

Capítulo 2

Estado del arte

2.1. Aplicaciones de reciclaje

Debido a los distintos materiales que se encuentran en los residuos del día a día de la mayoría de las personas y la cantidad de estos que se generan, ha surgido una necesidad de ayuda para reciclar correctamente todos estos residuos.

Esta ayuda ha llegado, sobre todo, en forma de aplicación móvil. La generalización del uso de los dispositivos móviles en la población favorece a que esta sea la mejor forma de encontrar esta ayuda.

Cada nación cuenta con un sistema diferente de gestión de residuos y reciclaje, debido a la cantidad de opciones que hay en este trabajo se enfocará al territorio nacional. En España, en aplicaciones de reciclaje que identifiquen objetos contamos con dos. La primera es “AIRE Asistente Inteligente de Reciclaje”¹, disponible tanto para iOS, Android y en el navegador, es una aplicación publicada por “ecoembes”, empresa encargada del reciclaje de los residuos de los contenedores amarillo y azul en España. Se trata de un asistente virtual con el que se tienen una conversación de chat, el usuario escribe el material o el objeto que desea reciclar y el *bot* le responde explicando la forma idónea de desecharlo. También se le pueden enviar imágenes y audios para que identifique el objeto. Otra es “ReciclaYa”², aplicación desarrollada por Carrefour. En este caso la identificación se lleva a cabo mediante el código de barras del ticket de la compra que se haya hecho, según qué productos se hayan adquirido indica cómo reciclar cada uno sus envases. Esta funcionalidad está sólo disponible para algunas de las marcas principales, ya que son las propias empresas las que brindan esa información. Algunas de estas empresas sobre las que se disponen los datos son Carrefour, Coca Cola, Nestlé, Central Lechera Asturiana, Mahou San Miguel, PepsiCo, El Pozo o Pescanova, entre muchas otras.

¹<https://www.ecoembes.com/proyectos-destacados/chatbot-aire/>

²<https://www.reciclaya.app/es/como-funciona>

Además de aplicaciones de identificación de residuos, también se encuentran numerosas aplicaciones enfocadas más a la ayuda y el aprendizaje sobre reciclaje, algunos ejemplos son “Residus” de la Generalitat de Catalunya³ en colaboración con “ecomenbes” y “ecovidirio”. Esta aplicación cuenta con una base de datos de residuos con su respectivo contenedor adecuado y el proceso que sigue una vez es desechado en este. Cuenta con dos maneras de informarme sobre cada desecho. La primera es por contenedores, se selecciona el contenedor sobre el que se quiere tener más información, sobre este sale un listado de todos los residuos disponibles en la base de datos reciclables en él, y finalmente se selecciona uno de ellos. La segunda manera es introducir en el buscador el residuo. Cuando se tiene uno seleccionado, en ambas opciones, se explica el proceso que sufre este desecho desde el contenedor, pasando por la plantas de selección y reciclaje, y en qué se transforma finalmente. Además de informar sobre la forma adecuada de deshacerse de cada residuo, también cuenta con un buscador de puntos limpios a partir de la localización del dispositivo o mediante un buscador.

Otro ejemplo es “Recicla y suma” de Pensumo⁴. Es una herramienta pensada para incentivar el reciclaje animando al ciudadano a que se acostumbre a separar en casa los residuos que genera para luego acudir a los contenedores y reciclarlos. La función principal es la de generar el hábito en el ciudadano, para que, animado por el incentivo económico, acabe interiorizando la necesidad de reciclar y aprenda a depositar correctamente los residuos. Los usuarios sacan fotos en el momento que depositan los residuos en el contenedor correcto y reciben dinero por la acción.

Alejándose un poco del reciclaje como tal encontramos “The Planet App”⁵ por Clean Planet Ventures, SL. Es una aplicación de concienciación, aprendizaje y ayuda sobre la huella de carbono que genera cada uno. Permite al usuario calcular su huella de carbono y conocer las emisiones que genera separadas por categorías, además de comparar sus emisiones con las de distintos grupos poblacionales y otros usuarios y configurar un plan de sostenibilidad personalizado adquiriendo así hábitos y realizando acciones que reduzcan las emisiones que genera.

En el lado internacional existen muchas más aplicaciones, como “My Little Plastic Footprint”⁶, de Plastic Soup Foundation. Esta es muy similar a la aplicación anterior, pero en vez de ser respecto a la huella de carbono es sobre el plástico que genera el usuario. En este caso, la aplicación te indica tu PMI (índice de masa plástica), cuando se tiene esto se comienza la denominada “dieta de plástico” que consiste en ir adquiriendo como hábito alternativas al plástico que se utiliza normalmente. A medida que se van incorporan-

³<https://play.google.com/store/apps/details?id=cat.gencat.mobi.residus>

⁴<https://reciclaysuma.com/>

⁵<https://theplanetapp.com/>

⁶<https://mylittleplasticfootprint.org/>

do los consejos de la aplicación al perfil del usuario, el PMI de este se irá reduciendo.

“Grow Recycling”⁷, de Gro Play Digital, es un juego educativo para niños utilizado en colegios y guarderías en Estados Unidos y Europa para impartir educación sobre la sostenibilidad. El juego consiste en alimentar a distintos cubos de basura con los residuos correspondientes para que niños y niñas puedan aprender a cuidar del planeta. Otra aplicación educativa de reciclaje es “Recycle Coach”⁸, de Municipal Media Inc., con ella se puede aprender sobre el proceso de reciclaje, los diferentes materiales y el proceso de depósito y recogida de residuos en tu zona. Además, cuenta con *tests*, artículos y actividades para aprender apropiadamente sobre la disposición de residuos.

2.2. Generación sintética de imágenes

Respecto a generación de imágenes sintéticas se encuentran distintos acercamientos. Por un lado, están las redes generativas adversativas, relacionadas con el tratamiento de imágenes. Este modelo consta de dos redes neuronales denominadas generador y discriminador, respectivamente. El objetivo es generar datos similares a los que se han usado para el entrenamiento. La red generador, como su nombre indica, es la encargada de generar datos del tipo de los del entrenamiento. Por otro lado, la red discriminadora distingue entre los datos reales que se le proporcionan y los generados por la red anterior [14].

Acercamientos más similares al que se propone en este trabajo, son sobre imágenes generadas a partir de entornos virtuales. Por ejemplo, SYNTHIA [31] es un *dataset* de imágenes fotorrealistas generadas a partir de los frames obtenidos de la rederización de un mundo virtual formado a partir de modelos tridimensionales, además de las imágenes cuenta con las anotaciones semánticas a nivel de *pixel* que etiquetan los distintos objetos presentes en la escena, distinguiendo hasta 13 clases distintas. Este *dataset* está orientado al entrenamiento de coches autónomos, sistemas de conducción asistida y proyectos relacionados, como complemento para diferentes *datasets* reales (Camvid, KITTI, U-LabelMe, CBCL).

Otro ejemplo similar, es el planteamiento de generar imágenes efectivas para el entrenamiento de detección de objetos partiendo de un conjunto pequeño de imágenes reales y el modelo tridimensional del objeto a identificar [?]. Los parámetros pueden reutilizarse para generar un número ilimitado de imágenes de entrenamiento del objeto de interés en poses del modelo tridimensional arbitrarias. Por cada imagen real se computan cinco parámetros de la pose, tres de orientación y dos de traslación que permite colocar el modelo 3D en la posición deseada. Una vez se obtiene la imagen, para hacerla lo

⁷<https://www.groplay.com/apps/grow-recycling/>

⁸<https://recyclecoach.com/>

más similar a la imagen real, pasa por un postprocesamiento que involucra el añadir desenfoque, para simular la lejanía y el movimiento del objeto; ruido, que imita el ruido que añade la cámara al tomar la fotografía; y propiedades del material del objeto. Esto permite que las imágenes generadas sintéticamente sean más similares a las reales, obteniendo un *dataset* sintético muy similar a uno real.

Un proyecto basado en el anterior, es la detección de objetos en contextos industriales [7]. En este proyecto se parte del modelo renderizado de una pieza utilizada en el entorno industrial con el que se genera un *dataset*. Como en el caso anterior, para añadir más diversidad e identificar las piezas en un mayor número de situaciones, generan las imágenes con diferentes fondos creados a partir de una selección de imágenes de localizaciones reales; alterando la iluminación y añadiendo sobras de distintas formas e intensidades; y creando imágenes desenfocadas o en movimiento.

Perception [37] es un paquete de Unity en desarrollo que proporciona herramientas para generar *datasets* sintéticos a gran escala. Está planteado para tareas de aprendizaje automático, tales como detección de objetos, segmentación semántica y más. Este *dataset* está en formato por frames, estos son capturados utilizando sensores simulados y es lo que se utiliza para el entrenamiento y la validación del modelo. Perception, además de permitir al usuario que genere sus propios datos para la tarea que requiera, ofrece una cantidad de etiquetas comunes ya generadas para facilitar la generación de datos sintéticos. Además de la generación de *datasets* sintéticos, incluye herramientas para generar *keypoints*, poses y animaciones aleatorias. Este paquete puede utilizarse para diversidad de proyectos, desde simulación de coches autónomos o *smart cities*⁹, pasando por demostraciones de recoger y colocar objetos mediante un brazo robótico¹⁰, hasta detención de objetos utilizando únicamente imágenes sintéticas para el entrenamiento.

Este último es el proyecto SynthDet [19], utiliza Perception para crear el *dataset* utilizado en el entrenamiento de la red neuronal. Generan imágenes de productos habituales de supermercados con el fin de estudiar la viabilidad de los supermercados sin cajas registradoras. Este objetivo se basa en el supermercado Amazon Go [29], creado por Amazon en 2017. Se entrará más en profundidad sobre este proyecto y sus distintas características en la sección 2.4

2.3. Redes neuronales

Las redes neuronales artificiales surgen inspiradas por la funcionalidad sofisticada del cerebro humano, donde miles de millones de neuronas se encuentran interconectadas y procesan información de manera paralela. Tratan

⁹<https://github.com/Unity-Technologies/Unity-Simulation-Smart-Camera-Outdoor>

¹⁰<https://github.com/Unity-Technologies/Robotics-Object-Pose-Estimation>

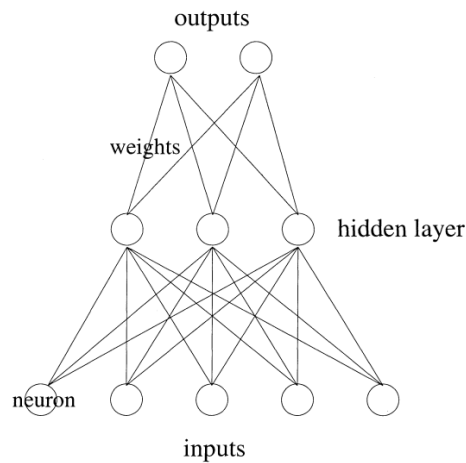


Figura 2.1: Estructura de una red neuronal artificial.

de emular el comportamiento del cerebro humano caracterizado por el aprendizaje a través de la experiencia y la extracción de conocimiento a partir de un conjunto de datos [10], además de simular las características básicas de las redes neuronales biológicas; el procesamiento paralelo, la memoria distribuida y adaptabilidad. Esto permite a las redes artificiales aprender y generalizar a partir de un conjunto de datos de relación matemática desconocida [16]. De esta forma, adquieren, almacenan y utilizan conocimientos basados en la experiencia en lugar de ser programados de manera explícita [41].

Una red neuronal artificial consta de una capa de entrada de neuronas, una o varias capas ocultas y una capa final de salida. En la figura 2.1 se muestra la estructura habitual de una red neuronal artificial.

Las neuronas están conectadas entre ellas mediante líneas, cada una de estas conexiones se asocia con un valor numérico llamado peso[38]. Durante el aprendizaje de la red se modifica el valor de los pesos asociados a las neuronas con el fin de que generen una salida a partir de unos datos de entrada. Se considera que el aprendizaje ha terminado cuando los valores de los pesos permanecen estables.

Aunque existen muchas definiciones para las redes neuronales, una de las más extendidas establece que una red neuronal artificial se define como un grafo dirigido, donde los nodos representan las neuronas y las aristas la sinapsis. Aquellos nodos que no cuentan con conexiones entrantes son consideradas las neuronas de entrada, y los que no tiene conexiones salientes las de salida. Por lo tanto, los nodos con conexiones entrantes y salientes corresponden a las neuronas ocultas [10, 13, 16, 26]

Se pueden seleccionar varios puntos de inicio de las redes neuronales

artificiales. Algunos comienzan hablando de cómo Ramón y Cajal en 1888 demuestra que el sistema nervioso está compuesto por una red de células individuales, las neuronas, ampliamente interconectadas entre sí y su funcionamiento [22]. Otros nombran a Alan Turing como pionero al estudiar el cerebro como una forma de ver el mundo de la computación. Independientemente, en todos los casos se toma a Warren McCulloch (neurobiólogo) y Walter Pitts (estadístico) como los primeros teóricos que concibieron fundamentos de la computación neuronal en el artículo “*A logical calculus of Ideas Imminent in Nervous Activity*” en 1943 [25]. Posteriormente, en 1956, tuvo lugar el congreso de Dartmouth, comúnmente considerado el nacimiento de la inteligencia artificial. Un año después Frank Rosenblatt comenzó el desarrollo de “Perceptron”.

Perceptron es considerada la red neuronal más antigua, siendo un sistema clasificador de patrones. Este modelo era capaz de reconocer patrones similares a los del entrenamiento pero que no había visto antes. Sin embargo, contaba con ciertas limitaciones, la más importante fue que no era capaz de resolver la función lógica OR exclusivo¹¹ ni clasificar clases no separables linealmente. Esto fue demostrado matemáticamente en los años 60 por Minsky y Papert, lo que desencadenó un fuerte desinterés en la computación neuronal.

No fue hasta principios de los años 80 cuando se consiguió devolver el interés en este campo con la publicación del artículo “*Hopfield Model o Cross-bar Associative Network*” de John Hopfield y el redescubrimiento de David Rumelhart y G. Hinton del algoritmo de aprendizaje de propagación hacia atrás (*backpropagation*) planteado por Paul Werbos en 1974 [6, 22, 24].

Existen dos tipos de redes neuronales principales, las redes convolucionales y las recurrentes. Las primeras se desarrollaron para el reconocimiento de imágenes y son útiles para distinguir elementos esenciales en una entrada de la red compleja [28]. Una particularidad de estas redes es la operación de convolución¹² que se realiza en las primeras capas utilizándose como filtro, esto permite que tenga muchas aplicaciones en el tratamiento de imágenes [23]. Por otro lado, las redes neuronales recurrentes son aquellas en las que la salida de una neurona vuelve en forma de entrada, por lo que puede haber ciclos en las conexiones entre las neuronas de cada capa. Cuentan con la peculiaridad de que no sólo aprenden de los datos sino también de secuencias de estos. Todo esto las convierte en buenas candidatas para trabajar con datos con dependencias estructurales en una dimensión, como el texto, o el audio [28, 15].

También se lleva a cabo una distinción en las redes neuronales según el

¹¹XOR u OR exclusiva es una puerta lógica cuya salida es verdadera si una, y sólo una, de las entradas es verdadera. Si ambas son falsas o ambas son verdaderas la salida es falsa.

¹²Una convolución es un operador matemático que transforma dos funciones f y g en una tercera que representa la magnitud en la que se superponen f y una versión trasladada e invertida de g .

paradigma de aprendizaje que siguen, existen dos tipos principales, las redes neuronales supervisadas y las no supervisadas. Las redes neuronales supervisadas se caracterizan por la presencia de un agente externo que controla el entrenamiento, este determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor controla la salida de la red y en caso de que ésta no coincida con la deseada se modifican los pesos de las conexiones con el fin de conseguir una salida aproximada a la esperada. En cambio, las redes con aprendizaje no supervisado no requieren influencia externa para ajustar los pesos de las conexiones entre sus neuronas. La red no recibe ninguna información por parte del entorno que le indique si la salida generada en respuesta a una determinada entrada es o no correcta [24].

Algunas de las Redes Neuronales más expandidas actualmente son AlexNet, VGGNet, MobileNet e InceptionNet.

AlexNet [21] es una red ampliamente utilizada que se dió a conocer en la competición de ImageNet, donde consiguió el primer puesto con una tasa de error del 15.3 % (2012). Su estructura se divide en dos GPUs y fue la primera red neuronal en utilizar ReLu¹³ como función de activación¹⁴ en vez de la función sigmoide¹⁵. Este cambio mejoró notablemente la velocidad de entrenamiento de las redes de aprendizaje profundo [8, 18, 40].

La segunda red destacada es VGGNet, esta también saltó a la fama gracias a ImageNet en 2014. La diferencia más destacable con AlexNet es la gran cantidad de capas convolucionales más que tiene. Además de una arquitectura más profunda presenta también otros nuevos conceptos [35, 40].

InceptionNet, desarrollada por Google, fue de las primeras redes que no se basó en añadir más capas convolucionales para mejorar la red, sino que cambió la estructura de estas usando filtros de varios tamaños. El objetivo era tener filtros de diferentes tamaños para un mismo objeto. Esto provocó que la red fuera más “ancha” en vez de más profunda [36].

La última red es MobileNet, esta también fue desarrollada por Google. Surgió con la intención de adaptar InceptionNet a dispositivos móviles. Su objetivo fue reducir el número de computaciones y parámetros pero manteniendo unos resultados similares [34, 17].

2.4. Identificación de objetos e imágenes

La detección e identificación de objetos es un campo de la visión artificial y el procesamiento de imágenes. El objetivo de la visión computacional es desarrollar algoritmos que reciban una imagen de entrada y produzcan una

¹³ReLu transforma los valores de entrada anulando los negativos y manteniendo los positivos.

¹⁴la función de activación en redes neuronales devuelve una salida que es generada por una o varias entradas. Cada capa de la red cuenta con una función de activación.

¹⁵La función sigmoide transforma los valores de entrada a una escala entre 0 y 1.

interpretación describiendo los objetos presentes, en qué posición y la relación espacial tridimensional de los objetos presentes. A pesar de que actualmente ya es posible detectar y reconocer objetos en conjuntos de miles de imágenes complejas, todavía no se cuenta con un paradigma aceptado y dominante con el que la mayoría de investigadores trabajen [4].

La detección, seguimiento y reconocimiento de objetos en imágenes es uno de los problemas claves en visión computacional [9]. Normalmente, en cada imagen sólo aparece una pequeña cantidad de los objetos posibles, pero estos pueden encontrarse en una gran cantidad de posiciones y escalas que se deben tener en cuenta [5].

Esto es algo utilizado en diversos ámbitos, desde objetos, personas y caras, hasta campos mucho más específicos. Un ejemplo es VOCUS (Visual Object detection with a CompUtational attention System) [12], un sistema capaz de seleccionar automáticamente secciones de interés en imágenes y detectar objetos específicos. Cuenta con dos modos de trabajo, un primer modo de exploración en el que no se especifica un objetivo y en el que se buscan zonas de interés. Son consideradas zonas de interés aquellas en las que haya grandes contrastes o características únicas (por ejemplo una oveja negra en un rebaño de ovejas blancas). El segundo modo es de búsqueda con un objetivo definido, en este modo se utiliza información previamente aprendida sobre el objetivo para seleccionar las computaciones más destacadas respecto a ello.

Otro ejemplo es el reconocimiento de matrículas de los coches presentes en una imagen [27]. En este caso se presenta una doble identificación, ya que primero debe localizarse la matrícula en la imagen y posteriormente sobre esta se tiene que llevar a cabo la segmentación e identificación de los caracteres que la componen.

En el apartado 2.2 se mencionaron varios proyectos de generación de *datasets* sintéticos. Todos ellos, además, tienen también como objetivo la identificación de objetos en imágenes.

SynthDet [19] trata de identificar los distintos productos contenidos en carritos de la compra con el objetivo de facturar automáticamente las adquisiciones de los clientes en supermercados. Está basado en Amazon Go[29], supermercado de la empresa Amazon, ubicado en Seattle (Estados Unidos). El modelo de Amazon Go se basa en un sistema de cámaras y sensores distribuidos por el establecimiento que analizan los movimientos de los clientes. Estos, además de obtener la información necesaria para realizar los cobros, también extraen información sobre las pautas de comportamiento de los consumidores [30]. SynthDet propone realizar la misma tarea de identificación y facturación de productos, pero entrenando el modelo a partir de imágenes sintéticas.

El proyecto sobre detección de drones [32] y el de identificación de piezas industriales [7], cuentan con un trasfondo similar. La generación del *dataset*

surge por la necesidad de entrenar el modelo para para finalmente llevar a cabo una identificación de objetos.

Pero la inclusión y el desarrollo de la detección y la identificación de objetos se limita al entorno de la investigación; sino que en la vida cotidiana también se percibe un crecimiento de este ámbito. Actualmente el uso de la detección e identificación de objetos se está expandiendo por más aplicaciones, una de las más conocidas y extendidas es Google Lens¹⁶, la cual tiene diversas funcionalidades dentro del mundo de la detección de objetos. Esta aplicación es una de las más completas en este dominio, a partir de una imagen o de la cámara del dispositivo escanea y traduce textos; identifica objetos y busca en Google otros similares; reconoce lugares y edificios, ofreciendo información sobre ellos; detecta animales y plantas; y registra operaciones matemáticas sobre las que ofrece explicaciones y resultados de la web.

Otra aplicación de Google con identificación de objetos es Google Fotos¹⁷, una aplicación de intercambio y almacenamiento de fotografía y vídeo. Puede utilizarse en conjunto con Google Lens para tener todas las funcionalidades mencionadas anteriormente sobre las imágenes almacenadas en la aplicación, pero además Google Fotos cuenta con su propia funcionalidad de identificación de texto, objetos y personas. En la aplicación pueden realizarse búsquedas que devuelven una selección de imágenes y vídeos en los que aparece la persona, objeto o texto buscado, además de funciones con posturas y acciones.

Otras aplicaciones similares, aunque algo más rústicas, son las recogidas en el paquete de Microsoft Office. En este caso la identificación se presenta como una funcionalidad de accesibilidad para personas con discapacidad visual. Se trata del texto alternativo, disponible para imágenes, formas, vídeos, gráficos y tablas, entre otros. En algunas de sus aplicaciones Microsoft va un paso más allá y genera el texto alternativo de las imágenes automáticamente, identificando personas, objetos y situaciones.

¹⁶<https://lens.google/intl/es-419/>

¹⁷<https://www.google.com/intl/es/photos/about/>

Bibliografía

- [1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattemberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] C. Alfonso, R. Estévez Estévez, J. M. Lobo, B. Lozano Diéguez, F. Prieto, J. Santamarta, and A. Gaerter. Emergencia climática en España. Diciembre 2016.
- [3] O. Alsing. Mobile object detection using tensorflow lite and transfer learning. Master's thesis, KTH, School of Electrical Engineering and Computer Science (EECS), 2018.
- [4] Y. Amit. *2D Object Detection and Recognition: Models, Algorithms, and Networks*. Mit Press. MIT Press, 2002.
- [5] Y. Amit, P. Felzenszwalb, and R. Girshick. *Object Detection*. Springer International Publishing, Cham, 2020.
- [6] X. Basogain Olabe. Redes neuronales artificiales y sus aplicaciones. *Dpto. Ingeniería de Sistemas y Automática, Escuela Superior de Ingeniería Bilbao. Open Course Ware.*[En línea] disponible en http://ocw.ehu.es/enseñanzas-tecnicas/redes-neuronales-artificiales-y-sus-aplicaciones/Course_listing. [Consultada 20-09-2012], 2008.
- [7] J. Cohen, C. F. Crispim-Junior, C. Grange-Faivre, and L. Tougne. CAD-based Learning for Egocentric Object Detection in Industrial Context. In *15th International Conference on Computer Vision Theory and Applications*, volume 5, Valletta, Malta, Feb. 2020. SCITEPRESS - Science and Technology Publications.

- [8] G. Cortina Fernández. Técnicas inteligentes para su integración en un vehículo autónoma. Trabajo de Fin de Grado en Ingeniería del Software, Facultad de Informática UCM, Departamento de Ingeniería del Software e Inteligencia Artificial, Curso 2019/2020., 2020.
- [9] B. Cyganek. *Object Detection and Recognition in Digital Images: Theory and Practice*. Wiley, 2013.
- [10] R. Flórez López, J. M. Fernández, and J. M. Fernández Fernández. *Las Redes Neuronales Artificiales*. Metodología y Análisis de Datos en Ciencias Sociales. Netbiblo, 2008.
- [11] R. Fonfría, R. Sans, and J. de Pablo Ribas. *Ingeniería ambiental: contaminación y tratamientos*. Colección productiva. Marcombo, 1989.
- [12] S. Frintrop. *VOCUS: A visual attention system for object detection and goal-directed search*, volume 3899. Springer, 2006.
- [13] G. A. Gómez Rojas, J. C. Henao López, and H. Salazar Isaza. Entrenamiento de una red neuronal artificial usando el algoritmo simulated annealing. *Scientia Et Technica*, 2004.
- [14] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks, 2014.
- [15] G. Guridi Mateos et al. Modelos de redes neuronales recurrentes en clasificación de patentes. B.S. thesis, 2017.
- [16] J. R. Hilera and V. J. Martínez Hernando. *Redes neuronales artificiales: fundamentos, modelos y aplicaciones*. 01 1995.
- [17] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weand, M. Andreetto, and H. Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017.
- [18] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5mb model size, 2016.
- [19] Y.-C. Jhang, A. Palmar, B. Li, S. Dhakad, S. K. Vishwakarma, J. Hoggins, A. Crespi, C. Kerr, S. Chockalingam, C. Romero, A. Thaman, and S. Ganguly. Training a performant object detection ML model on synthetic data using Unity Perception tools, Sep 2020.
- [20] R. Karim. *TensorFlow: Powerful Predictive Analytics with TensorFlow*. Packt Publishing, Limited, 2018.

- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.
- [22] P. Larranaga, I. Inza, and A. Moujahid. Tema 8. redes neuronales. *Redes Neuronales, U. del P. Vasco*, 12, 1997.
- [23] M. A. López Pacheco. Identificación de sistemas no lineales con redes neuronales convolucionales. *Cuidad de Mexico: Centro de investigación y de estudios avanzados*, 2017.
- [24] D. J. Matich. Redes neuronales: Conceptos básicos y aplicaciones. *Universidad Tecnológica Nacional, México*, 41, 2001.
- [25] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 1943.
- [26] B. Müller, J. Reinhardt, and M. Strickland. *Neural Networks: An Introduction*. Physics of Neural Networks. Springer Berlin Heidelberg, 1995.
- [27] C. Parra Ramos and D. Regajo Rodríguez. Reconocimiento automático de matrículas. *Universidad Carlos III de Madrid*, 2006.
- [28] R. Pavón Benítez. Técnicas de deep learning para el reconocimiento de movimientos corporales. Trabajo de Fin de Grado en Ingeniería del Software, Facultad de Informática UCM, Departamento de Ingeniería de Software e Inteligencia Artificial, Curso 2019/2020, 2020.
- [29] A. Polacco and K. Backes. The amazon go concept: Implications, applications, and sustainability. *Journal of Business and Management*, 24(1), 2018.
- [30] A. Polacco and K. Backes. The amazon go concept: Implications, applications, and sustainability. *Journal of Business and Management*, 24(1), 2018.
- [31] G. Ros, L. Sellart, J. Materzynska, D. Vazquez, and A. M. Lopez. The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [32] A. Rozantsev, V. Lepetit, and P. Fua. On rendering synthetic images for training an object detector. *Computer Vision and Image Understanding*, 137, 11 2014.
- [33] M. Sánchez and J. Castro. *Gestión y Minimización de Residuos*. Fundación Confemetal, 2007.

- [34] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [35] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [36] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015.
- [37] Unity Technologies. Unity Perception package, 2020.
- [38] S.-C. Wang. *Artificial Neural Network*. Springer US, Boston, MA, 2003.
- [39] P. Warden and D. Situnayake. *TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers*. O'Reilly Media, 2019.
- [40] W. Yu and Y. Bai. Visualizing and comparing alexnet and vgg using deconvolutional layers. 2016.
- [41] J. Zurada. *Introduction to Artificial Neural Systems*. West, 1992.

