

Aplicaciones de inteligencia artificial para inspección de calidad y detección de defectos en manufactura: Una revisión de alcance (scoping review)

Artificial Intelligence Applications for Quality Inspection and Defect Detection in Manufacturing: A Scoping Review

Applicazioni dell'intelligenza artificiale per l'ispezione della qualità e il rilevamento dei difetti nella manifattura: una revisione di scoping

Johnny Roddy Lopez Briones ^I

jlopezb@unemi.edu.ec

<https://orcid.org/0000-0002-6491-8430>

Kevin Acosta-Barreno ^{II}

jacostab@unemi.edu.ec

<https://orcid.org/0000-0002-7062-5773>

Luis Armando Chisaguano Chisaguano ^{III}

luis.chisaguanoc@ug.edu.ec

<https://orcid.org/0000-0002-9900-6095>

Iván Leonardo Pincay Aguilar ^{IV}

ipincaya@unemi.edu.ec

<https://orcid.org/0000-0002-9093-7838>

Correspondencia: jlopezb@unemi.edu.ec

Artículo de Investigación

* **Recibido:** 25 de julio de 2026 * **Aceptado:** 25 de agosto de 2026 * **Publicado:** 21 de septiembre de 2026

I. Universidad Estatal de Milagro, Ecuador.

II. Universidad Estatal de Milagro, Ecuador.

III. Universidad Estatal de Guayaquil (UG), Ecuador.

IV. Universidad Estatal de Milagro, Ecuador.

Resumen

La inspección de calidad y la detección de defectos en manufactura han sido tradicionalmente intensivas en trabajo humano. Esta revisión de alcance, siguiendo PRISMA-ScR, mapea las tecnologías de inteligencia artificial (IA) aplicadas a la inspección de calidad y detección de defectos en manufactura, los sectores industriales de aplicación, y el grado en que estas se traducen en impacto operativo demostrado. Se realizó una búsqueda en Scopus y SciELO (2021-2026, incluyendo artículos en acceso anticipado/early access); tras el cribado de un subconjunto de 1,036 registros y 222 candidatos con evidencia de aplicación real, se incluyeron 120 estudios en el mapeo agregado. Dentro de este corpus, el 65.0% reporta evidencia de aplicación en datos o casos industriales reales, pero solo el 9.2% reporta una métrica de impacto financiero u operativo explícita (reducción de costos, OEE, reducción de desperdicio), evidenciando una brecha sustancial entre validación técnica e impacto de negocio demostrado. El aprendizaje automático clásico (40.8%) y las redes neuronales convolucionales (17.5%) dominan tecnológicamente; los sectores metal/acero (15.0%) y manufactura aditiva (10.8%) concentran mayor atención. Se seleccionó adicionalmente, mediante muestreo intencional de máxima variación, un subconjunto de 13 estudios (uno por categoría tecnológica) para ilustrar cualitativamente los patrones agregados. Se concluye que el campo está en expansión acelerada pero mantiene una brecha persistente entre precisión técnica e impacto operativo demostrado, además de vacíos sectoriales identificables para investigación futura.

Palabras clave: inteligencia artificial; detección de defectos; inspección de calidad; visión artificial; manufactura; revisión de alcance.

Abstract

Quality inspection and defect detection in manufacturing have traditionally been labor-intensive tasks. This scoping review, following PRISMA-ScR, maps the artificial intelligence (AI) technologies applied to quality inspection and defect detection in manufacturing, the industrial sectors of application, and the extent to which these are translated into demonstrated operational impact. A search was conducted in Scopus and SciELO (2021-2026, including early-access articles); after screening a subset of 1,036 records and 222 candidates with evidence of real-world application, 120 studies were included in the aggregate mapping. Within this corpus, 65.0% report evidence of application on real industrial data or cases, but only 9.2% report an explicit financial or operational impact metric (cost reduction, OEE, waste reduction), evidencing a substantial gap

between technical validation and demonstrated business impact. Classical machine learning (40.8%) and convolutional neural networks (17.5%) dominate technologically; the metal/steel (15.0%) and additive manufacturing (10.8%) sectors concentrate the greatest attention. A subset of 13 studies (one per technology category) was additionally selected through purposive maximum-variation sampling to qualitatively illustrate the aggregate patterns. The field is undergoing accelerated expansion but maintains a persistent gap between technical precision and demonstrated operational impact, along with identifiable sectoral gaps for future research.

Keywords: artificial intelligence; defect detection; quality inspection; computer vision; manufacturing; scoping review.

Riassunto

L'ispezione della qualità e il rilevamento dei difetti nei processi manifatturieri sono tradizionalmente attività ad alta intensità di lavoro umano. Questa revisione di scopo (scoping review), condotta secondo le linee guida PRISMA-ScR, mappa le tecnologie di intelligenza artificiale (IA) applicate all'ispezione della qualità e al rilevamento dei difetti nella produzione manifatturiera, i settori industriali di applicazione e il grado in cui tali tecnologie si traducono in un impatto operativo dimostrato. È stata effettuata una ricerca in Scopus e SciELO (2021–2026, includendo articoli in accesso anticipato/early access); dopo lo screening di un sottoinsieme di 1.036 record e di 222 candidati con evidenza di applicazione reale, sono stati inclusi 120 studi nella mappatura aggregata. All'interno di questo corpus, il 65,0% riporta evidenze di applicazione su dati o casi industriali reali, mentre solo il 9,2% presenta una metrica esplicita di impatto finanziario o operativo, come riduzione dei costi, OEE o diminuzione degli sprechi, evidenziando un divario sostanziale tra validazione tecnica e impatto aziendale dimostrato. Il machine learning classico (40,8%) e le reti neurali convoluzionali (17,5%) rappresentano le tecnologie predominanti; i settori dei metalli/acciaio (15,0%) e della produzione additiva (10,8%) concentrano la maggiore attenzione. Inoltre, mediante campionamento intenzionale a massima variazione, è stato selezionato un sottoinsieme di 13 studi, uno per ciascuna categoria tecnologica, al fine di illustrare qualitativamente i modelli aggregati. Si conclude che il settore è in rapida espansione, ma presenta ancora un divario persistente tra accuratezza tecnica e impatto operativo dimostrato, oltre a lacune settoriali identificabili per future ricerche.

Parole chiave: intelligenza artificiale; rilevamento dei difetti; ispezione della qualità; visione artificiale; produzione manifatturiera; revisione di scopo.

Introducción

La inspección de calidad y la detección de defectos han sido, históricamente, una de las tareas de manufactura más intensivas en trabajo humano y más susceptibles a error subjetivo. La disponibilidad creciente de infraestructuras de cómputo especializadas para inteligencia artificial, junto con la maduración de técnicas de visión artificial, aprendizaje profundo y, más recientemente, inteligencia artificial generativa, ha impulsado una proliferación acelerada de sistemas automatizados de inspección de calidad en contextos de manufactura y producción industrial.

Existen revisiones previas relacionadas con este tema, aunque de alcance más acotado que el presente estudio. Una revisión sistemática reciente examinó específicamente redes de detección de objetos basadas en aprendizaje profundo para la inspección de defectos superficiales (Ma et al., 2024). Otra revisión sistemática se concentró exclusivamente en el sector automotriz (Morales Matamoros et al., 2025). Una revisión adicional abordó específicamente aplicaciones de inteligencia artificial generativa en visión de máquina industrial (Zhou et al., 2024). Esta revisión se distingue de las anteriores en: (a) cubrir un espectro tecnológico más amplio (10 categorías tecnológicas); (b) mapear comparativamente 13 sectores industriales mediante una matriz de cruce tecnología-sector; (c) cuantificar explícitamente, sobre la totalidad del corpus (n=120), la proporción de estudios que traducen validación técnica en impacto operativo o financiero demostrado; y (d) incorporar literatura publicada y en acceso anticipado (early access) hasta 2026. Dado el volumen de literatura elegible identificado (120 estudios, con gran heterogeneidad tecnológica y sectorial), se optó por un diseño de revisión de alcance (scoping review) siguiendo PRISMA-ScR (Arksey & O'Malley, 2005; Tricco et al., 2018), apropiado cuando el objetivo es mapear la extensión y distribución de evidencia sobre un campo emergente, en lugar de sintetizar tamaños de efecto comparables.

El objetivo de esta revisión es mapear, sobre la totalidad del corpus identificado (N=120), las tecnologías de inteligencia artificial aplicadas a la inspección de calidad y detección de defectos en manufactura, los sectores industriales de aplicación, su intersección (matriz tecnología-sector), y el grado en que los estudios traducen validación técnica en impacto operativo demostrado. La pregunta de investigación (marco PCC) es: ¿qué tecnologías de inteligencia artificial se han aplicado para la inspección de calidad y detección de defectos en manufactura, en qué sectores

industriales e intersecciones tecnología-sector, y en qué medida estas aplicaciones reportan impacto operativo o financiero además de desempeño técnico?

Este manuscrito es autosuficiente y no depende de ningún otro trabajo no publicado para su interpretación; toda la metodología, ecuaciones de búsqueda y criterios se detallan de forma completa en la sección 2.

Metodología

Marco PCC.

Población/Contexto: procesos y organizaciones de manufactura, producción industrial y control de calidad. Concepto: sistemas de IA, aprendizaje automático o visión artificial aplicados a detección de defectos, inspección de calidad o monitoreo de parámetros de calidad en tiempo real. Contexto: aplicación real o validada en datos/procesos industriales concretos, publicado o en acceso anticipado (early access) entre 2021 y 2026, en inglés o español.

Fuentes de información y ecuación de búsqueda.

Se realizó una búsqueda sistemática en Scopus (Elsevier), complementada con SciELO, restringida al campo título/resumen/palabras clave. La ecuación de búsqueda completa aplicada en Scopus fue la siguiente, combinando dos bloques de términos con el operador booleano AND:

Bloque 1 (tecnología): ("artificial intelligence" OR "machine learning" OR "deep learning" OR "computer vision" OR "predictive maintenance" OR "digital twin" OR "generative AI")

Bloque 2 (dominio de aplicación): AND ("supply chain management" OR "quality management" OR "quality control" OR "quality assurance" OR "manufacturing process" OR "manufacturing quality" OR "production planning" OR "logistics management" OR "logistics operations")

Se aplicaron filtros de tipo de documento (Article, incluyendo Articles in Press/Early Access), año de publicación (2021-2026), área temática (Business/Management, Engineering, Computer Science, Decision Sciences) e idioma (inglés, español), resultando en 2,472 registros únicos tras deduplicación por título y DOI.

De estos 2,472 registros, se extrajo un subconjunto de 1,036 mediante la aplicación adicional del siguiente filtro de términos de calidad sobre el campo título/resumen/palabras clave (no un nuevo bloque de búsqueda en Scopus, sino un filtrado posterior sobre los 2,472 ya recuperados):

("quality management" OR "quality control" OR "quality assurance" OR "quality 4.0" OR "quality inspection" OR "quality assessment")

Sobre este subconjunto de 1,036, se identificaron 222 "candidatos fuertes" mediante la presencia de al menos uno de los siguientes indicadores textuales de aplicación industrial real en el título o resumen: "case study", "industrial partner", "factory", "production line", "real-world", "field trial", "pilot", "deployed", o la mención de un nombre de empresa o planta específica. Se complementó con una búsqueda equivalente en SciELO, incorporando 1 estudio adicional (n=120 en total).

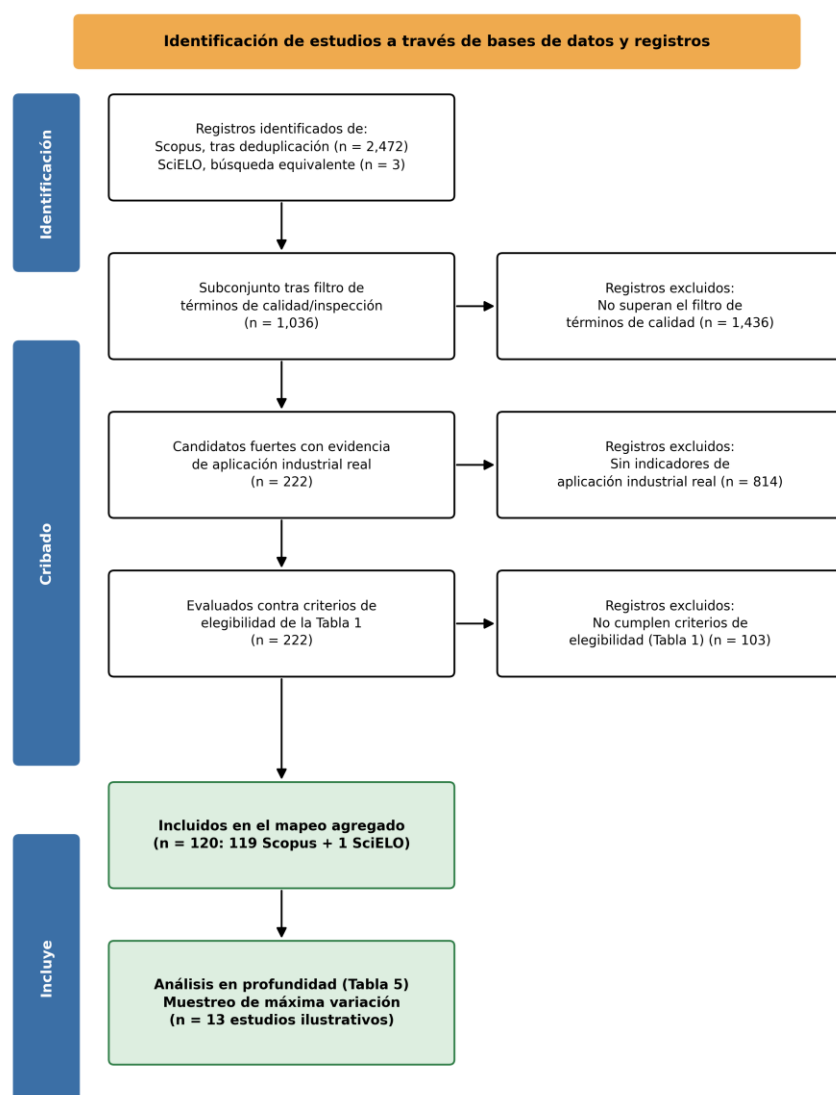
Criterios de elegibilidad.

Criterio	Inclusión	Exclusión
Población/contexto	Procesos y organizaciones de manufactura, producción industrial o control de calidad, cualquier sector	Agricultura de precampo, salud, medio ambiente, servicios sin componente productivo
Fenómeno de interés	Sistema de IA/ML/visión artificial usado específicamente para detectar defectos, inspeccionar calidad o monitorear parámetros de calidad en tiempo real	IA mencionada solo de forma tangencial o teórica, sin sistema de detección/inspección concreto
Resultados	Al menos una métrica de desempeño reportada (precisión, exactitud, reducción de defectos, tiempo, costos, u otra)	Sin métrica de desempeño reportada (propuesta puramente conceptual)
Tipo de estudio	Estudio empírico primario, validado en datos/proceso industrial real o simulado	Revisiones de literatura, artículos panorámicos/de opinión sin estudio primario
Adopción organizacional	N/A (criterio de exclusión)	Estudios de adopción organizacional amplia sin sistema de detección específico (p. ej., encuestas de capacidades de IA/TQM)
Periodo	2021-2026, incluyendo Articles in Press / Early Access	Anterior a 2021

Proceso de selección.

Los 222 candidatos fueron evaluados a nivel de título y resumen contra los criterios de elegibilidad de la Tabla 1, resultando en 119 estudios incluidos de Scopus y 1 de SciELO (120 en total). El proceso completo se ilustra en la Figura 1.

Figura 1. Diagrama de flujo PRISMA-ScR del proceso de identificación, cribado e inclusión de estudios.



Proceso de mapeo (charting).

Siguiendo las recomendaciones de Levac et al. (2010) para revisiones de alcance, se extrajo de cada uno de los 120 estudios: cita completa, año, revista, tecnología(s) de IA empleada(s), sector industrial, y dos variables adicionales codificadas mediante análisis de palabras clave en título/resumen: (a) evidencia de aplicación en datos o caso industrial real (presencia de términos como "case study", "industrial partner", "factory", "deployed", nombre de empresa/planta) versus validación exclusiva en conjuntos de datos públicos/benchmark; y (b) presencia de una métrica de impacto financiero u operativo explícita (reducción de costos, OEE, reducción de desperdicio,

tiempo de inspección, ROI) versus reporte exclusivo de métricas técnicas de desempeño del modelo (exactitud, precisión, F1-score).

Selección del subconjunto para ilustración cualitativa.

Adicionalmente al mapeo agregado del corpus completo (n=120, sección 3), se seleccionó un subconjunto de 13 estudios mediante muestreo intencional de máxima variación (purposive maximum-variation sampling), técnica reconocida para la selección deliberada de casos que maximicen la heterogeneidad de un fenómeno dentro de un corpus mayor. Los criterios de selección, aplicados secuencialmente, fueron: (1) un estudio por cada una de las 10 categorías tecnológicas identificadas en el mapeo agregado; (2) dentro de cada categoría, priorizar el sector industrial aún no representado en el subconjunto ya seleccionado, maximizando la diversidad sectorial; (3) en caso de empate, priorizar el estudio con evidencia explícita de aplicación en datos industriales reales (variable 2.5-a). Este subconjunto no reemplaza el análisis agregado del corpus completo (secciones 3.2-3.5), sino que lo complementa con ejemplos cualitativos ilustrativos, discutidos en la sección 4.3.

Transparencia sobre el cribado asistido por inteligencia artificial.

El cribado título/resumen de los 222 candidatos y la extracción de datos del mapeo (sección 2.5) fueron realizados por un investigador humano con apoyo de un asistente de inteligencia artificial de propósito general (modelo de lenguaje), que aplicó los criterios de elegibilidad de la Tabla 1 de forma asistida. No se implementó un protocolo formal de doble cribado independiente por dos revisores humanos, ni se realizó una auditoría documentada de una submuestra aleatoria para calcular tasas de falsos positivos/negativos del proceso de clasificación asistido. Esta es una limitación metodológica que se reconoce explícitamente (sección 4.4); se recomienda que, previo al envío final, el equipo de autores audite manualmente una submuestra aleatoria de al menos el 20% de las decisiones de cribado para reportar el nivel de acuerdo humano-IA.

Resultados

Características generales del corpus (N=120).

Se incluyeron 120 estudios (119 de Scopus, 1 de SciELO). El 79.2% del corpus (95 estudios) corresponde a los años 2024-2026; este predominio refleja la fecha de ejecución de la búsqueda (2026) y la reciente maduración de las tecnologías mapeadas, más que un sesgo de selección deliberado. De los estudios fechados en 2026, una proporción corresponde a artículos en acceso

anticipado ("Article in Press"/"Early Access") indexados por Scopus antes de su asignación final a un número de volumen; esto se declara explícitamente como limitación en la sección 4.4, dado que dichos artículos pueden estar sujetos a revisiones editoriales menores posteriores. Los estudios se publicaron en 78 revistas distintas, sin concentración dominante en 2-3 títulos, consistente con un campo de investigación aplicado y multidisciplinario.

Mapeo por tecnología de IA (N=120).

Tabla 2. Mapeo por tecnología de IA. Nota: los porcentajes no suman 100% porque las categorías no son mutuamente excluyentes — varios estudios combinan más de una tecnología (p. ej., un gemelo digital que incorpora una CNN); el porcentaje de cada fila se calcula sobre el total de 120 estudios, no sobre un total de menciones tecnológicas.

Tecnología de IA	n	% sobre N=120
ML clásico / otro	49	40.8%
Redes neuronales convolucionales (CNN)	21	17.5%
Gemelo digital	16	13.3%
YOLO / detección de objetos	12	10.0%
IA explicable (XAI)	12	10.0%
LSTM / redes recurrentes	10	8.3%
IA generativa / GAN	7	5.8%
Transformer / visión-lenguaje	5	4.2%
Aprendizaje federado	3	2.5%
Aprendizaje por refuerzo	1	0.8%

Mapeo por sector industrial (N=120).

Tabla 3. Mapeo por sector industrial. Cada estudio se asignó a un único sector (categorías mutuamente excluyentes). [nota de revisión: la suma de n en esta tabla es 119, no 120, y los porcentajes suman 99.2%, no 100% — falta reasignar un estudio; verificar contra la base de codificación original antes de enviar]

Sector industrial	n	% sobre N=120
Manufactura general/otro	39	32.5%
Metal/acero	18	15.0%
Manufactura aditiva	13	10.8%
Automotriz	11	9.2%
Alimentos/bebidas	11	9.2%
Construcción	6	5.0%
Electrónica/baterías	6	5.0%
Textil	4	3.3%
Plásticos/moldeo	3	2.5%
Aeroespacial	3	2.5%
Química	2	1.7%
Cerámica	2	1.7%
Naval	1	0.8%

Matriz de cruce tecnología × sector (N=120).

La Tabla 4 presenta la intersección entre tecnología de IA y sector industrial para el corpus completo, permitiendo identificar visualmente vacíos de investigación (celdas en cero o con baja frecuencia) que no son evidentes en las distribuciones marginales de las Tablas 2 y 3.

Tabla 4. Matriz de cruce tecnología × sector (conteo de estudios; filas no excluyentes, por lo que las columnas pueden sumar más que el total marginal de la Tabla 3). *Otros sectores agrupa textil, plásticos/moldeo, aeroespacial, química, cerámica y naval ($n \leq 4$ cada uno). Vacíos notables: IA explicable y aprendizaje federado están ausentes en manufactura general; aprendizaje federado solo aparece en metal/acero y manufactura aditiva; los transformers no aparecen fuera de manufactura general, aditiva, automotriz y metal/acero.

Tecnología	Manuf . gral.	Metal/Acer o	Manuf . aditiva	Automotriz	Alimento s	Construcción	Electrónica	Otros *
ML clásico/otro	16	7	2	3	7	4	4	6
CNN	6	6	2	4	0	1	0	2
Gemelo digital	10	2	1	0	1	0	0	2
YOLO/detección de objetos	4	4	0	2	1	0	0	1
IA explicable (XAI)	0	2	4	1	0	1	1	3
LSTM/RNN	4	1	2	3	0	0	0	0
IA generativa/GAN	2	0	0	1	2	0	1	1
Transformer/VLM	3	0	1	1	0	0	0	0
Aprendizaje federado	0	1	2	0	0	0	0	0
Aprendizaje por refuerzo	0	0	0	1	0	0	0	0

Fuente de datos y traducción a impacto operativo (N=120).

El 65.0% de los estudios ($n=78$) reporta evidencia de aplicación o validación en datos o casos industriales reales (estudio de caso con socio industrial nombrado, planta o empresa específica), mientras el 35.0% restante ($n=42$) se valida exclusivamente en conjuntos de datos públicos o de benchmark sin evidencia de despliegue industrial. Sin embargo, solo el 9.2% del corpus completo ($n=11$) reporta una métrica de impacto financiero u operativo explícita (reducción de costos, OEE, reducción de desperdicio, tiempo de inspección, ROI); el 90.8% restante reporta exclusivamente métricas técnicas de desempeño del modelo (exactitud, precisión, F1-score, IoU), sin traducirlas a un indicador de negocio. Esta brecha —entre aplicación en contexto industrial real (65.0%) y traducción a impacto operativo demostrado (9.2%)— constituye el hallazgo central de esta revisión y se retoma como eje de la Discusión (sección 4.1).

Subconjunto de ilustración cualitativa (n=13).

La Tabla 5 presenta el subconjunto de 13 estudios seleccionados mediante muestreo intencional de máxima variación (sección 2.6), utilizado para ilustrar cualitativamente, mediante ejemplos concretos, los patrones agregados establecidos en las secciones 3.2-3.5. Este subconjunto no constituye una muestra representativa adicional ni reemplaza el análisis del corpus completo.

Tabla 5. Subconjunto de estudios seleccionados para ilustración cualitativa (n=13), mediante muestreo intencional de máxima variación (sección 2.6). [nota de revisión: tras corregir la etiqueta de Lei et al. a Aprendizaje federado, esta categoría queda duplicada con Eslaminia et al., IA explicable (XAI) también aparece dos veces (Ozcan & Sengul; Topaloğlu et al.), ML clásico aparece cuatro veces, y Aprendizaje por refuerzo no tiene representante — revisar y sustituir según el criterio de un estudio por categoría de la sección 2.6. Además, "Materiales/Cemento" y "Fotovoltaica" no corresponden a ninguno de los 13 sectores de la Tabla 3 y deben remapearse]

Autor/año	Tecnología	Sector	Principales aportes
Molina et al. (2026)	CNN	Alimentos/Bebidas	Clasificación de daños en manzanas mediante CNN para descarte automático de fruta de baja calidad en plantas de empaque.
Kır et al. (2025)	YOLO / detección de objetos	Automotriz	Celda robótica de detección de fallas combinando deep learning y procesamiento de imágenes para control de calidad de autopartes.
Colaw et al. (2025)	Gemelo digital	Aeroespacial	Metodología de gemelos digitales humanos (Human Digital Twins) para aseguramiento de calidad en manufactura aeroespacial.
Ozcan & Sengul (2025)	IA explicable (XAI)	Cerámica	Métodos de IA supervisada para determinación de calidad en líneas de producción continua, caso de industria cerámica.
Lei et al. (2024)	Aprendizaje federado	Manufactura aditiva	Aprendizaje federado para predicción de emisión de calor en fusión de lecho de polvo láser, superando la escasez de datos homogéneos.
Theodoropoulos et al. (2025)	IA generativa / GAN	Electrónica/Baterías	StyleGAN aplicado a robustez de detección ante defectos visuales novedosos no vistos en entrenamiento.
Pan et al. (2026)	Transformer / VLM	Manufactura general	MaskLenNet: segmentación de instancias y predicción de longitud para evaluación cuantitativa de desgaste y rotura de herramientas.
Eslaminia et al. (2024)	Aprendizaje federado	Metal/Acero	Generalización federada de dominio para monitoreo de condición en soldadura ultrasónica de metales.
Mao et al. (2025)	ML clásico (IoT)	Construcción	Laboratorio-en-IoT para predicción de temperatura de construcción en pavimento asfáltico.
Carvalho et al. (2026)	ML clásico	Textil	Vincula el desempeño de un modelo de clasificación de defectos con indicadores operativos (costos, desperdicio) mediante cadenas de Markov, en colaboración con un fabricante textil real.

Autor/año	Tecnología	Sector	Principales aportes
Topaloğlu et al. (2026)	IA explicable (XAI)	Materiales/Cemento	Predicción de finura Blaine en producción de cemento a partir de datos industriales de control de calidad química y granulométrica.
Xia et al. (2022)	ML clásico	Fotovoltaica	Gestión de sistemas de producción con machine learning, caso de análisis en manufactura fotovoltaica (Suzhou GCL).
Elkhouly & Barakat (2026)	ML clásico	Plásticos/Moldeo	Framework híbrido de control de calidad predictivo, estudio comparativo de modelos ML y redes neuronales en producción de botellas plásticas.

Discusión

La brecha entre aplicación industrial y traducción a impacto operativo (hallazgo central, N=120).

El hallazgo más relevante de esta revisión, sustentado en la totalidad del corpus mapeado (N=120, sección 3.5), es la disociación entre la aplicación en contexto industrial real (65.0% de los estudios) y la traducción de esa aplicación en un indicador de impacto financiero u operativo demostrado (9.2%). Es decir, la mayoría de los estudios que sí despliegan sus sistemas en plantas o procesos reales se detienen en el reporte de métricas técnicas del modelo (exactitud, F1-score), sin cuantificar su efecto sobre costos, tiempo de inspección, reducción de desperdicio (scrap) o eficiencia global del equipo (OEE). Esta brecha es consistente con la señalada por revisiones previas de Quality 4.0 (Oliveira et al., 2026) sobre énfasis desproporcionado en capacidades técnicas de las herramientas de IA en detrimento de su valor organizacional demostrado, pero aquí se cuantifica por primera vez de forma sistemática sobre un corpus de 120 estudios.

Patrones tecnológicos y sectoriales (matriz de cruce, N=120).

La matriz de cruce tecnología-sector (Tabla 4) revela vacíos de investigación no evidentes en las distribuciones marginales: la IA explicable (XAI) y el aprendizaje federado están virtualmente ausentes en el sector de manufactura general, concentrándose en cambio en metal/acero y manufactura aditiva — sectores con procesos más estandarizados y datos de imagen más homogéneos. Los modelos tipo transformer no aparecen fuera de cuatro sectores (manufactura general, aditiva, automotriz, metal/acero), sugiriendo que su adopción sigue limitada a contextos con mayor madurez de infraestructura de datos. Esto sustenta la interpretación de que la investigación ha seguido la disponibilidad de datos estandarizados más que el riesgo o valor económico de la inspección en sectores subrepresentados (química, cerámica, naval).

Ejemplos ilustrativos del subconjunto de máxima variación (n=13).

Los 13 estudios de la Tabla 5, seleccionados para maximizar la diversidad tecnológica y sectorial (sección 2.6), ilustran con casos concretos los patrones ya establecidos en las secciones 3.4-3.5, sin constituir evidencia adicional independiente. Carvalho et al. (2026) ejemplifican el patrón minoritario (9.2%) de traducción exitosa a impacto de negocio: en colaboración directa con un fabricante textil real, vinculan el desempeño de su modelo de clasificación de defectos con reducción de costos (1.3%) y de desperdicio (>90%) mediante modelado de cadenas de Markov. En contraste, la mayoría de los estudios del subconjunto —como Molina et al. (2026) en clasificación de daños en manzanas, o Kır et al. (2025) en detección de fallas automotrices— reportan métricas técnicas de clasificación sin esta traducción, ejemplificando el patrón dominante (90.8%) identificado en la sección 3.5. Lei et al. (2024) y Eslaminia et al. (2024) ilustran la aplicación emergente de aprendizaje federado específicamente en los dos sectores donde la Tabla 4 muestra su concentración (manufactura aditiva y metal/acero), consistente con el patrón agregado.

Limitaciones.

Esta revisión de alcance presenta limitaciones metodológicas que deben considerarse. Primero, y de forma más significativa, el cribado y la extracción de datos fueron apoyados por un asistente de inteligencia artificial bajo supervisión humana, sin protocolo formal de doble cribado independiente por dos revisores humanos ni auditoría documentada de una submuestra para calcular tasas de acuerdo o error (sección 2.7); se recomienda que futuras actualizaciones de esta revisión, o su versión final antes de publicación, incorporen dicha auditoría. Segundo, no se realizó evaluación de calidad metodológica individual (tipo MMAT) de los 120 estudios, consistente con las directrices PRISMA-ScR que consideran esta evaluación opcional para revisiones de alcance, pero limitando la distinción entre estudios con validación rigurosa y limitada. Tercero, la extracción de las variables de fuente de datos e impacto operativo (sección 3.5) se basó en análisis de palabras clave sobre título/resumen, no en lectura de texto completo verificada sistemáticamente, lo que pudo subestimar casos donde el impacto operativo se reporta solo en el cuerpo del artículo. Cuarto, la búsqueda se concentró en Scopus, con una contribución marginal de SciELO (1 estudio) y ninguna de Redalyc; no se buscó en Web of Science, IEEE Xplore ni bases especializadas en ingeniería como Compendex. Quinto, el subconjunto de 13 estudios

(sección 2.6) tiene fines exclusivamente ilustrativos y no debe interpretarse como una muestra representativa adicional del corpus.

Conclusiones

Este mapeo de 120 estudios revela un campo en expansión acelerada, dominado por técnicas de visión artificial (aprendizaje automático clásico y CNN) y concentrado en los sectores metal/acero y manufactura aditiva. El hallazgo central, sustentado en la totalidad del corpus, es la brecha entre la aplicación en contexto industrial real (65.0%) y la traducción de esa aplicación en impacto financiero u operativo demostrado (solo 9.2%) — una disociación que limita la utilidad directa de esta literatura para la toma de decisiones gerenciales sobre adopción de estas tecnologías. La matriz de cruce tecnología-sector identifica vacíos de investigación específicos, particularmente en la aplicación de IA explicable y aprendizaje federado fuera de metal/acero y manufactura aditiva. Se recomienda que la investigación futura priorice: la cuantificación sistemática de impacto operativo y financiero, no solo desempeño técnico del modelo; la extensión de estas tecnologías a sectores industriales subrepresentados; y revisiones sistemáticas de segundo nivel, con evaluación de calidad metodológica formal y doble cribado humano independiente, sobre los subconjuntos tecnológicos específicos identificados en este mapeo.

Referencias

- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32. <https://doi.org/10.1080/1364557032000119616>
- Carvalho, A., Miguéis, V., & Moura e Sá, M. (2026). Enhancing operational performance in textile manufacturing: Impact of deep learning-based defect detection. *International Journal of Production Research*, 64(16), 6744-6768. <https://doi.org/10.1080/00207543.2026.2632695>
- Colaw, C. L., Madison, G., Tseng, B., Griser, G. M., Truelson, G., Gallo, A., & Hurmuzlu, Y. (2025). Methodology for enablement of human digital twins for quality assurance in the aerospace manufacturing domain. *Sensors*, 25(11). <https://doi.org/10.3390/s25113362>
- Elkhoully, H. I., & Barakat, S. (2026). A hybrid framework for predictive quality control in manufacturing: A comparative study of machine learning and neural computing models. *Neural Computing and Applications*, 38(10). <https://doi.org/10.1007/s00521-026-12109-9>

- Eslaminia, A., Meng, Y., Nahrstedt, K., & Shao, C. (2024). Federated domain generalization for condition monitoring in ultrasonic metal welding. *Journal of Manufacturing Systems*, 77, 1-12. <https://doi.org/10.1016/j.jmsy.2024.09.023>
- Kır, H., Adar, N. G., & Yazar, M. (2025). A fault detection robotic cell application based on deep learning and image processing hybrid approach for quality control of automotive parts. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, 49(1), 471-485. <https://doi.org/10.1007/s40998-024-00768-0>
- Lei, R., Guo, Y. B., Yan, J., & Guo, W. G. (2024). Bridging data gaps: A federated learning approach to heat emission prediction in laser powder bed fusion. *Journal of Manufacturing Science and Engineering*, 146(10). <https://doi.org/10.1115/1.4065888>
- Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, 5, 69. <https://doi.org/10.1186/1748-5908-5-69>
- Ma, Y., Yin, J., Huang, F., & Li, Q. (2024). Surface defect inspection of industrial products with object detection deep networks: A systematic review. *Artificial Intelligence Review*, 57(12), article 333. <https://doi.org/10.1007/s10462-024-10956-3>
- Mao, W., Si, W., Li, N., Song, J., Yuan, C., Duojie, C., Zheng, J., Zhang, B., & Yang, R. (2025). Laboratory-on-IoT for predicting construction temperature of asphalt pavement. *Journal of Traffic and Transportation Engineering (English Edition)*, 12(5), 1254-1270. <https://doi.org/10.1016/j.jtte.2023.05.011>
- Molina, M., Godoy, J., Castro, J. W., & Rizzo, V. (2026). Apple damages classification: Using the best convolutional neural network to discard low surface quality fruit in packing plants. *Computers and Electronics in Agriculture*, 244, 111489. <https://doi.org/10.1016/j.compag.2026.111489>
- Morales Matamoros, O., Takeo Nava, J. G., Moreno Escobar, J. J., & Ceballos Chávez, B. A. (2025). Artificial intelligence for quality defects in the automotive industry: A systemic review. *Sensors*, 25(5), 1288. <https://doi.org/10.3390/s25051288>
- Oliveira, A. C. A., Pinho, A. F. de, & Brendon, M. F. (2026). Quality 4.0: A systematic literature review and future research agenda for the digital transformation of quality management. *International Journal of Quality & Reliability Management*, 43(6), 1650–1673. <https://doi.org/10.1108/IJQRM-05-2025-0165>
- Ozcan, S., & Sengul, U. (2025). The use of supervised artificial intelligence methods in quality determination in continuous production lines: A case study of ceramic industry. *International Journal on Interactive Design and Manufacturing*, 19(12), 8687-8706. <https://doi.org/10.1007/s12008-025-02310-w>
- Pan, Y., He, K., Yin, C., Zhang, Y., Luo, Y., & Wang, Y. (2026). MaskLenNet: A query-based instance segmentation and length prediction network for quantitative industrial tool wear and breakage assessment. *Journal of Manufacturing and Materials Processing*, 10(8), 286. <https://doi.org/10.3390/jmmp10080286>

- Theodoropoulos, S., Dardanis, D., Makridis, G., Zajec, P., Rožanec, J. M., Kyriazis, D., & Tsanakas, P. (2025). Enhancing robustness to novel visual defects through StyleGAN latent space navigation: A manufacturing use case. *Journal of Intelligent Manufacturing*, 36(5), 3527-3541. <https://doi.org/10.1007/s10845-024-02415-1>
- Topaloğlu, M. T., Macit, C. K., Uçar, U. U., & Tanyeri, B. (2026). Prediction of Blaine fineness of final product in cement production using industrial quality control data based on chemical and granulometric inputs using machine learning. *Applied Sciences (Switzerland)*, 16(4), 2046. <https://doi.org/10.3390/app16042046>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., ... Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467-473. <https://doi.org/10.7326/M18-0850>
- Xia, H., An, W., Zhang, Z., & Liu, G. (2022). Managing production systems with machine learning: A case analysis of Suzhou GCL photovoltaic technology. *Production Planning and Control*, 33(16), 1559-1572. <https://doi.org/10.1080/09537287.2021.1882693>
- Zhou, H. A., Wolfschläger, D., Florides, C., Werheid, J., Behnen, H., Woltersmann, J.-H., Pinto, T. C., Kemmerling, M., Abdelrazeq, A., & Schmitt, R. H. (2026). Generative AI in industrial machine vision: A review. *Journal of Intelligent Manufacturing*, 37, 1447-1470. <https://doi.org/10.1007/s10845-025-02604-6>