

DETECCIÓN EN TIEMPO REAL DE PAQUETES DEFECTUOSOS EN LÍNEAS DE PRODUCCIÓN DE PAPEL HIGIÉNICO MEDIANTE REDES NEURONALES CONVOLUCIONALES POR ANÁLISIS DE CUADRANTES

REAL – TIME DETECTION OF DEFECTIVE PACKAGES IN TOILET PAPER PRODUCTION LINES USING CONVOLUTIONAL NEURAL NETWORKS BY QUADRANT ANALYSIS

Israel Rodrigo Hoyos Borja¹, Marlon Danilo Basantes Valverde²

{hoyosisrael2015@gmail.com¹, mbasantes@unach.edu.ec²}

Fecha de recepción: 07/07/2026 / Fecha de aceptación: 24/07/2026 / Fecha de publicación: 05/08/2026

RESUMEN: Los rollos de papel higiénico son un producto de alta producción y comercialización en la industria ecuatoriana; la ausencia de rollos dentro de un paquete en una línea de producción genera pérdidas operativas. El presente artículo propone la implementación de un sistema de visión artificial basado en redes neuronales convolucionales con transfer learning sobre MobileNetV2, orientado a la detección automática de paquetes de papel higiénico en una jornada laboral. El núcleo metodológico presenta un enfoque de clasificación por cuadrantes de tres filas por dos columnas, lo que permite analizar individualmente los seis rollos que componen cada cuadrante. A partir de 100 imágenes originales, se construyó un conjunto de imágenes de 572 cuadrantes etiquetados, con una distribución balanceada. El modelo entrenado alcanzó un desempeño satisfactorio con una métrica de AUROC igual a 0.9978 y una exactitud del 97% en la validación durante el entrenamiento. Durante la fase de validación en condiciones representativas de producción y en un ambiente controlado, se diseñaron tres escenarios: uno con datos completamente defectuosos, otro con datos totalmente conformes y un tercero con selección aleatoria. Estos escenarios se ejecutaron mediante 38 simulaciones, número que corresponde a la capacidad productiva diaria (38 tirajes). En total, se analizaron 1996 imágenes, obteniendo resultados sobresalientes en recall de 1.00 y precisión de 1.00. Se realizó un análisis estadístico donde se obtuvo $p < 0.000001$, lo cual corroboró que el número de defectos detectados por el modelo

¹Universidad Nacional de Chimborazo, (UNACH) Dirección de Posgrado, Programa de Maestría en Matemática aplicada con mención en Matemática Computacional, <https://orcid.org/0009-0004-9333-3257>.

²Universidad Nacional de Chimborazo, (UNACH) Facultad de Ciencias, Escuela de Telecomunicaciones, grupo de Investigación GI(CT), <https://orcid.org/0000-0001-6011-1659>

coincide exactamente con el número real de defectos introducidos, validando con un nivel de confianza del 95%.

Palabras clave: *Redes neuronales convolucionales, MobileNetV2, Detección de defectos, Control de calidad industrial y Clasificación por cuadrantes*

ABSTRACT: Toilet paper rolls are a mass-produced and widely commercialized product in the Ecuadorian industry; the absence of rolls within a package on a production line generates operational losses. This article proposes the implementation of a computer vision system based on convolutional neural networks using transfer learning with MobileNetV2, aimed at the automatic detection of toilet paper packages during a work shift. The methodological core presents a quadrant-based classification approach with a three-row by two-column grid, allowing for the individual analysis of the six rolls that make up each quadrant. Based on 100 original images, a dataset of 572 labeled quadrants was constructed with a balanced distribution. The trained model achieved satisfactory performance with an AUROC metric of 0.9978 and 97% accuracy during training validation. During the validation phase, under representative production conditions and in a controlled environment, three scenarios were designed: one with completely defective data, another with totally conforming data, and a third with random selection. These scenarios were executed through 38 simulations, a number corresponding to the daily production capacity (38 runs). A total of 1996 images were analyzed, obtaining outstanding results with a recall and precision of 1.00. A statistical analysis was performed, obtaining $p < 0.000001$, which corroborated that the number of defects detected by the model exactly matches the actual number of defects introduced, validating the results with a 95% confidence level.

Keywords: *Convolutional neural networks, MobileNetV2, Defect detection, Industrial quality control and Quadrant-based classification*

INTRODUCCIÓN

El papel higiénico es un producto de alta rotación, y la presencia de paquetes incompletos o con defectos de sellado compromete su calidad establecida por la NORMA INEN 1430 y genera pérdidas costosas. La inspección manual tardía de paquetes incompletos en líneas de empaquetamiento a velocidad constante genera retrasos y, además, el riesgo de que los productos lleguen al mercado ecuatoriano con paquetes faltantes (1,2).

Ante esta problemática, los sistemas de visión artificial han demostrado resultados superiores a la inspección manual en múltiples escenarios industriales, procesando imágenes en tiempo real con consistencia que supera el desempeño humano (3, 4, 5). Sin embargo, la mayoría de los

estudios requieren miles de imágenes para entrenar modelos robustos (1, 3, 6), lo que representa una barrera para empresas con acceso limitado a datos etiquetados.

La visión artificial comprende técnicas computacionales que permiten interpretar información visual del entorno. Un sistema de visión industrial integra una cámara de captura, un módulo de preprocesamiento y de la decisión basada en la Norma Técnica Ecuatoriana (NTE) INEN 1430 y el Reglamento Técnico Ecuatoriano (RTE) INEN 132 que determina la conformidad del producto (2,7). OpenCV provee las primitivas de redimensionamiento, normalización y segmentación de imágenes en tiempo real (8).

Las CNN procesan datos con estructura espacial mediante operaciones de convolución, activación ReLU, pooling y capas densas. La pérdida de Binary Cross-Entropy para clasificación binaria se optimiza mediante el algoritmo Adam (9,10).

El transfer learning aprovecha representaciones aprendidas previamente para acelerar el entrenamiento con datos limitados. MobileNetV2 emplea convoluciones separables en profundidad y bloques residuales invertidos, preentrenada en ImageNet (11). En ambientes controlados, los filtros de bordes y texturas son transferibles al dominio de inspección (9,12).

Subdividir la imagen en regiones de interés y clasificar cada región independientemente reduce la varianza del problema, aumenta el número efectivo de muestras y permite localizar espacialmente el defecto (13).

El aumento excesivo sobre un conjunto base reducido puede inducir un sesgo de distribución sintética, donde la red memoriza transformaciones sin aprender representaciones genuinas (14,15). El reequilibrio natural de clases evita técnicas agresivas de sobremuestreo (16,17).

La pregunta científica que orienta el estudio es: ¿Cómo contribuirá el desarrollo de un sistema con visión artificial para la detección de paquetes defectuosos de empaquetado de papel higiénico a la calidad de su distribución al mercado?

El objetivo es diseñar un sistema de visión por cuadrantes basado en CNN para detectar paquetes defectuosos, entrenar el modelo de red neuronal con un conjunto de datos representativo de paquetes conformes y defectuosos y validar el sistema en condiciones reales de líneas de producción.

Esta investigación propone segmentar la imagen del paquete en cuatro cuadrantes, cada uno con una subgrilla de 3 filas $\times$ 2 columnas de rollos (6 rollos por cuadrante), lo que simplifica el problema de clasificación, multiplica el número de muestras reales y genera un dataset naturalmente balanceado (13).

MATERIALES Y MÉTODOS

La estrategia es de carácter cuasi-experimental, se basa en las condiciones reales de funcionamiento de la línea de producción en una jornada. Un tiraje produce 38 unidades, equivalentes a un paquete de 24 rollos. De acuerdo con los criterios de control de calidad establecidos, se debe inspeccionar un mínimo de 30 unidades por lote. Este proceso de revisión tiene una duración cercana a 40 segundos por lote, lo que corresponde a aproximadamente 0.66 minutos.

En este tiempo, la revisión de un tiraje completo de 38 unidades, implica alrededor de 25.33 minutos, lo que corresponde aproximadamente a 0.4222 horas. Si se considera un esquema laboral de dos turnos diarios de 8 horas cada uno (16 horas en total), se puede estimar que el sistema permite procesar aproximadamente 38 tirajes por día (resultado de dividir 16 horas entre 0.4222 horas por tiraje).

Los valores mencionados se sintetizan en la Tabla 1 y sirven como base para definir, por un lado, el tamaño de las muestras empleadas en cada validación (entre 31 y 38 imágenes) y, por otro, el número total de simulaciones realizadas. En este sentido, se ejecutan 38 repeticiones en el escenario de validación aleatoria, lo cual permite reproducir de manera coherente el comportamiento del sistema a lo largo de una jornada completa de producción.

Tabla 1. Justificación operativa del número de tirajes diarios (38) usado en el diseño experimental.

Parámetros	Valor
Unidades por tiraje	38 unidades (24 rollos cada uno)
Tiempo de inspección por tiraje	Aprox. 40 segundos (0.66 min)
Tiempo total de inspección (lote control)	25.33 min equivalente a 0.4222 h (30 unidades)
Jornada laboral	Dos turnos diarios de 8 h (16 horas)
Tirajes por jornada	Aprox. 38 tirajes por día
Producción diaria estimada	34656 ternas = 103968 rollos por día

Según la Tabla 1, el procesamiento diario de 38 tirajes equivale a inspeccionar un volumen de 103968 rollos de papel higiénico individuales. Estas cifras permiten dimensionar el alto volumen de trabajo que se debe gestionar en el sistema de inspección automática, el cual necesita mantener un desempeño estable y confiable a lo largo de toda la jornada de 16 horas.

Cada imagen completa fue segmentada en 4 cuadrantes (cuadrícula 2×2), cada uno con 3 filas y 2 columnas, es decir, 6 rollos por cuadrante). Para preservar la proporción real, cada imagen fue redimensionada a 448×672 px antes del corte, generando cuadrantes de 224×336 px escalados a 224×224 px para MobileNetV2. La clasificación del paquete completo se obtuvo como: paquetes conformes y, en caso contrario, paquetes defectuosos.

Cada una de las 100 imágenes originales fue segmentada en 4 cuadrantes, obteniendo 400 cuadrantes. Posteriormente se logró descartar manualmente cuadrantes no representativos, aquellos en los que no se evidenciaban los 6 rollos por cuadrante; se conservaron 143 cuadrantes base (80 conformes y 63 defectuosos). Sobre estos se aplicaron 4 transformaciones geométricas (original, rotación, espejo horizontal y espejo vertical), generando un conjunto final de 572 cuadrantes etiquetados (320 conformes y 252 defectuosos), que presentó una ratio de 1.27:1, prácticamente balanceada (18).

La arquitectura implementada con MobileNetV2 preentrenada en base de imágenes de ImageNet. El procesamiento se estructuró mediante una capa de agrupación global promedio (GlobalAveragePooling2D), seguida por capas densas conectadas con la función de activación ReLU. Para mitigar el sobreajuste, se configuraron con tasas de descarte (Dropout) de 0.4 y 0.3 respectivamente y finalmente se empleó la capa densa con activación Sigmoid.

El proceso de entrenamiento se ejecutó en dos etapas:

- Primera etapa: Consistió en tasa de aprendizaje $= 1 \times 10^{-4}$ durante 10 épocas
- Segunda etapa: fine-tuning (tasa de aprendizaje $= 1 \times 10^{-5}$, últimas 30 capas, EarlyStopping monitoreando val_recall).

Se diseñaron tres escenarios de validación sobre lotes de 31 a 38 imágenes (P1 a P8), compuestos por 30 imágenes conformes fijas más N imágenes adicionales:

- Validación 1: con N imágenes exclusivamente defectuosas (1 a 8)
- Validación 2: con N imágenes exclusivamente conformes (1 a 8)
- Validación 3: Se realizaron 38 simulaciones independientes (S01–S38) en las cuales N imágenes adicionales fueron seleccionadas de manera aleatoria a partir de un conjunto de 40 imágenes conformadas por 20 casos conformes y 20 defectuosos. Este procedimiento busca reproducir la variabilidad presente en los 38 tirajes que se generan diariamente en la producción real.

Para la evaluación del desempeño del modelo se consideraron los indicadores accuracy, precisión $= \frac{TP}{TP+FP}$, recall $= \frac{TP}{TP+FN}$, F1, así como las áreas bajo las curvas AUC-ROC y AUC-PR.

Adicionalmente, en el escenario de validación con selección aleatoria, se empleó la prueba no paramétrica de Wilcoxon de rangos de signo (19). Esta prueba permitió contrastar la hipótesis nula H_0 : el desempeño del modelo es menor o igual al observado en condiciones reales, frente a la hipótesis alternativa H_1 : el modelo presenta un desempeño superior, utilizando un nivel de significancia de $\alpha = 0.05$

RESULTADOS

Composición del dataset: El conjunto de datos final, compuesto por 572 cuadrantes, se dividió siguiendo una proporción 80/20 para los subconjuntos de entrenamiento y prueba (véase Tabla 2). Esta partición dio lugar a una relación de clases cercana a 1.27:1, lo que indica una distribución prácticamente equilibrada entre las categorías analizadas.

Tabla 2. Distribución de cuadrantes por conjunto.

Conjunto	Conformes	Defectuosos	Total	Proporción
Entrenamiento	256	202	458	80%
Validación	64	50	114	20%
Total	320	252	572	100%

Como se ilustra en la Figura 1, se presenta una distribución del conjunto de datos, conformado por 320 cuadrantes clasificados como conformes (55.9%) y 252 catalogados como defectuosos (44.1%). Esta composición corresponde a una relación cercana a 1.27:1, lo que refleja un equilibrio adecuado entre las clases y hace no necesario aplicar técnicas de sobremuestreo.

Figura 1. Distribución del Dataset de Cuadrantes
(3 filas × 2 columnas = 6 rollos por cuadrante)

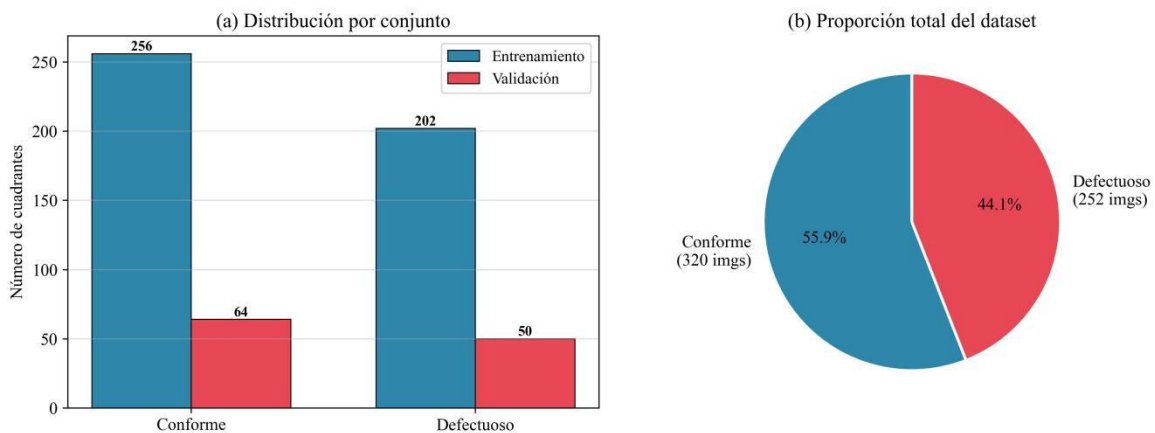


Figura 1. Distribución de cuadrantes: (a) por conjunto entrenamiento y validación, (b) proporción total.

Entrenamiento y métricas de evaluación: Durante la primera fase, que contempló 10 épocas con una tasa de aprendizaje de 1×10^{-4} , el modelo alcanzó un *val_accuracy* y *val_recall* de 97.37%. Posteriormente, en la Fase 2 de ajuste fino (*fine-tuning*), el uso de la estrategia *EarlyStopping* evidenció una convergencia en la época 8, confirmando que la arquitectura base MobileNetV2 había logrado transferir representaciones suficientemente discriminativas para la tarea.

La evaluación final realizada sobre el conjunto de validación compuesto por 114 cuadrantes, presentó los resultados según la Tabla 2.

Tabla 2. Métricas de evaluación del modelo por clase ($n = 114$).

Clase	Precisión	Recall	F1-Score	Soporte
Conforme	0.97	0.98	0.98	64
Defectuoso	0.98	0.96	0.97	50
Promedio ponderado	0.97	0.97	0.97	114
AUC-ROC	—	—	0.9978	—
AUC-PR	—	—	0.9973	—

La Figura 2 muestra la matriz de confusión donde el modelo clasificó correctamente 63 de los 64 cuadrantes conformes ($FP = 1$), mientras que 48 de las 50 eran defectuosos ($FN = 2$). Esto evidencia un margen de error global de apenas el 2.6% para todo el conjunto de prueba.

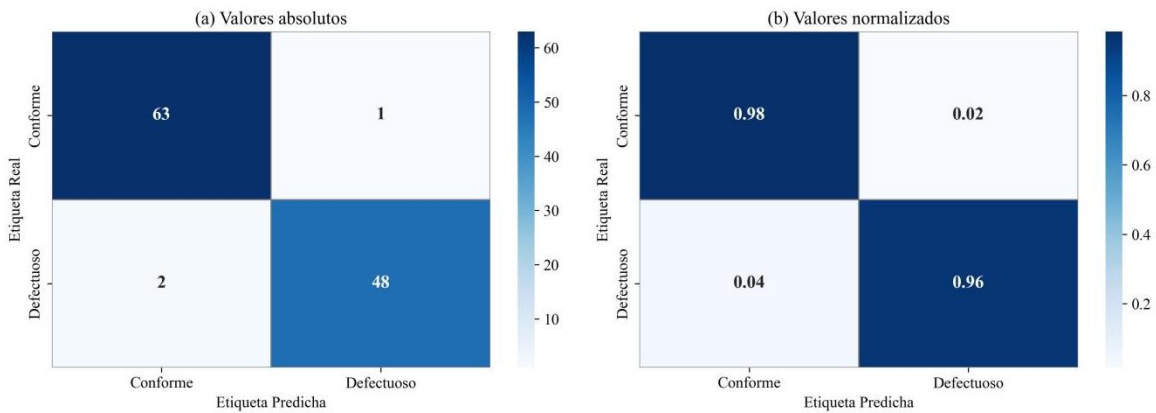


Figura 2. Matriz de confusión del modelo 3x2 donde literales: (a) valores absolutos y (b) valores normalizados.

Los gráficos estadísticos presentados en la Figura 3. Valores de 0.9978 para AUC-ROC y de 0.9973 para AUC-PR confirman la capacidad discriminativa excepcional. El umbral óptimo de entrenamiento (0.7641) se calibró a 0.5500 para producción real.

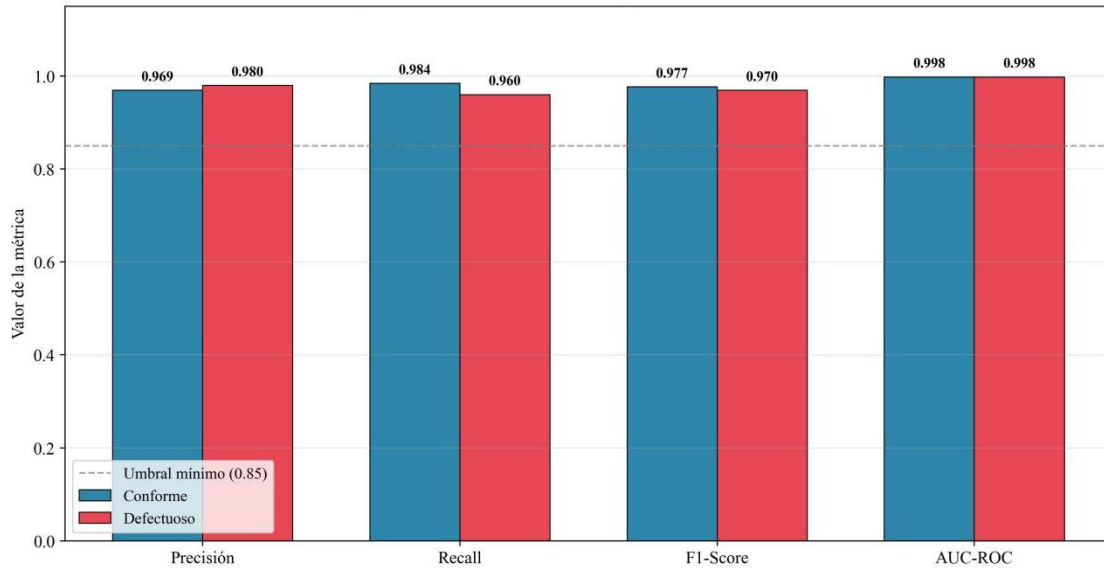


Figura 3. Métricas de evaluación por clase: precisión, recall, F1-Score y AUC-ROC para ambas clases.

Calibración del umbral de producción: La distribución de probabilidades sobre imágenes de producción evidenció separación perfecta de 0.17 a 0.53; defecto: 0.61 a 0.87. La Figura 4 muestra que con el umbral calibrado de 0.55 se alcanza Recall igual a 1.00 y Precisión de 1.00 simultáneamente, con $FP=FN=0$, mientras que el umbral de entrenamiento (0.7641) genera falsos negativos.

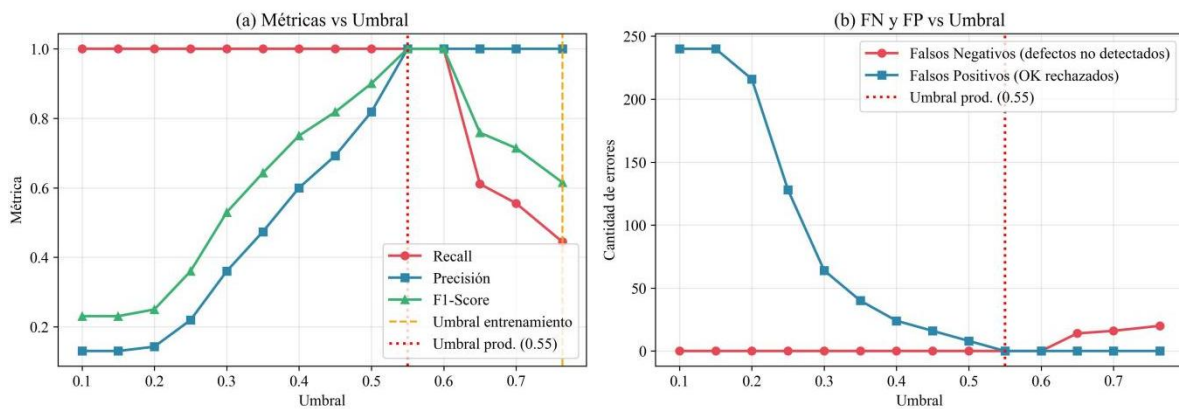


Figura 4. Calibración del umbral de clasificación para producción real: (a) métricas vs. umbral, (b) falsos negativos y positivos vs. Umbral.

Validación 1: Escenario con paquetes defectuosos exclusivamente

El primer escenario evaluó la detección cuando todas las imágenes adicionales a las 30 conformes fijas son defectuosas (1 a 8 por prueba, Tabla 3).

Tabla 3. Validación 1: resultados por prueba con paquetes defectuosos.

Prueba	Totales de imágenes	Defectuosas	TP	TN	FP	FN	Recall
prueba_1	31	1	1	30	0	0	1.00
prueba_2	32	2	2	30	0	0	1.00
prueba_3	33	3	3	30	0	0	1.00
prueba_4	34	4	4	30	0	0	1.00
prueba_5	35	5	5	30	0	0	1.00
prueba_6	36	6	6	30	0	0	1.00
prueba_7	37	7	7	30	0	0	1.00
prueba_8	38	8	8	30	0	0	1.00
TOTAL	276	36	36	240	0	0	1.00

El sistema identificó correctamente los 36 paquetes defectuosos ($TP = 36, FN = 0$) sin falsos positivos. Tal como se detalla en la Figura 5, se confirma el 100% de detección en las 8 pruebas (276 paquetes), con 36 defectuosos (13,0%) y 240 unidades en buen estado (87,0%).

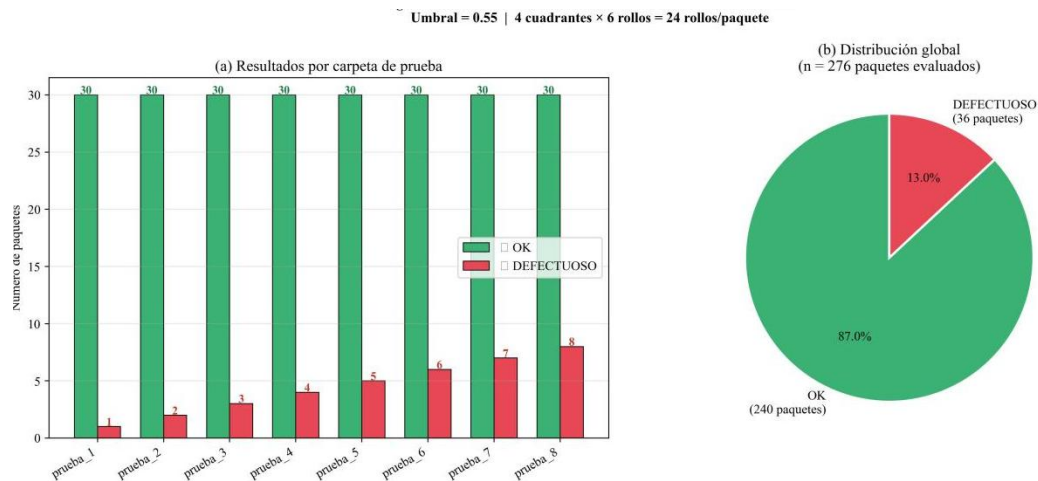


Figura 5. Resultados por prueba y distribución global (n = 276 paquetes).

Validación 2: Escenario con paquetes conformes exclusivamente

El segundo escenario evaluó la especificidad del sistema cuando todas las imágenes adicionales son conformes (Tabla 4).

Tabla 4. Validación 2: resultados por prueba con paquetes conformes.

Prueba	Totales de imágenes	Solo OK	TP	TN	FP	FN	Precisión
prueba_1	31	31	—	31	0	—	1.00
prueba_2	32	32	—	32	0	—	1.00
prueba_3	33	33	—	33	0	—	1.00
prueba_4	34	34	—	34	0	—	1.00
prueba_5	35	35	—	35	0	—	1.00
prueba_6	36	36	—	36	0	—	1.00
prueba_7	37	37	—	37	0	—	1.00
prueba_8	38	38	—	38	0	—	1.00
TOTAL	276	276	—	276	0	—	1.00

Respecto al segundo escenario, el sistema clasificó correctamente los 276 paquetes conformes (TN=276, FP=0). La Figura 6 confirma 100% de paquetes conformes (OK) sin ningún rechazo, con 0% de defectuosos detectados, garantizando que la eficiencia de la línea de producción no se ve comprometida por rechazos innecesarios.

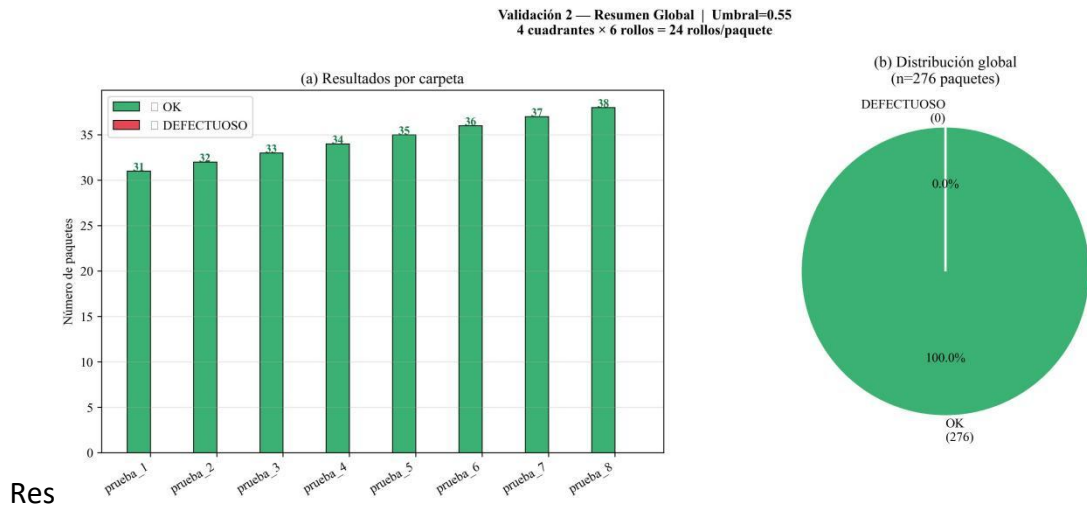


Figura 6. Resultados por prueba y distribución global paquetes, 100% OK).

Validación 3: Escenario aleatorio con 38 simulaciones

El tercer escenario, el más exigente, replicó los 38 tirajes diarios de producción mediante 38 simulaciones independientes (S01–S38). En cada simulación, las 8 imágenes adicionales de P8 se seleccionaron aleatoriamente de las imágenes de 40 (20 OK + 20 defectuosas), generando combinaciones distintas. La Figura 7 muestra la composición exacta de P8 en cada simulación: el

total se mantiene en 38 imágenes (30 fijas + 8 aleatorias), con una proporción variable de OK/defectuosas (panel b).

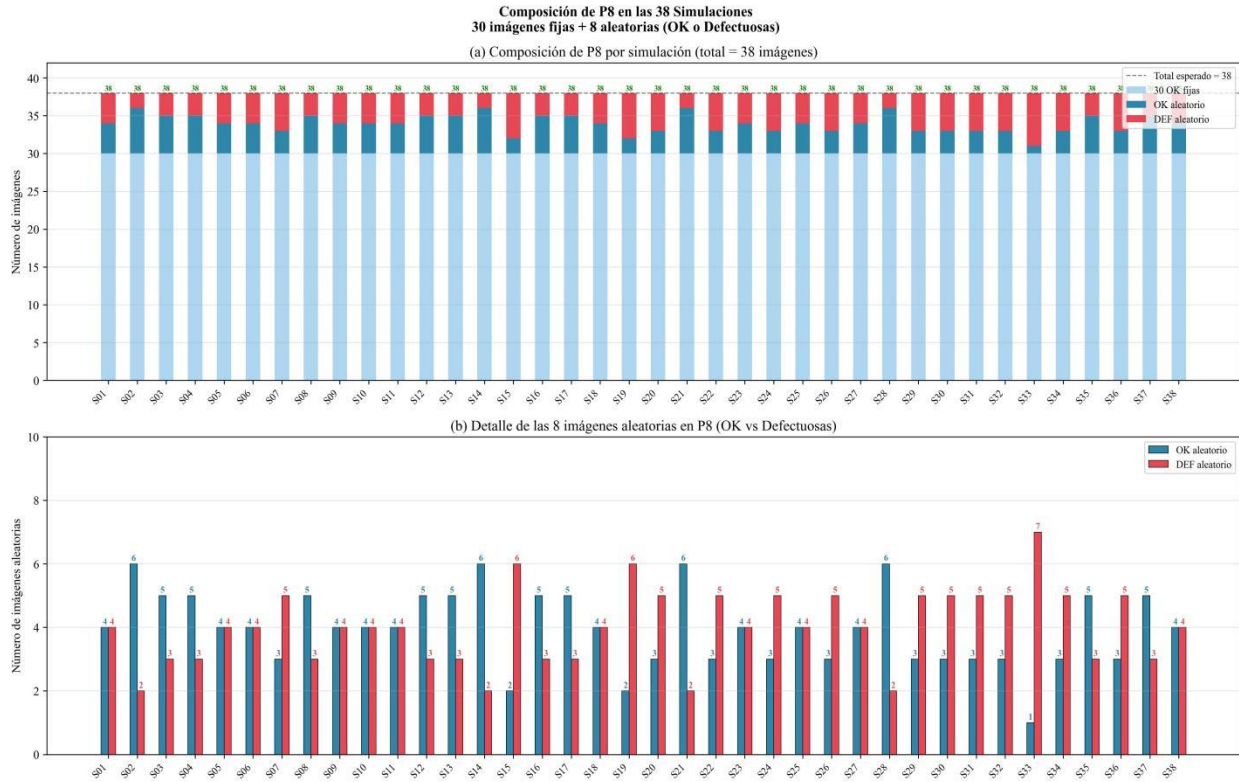


Figura 7. Composición de P8 en las 38 simulaciones: (a) total de 38 imágenes por simulación, (b) detalle OK vs. defectuosas aleatorias.

La Figura 8 presenta el mapa de calor del número de imágenes defectuosas por prueba (P1–P8) en cada una de las 38 simulaciones, evidenciando una distribución creciente y variable de 0 a 8 defectuosas que cubre ampliamente el espacio de combinaciones posibles.

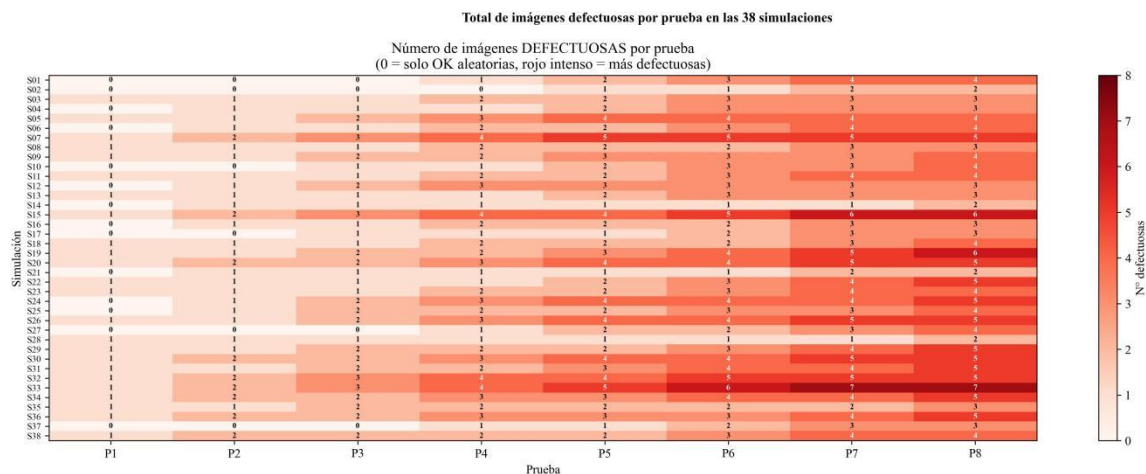


Figura 8. Mapa de calor del número de imágenes defectuosas por prueba (P1–P8) en las 38 simulaciones.

La Tabla 5 resume los resultados del experimento estadístico aplicado a las 38 simulaciones, comparando el número de defectuosas detectadas por el modelo frente al número real plantado (#Real) en cada simulación.

Tabla 5. Resultados del experimento estadístico — Validación 3 (38 simulaciones).

Métrica	Valor	Descripción
Simulaciones evaluadas	38	S01 a S38, cada una con P8 de 38 imágenes
Imágenes Aleatorias	30 fijas + 20 OK + 20 DEF	40 imágenes aleatorias disponibles por simulación
Simulaciones perfectas	38/38 (100%)	FP=0 y FN=0 en todas las simulaciones
Recall global	1.00	Todos los defectos detectados en las 38 simulaciones
Precisión global	1.00	Ningún paquete OK rechazado
Sumatoria Real	153	Totales de defectuosas plantadas en las 38 SIM (P8)
Sumatoria del Modelo	153	Totales de defectuosas detectadas por el modelo
Diferencia de media	0.0000	Modelo – Real, idéntica en cada simulación
Correlación r	1.0000	Correlación perfecta Modelo vs Real
Wilcoxon	W=0.0000 p<0.000001	Rechaza H ₀ por lo tanto H ₁ aceptada ($\alpha = 0.05$)

La Figura 6 confirma que Recall y Precisión se mantuvieron en 1.00 en las 38 simulaciones, sin falsos positivos ni negativos, con el 100% de simulaciones perfectas (38/38).

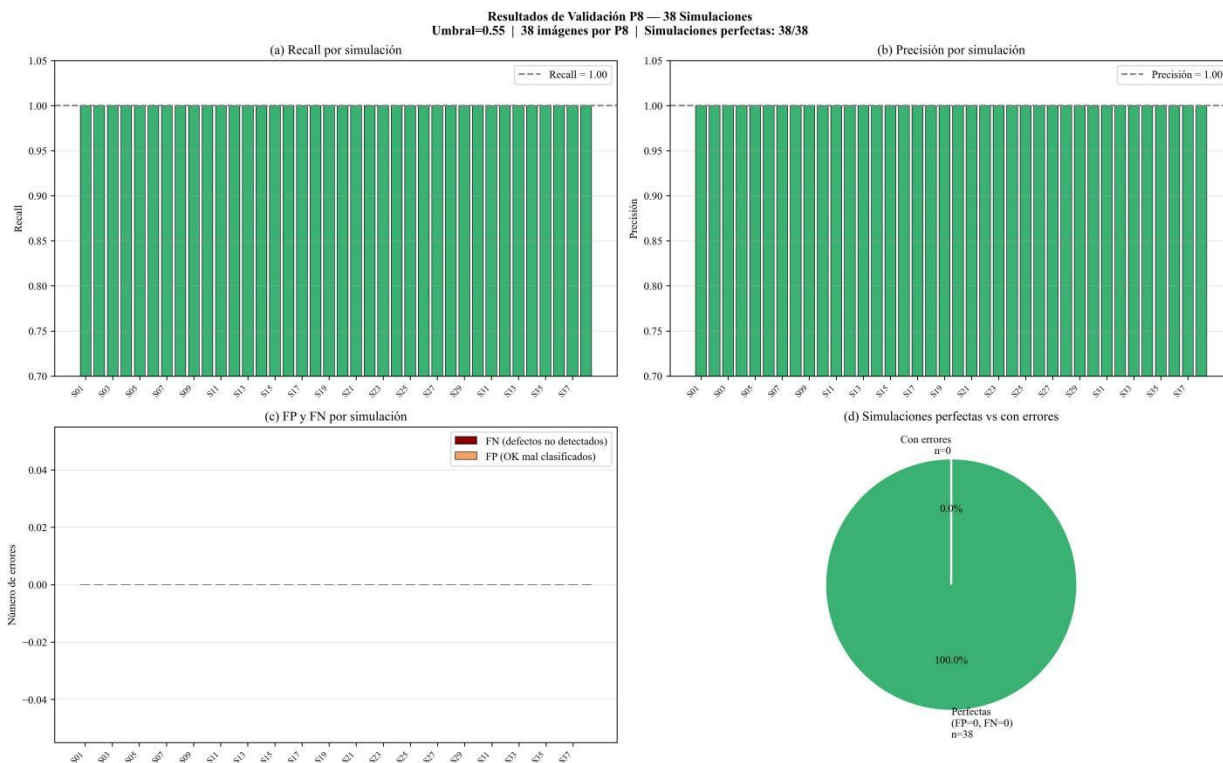


Figura 6. Resultados de validación P8 en las 38 simulaciones: (a) Recall por simulación, (b) Precisión por simulación, (c) FP y FN y (d) simulaciones perfectas vs con errores.

La sumatoria de defectuosas reales (Real=153) coincidió exactamente con la sumatoria detectada por el modelo (Modelo=153), con diferencia media de 0.0000 y correlación perfecta ($r = 1.0000$). La prueba de Wilcoxon de rangos con signo confirmó el rechazo de H_0 ($W = 0.0000, p < 0.000001$), aceptando H_1 : el modelo detecta consistentemente el número exacto de paquetes defectuosos plantados, validando estadísticamente al 95% de confianza la capacidad discriminativa del sistema.

Resumen comparativo de los tres escenarios de validación

Tabla 6. Comparación de los tres escenarios de validación del sistema.

Validación	n imágenes	Escenario	FP	FN	Recall y Precisión
Validación 1 (solo defectuosas)	276	P1 a P8 solo DEF	0	0	1.00
Validación 2 (solo OK)	276	P1 a P8 solo OK	0	—	1.00
Validación 3 (38 simulaciones aleatorias)	1444	30 fijas con (20 OK y 20 defectuosas)	0	0	1.00
TOTAL	1996	3 escenarios	0	0	1.00

La validación de los tres escenarios alcanzó de forma consistente Recall y Precisión iguales a 1.00, sin falsos positivos ni negativos en las 1996 imágenes acumuladas, lo que valida la robustez del sistema para su aplicación en la producción real.

DISCUSIÓN

El proceso metodológico partió de 100 imágenes originales; la segmentación en cuadrantes y la limpieza manual posterior conformaron un conjunto base de 143 imágenes representativas por cuadrante. Sobre este conjunto se aplicaron cuatro transformaciones geométricas (original, rotación, espejo horizontal y espejo vertical), generando un conjunto final de 572 cuadrantes etiquetados (320 conformes y 252 defectuosos), con una distribución prácticamente balanceada de 1.27: 1. Este enfoque permitió multiplicar el número de muestras y definir una tarea de clasificación más ordenada, centrada en verificar la presencia adecuada de seis rollos en cada cuadrante sin aplicar técnicas de sobremuestreo (13).

Tabla 9. Comparación con estudios relacionados de detección de defectos industriales.

Estudio	Arquitectura	Dataset	Accuracy	AUC-ROC
Sundaram & Zeid (2023)	ResNet50	Aprox 10000	96.0 %	0.971
Hou et al. (2024)	CNN custom	Aprox 5000	98.5 %	0.982
Mjahad & Rosado (2025)	ResNet50	Mayor a 140000	96.0 %	0.970
Presente estudio (2025)	MobileNetV2	572	97.0 %	0.9978

En el ámbito de la detección de defectos industriales, se observa que el presente estudio alcanza desempeños competitivos (20). Si se considera que se trabajó con un conjunto de datos mucho más reducido (572 cuadrantes) en comparación con otros estudios que emplean entre 5 000 y más de 140 000 imágenes, como se muestra en la Tabla 9 (3,6,21,22).

Este trabajo se ejecutó bajo un entorno controlado; esto contribuyó a reducir la variabilidad causada por factores externos. La estrategia de aprendizaje por transferencia (*transfer learning*) mediante MobileNetV2 (11) permitió aprovechar conocimientos previamente adquiridos por el modelo, optimizando el proceso de entrenamiento y el enfoque de segmentación por cuadrantes, que resultó una ventaja estructural al descomponer el problema en unidades más simples y homogéneas, lo que favoreció la estabilidad del modelo.

Durante el proceso de validación se realizó con 38 simulaciones; estas no responden a una elección aleatoria, sino que se encuentran vinculadas con la capacidad operativa diaria del sistema de producción. Este número representa, el máximo de tirajes que pueden procesarse en una jornada completa de 16 horas, considerando el tiempo requerido por cada unidad de inspección.

Desde el punto de vista metodológico, esto contribuye a la solidez de este estudio. Cada simulación representa un tiraje real dentro del proceso productivo. Dado el conjunto de 38 simulaciones, se presentan de manera consistente en condiciones de operación que el sistema enfrenta en la línea de producción, asegurando que los resultados obtenidos no solo sean estadísticamente sólidos, sino también representativos del contexto de la industria ecuatoriana.

El nivel de desempeño resulta importante y superior a los métodos tradicionales de inspección manual, en los cuales se reportan tasas de detección de errores que varían entre el 20% y el 30% (23). En este contexto, no haber registrado falsos positivos en los distintos escenarios presentados adquiere un valor significativo, ya que evita descartar innecesariamente productos que contribuyen a mantener la eficiencia operativa del sistema de producción.

Los resultados obtenidos son muy satisfactorios; sin embargo, es necesario reconocer algunas limitaciones durante este estudio: el conjunto de datos de imágenes se generó en condiciones controladas, utilizando una única configuración de captura de imágenes y la misma iluminación durante la jornada de producción. Esto implica que el modelo aún no ha sido puesto a prueba frente a la variabilidad de los entornos industriales reales y, por último, que el umbral de decisión definido podría requerir ajustes periódicos si cambian las condiciones del entorno, como la iluminación o la posición de los objetos. Considerando estos factores, será fundamental en futuras etapas asegurar el desempeño consistente del sistema.

CONCLUSIONES

Los resultados obtenidos evidencian un comportamiento estable del modelo en los tres escenarios de validación analizados, evaluando un total de 1996 imágenes que simulan una jornada productiva completa de 38 tirajes. Bajo estas condiciones, el sistema alcanzó niveles óptimos de desempeño, con valores de recall y precisiones iguales a 1.00, sin evidenciar errores de clasificación.

Desde la metodología de este trabajo, se presentó un enfoque basado en la división de imágenes en cuadrantes, sobre el cual se entrenó un modelo MobileNetV2 utilizando 572 muestras con una distribución balanceada. Este modelo logró un AUC-ROC aproximado de 0.9978 y fue sometido a diferentes validaciones en tres condiciones: datos completamente defectuosos, completamente conformes y escenarios con selección aleatoria ajustados a la dinámica real de producción. Adicionalmente, se realizó un contraste estadístico mediante la prueba de Wilcoxon; p es menor que 0.000001, lo que respalda la capacidad del modelo para detectar con exactitud los defectos introducidos.

Este estudio destaca la aplicación de segmentación estructurada como estrategia para simplificar el aprendizaje del modelo, seguida de la definición del umbral de decisión (0.55), respaldada por

las condiciones controladas, la formulación de un esquema de validación alineado con la capacidad operativa diaria mediante 38 simulaciones, y la validación estadística de los resultados, que confirma la correspondencia entre las predicciones y los defectos reales observados.

Como líneas futuras, se plantea avanzar hacia la integración del sistema con un controlador lógico programable (PLC), lo que permitiría automatizar la separación de productos defectuosos en tiempo real, así como su implementación en plataformas de hardware embebido. Donde se considera pertinente extender esta metodología a otros procesos de empaquetado que presenten estructuras similares.

REFERENCIAS BIBLIOGRÁFICAS

1. Hou M, Li P, Cheng S, Yv J. CNN-based defect detection in manufacturing. *Adv Control Appl Eng Ind Syst.* 2024;6(4):e196. doi:10.1002/ad2.196
2. Instituto Ecuatoriano de Normalización. NTE INEN 1430: Papeles y cartones. Papel higiénico. Requisitos [Internet]. 2a rev. Quito: INEN; 2015 [citado 2026 Jul 12]. Disponible en: <https://es.scribd.com/document/597487555/1430-NORMA-ECUATORIANA-PAPEL-HIGIENICO-REQUISITOS>
3. Hütten N, Alves Gomes M, Hölken F, Andricevic K, Meyes R, Meisen T. Aprendizaje profundo para la inspección visual automatizada en la fabricación y el mantenimiento: una revisión de artículos de acceso abierto. *Appl Syst Innov.* 2024;7(1):11. doi:10.3390/asi7010011
4. Zhou S, Yao S, Shen T, Wang Q. A novel end-to-end deep learning framework for chip packaging defect detection. *Sensors.* 2024;24(17):5837. doi:10.3390/s24175837
5. Lee J, Kim S, Park I. Computer-vision-based product quality inspection and novel counting system. *Appl Syst Innov.* 2024;7(6):127. doi:10.3390/asi7060127
6. Sundaram S, Zeid A. Artificial intelligence-based smart quality inspection for manufacturing. *Micromachines.* 2023;14(3):570. doi:10.3390/mi14030570
7. Subsecretaría de la Calidad. Reglamento técnico ecuatoriano RTE INEN 132 (1R): Productos absorbentes de papel [Internet]. Quito: Ministerio de Producción, Comercio Exterior, Inversiones y Pesca; 2019 [citado 2026 Mar 17]. Resolución N° 2019-068. Disponible en: <http://extranet.comunidadandina.org/sirt/sirtDocumentos/ECOTCP19090.pdf>
8. Zeng L, Wan F, Zhang B, Zhu X. Inspección visual automatizada para la detección y clasificación precisa de defectos en insertos de CBN. *Sensors.* 2024;24(23):7824. doi:10.3390/s24237824
9. Guan J, Fei J, Li W, Jiang X, Wu L, Liu Y, et al. Improving industrial quality control: a transfer learning approach to surface defect detection. *Sensors.* 2025;25(2):527. doi:10.3390/s25020527
10. Liu J, Wang C, Gao J, Guo C. Review of image classification algorithms based on convolutional neural networks. *Remote Sens.* 2021;13(22):4712. doi:10.3390/rs13224712

11. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: inverted residuals and linear bottlenecks. En: Proceedings CVPR; 2018. p. 4510-4520. doi:10.1109/CVPR.2018.00474
12. Abdul Majeed A, Ateeq M, Hu B, Isa W, Omar Z, Chen W. Feature-based transfer learning for IoT-enabled defect detection. En: Advances in Intelligent Manufacturing and Robotics. Singapore: Springer; 2024. p. 455-467. doi:10.1007/978-981-99-8498-5_36
13. Gour A, Bhanodia PK, Sethi KK, Rajput S. Novel framework for image classification based on a patch-based CNN model. En: Advances in Data and Information Sciences. Singapore: Springer; 2024. p. 235-245. doi:10.1007/978-981-99-6547-2_25
14. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data. 2019;6(1):60. doi:10.1186/s40537-019-0197-0
15. Yang S, Xiao W, Zhang M, Guo S, Zhao J, Furao S. Image data augmentation for deep learning: a survey. ArXiv [Internet]. 2022 [citado 2026 Jul 12]; abs/2204.08610. Disponible en: <https://arxiv.org/abs/2204.08610>
16. Bochnak P, Kozłowski M. On-the-fly image-level oversampling for imbalanced datasets of manufacturing defects. Mach Learn. 2024;113:1245-1268. doi:10.1007/s10994-023-06498-4
17. Lango M, Stefanowski J. Effective class-imbalance learning based on SMOTE and convolutional neural networks. Appl Sci. 2023;13(6):4006. doi:10.3390/app13064006
18. Wu Z, Chen L. Image data augmentation techniques based on deep learning: a survey. Math Biosci Eng. 2024;21(3):272-295. doi:10.3934/mbe.2024272
19. García S, Herrera F, Luengo J. Pruebas no paramétricas avanzadas para comparaciones múltiples en el diseño de experimentos en inteligencia computacional y minería de datos. Inf Sci. 2023;620:634-660. doi:10.1016/j.ins.2022.11.076
20. Reddy KS, Prasad V. Using CNN for material surface quality inspection and classification. Int J Comput Intell Syst. 2025;18:951. doi:10.1007/s44196-025-00951-z
21. Bajaan L, Paladines J. Diseño e implementación de un prototipo clasificador de granos de cacao usando visión artificial y machine learning [tesis de grado]. Guayaquil: Universidad Politécnica Salesiana; 2021.
22. Mjahad A, Rosado-Muñoz A. Robust industrial surface defect detection using statistical feature extraction and capsule network architectures. Sensors. 2025;25(19):6063. doi:10.3390/s25196063
23. Cheong LK, Hodgson AJ, Cheng AHC, Lee A, Wong A. SolderNet: hacia una inspección visual confiable de uniones de soldadura en la fabricación de productos electrónicos mediante inteligencia artificial explicable. En: Proceedings of the AAAI Conference on Artificial Intelligence. 2023;37(13):15725-15732. doi:10.1609/aaai.v37i13.26865