (2020, 23 julio). Pewtrust. <https://www.pewtrusts.org/-/media/assets/2020/10/breakingtheplasticwave_mainreport.pdf>

El continente de plástico que flota en las aguas del Pacífico. (2019, 25 noviembre). Iberdrola. <https://www.iberdrola.com/sostenibilidad/isla-de-plastico-pacifico-septimo-continente>

Conversation. (2021, 19 marzo). National Geographic. www.nationalgeographic.com.es. <https://www.nationalgeographic.com.es/mundo-ng/nuevas-tecnicas-deteccion-y-contabilizacion-basura-que-flota-mar_16717>

Euroinnova Business School. (2022, 28 abril). Certificado de control de plagas. <https://www.euroinnova.edu.es/blog/que-es-la-contaminacion-marina#iquestqueacute-es-la-contaminacioacuten-marina>

Descubre los principales beneficios del Machine Learning. (2022). Iberdrola. <https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico#:%7E:text=El%20Machine%20Learning%20es%20una,elaborar%20predicciones%20(an%C3%A1lisis%20predictivo)>.

DJI. (2022). DJI Matrice 600 Pro. <https://www.dji.com/matrice600-pro>

Garcia-Garin, O., Monleón-Getino, T., López-Brosa, P., Borrell, A., Aguilar, A., Borja-Robalino, R., Cardona, L., & Vighi, M. (2021). Automatic detection and quantification of floating marine macro-litter in aerial images: Introducing a novel deep learning approach connected to a web application in R. Environmental Pollution, 273, 116490. <https://doi.org/10.1016/j.envpol.2021.116490>

Grafana Labs. (2022). Grafana: The open observability platform. <https://grafana.com>

Kolungade, S. (2021, 15 diciembre). Object Detection, Image Classification and Semantic Segmentation using AWS Sagemaker. Medium. <https://medium.com/@kolungade.s/object-detection-image-classification-and-semantic-segmentation-using-aws-sagemaker-e1f768c8f57d>

Lebreton, L. (2018, 22 marzo). Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic. Nature. <https://www.nature.com/articles/s41598-018-22939-w>

Llamas, L. (2019, 17 abril). ¿Qué es MQTT? Su importancia como protocolo IoT. https://www.luisllamas.es/que-es-mqtt-su-importancia-como-protocolo-iot/

MARLIT - Coastal risk management. (2021, 29 octubre). AZTI. <https://www.azti.es/en/proyectos/marlit/>

Méndez, I. (2020, 7 junio). La detección por satélite de vertidos creada por un malagueño llega a la Agencia Espacial Europea. Diario Sur. <https://www.diariosur.es/sociedad/deteccion-satelite-vertidos-20200607213229-nt.html>

Orca Robotics. (2022). Sourceforge. <http://orca-robotics.sourceforge.net>

The Orocos Project – Smarter control in robotics & automation! (2022). Orocos. https://orocos.org

Oracle España. (2022). Oracle. https://www.oracle.com/es/big-data/what-is-big-data/ROS: Home. (2022). ROS. <https://www.ros.org>

RanMarine Technology. (2022, 30 marzo). WasteShark technical details. https://www.ranmarine.io/products/wasteshark/

Rule, S. (2018, 28 noviembre). UN DRON ACUÁTICO QUE RECOGE BASURA Y DATOS. SURFER RULE • Más que surf, olas gigantes y tendencias. <https://www.surferrule.com/dron-acuatico-recoge-basura-datos/>

Statista. (2021, 26 octubre). Producción mundial de plástico 1950–2019. <https://es.statista.com/estadisticas/636183/produccion-mundial-de-plastico/>

Tecnología, A. (2017, 19 septiembre). Los drones «aprenden» a localizar vertidos en el mar. abc. <https://www.abc.es/tecnologia/informatica/hardware/abci-drones-aprenden-localizar-vertidos-mar-201709181938_noticia.html>

Seasam (2021, 23 septiembre). Automous solution to collect underwater data. <https://seasam.notiloplus.com>

1. Son redes de pesca que han sido perdidas o abandonadas por pescadores [↑](#footnote-ref-2)
2. <https://github.com/ralvarezmar/tfm> [↑](#footnote-ref-3)
3. <https://www.warmtebeeldcamera.nl/en/flir-duo-pro-r-336.html> [↑](#footnote-ref-4)
4. <https://es.omega.com/pptst/OM-90_Series.html> [↑](#footnote-ref-5)
5. <https://thepihut.com/products/gravity-pm2-5-air-quality-sensor> [↑](#footnote-ref-6)
6. https://github.com/amonleong/MARLIT [↑](#footnote-ref-7)
7. <https://colab.research.google.com/?hl=es> [↑](#footnote-ref-8)
8. <https://github.com/ralvarezmar/tfm> [↑](#footnote-ref-9)
9. Proceso para recopilar automáticamente información de determinada página web [↑](#footnote-ref-10)
10. <https://github.com/OlafenwaMoses/ImageAI> [↑](#footnote-ref-11)