

Intelligent system based on Yolo v11 for detecting defects in bottle filling at a bottling plant

Luis Castro Zevallos¹; Oriol Tolentino Tucto²; Arnold Huamán Aguirre³; José Briones Zuñiga⁴

^{1,2,3,4}Universidad Tecnológica del Perú, U19207800@utp.edu.pe, 1220663@utp.edu.pe, C19532@utp.edu.pe, C19980@utp.edu.pe

Abstract– This study addressed the challenges of efficiency and accuracy in quality inspection at bottling plants, where manual methods lead to errors, high costs, and risks to brand reputation. An automated system based on artificial vision was implemented and evaluated to identify filling defects in water bottles at the Cielo brand plant in Iquitos during 2025. The objective was to determine the relationship between the implementation of the system and the improvement in defect detection. A quantitative and experimental approach was used, collecting and preprocessing a set of image data to train and validate a model from the YOLO family in Python. The model's performance was evaluated using metrics such as Mean Average Precision (mAP), accuracy, and recall. The results demonstrated high effectiveness in identifying bottles in good condition (87% recall) and moderate ability to detect defects (52% recall), reaching an inference speed of 30 FPS, suitable for real-time operation. It is concluded that the system is technically viable for improving the speed and accuracy of quality control, although further training is required to reduce the false negative rate in defective bottles.

Keywords– Yolo v11, Quality Control, Defect Detection, Bottling Industry, Industrial Automation.

Sistema inteligente basado en Yolo v11 para la detección de defectos en el llenado de botellas en una planta embotelladora

Luis Castro Zevallos¹; Oriol Tolentino Tucto²; Arnold Huamán Aguirre³; José Briones Zuñiga⁴

^{1,2,3,4}Universidad Tecnológica del Perú, U19207800@utp.edu.pe, 1220663@utp.edu.pe, C19532@utp.edu.pe, C19980@utp.edu.pe

Resumen– Este estudio abordó los desafíos de eficiencia y precisión en la inspección de calidad en plantas embotelladoras, donde los métodos manuales conducen a errores, costos elevados y riesgos para la reputación de la marca. Se implementó y evaluó un sistema automatizado basado en visión artificial para identificar defectos de llenado en botellas de agua en la planta de la marca Cielo en Iquitos durante 2025. El objetivo fue determinar la relación entre la implementación del sistema y la mejora en la detección de defectos. Se utilizó un enfoque cuantitativo y experimental, recopilando y preprocesando un conjunto de datos de imágenes para entrenar y validar un modelo de la familia YOLO en Python. El rendimiento del modelo se evaluó con métricas como la Precisión Media Promedio (mAP), precisión y recall. Los resultados demostraron una alta eficacia para identificar botellas en buen estado (87% de recall) y una capacidad moderada para detectar defectos (52% de recall), alcanzando una velocidad de inferencia de 30 FPS, adecuada para operación en tiempo real. Se concluye que el sistema es técnicamente viable para mejorar la velocidad y precisión del control de calidad, aunque se requiere un mayor entrenamiento para reducir la tasa de falsos negativos en botellas defectuosas.

Palabras clave– Yolo v11, Control de Calidad, Detección de Defectos, Industria Embotelladora, Automatización Industrial.

I. INTRODUCCIÓN

La creciente demanda de agua embotellada en Perú, impulsada por el consumo masivo y la disponibilidad de recursos hídricos, ha generado un mercado significativo con actores relevantes como la planta embotelladora de la marca Cielo en Iquitos. Sin embargo, este crecimiento también ha puesto de manifiesto diversos desafíos industriales, especialmente en lo que respecta a los procesos de control de calidad y eficiencia productiva en las líneas de llenado. En este contexto, la inspección manual de botellas, utilizada comúnmente, se ha mostrado ineficaz y propensa a errores humanos, lo que provoca defectos como niveles de llenado incorrectos, incrementando los costos operativos y el desperdicio de recursos[1].

Los sistemas de inspección manual en la industria embotelladora enfrentan limitaciones importantes para cumplir con los estándares modernos de calidad debido a su dependencia de la intervención humana. Por ello, se han explorado alternativas tecnológicas como los sistemas automatizados de inspección basados en Deep Learning y Redes Neuronales Convolucionales (CNN). Estas tecnologías han demostrado ser herramientas confiables para identificar defectos superficiales en múltiples sectores industriales[2]. En particular, la implementación de CNN en líneas de producción

de bebidas permite una inspección precisa y en tiempo real, mejorando la detección de defectos y optimizando la eficiencia operativa[2].

Estudios recientes han reportado casos de éxito en el uso de técnicas de Deep Learning para el control de calidad, destacando su capacidad para adaptarse a las características específicas de los datos y procesos productivos[2]. Además, el uso de metodologías como Lean Manufacturing y TPM (Total Productive Maintenance) también ha permitido incrementar la eficiencia de las líneas de producción mediante la reducción de tiempos inactivos y la mejora de la disponibilidad operativa[3], [4]. No obstante, persisten desafíos relacionados con la integración de estas tecnologías en los sistemas existentes y su impacto en la sostenibilidad económica de las empresas[5].

Ante el panorama descrito, esta investigación se centra en evaluar el impacto de la automatización de la inspección mediante sistemas basados en Redes Neuronales Convolucionales (CNN) en la industria embotelladora peruana. Dado que la producción de agua embotellada enfrenta presiones crecientes para reducir costos, minimizar defectos y cumplir con estándares de calidad, la implementación de estas tecnologías representa una oportunidad para transformar los procesos de control de calidad[2], [5]. Además, se busca analizar cómo estas soluciones pueden contribuir a la competitividad del sector, promoviendo prácticas más sostenibles y eficientes, en línea con las demandas del mercado y la conciencia ambiental.

Por lo tanto, la investigación tiene por objetivo determinar la efectividad de un sistema inteligente automatizado de inspección basado en YOLOV11 para optimizar la detección de defectos en las líneas de llenado de botellas. Los hallazgos contribuirán a establecer directrices para la adopción de sistemas automatizados en la industria embotelladora, fomentando su competitividad y sostenibilidad en el mercado peruano.

II. ESTADO DEL ARTE

La industria del agua embotellada en Perú representa un sector dinámico que responde a una constante demanda de los consumidores. Dentro de este contexto, asegurar la calidad del producto final es primordial, siendo el proceso de llenado una etapa crítica. Tradicionalmente, la inspección de calidad en las líneas de embotellado ha dependido en gran medida de métodos manuales; sin embargo, esta aproximación enfrenta

desafíos significativos, particularmente en entornos de alta producción como la planta embotelladora de la marca Cielo en Iquitos. La inspección manual es inherentemente propensa a errores humanos y a menudo resulta ineficaz para detectar consistentemente defectos sutiles [1], [2].

Esta situación puede derivar en la comercialización de productos no conformes, lo que no solo incrementa los costos operativos debido al desperdicio de producto y recursos — pudiendo generar un impacto económico adverso que, en algunos casos estudiados, alcanza el 25.75% de la facturación (equivalente a S/164,172.09) debido a una baja tasa de producción real— [6], sino que también puede impactar negativamente la reputación de la marca y la lealtad del consumidor [5]. Las consecuencias de fallos en el control de calidad, como niveles de llenado incorrectos o presencia de irregularidades en el envase, ponen en riesgo la rentabilidad y la posición competitiva de la empresa; por ejemplo, las paradas no planificadas en líneas de envasado han llegado a sumar 1,006 horas de indisponibilidad en un año, representando pérdidas valorizadas en S/. 994,010. [5]. Las estrategias convencionales, aunque necesarias, demuestran ser insuficientes para garantizar la calidad exigida en el mercado actual [4], donde la implementación de mejoras ha demostrado poder elevar los niveles de calidad de un 76.87% a un notable 99%..

Ante esta problemática, surge la necesidad de explorar e implementar alternativas tecnológicas avanzadas para la automatización del control de calidad. En este sentido, la literatura reciente destaca el potencial de los sistemas de visión artificial, y en particular, las Redes Neuronales Convolucionales (CNN), como una solución robusta y eficiente [7], [8], [9]. Las CNN, un tipo de red neuronal supervisada inspirada en la corteza visual humana, son especialmente aptas para el análisis y reconocimiento de patrones en imágenes [2], [10]. Estas redes emplean capas convolucionales jerárquicas para extraer características relevantes de las imágenes, desde bordes y texturas simples hasta formas complejas, permitiendo clasificar objetos con alta precisión, alcanzando en aplicaciones específicas de detección de envases una exactitud superior al 99.9% [2], [8], [11].

La aplicación de CNN en la inspección de botellas permite analizar imágenes en tiempo real para identificar defectos como niveles de llenado inadecuados, deformaciones o daños superficiales. Diversos estudios demuestran la capacidad de las CNN para adaptarse a variaciones en las condiciones de producción, como cambios de iluminación, asegurando una detección robusta [10]. El desarrollo de estos sistemas implica típicamente la recolección de un gran volumen de datos (imágenes de botellas conformes y defectuosas), técnicas de preprocesamiento para optimizar la entrada, el entrenamiento del modelo ajustando hiperparámetros, y la validación mediante métricas como precisión y recall. La implementación de estos modelos en sistemas de monitoreo en tiempo real facilita una respuesta inmediata a los fallos detectados, mejorando la eficiencia operativa; por ejemplo, la optimización de procesos mediante tecnología ha logrado

reducir el tiempo de máquina detenida en un 45.2% y, en otros casos, incrementar la disponibilidad de la línea en un 8% tras la implementación de mejoras [2], [8]. Por lo tanto, la adopción de tecnología CNN se presenta como una vía prometedora para superar las limitaciones de la inspección manual, garantizando la calidad del producto, optimizando la producción —logrando, por ejemplo, incrementar la tasa de producción real de un 74.05% a un 92.37%— y fortaleciendo la satisfacción del cliente en la planta embotelladora [12], [13], [14].

III. METODOLOGÍA

La metodología se organiza en las siguientes fases:

A. *Recolección de Datos*

Se recopiló un conjunto de datos de imágenes de botellas de agua directamente desde la línea de producción de la planta embotelladora. La adquisición de imágenes se realizó mediante una cámara digital con resolución de 1920×1080 píxeles y una tasa de captura de 30 cuadros por segundo (FPS), instalada a una distancia aproximada de 50 cm respecto a la línea de transporte y con un ángulo de captura perpendicular al flujo de botellas, con el fin de obtener una visualización clara del nivel de llenado.

Las condiciones de iluminación fueron parcialmente controladas mediante iluminación artificial constante, con el propósito de reducir variaciones ocasionadas por sombras, reflejos y cambios en el entorno industrial. Asimismo, se mantuvieron ciertas condiciones reales de operación para favorecer la capacidad de generalización del modelo en escenarios de producción.

El conjunto de datos estuvo conformado por un total de 1800 imágenes, distribuidas en dos clases: botellas en buen estado y botellas en mal estado. La clase de botellas defectuosas incluyó principalmente casos de nivel de llenado insuficiente, exceso de llenado y defectos visibles en el envase. La distribución de las imágenes presentó un ligero desbalance, con 1200 imágenes correspondientes a botellas en buen estado y 600 imágenes correspondientes a botellas defectuosas, condición que influye directamente en el desempeño del modelo para la detección de la clase minoritaria.

Posteriormente, el conjunto de datos fue dividido en tres subconjuntos: 70% para entrenamiento, 20% para validación y 10% para prueba, garantizando la representación de ambas clases en cada partición. Esta división permitió entrenar el modelo, ajustar sus parámetros y evaluar su capacidad de detección en imágenes no utilizadas durante el proceso de aprendizaje.

Durante la etapa de inferencia, el modelo YOLO generó automáticamente cuadros delimitadores (bounding boxes) sobre cada botella detectada, asignando una etiqueta y un color según la clase predicha, lo que facilita la identificación visual del estado de cada botella en tiempo real dentro de la línea de producción.

B. Preprocesamiento de Imágenes

Las imágenes recolectadas se prepararon para el modelo. Esto incluye ajustar su tamaño a un formato uniforme, normalizar los valores de los píxeles y aplicar técnicas de aumentación de datos (como rotaciones, cuadros negros o cambios de brillo) para mejorar la robustez y capacidad de generalización del modelo. (Fig. 1)

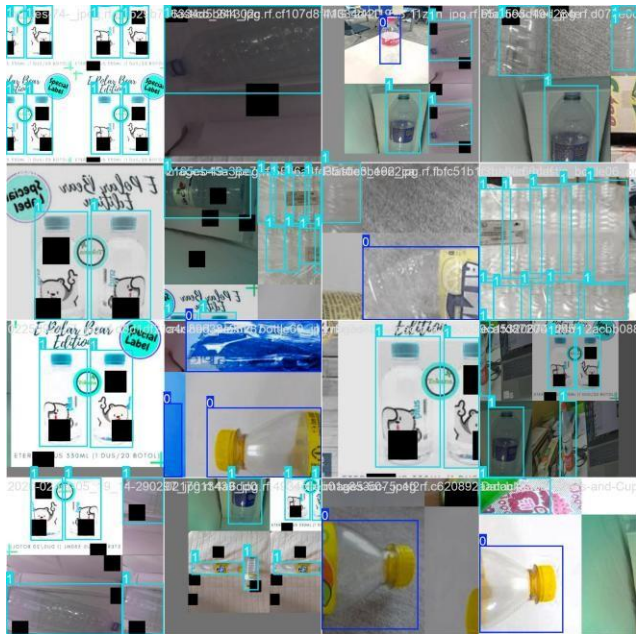


Fig. 1. Preprocesamiento de datos

C. Desarrollo y Entrenamiento del Modelo

Se empleó un modelo de detección de objetos basado en la arquitectura YOLOv11 (You Only Look Once), seleccionado por su capacidad para realizar detección en tiempo real con alta eficiencia computacional. Esta arquitectura permite localizar y clasificar objetos en una sola etapa, lo que resulta adecuado para aplicaciones industriales donde se requiere baja latencia y alta velocidad de procesamiento.

Para el desarrollo del sistema, se utilizó un modelo preentrenado como punto de partida, aplicando la técnica de fine-tuning con el objetivo de adaptarlo a la tarea específica de clasificación de botellas en buen estado y mal estado. Este enfoque permite aprovechar características previamente aprendidas en grandes volúmenes de datos, mejorando la capacidad de generalización del modelo y reduciendo el tiempo de entrenamiento.

El entrenamiento fue implementado en el entorno de programación Python, utilizando las librerías correspondientes para la ejecución del modelo YOLO. Las imágenes de entrada fueron redimensionadas a un tamaño uniforme de 640×640 píxeles, permitiendo mantener un equilibrio adecuado entre precisión de detección y velocidad de inferencia.

Durante el proceso de entrenamiento, se definieron los siguientes hiperparámetros: 50 épocas (epochs), un tamaño de lote (batch size) de 16, una tasa de aprendizaje (learning rate)

de 0.001 y el uso del optimizador Adam. Estos parámetros fueron seleccionados con el fin de lograr una convergencia estable del modelo y minimizar el riesgo de sobreajuste.

El modelo fue entrenado de manera iterativa, ajustando sus parámetros internos para minimizar la función de pérdida, la cual integra componentes de localización, clasificación y ajuste de las cajas delimitadoras. De forma paralela, se utilizó el conjunto de validación para monitorear el desempeño del modelo durante el entrenamiento, permitiendo identificar el punto óptimo de aprendizaje.

El proceso general de entrenamiento sigue un flujo estructurado que inicia con la preparación del conjunto de datos, incluyendo su recopilación, clasificación y partición en subconjuntos de entrenamiento, validación y prueba. Posteriormente, el modelo es entrenado de forma iterativa hasta alcanzar un desempeño adecuado, el cual es evaluado mediante métricas como la Precisión Media Promedio (mAP).

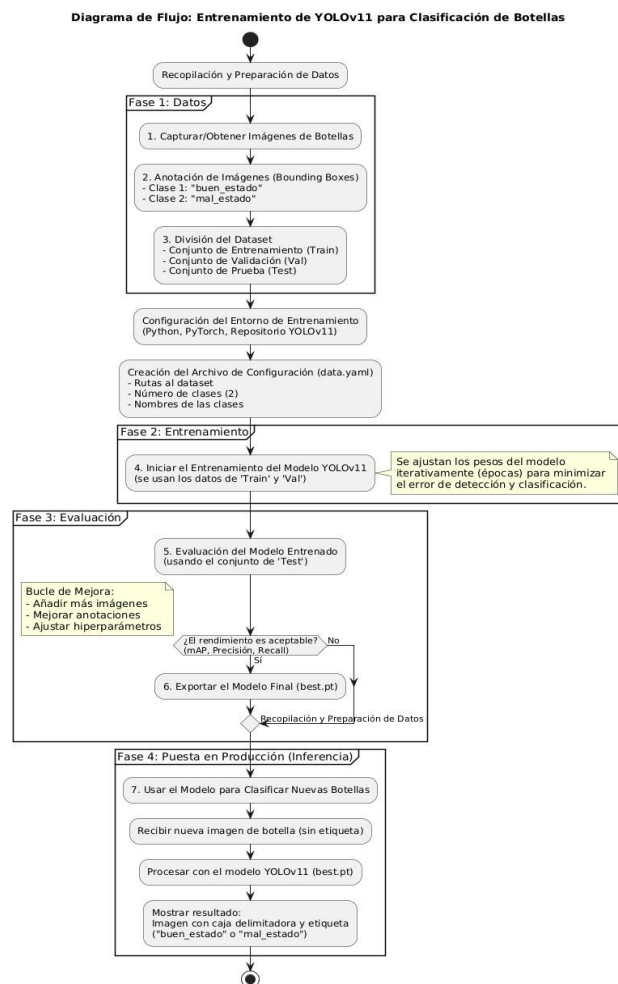


Fig. 2. Secuencia de entrenamiento del modelo

Finalmente, se seleccionó el modelo con mejor rendimiento obtenido durante el entrenamiento, almacenado como archivo best.pt, el cual fue utilizado en la fase de inferencia para la detección automática de botellas en tiempo real dentro del sistema propuesto. Este flujo de entrenamiento y despliegue se

resume en la Fig. 2.

D. Implementación del Sistema

El modelo YOLO validado fue integrado en un sistema de monitoreo capaz de operar en tiempo real sobre la línea de producción. Dicho sistema procesa de forma continua las imágenes adquiridas por la cámara, permitiendo la detección y clasificación automática del estado de cada botella mediante técnicas de visión artificial.

De manera complementaria, se incorporaron sensores de nivel como mecanismo de verificación redundante del proceso de llenado, con el objetivo de contrastar la información obtenida por el modelo de visión con mediciones físicas y aumentar la confiabilidad del sistema.

La salida del modelo es transmitida a un sistema de control basado en un microcontrolador tipo Arduino, seleccionado por su bajo costo, facilidad de implementación y flexibilidad en entornos de prototipado. Este dispositivo ejecuta la lógica de decisión del sistema; cuando se detecta una botella en condición defectuosa, se genera una señal de control que activa un actuador neumático encargado de retirar automáticamente el producto de la línea de producción.

Este enfoque permite la automatización del proceso de inspección mediante una arquitectura accesible y escalable, manteniendo la posibilidad de integración futura con sistemas de control industrial basados en PLC. (Fig. 3).

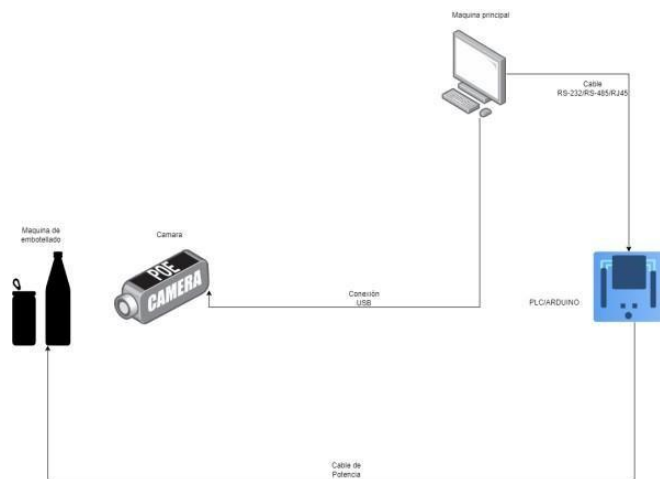


Fig. 3. Instalación en planta embotelladora

E. Análisis de Resultados

Se evaluó el desempeño del sistema automatizado en términos de precisión, velocidad y consistencia, comparándolo con los resultados obtenidos mediante los métodos de inspección manual. El rendimiento del modelo fue analizado utilizando las métricas definidas en la Tabla I, lo que permitió cuantificar su capacidad de detección y clasificación en condiciones reales de operación.

TABLA I. VARIABLES DE RESULTADOS MÁS RELEVANTES PARA EVALUAR UN MODELO YOLO

Variable de Resultado (Métrica)	Descripción
mAP (mean Average Precision - Precisión Media Promedio)	Métrica estándar que promedia la Precisión Media (AP) a través de todas las clases. Considera la calidad de localización (IoU) y clasificación.
Precisión (Precision)	De todas las detecciones que el modelo hizo para una clase específica (o en general), ¿qué proporción fue correcta? Mide la exactitud de las predicciones positivas.
Recall (Sensibilidad, Exhaustividad)	De todos los objetos reales de una clase específica (o en general) presentes en los datos, ¿qué proporción detectó correctamente el modelo? Mide la capacidad del modelo para encontrar todos los objetos relevantes.
F1-Score	Media armónica de Precisión y Recall, que mide el balance entre ambas.
Velocidad de Inferencia (FPS / Latencia)	El tiempo que tarda el modelo en procesar una sola imagen y generar las predicciones. Se expresa comúnmente en cuadros por segundo (FPS) o milisegundos por imagen (latencia).

IV. RESULTADOS Y DISCUSIÓN

Tras el entrenamiento del modelo YOLO, se obtuvieron diversas métricas y visualizaciones que permiten evaluar su desempeño. Estos resultados corresponden a la capacidad del modelo para identificar y localizar correctamente las botellas "Botella-buen-estado" y "Botella-mal-estado".

A. Rendimiento del Entrenamiento del Modelo

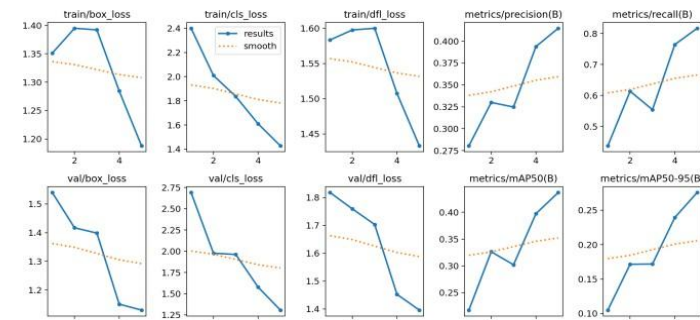


Fig. 4. Curva de progresión de aprendizaje.

Las curvas de aprendizaje (Fig. 4) muestran la evolución de las métricas a lo largo del proceso de entrenamiento. En ellas, la línea continua representa los valores correspondientes al conjunto de entrenamiento, mientras que la línea discontinua corresponde al conjunto de validación.

Se observa una tendencia decreciente en las funciones de pérdida (box_loss, cls_loss, dfl_loss) en ambos conjuntos, lo que indica una mejora progresiva en la capacidad del modelo para ajustar las predicciones. De manera simultánea, las métricas de rendimiento como precisión, recall y mAP presentan una tendencia ascendente, evidenciando una mejora

en la capacidad de detección.

Este comportamiento sugiere que el modelo logró un aprendizaje adecuado a partir de los datos, sin presentar indicios significativos de sobreajuste (overfitting) durante las épocas evaluadas, mostrando una adecuada capacidad de generalización frente a datos no vistos.

- **train/box_loss** (Pérdida de las cajas delimitadoras)
 - En la época 1, el valor de pérdida comienza en torno a 1.35.
 - Aumenta ligeramente en la época 2, lo cual puede deberse a un ajuste en el aprendizaje o ruido inicial.
 - A partir de la época 3 hasta la 5, se observa una disminución constante, bajando hasta aproximadamente 1.20.
 - Esto indica que el modelo aprende progresivamente a ajustar mejor las cajas alrededor de los objetos.
- **train/cls_loss** (Pérdida de clasificación)
 - Disminuye rápidamente desde 2.4 en la época 1 hasta cerca de 1.2 en la época 5.
 - Esta reducción constante refleja una mejora clara en la capacidad del modelo para clasificar correctamente los objetos detectados.
- **train/df_l_loss** (Distribution Focal Loss) Comienza en aproximadamente 1.6, mantiene ese valor en las épocas 1-2, pero luego desciende de manera constante hasta cerca de 1.42.
 - Representa que el modelo mejora en la precisión fina del ajuste de los bordes de las cajas, un aspecto importante en tareas de detección precisa.
- **metrics/precision(B)** (Precisión)
 - Aumenta desde un valor inicial bajo (~ 0.275) hasta llegar a más de 0.40 en la época 5.
 - Muestra una mejora en la proporción de predicciones correctas entre todas las predicciones hechas (menos falsos positivos).
 - La línea punteada muestra una tendencia ascendente clara.
- **metrics/recall(B)** (Sensibilidad o Recall)
 - Sube desde aproximadamente 0.40 hasta 0.82.
 - Esto indica que el modelo detecta cada vez más objetos presentes en la imagen, reduciendo falsos negativos.
- **Gráficos de Validación (Fila Inferior)**

- **val/box_loss**
 - Similar al entrenamiento: desciende desde ~ 1.5 hasta cerca de 1.15.
 - Indica que la capacidad del modelo para predecir correctamente las cajas en datos nuevos mejora con las épocas.
- **val/cls_loss**
 - Baja de forma significativa desde 2.7 hasta ~ 1.25 .
 - Reafirma que el modelo generaliza bien la clasificación en datos no vistos.
- **val/df_l_loss**
 - Disminución progresiva desde 1.8 a 1.4.
 - Esto confirma que no solo en entrenamiento, sino también en validación, el modelo refina la predicción de los bordes de las cajas.
- **metrics/mAP50(B)**
 - La media de precisión para un $\text{IoU} \geq 0.5$ aumenta, con un ligero salto desde la época 2 a la 3, y una mejora clara hasta ~ 0.55 en la época 5.
 - Esta métrica combina precisión y recall, mostrando buen progreso en detección general.
- **metrics/mAP50-95(B)**
 - Es la métrica más exigente, ya que considera múltiples umbrales de IoU (de 0.5 a 0.95).
 - Aunque muestra más variabilidad, termina en ascenso, pasando de ~ 0.20 a ~ 0.28 , reflejando que el modelo mejora no solo en detectar, sino en hacerlo con precisión milimétrica.

B. Métricas de Detección en el Conjunto de Prueba

El análisis de la curva F1-Confidence (Fig. 5) muestra que el rendimiento óptimo del modelo, con un F1-Score promedio de 0.55, se alcanza con un umbral de confianza de aproximadamente 0.186. Este resultado indica que el mejor equilibrio entre precisión y recall se obtiene utilizando un umbral inferior al valor estándar (0.5).

Asimismo, se observa que el desempeño varía entre clases, siendo mayor en la clase con características más definidas, lo que sugiere que la capacidad de detección del modelo depende tanto de la calidad como de la representatividad de los datos de entrenamiento.

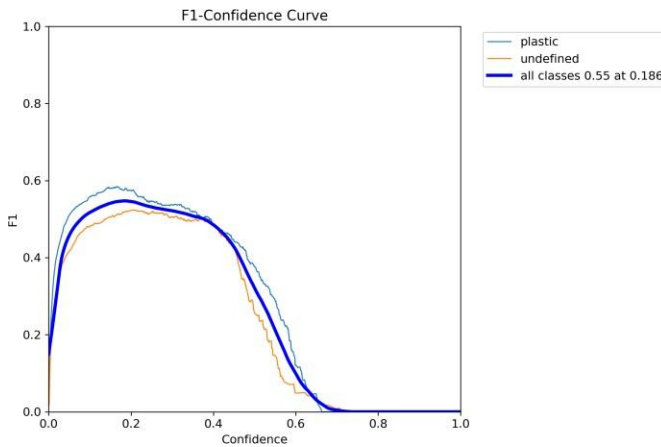


Fig. 5. Curva F1-Confidence del modelo.

Esta gráfica revela que el rendimiento óptimo general del modelo, medido por el F1-Score promedio para todas las clases, es de 0.55. Este pico se alcanza cuando el umbral de confianza se establece en aproximadamente 0.186. Esto sugiere que este es el punto de operación ideal si se busca el mejor balance general entre no omitir detecciones importantes y evitar un exceso de falsas alarmas.

Así como la representación de cómo varía el F1-score (una medida de equilibrio entre precisión y recall) en función del umbral de confianza para las predicciones del modelo, separado por clase:

- Eje X (Confidence): Probabilidad mínima que debe tener una predicción para ser considerada válida.
- Eje Y (F1): F1-score calculado para ese nivel de confianza.

Curvas por Clase:

Línea azul claro (plastic): Representa la clase "plastic". Tiene el mayor rendimiento F1, alcanzando un pico alrededor de 0.55. Su punto óptimo ocurre cuando el modelo tiene baja confianza (~0.15 a 0.25), lo cual indica que esta clase es detectada correctamente incluso cuando las predicciones no son altamente seguras. Después del pico, el F1-score disminuye rápidamente a medida que se eleva el umbral de confianza, ya que se descartan muchas verdaderas positivas por ser consideradas "inseguras".

Línea naranja (undefined): Representa la clase "undefined". Su F1-score es inferior al de la clase "plastic", alcanzando un máximo cerca de 0.47. También sigue una forma de campana, con descenso más pronunciado después de su pico, indicando que esta clase es más difícil de detectar con alta precisión y confianza. Línea azul oscuro (all classes): Representa el promedio para todas las clases. El pico máximo de $F1 = 0.55$ se alcanza en un umbral de confianza de 0.186, como se indica en la leyenda. Este valor sugiere que el mejor equilibrio entre

precisión y recall para el modelo ocurre cuando se consideran predicciones con al menos 18.6% de confianza.

El mejor umbral de confianza para el modelo es 0.186. Más allá de ese punto, el F1-score cae debido a la eliminación de muchas verdaderas positivas (aumentan falsos negativos). Esto sugiere que ajustar el umbral de decisión por debajo de los típicos 0.5 puede mejorar el rendimiento si el objetivo es maximizar F1. La diferencia entre clases indica que la clase "plastic" se detecta mejor que "undefined", probablemente por características visuales más definidas o mayor cantidad de ejemplos en el entrenamiento.

C. Análisis de la Matriz de Confusión

La matriz de confusión normalizada permite analizar el desempeño del modelo por clase. Para una evaluación práctica, nos centramos en la capacidad del modelo para diferenciar entre "Botella-buen-estado" y "Botella-mal-estado".

TABLA II. PORCENTAJE DE DETECCIÓN CORRECTA (RECALL) POR CLASE DE BOTELLA

Tipo de Botella	Porcentaje de detección
Buen Estado	87%
Mal Estado	52%
Tasa de Error	31%

El modelo demostró una alta capacidad para identificar correctamente las botellas en buen estado, con un recall del 87%. Sin embargo, la detección de botellas defectuosas fue significativamente más baja, logrando solo un recall del 52%. El error más crítico es que el 10% de las botellas en mal estado fueron clasificadas incorrectamente como en buen estado, un fallo que permitiría que productos defectuosos avancen en la línea de producción. Adicionalmente, el 38% de las botellas en mal estado fueron ignoradas por el modelo (clasificadas como fondo), lo que también representa una falla grave en el control de calidad.

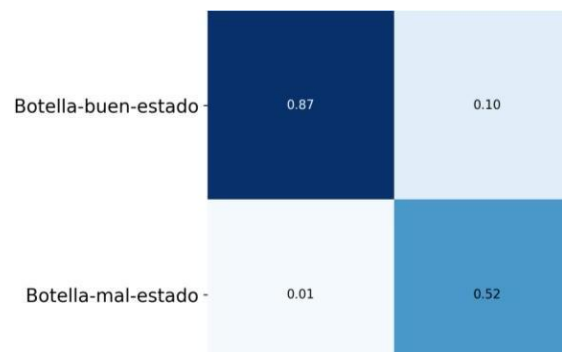


Fig. 6. Matriz de confusión Normalizada

La matriz de confusión (Fig. 6) confirma que el modelo es fiable para identificar la clase "Botella-buen-estado" (87% de acierto). No obstante, su rendimiento decae drásticamente con la clase "Botella-mal-estado", que solo se clasifica correctamente en un 52% de las ocasiones. La principal debilidad del modelo es su tendencia a omitir detecciones (clasificar como fondo), lo que ocurre en el 38% de las botellas defectuosas y en el 12% de las

botellas en buen estado. En resumen, si bien el modelo es prometedor, su actual incapacidad para identificar de manera fiable los defectos y la alta tasa de omisión lo hacen insuficiente para un sistema de control de calidad final sin un mayor refinamiento.

Al comparar estos resultados con otras investigaciones, se evidencia un margen de mejora. Trabajos como los citados en [2], [8] y [11] reportan precisiones superiores al 99% en tareas de detección de defectos en envases utilizando arquitecturas similares. Asimismo, la mejora de niveles de calidad de un 76.87% a un 99% documentada en [4] establece un punto de referencia industrial. El rendimiento de nuestro modelo (mAP50 de 0.55 y recall del 52% para defectos) es inferior, lo que subraya que nuestro sistema es un prototipo funcional que requiere un mayor refinamiento. La diferencia puede atribuirse a un conjunto de datos limitado en variedad y cantidad de defectos, así como a la necesidad de un mayor ajuste de hiperparámetros.

D. Aplicación en Tiempo Real y Propuesta de Automatización

Para evaluar su aplicabilidad práctica, el modelo se implementó en un script de Python utilizando OpenCV, procesando video en tiempo real a partir de una cámara web (Fig. 7). Las pruebas se realizaron en condiciones controladas de laboratorio, obteniéndose una velocidad de inferencia promedio de 30 FPS, lo que equivale aproximadamente a 33 ms por imagen procesada. Este resultado evidencia que el modelo mantiene un desempeño estable en escenarios de operación continua, cumpliendo con los requisitos de baja latencia necesarios para sistemas de inspección automatizada en líneas de producción de 30 FPS. Esto confirma que el modelo es lo suficientemente rápido para su integración en una línea de producción.

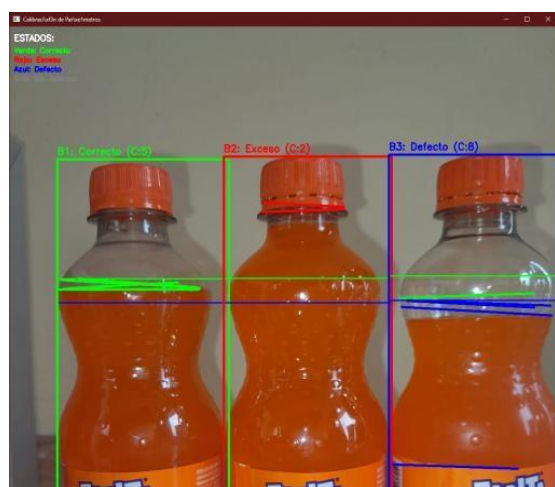


Fig. 7. Ejemplos de detección en tiempo real con cámara web

La arquitectura del sistema de control propuesto se estructura en tres niveles jerárquicos: adquisición de datos, procesamiento e inferencia, y ejecución del control. En el primer nivel, una cámara web o cámara industrial captura imágenes en tiempo real de la línea de producción, las cuales son transmitidas a un sistema de cómputo donde se ejecuta el modelo de visión

artificial.

En el segundo nivel, el modelo YOLOv11 procesa cada imagen en tiempo real utilizando Python y OpenCV, generando como salida la clase predicha de cada botella ("buen estado" o "mal estado") junto con su nivel de confianza. Este procesamiento se realiza de forma continua, permitiendo la detección instantánea de defectos en el flujo de producción.

La comunicación entre el sistema de visión artificial y el sistema de control se realiza mediante comunicación serial (UART/USB), enviando señales digitales desde el entorno de procesamiento hacia un controlador embebido. El controlador utilizado corresponde a un Arduino UNO o un PLC industrial Siemens S7-1200, dependiendo del nivel de implementación del sistema.

En el tercer nivel, el controlador ejecuta la lógica de decisión definida por el sistema. Si la clase detectada corresponde a una botella en condición defectuosa, se activa una salida digital que energiza un relé o válvula solenoide, la cual acciona un pistón neumático encargado de expulsar la botella de la línea de producción. En caso contrario, la botella continúa su recorrido sin intervención.

El tiempo de respuesta total del sistema, desde la detección por el modelo hasta la activación del actuador, se mantiene dentro de un rango de baja latencia adecuado para aplicaciones industriales en tiempo real, garantizando la viabilidad del sistema para su integración en procesos automatizados de control de calidad.

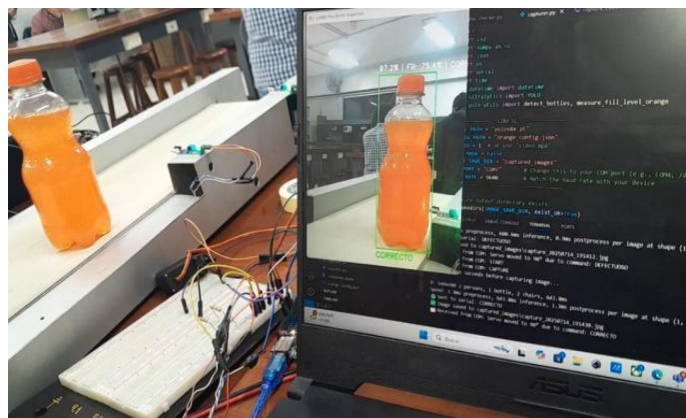


Fig. 8. Conexión de control automatizado

V. CONCLUSIONES

La investigación generó evidencia de la viabilidad técnica de utilizar un modelo YOLOv11 para la detección automatizada del estado de botellas en una línea de producción. El modelo exhibe un rendimiento prometedor en la identificación de "Botella-buen-estado" (87% de recall), validado mediante pruebas en tiempo real que confirman su capacidad para operar a velocidades industriales (30 FPS). No obstante, la detección de "Botella-mal-estado" (52% de recall) y la confusión crítica donde botellas defectuosas son aceptadas como buenas, representan el principal desafío a superar para garantizar un control de calidad robusto. La conceptualización del conexionado con sistemas de control PLC/Arduino establece una clara hoja de ruta para la automatización futura. Como trabajo futuro, se propone: 1) ampliar y diversificar significativamente el conjunto de datos, con un enfoque en la

recolección de más ejemplos de "Botella-mal-estado" que cubran una mayor variabilidad de defectos; 2) Explorar arquitecturas de detección más avanzadas para mejorar la precisión en la clase minoritaria; y 3) implementar y validar la interfaz física y lógica con el sistema de control en condiciones reales de producción. Estos pasos son cruciales para evolucionar de un prototipo funcional a un sistema de inspección automatizado, confiable y eficaz.

REFERENCIAS

- [1] E. Peralta Pari y C. H. Odar Olazabal, «Propuesta para incrementar la disponibilidad en la línea de producción en una planta embotelladora aplicando metodología TPM», Univ. Peru. Cienc. Apl. UPC, nov. 2024, Accedido: 18 de abril de 2025. [En línea]. Disponible en: <https://repositorioacademico.upc.edu.pe/handle/10757/682189>
- [2] L. E. Aparicio Pico, P. Devia Lozano y O. J. Amaya Marroquin, "Application of Deep Learning for the Identification of Surface Defects Used in Manufacturing Quality Control and Industrial Production: A Literature Review," *Rev. Ing.*, vol. 28, no. 1, 2023, doi: 10.14483/23448393.18934.
- [3] M. G. Céspedes Gómez y W. L. Alcazar Jauregui, «Mejora continua del proceso de producción aplicando Lean Manufacturing para incrementar la tasa de producción real de una planta embotelladora de agua de mesa», Univ. Peru. Cienc. Apl. UPC, nov. 2023, Accedido: 18 de abril de 2025. [En línea]. Disponible en: <https://repositorioacademico.upc.edu.pe/handle/10757/672088>
- [4] J. G. Madrid Pita, "Implementación de herramientas SMED y mantenimiento autónomo para incrementar disponibilidad en la línea de envasado 22 empresa AJEPER, periodo 2018 – 2019," Tesis de pregrado, Univ. San Ignacio de Loyola, Lima, Prov. Lima, Perú, 2021. <https://repositorio.usil.edu.pe/entities/publication/021aa84a-8137-4fa2-99e0-e658fa9a7c0d>
- [5] F. P. Vargas Portal, S. C. Reategui Ramirez y D. A. M. Caverro-Egusquiza Espinoza, "Participación de los líderes de AJE e ISM en LinkedIn y su impacto en el posicionamiento y diferenciación de las organizaciones. 2023-2024," Tesis de pregrado, Univ. Peruana de Ciencias Aplicadas (UPC), Lima, Prov. Lima, Perú, 2024.: <https://repositorioacademico.upc.edu.pe/handle/10757/682115>
- [6] W. J. M. de la Cruz y A. del J. L. Pillasagua, «rentabilidad financiera y su incidencia en la toma de decisiones de empresa embotelladoras de agua», *Rev. Ciencia y Desarrollo.*, vol. 27, n.o 3, Art. n.o 3, jul. 2024, doi: 10.21503/cyd.v27i3.2667.
- [7] I. P. M. S. Medina Haro, J. A. Vázquez López, M. I. M. Darío Hernández Ripalda, R. Rodríguez Vázquez, R. A. Grunberguer Reynoso y M. Á. Alvarado Vera, "Importancia del Empleo de las Redes Neuronales Artificiales para la Detección de Patrones en Gráficos de Control de Variables para Promedios (tipo $\bar{X}$)," *Rev. Ing. Ind.*, vol. 18, no. 1, p. 16, 2024. <https://static1.squarespace.com/static/55564587e4b0d1d3fb1eda6b/t/67462aeb46d64544fd8a1df4/1732651755631/MedinaHaro+--+II2024+---+16-24.pdf>
- [8] A. Orrasco y A. M. Negro Macho, *La implantación de la industria 4.0 en la viticultura: estudio de caso dentro de una bodega en la Ribera del Duero*. Valladolid, España: Universidad de Valladolid, Facultad de Comercio, 2023. [En línea]. Disponible en: <https://uvadoc.uva.es/handle/10324/62934>
- [9] J. C. Mendieta Villamizar, M. A. Olarte Daza, y H. González Acevedo, «Diseño de un sistema de control inteligente para un pasteurizador tipo túnel», *Prospectiva*, vol. 10, n.o 2 (julio-diciembre), pp. 81-88, 2012.
- [10] R. Rivero Mier, "Monitorizado del proceso de trefilado mediante adquisición de imágenes y uso de redes neuronales," Tesis de grado, Univ. de Cantabria, Santander, España, feb. 2025.. Disponible en: <https://repositorio.unican.es/xmlui/handle/10902/35718>
- [11] M. J. C. Samonte, T. Dong, M. Wu, y Y. Wang, «Detection of Defects in Bottled Liquor Using Deep Learning», en *Proceedings of the 2023 5th World Symposium on Software Engineering*, en WSSE '23. New York, NY, USA: Association for Computing Machinery, dic. 2023, pp. 292-297. doi: 10.1145/3631991.3632039.
- [12] M. Mao y M. Hong, «YOLO Object Detection for Real-Time Fabric Defect Inspection in the Textile Industry: A Review of YOLOv1 to YOLOv11», *Sensors*, vol. 25, n.o 7, p. 2270, abr. 2025, doi: 10.3390/s25072270.
- [13] L. Remmelzwaal, «Object Detection and Tracking for Crate and Bottle Identification in a Bottling Plant Using Deep Learning», *Artif. Intell. Appl.*, vol. 1, n.o 3, Art. n.o 3, abr. 2023, doi: 10.47852/bonviewAIA3202798.
- [14] S. Kang, Z. Hu, L. Liu, K. Zhang y Z. Cao, "Algoritmos YOLO de detección de objetos y sus aplicaciones industriales: descripción general y análisis comparativo," *Electronics*, vol. 14, no. 6, p. 1104, mar. 2025, doi: 10.3390/electronics14061104.. Disponible en: <https://www.mdpi.com/2079-9292/14/6/1104>