

Desarrollo de un Sistema Inteligente para la Clasificación Automática de Botellas Plásticas por Tamaño en la Empresa INPLASET SAC.

Ryan A. LEÓN LEÓN

La Libertad, Universidad Privada del Norte
Trujillo, 13001, Perú

Maricarmen A. LÓPEZ SANCHEZ

La Libertad, Universidad Privada del Norte
Trujillo, 13001, Perú

Ronaldo C. REYES CÁCERES

La Libertad, Universidad Privada del Norte
Trujillo, 13001, Perú

Development of an Intelligent System for the Automatic Classification of Plastic Bottles by Size at the Company INPLASET SAC.

ABSTRACT: This study aims to develop an intelligent system for the automatic classification of plastic bottles by size at INPLASET SAC, with a minimum expected accuracy of 90%, aiming to improve production efficiency and reduce error margins. The methodology included the collection and processing of a dataset of 2,649 images representing three bottle sizes. The images were labeled and preprocessed using Roboset to normalize their quality and resolution, facilitating proper model training. The YOLOv8 architecture was used in Google Colab, optimizing its accuracy and detection speed in identifying plastic bottles. After training, the system was implemented in PyCharm to perform real-time detection, evaluating the height and width of the bottles using computer vision techniques, including segmentation and edge detection, to classify the sizes. The results showed that the YOLOv8 model achieved 93% accuracy in bottle classification, displaying an optimal precision curve with high confidence levels. in the final printed copies of the Proceedings. Please keep in mind that the paper you prepare will be printed as it is received.

Keywords: YOLOv8, plastic industry, object detection, computer vision, PyCharm.

RESUMEN: Este estudio tiene como objetivo desarrollar un sistema inteligente para la clasificación automática de botellas plásticas por tamaño en la empresa INPLASET SAC, con una precisión mínima esperada del 90%, buscando mejorar la eficiencia en la producción y reducir el margen de error. La metodología incluyó la recolección y procesamiento de un conjunto de 2,649 imágenes, representando tres tamaños de botellas. Las imágenes fueron etiquetadas y preprocesadas usando Roboset para normalizar su calidad y resolución, facilitando un entrenamiento adecuado del modelo. Se utilizó la arquitectura YOLOv8 en Google Colab, optimizando su precisión y velocidad de detección en la identificación de botellas plásticas. Tras el entrenamiento, el sistema fue implementado en PyCharm para realizar la detección en tiempo real, evaluando la altura y el ancho de las botellas mediante técnicas de visión artificial, incluyendo segmentación y detección de bordes, para clasificar los tamaños. Los resultados indicaron que el modelo YOLOv8 alcanzó una precisión del 93% en la clasificación de botellas, mostrando una curva de precisión óptima con altos niveles de confianza.

Palabras Claves: YOLOv8, industria plástica, detección de

objetos, visión artificial, PyCharm.

1. INTRODUCCIÓN

Actualmente, la industria plástica en Perú tiene un impacto económico relevante, contribuyendo al crecimiento económico, la generación de empleo, el comercio exterior y la recaudación fiscal. Sin embargo, uno de los desafíos principales en este sector es el proceso de clasificación de botellas plásticas por tamaño. Este proceso resulta crucial para optimizar la eficiencia de producción y garantizar la calidad del producto final. No obstante, al realizarse de manera manual, enfrenta limitaciones como la lentitud, una mayor probabilidad de errores y un aumento en los costos operativos, lo cual reduce la competitividad y eficiencia de la industria. La clasificación manual no solo es un proceso lento y propenso a errores, sino que también aumenta los gastos operativos debido a la necesidad de mano de obra intensiva. Además, la inconsistencia en la clasificación de las botellas puede afectar la calidad del envasado y causar retrasos en la cadena de producción. La falta de automatización en este proceso limita la capacidad para cumplir con los estándares de calidad y eficiencia requeridos por el mercado. A medida que la demanda de productos de plástico sigue creciendo, es imperativo que se implemente soluciones tecnológicas que permitan aumentar su competitividad y sostenibilidad. La automatización podría reducir costos en un 30-40% y permitiría recuperar la inversión en 3 a 5 años. Esto subraya la necesidad de implementar soluciones tecnológicas para optimizar la competitividad y sostenibilidad del sector. Por lo anteriormente mencionado, surge la necesidad de desarrollar un sistema inteligente de clasificación de botellas plásticas de acuerdo al tamaño, que permita optimizar sus procesos, reducir costos y mejorar la calidad de su producción.

El desarrollo de sistemas inteligentes para la clasificación automática de objetos ha avanzado notablemente en la última década, impulsado por el crecimiento de las tecnologías de visión artificial y el aprendizaje profundo [1,2]. Es así, que las redes neuronales convolucionales (CNN) han sido ampliamente estudiadas y aplicadas en la clasificación y detección de objetos en imágenes, destacándose en diversos estudios la arquitectura YOLO como una de las soluciones más efectivas en este campo [3,4]. Dentro de la industria del plástico la automatización de la clasificación de botellas plásticas no solo mejora la eficiencia de la producción, sino que también reduce el error humano, ayuda en el reciclaje y optimiza el uso de recursos [5,6]. En este contexto, la capacidad avanzada que posee YOLO para procesar

imágenes en tiempo real la convierte en una opción ideal [7,8]. Su alta velocidad y precisión permiten optimizar el proceso de detección en comparación con métodos tradicionales basados en CNN, ofreciendo una solución eficaz para la clasificación automática. [9, 10]. Diferentes investigaciones han utilizado las diversas versiones de YOLO para resolver problemas de detección de objetos en tiempo real. Por ejemplo, un estudio desarrolló un sistema de reconocimiento de botellas de plástico basado en un modelo mejorado de YOLOv3. La captura y reconocimiento de imágenes se realizaron en OpenCV. Según los resultados se logró una tasa de precisión del 91% [11]. Además, investigaciones previas clasificaron botellas de plástico en función de su tamaño utilizando YOLOv5, logrando una precisión de hasta el 84% en tiempo real [12,13]. Así mismo, Matta et al. [14] utilizaron arquitecturas como YOLOv5 y YOLOv8 para detectar botellas de plástico en imágenes. Aunque YOLOv8 tiene más parámetros y requiere más tiempo de cálculo, se descubrió que superó a YOLOv5 en la identificación precisa de botellas de plástico. Por ello, el método propuesto se implementa con la última versión de YOLOv8, que ofrece mejoras significativas y ventajas para el reconocimiento y clasificación precisa de objetos, incluyendo botellas de plástico según su tamaño. YOLOv8 se destaca en condiciones de iluminación variable gracias a su arquitectura mejorada de extracción de características, que permite identificar detalles incluso en imágenes con poca luz, sombras profundas o brillo excesivo. Su entrenamiento con técnicas avanzadas de aumento de datos, como ajustes de brillo y contraste, lo prepara para responder eficazmente a diversas condiciones lumínicas, haciéndolo ideal para entornos con cambios constantes de luz, como al aire libre. En escenarios con superposición de objetos, YOLOv8 utiliza detección a múltiples escalas, mejorando la identificación de objetos parcialmente ocultos o en contacto. Estas capacidades posicionan a YOLOv8 por encima de modelos como Faster R-CNN y SSD. Mientras Faster R-CNN prioriza precisión a costa de velocidad y SSD ofrece rapidez con menor precisión en objetos pequeños o superpuestos, YOLOv8 combina velocidad y exactitud, consolidándose como la mejor opción para aplicaciones desafiantes en tiempo real.

La clasificación automática de botellas de plástico tiene la ventaja de una clasificación no intrusiva, velocidad, consistencia, rentabilidad a largo plazo e incluso evita riesgos para la salud de los trabajadores que trabajan en entornos de reciclaje. Como resultado, es imperativo reemplazar los sistemas de clasificación humanos por sistemas automatizados inteligentes. Este estudio tiene como objetivo desarrollar un sistema inteligente para la clasificación automática de botellas plásticas por tamaño en la empresa INPLASET SAC, con una precisión mínima esperada del 90% para reducir el margen de error comparado con la clasificación manual. Para ello, se plantean los siguientes objetivos específicos: integrar el sistema en la línea de producción asegurando su compatibilidad técnica y operativa con los equipos actuales; validar el sistema bajo condiciones industriales reales, considerando factores como iluminación variable, desgaste de las botellas y posibles interrupciones en el flujo de producción; y optimizar iterativamente el sistema para mejorar su precisión, velocidad y capacidad de reducción de errores, adaptándolo a las variaciones propias del entorno industrial. Finalmente, se pretende evaluar y optimizar el rendimiento del sistema en términos de precisión, velocidad y reducción de errores, realizando ajustes iterativos para adaptarse a variaciones en las condiciones de producción, como iluminación y desgaste de las botellas.

2. METODOLOGÍA

El proyecto de clasificación automática de botellas plásticas por tamaño comenzó con la selección y gestión de un conjunto de 2,649 imágenes de botellas, capturadas desde distintos ángulos y bajo diversas condiciones de iluminación. Las imágenes representaban tres categorías de tamaño específicas: pequeña (650 ml), mediana (2.5 litros) y grande (20 litros), asegurando que cada una de estas clases estuviera bien representada en el conjunto de datos. Se empleó una cámara Sony ZV-1F de 12 MP con alta resolución y sensor CMOS, ideal para capturar detalles precisos de las botellas. Las imágenes se obtuvieron bajo luz natural, y preferiblemente durante horas de luz suave para reducir sombras fuertes. La cámara se colocó a una distancia de entre 50 cm y 1 metro, montada en un trípode para asegurar estabilidad y uniformidad en los ángulos de captura (frontales y laterales), lo cual facilitó una recolección de datos consistente.

Para gestionar y procesar este volumen de imágenes, se utilizó la plataforma Roboset, una herramienta avanzada que facilita tanto el etiquetado preciso como la organización eficiente de datos visuales. En la primera fase, las imágenes fueron importadas y sometidas a un proceso de preselección donde Roboset ayudó a eliminar duplicados y a filtrar aquellas que no cumplían con los estándares de calidad, como imágenes borrosas o con mala iluminación. Esto permitió conformar un conjunto de datos limpio y representativo para el entrenamiento del modelo. Las imágenes incluyen botellas en diferentes contextos y fondos, tales como en superficies lisas, estanterías, o incluso mezcladas con otros objetos. Esta diversidad ayuda al modelo a enfocarse en las características relevantes de las botellas, evitando que aprenda patrones específicos relacionados únicamente con un fondo particular. Posteriormente, cada imagen fue etiquetada manualmente en Roboset según su categoría de tamaño, asegurando que el conjunto de datos estuviera perfectamente clasificado. La plataforma no solo facilitó el etiquetado, sino que también permitió aplicar un preprocesamiento básico para normalizar la resolución de las imágenes en 640x640 px y corregir cualquier discrepancia visual, lo que mejoró la consistencia del conjunto de datos y la eficacia del modelo de clasificación.

En cuanto a la distribución de los datos, se asignó el 90% de las imágenes para el entrenamiento, mientras que el 10% restante fue utilizado para la validación, lo que permitió entrenar un modelo robusto y confiable para la clasificación de botellas plásticas por tamaño.

Se utilizó YOLOv8, la versión más reciente de esta tecnología, destacada por su notable precisión y eficiencia. YOLOv8 se basa en una arquitectura de red neuronal convolucional (Figura. 1), que permite realizar la detección de objetos en una sola pasada por la red, a diferencia de otros métodos que requieren varias etapas de procesamiento. La selección de este modelo está respaldada por estudios recientes que demuestran la superioridad de YOLOv8 en cuanto a precisión y velocidad, comparado tanto con versiones previas como con otros modelos de detección de objetos [16,17]. En el sistema desarrollado para la clasificación automática de botellas plásticas por tamaño ver (Figura. 2), YOLOv8 emplea la función de activación SiLU (Sigmoid Linear Unit) en sus capas internas. Esta función juega un papel fundamental en el aprendizaje del modelo, ya que permite procesar y resaltar características críticas de las imágenes, como bordes y formas, independientemente de su tamaño.

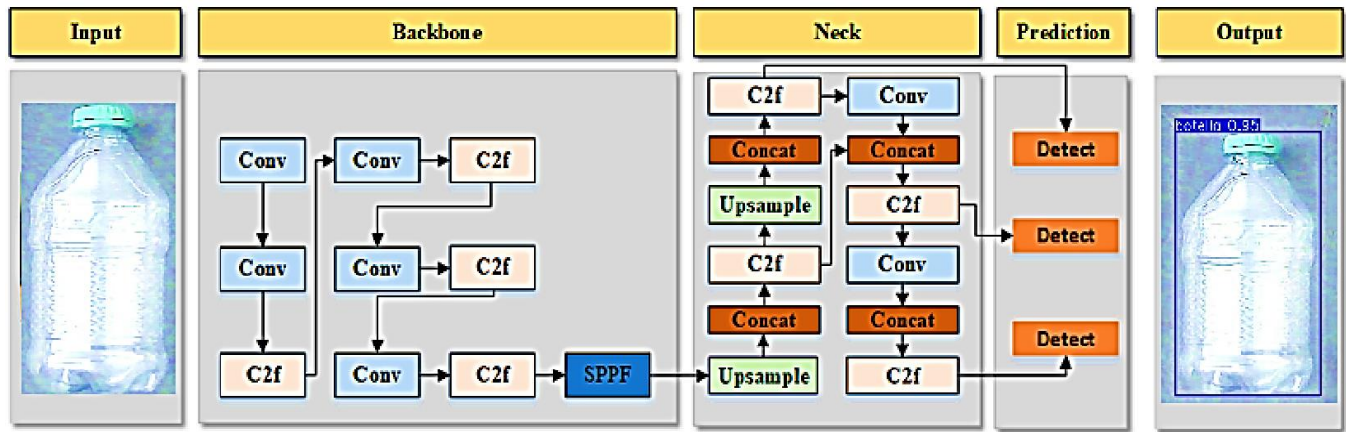


Fig.1 Arquitectura yolov8



Fig.2 Etiquetado botella de 650 ml – Roboset

A diferencia de funciones como ReLU, que descartan valores negativos, SiLU (Figura. 3) los activa de manera suave, conservando información importante que podría ser clave para identificar diferencias sutiles entre botellas de distintos tamaños. Además, SiLU Ec. (1) para proporciona una transición suave en valores cercanos a cero y un crecimiento casi lineal en valores positivos, lo que contribuye a una propagación eficiente de la información y a una representación más rica de las características visuales.

$$\text{SiLU}(x) = x \cdot \sigma(x) \quad (1)$$

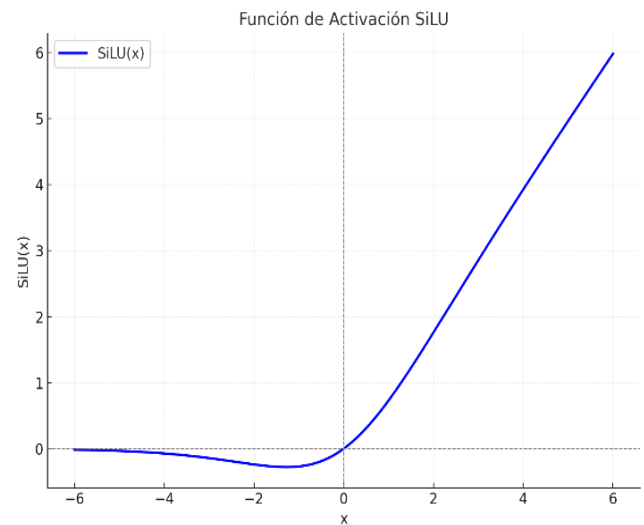


Fig. 3 Función de activación SiLU

Primero, la imagen de la botella es ingresada al sistema, donde se aplican varias operaciones en diferentes etapas de la arquitectura de YOLOv8. En la fase de Backbone, se utilizan capas convolucionales para extraer características clave como bordes, texturas y formas, lo que permite resaltar los aspectos más importantes de la botella (Figura. 4), el filtro se desplaza por la imagen de entrada, generando un valor de salida en cada posición mediante una suma ponderada entre los valores de la imagen y del filtro [18,19]. La convolución aplica un filtro (o kernel) (Figura. 5) a la imagen de entrada para generar un mapa de características.

Un kernel es una matriz pequeña (3×3) cuyos valores son aprendidos durante el entrenamiento. Este filtro se desliza sobre la imagen (operación conocida como stride), multiplicando sus valores por los de la imagen en cada posición y sumándolos para producir un nuevo valor en la posición correspondiente del mapa de características. Las capas C2f (Cross Stage Partial Networks) optimizan el intercambio de características entre las diferentes etapas de la red, permitiendo una mejor representación visual de la imagen. Además, se incluye un bloque SPPF (Spatial Pyramid Pooling Fast), el cual reduce la dimensionalidad de las características extraídas, mejorando la eficiencia computacional sin pérdida significativa de información.

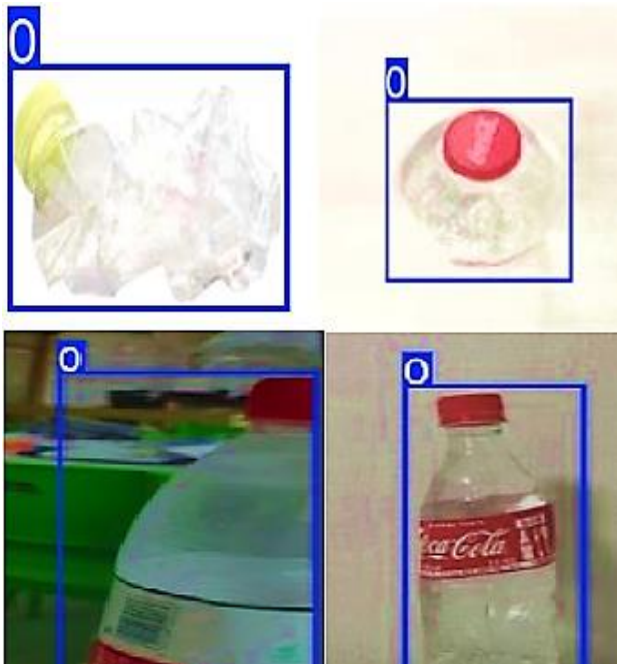


Fig. 4 Operaciones convolucionales

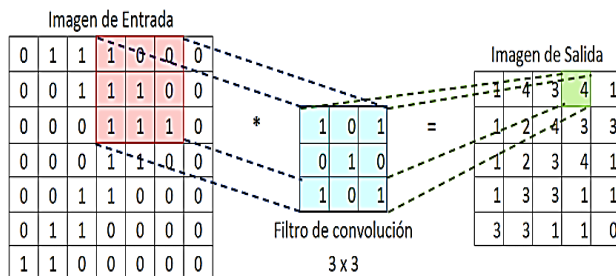


Fig. 5 Kernel – Convolución

En la fase de Neck, se integran y combinan las características extraídas a través de operaciones de upsampling (aumento de resolución) y concatenación para asegurar que la red aproveche tanto las características de bajo nivel (detalles finos) como las de alto nivel (características más abstractas). Estudios previos han demostrado su eficacia en la clasificación y detección de diversas características en objetos, como frutas, verduras y plásticos, subrayando su utilidad en el análisis de productos reciclables [20-22]. Este proceso permite optimizar la detección de las botellas, especialmente aquellas que puedan estar en posiciones complejas o parcialmente ocultas.

Finalmente, en la fase de Prediction, el modelo realiza la detección de las botellas en la imagen. Se generan predicciones que incluyen las coordenadas del cuadro delimitador de la botella, la clase (botella de plástico) y la confianza asociada a dicha predicción. La salida del sistema es una imagen con un cuadro delimitador alrededor de la botella, acompañado de una etiqueta que indica el tipo de objeto detectado y el nivel de confianza de la detección.

Este proceso es crucial para asegurar la correcta identificación de las botellas antes de proceder a la clasificación por tamaño. La arquitectura YOLOv8 permite realizar detecciones rápidas y precisas, lo cual es fundamental en la implementación de sistemas de clasificación automática de botellas plásticas por tamaño en entornos industriales.

Para entrenar el modelo YOLOv8 en la tarea de detección y clasificación de botellas plásticas por tamaño, se utilizó el optimizador AdamW, una variante del optimizador Adam Ec. (2) para incorporar regularización mediante decaimiento de los pesos (Weight Decay). Este optimizador combina lo mejor de los métodos de optimización basados en gradiente, ofreciendo una convergencia más rápida y estable al ajustar dinámicamente la tasa de aprendizaje de cada parámetro.

$$w_{t+1} = w_t - \eta \cdot \frac{m_t}{\sqrt{v_t + \epsilon}} + \lambda \cdot w_t \quad (2)$$

Esta investigación empleó Google Colab para entrenar el modelo YOLOv8.m, aprovechando su capacidad computacional y recursos especializados para el entrenamiento de redes neuronales, incluidas las GPUs. En concreto, se utilizó una GPU T4 con 15 GB de memoria, compatible con CUDA, lo que facilitó una notable aceleración del proceso de entrenamiento. La potencia de procesamiento de esta GPU garantizó una rápida convergencia del modelo y tiempos de entrenamiento eficientes. Después de configurar el entorno en Google Colab, se importaron las bibliotecas necesarias, como torch y yolov8, para el desarrollo y entrenamiento del modelo. PyTorch, conocida por su facilidad de uso y flexibilidad, es una popular biblioteca de código abierto para el aprendizaje profundo, ampliamente respaldada por revisiones por su eficacia en la investigación y el desarrollo de redes neuronales [23]. Para el entrenamiento en Google Colab se descargaron los datos etiquetados y se cargaron en el entorno de trabajo. Seguidamente, se configuró el modelo YOLOv8 ajustando parámetros como el tamaño del lote, la tasa de aprendizaje y el número de épocas, fijando este último en 450. Durante el entrenamiento, se supervisaron métricas como la pérdida y la precisión en el conjunto de validación, lo que permitió realizar ajustes en los hiperparámetros para mejorar el rendimiento del modelo (Figura. 6).

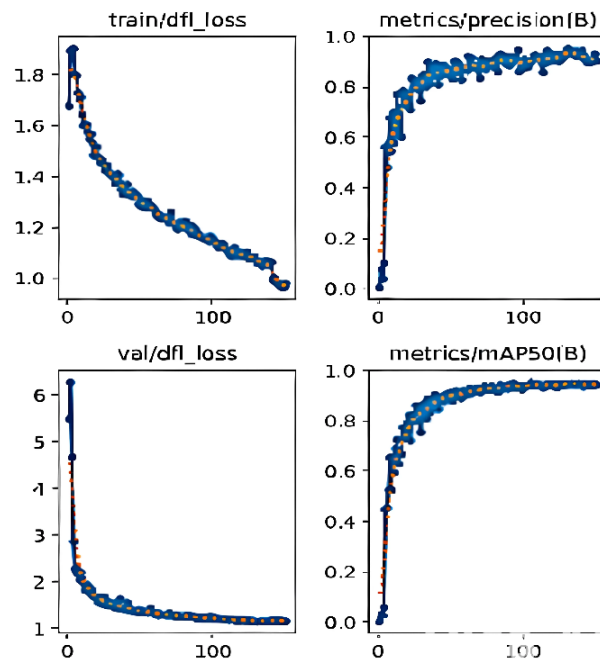


Fig. 6 Redimiendo del modelo

El gráfico superior izquierdo muestra la pérdida (train/dfloss) en los datos de entrenamiento, que disminuye consistentemente con las epochs, indicando que el modelo está aprendiendo. El

superior derecho presenta la precisión (metrics/precision(B)), que mide la proporción de detecciones correctas, aumentando rápidamente y estabilizándose cerca de 1.0, lo que refleja un buen rendimiento. El inferior izquierdo muestra la pérdida en los datos de validación (val/df1_loss), también en descenso, señalando que el modelo generaliza bien a datos no vistos. Finalmente, el inferior derecho ilustra el Mean Average Precision al 50% de IoU (metrics/mAP50(B)), una métrica de calidad Ec. (3) para la detección de objetos, que crece rápidamente y se estabiliza, evidenciando un rendimiento sólido.

$$\text{IoU} = \frac{\text{Área de Intersección}}{\text{Área de Unión}} \quad (3)$$

La función de pérdida es el elemento central que guía el entrenamiento de un modelo como YOLOv8. En este proyecto, su objetivo es minimizar los errores cometidos al predecir cada clase, optimizando así la precisión del modelo. Esta función cuantifica las discrepancias entre las predicciones del modelo y las etiquetas reales, lo que permite ajustar los pesos de la red neuronal para mejorar el rendimiento.

$$L = L_{\text{coord}} + L_{\text{conf}} + L_{\text{cls}} \quad (4)$$

La función de pérdida en YOLOv8 está compuesta por tres términos principales, cada uno con una contribución específica al aprendizaje del modelo. $L_{\text{"coord"}}$ mejora la precisión espacial al garantizar que las cajas delimitadoras predichas se ajusten correctamente al objeto. $L_{\text{"conf"}}$ refuerza la capacidad del modelo para diferenciar entre regiones con y sin objetos, reduciendo así falsos positivos y negativos. Finalmente, $L_{\text{"cls"}}$ optimiza la clasificación de los objetos, asegurando que las etiquetas asignadas a cada detección sean precisas. Estos tres componentes trabajan en conjunto Ec. (4) para maximizar la efectividad del modelo en tareas complejas de detección y clasificación.

La pérdida de coordenadas evalúa qué tan bien las cajas delimitadoras predichas se alinean con las cajas reales. En la Figura 6 se muestra la gráfica llamada Train/df1_loss. Para ello, utiliza métricas como IoU (Intersection over Union), que no solo miden la superposición entre las cajas, sino también la distancia entre los centros de las mismas y sus proporciones de aspecto. En este proceso, se consideran las coordenadas x,y, que representan el centro de la caja delimitadora normalizado respecto al tamaño de la celda en la cuadrícula, y las dimensiones w,h, que son proporciones del tamaño total de la imagen.

Al finalizar el proceso de entrenamiento, se evaluó el modelo utilizando un conjunto de datos de prueba separado, lo que permitió verificar que no hubiera sobreajuste y validar su capacidad para generalizar a datos no vistos. Este enfoque garantizó precisión del modelo en la detección de botellas plásticas (Figura. 7).

Para implementar el algoritmo de detección y clasificación, se utilizó Python en el entorno de desarrollo PyCharm [24]. El proceso comenzó configurando un entorno virtual en PyCharm, donde se instalaron las dependencias necesarias, como opencv-python para el procesamiento de imágenes y pytorch para el uso del modelo YOLOv8.

El modelo YOLOv8, previamente entrenado para detectar botellas plásticas, fue cargado en un script de Python. Este script procesó nuevas imágenes de botellas, ya sea capturando

imágenes desde una cámara en tiempo real o cargándolas desde el sistema de archivos. Una vez cargadas las imágenes, el modelo YOLOv8 detectaba las botellas en las mismas y generaba cuadros delimitadores con las etiquetas correspondientes.

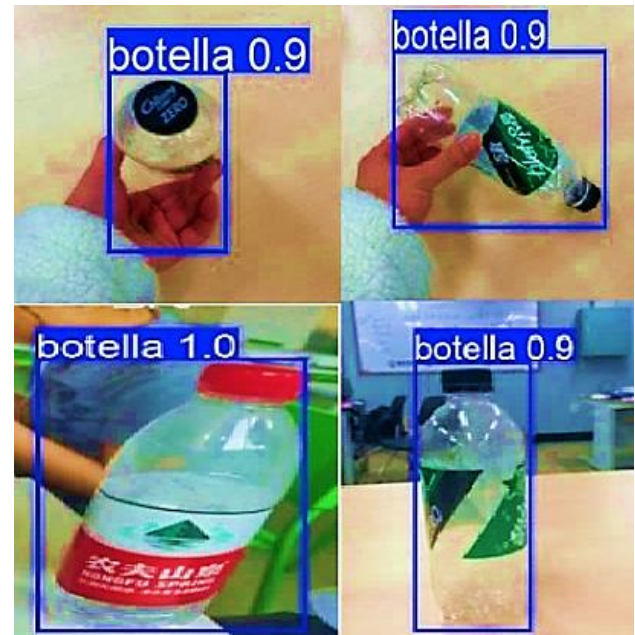


Fig. 7 Detección de botellas plásticas

Tras la detección de las botellas, se realizó una extracción de características para medir sus dimensiones. Utilizando las coordenadas devueltas por el modelo YOLOv8, se segmentaron las botellas y se aplicaron varias técnicas de visión artificial para medir las dimensiones clave, como la altura y el ancho. Entre las técnicas empleadas se utilizó la segmentación por umbralización Ec. (5) para separar la botella del fondo, facilitando la medición directa de sus dimensiones. Asimismo, las dimensiones del cuadro delimitador generado por YOLOv8 sirvieron como una aproximación inicial del tamaño de la botella. Se empleó también la detección de bordes para identificar los contornos de la botella, lo que permitió calcular con mayor precisión sus dimensiones. A partir del análisis de contornos se extrajeron características clave, como la altura y el ancho, y en algunos casos se aplicó la transformada de Hough Ec. (6) para detectar la forma cilíndrica de las botellas, obteniendo así dimensiones más precisas.

$$I_{\text{bin}}(x,y) = \begin{cases} 1 & \text{si } I(x,y) \geq T \\ 0 & \text{si } I(x,y) < T \end{cases} \quad (5)$$

$$(x - a)^2 + (y - b)^2 = r^2 \quad (6)$$

Además, en ciertos escenarios se utilizó una referencia física de tamaño conocido para calibrar las mediciones, convirtiendo los valores en píxeles a medidas físicas como centímetros o milímetros. Estas técnicas en conjunto permitieron calcular de manera efectiva el tamaño de las botellas según parámetros predefinidos.

Finalmente, las botellas fueron clasificadas en diferentes categorías de tamaño (pequeño, mediano, grande) de acuerdo con las dimensiones obtenidas. Los resultados se visualizaron en las imágenes originales, donde se mostraron las botellas con sus respectivas etiquetas de tamaño y cuadros delimitadores (Figura. 8).

3. RESULTADOS

La matriz de confusión normalizada revela que el modelo tiene un alto desempeño en la clasificación de las clases "botella" y "background". Para la clase "botella", el modelo alcanzó una precisión del 93%, calculada según la Ec. (8), clasificando correctamente la mayoría de los ejemplos como "botella". Además, el desempeño general del modelo en la detección de objetos, medido a través del Mean Average Precision Ec. (7), indica una alta capacidad del modelo para clasificar correctamente las instancias en diferentes umbrales de confianza; sin embargo, un 7% de las botellas fueron incorrectamente identificadas como "background". En general, el modelo demuestra una gran efectividad en la identificación de ambas clases, aunque existe un margen de error menor en la detección de botellas que podría ser optimizado.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (7)$$

$$P = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \quad (8)$$

La curva de precisión-confianza varía la precisión del modelo en función del nivel de confianza en la clase "botella" y en todas las clases combinadas. A medida que aumenta la confianza, la precisión también mejora, alcanzando el valor máximo de 1.00 a un nivel de confianza de 0.925. Esto indica que el modelo mantiene una alta precisión a medida que se vuelve más selectivo en sus predicciones, lo cual es ideal para aplicaciones en las que es importante minimizar falsos positivos al trabajar con niveles de confianza elevados. Complementariamente, la curva de recall en función del nivel de confianza permite analizar cómo la sensibilidad del modelo varía con diferentes umbrales de decisión, proporcionando un equilibrio entre la detección de objetos y la reducción de falsos positivos.

4. DISCUSIÓN

La matriz de confusión normalizada y la curva de precisión en función del nivel de confianza brindan una visión detallada del desempeño del modelo en la clasificación de las clases "botella" y "background". La matriz de confusión (Fig. 9) destaca un alto nivel de precisión del 93% en la detección de la clase "botella", lo cual es alentador para un sistema que busca optimizar la eficiencia en la clasificación de objetos. No obstante, el 7% de las botellas que fueron clasificadas incorrectamente como "background" refleja un margen de error bajo. El margen de error en la clasificación del modelo puede deberse a varios factores, principalmente relacionados con la variabilidad en las condiciones de captura de las imágenes y las limitaciones inherentes al conjunto de datos de entrenamiento. Cambios en la iluminación durante la captura de imágenes pueden afectar la claridad con la que se visualizan los bordes y características distintivas de las botellas, causando que el modelo confunda algunas de estas con el "background". Asimismo, la variabilidad en los ángulos y posición de las botellas genera dificultades cuando estas se presentan en perspectivas distintas a las utilizadas en el entrenamiento. Para mejorar, se podrían ajustar los pesos de las clases durante el entrenamiento, aumentar datos de fondo en el dataset, mejorar la calidad de las imágenes con preprocesamiento (e.g., eliminación de ruido) y ajustar hiperparámetros como el umbral de confianza del modelo. Una validación más robusta con técnicas como cross-validation también ayudaría a evaluar su estabilidad.

Al comparar nuestro sistema YOLOv8 con el sistema basado en YOLOv3, que alcanzó una precisión del 91% para detección de botellas plásticas [11], nuestro modelo presenta una ventaja porcentual en la tasa de precisión. Si bien esta diferencia podría parecer pequeña, resulta relevante en contextos industriales, donde incluso mejoras mínimas en la precisión pueden reducir errores en la clasificación y, por ende, aumentar la eficiencia operativa y minimizar desperdicios. La mejora puede atribuirse al uso de la última versión de YOLOv8, que optimiza tanto la detección como la precisión en entornos de producción reales. Asimismo, el modelo supera notablemente a las investigaciones previas que emplearon YOLOv5, las cuales alcanzaron una precisión de hasta el 84% en la clasificación de botellas según su tamaño [12,13]. La ventaja de 9 puntos porcentuales sobre YOLOv5 evidencia un avance considerable en la capacidad del modelo para adaptarse a la variabilidad en las condiciones de captura, como cambios en iluminación y posiciones de las botellas, lo que se traduce en una menor tasa de error en la detección y clasificación de objetos. Esto es especialmente valioso en aplicaciones donde se requiere una alta consistencia en la detección para asegurar la calidad del producto final.

Por último, al considerar el estudio de Matta et al. [14], donde se compararon YOLOv5 y YOLOv8, se descubrió que, aunque YOLOv8 tiene más parámetros y requiere más tiempo de cálculo, supera a YOLOv5 en la precisión de identificación de botellas de plástico. Nuestro proyecto, que utiliza YOLOv8, confirma estos hallazgos al mejorar tanto la precisión como la fiabilidad del modelo en un entorno industrial. A pesar del mayor costo computacional, el rendimiento superior de YOLOv8 justifica su implementación, ya que la alta precisión y la capacidad de adaptación a diferentes condiciones de producción compensan con creces el aumento en el tiempo de cálculo. Desde un punto de vista práctico, la aplicabilidad de estos resultados es significativa en diversos sectores industriales que dependen de la clasificación automática de objetos, como el reciclaje de plásticos, el control de calidad en líneas de producción y la logística de empaques. Por ejemplo, en plantas de reciclaje, la mejora en la precisión del modelo permite una mejor separación automatizada de botellas de plástico, reduciendo errores en la clasificación y aumentando la eficiencia del proceso. En líneas de producción de bebidas embotelladas, la detección más precisa de botellas defectuosas o mal posicionadas puede reducir pérdidas económicas y garantizar un control de calidad más estricto. Además, la capacidad de YOLOv8 para detectar objetos con alta precisión incluso en entornos con variabilidad de iluminación y ángulos diversos lo hace ideal para aplicaciones industriales automatizadas, donde la estabilidad del sistema es crítica. Su implementación en hardware optimizado, como GPUs industriales, facilita su uso en sistemas de visión por computadora en tiempo real, permitiendo respuestas rápidas y eficientes en entornos de producción.

Este proyecto no solo demuestra una precisión superior en la detección de botellas plásticas, sino que también subraya la eficacia de YOLOv8 como una herramienta robusta para mejorar la consistencia y reducir los errores en aplicaciones industriales. YOLOv8 representa una evolución significativa en la serie de arquitecturas YOLO, optimizando tanto la precisión como la velocidad en la detección de objetos en tiempo real. Una de sus principales ventajas es su capacidad para procesar y detectar múltiples objetos con alta precisión, incluso en entornos con variabilidad de iluminación, ángulos y complejidad de fondo. Esto lo hace ideal para aplicaciones industriales, donde la adaptabilidad y la consistencia son esenciales [25]. Además,

YOLOv8 incorpora mejoras en el manejo de anclajes adaptativos y predicción de bounding boxes (cajas delimitadoras) en comparación con versiones anteriores. Esto permite un mejor ajuste y localización de los objetos, lo que reduce los errores de clasificación y mejora la capacidad del modelo para distinguir objetos cercanos o parcialmente ocultos. Su diseño modular también facilita la optimización en distintos tipos de hardware, permitiendo que sea implementado en sistemas con recursos limitados o en configuraciones más avanzadas, como GPU industriales [26].

5. CONCLUSIONES

El desarrollo de un sistema inteligente para la clasificación automática de botellas plásticas según su tamaño en la empresa INPLASET SAC representa un avance significativo hacia la automatización de procesos industriales críticos. Este proyecto ha demostrado que la integración de tecnologías avanzadas de visión artificial, específicamente mediante el uso de arquitecturas de redes neuronales convolucionales como YOLOv8, puede mejorar considerablemente la eficiencia, precisión y velocidad en la clasificación de botellas. La implementación de este sistema no solo reduce la dependencia de mano de obra intensiva y disminuye el error humano, sino que también optimiza los recursos y fortalece la sostenibilidad operativa al mejorar la consistencia y calidad del producto final. Además, el sistema permite la adaptación continua y el ajuste a las condiciones cambiantes de la producción, asegurando su viabilidad y efectividad a largo plazo.

No obstante, el sistema aún presenta ciertas limitaciones que deben ser consideradas en futuras investigaciones. Uno de los principales desafíos es la sensibilidad a las condiciones de iluminación y variabilidad del entorno, lo que puede afectar la precisión del modelo en escenarios donde las imágenes presentan sombras, reflejos o iluminación no uniforme. Asimismo, el sistema enfrenta retos en la clasificación de botellas que aparecen en ángulos poco comunes o en posiciones parcialmente ocluidas. Además, el costo computacional asociado con la implementación de YOLOv8 en tiempo real puede representar una barrera en entornos industriales con restricciones de hardware, requiriendo GPUs de alto rendimiento para su ejecución eficiente. A esto se suma la dependencia del conjunto de datos de entrenamiento, ya que la diversidad y representatividad de las imágenes utilizadas influyen directamente en la robustez del modelo, pudiendo afectar su rendimiento en condiciones no vistas durante la fase de entrenamiento. Para abordar estas limitaciones, se plantean varias estrategias que podrían mejorar la aplicabilidad y eficiencia del sistema. En primer lugar, la optimización de modelos y la exploración de arquitecturas más ligeras, como versiones avanzadas de YOLO o modelos basados en Transformers, permitiría reducir la latencia sin comprometer la precisión. Además, el uso de técnicas de aumento de datos que simulen variaciones en iluminación, ángulos y condiciones adversas podría fortalecer la robustez del sistema ante cambios en el entorno.

Otra dirección prometedora es la implementación de aprendizaje continuo y autocalibración, lo que permitiría que el modelo se adapte a nuevos patrones de clasificación sin necesidad de reentrenamiento manual. Esto facilitaría su ajuste dinámico en entornos industriales cambiantes y su capacidad para mejorar el rendimiento con el tiempo. Asimismo, la expansión del sistema hacia la clasificación de otros tipos de materiales plásticos o tamaños de botellas incrementaría su versatilidad y aplicabilidad

en distintos sectores productivos. Finalmente, para evaluar la viabilidad de una adopción más amplia, sería beneficioso realizar un análisis económico detallado que considere costos de implementación, mantenimiento y escalabilidad en distintos escenarios industriales. La integración del sistema con plataformas de control de calidad y logística automatizada podría potenciar aún más su valor dentro de la cadena de producción.

6. REFERENCIAS

- [1] Wang, Zhaokun, et al. Classification for plastic bottles recycling based on image recognition. *Waste management*, 2019, vol. 88, p. 170-181. <https://doi.org/10.1016/j.wasman.2019.03.032>
- [2] L. R. Kambam and R. Aarthi, "Classification of plastic bottles based on visual and physical features for waste management," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-6, doi: 10.1109/ICECCT.2019.8869191.
- [3] Lee, Kok Jin, et al. Deep learning applied solid waste recognition system targeting sustainable development goal. *En Machine Intelligence in Mechanical Engineering*. Academic Press, 2024. p. 393-421. <https://doi.org/10.1016/B978-0-443-18644-8.00005-8>
- [4] A. Gizaw and T. Kebebew, "Water Bottle Defect Detection System Using Convolutional Neural Network," 2022 International Conference on Information and Communication Technology for Development for Africa (ICT4DA), Bahir Dar, Ethiopia, 2022, pp. 19-24, doi: 10.1109/ICT4DA56482.2022.9971271.
- [5] G. Aldric Sio, D. Guantero and J. Villaverde, "Plastic Waste Detection on Rivers Using YOLOv5 Algorithm," 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2022, pp. 1-6, doi: 10.1109/ICCCNT54827.2022.9984439.
- [6] Tan, Zhiying, et al. Identification of Recycled Plastic Bottles by Convolutional Neural Network Based Polarization Information: Leveraging Polarization Information for Accurate Sorting and Recycling of Plastic Bottles. *En Proceedings of the 15th International Conference on Digital Image Processing*. 2023. p. 1-10. <https://doi.org/10.1145/3604078.3604081>
- [7] Z. Tian, D. Sun, C. Yu, J. Li, G. Ma and H. Xia, "Garbage Classification Algorithm Based on Deep Learning," 2021 40th Chinese Control Conference (CCC), Shanghai, China, 2021, pp. 8199-8203, doi: 10.23919/CCC52363.2021.9549336.
- [8] Akbar, Fathi Sei Pahangai, et al. Object Detection on Bottles Using the YOLO Algorithm. *En 2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)*. IEEE, 2022. p. 1-5. doi: 10.1109/ICORIS56080.2022.10031554.
- [9] Wen, Shengping; Yuan, Yue; Chen, Jingfu. A vision detection scheme based on deep learning in a waste plastics sorting system. *Applied Sciences*, 2023, vol. 13, no 7, p. 4634. <https://doi.org/10.3390/app13074634>
- [10] H. Cho, "Transparent Plastic Bottle Detection and Depth Decision Method using YOLOv8 in Recyclable Waste Segregation Systems," 2024 IEEE International Conference on Omni-layer Intelligent Systems (COINS), London, United Kingdom, 2024, pp. 1-4, doi: 10.1109/COINS61597.2024.10622125.
- [11] J. Xiao, Y. Tang, Y. Zhao and Y. Yan, "Design of Plastic Bottle Image Recognition System Based on Improved

- YOLOv3," 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Harbin, China, 2020, pp. 2047-2050, doi: 10.1109/ICMCCE51767.2020.00445.
- [12] Karyanto, et. al, Detection of the Size of Plastic Mineral Water Bottle Waste Using the YOLOv5 Method. 2024; 7(2): 123-130. <https://doi.org/10.33387/jiko.v7i2.8535>
- [13] K. Chen, J. An, Y. Fang and T. Bu, "Research on solid waste plastic bottle cognitive based on YOLOv5s and deep stochastic configuration network," 2022 7th International Conference on Automation, Control and Robotics Engineering (CACRE), Xi'an, China, 2022, pp. 275-280, doi: 10.1109/CACRE54574.2022.9834213.
- [14] Matta, V.D., Mudunuri, K.A.V., Sai Baba, B.S., Kiran, K.B., Veenadhari, C.H.L., Prasanthi, B.V. Single Use Plastic Bottle Recognition and Classification Using Yolo V5 and V8 Architectures. Cognitive Computing and Cyber Physical Systems. IC4S 2023. vol 537. https://doi.org/10.1007/978-3-031-48891-7_8
- [15] Somantri, A. Sujjada, YS Bayan, VR Ulum, FA Ilhami y GP Insany, "Detección y clasificación mejoradas de plásticos: avances en el reconocimiento de variedades de plásticos mediante YOLOv8", 2023 IEEE 9th International Conference on Computing, Engineering and Design (ICCED). 2023, págs. 1-6, doi: 10.1109/ICCED60214.2023.10425093.
- [16] V. B. K. Reddy, S. S. Basha and J. Swarnalakshmi, "Augmenting Underwater Plastic Detection: A Study with YOLO-V8m on Enhanced Datasets," 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC). 2024, pp. 1309-1315, doi: 10.1109/ICAAIC60222.2024.10575159.
- [17] Shukhratov, Islomjon, et al. Optical detection of plastic waste through computer vision. Intelligent Systems with Applications, 2024, vol. 22, p. 200341. <https://doi.org/10.1016/j.iswa.2024.200341>
- [18] Z. Li, F. Liu, W. Yang, S. Peng and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," in IEEE Transactions on Neural Networks and Learning Systems. 2022; vol. 33, no. 12, pp. 6999-7019. doi: 10.1109/TNNLS.2021.3084827.
- [19] Traore, B. B., Kamsu-Foguem, B., & Tangara, F. Deep convolution neural network for image recognition. Ecological informatics. 2018; 48: 257-268. <https://doi.org/10.1016/j.ecoinf.2018.10.002>
- [20] Krichen, Moez. Convolutional neural networks: A survey. Computers, 2023, vol. 12, no 8, p. 151. <https://doi.org/10.3390/computers12080151>
- [21] R. Maliks and R. Kadikis, "Multispectral data classification with deep CNN for plastic bottle sorting," 2021 6th International Conference on Mechanical Engineering and Robotics Research (ICMERR), Krakow, Poland, 2021, pp. 58-65, doi: 10.1109/ICMERR54363.2021.9680850.
- [22] A. N. Kokoulin, A. I. Tur and A. A. Yuzhakov, "Convolutional neural networks application in plastic waste recognition and sorting," 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering. 2018, pp. 1094-1098, doi: 10.1109/EIConRus.2018.8317281.
- [23] Wang, Z., Liu, K., Li, J. et al. Various Frameworks and Libraries of Machine Learning and Deep Learning: A Survey. Arch Computat Methods Eng 31, 1–24 (2024). <https://doi.org/10.1007/s11831-018-09312-w>
- [24] Dokl, Monika, et al. A waste separation system based on sensor technology and deep learning: A simple approach applied to a case study of plastic packaging waste. Journal of Cleaner Production, 2024, vol. 450, p. 141762. <https://doi.org/10.1016/j.jclepro.2024.141762>
- [25] Hussain, Muhammad. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. Machines, 2023, vol. 11, no 7, p. 677. <https://doi.org/10.3390/machines11070677>
- [26] Sohan, Mupparaju, et al. A review on yolov8 and its advancements. En International Conference on Data Intelligence and Cognitive Informatics. Springer, Singapore, 2024. p. 529-545. https://doi.org/10.1007/978-981-99-7962-2_39