

Clasificación automatizada de residuos sólidos mediante visión por computador y el modelo YOLOv8 para gestión ambiental

Automated solid waste classification through computer vision and the YOLOv8 model for environmental management

Thomás Ricardo Borja Saltos ^{1*}, Claudio Xavier Borja Saltos ², Jorge Wilson Tamami Pachala ³, Cristhian Geovanny Villamarin Arroba ⁴

¹ Laboratorio de Investigación en Inteligencia y Visión Artificial, Departamento de Informática y Ciencias de la Computación, Escuela Politécnica Nacional, Quito 170525, Ecuador. ROR: <https://ror.org/01gb99w41>

² Departamento de Desarrollo de Software. SOTTECAPPS S.A.S. Guaranda 020101, Ecuador.

³ Facultad de Ciencias de la Educación, Universidad Estatal de Bolívar. Guaranda 020150, Ecuador. ROR: <https://ror.org/005cgg117>

⁴ Unidad de Posgrados. Universidad Estatal de Milagro. Milagro 091050, Ecuador. ROR: <https://ror.org/0ogd7ns03>

✉ thomas.borja@epn.edu.ec

| ORCID: <https://orcid.org/0000-0002-8693-6111>

✉ claudio-borja@hotmail.com

| ORCID: <https://orcid.org/0009-0008-6938-9399>

✉ jtamami@ueb.edu.ec

| ORCID: <https://orcid.org/0000-0002-3470-7894>

✉ cwillamarina@unemi.edu.ec

| ORCID: <https://orcid.org/0009-0000-1905-4015>

E-mail de correspondencia: thomas.borja@epn.edu.ec

Serie Monográfica

Mercado, Tecnología y Ciudadanía.

e-ISSN: 3103-117X

Vol. 2(1) enero - abril 2026

Desafíos de la sociedad contemporánea

ISBN: 978-9942-7391-7-9

Editor académico

Carlos Vásquez

Universidad Técnica de Ambato. Ecuador.

Tipo de revisión

Capítulo de libro revisado por dos pares expertos en modalidad doble ciego.

Como citar este capítulo

Borja Saltos, T. R., Borja Saltos, C. X., Tamami Pachala, J. W., & Villamarin Arroba, C. G. (2026). Clasificación automatizada de residuos sólidos mediante visión por computador y el modelo YOLOv8 para gestión ambiental. En *Serie Monográfica Mercado, Tecnología y Ciudadanía* (Vol. 2, Núm. 1, Cap. I, e1). <https://doi.org/10.63804/mtc.v2i1.e1>

Copyright

© 2026 Los autores. Este es un capítulo de libro de acceso abierto distribuido bajo los términos y condiciones de la licencia Creative Commons Attribution 4.0 International (CC BY-NC-SA 4.0). Se autoriza el uso, distribución y reproducción de este contenido en cualquier medio, de forma irrestricta, siempre que se otorgue el crédito a los autores originales y se cite debidamente la fuente primaria de publicación.

Recibido: 11 de noviembre de 2025

Revisado: 18 de diciembre de 2025

Aceptado: 10 de enero de 2026

Publicado: 28 de febrero de 2026

Resumen

La gestión sostenible de residuos sólidos urbanos enfrenta desafíos importantes debido a la escasa eficiencia de los procesos de clasificación manual y la baja tasa de recuperación de materiales reciclables. Este estudio tuvo como objetivo evaluar la efectividad de un modelo de visión artificial basado en la arquitectura You Only Look Once (YOLOv8) para la identificación y clasificación automatizada de residuos. La metodología comprendió el procesamiento de una versión optimizada del conjunto de datos TrashNet, conformada por 2.390 imágenes tras un riguroso proceso de depuración y balanceo. El dataset se categorizó en cinco clases fundamentales: cartón, papel, plástico, metal y vidrio. Se emplearon técnicas de transferencia de aprendizaje (*transfer learning*) y segmentación estratificada para garantizar la robustez del entrenamiento en un entorno de computación de alto rendimiento. Los resultados experimentales demostraron un desempeño superior, alcanzando una exactitud global de 0,94 en el conjunto de validación. Las métricas de precisión, sensibilidad y F1-score evidenciaron una alta consistencia interclase, manteniendo la estabilidad operativa frente a variaciones de iluminación y ruido visual. En la fase de discusión e impacto, se validó la aplicabilidad tecnológica mediante el despliegue del modelo en aplicaciones funcionales de escritorio y entorno web. Estas herramientas permitieron el procesamiento en tiempo real sin dependencia de servidores externos, lo que facilitó la descentralización del monitoreo ambiental. Se concluye que la integración de redes neuronales convolucionales de una sola etapa representa una solución escalable y de bajo costo computacional, con un potencial significativo para transformar los sistemas de economía circular y optimizar la recuperación de recursos en contextos urbanos e industriales.

Palabras clave: aprendizaje profundo; conservación de recursos; gestión de desechos; inteligencia artificial; tecnología ambiental.

ABSTRACT

Sustainable urban solid waste management faces significant challenges due to the low efficiency of manual classification processes and the limited recovery rate of recyclable materials. This study aimed to evaluate the effectiveness of a computer vision model based on the You Only Look Once (YOLOv8) architecture for the automated identification and classification of waste. The methodology involved processing an optimized version of the TrashNet dataset, consisting of 2,390 images after a rigorous cleaning and balancing procedure. The dataset was categorized into five fundamental classes: cardboard, paper, plastic, metal, and glass. Transfer learning techniques and stratified segmentation were employed to ensure robust training within a high-performance computing environment. Experimental results demonstrated superior performance, achieving an overall validation accuracy of 0.94. Precision, sensitivity, and F1-score metrics revealed high inter-class consistency, maintaining operational stability under variations in lighting and visual noise. In the discussion and impact phase, technological applicability was validated through the deployment of the trained model in functional desktop and web-based applications. These tools enabled real-time processing without reliance on external servers, facilitating the decentralization of environmental monitoring. It is concluded that the integration of single-stage convolutional neural networks represents a scalable and computationally efficient solution, with significant potential to transform circular economy systems and optimize resource recovery in urban and industrial contexts.

Keywords: deep learning; resource conservation; waste management; artificial intelligence; environmental technology.

INTRODUCCIÓN

La generación creciente de residuos sólidos representa uno de los principales desafíos ambientales contemporáneos, con efectos directos sobre la calidad del suelo, cuerpos de agua, salud pública y eficiencia de los sistemas de recolección y tratamiento. En contextos urbanos, el aumento del consumo y la disposición inadecuada intensifican la presión sobre la infraestructura disponible, mientras que la recuperación de materiales reciclables se ve limitada por fallas en la separación en origen y por procesos de clasificación con baja estandarización.

En consecuencia, la clasificación correcta de residuos se reconoce como una etapa determinante para optimizar el tratamiento de desechos, reducir pérdidas de recursos y favorecer prácticas alineadas con el desarrollo sostenible [1]. A pesar de su importancia, la clasificación manual continúa predominando en numerosos escenarios, especialmente en ámbitos comunitarios, educativos e industriales con recursos limitados.

Este enfoque presenta restricciones vinculadas con la variabilidad operativa, la dependencia de la capacitación de los usuarios, fatiga y subjetividad en la toma de decisiones, lo que incrementa la probabilidad de errores y reduce la eficiencia del proceso. La heterogeneidad de los residuos y las condiciones cambiantes de iluminación, fondo y oclusión dificultan la implementación de procedimientos consistentes y reproducibles, lo que plantea la necesidad de soluciones tecnológicas que mejoren la separación y el aprovechamiento de materiales [2].

En los últimos años, la visión por computador se ha empleado como una alternativa útil para automatizar tareas de reconocimiento visual mediante el análisis de imágenes, lo que permite identificar patrones de forma, textura y color asociados a diferentes materiales. En particular, el aprendizaje profundo ha demostrado avances significativos en problemas de clasificación multiclase al aprender representaciones discriminativas directamente a partir de datos, superando enfoques tradicionales basados en identificación manual [3].

En el dominio de residuos sólidos, diversos estudios han reportado resultados favorables al emplear redes neuronales convolucionales entrenadas con conjuntos de datos estandarizados, lo que respalda la viabilidad técnica de implementar sistemas automáticos de preclasificación y clasificación asistida [4], [5].

Dentro de este campo, la arquitectura *You Only Look Once* (YOLO), versión 8, se utiliza ampliamente por su eficiencia computacional y su capacidad para operar con baja latencia, características relevantes cuando se requiere un desempeño adecuado en aplicaciones prácticas.

En tareas de clasificación de imágenes, variantes ligeras de esta arquitectura ofrecen un equilibrio entre exactitud y consumo de recursos, esto favorece su despliegue en equipos sin unidades de procesamiento especializado y en entornos de ejecución orientados a usuario final, tales como aplicaciones de escritorio o soluciones basadas en navegador [6]. No obstante, la adopción de estos modelos exige evaluaciones experimentales reproducibles que integren preparación de datos, entrenamiento controlado y métricas de desempeño interpretables para sustentar su robustez y utilidad en condiciones realistas [7].

La evaluación del rendimiento debe considerar por una parte la exactitud global, y por otra, las medidas por clase que reflejen el comportamiento del modelo ante errores de confusión entre

categorías, así como métricas que permitan analizar sensibilidad, especificidad y equilibrio entre aciertos y fallos. En virtud de lo anterior, el presente estudio tuvo como objetivo evaluar experimentalmente la efectividad de un modelo de visión por computador basado en la arquitectura *You Only Look Once*, versión 8, para la clasificación automatizada de residuos sólidos.

METODOLOGÍA

Enfoque, alcance y diseño del estudio

El estudio se desarrolló bajo un enfoque cuantitativo, de alcance aplicado y con diseño experimental orientado a la evaluación del rendimiento de un modelo de clasificación de imágenes para residuos sólidos. La finalidad consistió en estimar, de manera reproducible, el desempeño de un modelo de visión por computador en la identificación automática de residuos: cartón, papel, plástico, metal y vidrio. Para garantizar objetividad, la evaluación se estructuró mediante un banco de pruebas independiente del entrenamiento, utilizando métricas estandarizadas para clasificación multiclase y procedimientos de validación que reducen el riesgo de sesgo por sobreajuste [8], en la Tabla 1 se muestran los detalles de los parámetros de validación.

Tabla 1. Matriz de diseño experimental y parámetros de validación para el modelo de clasificación de residuos.

Componente del diseño	Especificación técnica
Enfoque y alcance	Cuantitativo, de carácter aplicado con diseño experimental funcional.
Unidad de análisis	Imágenes multiclase de residuos sólidos (N = 2.390).
Variable independiente	Categoría de material (cartón, papel, plástico, metal y vidrio).
Variables dependientes	Métricas de desempeño: mAP, precisión, recall, especificidad y F1-score.
Arquitectura del modelo	Red neuronal convolucional <i>You Only Look Once</i> (YOLOv8, Ultralytics Inc., Frederick, MD, EE. UU.).
Partición del dataset	Distribución estratificada: 80 % entrenamiento, 10 % validación y 10 % test.
Umbral de aceptabilidad	Rendimiento predictivo ≥ 90 % en exactitud global y medida F1.

Nota. mAP (mean Average Precision): promedio de la precisión media que evalúa la exactitud del modelo en la detección y localización de objetos; Recall (Sensibilidad): proporción de casos positivos reales identificados correctamente por el algoritmo; F1-score: medida armónica que equilibra la precisión y la sensibilidad, proporciona una métrica robusta de la exactitud del modelo.

Conjunto de datos y preparación

El conjunto de datos empleado en esta investigación corresponde a una versión adaptada del dataset *TrashNet*, originalmente propuesto en el Massachusetts Institute of Technology como un recurso abierto para la clasificación automática de residuos [9]. En su forma original, contiene imágenes de seis categorías (cartón, papel, plástico, metal, vidrio y basura general) capturadas bajo condiciones controladas de iluminación y fondo uniforme. Para el presente estudio se efectuó un proceso de depuración y reorganización destinado a mejorar la coherencia visual y la calidad de las clases utilizadas en el entrenamiento del modelo.

El conjunto fue reducido a cinco categorías cartón, papel, plástico, metal y vidrio, eliminando la clase “basura general” debido a su ambigüedad semántica y a la heterogeneidad de sus muestras. Posteriormente, las imágenes fueron renombradas y redistribuidas en carpetas individuales por clase, asegurando uniformidad en la nomenclatura y trazabilidad durante el proceso de carga del modelo, se descartaron las imágenes duplicadas o con baja resolución, lo que resultó en un conjunto final compuesto por 2.390 imágenes válidas (Tabla 2).

Tabla 2. Composición y distribución de las imágenes seleccionadas por categoría de residuo y subconjuntos de datos.

Clase	Entrenamiento	Validación	Cantidad de imágenes	Distribución de imágenes, %
Cartón	323	80	403	16,86
Papel	476	118	594	24,85
Plástico	386	96	482	20,17
Metal	328	82	410	17,15
Vidrio	401	100	501	20,96
Total general	1.914	476	2.390	100,00

Para garantizar la representatividad de los datos, se aplicó un muestreo estratificado, dividiendo el conjunto en dos subconjuntos principales: entrenamiento y validación, que contienen 1.914 y 476 imágenes respectivamente, lo que equivale a una proporción aproximada de 80 % y 20 %.

El proceso de organización de datos fue automatizado mediante un script desarrollado en el lenguaje de programación Python (versión 3.10, Python Software Foundation, Wilmington, DE, EE. UU.). Se emplearon las bibliotecas estándar `os` y `shutil` para la gestión programática de directorios y la transferencia de archivos. La partición de las muestras se realizó mediante una asignación aleatoria controlada, lo que permitió preservar la proporción de clases y garantizar la integridad estructural del conjunto de datos original TrashNet. Este procedimiento aseguró una distribución equilibrada de las categorías analizadas (cartón, papel, plástico, metal y vidrio), facilitando la reproducibilidad del experimento. El esquema general del diseño del estudio, incluyendo las variables y el umbral de aceptabilidad técnica, se detalla en la Tabla 1.

El esquema de partición de datos aplicado en este estudio es coherente con prácticas reportadas en investigaciones recientes que emplean TrashNet o conjuntos derivados como base para el entrenamiento y evaluación de modelos de clasificación de residuos. En particular, diversos trabajos documentan la conveniencia de utilizar este tipo de conjuntos estandarizados para evaluar clasificadores multiclase y compararlos bajo métricas consistentes, destacando su compatibilidad con arquitecturas contemporáneas basadas en YOLO y EfficientNet [10], [11].

Arquitectura y entorno de entrenamiento

Se seleccionó una variante ligera de la arquitectura You Only Look Once, versión 8 (Ultralytics Inc., Frederick, MD, EE. UU.), especializada en clasificación, debido a su equilibrio entre exactitud y eficiencia computacional, lo que facilita su entrenamiento y despliegue en equipos sin aceleración por unidad de procesamiento gráfico [12].

El entrenamiento se desarrolló bajo un esquema de transferencia de aprendizaje, utilizando pesos preentrenados como punto de partida para acelerar la convergencia y mejorar la generalización en el dominio de residuos sólidos. La configuración detallada de los hiperparámetros, seleccionada para optimizar la convergencia y evitar el sobreajuste, se presenta en la Tabla 3. El modelo se configuró con los hiperparámetros, definidos en el archivo `runs/classify/basura_cls_run/args.yaml` y aplicados automáticamente durante la ejecución.

Tabla 3. Configuración de hiperparámetros y entorno de ejecución para el entrenamiento del modelo.

Parámetro técnico	Valor asignado
Arquitectura de red	YOLOv8n-cls (clasificador de imágenes, preentrenado en ImageNet)
Épocas de entrenamiento	25
Tamaño de lote (batch size)	32
Resolución de entrada	224×224 píxeles
Tasa de aprendizaje inicial (η)	0,001
Optimizador	AdamW (optimizer='auto' en YOLOv8; seleccionado automáticamente en este experimento)
Función de pérdida	Cross-Entropy para clasificación de una sola etiqueta
Dispositivo de cómputo (hardware)	CPU sin aceleración por GPU

Nota. Los valores fueron seleccionados mediante un proceso de ajuste empírico (*fine-tuning*) basado en múltiples iteraciones de entrenamiento para maximizar la convergencia del modelo.

Durante la ejecución, el script verificó la existencia de las rutas *datos/preparados/{train,val}*, los pesos pre-entrenados y los directorios de salida mediante una validación previa de dependencias. Los resultados del entrenamiento se centralizaron en la ruta *runs/classify/basura_cls_run/*, donde se generaron automáticamente registros tabulares *results.csv*, *results.png*, *confusion_matrix.png*, *confusion_matrix_normalized.png* y los mosaicos visuales de entrenamiento y validación *train_batch*.jpg*, *val_batch*_labels.jpg*, *val_batch*_pred.jpg*, que se pueden verificar en el repositorio.

Al finalizar, el modelo entrenado se guardó como *modelo_entrenado.pt* en el directorio *modelo/entrenado/*, mientras que las métricas se consolidaron en *modelo/métricas/*. La evaluación final se realizó sobre el subconjunto de validación independiente utilizando la función de inferencia *predict()* de YOLO, considerando todas las imágenes de las subcarpetas *val/*. Las métricas se calcularon a partir de las predicciones y etiquetas reales, obteniendo precisión global (accuracy), precisión por clase (precision), sensibilidad (recall), F1-score y matriz de confusión, de acuerdo con el marco metodológico propuesto por Powers [13].

Evaluación del desempeño

La evaluación del desempeño del clasificador se realizó sobre el subconjunto de validación generado durante la preparación del conjunto de imágenes, con el fin de estimar la capacidad del modelo para generalizar a muestras no utilizadas en el entrenamiento. Para ello, se ejecutó el flujo de evaluación implementado en el proyecto, el cual obtiene las predicciones del modelo para todas las imágenes del directorio de validación y las contrasta con sus etiquetas reales, preservadas por la organización por carpetas de clase.

El desempeño se cuantificó mediante exactitud global y métricas por clase obtenidas a partir de un reporte de clasificación multiclase, incluyendo precisión predictiva, sensibilidad y medida armónica entre precisión predictiva y sensibilidad. Estas métricas se calcularon de forma automática utilizando funciones estándar de bibliotecas de aprendizaje automático, con el propósito de asegurar consistencia en el cálculo y trazabilidad de resultados.

Como evidencia de reproducibilidad, el proceso de evaluación conserva artefactos de salida asociados al entrenamiento y a la validación, incluyendo:

- Reporte de métricas en formato de texto.
- Matriz de confusión en formato de imagen.
- Los registros del entrenamiento en el directorio de ejecución del experimento.

RESULTADOS

Entrenamiento y convergencia del modelo

El entrenamiento se ejecutó durante 25 épocas, registrando en cada una la pérdida en entrenamiento y validación, la exactitud de validación y la tasa de aprendizaje. La Figura 1 sintetiza la evolución conjunta de la pérdida medida como entropía cruzada, una métrica que cuantifica qué tan lejos están las probabilidades predichas por el modelo de las etiquetas verdaderas y que disminuye a medida que el modelo acierta con mayor seguridad, y de la exactitud, lo que permite evaluar la dinámica de convergencia y la estabilidad del proceso de entrenamiento.

Durante las primeras épocas se observó una mejora acelerada del desempeño: la pérdida de entrenamiento disminuyó de 1,27131 (época 1) a 0,21107 (época 7), mientras la exactitud de validación aumentó de 0,75 a 0,90966 en el mismo intervalo.

A partir de la época 9 el modelo ingresó a un régimen de aprendizaje estable, con exactitud de validación $\geq 0,91807$ y valores recurrentes cercanos a 0,93 - 0,94. En particular, desde la época 15 se evidenció un comportamiento de meseta: la exactitud de validación osciló dentro de un rango estrecho (0,92017 - 0,94118), mientras la pérdida de validación se mantuvo acotada (aprox. 0,20155 - 0,26557), lo cual es consistente con un proceso de convergencia sin inestabilidad marcada.

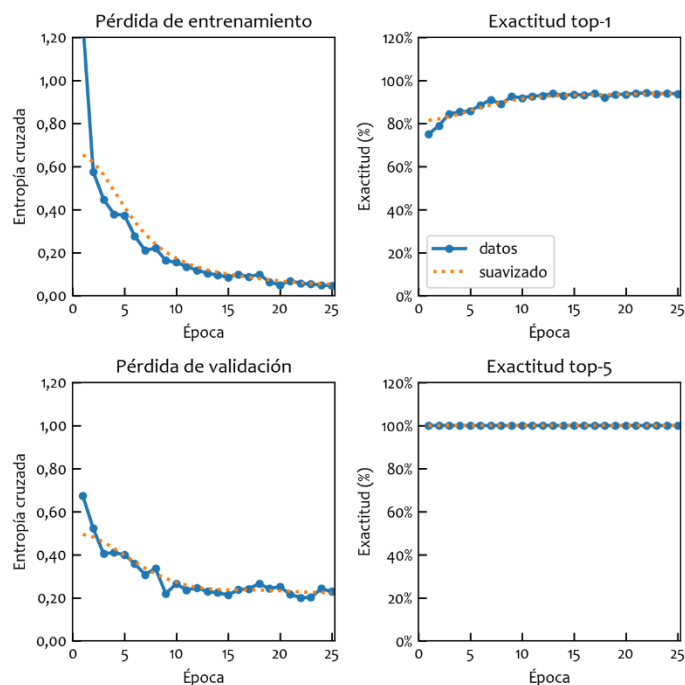


Figura 1. Evolución de la pérdida y la exactitud de validación durante el entrenamiento.

Nota. Figura elaborada a partir del archivo de registro results.csv, con tasas de aprendizaje decrecientes según la configuración del experimento.

En términos puntuales, el mejor desempeño registrado en validación correspondió a una exactitud de 0,94118 y una pérdida de validación de 0,20155 en la época 22, con valores posteriores similares. Paralelamente, la pérdida de entrenamiento continuó descendiendo hasta 0,04698 al final del entrenamiento (época 25), con fluctuaciones menores intermedias, lo que sugiere refinamiento del ajuste sin un deterioro progresivo del desempeño de validación.

Adicionalmente, la reducción gradual de la tasa de aprendizaje (de 0,000364161 en la época 1 a 0,0000551056 en la época 25) contribuye a explicar la estabilización observada en la fase final, al favorecer actualizaciones más pequeñas y un ajuste fino del modelo. En conjunto, la evidencia mostrada en la Figura 1 respalda que el proceso alcanzó convergencia operativa a partir de la época 15 y mantuvo un desempeño estable en validación, lo cual fundamenta la presentación posterior de resultados globales y por clase en las subsecciones siguientes.

Desempeño global en validación

El desempeño global del sistema se estimó mediante la exactitud global obtenida sobre el subconjunto de validación, compuesto por 476 imágenes.

En este marco, se reporta una exactitud global de 0,94, lo que implica que el modelo asignó correctamente la categoría en aproximadamente el 94 % de los casos evaluados en dicho subconjunto.

Este valor resulta relevante para una tarea de clasificación multiclase con las cinco categorías seleccionadas (cartón, papel, plástico, metal y vidrio), debido a que resume la proporción de aciertos del modelo frente a todas las muestras evaluadas y constituye una medida global directa del desempeño del sistema.

No obstante, dado que la exactitud es un indicador agregado [14], su interpretación debe complementarse con el análisis por clase y con la matriz de confusión, ya que un mismo valor global puede coexistir con diferencias de rendimiento entre categorías y con patrones específicos de confusión entre materiales. En consecuencia, el valor 0,94 coincide con el comportamiento encontrado en las épocas finales del entrenamiento, donde la exactitud de validación se estabiliza alrededor de 0,93 - 0,94 (por ejemplo, un máximo registrado de 0,94118), por lo que 0,94 corresponde a una presentación coherente del rendimiento global en validación.

Desempeño por clase y análisis comparativo entre categorías

El rendimiento del modelo se analizó a nivel de categoría mediante métricas por clase obtenidas sobre el subconjunto de validación, con el fin de identificar diferencias de desempeño que pueden quedar ocultas al utilizar únicamente la exactitud global. En particular, se consideraron precisión predictiva, sensibilidad y medida armónica entre precisión predictiva y sensibilidad para cada una de las clases evaluadas [15]. Los valores consolidados se presentan en la Tabla 4, lo que permite comparar de forma directa el comportamiento del clasificador entre materiales.

Tabla 4. Métricas de desempeño por clase en validación.

Clase	Precisión predictiva	Sensibilidad	Medida armónica	Soporte (n)
Cartón	0,97	0,97	0,97	80
Metal	0,96	0,90	0,93	82
Papel	0,96	0,95	0,95	118
Plástico	0,91	0,93	0,92	96
Vidrio	0,91	0,95	0,93	100
Exactitud global	-	-	0,94	476
Promedio macro	0,94	0,94	0,94	476
Promedio ponderado	0,94	0,94	0,94	476

Nota. Las métricas corresponden al reporte de clasificación calculado sobre el subconjunto de validación.

En términos comparativos, la clase cartón registró el desempeño más alto y equilibrado (0,97 en las tres métricas), lo cual sugiere una capacidad consistente del modelo para identificar esta categoría con baja ambigüedad en el conjunto de validación. Las clases papel y metal mostraron también valores elevados en precisión predictiva (0,96 en ambos casos); sin embargo, se observó una diferencia en sensibilidad para metal (0,90), inferior a la registrada para papel (0,95).

Este patrón implica que, aunque las predicciones de metal tienden a ser correctas cuando el modelo selecciona esa clase, existe una proporción comparativamente mayor de casos de metal que el modelo asigna a otras categorías, reduciendo la cobertura de identificación de dicha clase dentro del subconjunto evaluado.

Por su parte, plástico y vidrio presentaron los valores más bajos de precisión predictiva (0,91 en ambos casos), aunque con sensibilidades relativamente altas (0,93 para plástico y 0,95 para vidrio). En términos prácticos, este comportamiento es compatible con un escenario en el que el modelo logra recuperar una fracción considerable de ejemplos reales de estas clases (sensibilidad), pero una parte de las predicciones emitidas como “plástico” o “vidrio” corresponde en realidad a otras categorías, lo que reduce la precisión predictiva.

En consecuencia, la medida armónica se mantiene en niveles competitivos (0,92 para plástico y 0,93 para vidrio), pero evidencia que estas clases requieren mayor atención en el análisis de confusiones y en posibles ajustes del proceso de preparación de datos o de entrenamiento, en conjunto, los resultados por clase muestran un desempeño alto y relativamente homogéneo.

Matriz de confusión y patrones de error

La caracterización de los errores de clasificación se realizó mediante la matriz de confusión, presentada en la Figura 2.

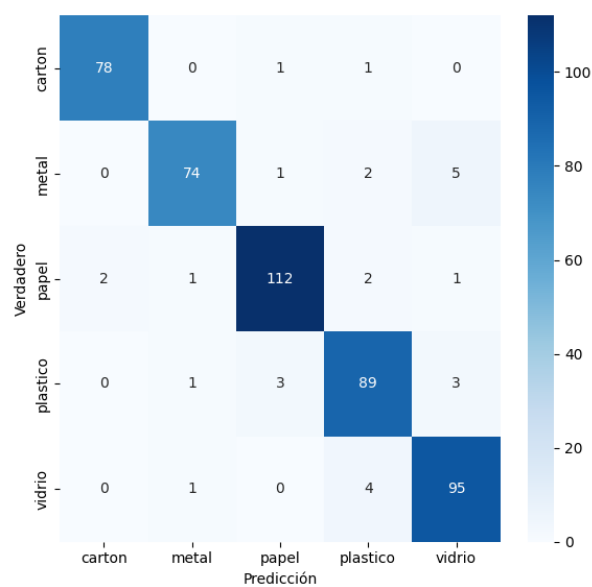


Figura 2. Matriz de confusión del modelo YOLOv8 para la clasificación de cinco categorías de residuos sólidos.

Nota. La matriz de confusión se generó a partir de las predicciones del modelo sobre el subconjunto de validación.

El desempeño del modelo se evaluó mediante una matriz de confusión (Figura 2), recurso que permitió identificar, de forma desagregada, la tasa de verdaderos positivos (diagonal principal) y los patrones sistemáticos de error entre categorías. Este análisis es determinante, ya que ofrece una visión de la robustez del modelo que no puede inferirse únicamente a partir de la exactitud global [16].

Desde una perspectiva interpretativa, los errores fuera de la diagonal señalan qué clases tienden a confundirse y, por tanto, dónde se concentra el riesgo de clasificación incorrecta en escenarios de uso. En tareas de clasificación de residuos, dichas confusiones suelen asociarse a similitudes visuales entre materiales (por ejemplo, superficies reflectantes, transparencias, texturas lisas, degradación del material o presencia de etiquetas), así como a condiciones de captura que alteran la percepción de color y brillo.

En términos de implicación práctica, la identificación de estos patrones orienta decisiones de mejora tales como:

- Reforzar la variabilidad del conjunto de entrenamiento en las clases con mayor confusión.
- Revisar la calidad y consistencia del etiquetado en categorías visualmente cercanas.
- Incorporar criterios adicionales de captura o preprocesamiento para reducir ambigüedades visuales.

Finalmente, la matriz de confusión constituye la base para discutir con mayor precisión el desempeño diferencial encontrado por clase en la Tabla 4, ya que permite explicar si valores relativamente menores de alguna métrica responden a confusiones concentradas con una clase específica o a errores distribuidos entre varias categorías.

Evidencia de validación tecnológica: ejecución del modelo en aplicaciones

Con el fin de complementar la evaluación cuantitativa del desempeño del clasificador, se documentó la ejecución funcional del modelo en dos modalidades de aplicación.

Esta validación tecnológica se entiende como una verificación de operatividad, orientada a comprobar que el modelo entrenado puede ser cargado y utilizado para generar predicciones a partir de entradas reales (imagen o cámara) dentro de interfaces orientadas al usuario, sin que ello implique, por sí mismo, una medición comparativa de rendimiento computacional o latencia (Figura 3).

En términos operativos, estas aplicaciones facilitan el uso del modelo en escenarios donde se requiere interacción directa con el usuario, visualización inmediata de la clase estimada y verificación práctica del flujo de predicción sin depender de servicios externos. Los puntos de entrada para esta modalidad corresponden a los scripts `app_seleccion.py` (clasificación por imagen) y `app_webcam.py` (clasificación por cámara), los cuales consumen el modelo entrenado almacenado en el directorio del proyecto.



Figura 3. Interfaz de usuario para la validación tecnológica del modelo.

Nota. Aplicación de escritorio desarrollada mediante el *framework* PyQt6. La interfaz integra el motor de inferencia YOLOv8 para permitir la clasificación de residuos mediante la carga de archivos locales o la captura de video en tiempo real a través de periféricos ópticos.

En la modalidad de escritorio, el proyecto incorpora una interfaz gráfica desarrollada con PyQt6, que permite realizar inferencia sobre imágenes seleccionadas desde el sistema de archivos y, adicionalmente, ejecutar clasificación en tiempo real mediante captura desde webcam.

En conjunto, estas implementaciones constituyen evidencia de portabilidad y uso práctico del modelo entrenado dentro de aplicaciones funcionales, mostrando que el clasificador puede integrarse en flujos de interacción habituales (selección de imagen y captura por cámara) y ofrecer resultados en tiempo real o cuasi tiempo real para propósitos de validación funcional.

Discusión

Los resultados de la validación evidencian que un enfoque de clasificación de residuos mediante visión por computador y aprendizaje profundo permite alcanzar niveles de desempeño compatibles con escenarios de separación asistida, especialmente cuando se trabaja con categorías reciclables bien definidas. Este comportamiento es coherente con la tendencia reportada en la literatura, donde conjuntos de datos como TrashNet se emplean de manera recurrente para evaluar modelos de clasificación de residuos y comparar arquitecturas de redes neuronales en contextos multiclase [1].

Desde la perspectiva de desempeño global, una exactitud aproximada de 0,94 en validación sugiere que el modelo logra una discriminación efectiva entre cinco categorías (cartón, papel, plástico, metal y vidrio) bajo las condiciones del subconjunto de evaluación.

No obstante, en clasificación multiclase, la exactitud por sí sola no describe la distribución de errores, por lo que su interpretación requiere complementarse con el análisis por categoría y con la matriz de confusión para identificar clases con mayor propensión a confusión, lo cual es una práctica recomendada en reportes comparables de clasificación de residuos.

El análisis por clase muestra un rendimiento alto y relativamente homogéneo, aunque con diferencias sistemáticas: cartón presenta el comportamiento más consistente, mientras que plástico y vidrio exhiben valores relativamente menores de precisión predictiva en comparación con otras categorías. Este patrón es importante en problemas de clasificación de residuos, donde ciertas clases pueden compartir rasgos visuales (brillos, transparencias, etiquetas, deformaciones) que incrementan la ambigüedad, en particular cuando se trabaja con imágenes capturadas bajo condiciones no estrictamente uniformes.

En este sentido, la matriz de confusión resulta determinante para establecer si el decremento de métricas responde a confusiones focalizadas entre pares de clases o a errores dispersos, lo que orienta decisiones de mejora sobre datos (variabilidad, balance, limpieza) y sobre configuración de entrenamiento (estrategias de regularización o ajustes de hiperparámetros) [17].

En relación con el comportamiento del entrenamiento, la estabilización observada hacia la segunda mitad del proceso es consistente con un esquema de transferencia de aprendizaje, donde una fase inicial de mejora rápida suele ser seguida por una meseta asociada al ajuste fino. Esto concuerda con prácticas habituales al entrenar clasificadores modernos con pesos preentrenados y ciclos de entrenamiento moderados, donde la lectura de curvas de pérdida y exactitud se utiliza como evidencia de convergencia y como criterio para evitar prolongar el entrenamiento sin beneficios sustantivos [18].

Una contribución significativa de este estudio reside en la validación tecnológica mediante el despliegue del modelo en entornos multiplataforma. La integración en aplicaciones de escritorio y el desarrollo de una interfaz web basada en el formato ONNX (Open Neural Network Exchange) [19] demuestran la viabilidad de ejecutar clasificadores de aprendizaje profundo en dispositivos finales con recursos limitados. Este enfoque técnico permite la inferencia en el lado del cliente (*client-side inference*), eliminando la dependencia de infraestructuras de servidor costosas y latencias de red. Dicha arquitectura se alinea con el paradigma *Local-first AI*, el cual prioriza la privacidad de los datos, la portabilidad y la eficiencia operativa al procesar la información directamente en el navegador

mediante ONNX Runtime. Esta estrategia de implementación simplifica la arquitectura del sistema al tiempo que valida la escalabilidad del modelo YOLOv8 para aplicaciones de monitoreo ambiental en tiempo real y entornos de computación distribuida.

En cuanto al marco tecnológico, el uso de la familia YOLO en implementaciones prácticas se sustenta en su orientación a la eficiencia y en la disponibilidad de herramientas de entrenamiento, validación y exportación en el ecosistema de Ultralytics, lo que facilita reproducibilidad operativa y despliegue en diferentes formatos.

Deben reconocerse limitaciones metodológicas para acotar el alcance de las conclusiones. Primero, el desempeño reportado corresponde al subconjunto de validación del flujo actual; por tanto, las inferencias deben restringirse a ese marco evaluativo y no extenderse, sin evidencia adicional, a escenarios no observados [20].

Segundo, aunque la integración en aplicaciones confirma operatividad, el trabajo no reporta de forma sistemática mediciones de latencia, consumo de memoria o pruebas de robustez controladas; por consiguiente, tales dimensiones deben considerarse como líneas de evaluación futura. En esa dirección, investigaciones posteriores pueden incorporar un subconjunto de prueba independiente, ampliar la variabilidad visual de los datos y establecer protocolos de evaluación computacional por dispositivo, de modo que el desempeño se caracterice por exactitud y por estabilidad y viabilidad en condiciones de despliegue reales, siguiendo tendencias recientes en clasificación automatizada de residuos.

CONCLUSIONES

Este trabajo evaluó la eficacia de un modelo de visión artificial basado en la arquitectura YOLOv8 para la clasificación automatizada de residuos sólidos urbanos. Los resultados experimentales confirman que el modelo clasifica de manera consistente cinco categorías fundamentales (cartón, papel, plástico, metal y vidrio), logrando una convergencia estable durante el entrenamiento y un desempeño robusto en tareas de clasificación multiclase. Se determinó que la integración de técnicas de transferencia de aprendizaje (*transfer learning*) sobre un dataset depurado es una estrategia eficaz para alcanzar precisiones superiores al 94 %, validando la idoneidad de las redes neuronales de una sola etapa para la gestión ambiental.

Desde una perspectiva técnica, la investigación demuestra la viabilidad de emplear enfoques de *Deep Learning* como soporte en procesos de separación asistida. La metodología propuesta ofrece una solución reproducible que abarca desde la ingeniería de datos hasta el despliegue de artefactos funcionales. El análisis detallado mediante matrices de confusión permitió trascender las métricas agregadas, identificando patrones de error asociados a similitudes visuales intrínsecas en materiales translúcidos y reflectantes. Este resultado proporciona criterios objetivos para la optimización de futuros modelos y la mejora de la calidad de los conjuntos de datos en aplicaciones de economía circular.

Finalmente, la validación tecnológica mediante el despliegue en interfaces de escritorio y entornos web basados en el paradigma *Local-first AI* confirma la portabilidad y escalabilidad del sistema. La capacidad del modelo para realizar inferencias en tiempo real, sin dependencia de infraestructuras de servidor, respalda su pertinencia en escenarios operativos que exigen retroalimentación

inmediata. En conjunto, este trabajo establece una base técnica para automatizar la clasificación de residuos.

Contribución y autoría

T.B. y **C.B.:** Desarrollo de software, implementación y despliegue, validación tecnológica, visualización, curación de datos, administración del proyecto y redacción: revisión y edición. **W.T.** y **C.V.:** Conceptualización, metodología, investigación, curación de datos, análisis formal, validación y redacción: preparación del borrador original. Todos los autores han leído y aceptado la versión publicada del manuscrito.

T.B. and **C.B.:** Software development, implementation and deployment, technological validation, visualization, data curation, project administration, and writing-review and editing. **W.T.** and **C.V.:** Conceptualization, methodology, investigation, data curation, formal analysis, validation, and writing-original draft preparation. All authors have read and agreed to the published version of the manuscript.

Financiamiento

Esta investigación no recibió financiación específica de agencias públicas, comerciales o sin ánimo de lucro. El estudio fue realizado de forma independiente como resultado del esfuerzo intelectual de los autores, contando únicamente con el apoyo institucional de SOFTECAPPS S.A.S. en el marco de sus actividades de investigación.

Declaración ética

El trabajo se basó en el uso de un conjunto de datos de imágenes de acceso público con fines académicos, y las actividades realizadas se limitaron al entrenamiento, evaluación e implementación de un modelo computacional, sin intervención sobre sujetos u entornos sensibles.

Uso de inteligencia artificial

La concepción del estudio, el diseño experimental, el análisis e interpretación de los resultados, así como la redacción y revisión crítica del manuscrito, fueron realizados de manera autónoma por los autores, quienes asumen la responsabilidad plena por el contenido final del artículo. Los autores declaran que se utilizaron herramientas de inteligencia artificial generativa de forma asistida exclusivamente para apoyar tareas técnicas de programación, en particular la optimización y refactorización del código fuente mediante un agente de copilot con modelos como Claude Sonnet 4.6, Opus 4.6 y GPT-Codex-Max 5.2.

Disponibilidad de datos

Los datos y los artefactos de investigación que respaldan los resultados de este estudio están disponibles de forma abierta en el repositorio Zenodo (DOI: [10.5281/zenodo.17566882](https://doi.org/10.5281/zenodo.17566882)). Asimismo, con el objetivo de garantizar la reproducibilidad computacional, el código fuente completo, los *scripts* de preprocesamiento, configuración del entrenamiento y módulos de despliegue han sido depositados en el repositorio público de GitHub: https://github.com/cxborjas/clasificacion_residuos.

Conflicto de interés

Los autores declaran que no existe ningún conflicto de interés de carácter financiero, personal, académico o profesional que pueda haber influido en el diseño del estudio, la obtención y análisis de los resultados, su interpretación o la preparación del manuscrito.

REFERENCIAS

- [1] X. Zhang, C. Liu, Y. Chen, G. Zheng, and Y. Chen, "Source separation, transportation, pretreatment, and valorization of municipal solid waste: A critical review," *Environment, Development and Sustainability*, vol. 24, pp. 11471-11513, 2022, doi: 10.1007/s10668-021-01932-w.
- [2] J. Aberger et al., "recAIdc: An Intelligent Assistance System for Manual Waste Sorting-Validation and Scalability," *Recycling*, vol. 10, no. 6, Art. no. 221, 2025, doi: 10.3390/recycling10060221.
- [3] P. Dhiman, S. K. Vishwakarma, and A. Tiwari, "Deep learning models for digital image processing: A review," *Artificial Intelligence Review*, 2023, doi: 10.1007/s10462-023-10631-z.
- [4] R. A. Aral, Ş. R. Keskin, M. Kaya, and M. Hacıömeroğlu, "Classification of TrashNet Dataset Based on Deep Learning Models," in *Proc. IEEE Int. Conf. on Big Data (Big Data)*, Seattle, WA, USA, 2018, pp. 2058–2062, doi: 10.1109/BigData.2018.8622212.
- [5] M. M. Islam, S. M. M. Hasan, M. R. Hossain, M. P. Uddin, and M. A. Mamun, "Towards sustainable solutions: Effective waste classification framework via enhanced deep convolutional neural networks," *PLOS ONE*, vol. 20, no. 6, e0324294, Jun. 2025, doi: 10.1371/journal.pone.0324294.
- [6] Z. Song, X. Zhang, and P. Tan, "YOLO-Fast: a lightweight object detection model for edge devices," *The Journal of Supercomputing*, vol. 81, Art. no. 724, 2025, doi: 10.1007/s11227-025-07172-3.
- [7] W. Ahmed, V. K. Kommineni, B. König-Ries, J. Gaikwad, L. Gadelha, and S. Samuel, "Evaluating the method reproducibility of deep learning models in biodiversity research," *PeerJ Computer Science*, vol. 11, e2618, 2025, doi: 10.7717/peerj-cs.2618.
- [8] G. Varoquaux, "Cross-validation failure: Small sample sizes lead to large error bars," *NeuroImage*, vol. 180, pp. 68-77, 2018, doi: 10.1016/j.neuroimage.2017.06.061.
- [9] G. Thung and M. Yang, TrashNet: Dataset of images of trash (cardboard, glass, metal, paper, plastic, trash) [Online]. Available: <https://github.com/garythung/trashnet>
- [10] D. Vieri, R. Susanto, E. S. Purwanto, and M. K. Ario, "Enhancing Waste Classification with YOLOv8 Models for Efficient and Accurate Sorting," *Procedia Computer Science*, vol. 245, pp. 889-895, 2024, doi: 10.1016/j.procs.2024.10.316.
- [11] G. Celik, "Multi-layer feature fusion for high-accuracy solid waste classification using a hybrid deep learning model," *The Visual Computer*, 2025, doi: 10.1007/s00371-025-04031-3.

- [12] A. El Ichi, W. Kaddah, M. El Bouz, and I. Badoc, "Optimizing YOLOv8n using hOSVD for efficient object detection: towards real-time video surveillance applications," *Signal, Image and Video Processing*, vol. 19, Art. no. 1066, 2025, doi: 10.1007/s11760-025-04301-6.
- [13] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv*, 2020, doi: 10.48550/arXiv.2010.16061.
- [14] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009, doi: 10.1016/j.ipm.2009.03.002.
- [15] C. Goutte and É. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," in *Advances in Information Retrieval (ECIR 2005)*, D. E. Losada and J. M. Fernández-Luna, Eds. Berlin, Heidelberg: Springer, 2005, pp. 345-359, doi: 10.1007/978-3-540-31865-1_25.
- [16] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sensing of Environment*, vol. 37, no. 1, pp. 35-46, 1991, doi: 10.1016/0034-4257(91)90048-B.
- [17] A. Theissler, M. Thomas, M. Burch, and F. Gerschner, "ConfusionVis: Comparative evaluation and selection of multi-class classifiers based on confusion matrices," *Knowledge-Based Systems*, vol. 247, Art. no. 108651, Jul. 2022, doi: 10.1016/j.knosys.2022.108651.
- [18] L. Prechelt, "Early Stopping - But When?," in *Neural Networks: Tricks of the Trade*, 2nd ed., G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Berlin, Heidelberg: Springer, 2012, pp. 53-67, doi: 10.1007/978-3-642-35289-8_5.
- [19] ONNX, "Open Neural Network Exchange," ONNX.ai, 2017. [En línea]. Disponible en: <https://onnx.ai>.
- [20] E. Vabalas, G. Gowen, E. Poliakoff, and A. J. Casson, "Machine learning algorithm validation with a limited sample size," *PLOS ONE*, vol. 14, no. 11, e0224365, 2019, doi: 10.1371/journal.pone.0224365.

Biografía de los autores

Thomás Ricardo Borja Saltos

Ingeniero en Electrónica y Telecomunicaciones y Magíster en Software con mención en Seguridad por la Escuela Politécnica Nacional, se desempeña como docente a tiempo completo en la carrera de Desarrollo de Software de la Facultad de Ingeniería en Sistemas de la misma institución. Su actividad académica se orienta al fortalecimiento de competencias en ingeniería de software, seguridad informática y desarrollo de soluciones tecnológicas aplicadas, con énfasis en métodos reproducibles y evaluación experimental de sistemas basados en inteligencia artificial.

Holds a B.Sc. in Electronics and Telecommunications Engineering and an M.Sc. in Software Engineering with a specialization in Security from Escuela Politécnica Nacional. He is a full-time lecturer in the Software Development program at the Faculty of Systems Engineering of the same institution. His academic work focuses on software engineering, cybersecurity, and the experimental evaluation of artificial intelligence-based systems, emphasizing reproducible methods and applied technological solutions.

Claudio Xavier Borja Saltos

Tecnólogo en Informática con mención en Análisis de Sistemas y Tecnólogo en Desarrollo de Software por el Instituto Tecnológico El Libertador. Es desarrollador de software y se desempeña como Gerente de Ventas en SOFTECAPPS S.A.S. Su experiencia se concentra en implementación de aplicaciones, integración de soluciones informáticas y despliegue de sistemas orientados a usuario final, combinando criterios técnicos y de gestión para la adopción de herramientas digitales.

Holds a degree as Information Technology Technologist with a specialization in Systems Analysis and a degree as Software Development Technologist from Instituto Tecnológico El Libertador (Ecuador). He works as a software developer and serves as Sales Manager at SOFTECAPPS S.A.S. His professional experience focuses on application implementation, software integration, and the deployment of end-user-oriented systems, combining technical execution with management-oriented decision-making.

Jorge Wilson Tamami Pachala

Licenciado en Ciencias de la Educación con mención en Informática Educativa, Especialista en Redes de Comunicación de Datos y Magíster en Informática Empresarial. Ha ejercido docencia en instituciones educativas y en docencia universitaria, además de desempeñar funciones públicas en la provincia de Bolívar, incluyendo Teniente Político, Jefe Político, Gobernador y cargos de gestión en el Distrito 02D01 de Educación. Su trayectoria integra educación, gestión institucional y tecnologías de la información aplicadas.

Holds a B.Ed. in Education with a specialization in Educational Computing, a postgraduate specialization in Data Communication Networks, and an M.Sc. in Business Informatics. He has taught in multiple educational institutions and in higher education, and has held public administration roles in Bolívar Province, including local and provincial leadership positions and educational district management. His background integrates education, institutional management, and applied information technologies.

Cristhian Geovanny Villamarín Arroba

Licenciado en Informática Educativa por la Universidad Estatal de Bolívar, Tecnólogo en Desarrollo de Software por el Instituto Tecnológico El Libertador y Magíster en Tecnologías de la Información por la Universidad Estatal de Milagro. Se desempeña como docente e investigador, con interés en informática aplicada, ciencia de datos y desarrollo de software. Su trabajo se orienta a la aplicación de enfoques computacionales en la mejora de procesos educativos y tecnológicos.

Holds a B.Ed. in Educational Computing from Universidad Estatal de Bolívar, a degree as Software Development Technologist from Instituto Tecnológico El Libertador, and an M.Sc. in Information Technologies from Universidad Estatal de Bolívar. He works as a lecturer and researcher with interests in applied computing, data science, and software development. His work emphasizes computational approaches to support educational and technological improvement initiatives.

Descargo de responsabilidad

Los libros y capítulos de libros publicados en la Editorial Unión Científica representan únicamente las opiniones de los autores. La Editorial Unión Científica, su equipo editorial y sus revisores no se hacen responsables del contenido, las interpretaciones o las consecuencias derivadas de la aplicación de los métodos o conclusiones incluidos en los trabajos. Todas las publicaciones se rigen por las políticas éticas de la editorial.