
Aplicación móvil de identificación de objetos para reciclaje



TRABAJO DE FIN DE GRADO

Celia Castaños Bornaechea

Grado en Desarrollo de Videojuegos

Facultad de Informática

Universidad Complutense de Madrid

Junio 2021

Documento maquetado con T_EX!S v.1.0.

Este documento está preparado para ser imprimido a doble cara.

Aplicación móvil de identificación de objetos para reciclaje

Memoria que presenta para optar al título de Doctor en Informática

Celia Castaños Bornaechea

Dirigida por el doctor: [0.3em]

**Grado en Desarrollo de Videojuegos
Facultad de Informática
Universidad Complutense de Madrid**

Junio 2021

Copyright © Celia Castaños Bornaechea

Al duque de Béjar
y
a tí, lector carísimo

*I can't go to a restaurant and
order food because I keep looking
at the fonts on the menu.
Donald Knuth*

Agradecimientos

Inicio de las Leyes Orgánicas. Juan
Carlos I

Resumen

...

...

...

Índice

Agradecimientos	IX
Resumen	XI
1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	1
1.3. Plan de Trabajo	2
2. Estado del Arte	3
2.1. Aplicaciones de Reciclaje	3
2.2. Redes Neuronales	3
2.3. Generación de Imágenes	4
2.4. Identificación de Objetos e Imágenes	4
3. Generación de Imágenes	5
3.1. Introducción	5
3.2. Unity	6
3.3. Modelos	7
3.4. Capturas	7
4. Red Neuronal	9
4.1. Tensorflow	9
4.2. Tensorflow Lite	9
4.3. Entrenamiento	9
4.4. Resultados	10
4.5. Comparación	10
5. Aplicación de Identificación de Objetos	11
5.1. Introducción	11
5.2. Tensorflow Lite	11
5.3. Identificación de Objetos	11

6. Conclusiones y Trabajo Futuro	13
6.1. Conclusiones	13
6.2. Trabajo Futuro	13
A. Así se hizo...	15
A.1. Introducción	15

Índice de figuras

Índice de Tablas

Capítulo 1

Introducción

1.1. Motivación

La idea de este Trabajo de Fin de Grado surge de una necesidad real de ayuda y aprendizaje para reciclar distintos residuos. Con la cantidad de diferentes materiales existentes en el día a día, actualmente muchas veces se hace complicado saber cómo reciclarlos todos adecuadamente. Así, surgió una necesidad de hacer algo al respecto que no sólo fuera útil, sino además fácil de utilizar.

Como seres humanos la forma principal y más sencilla, para la mayoría, de identificar objetos es la vista; así que por qué no trasladar esto a la identificación de residuos.

Por otro lado, para llevar a cabo una aplicación de identificación, ya sea imágenes u objetos, son necesarias redes neuronales y para entrenar estas son necesarias imágenes. En el caso de este TFG miles de fotos para contar con diversidad de objetos y materiales diferentes. Para agilizar la obtención de estas imágenes se ha querido explorar, además, la posibilidad de generarlas de manera automática a partir de un modelo 3D y así no tener que hacer todas las fotos, una a una, personalmente.

1.2. Objetivos

Los objetivos de este proyecto son desarrollar un prototipo de aplicación móvil de identificación de objetos para reciclaje, haciendo más accesible la información a los ciudadanos; y encontrar una alternativa a tener que hacer cientos de fotografías a distintos objetos para el entrenamiento de la aplicación. Para evitar esto el objeto es desarrollar otra aplicación, esta para ordenador, que permita generar imágenes para el entrenamiento de redes neuronales a partir de modelos 3D. Además, que dicha aplicación pueda ser distribuida y los usuarios puedan introducir sus propios modelos para generar

imágenes de lo que requieran.

1.3. Plan de Trabajo

El trabajo se divide en tres partes: la generación de imágenes, el entrenamiento de la red neuronal y la identificación de objetos.

En la primera parte se debe llevar a cabo una aplicación que cargue modelos 3D, separados por material, y realice numerosas imágenes a cada uno cambiándoles la posición y rotación y, además, cambiar el fondo para obtener diversidad en las imágenes. Estas imágenes deberán ser guardadas separadas por el material al que corresponden, igual que lo estaban los modelos al cargarlos.

Con las imágenes generadas del paso anterior tiene lugar el entrenamiento de la Red Neuronal. Para llevar a cabo esto se deberá investigar sobre los distintos tipos de Redes Neuronales y elegir la opción más adecuada para que el resultado, el modelo entrenado, sea utilizado desde una aplicación móvil.

Por último, como se ha mencionado antes, tendrá lugar el desarrollo de una aplicación para dispositivos móviles que haciendo uso de la cámara y el modelo entrenado identifique el material del objeto al que se está enfocando y después indique al usuario cómo se debe reciclar dicho material.

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. Con la generalización del uso de los dispositivos móviles en la población favorece a que esta sea la mejor forma de encontrar esta ayuda.

En España, para empezar, existe la aplicación de “AIRE Asistente Inteligente de Reciclaje” disponible tanto para iOS¹ como Android² aplicación publicada por “ecoembes”, empresa encargada del reciclaje de los residuos de los contenedores amarillo y azul en España.

Carrefour, por ejemplo, también cuenta con una aplicación de ayuda para reciclaje³, esta a partir del código de barras del ticket de la compra que se haya hecho indica cómo reciclar los envases de los productos de algunas marcas.

En el lado internacional se encuentran aplicaciones como RecycleRight, Brisbane Bin and Recycling, Grow Recycling o Recycle Coach, aunque distintas entre ellas todas estas aplicaciones cuentan con un factor de explicación y aprendizaje sobre reciclaje para el usuario.

2.2. Redes Neuronales

Redes neuronales más usadas. En qué se usan. Aplicaciones que las utilicen.

¹<https://apps.apple.com/es/app/air-e/id1442582214>

²<https://play.google.com/store/apps/details?id=com.ecoembes.aire&hl=en>

³<https://www.reciclaya.app/es/como-funciona>

2.3. Generación de Imágenes

Esta parte, de generación de imágenes, se encuentra menos explorada que las demás del capítulo. Aún así encontramos numerosos acercamientos desde distintas áreas de investigación.

En este ámbito se encuentran las Redes Generativas Adversativas, que se relaciona con el tratamiento de imágenes. Este modelo consta de dos redes neuronales denominadas generador y discriminador, y 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 discriminador distingue entre los datos reales que se le proporcionan y los generados por la red anterior. Esto fue propuesto en 2014 en el artículo “Generative Adversarial Networks”⁴.

2.4. Identificación de Objetos e Imágenes

Google Lens, Google Photos, Microsoft Office ->ejemplos más conocidos.

⁴<https://arxiv.org/abs/1406.2661>

Capítulo 3

Generación de Imágenes

RESUMEN: Para evitar tener que hacer o descargar miles de fotos para ser utilizadas para el entrenamiento de la red neuronal se tomó la decisión de tratar de hacer un generador de imágenes a partir de modelos 3D.

3.1. Introducción

Como el objetivo del proyecto es reconocer e identificar distintos objetos y su material para llevarlo a cabo es necesario realizar el entrenamiento de una red neuronal y para esto son necesarias cientos de fotografías para cada material diferente que se quiera introducir.

Conseguir tantas imágenes distintas y a la vez claras para llevar a cabo el entrenamiento supone un trabajo costoso, lento y tedioso; sin tener en cuenta el almacenaje de estas. Puede llevarse a cabo descargando imágenes de la red, pero para esto hay que tener en cuenta los derechos de autor y uso de cada una de ellas. También se plantea la idea de llevar a cabo un vídeo de objetos de ese material y luego separar los frames de este y usarlos como imágenes para el entrenamiento. Esta opción presenta varios problemas también, ya que si se graban a todos los objetos a la vez podría darse que no hubiera contenido suficiente para el entrenamiento, o que este no se mostrara lo suficientemente claro. Si se graban los distintos objetos habría que supervisar el descarte del paso intermedio entre un objeto y el siguiente. Para esta opción como para hacer fotografías a mano encontramos un problema importante ya que habría que poseer todo aquello con lo que se quiere entrena, o en su defecto hacer las capturas en un establecimiento que contara con numerosos modelos distintos del mismo material. Pero en este caso hay que tener en cuenta la legislación para llevar esto a cabo. Además, con estas opciones existe un riesgo a perder imágenes, en el caso que ocurriera

sería necesario volver a pasar por todo el proceso que se decidió como el definitivo. Para evitar todos estos inconvenientes se plantea la posibilidad de llevar a cabo una aplicación con la que automatizar este proceso.

Actualmente se va viendo una gran mejoría en los gráficos de contenidos digitales audiovisuales, ya sean películas, videojuegos, videoclips, anuncios, etc. Esta mejora llega hasta tal punto que a veces resulta difícil distinguir entre realidad y CGI (Computer-generated imagery). Teniendo esto en cuenta no se ha querido desaprovechar la oportunidad de enfocarlo desde un punto más cercano a este; y es esta es la opción que se decidió para realizar las imágenes de entrenamiento de la red neuronal. Consiste en llevar a cabo una aplicación en la que se vayan cargando distintos modelos de objetos y residuos; a cada objeto, tras ser cargado, se le harán numerosas capturas en distintas posiciones y rotaciones. Como resultado se contará con el número de capturas que se quieran de cada objeto y cada material pudiendo aumentar esta cantidad siempre que se desee o necesite.

3.2. Unity

Para llevar a cabo la generación de imágenes se ha utilizado Unity¹. Unity es una herramienta de creación de videojuegos desarrollada por Unity Technologies. Unity no se limita al ámbito de los videojuegos únicamente, sino que su uso está más extendido; es una herramienta utilizada en muchos otros sectores, como el cinemático, transporte y producción o en arquitectura, ingeniería y construcción. Se ha elegido Unity como herramienta ya que simplifica el trabajo que se quiere realizar, de esta forma no hay que crear y llevar a cabo una estructura compleja, que ya ha sido desarrollada en más ocasiones, que es necesaria para lo que se quiere llevar a cabo. Como Unity hay otros programas que se podrían haber utilizado en su lugar tales como Unreal Engine ² o incluso Blender ³. La decisión de utilizar Unity fue tomada debido a que es el programa sobre el que se tiene más conocimiento y más se ha utilizado previamente.

Para este proyecto, Unity las herramientas más necesarias tales como la cámara, el sistema de iluminación o las escenas. Además, hay que contar con la amplia documentación que existe sobre esta herramienta y la gran cantidad de cosas que permite llegar a desarrollar.

A pesar de que Unity facilita el trabajo igualmente hay que plantear y desarrollar el funcionamiento de la aplicación. El funcionamiento básico de esta consiste en ir cargando los modelos guardados, uno a uno, que están agrupados por el material al que pertenecen. En cada frame al ejecutar la aplicación se cambia la posición y la rotación del objeto actual en la escena y

¹<https://unity.com/es>

²<https://www.unrealengine.com/en-US/>

³<https://www.blender.org/>

el fondo. La nueva posición y rotación serán aleatorias, controlando siempre que no queden fuera de los límites de la cámara. Una vez situado el objeto y generado el fondo se realiza una captura de la pantalla. Cuando se hayan recorrido todos los modelos de todas las carpetas de materiales la aplicación terminará su ejecución.

Se exploró la opción de hacer de esto una aplicación ejecutable en la que el usuario pudiera introducir, además, sus propios modelos para poder generar más imágenes para entrenar la red neuronal y hacer la identificación de objetos más completa. Esta idea se tuvo que descartar debido a la dificultad de importar modelos. Se investigó al respecto pero los resultados no fueron exitosos. La introducción de modelos 3D es algo que no suele utilizarse en videojuegos y Unity no cuenta con una manera fácil de llevarlo a cabo. Otro factor importante para el fracaso de esto fueron los numerosos problemas que surgieron al introducir modelos correctamente al editor de Unity. Viendo que esto ya consumía mucho tiempo y esfuerzo se decidió descartar la aplicación para más usuarios en la que poder introducir nuevos modelos.

3.3. Modelos

La generación de imágenes se lleva a cabo a partir de modelos 3D. Los modelos están agrupados en carpetas según su material. La aplicación recorre todas las carpetas de materiales disponibles. Todos los modelos se cargan al iniciar la aplicación, se guarda cuántos materiales hay. Estos modelos son de distintos objetos y materiales y están ordenados y separados por esto último. De cada uno de los objetos se toman por defecto diez capturas y en cada una tendrán una posición y rotación diferente, siendo esta y aleatoria.

3.4. Capturas

Capítulo 4

Red Neuronal

4.1. Tensorflow

4.2. Tensorflow Lite

Para la red neuronal se utiliza TensorFlow Lite, un framework de código abierto de aprendizaje profundo para dispositivos móviles. TensorFlow Lite Permite ejecutar modelos de aprendizaje profundo en dispositivos móviles. A través de él se puede utilizar modelos entrenados en una computadora en un dispositivo móvil sin necesidad de utilizar un servidor. Utiliza MobileNet, que está diseñada y optimizada para imágenes en móviles, incluyendo detección y clasificación de objetos, detección de caras y reconocimiento de **lugares**

modelo de TensorFlow Lite (un formato FlatBuffer optimizado que se puede identificar mediante la extensión de archivo .tflite TensorFlow Lite incorpora (esto significa que se ejecuta en el dispositivo móvil) TensorFlow en dispositivos móviles. Anunciado en 2017, el paquete de software TFLite está diseñado específicamente para el desarrollo móvil. TensorFlow Lite “Micro” es, por otro lado, una versión especialmente para microcontroladores, que se fusionó recientemente con uTensor de ARM.

4.3. Entrenamiento

A partir del blog de TensorFlow y sus ejemplos se ha generado un script sencillo de carga de imágenes, separación de etiquetas y entrenamiento con todo esto que genera el modelo entrenado, un .tflite ya que se utiliza la extensión de TensorFlow: TensorFlow Lite, para dispositivos móviles. Y un documento de texto plano con las distintas etiquetas disponibles. Dicho script tiene dos maneras de utilizar las imágenes, la primera es que carga todas las imágenes de una carpeta, cada tipo de imagen estará separada en subcarpetas las cuales serán las distintas etiquetas. Una vez cargadas todas las

imágenes se separan en las de entrenamiento y las de test, siendo un 90 % para lo primero. La otra opción es para comprobar qué tal detectaba los objetos si se entrenaba con las imágenes generadas (spoiler: sale mal), se cargan primero las imágenes de entrenamiento de la carpeta correspondiente, la cual tiene la separación de materiales en subcarpetas; y después se cargan las imágenes de test, que serán imágenes reales de objetos de ese material, también estarán organizadas en subcarpetas según el material correspondiente, y debe haber tantas como en las de entrenamiento. Después de entrenar la red se comprueba con cuánta precisión detecta los objetos en las imágenes de test.

“Todas las distintas opciones que he hecho de elegir imágenes, cuáles para entrenamiento cuáles para test; cuáles reales y cuáles generadas. Pruebas con las latas, pruebas con la madera”

4.4. Resultados

4.5. Comparación

Comparaciones de cómo de bien/mal sale el resultado]

Capítulo 5

Aplicación de Identificación de Objetos

5.1. Introducción

Se trata de una aplicación móvil para Android a través de la cual, apuntando un objeto se identifica el material principal de este y se informa al usuario de cómo se debería reciclar el material identificado.

5.2. Tensorflow Lite

Por qué, cómo se usa, qué cosas proporciona/facilita.]

5.3. Identificación de Objetos

Qué hace ->apuntas a un objeto saca con qué porcentaje considera que es cierto material.]

...

...

Capítulo 6

Conclusiones y Trabajo Futuro

6.1. Conclusiones

Conclusiones de la generación de imágenes, si es útil si ahorra tiempo. ¿Encuesta sobre si la aplicación sería útil? Para escribir algo sobre que sí es algo que la gente considere necesario. Y qué opciones prefieren.]

6.2. Trabajo Futuro

Añadir más materiales, más objetos, hacer una interfaz propia, accesibilidad, modelos más realistas para el entrenamiento?, Mejor iluminación?]

Apéndice A

Así se hizo...

...

...

RESUMEN: ...

A.1. Introducción

...

*—¿Qué te parece desto, Sancho? — Dijo Don Quijote —
Bien podrán los encantadores quitarme la ventura,
pero el esfuerzo y el ánimo, será imposible.*

*Segunda parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes*

*—Buena está — dijo Sancho —; fírmela vuestra merced.
—No es menester firmarla — dijo Don Quijote—,
sino solamente poner mi rúbrica.*

*Primera parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes*

