

RECONOCIMIENTO Y DETECCIÓN DE FRAGMENTOS DE VIDRIO MEDIANTE EL USO DE YOLOv8

GLASS FRAGMENT RECOGNITION AND DETECTION USING YOLOv8

Luis Edgar Alanís Carranza

Tecnológico Nacional de México / TES de Coacalco, México
luis.alanis@tesco.edu.mx

Brandon Misael Muñoz Ochoa

Tecnológico Nacional de México / TES de Coacalco, México
202210254@tesco.edu.mx

José Emmanuel Sánchez Sánchez

Tecnológico Nacional de México / TES de Coacalco, México
Jose.emmanuel.sanchez.mct@tesco.edu.mx

Cristian Adrián Aguilar Contreras

Tecnológico Nacional de México / TES de Coacalco, México
202210245@tesco.edu.mx

José Roberto Pérez Torres

Tecnológico Nacional de México / TES de Coacalco, México
jose.roberto@tesco.edu.mx

Grethell Georgina Pérez Sánchez

Universidad Autónoma Metropolitana, Unidad Azcapotzalco, México
ggps@azc.uam.mx

Recepción: 24/noviembre/2023

Aceptación: 24/abril/2024

Resumen

En este artículo se presenta el desarrollo de un algoritmo para el reconocimiento y detección de fragmentos de vidrio, impulsado por la red neuronal convolucional del modelo YOLO versión 8; empleando las funciones para segmentación de imágenes y detección de objetos en tiempo real que ofrece dicho modelo. Como punto de partida se plantea la necesidad de optimizar los procesos manuales de clasificación utilizados en las plantas de reciclaje de vidrio mediante un sistema automatizado. El trabajo desarrollado permite concluir que la red convolucional YOLO posee características idóneas de rapidez y precisión para la clasificación de

vidrio de los colores verde, ámbar y transparente. Se describen los procesos de recolección y etiquetado de datos, haciendo uso del software Anaconda y la herramienta Labelimg, y finalmente se destaca la obtención de una precisión promedio del 88% en el reconocimiento de imágenes tras el entrenamiento con una duración de 400 épocas.

Palabras Clave: Aprendizaje de máquina, Aprendizaje profundo, Reconocimiento de objetos, Vidrio, YOLO.

Abstract

This article presents the development of an algorithm driven by the convolutional neural network of the YOLO model in its version 8 for the recognition and detection of fragments using the functions for image segmentation and object detection in real time offered by such model. As a starting point, it is considered the need to optimize the classification processes in glass recycling plants through an automated system. The work carried out allows us to conclude that the YOLO convolutional network has the optimal characteristics for carrying out the project due to its speed and precision. The data collection and labeling processes are described, using the Anaconda software and the Labelimg tool, and finally, obtaining an average accuracy of 88% in image recognition after training with a duration of 400 epochs is highlighted.

Keywords: Deep learning, Glass, Machine learning, Object recognition, YOLO.

1. Introducción

La contaminación es un problema de importancia a nivel mundial, por lo que se deben tomar acciones para reducir sus consecuencias, como el cambio climático y el calentamiento global. Ante esta problemática, la respuesta internacional se ha enfocado en el desarrollo de métodos para reducir el impacto ambiental [Alcaraz-Londoño, 2022], [Giraldo, 2022], por lo que se ha puesto atención en el desarrollo de sistemas que permitan la clasificación y separación de materiales [Expósito, 2022].

El vidrio es un material versátil con muchas aplicaciones diferentes [He, 2021], [Kaspi, 2022], [Tao, 2023], [Zhou, 2007]. De acuerdo con la Agencia de Protección

Ambiental de Estados Unidos, en el año 2017 sólo la cuarta parte del vidrio fabricado fue reciclado [Bohn, 2021]. El desperdicio de vidrio que se genera es alto, por lo que la industria de reciclaje de dicho material es amplia y sigue creciendo [Bristogianni, 2022]. La separación del vidrio de resto de desperdicio es un reto que, si bien se ha resuelto con soluciones automáticas, estas son costosas y están al alcance de las grandes industrias. Las pequeñas y medianas empresas del giro emplean mano de obra humana para hacer la separación manual de los desperdicios.

Inteligencia Artificial (IA)

La clasificación automática de objetos ha atraído mucha atención en el campo de la inteligencia artificial (IA), especialmente en situaciones complejas como la clasificación y separación de fragmentos de vidrio en la industria del reciclaje. El proceso de clasificación enfrenta varios problemas, como cambios en la forma, el tamaño y el color del fragmento de vidrio, o defectos en la muestra, por lo que se requieren métodos que puedan llevar a cabo una clasificación precisa y eficiente. En [Ponce, 2010] se aborda a la Inteligencia Artificial y su aplicación en la ingeniería, que por medio del uso de simbología y procedimientos no algorítmicos ayuda en la resolución de problemas.

Aprendizaje de Máquina (Machine Learning)

Por otro lado, el aprendizaje de máquina es una rama de la IA cuyo objetivo es el desarrollo de técnicas que permitan a las computadoras aprender la resolución de problemas específicos, empleando algoritmos que reconocen patrones y comportamientos a partir de datos que sirven como ejemplo [Centeno, 2019]. Dentro del aprendizaje de máquina existen dos enfoques para llevar a cabo el aprendizaje para la resolución de problemas: el aprendizaje supervisado y el no supervisado. Este trabajo se enfoca en el primero, donde se dispone de un conjunto de datos, y se sabe cuál es la respuesta correcta a la que debe llegar el ordenador, pero se desconoce la forma en la que se llega de los datos a la respuesta. En [Centeno, 2019] se establece que en el aprendizaje supervisado se dispone de un conjunto de datos, y se sabe cuál es la respuesta correcta a la que debe llegar

el ordenador, pero se desconoce la forma en la que se llega de los datos a la respuesta. A su vez, el aprendizaje supervisado se centra en dar solución a los retos que presentan la regresión y la clasificación. El primero tiene como objetivo la predicción de resultados que pueden surgir dentro de un rango continuo de valores a partir de un conjunto de datos de entrada. Por ejemplo, determinar la edad de una persona dada una fotografía suya. Por otro lado, la clasificación se basa en la predicción de resultados discretos. Por ejemplo, reconocer una letra escrita a mano.

Aprendizaje Profundo (Deep Learning)

Una rama del aprendizaje de máquina es el aprendizaje profundo que emplea una estructura de redes neuronales artificiales (RNA) semejantes a la estructura neuronal del cerebro humano. Su principal característica es que, a diferencia de los modelos clásicos de aprendizaje de máquina que se extraen manualmente las características, esta técnica tiene la capacidad de usar modelos analíticos para originar predicciones [Janiesch, 2021].

El aprendizaje profundo permite el entrenamiento a partir de diversos elementos, como pueden ser imágenes, texto o sonido, para luego entonces emplear lo aprendido para llevar a cabo tareas de clasificación. El entrenamiento se lleva a cabo empleando grandes cantidades de datos etiquetados, así como redes neuronales de muchas capas, por lo que requiere de gran potencia de cálculo.

Redes Neuronales Artificiales (RNA)

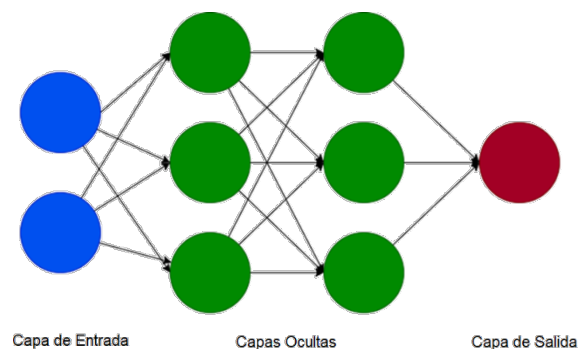
“Una Red Neuronal Artificial es un modelo matemático inspirado en el comportamiento biológico del cerebro humano” [Centeno, 2019]. Se compone de varios elementos simples denominados neuronas, donde cada una tiene asociados datos de entrada con los que realiza un cálculo y como salida genera una única respuesta.

Las neuronas forman grupos dentro de la red, pudiendo agruparse en niveles o capas que dependiendo de su ubicación se clasifican en:

- La capa que recibe la información de manera directa desde el exterior, interiorizándola en la red, conocida como capa de entrada.

- Las capas intermedias que se encuentran ocultas internamente dentro de la red, y que llevan a cabo el procesamiento de los datos de entrada.
- La capa que transmite la información de la red hacia el exterior de esta, conocida como capa de salida.

En la figura 1 se presenta una red de 4 capas (la capa de entrada, 2 capas ocultas más la capa de salida). Las flechas indican el flujo de datos dentro de la red y cada círculo representa una neurona individual.



Fuente: elaboración propia.

Figura 1 Diagrama de una red neuronal artificial multicapa.

Redes Neuronales Convolucionales

La convolución es una operación matemática con dos funciones cualesquiera que arroja como salida una tercera función [Schlumberger, s.f.]. En el caso de las redes neuronales convolucionales, la convolución es usada para extraer características importantes de los datos de entrada para establecer relaciones. Las capas convolucionales de una red aplican filtros a los datos de entrada y generan mapas de características, resaltando así datos relevantes.

YOLO

“YOLO (You Only Look Once) es una red convolucional que predice múltiples cuadros delimitadores y probabilidades de clase para cada uno” [Miranda, 2019]. En otras palabras, YOLO puede analizar una imagen una sola vez y generar predicciones precisas de la ubicación y clase de un objeto dentro de esta en tiempo

real. En [López, 2022] se emplea el término de *precisión* como parámetro para medir el rendimiento del modelo. Esta característica hace que YOLO sea un modelo muy eficiente y rápido en la detección de objetos dentro de imágenes y videos. Una clase es un conjunto de datos que comparte una serie de características en común que las identifica como similares.

YOLO v8 es la última versión de la red convolucional YOLO desarrollada por Ultralytics, y que fue lanzada en enero de 2023. Esta versión es considerablemente más precisa e intuitiva que sus versiones anteriores, por lo que es una gran opción para los sistemas de clasificación y reconocimiento de objetos [Solawetz, 2023].

Descripción de algunos trabajos previos en los que se ha empleado YOLO:

- [Giraldo, 2022] desarrolló un sistema clasificador de residuos sólidos mediante la creación de una red neuronal convolucional basada en el modelo de Efficientnet [Fu, 2023]. Concluyó con una precisión del 80% en el reconocimiento de imágenes desconocidas para la red y un 98% en imágenes conocidas.
- [Estrella, 2020] desarrolló una aplicación clasificadora de dígitos escritos a mano mediante el desarrollo de una red neuronal convolucional. Concluyó con una precisión del 91.07% en el reconocimiento de imágenes desconocidas para la red.
- [Capistran, 2020] presentó un sistema de análisis y detección de tráfico vehicular utilizando YOLO y comparándolo con otros métodos de detección de objetos. Concluyó que YOLO es un método más preciso y seguro que el resto, por lo que recomienda su uso.
- [López, 2022] desarrolló un sistema de detección y control de estudiantes universitarios utilizando YOLO v5. Concluye con una precisión del 87.5% tras un entrenamiento de 120 épocas.
- [Expósito, 2022] desarrolló un sistema de visión para vehículos autónomos usando YOLO v5. Concluye con una precisión en torno al 60%, pero agrega que con un entrenamiento más extenso se puede aumentar este valor, por lo que considera que YOLO v5 es adecuado para su uso en la conducción autónoma.

2. Métodos

Para poder llevar a cabo la identificación del color de los fragmentos de vidrio en una imagen buscando que dicha tarea pueda ser realizada de manera autónoma, rápida y precisa, se decidió emplear el aprendizaje profundo. Dicha área de aprendizaje de máquina emplea redes neuronales convolucionales para que de manera autónoma se extraigan las características necesarias para identificar el color de los fragmentos de vidrio en una imagen. Esta elección fue hecha considerando la capacidad del aprendizaje profundo que ofrece dicha técnica para extraer características de alto nivel que permitan una clasificación más precisa y eficiente de los fragmentos de vidrio por color en imágenes.

YOLO v8

Se optó por utilizar la red convolucional YOLO en su versión 8 debido a su destacada eficiencia en la detección de objetos en tiempo real. Esta red neuronal convolucional profunda fue entrenada en una computadora con GPU NVIDIA GeForce RTX 1060 utilizando 1200 imágenes con una resolución de 4160 x 1968 píxeles para el conjunto de datos total. Las muestras están conformadas de 3 clases (vidrio transparente, ámbar y verde) con 400 imágenes para cada clase, como se muestra en la figura 2.



Fuente: elaboración propia.

Figura 2 Muestras que forman parte de la base de datos.

Se eligieron dichas clases dado que el material obtenido de la empresa contiene en su mayoría estos colores de fragmentos de vidrio. Las condiciones de iluminación de las imágenes son las que presenta el ambiente donde se recicla el vidrio, que actualmente no tienen iluminación controlada. Los datos fueron divididos en una

proporción de 80% para el conjunto de entrenamiento, 10% el de validación y 10% para el de prueba, respectivamente, siguiendo las proporciones que se han empleado en trabajos anteriores descritos en la página anterior.

La red neuronal puede ser nutrida agregando más objetos a la categoría o clase, por lo que fueron empleadas fotografías que contienen una variedad de fragmentos de vidrio, con diferentes ángulos e iluminación, tal y como se muestra en la figura 3. Esto permitió el crecimiento de la base de datos.

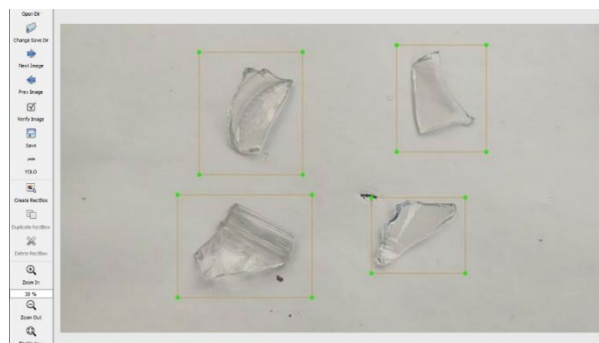


Fuente: elaboración propia.

Figura 3 Fragmentos de vidrio con diferentes formas, ángulos e iluminación.

Dado que los fragmentos de vidrio fueron obtenidos de una empresa dedicada al reciclaje de dicho material, los fragmentos provienen mayormente de botellas y focos, aunque también se incluyen fragmentos de otros objetos hechos de vidrio. Los fragmentos empleados no superan los 10 centímetros de tamaño.

Después de la recopilación de imágenes, todos los objetos pertenecientes a cada clase fueron marcados manualmente dentro de ellas. Este proceso de etiquetado manual consiste en etiquetar individualmente los objetos de interés de cada imagen y asignarles la clase correspondiente como se muestra en la figura 4.



Fuente: elaboración propia.

Figura 4 Etiquetado de objetos en imagen con el programa Labelimg.

Para cada etiquetado, el programa genera archivos de formato .txt que almacenan las coordenadas de localización del objeto dentro de la imagen y la clase a la que pertenece. Dentro del archivo, cada línea contiene información como la que se muestra a continuación: “2 0.518389 0.482978 0.173317 0.608232”. El primer dígito representa el número de clase a la que pertenece el objeto etiquetado (0-verde, 1-ámbar y 2-transparente), el segundo y tercer valores representan el alto y el ancho respecto al tamaño de la región seleccionada de la imagen, y el cuarto y quinto valores representan las coordenadas x y y desde la esquina superior, en las que se encuentra el cuadro en toda la imagen. La selección manual de objetos se llevó a cabo en las 1200 imágenes, ya que el modelo de entrenamiento solo debe extraer características del objeto de interés y no de la imagen completa. Este proceso se realizó utilizando la herramienta de Labelimg [Labelimg, 2023] que es una herramienta para etiquetar manualmente objetos en imágenes.

Selección y descripción del modelo

En este trabajo la base de datos fue entrenada por medio del modelo de Yolov8. Dicho modelo está diseñado para detectar objetos en una sola revisión de las imágenes y fue presentado en enero del 2023 por Ultralytics [Ultralytics, 2023], la misma compañía que desarrollo Yolov5 [Estrella, 2020]. Yolov8 tiene 5 versiones: YOLOv8n (nano), YOLOv8s (pequeño), YOLOv8m (mediano), YOLOv8l (largo) y YOLOv8x (extralargo). El algoritmo admite múltiples tareas de visión para detectar y segmentar simultáneamente los objetos dentro de una imagen [YOLOv5, 2023]. El modelo cuenta con capas convolucionales mejoradas con una etapa de detección mucho más avanzada, lo que la convierte en la mejor opción para la detección de objetos en tiempo real.

La versión más reciente de YOLO ofrece soporte para los últimos algoritmos de visión por computadora en la detección de múltiples objetos en una imagen. Este utiliza una red Darknet-53 más rápida y precisa que la red anterior de YOLOv7 [Wang, 2023]. La red neuronal convolucional de dicho modelo es más efectiva que las versiones anteriores debido a su mapa de características más grande y su red neuronal convolucional mejorada, incrementando la precisión y velocidad. También,

incorpora redes piramidales de características para reconocer objetos de diferentes tamaños. Este modelo ya se ha implementado en otras aplicaciones.

Etapa de entrenamiento

En el entrenamiento se utilizó la versión larga de Yolov8 (Yolov8l) para esta primera etapa. En una segunda etapa se planea entrenar la base de datos con las demás versiones para tener una comparación entre todos los modelos y seleccionar la que tenga mejor rendimiento para el objetivo establecido en el presente trabajo. El *tamaño de imagen* que entra a la primera capa convolución es escalado a 640x640 que es la medida que maneja por default por el modelo, por lo que las imágenes fueron redimensionadas. El *device* es igual a 0, esto permitió utilizar los GPU que tiene la tarjeta NVIDIA GeForce RTX 1060 de la computadora y acelerar el entrenamiento. La RAM de la computadora que entrenó este modelo es de gama baja, por lo que los lotes de imágenes (*Batch*) tenían que ser pocos (igual a 2) y los *Workers* igual a 0 por no tener la posibilidad de trabajar en paralelo. Finalmente, el número de *épocas* de entrenamiento fue de 400 para esta primera etapa. El modelo de Yolov8 fue entrenado con los hiperparámetros de la tabla 1, de los cuales el tamaño de la imagen es el que por defecto usa el modelo, y el número de clases, el batch y las épocas fueron optimizados. El entrenamiento fue realizado con Anaconda y utilizando el lenguaje de programación Python.

Tabla 1 Hiperparámetros utilizados en el entrenamiento.

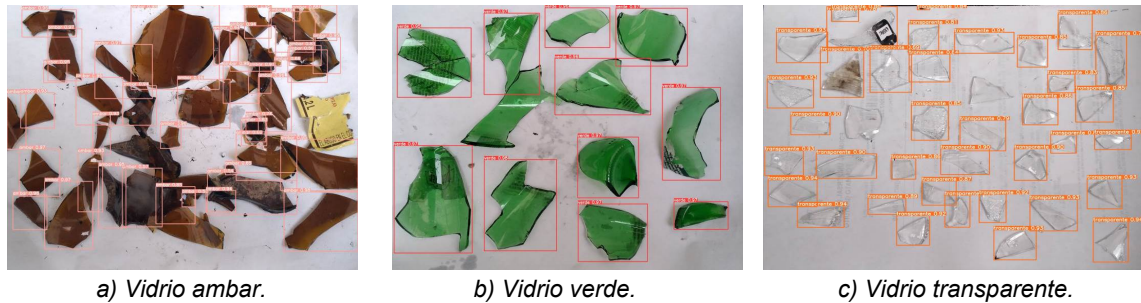
Hiperparámetro	Dato
Versión	Yolov8l.pt
Número de clases	3
Tamaño de imagen	640x640
Device	0
Batch	2
Workers	0
Número de épocas	400

Fuente: elaboración propia.

3. Resultados

Dentro del conjunto de imágenes de prueba y validación se incluyeron aproximadamente 121 de ellas, conteniendo 5 fragmentos o más de vidrio. El

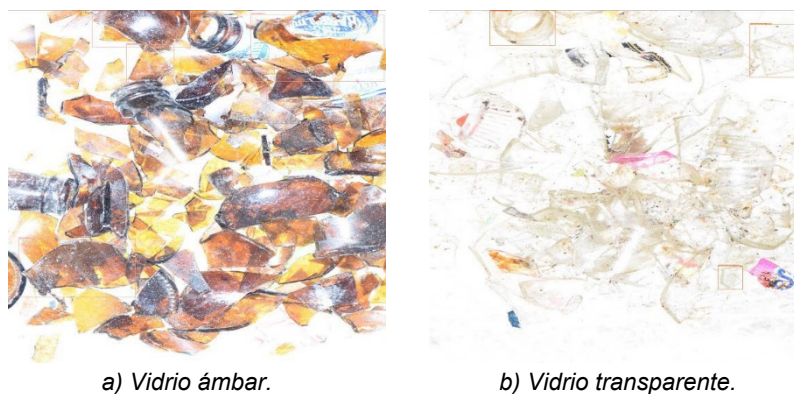
modelo entrenado debe ser capaz de poder detectar y reconocer vidrios de color verde, ámbar y transparente, a partir de las muestras de prueba. Las detecciones del modelo entrenado en las imágenes de prueba son mostradas en la figura 5.



Fuente: elaboración propia.

Figura 5 Resultados de detección aplicando el modelo en diferentes tipos de vidrio.

Los cuadros delimitadores alrededor de los objetos encontrados muestran el ID de la clase a la que pertenece, así como el porcentaje de probabilidad de que pertenezca a dicha clase. Este modelo fue probado en imágenes con alta iluminación y ruido (imágenes no integradas en la base de entrenamiento). Los resultados para el vidrio ámbar y transparente son mostrados en la figura 6. En ella se observa que el algoritmo no es capaz de detectar todos los objetos en la imagen debido a la gran cantidad de iluminación, sin embargo, clasifica correctamente a la clase que pertenece el objeto detectado. En condiciones óptimas el modelo es capaz de identificar y reconocer los fragmentos de vidrio mostrados en la figura 5.



Fuente: elaboración propia.

Figura 6 Detección del modelo en imágenes con alta iluminación.

Para evaluar el rendimiento del modelo utilizando el conjunto de datos de prueba se utiliza la métrica mAP50, la cual es la precisión media que se utiliza para medir el rendimiento de los modelos de detección de objetos. Este valor se obtiene a partir de la relación que hay entre la precisión y la sensibilidad del modelo con un umbral de 0.5 [Terven, 2023]. La precisión mide el porcentaje de predicciones positivas del modelo, mientras que la exhaustividad o sensibilidad (Recall) mide la proporción de casos positivos reales que el modelo identifica correctamente. Es importante destacar que la clasificación es multiclase y no binaria, por lo que las ecuaciones 1 y 2, son las necesarias para obtener dichas métricas [Markoulidakis, 2021].

$$TPR(C_i) = \frac{TP(C_i)}{TP(C_i) + FN(C_i)} \quad (1)$$

$$PPV(C_i) = \frac{TP(C_i)}{TP(C_i) + FP(C_i)} \quad (2)$$

Donde:

$TP(C_i)$: Verdaderos positivos de la categoría i

$FP(C_i)$: Falsos positivos de la categoría i

$FN(C_i)$: Falsos negativos de la categoría i

$TPR(C_i)$: Sensibilidad de la categoría i

$PPV(C_i)$: Precisión de la categoría i .

Los verdaderos positivos $TP(C_i)$ son la cantidad de objetos que el modelo pudo clasificar correctamente de esa categoría, mientras que los falsos positivos $FP(C_i)$ son los objetos que no detecto correctamente. Los falsos negativos $FN(C_i)$ son detecciones que el modelo realizó, pero que no existía algún objeto de alguna categoría. Las métricas mencionadas se obtienen por cada una de las categorías, debido a que se trata de una clasificación multiclase. Finalmente, se obtiene la sensibilidad (Ecuación 1) y la precisión por categoría (Ecuación 2), obteniendo un total de 6 métricas utilizando las 3 categorías. Al igual que las métricas previas, en relación con la clasificación multiclase, cada categoría obtuvo un mAP50.

En el conjunto de datos de prueba, el modelo Yolov8 obtuvo una puntuación promedio de las tres categorías de mAP50 del 0.863%, obteniendo una velocidad

de inferencia de 37.8 ms por imagen de prueba. Esta velocidad de inferencia podría ser adecuada para procesamiento en tiempo real, por ejemplo, en video.

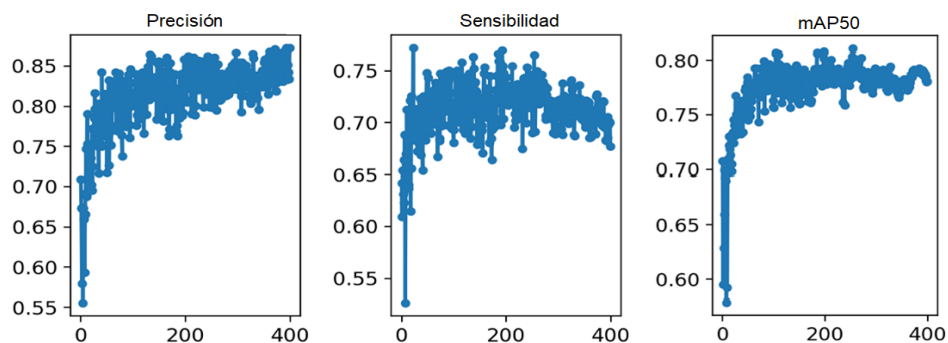
La tabla 2 muestra los resultados obtenidos al aplicar el algoritmo al conjunto de imágenes de prueba donde la clase transparente y ámbar tienen una precisión media y sensibilidad cercanas o superiores al 90%; sin embargo, la clase de los verdes obtuvo un porcentaje más bajo, probablemente debido a que dentro de la base de datos se encontraban imágenes desenfocadas o incorrectamente etiquetadas en el Labelimg. Esto provocó que el rendimiento del algoritmo llegara aproximadamente al 88%.

Tabla 2 Resultados por categoría para imágenes de prueba y el promedio de las 3 clases.

Clase	Precisión	Sensibilidad (Recall)	mAP50
Transparente	0.979	0.928	0.961
Ámbar	0.927	0.873	0.929
Verde	0.703	0.649	0.699
Promedio	0.87	0.817	0.863

Fuente: elaboración propia.

La figura 7 presenta 3 gráficas, de precisión, de sensibilidad y del mAP50 total de las tres clases durante el entrenamiento de las 400 épocas. La mejor precisión se obtuvo alrededor de la época 350, mientras que la mejor sensibilidad se obtuvo en la época 300. Finalmente, el mejor mAP50 se obtuvo en la época 310. Aumentado el número de épocas se puede aumentar la precisión y la sensibilidad del modelo, dando como resultado un mejor mAP50. Sin embargo, para esta etapa del trabajo se llegó hasta la época 400 con resultados cercanos al 90%.



Fuente: elaboración propia.

Figura 7 Resultados del entrenamiento.

4. Discusión

La clasificación de vidrio para reciclaje es una tarea de importancia para reducir el impacto de contaminación en el medio ambiente. La automatización del proceso es crucial para poder mejorar la calidad y velocidad en la que el vidrio es separado del resto de desperdicio. La implementación de aprendizaje profundo mediante el algoritmo YOLOv8 para la tarea de entrenamiento e identificación de fragmentos de vidrio, ha mostrado resultados prometedores para el desarrollo de sistemas automatizados para la clasificación y separación de vidrio para reciclaje.

En el presente trabajo se muestran resultados con un grado de precisión cercano al 90% en la detección de fragmentos de vidrio de color transparente y ámbar, y de más del 80% para el caso del vidrio color azul.

Como trabajo futuro se pueden aumentar las épocas evitando el sobreajuste, con lo que se espera que el nivel de precisión en la detección del vidrio aumente. Por otro lado, dada la rapidez del algoritmo, su implementación para el procesamiento de video en tiempo real haría factible el desarrollo de sistemas para la detección y clasificación de vidrio en tiempo real para sistemas de reciclaje de dicho material.

Actualmente muchos de los sistemas de separación de vidrio requieren la separación previa del mismo en colores, por lo que parte del trabajo futuro puede incluir pruebas de funcionamiento y precisión del algoritmo para vidrio de cualquier color.

Al momento de someter este trabajo, no se tiene conocimiento de qué el modelo haya sido implementado para la identificación de fragmentos de vidrio para su reciclaje.

5. Conclusiones

Este artículo ofrece una descripción del proceso de recolección de datos, así como del entrenamiento de la red convolucional YOLO en su versión 8 y las pruebas subsecuentes hechas al modelo. Para los diferentes colores de vidrio (verde, ámbar y transparente), la detección de fragmentos de dicho material después de entrenar al modelo alcanzó una precisión cercana al 90%, lo que significa que ha proporcionado buenos resultados para su implementación en la tarea de detección

de fragmentos de virio. La implementación de dicha funcionalidad en sistemas para la detección y separación de vidrio para reciclaje podría resultar en una mejora sustancial sobre los sistemas de reciclaje de vidrio que han sido implementados hasta el día de hoy.

Dada la rapidez en las pruebas del modelo, este podría ser implementado en el procesamiento de vidrio en tiempo real por medio de video (el cual es una sucesión de imágenes).

YOLO proporcionó buenos resultados de inicio, por lo que provee de un sistema óptimo para ser implementado en una planta de reciclaje tras un incremento en la duración del entrenamiento y la creación de clases adicionales.

Para futuros trabajos se pueden incrementar las clases del modelo, así como mejorar el entrenamiento procurando un mayor número de épocas durante dicho proceso.

6. Bibliografía y Referencias

- [1] Alcaraz-Londoño, L., et. al. Review on techniques of automatic solid waste separation in domestic applications. *Bulletin of Electrical Engineering and Informatics*. ISSN 2302-9285. Vol 11. No.1 128-133 <http://dx.doi.org/10.11591/eei.v11i1.3448>, 2022.
- [2] Bohn, B., et. al. A novel method to produce a ceramic paver recycling waste glass. *Cleaner Engineering and Technology*. Vol. 2. ISSN 2666-7908. <https://doi.org/10.1016/j.clet.2021.100043>, 2021.
- [3] Bristogianni, T., et. al. Glass up-casting: a review on the current challenges in glass recycling and a novel approach for recycling “as-is” glass waste into volumetric glass components. *Glass Structures & Engineering*. Springer. 10.1007/s40940-022-00206-9, 2022.
- [4] Capistran, O. Sistema de Aprendizaje Automático para Detección y Análisis de Tráfico Vehicular. [Tesis, Centro de Ingeniería y Desarrollo Industrial]. <http://cidesi.repositorioinstitucional.mx/jspui/handle/1024/448>, 2020.
- [5] Centeno, A. (2019). Deep Learning. (Trabajo Fin de Doble Grado) Universidad de Sevilla, Sevilla. <https://hdl.handle.net/11441/90004>, 2019.

- [6] Estrella, A. Aplicaciones basadas en aprendizaje automático (Machine Learning) en plataformas de bajo consumo. [Proyecto Fin de Grado, Escuela Técnica Superior de Ingeniería y Sistemas de Telecomunicación]. <https://oa.upm.es/66520/>, 2020.
- [7] Expósito, M. Desarrollo de aplicación basada en CNN para algoritmos de visión en coches autónomos. [Trabajo Fin de Grado, Universidad Complutense de Madrid]. <https://hdl.handle.net/20.500.14352/10479>. 2022.
- [8] Fu, Y. Image classification via fine-tuning with EfficientNet. Keras. https://keras.io/examples/vision/image_classification_efficientnet_fine_tuning. Actualizado en 2023/07/10, 2023.
- [9] Giraldo, J. Clasificador de Residuos Sólidos Haciendo Uso de Deep Learning. [Proyecto Fin de Grado, Universidad Autónoma de Occidente]. <https://hdl.handle.net/10614/14209>, 2022.
- [10] He, Hao & Li, Xiangtai & Cheng, Guangliang & Shi, Jianping & Tong, Yunhai & Meng, Gaofeng & Prinet, Véronique & Weng, LuBin. Enhanced Boundary Learning for Glass-like Object Segmentation. 15839-15848. 10.1109/ICCV48922.2021.01556, 2021.
- [11] Janiesch, C. & Zschech, P. & Heinrich, K. Machine learning and deep learning. *Electronic Markets*. Vol. 31. 1422-8890. 10.1007/s12525-021-00475-2, 2021.
- [12] Kaspi, Omer & Israelsohn-Azulay, Osnat & Yigal, Zidon & Rosengarten, Hila & Krmpotić, Matea & Gouasmia, Sabrina & Radovic, Iva & Jalkanen, Pasi & Liski, Anna & Mizohata, Kenichiro & Räisänen, Jyrki & Kasztovszky, Zs & Harsányi, Ildikó & Acharya, Raghunath & Pujari, P. & Mihály, Molnár & Braun, Mihaly & Shabi, Nahum & Girshevitz, Olga & Senderowitz, Hanoach. Toward Developing Techniques—Agnostic Machine Learning Classification Models for Forensically Relevant Glass Fragments. *Journal of Chemical Information and Modeling*. 10.1021/acs.jcim.2c01362, 2022.
- [13] LabelImg.Heartexlabs/labelImg:<https://github.com/heartexlabs/labelImg>, 2023.
- [14] López, J., et. al. Sistema de Detección de Control de Estudiantes Universitarios Basado en Machine Learning e Inteligencia Artificial. Perú. ISSN 2414-6390. <http://dx.doi.org/10.18687/LEIRD2022.1.1.178>, 2022.

- [15] Markoulidakis, J., et. al. Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. Vol. 9, pp. 81. 10.3390/technologies9040081, 2021.
- [16] Miranda, R., et. al. Introducción al Aprendizaje Automático con YOLO. Instituto Tecnológico de Costa Rica. Revista de la Facultad de Ingenierías y Tecnologías de la Información y Comunicación. Vol. 2 No. 6, 52-58. ISSN 1659-0775, 2019.
- [17] Ponce, P. Inteligencia Artificial con Aplicaciones a la Ingeniería. Alfaomega. ISBN: 978-607-7854-83-8. México, 2010.
- [18] Schlumberger. Energy Glossary en español. Convolución. <https://glossary.slb.com/es/terms/c/convolution>, (s.f.).
- [19] Solawetz, J. What is YOLOv8? The ultimate Guide. <https://blog.roboflow.com/whats-new-in-yolov8/>, 2023.
- [20] Tao, Yong & Gao, He & Wen, Yufang & Duan, Lian & Lan, Jiangbo. Glass Recognition and Map Optimization Method for Mobile Robot Based on Boundary Guidance. Chinese Journal of Mechanical Engineering. 36. 10.1186/s10033-023-00902-9, 2023.
- [21] Terven, J. et. al. A Comprehensive Review of YOLO: From YOLOv1 and Beyond. Sometido a ACM Computing Surveys. <https://arxiv.org/abs/2304.00501>, 2023.
- [22] YOLOv5, (2023). YOLO Version 5 <https://github.com/ultralytics/yolov5/releases>
- [23] Ultralytics. NEW - YOLOv8. GitHub. <https://github.com/ultralytics/ultralytics>, 2023.
- [24] Wang, C., et. al. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.
- [25] Zhou, X. & Liu, X. The Detection and Recognition Algorithm of Safety Glass Fragment. International Conference on Mechatronics and Automation, Harbin, China, pp. 963-967, doi: 10.1109/ICMA.2007.4303677, 2007.