

Capítulo 6

Conclusiones y trabajo futuro

6.1. Recapitulación

El resultado del proyecto es un prototipo de una aplicación de ayuda al reciclaje para dispositivos con sistema operativo Android. Muestra aquello a lo que se apunta con la cámara del dispositivo e identifica y ofrece información sobre el material que compone el objeto además de indicar la manera adecuada de desecharlo. Como se ha comentado a lo largo del documento, para llevarla a cabo es necesario utilizar técnicas de visión artificial, las que actualmente conllevan la generación de un modelo entrenado por una red neuronal. Esto supone la necesidad de establecer varios subobjetivos.

El primer subobjetivo que se planteó es la obtención de un *dataset* amplio, claro y variado con imágenes de todos los materiales que se vayan a tener disponibles. Buscando facilitar la obtención de *datasets*, se ha desarrollado una aplicación de generación de imágenes sintéticas a partir de modelos tridimensionales. Esta está generada en la plataforma Unity, motor de videojuegos multiplataforma creado por Unity Technologies. El funcionamiento de esta aplicación consiste en ir cargando los distintos modelos 3D disponibles en los recursos, separados por material, y realizar numerosas capturas a cada uno hasta que se recorren todos. En cada captura tomada tanto la posición y rotación del objeto, como el fondo de la imagen, se establecen de manera aleatoria, proceso necesario para contar con una alta diversidad en las imágenes. Debido a los inconvenientes surgidos durante la obtención de los modelos, finalmente se obtuvo un *dataset* compuesto por tres materiales diferentes (metal, vidrio y plástico), donde para metal se mezclan imágenes reales con sintéticas.

Una vez generado el *dataset* se dio paso al segundo propósito para realizar la aplicación. Este se divide en dos pasos diferentes. El primero, y primordial, es el entrenamiento de la red neuronal y la generación del modelo entrenado utilizando las imágenes generadas previamente. El resultado del entrenamiento es incorporado posteriormente en Android Studio para su uso en la

aplicación final. Para este apartado se ha utilizado una librería de TensorFlow Lite que permite el entrenamiento por transferencia utilizando la red EfficientNet y proporciona herramientas para llevar a cabo el entrenamiento de manera fácil e intuitiva.

Como segundo paso, una vez finalizado el entrenamiento de la red neuronal, se realizan diversas pruebas y comparaciones con la finalidad de encontrar la proporción entre imágenes sintéticas y reales que ofrecen el mejor equilibrio entre facilidad de obtención del *dataset* y exactitud del modelo entrenado. La conclusión de las pruebas fue que a partir del 30 % de imágenes reales la precisión sufre muy pocos cambios y se mantiene siempre por encima del 90 % de confianza. Es decir, la precisión del modelo cuando todas las imágenes son reales y cuando se utiliza en uno de los materiales sólo un 30 % de estas, es muy similar. De esta forma, se tomó la decisión de utilizar el modelo con el 70 % de imágenes generadas y el 30 % restante reales para el metal, que cuenta con un 93 % de confianza.

Finalmente, tiene lugar el desarrollo de la aplicación para dispositivos Android mediante la herramienta Android Studio y utilizando TensorFlow Lite. La aplicación, a través de la cámara, recibe imágenes de los *frames* capturados y consultando el modelo importado trata de identificar el objeto que aparece en la imagen. Devuelve como resultado las etiquetas disponibles ordenadas según el porcentaje de seguridad con el que considera que se trata de ese objeto, informando al usuario del material, el contenedor donde desecharlo y la confianza con la que considera que se trata de dicho material.

Para probar su funcionamiento se hicieron ensayos sobre objetos del mundo real. Para ello, se realizaron cambios en la aplicación dando lugar a otra de *test*. A través de ella se compara el funcionamiento de varios modelos simultáneamente. Con los datos obtenidos, se corroboró el correcto funcionamiento del modelo entrenado con el *dataset* seleccionado.

6.2. Conclusiones

En el proceso se han ido observando diversas dificultades y conclusiones importantes que han afectado al desarrollo del trabajo y que se deben tener en cuenta en posibles ampliaciones futuras o el desarrollo de otros proyectos similares o relacionados.

Se ha observado la dificultad real de obtener *datasets* amplios sobre objetos concretos si no se cuenta con muchos recursos, como se ha experimentado en el desarrollo del proyecto, y lo que ha provocado limitarse a contar con solamente tres materiales. Una solución a este problema puede ser el uso de imágenes sintéticas para completarlos. Este trabajo ha demostrado que, al menos en el contexto planteado, esta opción tiene resultados altamente positivos, ya que poblando una parte significativa del *dataset* con imágenes sintéticas, la precisión apenas se ve afectada. Tras la realización de pruebas

en el entorno real, se corrobora el correcto funcionamiento de la aplicación obteniendo resultados acertados en la identificación y con confianza considerablemente positiva. Por este motivo, se considera que la aplicación de generación de imágenes es un aporte útil para muchos proyectos que se salen del estándar en el ámbito del reconocimiento de objetos e imágenes.

No obstante, para que esto se cumpla, conseguir imágenes sintéticas tiene que resultar más sencillo que conseguir fotografías, ya que sigue existiendo cierta dificultad para generar el *dataset*. Esto es provocado por la obtención de modelos tridimensionales, los cuales es importante que cuenten con un alto nivel de realismo en las texturas y materiales. Esto se debe a que la manera en que la luz incide y se refleja sobre ellos puede resultar poco realista y generar problemas en el entrenamiento de la red neuronal, haciendo que el problema de la generación de *datasets* no quede solventado en su totalidad.

La iluminación es otro punto decisivo, a pesar de las facilidades que ofrece Unity para crear distintos tipos de iluminación, si no se tiene suficientes conocimientos y experiencia, el resultado puede terminar siendo pobre y, de nuevo, poco realista. Esto es un problema ya que la aplicación va a utilizarse sobre objetos reales y si no hay un equilibrio entre lo que se utiliza para el entrenamiento y el uso final, la precisión de la aplicación queda notablemente disminuida. Esto pudo observarse en los entrenamientos en los que el *dataset* contaba con más de un 80 % de imágenes sintéticas, al ocupar estas la mayor parte del entrenamiento y diferir cómo se ve en la realidad, los resultados fueron peores en comparación con el resto de proporciones.

En vista de los resultados obtenidos durante el trabajo, y con el apoyo de otros proyectos donde se ha trabajado con *datasets* sintéticos, se puede afirmar que si se va a utilizar un *dataset* compuesto por imágenes sintéticas es necesario que estas se encuentren entremezcladas con reales para un correcto funcionamiento. Esto se debe a que las cámaras virtuales y las reales son sensores diferentes.

Por último, una última observación a tener en cuenta, es que al llevar a cabo el entrenamiento respecto a tres materiales solamente, y donde las imágenes de sólo uno de ellos se mezclan con sintéticas, se observa que esto afecta no solamente al material que tiene las imágenes generadas, sino que también influye a los demás, generando errores en la identificación de objetos compuestos de materiales sobre los que se ha entrenado por completo con imágenes reales.

6.3. Trabajo futuro

El proyecto cuenta con diversos puntos ampliables. El primero es la obtención de un *dataset* más amplio, añadiendo más variedad de materiales y objetos identificables. Esto puede hacerse a partir de imágenes reales únicamente, o bien mezclándolas con imágenes sintéticas. Para esta segunda

opción, es necesario conseguir o generar modelos tridimensionales de todos los materiales y objetos que se quieran incorporar, los cuales deben tener una calidad bastante elevada. Otro factor ampliable, para permitir mayor diversidad en las imágenes generadas, es la extensión del número de imágenes disponibles para el fondo.

Con el objetivo de generar imágenes más realistas, sería necesario revisar la iluminación presente en la escena en la que tiene lugar las capturas, acercándolo a cómo se ven los objetos posteriormente en un entorno real. Asimismo, es necesario adquirir un *dataset* de imágenes reales de los materiales que se quieran agregar. La cantidad de imágenes de este estaría regido por si se va combinarse con imágenes sintéticas o no.

Una vez obtenido el *dataset*, simplemente es necesario entrenar y generar el modelo para, finalmente, importarlo en la aplicación mediante Android Studio. Todos los materiales añadidos además deben relacionarse con el contenedor o lugar de desecho apropiado, que es la información que busca el usuario.

Para mejorar la red y ampliar el *dataset*, podría desarrollarse la opción de que los usuarios puedan enviar sus propias fotografías etiquetadas con el material del que se trate, dando lugar a un *dataset* más completo que se encuentre en continuo crecimiento ofreciendo así un mejor servicio a los usuarios.

Otra mejora adicional interesante, es convertir la generación de imágenes en una aplicación ejecutable independiente de Unity, en la que los usuarios puedan generar *datasets* para sus proyectos a partir de modelos propios suyos que importen en esta, o aprovechar algunos de los que se ofrezcan por defecto, por ejemplo los utilizados en este trabajo. Aunque la importación de modelos en tiempo de ejecución es algo poco extendido en el mundo de los videojuegos, debido a la cantidad y diversidad de usuarios con los que cuenta Unity, esta funcionalidad ha sido ampliamente discutida y explorada. Pueden incluso encontrarse paquetes para llevar esto a cabo, por ejemplo “TriLib 2 - Model Loading Package” de Ricardo Reis¹, “Runtime OBJ Importer” de Dummiesman² u “OBJReader” de Starscene Software³, así como numerosas discusiones en foros sobre el tema.

Respecto a la aplicación de Android también existen varias posibilidades de trabajo futuro. Una de ellas es llevar a cabo una interfaz más personalizada que pueda ofrecer más información sobre los distintos cubos disponibles o el proceso de reciclaje. Otra, es desarrollar la misma aplicación para dispositivos iOS, permitiendo así su acceso a un mayor número de usuarios. Además, una ampliación necesaria de la aplicación, es la incorporación de

¹<https://assetstore.unity.com/packages/tools/modeling/trilib-2-model-loading-package-157548>

²<https://assetstore.unity.com/packages/tools/modeling/runtime-obj-importer-49547>

³<https://starscenesoftware.com/objreader.html#ObjReader>

accesibilidad con el objetivo de permitir su uso a todo tipo de público. Dichas mejoras corresponden, entre otras, a la introducción de una opción de voz, personalización del tamaño de la fuente o de los colores y contrastes.