UNIVERSIDAD DE BURGOS

Programa de Doctorado en Ingeniería y Tecnologías Industrial, Informática y Civil

Tesis Doctoral

Visión inteligente para la detección de defectos en la industria textil

Autora: Beatriz Gil Arroyo

Directores: Dr. Daniel Urda Muñoz y Dr. Ángel Arroyo Puente

Grupo de Investigación: Grupo de Inteligencia Computacional Aplicada (GICAP)

Departamento: Digitalización

Burgos, Diciembre de 2025

Agradecimentos

En primer lugar, deseo expresar mi sincero agradecimiento a la Universidad de Burgos, que ha sido un marco fundamental en mi trayectoria académica y profesional. Quiero reconocer especialmente al director del departamento, Álvaro Herrero, por su confianza y por brindarme la oportunidad de redescubrir la Universidad de Burgos, apreciando de nuevo su valor y las oportunidades que ofrece para la investigación.

Mi más profundo reconocimiento a mis directores de tesis, Daniel Urda y Ángel Arroyo, por su guía experta, su implicación y su paciencia. Gracias por ofrecerme siempre orientación clara, rigor y un acompañamiento constante en cada fase de este trabajo.

Extiendo mi agradecimiento a los colaboradores externos y coautores de los artículos, cuya contribución científica ha sido esencial. Quiero reconocer de manera especial a Juan Marcos Sanz, de Textil Santanderina, por la cesión del material textil y su implicación activa en el proyecto, incluyendo el exigente trabajo de revisión manual de imágenes.

Agradezco también a mis compañeros del edificio C por su cercanía, apoyo y por crear un ambiente laboral que ha hecho este proceso más llevadero. Cada conversación, gesto de ánimo y momento compartido han tenido un valor incalculable.

A mi familia y amigos, que me han acompañado en momentos de estrés, impaciencia y cansancio, especialmente durante las semanas más intensas de trabajo: gracias por vuestro cariño, comprensión y apoyo incondicional.

Por último, mi dedicatoria más especial es para mis padres. Me habéis transmitido valores que han guiado cada paso de mi vida: esfuerzo, constancia y respeto por los demás. Gracias por enseñarme a avanzar con firmeza, a enfrentar los retos con serenidad y aprender de cada experiencia. También quiero reconocer el privilegio que he tenido de poder formarme, una oportunidad que vosotros no tuvisteis, pero que siempre deseasteis para mí. Cada logro que alcanzo refleja vuestra fuerza, sacrificio y amor.

Resumen

La presente tesis doctoral, desarrollada en el marco del grupo de investigación Grupo de Inteligencia Computacional APlicada (GICAP) de la Universidad de Burgos (UBU), se centra en la aplicación de técnicas de Inteligencia Artificial (IA), concretamente *Deep Learning (DL)*, al control de calidad en la industria textil, con especial atención a la detección automática de defectos en tejidos Batavia y Sarga. Este trabajo se enmarca en la línea de investigación impulsada por el proyecto europeo *Deep lEarning for automatiC tExtile iNspecTion (DECENT)*, en el que participan Investigación Asesoramiento y Desarrollo Textil (Inade), la UBU y el *Digital Innovation Hub on Livestock, Environment, Agriculture & Forest (DIH-LEAF)* dentro de la segunda convocatoria DIH-WORLD del programa Horizon 2020. Dicho proyecto impulsó la línea de investigación y proporcionó los recursos necesarios para el desarrollo del dataset, la experimentación y el modelado de la visión artificial.

El objetivo general de esta tesis es **desarrollar, analizar y evaluar la aplicación de técnicas de DL para la detección automatizada de defectos en tejidos**, proporcionando un sistema más objetivo y consistente que la inspección manual, que sirva como base para futuras aplicaciones industriales.

La metodología se diseñó para garantizar resultados realistas y robustos. Se comenzó con la adquisición de 2.755 imágenes de tejido Batavia y 1.548 de tejido Sarga, ambos en escala de grises (2048 x 696 píxeles), capturadas en condiciones reales de producción. Tras el preprocesamiento (conversión de 16 a 8 bits y segmentación en más de 47.000 *patches*), se aplicaron dos estrategias de modelado:

1. **Estrategia *Convolutional Neural Network (CNN)* *Pura:*** Evaluación de diversas arquitecturas CNN, incluyendo Visual Geometry Group (VGG) (VGG16 y VGG19), Inception (InceptionV3), Residual Network (ResNet) (ResNet50V2 y ResNet101), Extreme Inception (Xception), Densely Connected Convolutional Network (DenseNet) (DenseNet121) y EfficientNet (B0 y B3) para identificar el mejor desempeño por cada tipo de tejido.

2. **Estrategia Híbrida *Autoencoder (AE)* + CNN:** Propuesta metodológica en dos etapas: en la primera fase, un AE no supervisado identificó imágenes potencialmente anómalas, facilitando la labor de etiquetado de los expertos y optimizando la eficiencia del dataset. En la segunda fase, los modelos CNN se entrenaron con todo el dataset, aunque los resultados se presentan sobre el subconjunto de imágenes anómalas, mostrando un desempeño notable en la detección

de defectos.

Los resultados, presentados en las tres publicaciones que conforman el compendio de la tesis doctoral, validan la hipótesis planteada: el uso de técnicas de DL es viable y supera a la inspección manual, proporcionando un sistema más objetivo, repetible y escalable. El primer artículo describe la creación y publicación del dataset curado, constituyendo un *benchmark* abierto. El segundo analiza la eficacia de distintas arquitecturas CNN en la detección automática de defectos. El tercero profundiza en la metodología híbrida AE + CNN, alcanzando métricas Area Under the Receiver Operating Characteristic Curve (AU-ROC) de hasta 0,97 en la referencia Sarga 43105, lo que evidencia la robustez del enfoque incluso en muestras complejas.

En conjunto, este trabajo impulsa la **transferencia de conocimiento** al sector industrial y sienta las bases para la implantación de sistemas de control de calidad más objetivos, rápidos y automatizados, contribuyendo así al avance de la industria textil hacia los estándares de la Industria 4.0.

Abstract

This doctoral thesis, developed within the research group GICAP at the UBU, focuses on the application of IA techniques, specifically DL, for quality control in the textile industry, with a particular emphasis on the automated detection of defects in Batavia and Sarga fabrics. This work is framed within the research line promoted by the European project DECENT, in which Inade (Textile Research, Advice and Development), UBU, and DIH-LEAF participate, under the second DIH-WORLD call of the Horizon 2020 program. This project promoted the research line and provided the necessary resources for dataset development, experimentation, and computer vision modeling.

The general objective of this thesis is to develop, analyze, and evaluate the application of DL techniques for the automated detection of defects in fabrics, providing a more objective and consistent system than manual inspection, which can serve as a basis for future industrial applications.

The methodology was designed to ensure realistic and robust results were obtained. The study began with the acquisition of 2,755 Batavia fabric images and 1,548 Sarga fabric images, both in grayscale (2048 $\times$ 696 pixels), captured under real production conditions. After preprocessing (conversion from 16 to 8 bits and segmentation into over 47,000 patches), two modeling strategies were applied:

Pure CNN Strategy: Evaluation of various CNN architectures, including VGG (VGG16 and VGG19), Inception (InceptionV3), ResNet (ResNet50V2 and ResNet101), Xception, DenseNet (DenseNet121), and EfficientNet (B0 and B3) to identify the best performance for each type of fabric.

Hybrid Autoencoder (AE) + CNN Strategy: Two-stage methodological proposal: in the first phase, an unsupervised AE identified potentially anomalous images, facilitating expert labeling and optimizing dataset efficiency. In the second phase, the CNN models were trained with the entire dataset, although the results were reported on a subset of anomalous images, showing remarkable performance in defect detection.

The results, presented across the three publications that comprise the thesis compendium, validate the proposed hypothesis: the use of DL techniques is feasible and outperforms manual inspection, providing a more objective, repeatable, and scalable system. The first article describes the creation and publication of a curated dataset, establishing an open benchmark. The second analyzes the effectiveness of different CNN architectures for automated defect detection. The third focuses on the hybrid AE + CNN methodology, achieving AU-ROC metrics of up to 0.97 on the Sarga 43105

reference, demonstrating the robustness of the approach even for complex samples.

Overall, this study promotes knowledge transfer to the industrial sector and lays the foundation for the implementation of more objective, fast, and automated quality control systems, thus contributing to the advancement of the textile industry towards Industry 4.0 standards.

Palabras clave / Keywords

Palabras clave:
Visión inteligente, detección de defectos, industria textil, deep learning, redes neuronales convolucionales, autoencoder, control de calidad, industria 4.0.

Keywords:
Intelligent vision, defect detection, textile industry, deep learning, convolutional neural networks, autoencoder, quality control, industry 4.0.

Índice general

1. Introducción

Antes de detallar los aspectos específicos de esta tesis, es importante contextualizar el problema, la motivación que impulsa este trabajo, así como el marco institucional y científico en el que se enmarca. Por último, se ofrece una visión general de la estructura de la tesis, con el fin de guiar al lector sobre el desarrollo y la organización de los contenidos que se presentan a lo largo del documento.

1.1. Estado del arte

La creciente competitividad en la industria ha impulsado a las empresas a optimizar sus procesos para garantizar productos de alta calidad que cumplan con las expectativas del mercado. En el sector textil, el aumento de la producción y la complejidad de los tejidos hace imprescindible mantener estrictos estándares de calidad, dado que los defectos en los materiales impactan directamente en su valor económico. Un estudio [1] realizado en 2020 en Bahir Dar Textile Share Company reportó que aproximadamente un 30 % del valor de la producción semanal se pierde debido a defectos durante el tejido, lo que representa una pérdida anual significativa para la empresa. Entre las principales causas de estos defectos se incluyen irregularidades en los hilos de urdimbre y trama, errores en la configuración y mantenimiento de las máquinas, calidad deficiente de los materiales auxiliares y descuidos del personal encargado del proceso.

Tradicionalmente, la inspección y control de calidad de los tejidos se ha realizado de manera **manual por operarios capacitados**. Este proceso resulta laborioso, subjetivo y propenso a errores, especialmente en tejidos con patrones complejos o elevada regularidad. Además, la velocidad de inspección suele ser inferior a la de producción, generando cuellos de botella y limitando la capacidad de mantener los niveles de calidad requeridos [2]. Por ejemplo, en Textil Santanderina (TS) [3] la inspección manual se realiza a 30 metros/minuto, mientras que la producción supera los 40 metros/minuto, lo que afecta a la productividad y al control de calidad.

Las técnicas clásicas de **control estadístico de procesos**, introducidas originalmente por Shewhart [4] en 1931, han sido ampliamente utilizadas para verificar y mantener la calidad de los productos industriales [5]. Sin embargo, estas herramientas están diseñadas fundamentalmente para detectar variaciones leves en parámetros numéricos y resultan poco eficaces frente a defectos complejos o datos de alta dimensionalidad, como los provenientes de imágenes textiles. Esta limitación ha motivado el desarrollo de métodos más automatizados y precisos basados en técnicas de visión por computador y DL.

La **visión artificial** [6] surge como alternativa a la inspección manual, automatizando la detección de defectos mediante procesamiento digital de imágenes y algoritmos clásicos que extraen características como bordes, texturas y formas según reglas definidas. Aunque mejora la consistencia y la velocidad respecto a la inspección humana, su desempeño se ve limitado en tejidos con patrones irregulares o deformaciones. Liu et al. (2011) [7] proponen un método de detección de defectos basado en un banco de filtros de Gabor cuyos parámetros dependen directamente de las características de textura del tejido. Mediante un conjunto de filtros con múltiples frecuencias y orientaciones, generan diversas imágenes filtradas que posteriormente se reconstruyen en una única representación para facilitar la detección y segmentación de defectos. Los resultados experimentales muestran que la mayoría de los defectos son detectados y segmentados correctamente, evidenciando la robustez del enfoque y la efectividad de los métodos clásicos de visión artificial en tejidos con patrones uniformes. De forma complementaria, Modrângă et al. (2017) [8] combinan transformada de Fourier y filtros de Gabor para detectar irregularidades en la periodicidad de la tela, reforzando la eficacia de la visión clásica en la inspección automatizada de tejidos.

La creciente complejidad de los tejidos y la alta velocidad de producción en la industria textil han impulsado la transición hacia sistemas automatizados de **visión inteligente** [6], que combinan procesamiento digital de imágenes con técnicas de IA para detectar defectos de manera rápida, precisa y consistente.

El desarrollo de los sistemas modernos de visión inteligente aplicados a la inspección automática de tejidos se apoya en una sólida base teórica en aprendizaje automático y redes neuronales artificiales. El entrenamiento de redes neuronales multicapa mediante algoritmos de optimización basados en gradiente se formalizó con el trabajo de Rumelhart et al. (1986) [9], quienes introdujeron el algoritmo de retropropagación del error como un método eficiente para ajustar los pesos de redes neuronales profundas. Posteriormente, Hornik et al. (1989) [10] demostraron que las redes neuronales multicapa con al menos una capa oculta poseen la capacidad de aproximar cualquier función medible con precisión arbitraria, estableciendo su carácter de aproximadores universales.

En el ámbito del modelado de dependencias temporales, Elman (1990) [11] introdujo las Redes Nueuronales Recurrentes (RNN), que incorporan realimentación interna para dotar a las redes de memoria dinámica y permitir el procesamiento de secuencias. No obstante, Bengio et al. (1994) [12] demostraron que el entrenamiento de RNN mediante descenso por gradiente presenta serias dificultades para capturar dependencias a largo plazo, debido a los problemas de desvanecimiento y explosión del gradiente. Como solución a esta limitación, Hochreiter y Schmidhuber (1997) [13] propusieron las redes Long Short-Term Memory (LSTM), que incorporan mecanismos de compuertas para preservar el flujo del gradiente y permitir el aprendizaje de dependencias temporales extensas.

Desde una perspectiva más general, Bishop (2006) [14] proporciona un marco unificado que integra reconocimiento de patrones y aprendizaje automático desde un enfoque probabilístico y bayesiano, abordando tanto modelos clásicos como técnicas modernas de inferencia aproximada. La consolidación del DL como paradigma dominante se refleja en el trabajo de LeCun et al. (2015) [15], quienes destacan el impacto de las arquitecturas profundas en tareas como visión por computador, reconocimiento de voz y procesamiento del lenguaje natural. Finalmente, Goodfellow et al. (2016) [16] presentan una visión integral de DL, cubriendo desde los fundamentos matemáticos hasta arquitecturas avanzadas y aplicaciones prácticas, convirtiéndose en una referencia esencial en el campo.

Los métodos supervisados requieren conjuntos de datos etiquetados, y han demostrado un rendimiento sobresaliente tanto en tejidos con patrón como en tejidos lisos. Biradar et al. (2021) [17] propusieron una CNN de tres capas para la detección de defectos en tejidos, capaz de extraer automáticamente características relevantes mediante capas de convolución y *pooling* antes de realizar la clasificación final. La red se evaluó tanto sobre el conjunto de datos público TILDA como sobre una base de datos propia. Bing Wei et al. (2022) [18] desarrollaron una Multilevel multi-attentional network (MLMA-Net), una red profunda con un mecanismo de atención multinivel diseñada específicamente para la detección multietiqueta en imágenes de defectos textiles, logrando mejoras significativas especialmente en escenarios con múltiples defectos presentes de forma simultánea. Li y Zhu (2024) [19] desarrollaron PEI-YOLOv5, un sistema de detección en tiempo real que integra *Particle Depthwise Convolution, Enhance-BiFPN* y una nueva función de pérdida denominada *IN loss*. Su método mostró un rendimiento notable en el conjunto GuangDong TianChi y también fue evaluado en NEU Surface Defect Database; aunque este último no es un conjunto textil, se emplea para evaluar la capacidad de generalización entre dominios. Yang et al. (2024) [20] propusieron AC-CTNet, un modelo que combina CNN y *Vision Transformer* para detectar defectos mediante un enfoque de saliencia, que resalta o identifica automáticamente las partes más relevantes o anómalas de una imagen. El método introduce un módulo de coordinación de contexto adyacente para mejorar la interacción multiescala y un módulo de agregación por contraste que resalta defectos poco visibles. Finalmente, Zhou et al. (2025) [21] propusieron DCFE-YOLO, una versión mejorada de YOLOv8 que incorpora *Dynamic Snake Convolution* y mecanismos de atención multiescala, permitiendo mejorar la detección de defectos sutiles en tejidos industriales.

Los métodos no supervisados han ganado protagonismo debido al alto coste de obtener muestras defectuosas etiquetadas y la dificultad de abarcar la gran variedad de defectos posibles. Estos enfoques se entrenan únicamente con muestras sin defecto, aprendiendo la distribución normal del tejido y detectando anomalías como desviaciones respecto a la reconstrucción. Koulali y Eskil (2021) [22] presentaron un método no supervisado basado en CNN que utiliza una estrategia dinámica de selección de características y evita el uso de retropropagación tradicional, permitiendo entrenar eficazmente con una única muestra sin defectos. Su modelo fue evaluado en el conjunto Patterned Fabric Dataset (PFD). Zhang et al. (2022) [23] desarrollaron un AE convolucional multiescala de tipo U para tejidos coloreados y con patrón, empleando mapas de residuo para la localización precisa del defecto y alcanzando resultados competitivos en el conjunto YDFID-1. Zhang et al.(2023) [24] propusieron QA-USTNet, una red *U-shaped* basada en *Swin Transformer* con un módulo de atención *Quadtree*, capaz de modelar el campo receptivo global y mejorar la reconstrucción de tejidos teñidos en hilo. Finalmente, Si y Kim (2024) [25] presentaron V-DAFT, que combina un AE convolucional con el análisis en el dominio de Fourier para identificar anomalías frecuenciales características de texturas defectuosas, logrando un rendimiento competitivo con una arquitectura ligera y un entrenamiento rápido.

En escenarios industriales reales, donde la disponibilidad de muestras defectuosas etiquetadas es limitada pero existe una gran cantidad de datos sin anotar, el DL semi-supervisado se presenta como una alternativa especialmente adecuada. Este paradigma permite aprovechar simultáneamente información etiquetada y no etiquetada, mejorando la robustez y la capacidad de generalización de los modelos. Un enfoque representativo es la estrategia Mean Teacher propuesta por Tarvainen y Valpola (2017) [26], que fomenta la consistencia entre un modelo estudiante y un modelo maestro cuyos pesos se obtienen como una media temporal, logrando un rendimiento competitivo incluso con un número reducido de muestras etiquetadas. En el contexto específico de la industria textil, Zhou et al. (2020) [27] aplicaron aprendizaje semi-supervisado para la detección de defectos en tejidos, combinando reconstrucción de imágenes mediante un AE variacional con estimación de densidad a través de un modelo de mezcla gaussiana. La integración de la información procedente tanto del dominio de la imagen reconstruida como del espacio latente permitió delimitar con mayor precisión las regiones defectuosas, reduciendo la dependencia de datos defectuosos etiquetados y demostrando la idoneidad de los enfoques semi-supervisados para sistemas de inspección textil automatizados.

En conjunto, el estado del arte muestra una clara evolución hacia modelos con mecanismos avanzados de atención, arquitecturas tipo *Transformer* y CNN optimizadas para su despliegue en tiempo real. Los métodos supervisados ofrecen habitualmente mayor precisión cuando se dispone de suficientes datos etiquetados, mientras que los métodos no supervisados resultan esenciales en escenarios donde los defectos son escasos, difíciles de etiquetar o altamente variables. Asimismo, algunos trabajos han demostrado la viabilidad de implementar sistemas de inspección basados en visión inteligente en entornos productivos reales. Hariharan y Sivaraman (2024) [28] desarrollan y despliegan una solución supervisada basada en CNN para la clasificación automática de defectos, logrando mejoras significativas en precisión, exhaustividad y tasa global de detección respecto a la inspección manual. Este tipo de aportaciones evidencia la madurez de las técnicas basadas en DL para su adopción industrial.

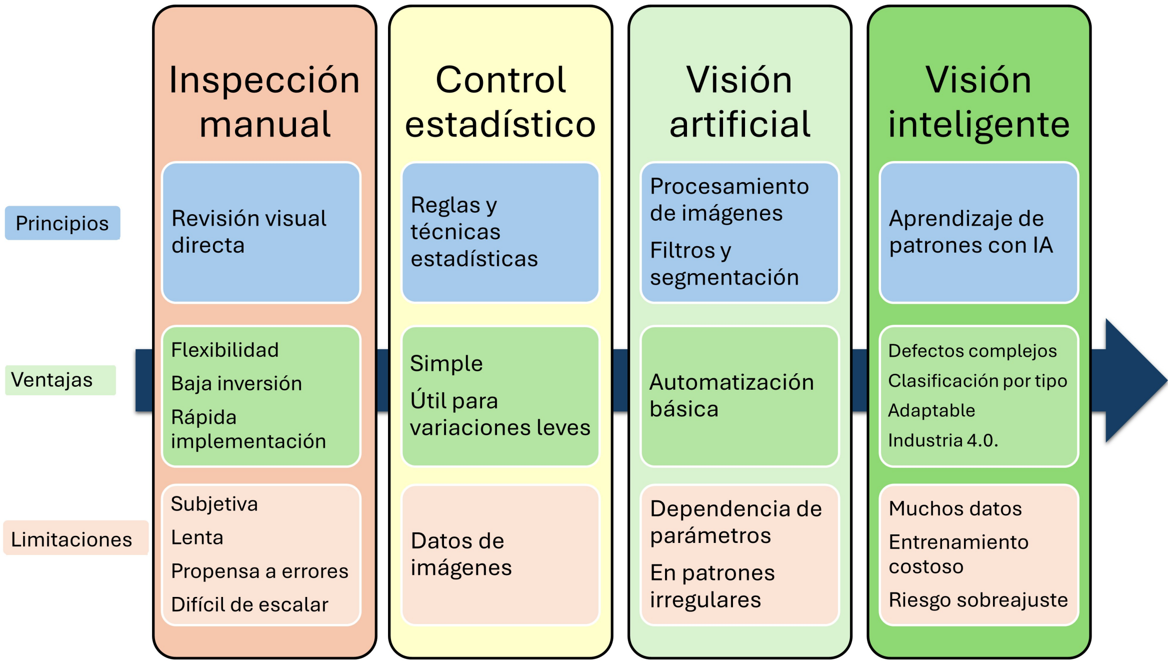
Para sintetizar los enfoques mencionados anteriormente de visión inteligente aplicada en industria textil, la Tabla 1 presenta los principales métodos, tipo de enfoque (supervisados, no supervisados o semi-supervisados), el año y el dataset utilizado. Esto permite comparar la diversidad de técnicas y la disponibilidad de los datos.

Tabla 1: Métodos de IA mencionados anteriormente para la detección de defectos textiles, agrupados por tipo de aprendizaje.

Método	Año	Dataset
Supervisado		
Biradar et al. (DCNN)	2021	TILDA (público) + con patrón (propio) + sin patrón (1000 imágenes) (propio) = 3000 imágenes
Bing Wei et al. (MLMA-Net)	2022	DHU-ML1000 (1000 imágenes) (propio)
Li & Zhu (PEI-YOLOv5)	2024	GuangDong TianChi (1637 imágenes) (público) + NEU Surface Defect (1800 imágenes) (público, metálico: se usa para validación)
Yang et al. (ACCTNet)	2024	Tejido liso (2200 train + 500 test) + Tejido con patrón (5948 train + 500 test) (propio)
Zhou et al. (DCFE-YOLO)	2025	TILDA (público) + Ali platform (propio) + GuangDong (público) = 3575 imágenes
No supervisado		
Koulali & Eskil	2021	PFD (50 imágenes) (público)
Zhang et al. (Autoencoder multiescala)	2022	YDFID-1 (3501 imágenes) (bajo petición) + MVTec AD (5000 imágenes) (público)
Zhang et al. (QA-USTNet)	2023	YDFID-1 (3501 imágenes) (bajo petición)
Si & Kim (V-DAFT)	2024	MVTec AD (5354 imágenes) (público)
Semisupervisado		
Zhou et al. (VAE + GMM)	2020	AITEX (245 imágenes, textil real, público) + DAGM 2007 (benchmark industrial)

Con el fin de proporcionar una visión global de los métodos utilizados en la inspección de defectos en la industria textil, en la Figura 1 se comparan los principales enfoques, resumiendo el principio de funcionamiento de cada método, sus ventajas y limitaciones, permitiendo apreciar la evolución de las técnicas y justificar la necesidad de soluciones automatizadas y adaptativas frente a los retos de productividad, heterogeneidad de tejidos y complejidad de defectos.

Figura 1: Comparativa de métodos de inspección de defectos en la industria textil.



Como han señalado diversos estudios [19, 29], la efectividad de los métodos automáticos de detección de defectos depende en gran medida de las propiedades del material analizado (estructura, textura, tejido, color, patrón), lo que complica su generalización para telas diversas. En esta tesis se consideran principalmente los tejidos Batavia y Sarga, cuyas características influyen directamente en la aparición y visibilidad de defectos. El tejido Sarga, con su patrón diagonal, proporciona resistencia y elasticidad, y se utiliza en prendas como pantalones, chaquetas y trajes, así como en tapicería. Por su parte, el tejido Batavia, una variante más uniforme y densa de la sarga, generalmente compuesto de algodón y lino, se emplea en camisas, blusas, vestidos y textiles para el hogar que requieren una superficie lisa. Estas diferencias de textura y densidad representan un desafío para los sistemas de inspección automática, ya que los métodos de detección deben ser capaces de adaptarse a variaciones en patrón, homogeneidad y densidad del tejido.

La adopción de métodos de IA para la inspección de tejidos se ve limitada por una restricción estructural: la **escasez de datasets públicos** amplios, diversos y bien anotados. Aunque existen algunos conjuntos disponibles —como como PFD, TILDA, ZJU-Leaper, MVTec AD o Lusitano—, muchos presentan limitaciones significativas. Algunos contienen un número reducido de muestras (PFD: 150; AITEX: 245; HKU

Fabric: 336; Fabric Stain: 466), escasa diversidad de tipos de defectos o tejidos muy homogéneos, lo que restringe su utilidad para entrenar modelos robustos y generalizar la detección a distintos tipos de tejidos y condiciones de producción [30]. Por otro lado, algunos conjuntos más grandes en cantidad de imágenes (ZJU-Leaper: 98.777; Lusitano: 34.684) carecen de anotaciones multiclase o de una clasificación detallada de defectos, lo que limita su aplicabilidad para métodos supervisados avanzados. Muchos datasets requieren solicitud previa o no son completamente públicos, y numerosos estudios dependen de conjuntos privados recolectados en fábricas, dificultando la reproducibilidad y la comparación directa de resultados. Esta combinación de factores evidencia que, a pesar de la existencia de varios datasets, sigue existiendo una carencia de datos normalizados y accesibles que sean lo suficientemente amplios, variados y bien anotados para entrenar y evaluar de manera robusta sistemas de inspección automática de tejidos.

En la Tabla 2 se presentan los principales datasets públicos textiles disponibles, indicando el número de muestras, la presencia de anotaciones multiclase, el número de tipos de defecto y la inclusión de imágenes sintéticas.

Tabla 2: Principales datasets públicos de defectos textiles, ordenados por número de muestras

Dataset	Muestras	Multiclase	Nº de clases	Sintéticas
ZJU-Leaper [31]	98777	No	-	No
Lusitano [30]	34684	No	-	No
Tianchi / Guangdong [32]	15436	Sí	15	No
MVTec AD [33]	5354	No	-	No
TILDA v2 [34]	896	Sí	8	No
Fabric Stain Dataset [35]	466	No	-	No
HKU Fabric [36]	336	Sí	6	Sí
AITEX [37]	245	Sí	13	No
PFD [38]	150	Sí	6	No

La motivación de esta investigación radica en desarrollar metodologías de visión inteligente capaces de manejar la heterogeneidad de patrones y el desbalance de clases, mejorar la trazabilidad de defectos y alinearse con los principios de Industria 4.0 [39]. Esta propuesta también busca servir como referencia para futuras investigaciones que requieran datasets industriales públicos y soluciones escalables.

1.2. Descripción del grupo GICAP y líneas de investigación

El grupo **GICAP** [40] es un grupo de investigación adscrito a la **UBU**, integrado por un amplio número de investigadores de perfiles diversos, interesados en el desarrollo de técnicas y modelos basados en IA.

El trabajo del grupo combina investigación teórica rigurosa con un marcado enfoque aplicado, orientado a la resolución de problemas reales. A continuación, se detallan sus líneas de investigación y ámbitos de aplicación.

Líneas de investigación

- **Métodos de IA**: estudio y desarrollo de sistemas capaces de percibir, razonar, aprender, comunicarse y actuar en entornos complejos. Incluye el trabajo con distintas arquitecturas de redes neuronales artificiales, supervisadas y no supervisadas, abarcando modelos convolucionales, densos y de DL.

- **Agentes y sistemas multiagente**: diseño y formalización de modelos de razonamiento para agentes inteligentes, especialmente arquitecturas deliberativas Beliefs Desires Intentions (BDI), que permiten construir sistemas autónomos y coordinados en entornos dinámicos.

Ámbitos de aplicación

- **Modelado y control de procesos industriales**: desarrollo de métodos para optimizar y supervisar operaciones, integrando técnicas de IA y metodologías de ingeniería.

- **Inspección automática de calidad en industria textil**: detección de defectos en tejidos mediante visión por computador y bases de datos de imágenes de alta resolución, facilitando la automatización del control de calidad.

- **Mantenimiento inteligente y control en mecanizado y fabricación industrial**: proyectos de modelado y control de maquinaria y procesos orientados a optimización y mejora del rendimiento.

- **Ciberseguridad e Internet of Things (IoT)**: investigación en la seguridad de dispositivos conectados, combinando IA y ciberseguridad para proteger sistemas en entornos interconectados.

- **Salud**: aplicación de técnicas de IA al análisis de datos clínicos y biomédicos, apoyo al diagnóstico, monitorización de pacientes y optimización de procesos sanitarios.

- **Agricultura de precisión**: desarrollo de soluciones basadas en IA para la monitorización de cultivos, predicción de rendimientos, detección temprana de plagas y optimización del uso de recursos mediante datos de sensores, imágenes y sistemas IoT.

- **Procesamiento de datos multidisciplinares y minería de datos**: aplicación de redes neuronales y aprendizaje automático a problemas en ámbitos diversos como gestión del conocimiento, física, química, ingeniería civil y seguridad de redes.

La presente tesis se enmarca en la línea de investigación **métodos de IA** y en el ámbito de aplicación **inspección automática de calidad en industria textil**.

Esta línea de investigación tuvo continuidad en el proyecto *Deep lEarning for automatiC tExtile iNspecTion (DECENT)* [41], en el que participaron **Investigación Asesoramiento y Desarrollo Textil (INADE)**, la **UBU** y **DIH-LEAF** [42]. El proyecto contó con **financiación europea** en el marco de la **segunda convocatoria DIH-WORLD** del programa **Horizon 2020**, orientada a impulsar la digitalización de las pymes mediante la adopción de tecnologías avanzadas de IA y visión artificial.

La participación del grupo GICAP en este proyecto consolidó su experiencia en la aplicación de redes neuronales a problemas industriales complejos y reforzó la transferencia de conocimiento entre el ámbito académico y el sector productivo.

1.3. Estructura general de la tesis

La tesis se organiza en cinco capítulos principales, acompañados de bibliografía y anexos, con el objetivo de presentar de manera clara y coherente la investigación realizada sobre visión inteligente para la detección de defectos en la industria textil.

- **Capítulo 1. Introducción:** Se contextualiza la investigación, se revisa el estado del arte, se describen los tejidos estudiados y las limitaciones de los datasets, y se presenta al grupo de investigación GICAP y su línea de trabajo.

- **Capítulo 2. Objetivos y Metodología General:** Se detallan los objetivos de la tesis y la metodología empleada, incluyendo el enfoque experimental, la selección de datasets, las arquitecturas de los modelos y los criterios de evaluación.

- **Capítulo 3. Publicaciones Incluidas en la Tesis:** Se resumen los artículos seleccionados, indicando la referencia completa, un resumen de cada trabajo y la contribución personal del autor.

- **Capítulo 4. Artículos Completos:** Se presentan los artículos científicos en su versión completa, tal como fueron publicados o aceptados.

- **Capítulo 5. Discusión General y Conclusiones:** Se integran y analizan los resultados, destacando las aportaciones científicas y tecnológicas, las limitaciones del estudio y las líneas de investigación futuras.

- **Bibliografía y Anexos:** Incluyen todas las referencias citadas, cartas de aceptación de los artículos y consentimientos de coautores.

Esta organización permite al lector seguir de forma lógica el desarrollo de la investigación, desde la motivación y el estado del arte hasta la discusión crítica y las conclusiones, mostrando de manera integral la contribución de la tesis al campo de la visión inteligente aplicada a la industria textil.

2. Objetivos y Metodología General

En este capítulo se presentan los objetivos y la metodología general que guían el desarrollo de la presente tesis. Se detallan los objetivos que permiten abordar de manera sistemática la investigación sobre la aplicación de técnicas de DL para la detección automatizada de defectos en tejidos industriales.

Asimismo, se describe la metodología seguida para alcanzar dichos objetivos, incluyendo la formulación del problema, la adquisición y procesamiento de datos, el diseño y entrenamiento de modelos, la evaluación de resultados y la difusión del conocimiento generado. Este planteamiento proporciona un marco estructurado y coherente que asegura la reproducibilidad y validez científica de los resultados obtenidos, facilitando la integración de los artículos publicados y la redacción de la memoria doctoral.

2.1. Objetivos

El objetivo general de esta tesis es **desarrollar, analizar y evaluar la aplicación de técnicas de DL para la detección automatizada de defectos en tejidos**, proporcionando un sistema más objetivo y consistente que la inspección manual, que sirva como modelo para futuras aplicaciones industriales.

Para alcanzar este objetivo general, se plantean los siguientes objetivos específicos:

OE1: **Disponer de un dataset** curado de imágenes de tejidos, capturado en condiciones reales de producción que sirva como base de referencia para la investigación y validación de modelos de detección de defectos.

OE2: **Diseñar y configurar modelos CNN preentrenados** para la detección de defectos en tejidos Batavia y Sarga.

OE3: **Diseñar y configurar una metodología híbrida basada en AE y CNN** para mejorar la detección de defectos en tejidos.

OE4: **Analizar el desempeño de los modelos** preentrenados CNN y del enfoque híbrido AE + CNN para la detección de defectos en tejidos Batavia y Sarga.

OE5: **Promover la difusión del conocimiento** científico generado y la **transferencia tecnológica** de los resultados hacia la empresa TS y la sociedad, mediante la consolidación de avances en la inspección automatizada de tejidos y el respaldo a las publicaciones derivadas de la tesis.

2.2. Metodología

La metodología seguida en esta tesis se estructura en varias etapas complementarias para garantizar un enfoque riguroso y reproducible en el desarrollo de sistemas de inspección automatizada de tejidos. Dicha estructura se divide en 12 fases clave, abarcando desde la formulación del problema hasta la difusión científica de los resultados, como se ilustra en el diagrama de flujo general de la metodología (Figura 2).

1. **Formulación del problema:** La investigación parte de la necesidad de mejorar los procesos de inspección de calidad textil, tradicionalmente basados en la observación manual, los cuales presentan limitaciones en precisión, consistencia y velocidad. Este desafío motiva el desarrollo de metodologías basadas en DL que permitan automatizar la detección de defectos en tejidos.

2. **Revisión bibliográfica:** Se llevó a cabo un análisis exhaustivo de la literatura relacionada con el control de calidad textil y las técnicas de DL, con especial atención a CNN y AE. Los hallazgos de esta revisión se reflejan en los apartados correspondientes de los artículos que componen la tesis. Para esta fase se consultaron las principales bases de datos científicas, priorizando los estudios más relevantes y recientes en la materia, con el fin de fundamentar sólidamente la metodología propuesta.

3. **Adquisición de imágenes:** Las imágenes fueron capturadas en noviembre de 2022 en la empresa TS, bajo condiciones reales de producción industrial. Se registraron un total de 2.755 imágenes de tejido Batavia y 1.548 imágenes de tejido Sarga, ambas en formato escala de grises de 16 bits y resolución 2048 $\times$ 696 píxeles.

4. **Preprocesamiento de imágenes:** Las imágenes originales de 16 bits se convirtieron a 8 bits para reducir el tamaño de los ficheros y facilitar su procesamiento computacional, manteniendo una calidad visual adecuada para la detección de defectos. Este procedimiento consistió en normalizar los valores de intensidad de cada imagen, que originalmente se encontraban en el rango de 0 a 65.535, escalándolos al rango de 0 a 255 propio de las imágenes de 8 bits. De esta manera, se mantiene una representación visual adecuada de las imágenes, suficiente para la detección de defectos, mientras que el tamaño en disco se reduce significativamente (aproximadamente de 1,5 MB a 1 MB por imagen). Esta conversión se realizó de manera automatizada mediante un *pipeline* en Python, asegurando que se conservaran las dimensiones originales de las imágenes y permitiendo su

posterior clasificación y procesamiento. Posteriormente, cada imagen se segmentó en 12 subimágenes (patches) de 365×365 píxeles con un pequeño solapamiento horizontal y vertical, siguiendo una matriz de 2×6 fragmentos. Este proceso se implementó mediante librerías de Python para manipulación de archivos e imágenes.

5. **Etiquetado por expertos:** Los patches resultantes fueron clasificados manualmente por especialistas de TS en dos categorías: imágenes con defectos (*cases*) e imágenes sin defectos (*controls*). Durante este proceso se eliminaron 4.367 imágenes ambiguas del tejido Batavia, en las que no se alcanzó consenso sobre la presencia de defecto. Este filtrado fue esencial para garantizar la fiabilidad del dataset y minimizar el ruido en el entrenamiento de modelos. Como resultado de este proceso de segmentación, se generó un conjunto de datos estructurado por tipo de tejido y clase. En el caso del tejido Batavia, se obtuvieron un total de 8.782 patches etiquetados como *case*, lo que representa aproximadamente el 30,6 % del conjunto, y 19.911 patches etiquetados como *controls*, correspondientes al 69,4 %. Para el tejido Sarga, el conjunto final está compuesto por 173 *cases*, que suponen alrededor del 0,9 % del total, y 18.403 *controls*, equivalentes al 99,1 %. Esta distribución evidencia un marcado desbalance de clases, especialmente acusado en el tejido Sarga, lo que supone un reto adicional para el entrenamiento de los modelos de aprendizaje automático y justifica la adopción de métricas de evaluación robustas y estrategias específicas para el tratamiento del desbalance durante el proceso de modelado.

6. **Documentación del dataset:** Se organizó y documentó sistemáticamente el conjunto de datos, incluyendo metadatos relevantes y la estructura jerárquica por tipo de tejido. Esta documentación permitió la publicación del dataset como recurso abierto, asegurando la reproducibilidad de los experimentos y facilitando la transferencia de conocimiento a la comunidad científica y a la industria. Además, este trabajo se materializó en el artículo 1 de la tesis, donde se describe detalladamente el dataset y su potencial aplicación en la inspección automatizada de tejidos.

7. **Selección de modelos IA:** Se identificaron y seleccionaron diferentes arquitecturas de CNN preentrenadas para evaluar su rendimiento en la detección de defectos en tejidos Batavia y Sarga. La elección de redes preentrenadas permitió aprovechar conocimientos previos de la visión por computador y reducir el tiempo de entrenamiento, adaptándolas al contexto específico del dataset. Además se utilizó un AE con el objetivo de un prefiltrado.

8. **Configuración de hiperparámetros:** Se ajustaron los parámetros clave de entrenamiento, como la tasa de aprendizaje, el tamaño de batch y el número de épocas, optimizando el rendimiento de cada modelo. Este proceso se realizó mediante pruebas iterativas y validación cruzada, garantizando un equilibrio entre precisión, generalización y eficiencia computacional.

9. **Experimentación:** Los modelos seleccionados se entrenaron utilizando el dataset curado, aplicando técnicas de validación cruzada, Data Augmentation (DA) y monitorización de métricas de desempeño para evitar sobreajuste.

10. **Evaluación y comparación:** Para evaluar los modelos de clasificación se siguió el enfoque clásico propuesto por Sokolova and Lapalme [43], quienes detallan un conjunto amplio de métricas derivadas de la matriz de confusión, incluyendo precisión, recall, F1-score entre otras. Además, se empleó la métrica AU-ROC, una medida umbral-independiente que resume la capacidad del modelo para discriminar entre clases positivas y negativas en toda la gama de posibles umbrales, tal como define Fawcett [44] en su introducción al análisis ROC.

11. **Interpretación de resultados:** Se evaluó el desempeño de las CNN y de la metodología híbrida AE + CNN de forma diferenciada para cada tipo de tejido, identificando las métricas más representativas de precisión, recall, F1-score y AU-ROC . Este análisis permitió determinar qué configuraciones y enfoques ofrecían el mejor equilibrio entre sensibilidad, robustez y capacidad de generalización para cada tejido, así como valorar la contribución del prefiltrado mediante AE.

12. **Conclusiones y difusión científica:** A partir de los resultados, se extrajeron conclusiones que respaldaron la publicación de los artículos 2 y 3 y la redacción de la memoria doctoral. Además, se promovió la transferencia de conocimiento a la empresa TS y se fomentó la difusión científica hacia la sociedad, contribuyendo a la consolidación de prácticas de inspección automatizada de tejidos y al avance del control de calidad textil mediante DL .

La Figura 2 muestra el diagrama de flujo general de la metodología seguida en esta tesis. En él se representan de manera esquemática las distintas fases del trabajo. Este esquema permite visualizar de forma clara la secuencia de etapas y su interrelación, facilitando la comprensión del enfoque metodológico adoptado.

Figura 2: Diagrama de flujo general de la metodología de la tesis.

3. Publicaciones Incluidas en la Tesis

En esta sección se presentan los artículos científicos que constituyen el compendio de la tesis doctoral. Cada publicación aporta un enfoque específico dentro del desarrollo de sistemas basados en DL para la detección automática de defectos en tejidos industriales Batavia y Sarga, abarcando desde la generación de datasets y el diseño de modelos de CNN hasta la implementación de metodologías combinadas.

Los artículos completos se reproducen en la sección 4, respetando su formato y numeración original de publicación.

3.1. Criterios y resumen de selección de artículos

La presente tesis doctoral se desarrolla bajo la modalidad de **compendio de publicaciones**, de acuerdo con lo establecido en la normativa vigente de la UBU. Esta modalidad se justifica por la naturaleza científica y aplicada del trabajo realizado, que ha dado lugar a resultados publicables y revisados por pares en revistas.

Los artículos incluidos abordan de manera complementaria los objetivos de la investigación: la creación de un dataset curado de imágenes textiles, el desarrollo y validación de modelos de DL para la detección de defectos, y la propuesta de metodologías para la inspección automatizada de tejidos. La coherencia temática y metodológica se asegura al compartir todos los trabajos un objetivo común: el desarrollo de sistemas de inspección automatizada mediante IA, centrados en la calidad textil, la detección de defectos y el análisis de imágenes. Metodológicamente, mientras que el primer artículo se centra en la publicación y descripción del dataset, el segundo y tercero emplean DL utilizando este dataset o subconjuntos del mismo, evaluando los resultados con métricas comparables, lo que garantiza un enfoque homogéneo y complementario.

Esta modalidad también permite una difusión más amplia de los resultados, potenciando el impacto científico y la transferencia de conocimiento hacia la industria textil, así como la consolidación de líneas de investigación activas en el Grupo GICAP de la UBU.

El compendio incluye tres artículos que reflejan los avances alcanzados en esta tesis en el control de calidad textil mediante técnicas DL y un dataset curado:

- **Artículo 1:** *Dataset for defect detection in textile manufacturing* [45]. Presenta un dataset público de imágenes de alta resolución de tejidos, recogido en condiciones reales de producción en la empresa TS. Este dataset constituye una base sólida para el desarrollo y la evaluación de modelos de DL orientados a la detección de defectos en la fabricación textil. DOI: https://doi.org/10.1016/j.dib.2025.111451

- **Artículo 2:** *Defect Detection in Batavia and Sarga Woven Fabrics by Means of Convolutional Neural Networks.* A partir del dataset generado en el artículo 1, se analiza el rendimiento de distintas arquitecturas CNN aplicadas a la detección de defectos en tejidos Batavia y Sarga. (Aceptado para su publicación en la revista Journal of Applied Logics).

- **Artículo 3:** *Deep Learning for Quality Control in Woven Fabrics* [46]. Propone una metodología combinada en dos etapas (AE + CNN) orientada a optimizar la detección de defectos en tejidos. DOI: https://doi.org/10.52152/D11517

A continuación, con el fin de evidenciar la correspondencia entre los objetivos de la tesis y las contribuciones científicas derivadas de la investigación, la Tabla 3 muestra cómo los tres artículos incluidos en el compendio contribuyen de manera conjunta al cumplimiento del objetivo general. Asimismo, se detalla la aportación de cada artículo a los objetivos específicos, ofreciendo una visión estructurada del progreso de la línea de investigación y de la manera en que cada publicación refuerza los distintos aspectos planteados.

Tabla 3: Relación entre los objetivos de la tesis y las publicaciones.

Publicación	**Objetivo**	**Descripción de la contribución al objetivo**
Artículo 1	General; OE1; OE5	Cumple el **Objetivo Específico 1**, construyendo y publicando un dataset de referencia para la investigación en detección de defectos en tejidos.
Artículo 2	General; OE2; OE4; OE5	Aborda el **Objetivo Específico 2**, mediante el desarrollo y evaluación de arquitecturasCNNaplicadas a tejidos Batavia y Sarga, y el **Objetivo Específico 4**, comparando el desempeño de los modelos según el tipo de tejido.
Artículo 3	General; OE3; OE4; OE5	Cumple el **Objetivo Específico 3**, proponiendo una metodología híbrida basada en AE y CNN, así como el **Objetivo Específico 4**, evaluando su desempeño.
Todos los artículos contribuyen al **Objetivo General** de la tesis y al **Objetivo Específico 5**, mediante la difusión de resultados científicos y la promoción de la transferencia tecnológica hacia la empresa TS.		

Además de estas publicaciones incluidas en el compendio que conforma la presente tesis, la investigación llevada a cabo ha generado las siguientes publicaciones en congresos, concretamente en International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO):

- Velasco-Pérez, N., Lozano-Juárez, S., Gil-Arroyo, B., Sanz, J. M., Basurto, N., Urda, D., & Herrero, Á. (2023). Defect Detection in Batavia Woven Fabrics by Means of Convolutional Neural Networks. In 18th International SOCO (pp. 205–215). Cham, Switzerland: Springer Nature Switzerland. ISBN 978-3-031-42536-3. https://doi.org/10.1007/978-3-031-42536-3_20 [47].

  Resumen:

  El control de calidad constituye una de las etapas clave en cualquier proceso de fabricación en general, y en el sector textil en particular. En la actualidad, el proceso de inspección en la industria textil se realiza principalmente mediante inspección visual llevada a cabo por personal cualificado, ya que las soluciones comerciales existentes presentan limitaciones significativas. Por ello, el presente estudio propone y valida modelos de DL aplicados al control automático de la calidad de tejidos producidos en condiciones reales de fabricación, caracterizadas por alta velocidad, complejidad y tecnología avanzada. En concreto, se validan CNN sobre imágenes reales obtenidas directamente de la línea de producción de tejidos Batavia. Los resultados satisfactorios obtenidos en los experimentos respaldan la aplicabilidad de estos modelos en esta compleja tarea de inspección automatizada.

- Gil-Arroyo, B., Velasco-Pérez, N., Sanz, J. M., Arroyo, Á., Urda, D., & Herrero, Á. (2025, junio). Impact of Data Augmentation on Woven Fabrics Defect Detection by means of Deep Learning. Comunicación presentada en 20th International SOCO, Salamanca, España. (Actas en prensa).

  Resumen:

  En este estudio se analizó el impacto de diversas técnicas de DA en la detección de defectos en tejidos Batavia, utilizando un conjunto de datos público compuesto por imágenes de alta resolución procedentes de procesos de fabricación textil. Se comparó un modelo base entrenado sin DA con modelos entrenados aplicando tanto técnicas individuales como combinadas, incluyendo volteos horizontales y verticales, rotaciones, blanqueamiento Zero-phase Component Analysis y ajustes de brillo. Los experimentos se llevaron a cabo mediante validación cruzada estratificada de cinco particiones (5-fold stratified cross-validation), y el rendimiento de la CNN DenseNet121 se evaluó utilizando métricas adecuadas para tareas de clasificación binaria desbalanceada. Los resultados muestran que no todas las estrategias de DA conducen a mejoras sustanciales: mientras que ciertas combinaciones ofrecieron el mejor equilibrio entre rendimiento y tiempo de entrenamiento, otras transformaciones tuvieron un impacto mínimo o incluso negativo. Además, los resultados obtenidos demuestran que el valor del AU-ROC se mantiene estable frente a la mayoría de las transformaciones simples, lo que sugiere que DA, si bien es útil, no siempre produce una mejora apreciable en el desempeño del modelo.

3.2. Artículo 1: "Dataset for defect detection in textile manufacturing"

3.2.1. Referencia completa

A continuación en la Tabla 4 se presenta la referencia completa del artículo relacionado con la generación del dataset utilizado en esta tesis. La información incluye autores, afiliaciones, revista, DOI, y métricas de impacto editorial.

Campo	**Información**
Autores	Beatriz Gil-Arroyo[1], Juan Marcos Sanz[2], Ángel Arroyo[1], Daniel Urda[1], Nuño Basurto[1] y Álvaro Herrero[1]
Afiliaciones	[1] Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Av. Cantabria s/n, 09006, Burgos, España. `http://www.ubu.es` [2] Textil Santanderina, Cabezón de la Sal, España. `http://www.textilsantanderina.com`
Título	Dataset for defect detection in textile manufacturing
Revista	Data in Brief. ISSN 2352-3409
Volumen	59
Página	111451
Editorial	Elsevier
Año de publicación	2025
DOI	`https://doi.org/10.1016/j.dib.2025.111451`
URL	`https://www.sciencedirect.com/science/article/pii/S2352340925001830`
Journal Impact Factor 2024	1.4 Q3 en la categoría MULTIDISCIPLINARY SCIENCES
Journal Citation Indicator 2024	0.32 Q3 en la categoría MULTIDISCIPLINARY SCIENCES

Tabla 4: Referencia completa del artículo: Dataset for defect detection in textile manufacturing.

3.2.2. Resumen

Este dataset, recopilado durante noviembre de 2022 en TS, un fabricante textil líder con sede en Cabezón de la Sal (Cantabria, España), comprende imágenes de alta resolución de tejidos Batavia y Sarga. Las imágenes se capturaron como parte de un proyecto destinado a documentar y analizar los complejos tejidos y patrones de estos materiales. Utilizando una cámara de alta resolución bajo condiciones de iluminación controladas, se obtuvieron imágenes detalladas para garantizar una calidad consistente y una representación precisa de la textura y el color del tejido.

El dataset se proporciona en formato procesado, en el que las imágenes han sido reducidas de 16 bits a 8 bits, recortadas y clasificadas en casos y controles. El principal potencial de reutilización de este conjunto de datos radica en su aplicación para modelos de IA y DL orientados a la detección de defectos en la fabricación textil. Al aprovechar estas imágenes procesadas de alta calidad, los investigadores y desarrolladores pueden entrenar modelos para identificar y clasificar distintos tipos de defectos en los tejidos, como inconsistencias en el tejido, variaciones de color e irregularidades superficiales. Este enfoque puede mejorar significativamente la eficiencia y precisión de los procesos de control de calidad en la producción textil.

Además, el dataset constituye un recurso valioso para la investigación académica en ingeniería textil y ciencia de materiales. Puede utilizarse para estudiar las propiedades y comportamientos de los tejidos Batavia y Sarga bajo diferentes condiciones, contribuyendo a avances en el diseño de telas y en técnicas de fabricación. La información visual detallada proporcionada por las imágenes procesadas también respalda el desarrollo de nuevas metodologías para la inspección automatizada y aseguramiento de la calidad textil.

Al poner este dataset a disposición, TS y la UBU buscan apoyar la innovación y mejora del control de calidad textil mediante soluciones impulsadas por IA, fomentando la colaboración y el desarrollo dentro de la industria.

Palabras clave: Fabricación textil, Industria textil, Tejidos Batavia y Sarga, Detección de defectos, Análisis de imágenes, Visión inteligente, Control de calidad.

3.2.3. Contribución personal

La contribución personal al artículo "Dataset for defect detection in textile manufacturing" se describe según el sistema Contributor Roles Taxonomy (CRediT):

- **Beatriz Gil-Arroyo:** Redacción del borrador original (Writing – Original Draft) y visualización de resultados (Visualization).

- **Juan Marcos Sanz:** Validación de resultados (Validation) y provisión de recursos (Resources).
- **Ángel Arroyo:** Análisis formal (Formal Analysis), investigación (Investigation), supervisión (Supervision) y revisión y edición del manuscrito (Writing – Review & Editing).
- **Daniel Urda:** Metodología (Methodology), desarrollo de software (Software), curación de datos (Data Curation), gestión del proyecto (Project Administration) y revisión y edición del manuscrito (Writing – Review & Editing).
- **Nuño Basurto:** Revisión y edición del manuscrito (Review & Editing).
- **Álvaro Herrero:** Conceptualización del proyecto (Conceptualization), supervisión (Supervision) y obtención de financiación (Funding Acquisition).

3.3. Artículo 2: "Defect Detection in Batavia and Sarga Woven Fabrics by Means of Convolutional Neural Networks"

3.3.1. Referencia completa

La Tabla 5 presenta la referencia del segundo artículo incluido en esta tesis. La información recopilada detalla los autores, sus afiliaciones, la publicación, así como las métricas de impacto editorial, facilitando la contextualización del artículo dentro del marco del presente trabajo. Este artículo está aceptado y está pendiente de publicación, por esta razón no aparece el volumen, páginas, DOI y URL.

3.3.2. Resumen

El aseguramiento de la calidad constituye una fase fundamental en todos los procesos de fabricación, particularmente en la industria textil. Actualmente, las inspecciones textiles dependen en gran medida de la evaluación visual humana debido a las deficiencias de las soluciones comerciales existentes. En respuesta, esta investigación propone la adopción de modelos de DL para automatizar el control de calidad de los tejidos en entornos de producción dinámicos y modernos.

En particular, se analizan CNN utilizando imágenes reales obtenidas de las líneas de producción de tejidos Batavia y Sarga. Los resultados experimentales identifican a DenseNet121 e InceptionV3 como los modelos más eficaces para los tejidos Batavia y

Campo	Información
Autores	Beatriz Gil-Arroyo[1], Nuria Velasco-Pérez[1], Juan Marcos Sanz[2], Abraham Casas[3], Ángel Arroyo[1] y Daniel Urda[1]
Afiliaciones	[1] Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Av. Cantabria s/n, 09006, Burgos, España. http://www.ubu.es [2] Textil Santanderina, Cabezón de la Sal, España. http://www.textilsantanderina.com [3] Centro Tecnológico de Componentes-CTC, Parque Científico y Tecnológico de Cantabria (PCTCAN), c/ Isabel Torres 1, 39011, Santander, España. https://centrotecnologicoctc.com/
Título	Defect Detection in Batavia and Sarga Woven Fabrics by Means of Convolutional Neural Networks
Revista	Journal of Applied Logics – IfCoLoG Journal of Logics and their Applications. ISSN 2631-9810 / eISSN 2631-9829
Volumen	-
Página	-
Editorial	College Publications
Año de publicación	-
DOI	-
URL	-
Journal Impact Factor 2024	0.2 Q4 en la categoría LOGIC
Journal Citation Indicator 2024	0.52 Q3 en la categoría LOGIC

Tabla 5: Referencia completa del artículo: Defect Detection in Batavia and Sarga Woven Fabrics by Means of Convolutional Neural Networks.

Sarga, respectivamente. DenseNet121 muestra un rendimiento equilibrado en las métricas clave para el tejido Batavia, mientras que InceptionV3 destaca en el tejido Sarga, especialmente en F1-score y AU-ROC .

Estos hallazgos subrayan el potencial de los modelos de DL para mejorar la precisión y eficiencia del control de calidad textil.

Palabras clave: Detección de defectos textiles, Control de calidad de tejidos, Industria 4.0, Deep Learning en manufactura, Redes neuronales convolucionales, Análisis y clasificación de imágenes.

3.3.3. Contribución personal

En el artículo "Defect Detection in Batavia and Sarga Woven Fabrics by Means of Convolutional Neural Networks", la contribución personal también se describe según CRediT:

- **Beatriz Gil-Arroyo:** Redacción del borrador original (Writing – Original Draft), curación de datos (data curation), investigación (Investigation) y visualización de resultados (Visualization).

- **Nuria Velasco-Pérez** Revisión y edición del manuscrito (Review & Editing).
- **Juan Marcos Sanz:** Validación de resultados (Validation) y provisión de recursos (Resources).
- **Abraham Casas** Revisión y edición del manuscrito (Review & Editing).
- **Ángel Arroyo:** Análisis formal (Formal Analysis), revisión y edición del manuscrito (Writing – Review & Editing) y obtención de financiación (Funding Acquisition).
- **Daniel Urda:** Conceptualización del proyecto (Conceptualization), Metodología (Methodology), desarrollo de software (Software), gestión del proyecto (Project Administration), supervisión (Supervision), revisión y edición del manuscrito (Writing – Review & Editing) y obtención de financiación (Funding Acquisition).

3.4. Artículo 3: "Deep Learning for Quality Control in Woven Fabrics"

3.4.1. Referencia completa

La Tabla 6 presenta la referencia completa del tercer artículo incluido en esta tesis, titulado *Deep Learning for Quality Control in Woven Fabrics*.

3.4.2. Resumen

La detección automatizada de defectos en los tejidos es un reto clave en el control de calidad dentro de la industria textil. Este estudio propone una metodología basada en DL para identificar defectos en los tejidos Batavia y Sarga. En la primera etapa, se utilizó un AE para filtrar imágenes anómalas, lo que permitió crear un dataset con número suficiente de casos defectuosos, que de otro modo serían difíciles de obtener en la producción textil.

Posteriormente, se entrenaron CNN (DenseNet121, EfficientNetB0/B3, Xception y glsvgg utilizando DA y validación cruzada estratificada. Para los tejidos Batavia, DenseNet121 alcanzó AU-ROC de 0,88 y AU-PR de 0,93, lo que demuestra una alta capacidad de detección. Para los tejidos Sarga, se consideraron tres referencias diferentes, que mostraron un rendimiento más variable entre los modelos y los datasets. No obstante, modelos como ResNet101 y Xception lograron resultados competitivos.

Campo	Información
Autores	Beatriz Gil-Arroyo[1], Marco Melgarejo[2], Abraham Casas[2], Alejandro López[2], Juan Marcos Sanz[3] y Daniel Urda[1]
Afiliaciones	[1] Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Av. Cantabria s/n, 09006, Burgos, España. http://www.ubu.es [2] Centro Tecnológico de Componentes–CTC, Parque Científico y Tecnológico de Cantabria (PCTCAN), c/ Isabel Torres 1, 39011, Santander, España. https://centrotecnologicoctc.com/ [3] Textil Santanderina, Cabezón de la Sal, España. http://www.textilsantanderina.com
Título	Defect Detection in Batavia and Sarga Woven Fabrics by Means of Convolutional Neural Networks
Título	Deep Learning for Quality Control in Woven Fabrics.
Revista	DYNA Ingeniería e Industria. ISSN-L 0012-7361
Volumen	100
Página	545–551
Editorial	Publicaciones DYNA S.L.
Año de publicación	2025
DOI	https://doi.org/10.52152/D11517
URL	https://www.revistadyna.com/search/deep-learning-for-quality-control-in-woven-fabrics
Journal Impact Factor 2024	0.7 Q3 en la categoría ENGINEERING, MULTIDISCIPLINARY
Journal Citation Indicator 2024	0.18 Q3 en la categoría ENGINEERING, MULTIDISCIPLINARY

Tabla 6: Referencia completa del artículo: Deep Learning for Quality Control in Woven Fabrics.

Los resultados indican que la combinación de AE y CNN facilita la generación de datasets equilibrados y permite una detección de defectos consistente, aunque el rendimiento depende del tipo de tejido y de la referencia específica, lo que sugiere que la selección del modelo debe adaptarse a las características de cada caso.

Palabras clave: Detección de defectos, Textil, Industria 4.0, Deep Learning, Redes neuronales convolucionales, Análisis de imágenes, Autoencoder.

3.4.3. Contribución personal

Para "Deep Learning for Quality Control in Woven Fabrics", la contribución personal también se presenta bajo el sistema CRediT:

- **Beatriz Gil-Arroyo:** Redacción del borrador original (Writing – Original Draft), curación de datos (data curation), análisis formal (Formal Analysis), investigación (Investigation) y visualización de resultados (Visualization).

- **Marco Melgarejo** desarrollo de software (Software), Revisión y edición del manuscrito (Review & Editing).
- **Abraham Casas** desarrollo de software (Software) y redacción del borrador original (Writing – Original Draft).
- **Alejandro López** desarrollo de software (Software) y redacción del borrador original (Writing – Original Draft).
- **Juan Marcos Sanz:** Validación de resultados (Validation) y provisión de recursos (Resources).
- **Daniel Urda:** Conceptualización del proyecto (Conceptualization), Metodología (Methodology), desarrollo de software (Software), gestión del proyecto (Project Administration), supervisión (Supervision), revisión y edición del manuscrito (Writing – Review & Editing) y obtención de financiación (Funding Acquisition).

4. Artículos Completos

4.1. Artículo 1: "Dataset for defect detection in textile manufacturing"

Data in Brief 59 (2025) 111451

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Dataset for defect detection in textile manufacturing

Beatriz Gil-Arroyo[a,*], Juan Marcos Sanz[b], Ángel Arroyo[a], Daniel Urda[a], Nuño Basurto[a], Álvaro Herrero[a]

[a] *Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Av. Cantabria s/n 09006, Burgos, Spain*
[b] *Textil Santanderina, Avda. Textil Santanderina S/N, 39500 Cabezón de la Sal (Cantabria), Spain*

A R T I C L E   I N F O

Article history:
Received 31 January 2025
Revised 25 February 2025
Accepted 3 March 2025
Available online 7 March 2025

Dataset link: Original and processed dataset of Batavia and Sarga woven fabric images (Original data)

Keywords:
Textile manufacturing
Textile industry
Batavia and Sarga fabric
Defect detection
Image analysis
Artificial Vision
Quality inspection

A B S T R A C T

This dataset, collected during November 2022 at Textil Santanderina, a leading textile manufacturer based in Cabezón de la Sal (Cantabria, Spain), comprises high-resolution images of Batavia and Sarga fabrics. The images were captured as part of a project to document and analyze the intricate weaves and patterns of these fabrics. Using a high-resolution camera under controlled lighting conditions, detailed images were obtained to ensure consistent quality and accurate representation of the fabric's texture and colour. The dataset is provided in processed format, where images have been downscaled from 16 bits to 8 bits, cropped, and classified into cases and controls. The primary reuse potential of this dataset lies in its application for Artificial Intelligence (AI) and Machine Learning (ML) models aimed at defect detection in textile manufacturing. By leveraging these high-quality processed images, researchers and developers can train models to identify and classify various types of fabric defects, such as weave inconsistencies, colour variations, and surface irregularities. This can significantly enhance the efficiency and accuracy of quality control processes in textile production. Additionally, the dataset serves as a valuable resource for academic research in textile engineering and material science. It can be used to study the properties and behaviours of Batavia and Sarga weaves under different conditions, con-

* Corresponding author.
E-mail address: bgarroyo@ubu.es (B. Gil-Arroyo).

https://doi.org/10.1016/j.dib.2025.111451

tributing to advancements in fabric design and manufacturing techniques. The detailed visual information provided by the processed images also supports the development of new methodologies for automated textile inspection and quality assurance. By making this dataset available, Textil Santanderina and University of Burgos aim to support innovation and improvement in textile quality control through AI-driven solutions, fostering collaboration and development within the industry.

Specifications Table

Subject	Computer Sciences
Specific subject area	Textil Santanderina: Experts in high-quality fabrics since 1923, committed to global innovation and sustainability.
Type of data	Processed images.
Data collection	The imaging system employs a Basler raL camera with the Awaiba DR-12k-3.5 CMOS sensor, with a resolution from 8 kHz to 12k [1]. Operating within the VIS-NIR bandwidth, with resolution of 10 pixels/mm. A LED array emitting at 850 nm is utilized, covering a field of 15 mm × 1510 mm, situated at a height of 20 cm above the tissue, comprises 360 infrared (IR) LEDs at 850 nm wavelength, each with an incident angle set at 15 degrees. This IR illumination prevents interference with existing standard D65 illumination systems commonly used in visual inspection.
Data source location	The data were collected in: Institution Textil Santanderina City/Town/Region: Cabezón de la Sal, Cantabria Country: Spain Geographical coordinates: 43.31164220450848, -4.225763758747999
Data accessibility	Repository name: Institutional Repository of the University of Burgos Data identification number: *DOI:* 10.36443/10259/9965 Direct URL to data: 10.36443/10259/9965 [2]
Related research article	B. Gil-Arroyo, N. Velasco-Pérez, J.M. Sanz, A. Casas, A. Arroyo, D. Urda, Defect detection in Batavia and Sarga woven fabrics by means of convolutional neural networks, J. Appl. Logics, submitted for publication, 2025 [3].

1. Value of the Data

- This dataset facilitates the integration of artificial intelligence and computer vision in the textile industry, serving as a valuable resource to train models for real-time defect detection in textile patterns. By incorporating machine learning models trained on this dataset into production lines, it enhances the efficiency and objectivity of defect detection, improving the ability to identify issues before final products leave the factory.
- The dataset provides a solid foundation for developing and training more advanced machine learning models such as categorizing defects by typology or classifying different twill weave patterns.
- The dataset could be valuable in training and coaching new employees in the textile industry.
- Production process experts can identify patterns and areas for improvement in machinery and product quality.
- Researchers exploring applications beyond the textile industry, such as image recognition and pattern analysis, can benefit from the diverse and well-labeled twill fabric dataset.

2. Background

In Spain, the textile industry plays a crucial role in both the economy and cultural heritage [4]. As a significant contributor to employment and Gross Domestic Product (GDP), it fosters innovation and sustainability in fabric production [5]. Machine learning (ML) is revolutionizing this sector by enhancing quality control and fault detection processes [6,7,8,9]. Algorithms can analyze vast datasets to predict and prevent defects in textiles, ensuring higher product standards and reducing waste. Moreover, ML applications extend beyond quality assurance; they optimize production processes by predicting demand, managing inventory efficiently, and streamlining supply chains. This predictive capability not only boosts productivity but also minimizes costs and environmental impact through better resource utilization. Looking ahead, advancements in ML-driven automation promise further improvements in textile manufacturing, from automated looms to smart factories. By integrating these technologies, Spain's textile industry can maintain its competitive edge globally, embracing sustainability and innovation while meeting evolving consumer demands. This convergence of tradition and technology underscores the pivotal role of textiles in Spain's industrial landscape, driving economic growth and sustainable development into the future.

3. Data Description

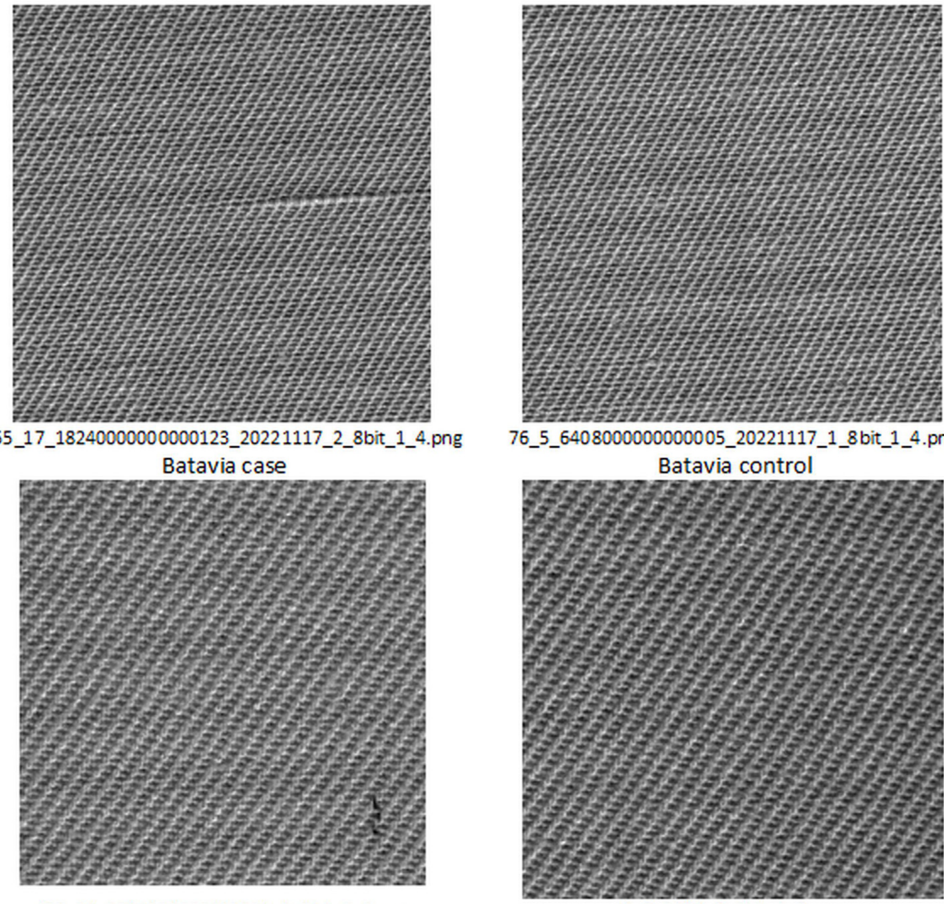
The dataset has two different weaves (Batavia and Sarga). These textiles were selected due to their distinctive characteristics and relevance in various textile applications.

Batavia fabric, a type of twill weave within Sarga fabrics, blends 85% cotton and 15% linen. In November 2022, 2755 high-resolution grayscale images (16-bit, 2048 × 696 pixels) were collected, focusing on a neutral greige fabric.

Sarga fabric, known for its diagonal weave pattern, offers durability and versatility for various textile applications. In November 2022, 1548 high-resolution grayscale images (16-bit, 2048 × 696 pixels) were gathered, featuring a neutral greige fabric to minimize colour variability.

The dataset is organized by two types of woven fabrics: Batavia and Sarga, each containing images in PNG format. Within each fabric category, there are specific directories:

1. Batavia:
    - Originals: Contains 2755 8-bit PNG images of size 2048 × 696 pixels.
    - Patches:
        - Cases: Includes 8782 cropped PNG images (365 × 365 pixels) with defects.
        - Controls: Includes 19,911 cropped PNG images (365 × 365 pixels) without defects.
        - info_patches.csv: A CSV file with columns for image and patch names, operator's and ground truth labels (0 for no defect, 1 for defect).
        - Note: In Batavia dataset, the total number of patches does not equal the number of originals multiplied by 12 due to a review by Textil Santanderina. During this process, where 4367 ambiguous patches were excluded as their labels (case/control) were unclear.
2. Sarga:
    - Originals: Contains 1548 8-bit PNG images of size 2048 × 696 pixels.
    - Patches:
        - Cases: Includes 173 cropped PNG images (365 × 365 pixels) with defects.
        - Controls: Includes 18,403 cropped PNG images (365 × 365 pixels) without defects.
        - info_patches.csv: A CSV file structured similarly to Batavia's file, providing image and patch details along with defect labels.
        - Note: In Sarga dataset, the total number of patches equals the number of originals multiplied by 12.



255_17_18240000000000123_20221117_2_8bit_1_4.png
Batavia case

76_5_6408000000000005_20221117_1_8bit_1_4.png
Batavia control

455_31_2624000000000031_2_8bit_2_3.png
Sarga case

0_0_9304_0_8bit_1_5.png
Sarga control

Fig. 1. Examples of patches images.

The directories patches contain the 365 × 365 cropped images with small overlapping areas on both the horizontal and vertical axis.

In Fig. 1, examples of patches can be observed. For each of the two fabrics, there is an example of an image with a defect (case) and another without a defect (control).

The PNG file naming convention follows:

- Original images: Identified by a numeric identifier, the original folder number from Textil Santanderina and followed by the suffix "8bit" For example: 0_0_9304_20221117_0_8bit.png
  - 0_0_9304 is the numeric ID.
  - 20221117 is the acquisition date of the image, in this example November 17, 2022. In Batavia, all the images were taken on November 17, 2022, while in Sarga, this date does not appear in the image name, but they were also taken in November 2022.
  - 0 is the original folder. The range goes from 0 to 2
  - 8bit is the suffix.
  - png is the file extension.
- Patches: Named using the original image name, row, and column indices, referencing the 2 × 6 matrix slicing of the original images into 12 segments of size 365 × 365 pixels, with a small overlap in both the x and y axes. For example: 0_0_9304_20221117_0_8bit_2_1.png:

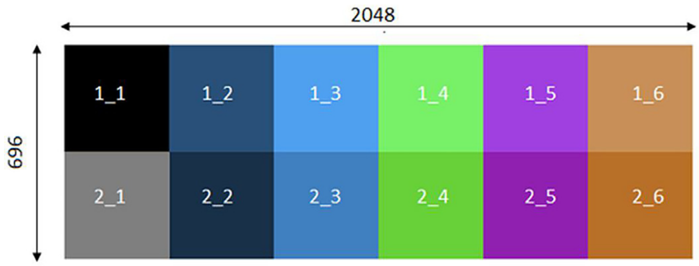


Fig. 2. 2 × 6 matrix slicing.

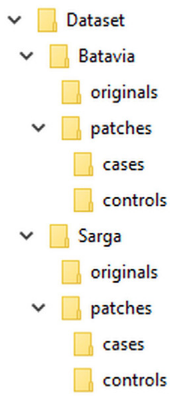


Fig. 3. Directory structure.

Table 1
Slicing details.

Patch	Coordinates (x,y)	Width (pixel)	Height (pixels)
1_1	(0,0)	365	365
1_2	(0, 336)	365	365
1_3	(0, 672)	365	365
1_4	(0, 1008)	365	365
1_5	(0, 1344)	365	365
1_6	(0, 1680)	365	365
2_1	(0, -330)	365	365
2_2	(336, -330)	365	365
2_3	(672, -330)	365	365
2_4	(1008, -330)	365	365
2_5	(1344, -330)	365	365
2_6	(1680, -330)	365	365

- 0_0_9304_20221117_0_8bit is the original image name.
- 2 is the row of the matrix.
- 1 is the column of the matrix.

In the process of slicing the original images, the 2048 × 696 pixel images were divided into 12 segments of 365 × 365 pixels each, following a 2 × 6 matrix with 2 rows and 6 columns. This slicing was carried out to facilitate analysis and experimentation. Fig. 2 illustrates this 2 × 6 matrix slicing of the original 2048 × 696 images into 12 patches of size 365 × 365 pixels each.

Table 1 presents the specific details of this slicing, including the starting coordinates of each patch within the original image. The point (0,0) is considered to be the top-left corner of the

Table 2
Dataset's folder structure and number of images.

Folder	N° images
Batavia\originals	2755
Batavia\patches\cases	8782
Batavia\patches\controls	19,911
Sarga\originals	1548
Sarga\patches\cases	173
Sarga\patches\controls	18,403

2048 × 696 image. The coordinates (x, y) correspond to the top-left corner of each patch. The slicing includes an overlap of 29 pixels in the x-axis and 35 pixels in the y-axis.

In Fig. 3, the directory structure of the dataset is shown. This fig. illustrates the hierarchical organization of directories and files comprising the dataset. The structured design facilitates efficient management and access to the dataset samples, essential for conducting any analysis. The directory structure is organized by fabrics (Batavia and Sarga). Within each fabric, there are original images of size 2,048 × 696 pixels and patches. Inside the 'patches' folder, there are two subfolders: 'cases' and 'controls', and a CSV file named 'info_patches' containing the detailed information mentioned above.

Table 2 presents an overview of the dataset's folder structure and the corresponding number of images included in each category. The dataset is organized into two main types of fabrics, Batavia and Sarga, each containing original images and patches. This systematic categorization is essential for subsequent research and detailed analysis of the dataset, enabling comprehensive exploration and understanding of the dataset's contents and distributions under different experimental conditions.

4. Experimental Design, Materials and Methods

The system was configured to capture images of Batavia and Sarga in November, 2022.

A total of 2755 images of Batavia weave and 1548 of Sarga weave were obtained. All images were 8-bit 2048 × 696 pixels, instead of 16 bits. This reduction was mainly done to decrease the file size; 16-bit images can represent 65,536 different colours, while 8-bit images have a lower colour depth limited to 256 colours. This reduction in the number of colours can affect the visual quality of the images, but if the additional colour information is not critical for the research, this conversion can simplify image processing.

Each original 8-bit 2048 × 696 image was divided into 12 parts of 365 × 365 pixels with a small overlap in the vertical and horizontal axes. Thus, for Batavia tissue, 28,693(the total number of patches does not equal the number of originals multiplied by 12 due to a review by Textil Santanderina) sliced images were obtained, of which 19,911 correspond to controls and 8782 to cases. Accordingly, 18,576 (12 × 1548) Sarga fabrics 365 × 365 patches were acquired, of which 18,403 controls and 173 cases.

We used different python libraries to image pre-processing, so we can obtain 12 patches from each 2048 × 696 8-bit image. The python libraries used have been:

- Os: for manipulating paths and directories [10].
- Glob: for finding file names that match a search pattern [11].
- PIL (Pillow): for opening, manipulating, and saving images [12].
- Shutil: for file and directory manipulation operations, such as copying and moving [13].

The raw data, pre-processed images, and image pre-processing code has been uploaded to GitHub, given that offers a centralized space for collaborative software development, providing essential tools for version control, project management, and collaboration between developers [14].

Limitations

Not applicable.

Ethics Statement

The authors have read and followed the ethical requirements for publication in Data in Brief and confirming that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

CRediT Author Statement

Beatriz Gil-Arroyo: writing - Original Draft, Visualization. **Juan Marcos Sanz**: Validation, Resources. **Ángel Arroyo**: Formal analysis, Investigation, supervising, writing – Review & Editing. **Daniel Urda**: Methodology, Software, Data Curation, Project administration, writing – Review & Editing. **Nuño Basurto**: Review & Editing. **Álvaro Herrero**: Conceptualization, Supervision, Funding acquisition.

Data Availability

Original and processed dataset of Batavia and Sarga woven fabric images (Original data) (University of Burgos. Public Repository).

Acknowledgements

The funding for this project was provided by the DECENT (Deep Learning for automatic Textile Inspection) initiative under the DIH-World 2nd Open Call framework. The authors express their gratitude to INADE for their collaboration in acquiring the images.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] The Basler racer raL12288-8gm [website], Basler, 2024. https://www.baslerweb.com/en/shop/ral12288-8gm/, 2024 (Accessed 20 February 2025).
[2] B. Gil-Arroyo, N. Velasco-Pérez, N. Basurto, J. M. Sanz, D. Urda, A. Herrero, Original and processed dataset of Batavia and Sarga woven fabric images, , UBU Repository, v1, 2025. https://doi.org/10.36443/10259/9965.
[3] B. Gil-Arroyo, N. Velasco-Pérez, J.M. Sanz, A. Casas, A. Arroyo, D. Urda, Defect detection in Batavia and Sarga woven fabrics by means of convolutional neural networks, J. Appl. Logics (2025) submitted for publication.
[4] H. Oevermann, B.M. Walczak, M. Watson, The Heritage of the Textile Industry, Lodz University of Technology Monographs, 2022, doi:10.34658/9788366741232.
[5] Ministerio de Industria, Comercio y Turismo, Informe sectorial: Textil y confección [website], 2024. https://www.mintur.gob.es/es-es/IndicadoresyEstadisticas/Presentaciones/20sectoriales/02./20Textil/20y/20confeccion.pdf, (Accessed 20 February 2025).
[6] P.R. Jeyaraj, E.R.S. Nadar, Effective textile quality processing and an accurate inspection system using the advanced deep learning technique, Text. Res. J. 90 (9-10) (2020) 971–980, doi:10.1177/0040517519884124.

[7] B. Wei, B. Xu, K. Hao, L. Gao, Textile defect detection using multilevel and attentional deep learning network (MLMA-Net), Text. Res. J. 92 (19-20) (2022) 3462–3477, doi:10.1177/00405175211073773.
[8] I. Koulali, M.T. Eskil, Unsupervised textile defect detection using convolutional neural networks, Appl. Soft Comput. 113 (2021) 107913, doi:10.1016/j.asoc.2021.107913.
[9] H. Zhang, W. Zhang, Y. Wang, S. Lu, L. Yao, X. Chen, Colour-patterned fabric-defect detection using unsupervised and memorial defect-free features, Color. Technol. 138 (6) (2022) 602–620, doi:10.1111/cote.12624.
[10] Python Software Foundation, os — Miscellaneous operating system interfaces [software], 2024. https://docs.python.org/3/library/os.html, (Accessed 20 February 2025).
[11] Python Software Foundation, glob — Unix style pathname pattern expansion [software], 2024. https://docs.python.org/3/library/glob.html, (Accessed 20 February 2025).
[12] Python Software Foundation, Pillow (PIL Fork) documentation [software], 2024. https://pillow.readthedocs.io/en/stable/, (Accessed 20 February 2025).
[13] Python Software Foundation, shutil — High-level file operations [software], 2024. https://docs.python.org/3/library/shutil.html, (Accessed 20 February 2025).
[14] Gicap-UBU, dif_textile [software], 2025. https://github.com/gicap-ubu/dif_textile#contents, (Accessed 20 February 2025).

4.2. Artículo 2: "Defect Detection in Batavia and Sarga Woven Fabrics by Means of Convolutional Neural Networks"

Defect Detection in Batavia and Sarga Woven Fabrics by Means of Convolutional Neural Networks

Beatriz Gil-Arroyo bgarroyo@ubu.es
Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Avenida Cantabria s/n, 09001 Burgos, Spain

Nuria Velasco-Pérez nuriavp@ubu.es
Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Avenida Cantabria s/n, 09001 Burgos, Spain

Juan Marcos Sanz juanmarcos@tsanta.es
Textil Santanderina, N-634, km 43, 39500 Cabezón de la Sal, Spain

Abraham Casas acasas@centrotecnologicoctc.com
Centro Tecnológico de Componentes-CTC, Scientific and Technological Park of Cantabria (PCTCAN), Calle Isabel Torres, 1, 39011 Santander, Spain

Ángel Arroyo aarroyop@ubu.es
Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Avenida Cantabria s/n, 09001 Burgos, Spain

Daniel Urda durda@ubu.es
Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Avenida Cantabria s/n, 09001 Burgos, Spain

Vol. \jvolume No. \jnumber \jyear
Journal of Applied Logics — IfCoLog Journal of Logics and their Applications

Abstract

Quality assurance stands as a pivotal phase across all manufacturing processes, particularly within the textile industry. Presently, textile inspections heavily rely on human visual assessment due to deficiencies in commercial solutions. In response, this research advocates for the adoption of Deep Learning models to automate fabric quality control within dynamic and contemporary production environments. Specifically, Convolutional Neural Networks are scrutinized using authentic images sourced from the Batavia and Sarga weave production lines. The experimental results identify DenseNet121 and InceptionV3 as the most effective models for Batavia and Sarga weaves, respectively. DenseNet121 demonstrates balanced performance across key metrics for Batavia weave, while InceptionV3 excels in Sarga weave, particularly in F1-score and AU-ROC. These findings underscore the potential of DL models to enhance the accuracy and efficiency of textile quality control.

Keywords: Textile defect detection, Fabric quality control, Industry 4.0, Deep Learning in manufacturing, Convolutional neural networks, Image analysis and classification.

1 Introduction and Previous Work

In recent years, various challenges in industrial sectors have been addressed through the application of soft computing methodologies [3]. Despite this progress, significant hurdles persist within the textile industry, which remains entrenched in traditional practices with limited technological integration. The sector's ability to withstand change can be attributed to its historical composition, primarily comprising long-standing entities adept at adapting to market shifts and competing effectively against low-cost alternatives. However, as the textile industry increasingly emphasizes competitiveness and elevates quality standards, the existing quality control mechanisms are proving inadequate and beset with limitations. Consequently, there arises an imperative to explore novel, automated, and intelligent inspection techniques.

Presently, quality inspection in the textile domain predominantly relies on visual scrutiny conducted by trained personnel, operating at an average pace of approximately 20 m/min, supplemented by sporadic quality audits. However, this inspection rate falls short of the manufacturing speed, which averages approximately 42 m/min, resulting in a bottleneck during the quality evaluation. While traditional visual inspection has served its purpose historically [17], its efficacy is waning, necessitating innovative solutions [19].

This study proposes and validates the efficacy of Deep Learning (DL) models in overseeing fabric quality at Textil Santanderina (TS) to address this challenge. The primary objective is to detect defects (including warp and weft irregularities, impurities, bowing, skewing, and color gradients) arising at various production stages (spinning, weaving, coating, and garment manufacturing). Specifically, Convolutional Neural Networks (CNNs) are advocated for their proven image processing capabilities in quality control applications [18]. Different CNN architectures will be evaluated to identify the most suitable model for the problem at hand. Despite the promise of artificial vision systems, particularly for inspection purposes, their industrial implementation remains constrained by persistent challenges, such as production interruptions following defect detection, delays, resource wastage, and over-reliance on human intervention. Although off-the-shelf image processing systems are available for textile applications, their lack of Artificial Intelligence (AI) integration hinders swift training and adaptation to the industry's high product variability (TS handles batches ranging from 500 m to 50,000 m, with over 700 new references annually).

The subsequent sections of this article are structured as follows: Firstly, a review of the relevant literature pertaining to the current investigation is presented in Section 2. Section 3 introduces the investigated issue and outlines the dataset generation process employed in this study. The methodological approaches utilized and the subsequent experimentation are delineated in Section 4. Section 5 provides a detailed exposition of the results of the experimentation. Finally, in Section 6, the best-performing models are identified, which also presents the main conclusions derived from the experimentation and the associated results.

2 Related Work

Several commercial systems currently in use include Uster® EVS FABRIQ VISION, Mahlo GmbH, and PEKAT VISION, each with its own set of limitations. For instance, the Uster® EVS FABRIQ VISION system is restricted to post-production inspection and is suitable only for uniform production batches. Similarly, the Mahlo GmbH system is limited to detecting defects that do not affect the fabric's structure. At the same time, PEKAT VISION is primarily designed to easily identify visible errors in long batches, lacking an AI learning system for pattern recognition and series analysis. In recent years, various AI techniques have been used for the monitoring of quality and manufacturing processes in industry in general [11] and in the textile industry in particular [20]. Focusing on AI techniques for image analysis [1], it can be seen that the techniques selected in this research work are of current relevance due to their validity in this field. DL [26] architectures have been extensively utilized for quality control across various industries[14][22]. Specifically, CNNs [15] have demonstrated significant potential in manufacturing and quality assurance [4][25][16]. While some researchers have employed traditional methods [29] for defect detection in the textile sector, others have explored DL-based approaches. Notably, early studies such as [12] employed supervised CNNs, albeit on outdated datasets that fail to meet current industry standards. Other DL models, like those described in [28], have focused on detecting multiple defects within a single fabric image, which does not represent real-world scenarios where simultaneous defects are rare. Alternatively, some researchers have proposed unsupervised learning methods to streamline hyperparameter initialization and labeling efforts. For example, [13] introduced techniques to reduce the time and effort required for model fine-tuning, while [30] proposed the use of Variational Autoencoders to label defect-free images automatically. However, in our study, manual labeling by human operators is currently employed, making these approaches less relevant to our context. Overall, previous studies have primarily been conducted offline and not validated under real-world conditions with high manufacturing speeds.

The work presented at 18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023) [27], demonstrated the feasibility of using DL for warp and weft defect detection. Although that study focused on the Batavia weave, this research delves into more fibers such as the Twill weave. Furthermore, a greater number of CNN pre-trained have been used in both tissues.

Therefore, our present research represents a significant advancement by validating CNNs under complex, high-speed production environments using up-to-date production conditions.

3 Case Study

In this study, CNNs are leveraged for analyzing images captured during the fabric production process at TS, a prominent European company specializing in yarns, fabrics, finishes, and garments made from various materials. Established in December 1960, TS boasts a rich history in the textile industry and currently offers a diverse range of products spanning the fashion/clothing, sportswear, healthcare, medical, military, public, and protective sectors.

Distinguished by its vertical supply chain, TS manages all aspects of fabric manufacturing, including spinning, weaving, coating finishing, garment production, and logistics. This comprehensive integration allows the company to deliver processed products at any stage of production with various finishes. With an annual output exceeding 40 million meters of fabrics and over 7 million garment units worldwide, TS is a key player in the textile market.

To support our research objectives, TS has equipped its facility with a Basler raL camera featuring the Awaiba DR-2k-7 CMOS sensor, capable of delivering 51 kHz at a resolution of 2k. Operating within the VIS-NIR bandwidth, the camera is equipped with Basler proprietary optics, enabling a resolution of 10 px/mm. For uniform illumination, a LED array emitting at 850nm, spanning a field of 15mm x 1510mm and comprising 360 infrared (IR) 850nm LEDs, is positioned 20 cm above the fabric at a 15º angle of incidence. This infrared (IR) illumination ensures compatibility with existing D65 standard illuminating systems for visual inspection. A depiction of the factory setup is provided in Figure 1.

The camera is synchronized with the fabric using a 10-bit inductive encoder to ensure high precision (<1/10 mm) positioning. CNNs are trained on the captured images following standard training, testing, and validation procedures outlined in Section 4. Lighting conditions, setup geometry, optics, and CMOS sensor specifications have been carefully chosen to meet additional requirements for defect inspection, including fabric width, batch length, and inspection speed, among others.

The experimentation has been conducted using two different weaves (Batavia and Sarga) in the research context. These textiles were selected due to their distinctive characteristics and relevance in various textile applications. The inclusion of both textiles in the study provides a comprehensive assessment of the efficacy of the proposed methods across a broader range of textile inspection scenarios. Additionally, it allows for comparing and contrasting the performance of machine learning algorithms under variable conditions, enriching the understanding of their ability to address specific challenges in the textile industry. To conduct this study, images provided by TS were utilized and subsequently compiled into a public dataset titled "Original and processed dataset of Batavia and Sarga woven fabric images" [6]. This dataset is further discussed in an article titled "Detailed description of the classified image dataset of a textile company in Spain"[5].

3.1 Batavia

For the Batavia fabric, it's important to note that Batavia is a specific type of twill weave, categorized under the broader classification of Sarga fabrics. Images of a Batavia fabric were gathered in November 2022 over several days. This fabric is a blend of cotton and linen (85/15) with a standard twill weave known as Batavia (refer to Figure 1). A neutral color (greige fabric) was selected for its widespread use in woven fabrics for the summer season and to minimize variability in color combinations and patterns. A total of 2,755 high-resolution grayscale images, each 16 bits, with dimensions of 2,048x696 pixels were collected (as illustrated in Figure

2).

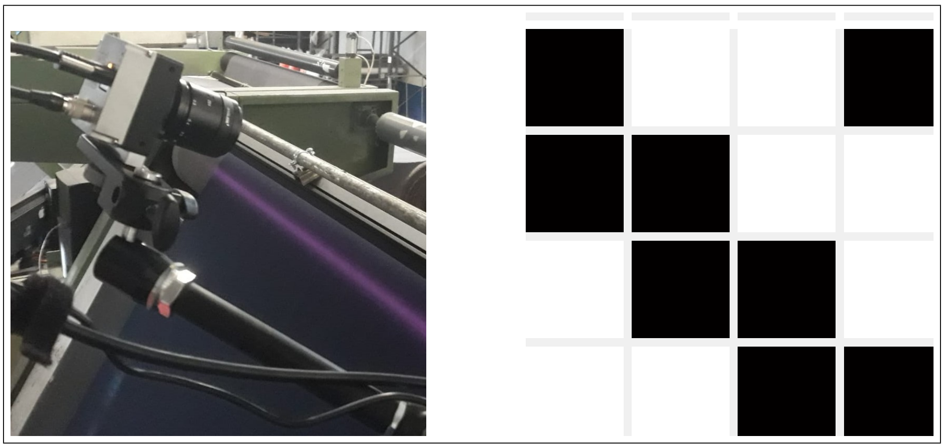


Figure 1: Image acquisition set up (left image) and Batavia fabric schematics with 4 yarns warp and 4 yarns weft (right image).

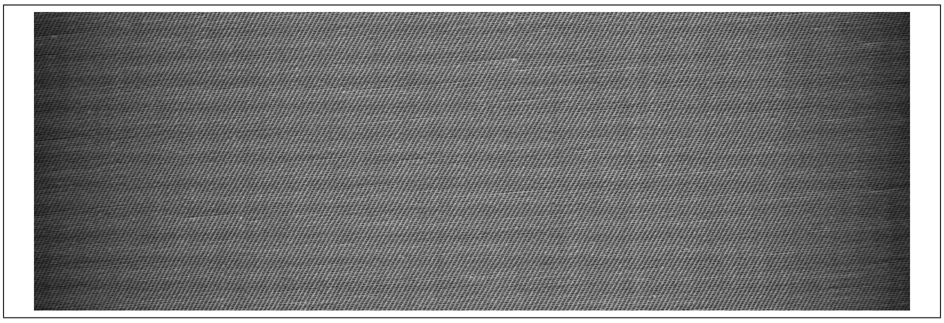


Figure 2: Example of 2,048X696 original Batavia image.

3.2 Sarga

The Sarga fabric, which is characterized by its distinctive diagonal weave pattern, offers durability and versatility, making it suitable for various textile applications. A neutral color (greige fabric) was selected to minimize variability in color combinations and patterns. A total of 1,548 high-resolution grayscale images, each 16 bits, with dimensions of 2,048x696 pixels were collected. In Figure 3 an image of Sarga fabric can be observed.

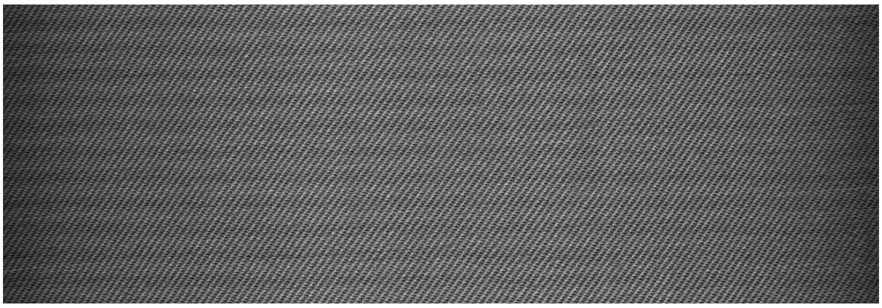


Figure 3: Example of 2,048X696 original Sarga image.

4 Methods and Experimentation

This section delineates the DL methodologies and strategies employed in this study to craft a precise model for fabric defect detection. Additionally, comprehensive insights into the experimental design are furnished within this section, with specific methodologies elaborated in subsection 4.1 and the experimentation outlined in subsection 4.2.

4.1 Methods

CNNs [7] have become increasingly recognized as a robust tool for tasks such as object detection and image analysis. As larger datasets have become more accessible, CNNs have demonstrated notable enhancements in performance, facilitating the creation of autonomous systems applicable across diverse domains. This study aims to harness the capabilities of CNNs to support quality control teams in promptly identifying fabric defects. Such assistance aims to enhance the quality of end products, lower operational expenses, and augment customer satisfaction.

To achieve this objective, the examination conducted in this study centers on the dataset delineated in the preceding section, wherein the gold standard entails the classification of fabric images meticulously evaluated by a quality control specialist. In this context, the authors suggest assessing the benefits of constructing two CNN models with analogous architectures applied to images within the specified dataset, along with employing transfer learning architectures to evaluate their efficacy compared to a standard labeling baseline.

Firstly, we are going to review the basis of CNNs and transfer learning. CNNs are a type of deep neural network that is designed to process data that has a grid-like structure (such as images) and use convolutional layers to extract relevant features from the image by applying a set of filters to the input. These filters can learn to detect specific patterns and structures in the images, such as edges, corners, and textures. CNNs typically include pooling layers, which down-sample the feature maps generated by the convolutional layers, as well as fully connected layers, which classify the features into specific categories. On the other hand, transfer learning is a DL technique in which a model of CNNs trained on one task is reused as a starting point for training a model on a different task. In other words, transfer learning involves using pre-trained neural network models as a starting point for training a new model on a smaller dataset or a different task. We can differentiate two types of transfer learning: in feature extraction, the pre-trained model is used as a fixed feature extractor, where only the last few layers of the model are modified to fit the new task and in fine-tuning, the pre-trained model is used as a starting point, but all of the layers are updated and retrained on the new task. In this study, we have focused on this last technique.

Next, the baseline and CNNs methods employed in this work are described.

4.1.1 Baseline

This entails the quality control team's real-time classification of products during the final stages of the production line. This performance benchmark aligns with the level of human accuracy we aim to achieve, or ideally surpass, through the utilization of DL-based methods for this intricate classification task.

The baseline methodology employed in this study does not necessitate the implementation of any specific model, as the dataset already contains labels assigned by the quality control team. Additionally, analogous labels are provided for the gold standard, derived offline by an expert who meticulously examines all images. Consequently, a performance metric can be readily computed to evaluate human proficiency in fabric defect detection.

4.1.2 Pre-trained CNNs

Popular pre-trained models have been adjusted for the given problem and fine-tuned on the datasets, under the "transfer learning" paradigm. The applied models are listed in chronological order of their introduction, reflecting the historical evolution of CNNs architectures:

- **VGG16**: It is a pre-trained CNN model that was developed by the Visual Geometry Group at the University of Oxford [21]. The model consists of 16 layers, including 13 convolutional layers and 3 fully connected layers, and was trained on the ImageNet dataset, which contains millions of labelled images. VGG16 is known for its simple and uniform architecture, which uses small 3x3 filters throughout the convolutional layers and max pooling layers to downsample the feature maps. The model has been shown to perform well on a wide range of image classification tasks, including object recognition and localization. One of the main advantages of VGG16 is its flexibility and ease of use.

- **VGG19**: It is a CNN model that also was developed by the Visual Geometry Group at the University of Oxford [21]. It is an extension of the VGG16 model and has 19 layers, including 16 convolutional layers and 3 fully connected layers. Like VGG16, VGG19 was trained on the ImageNet dataset. VGG19 is known for its deep architecture and uses a similar approach as VGG16, with small 3x3 filters and max pooling layers to downsample the feature maps. However, VGG19 has more layers and uses larger filters in the later layers, which allows it to capture more complex features in the input images.

- **InceptionV3**: It is a CNN model that was developed by Google researchers as part of the Inception architecture family [23]. The model consists of 48 convolutional layers, including traditional convolutional layers and Inception modules, which use multiple parallel convolutions of varied sizes to extract features from the input images. Some unique features of the network are that it uses batch normalization after every convolutional and fully connected layer to accelerate the training, and it uses factorization into smaller convolutions (such as 1x1 and 3x3 convolutions) in order to reduce the computational complexity of the model. InceptionV3 was trained on the ImageNet dataset with a top-5 error rate of less than 4%. The model is known for its efficiency and accuracy.

- **ResNet101**: It is a CNN model that was developed by researchers at Microsoft as part of the ResNet (Residual Network) family of architectures [8]. The model consists of 101 layers, with a total of over 44 million parameters. It makes use of residual connections to allow for the efficient training of very deep neural networks. These connections allow information from earlier layers of the network to be passed directly to later layers, bypassing the need to go through multiple non-linear transformations. In addition, ResNet101 uses a bottleneck architecture, which reduces the number of parameters in the model while increasing its depth. This is achieved by using 1x1 convolutions to reduce the number of channels in the input tensor before applying 3x3 convolutions. And, finally, ResNet101 uses pre-activation, which means that batch normalization and activation functions are applied before convolutions.

- **ResNet50V2**: It is a CNN architecture belonging to the ResNet family, designed to improve the efficient training of deep networks [9]. This model comprises 50 layers and leverages residual connections to mitigate the vanishing gradient problem, facilitating the training of very deep networks. ResNet50V2 adopts a pre-activation scheme, where batch normalization and activation functions are applied before convolutions in each residual block, enhancing gradient flow and training stability. Additionally, it employs a bottleneck architecture, combining 1x1, 3x3, and 1x1 convolutions to reduce the dimensionality of intermediate data while maintaining computational efficiency. This optimized design makes it particularly well-suited for complex computer vision tasks requiring both depth and efficiency.

- **Xception**: It is a CNN model that was developed by Google researchers in 2016 [2]. The model is based on the Inception architecture, which uses multiple parallel convolutions of different sizes to extract features from the input images. However, unlike Inception, Xception uses depthwise separable convolutions, which split the convolution operation into two separate operations: a depthwise convolution and a pointwise convolution. The use of depthwise separable convolutions in Xception allows the model to have a much smaller number of parameters than traditional CNNs. This makes Xception a computationally efficient model that is suitable for deployment on mobile devices and other resource-constrained environments.

- **DenseNet121**: It is a CNN model that was developed by researchers at Facebook AI Research [10]. The model consists of 121 layers and has a total of approximately 8 million parameters. One of the unique features of DenseNet121 is its use of dense connections, which allow information to be passed between all layers of the network. This helps to improve the flow of information through the network and makes it more efficient at learning complex representations of visual features. DenseNet121 also makes use of batch normalization and ReLu activation functions, which have been shown to improve the accuracy of the model. Additionally, the model makes use of a global average pooling layer, which reduces the number of parameters in the network and helps to prevent overfitting.

- **EfficientNetB0**: It is the baseline model in the EfficientNet family, comprising 5.3 million parameters [24]. It employs a compound scaling method that simultaneously adjusts the network's depth, width, and input resolution, optimizing both efficiency and performance. With an input resolution of 224x224, EfficientNetB0 achieves an optimal balance between accuracy and computational cost, making it particularly suitable for real-time computer vision tasks, such as image classification and object detection, in resource-constrained environments. This model outperforms traditional CNNs, such as ResNet50, with fewer parameters while maintaining competitive accuracy.

- **EfficientNetB3**: It extends the architecture of EfficientNetB0 by increasing the network's depth, width, and input resolution to 300x300, resulting in a model with 12 million parameters [24]. This enhancement allows EfficientNetB3 to capture more complex patterns and improve accuracy compared to the baseline model. The compound scaling approach ensures that the increase in model size does not lead to a disproportionate increase in computational cost. EfficientNetB3 is well-suited for applications that require higher accuracy, such as advanced image classification and object detection tasks, while maintaining computational efficiency relative to its performance.

Throughout the training procedure, specific hyperparameters were set as follows: a learning rate set to 0.002, 2 layers dense of neurons, with 8,192 neurons per layer, dropout per layer to 0, batches consisting of 128 images were used to train and update the parameters of DL models for up to 100 epochs, regulated by an early stopping system. Mechanism designed to stop training if no improvement in loss is observed on a validation set after 10 epochs.

4.2 Experimentation

The methodology employed in this investigation to develop and assess the efficacy of a CNN model for the specified quality control task encompasses several pivotal stages, which are delineated in further detail below.

4.2.1 Image preprocessing

Initial image preprocessing was conducted wherein the original 16-bit images underwent rescaling to their corresponding 8-bit versions. This transformation reduced the potential pixel values per image from 65,536 to 256. Subsequently, to facilitate meticulous fault detection and expand the dataset, the 8-bit images were partitioned into twelve segments, each measuring 365x365 pixels, with minor overlapping sections along both horizontal and vertical axes (refer to Figure 4).

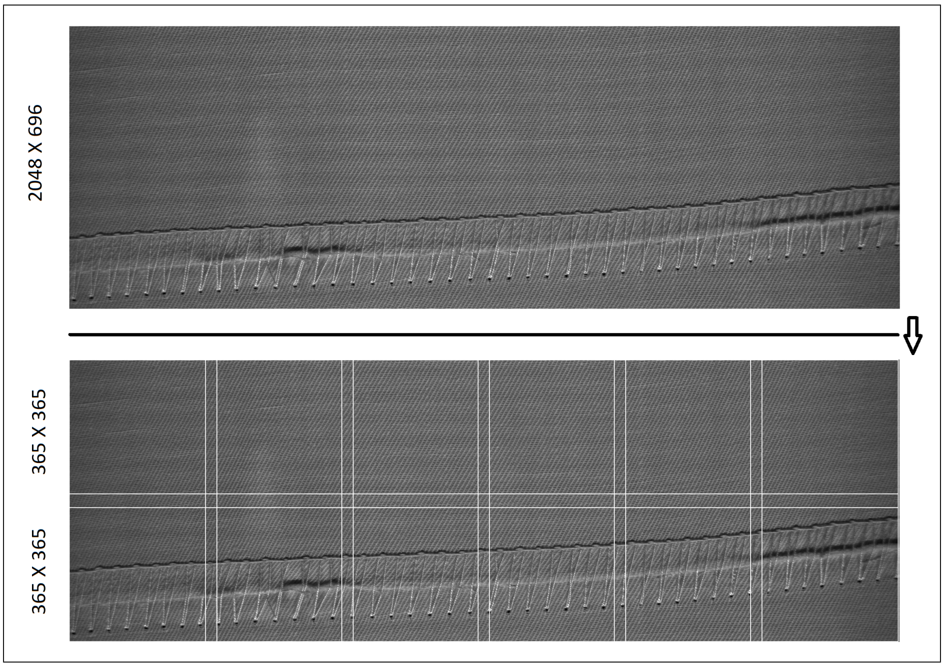


Figure 4: Small size patches extraction of 365x365 pixel from original 2048x696 images, with small overlapping areas in the horizontal and vertical axis.

At Batavia weave, the finalized dataset used to train the models comprised 28,693 images. After a thorough review by TS, 8,782 were confirmed as cases, while 19,911 were designated as controls and 4,367 were considered doubtful and eliminated from the data set. The number of patches is not equal to the number of originals multiplied by 12 due to the review carried out by Textil Santanderina. In the case of Sarga, a total of 18,576 Sarga fabric pieces, each measuring 365x365 units, were obtained. Among these, 18,403 were classified as controls and 173 as cases. Notably, this review process identified instances where a specific slice within an original image may constitute a control rather than a case, as the defect was absent in that particular segment. In Sarga dataset, the total number of patches equals the number of originals multiplied by 12. Lastly, pixel normalization was executed across each image, transitioning the pixel values from the [0-255] range to the [0-1] range through simple division by 255. This normalization approach is known to mitigate numerical instability issues during the model training phase.

Fabric	N	Controls	Cases
Batavia	28,693	19,911	8,782
Sarga	18,576	18,403	173

Table 1: The processed dataset includes information on fabrics (Batavia and Sarga), controls and cases.

The following Table 1 presents the processed dataset according to the type of fabric (Batavia and Sarga). It includes information on the total number of 365X365 images (N), the controls, and the cases.

4.2.2 Data augmentation

Addressing the class imbalance inherent in the dataset, on-the-fly data augmentation was implemented to augment the number of case samples available for CNN training. This widely employed technique in DL, particularly in computer vision, involves generating modified or transformed variants of existing samples to bolster the training set. By applying vertical and horizontal flips, along with alterations in color and brightness, to the training set, the dataset size was augmented by a factor of five. Consequently, this augmentation strategy augmented the training set with approximately a thousand cases, thereby enriching the diversity and breadth of examples and enhancing the accuracy and generalization capabilities of the models.

4.2.3 Evaluation strategy

The methodology adopted in this study involves training and evaluating the performance of CNN architectures using a 5-fold stratified cross-validation scheme. In this approach, the dataset is divided into five subsets of equal size, ensuring that the class distribution is maintained across each of them. In each iteration, one fold is reserved as the test set, while the remaining four folds are used for training and validation.

From these four folds, 75% of the data (i.e., 60% of the total dataset) is allocated to training, and the remaining 25% (20% of the total) is used for validation. This 60/20/20 (training/validation/test) split is maintained consistently throughout the five iterations. The validation set is used to monitor the model's performance during training (e.g., for early stopping), but not for the final evaluation. The test set is exclusively used for the final evaluation of performance once training is complete. This process is repeated across the five folds to ensure the robustness of the analysis and mitigate bias in sample selection.

To mitigate the impact of class imbalance during model training, we assigned specific weights to each class based on their relative frequency in the dataset. This was implemented through the class_weight parameter in the model.fit function of Keras, which adjusts the contribution of each class to the binary cross-entropy loss. This approach complements the use of evaluation metrics that are robust to class imbalance, such as the Area Under the ROC Curve (AU-ROC), the Geometric Mean (G-mean), and the Area Under the Precision-Recall Curve (AU-PR).

4.2.4 Metrics

In this study, various complementary performance metrics are analyzed. This approach is adopted due to the highly imbalanced class distribution inherent in the addressed problem. Quality control departments may express particular interest in enhancing the detection capabilities of models for either the negative class (controls or defect-free images), the positive class (cases or images with defects), or both. The assessment of model performance relies on metrics derived from True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

- Accuracy: This metric considers the TP, TN, FP, and FN derived from the predictions generated by a specific method. Its interpretation warrants careful consideration, particularly in the context of imbalanced classification challenges, where it may erroneously convey high performance by favoring the majority class for all input images. The computation of accuracy is delineated by the following 1:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{1}$$

- Precision: This metric quantifies the ratio of true positive predictions to all positive predictions generated by a specific method. Essentially, it provides insight into the etrexactness or quality of the method's positive predictions, particularly valuable in scenarios where the ramifications of a false positive are substantial (e.g., medical diagnosis, where a false positive outcome can trigger unnecessary treatment or patient anxiety). The calculation of precision is detailed in the following equation 2:

$$Precision = TP/(TP + FP) \tag{2}$$

- Recall: This metric evaluates the ratio of true positive predictions to all actual positive instances within the dataset. Essentially, it offers insight into the comprehensiveness or quantity of the model's positive predictions, particularly valuable in contexts where the consequences of a false negative are significant (e.g., in fraud detection, where a false negative can lead to financial losses). Its calculation is delineated in the equation 3:

$$Recall = TP/(TP + FN) \tag{3}$$

- F1-score: It represents the harmonic mean between the Precision and Recall metrics. A higher value closer to 1 indicates superior performance of the model. This metric is calculated according to the equation 4:

$$F1 = 2 * (Recall * Precision)/(Recall + Precision) \tag{4}$$

- Geometric Mean: G-mean is a metric used to evaluate the performance of a classification model, particularly in cases of imbalanced datasets. It is designed to balance the accuracy of both the positive and negative classes, ensuring that the model performs well on both sides of the classification task. The calculation is showed in the following equation 5:

$$\text{G-mean} = \sqrt{\frac{TP \times TN}{(TP + FN) \times (TN + FP)}} \tag{5}$$

- Area Under the Precision-Recall Curve (AU-PR): It is a performance metric primarily used in imbalanced classification problems. Unlike accuracy, which can be misleading in the case of imbalanced datasets, AU-PR focuses on evaluating the model's ability to correctly identify positive instances. The Precision-Recall curve plots precision against recall at various threshold levels, and the AU-PR represents the area under this curve. A higher AU-PR indicates a better trade-off between precision and recall, with values closer to 1 indicating superior model performance. This metric is especially useful in cases where the negative class is significantly more frequent than the positive class, such as rare disease detection or fraud detection.

- Area Under the Receiver Operating Characteristic Curve (AU-ROC): It is a widely used metric to evaluate the overall performance of a classification model, especially in binary classification tasks. The ROC curve plots the True Positive Rate (recall) against the False Positive Rate (1-specificity) at various classification thresholds. The area under this curve, known as AUROC, quantifies the model's ability to discriminate between the positive and negative classes. An AUROC value of 1 represents perfect classification, while a value of 0.5 suggests a model with no discriminative power, similar to random guessing.

AUROC is particularly useful for comparing models, as it is independent of the classification threshold, making it a robust metric across different decision thresholds.

These metrics were selected for their robustness in imbalanced classification scenarios, which complements the class reweighting strategy adopted during training (as described in Section 4.2.3).

4.2.5 Computational and implementation environment

All experiments were conducted on a GPU cluster funded by the InfraRed program of the Junta de Castilla y León. This infrastructure is designed for artificial intelligence and deep learning projects, supporting large-scale experimentation. It included 576 CPU cores (24 processors with 24 cores each) based on the AMD Epyc platform, 9.5 TB of DDR4 RAM, and 320 TB of storage. The cluster was equipped with 30 NVIDIA Tesla A100 GPUs, each with 40 GB of dedicated memory, providing a total computing power of 9360 TFLOPS in FP16 and 4680 TFLOPS in TF32 precision.

A software environment configured on the cluster was used, based on Python 3.9.18, which includes the main libraries required for model training and evaluation. These included TensorFlow, PyYAML, Pandas, NumPy, Scikit-learn, Matplotlib, and Pillow. These tools enable deep learning tasks to be performed using the GPU acceleration provided by the cluster infrastructure.

5 Results

The results obtained by the DL models that have been applied are described in this section, organized according to the weave types, Batavia and Sarga. The pretrained models have been fine-tuned for the specific task of defect detection in these weave types. The metrics presented to evaluate the models are accuracy, precision, recall, F1 score, G-mean, AU-PR and AU-ROC. Additionally, confusion matrices are provided to offer a detailed view of the classification outcomes and to better analyze the behavior of the models with respect to true and false positives and negatives for each class.

Radar plots, also known as spider or star plots, are graphical representations used to visualize multivariate data in a compact and informative manner. In the context of automatic textile inspection, radar plots offer a valuable tool for assessing the performance of various models across multiple metrics. Each axis of a radar plot represents a different metric, such as Accuracy, Precision, Recall, F1-score, AU-PR and AU-ROC allowing for a comprehensive evaluation of model performance. By

Red	Threshold	TN	FN	TP	FP
VGG16	0.375	8,346	2,129	6,653	11,565
VGG19	0.625	15,290	3,003	5,779	4,621
InceptionV3	0.500	11,430	2,539	6,243	8,481
ResNet101	0.450	15,078	3,649	5,133	4,833
ResNet50V2	0.375	13,927	2,479	6,303	5,984
Xception	0.275	6,798	1,456	7,326	13,113
DenseNet121	0.325	16,399	2,913	5,869	3,512
EfficientNetB0	0.450	15,601	3,848	4,934	4,310
EfficientNetB3	0.325	19,110	4,935	3,847	801

Table 2: Confusion matrices for the different models applied to the Batavia fabric.

plotting the values of these metrics on the corresponding axes and connecting them with lines or shapes, radar plots provide a visual summary of how well a model performs across different aspects of the inspection task. In our analysis, radar plots were generated to compare the different performance of pre-trained CNN models against baseline models. These plots allowed us to observe trends in performance metrics and identify areas where certain models excelled or underperformed. Furthermore, radar plots facilitate the comparison of multiple models by presenting their performance side by side. This enables researchers and practitioners to quickly assess the strengths and weaknesses of each model and make informed decisions regarding model selection. Overall, radar plots serve as powerful tools for evaluating the effectiveness of DL models in automatic textile inspection, providing insights that can inform future research and development efforts in this field.

The detailed outcomes for each weave type are provided in the subsequent sections, allowing for a thorough evaluation of the methodologies applied.Also, radar plots have been generated separately for the two types of weave under analysis, Batavia and Sarga.

5.1 Batavia weave

Prior to presenting the evaluation metrics, the confusion matrices corresponding to the Batavia fabric are included to provide a clearer understanding of the models' classification performance. These matrices illustrate the distribution of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), offering insight into model behavior under class imbalance conditions.

The confusion matrices for the different models applied to the Batavia fabric are shown in Table 2.

Following the implementation of DL models, including the baseline model, on Batavia fiber, the resulting metrics (accuracy, precision, recall, F1-score, G-mean,

Red	Accuracy	Precision	Recall	F1-score	G-mean	AU-PR	AU-ROC
Baseline	0.70	**1.00**	0.03	0.05	0.16	0.66	0.51
VGG16	0.52	0.37	0.76	0,49	0.56	0.54	0.68
VGG19	0.73	0.56	0.66	0.60	0.71	0.72	0.81
InceptionV3	0.62	0.42	0.71	0.53	0.64	0.61	0.73
ResNet101	0.70	0.52	0.58	0.55	0.67	0.64	0.75
ResNet50V2	0.71	0.51	0.72	0.60	0.71	0.68	0.79
Xception	0.49	0.36	**0.83**	0.50	0.53	0.65	0.76
DenseNet121	0.78	0.63	0.67	**0.65**	**0.74**	**0.80**	**0.87**
EfficientNetB0	0.72	0.53	0.56	0.55	0.66	0.70	0.79
EfficientNetB3	**0.80**	0.83	0.44	0.57	0.65	0.76	0.84

Table 3: Performance metrics of various pre-trained models fine-tuned on the Batavia weave dataset. The table compares the models across multiple metrics, including Accuracy, Precision, Recall, F1-score, G-mean, AU-PR (Area Under the Precision-Recall curve), and AU-ROC (Area Under the Receiver Operating Characteristic curve).

AU-PR and AU-ROC) detailed in Table 3.

Based on the performance metrics in the Table 3, DenseNet121 stands out as the most effective neural network model for the given task. DenseNet121 model exhibits the best overall performance, excelling across several key metrics. It achieves the highest AU-ROC (0.87), indicating the best ability to discriminate between positive and negative classes. Additionally, it demonstrates a relatively high F1-score of 0.65, indicating a good balance between precision and recall. The model also achieves a precision of 0.63, which is strong and well-balanced. Its recall is 0.67, showing that a significant proportion of positive instances are correctly identified. The AU-PR is 0.80, reflecting strong performance in dealing with imbalanced classes.

In contrast, the Baseline (human) model shows considerably lower performance across most metrics. The precision of the Baseline is perfect at 1.00, but this comes at the cost of a very low recall of 0.03, meaning it detects almost no positive instances. As a result, the F1-score is extremely low at 0.05, as the high precision is insufficient to compensate for the poor recall. The AU-ROC is 0.51, which indicates that the human model performs similarly to random guessing in distinguishing between classes. The AU-PR is 0.66, which is better than the AU-ROC but still reflects suboptimal performance compared to the neural networks.

Other models that perform well include EfficientNetB3, which achieves an excellent AU-ROC of 0.84, although its recall (0.44) and F1-score (0.58) are lower than those of DenseNet121. Additionally, VGG16 and VGG19 both perform well in terms of recall (0.76 for VGG16 and 0.66 for VGG19), but their F1-scores are lower than that of DenseNet121.

This section also presents a detailed analysis of the radar plots generated to

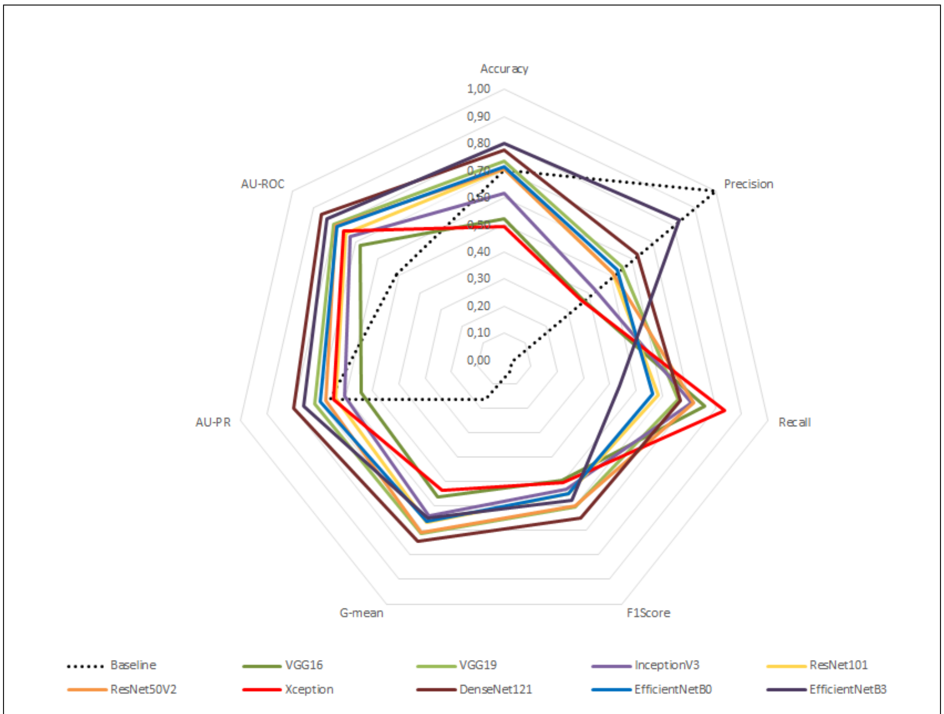


Figure 5: Radar plot illustrating the performance of various pre-trained models fine-tuned on the Batavia weave dataset. The plot compares the models across key metrics, including Accuracy, Precision, Recall, F1-score, G-mean, AU-PR, and AU-ROC.

evaluate the performance of the models applied to Batavia weave. The radar plot represents different performance metrics, offering a clear and concise visualization of how the models perform in key aspects of textile inspection. The trends observed in the data are examined in detail, including differences between pre-trained models and models trained from scratch, as well as the implications of these differences in terms of accuracy and defect detection capability.

The figure 5 shows a radar plot for each one of the studied metrics and the applied models on Batavia weave. Metrics associated to the baseline have also been included (represented with a dashed line).

Red	**Threshold**	**TN**	**FN**	**TP**	**FP**
VGG16	0.725	18,362	116	57	41
VGG19	0.950	18,187	119	54	216
InceptionV3	0.925	18,360	111	62	43
ResNet101	0.850	18,361	107	66	42
ResNet50V2	0.950	17,886	104	69	517
Xception	0.950	18,192	86	87	211
DenseNet121	0.600	14,510	53	120	3,893
EfficientNetB0	0.775	18,376	152	21	27
EfficientNetB3	0.550	18,244	148	25	159

Table 4: Confusion matrices for the different models applied to the Sarga fabric.

5.2 Sarga weave

The confusion matrices for the Sarga fabric are presented first, providing a clear overview of the model's performance in distinguishing between defect and non-defect cases.

Table 4 shows the confusion matrices for the different models applied to the Sarga fabric.

After applying DL models, including the baseline model, to Sarga fiber, the resulting metrics like accuracy, precision, recall, F1-score, G-mnean, AU-PR and AU-ROC are presented in Table 5.

Red	**Accuracy**	**Precision**	**Recall**	**F1-score**	**G-mean**	**AU-PR**	**AU-ROC**
Baseline	**0.99**	0.00	0.00	0.00	0.16	0.50	0.50
VGG16	**0.99**	0.58	0.33	0.42	0.57	0.42	0.86
VGG19	0.98	0.20	0.31	0.24	0.56	0.42	0.91
InceptionV3	**0.99**	0.59	0.36	0.45	0.60	0.52	**0.93**
ResNet101	**0.99**	**0.61**	0.38	**0.47**	0.62	**0.54**	0.90
ResNet50V2	0.97	0.12	0.40	0.18	0.62	0.39	**0.93**
Xception	0.98	0.29	0.50	0.37	0.71	0.40	0.88
DenseNet121	0.79	0.03	**0.69**	0.06	**0.74**	0.50	0.92
EfficientNetB0	**0.99**	0.44	0.12	0.19	0.35	0.29	0.78
EfficientNetB3	0.98	0.14	0.14	0.14	0.38	0.30	0.84

Table 5: Performance metrics of various pre-trained models fine-tuned on the Sarga weave dataset. The table compares the models across multiple metrics, including Accuracy, Precision, Recall, F1-score, G-mean, AU-PR (Area Under the Precision-Recall curve), and AU-ROC (Area Under the Receiver Operating Characteristic curve).

The analysis of the performance metrics reveals that the InceptionV3 model is the most suitable for the task of defect classification in Sarga weave threads. While ResNet101 achieves a slightly higher F1-score (0.47), InceptionV3 combines a competitive F1-score of 0.45 with the highest AU-ROC (0.93), reflecting superior

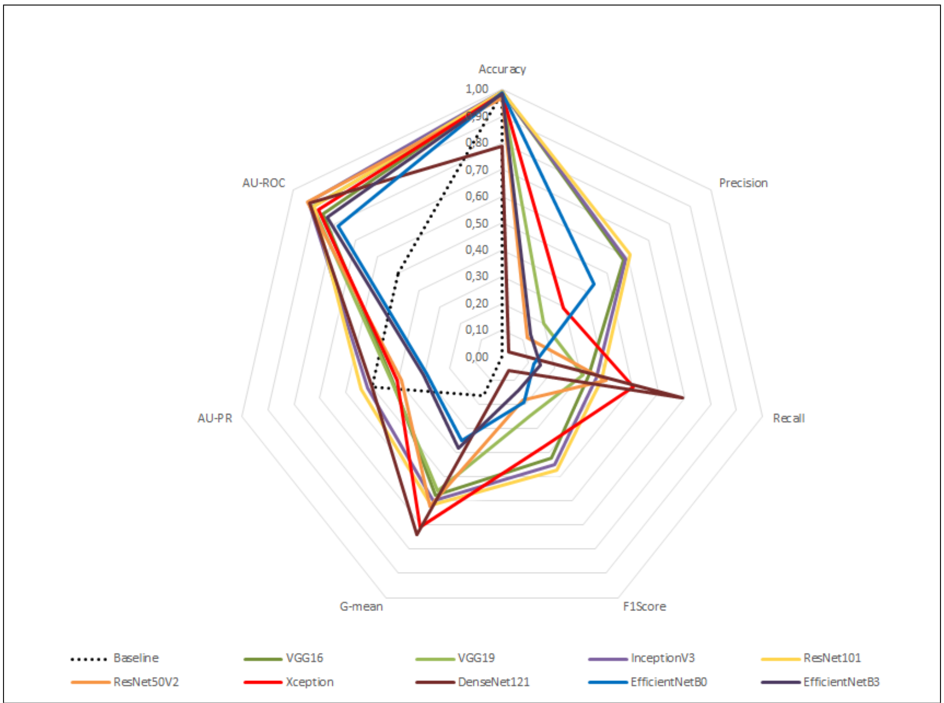


Figure 6: Radar plot illustrating the performance of various pre-trained models fine-tuned on the Sarga weave dataset. The plot compares the models across key metrics, including Accuracy, Precision, Recall, F1-score, G-mean, AU-PR, and AU-ROC.

overall discriminative ability across all thresholds. This model maintains a solid balance between precision (0.59) and recall (0.36), allowing it to identify a significant proportion of positive cases while effectively controlling false positives. In contrast, the human baseline, while showing a high Accuracy of 0.99 due to the dominance of the negative class, fails to detect any positive cases, with Precision and Recall values of 0. This highlights that, despite its seemingly high accuracy, the baseline provides no meaningful value in identifying defects. Compared to other models, such as EfficientNetB3 or DenseNet121, InceptionV3 demonstrates a more consistent performance, overcoming the challenges posed by class imbalance that adversely affect precision and recall in these alternatives.

Below, a comprehensive analysis is conducted of the radar plots generated to evaluate the performance of the models applied to Sarga weave.

The figure 6 shows a radar plot for each one of the studied metrics and the applied models for Sarga weave.

6 Conclusions

This study has addressed the critical challenge of quality inspection in the textile domain through the application of DL models to oversee fabric quality at TS. The investigation demonstrated that pretrained networks can significantly enhance the inspection process, effectively bridging the gap between the traditional inspection rate of 20 m/min and the production speed of 42 m/min. Our findings revealed that selecting appropriate CNN architectures was pivotal in optimizing defect detection and reducing reliance on human intervention. Consequently, the integration of these AI-based systems has shown great potential to mitigate production interruptions, reduce delays, minimize resource wastage, and manage the high variability of textile products handled by TS. The results of this study underscore the transformative potential of DL in revolutionizing textile quality control, offering a robust and scalable solution to meet the industry's evolving demands.

In conclusion, our study highlights key aspects regarding the effectiveness of DL models, particularly pretrained CNN architectures, for automated textile inspection. First, we emphasize the critical role of selecting appropriate evaluation criteria, as evidenced by the dependency of model performance on the chosen criterion. The key metrics for evaluating the performance of the networks on both fabrics are the F1-score, G-mean, and AU-ROC, as they provide a comprehensive assessment of the models' ability to balance precision and recall (F1-score), maintain a high true positive rate while minimizing false positives and negatives (G-mean), and effectively distinguish between positive and negative classes (AU-ROC). DenseNet121 stands out as the best network for Batavia weave, while InceptionV3 is the most suitable for Sarga weave. These networks exhibit balanced and superior performance across the key metrics, making them optimal for their respective fabrics

Furthermore, our findings highlight the effectiveness of pretrained CNN models in detecting defects across various weave types. By employing radar plots to visualize performance metrics, we provide a comprehensive view of the relative performance of different models. This study significantly contributes to advancing DL techniques in textile quality control, laying the foundation for future research in this field.

Additionally, our study demonstrates the suitability of the proposed DL architecture, specifically CNNs, for online fabric quality control, surpassing the capabilities of human operators. Notably, our analysis reveals superior results compared to the baseline, particularly in Precision and F1-score metrics, highlighting the limitations of traditional metrics such as Accuracy in imbalanced datasets. The integration of DL models into production machinery promises to revolutionize quality control and traceability processes, minimizing human intervention and improving defect detection coverage in textile manufacturing processes. Overall, our study predicts a positive technological impact on textile manufacturing systems, facilitating the transition towards automated quality control and enhanced traceability.

For future work, exploring novel DL architectures specifically designed for textile defect detection and incorporating advanced techniques such as attention mechanisms and reinforcement learning could further improve the performance of automated inspection systems. Additionally, investigating the transferability of models trained on one fabric type to another, as well as the robustness of the models to variations in lighting conditions and fabric textures, would be valuable for real-world applications. Furthermore, integrating real-time feedback mechanisms into inspection systems to enable continuous model refinement and adaptation to evolving manufacturing environments represents an exciting avenue for future research. Currently, the company's detection process is carried out offline on a system with adequate computing power. Nonetheless, we recognize the potential benefits of implementing real-time detection in the production line, and therefore, we plan to explore this possibility in the future. Achieving real-time detection will require overcoming challenges related to preprocessing and inference, ensuring that the system operates efficiently and reliably in a production environment. Another promising direction involves the use of semantic segmentation techniques, which would allow not only for detecting whether a defect exists but also for precisely localizing the defective area within the fabric image. This added spatial information could significantly enhance the interpretability and usability of automated inspection systems in industrial settings.

References

[1] Sampath Boopathi and Uday Kumar Kanike. Applications of artificial intelligent and machine learning techniques in image processing. In *Handbook of Research on Thrust Technologies' Effect on Image Processing*, pages 151–173. IGI Global, 2023.

[2] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017.

[3] Alberto Diez-Olivan, Javier Del Ser, Diego Galar, and Basilio Sierra. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0. *Information Fusion*, 50:92–111, 2019.

[4] Yiping Gao, Liang Gao, and Xinyu Li. A hierarchical training-convolutional neural network with feature alignment for steel surface defect recognition. *Robotics and Computer-Integrated Manufacturing*, 81:102507, 2023.

[5] Beatriz Gil-Arroyo, Juan Marcos Sanz, Ángel Arroyo, Daniel Urda, Nuño Basurto, and Álvaro Herrero. Dataset for defect detection in textile manufacturing. *Data in Brief*, 59:111451, 2025.

[6] Beatriz Gil-Arroyo, Nuria Velasco-Pérez, Nuño Basurto, Juan Marcos Sanz, Daniel Urda, and Alvaro Herrero. Original and processed dataset of batavia and sarga woven fabric images, 2025. Dataset.

[7] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, et al. Recent advances in convolutional neural networks. *Pattern recognition*, 77:354–377, 2018.

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*, pages 630–645. Springer, 2016.

[10] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[11] Zohaib Jan, Farhad Ahamed, Wolfgang Mayer, Niki Patel, Georg Grossmann, Markus Stumptner, and Ana Kuusk. Artificial intelligence for industry 4.0: Systematic review of applications, challenges, and opportunities. *Expert Systems with Applications*, 216:119456, 2023.

[12] Pandia Rajan Jeyaraj and Edward Rajan Samuel Nadar. Effective textile quality processing and an accurate inspection system using the advanced deep learning technique. *Textile research journal*, 90(9-10):971–980, 2020.

[13] Imane Koulali and M. Taner Eskil. Unsupervised textile defect detection using convolutional neural networks. *Applied Soft Computing*, 113:107913, 2021.

[14] J Judeson Antony Kovilpillai and S Jayanthy. An optimized deep learning approach to detect and classify defective tiles in production line for efficient industrial quality control. *Neural Computing and Applications*, 35(15):11089–11108, 2023.

[15] Moez Krichen. Convolutional neural networks: A survey. *Computers*, 12(8):151, 2023.

[16] Ha Quang Thinh Ngo. Design of automated system for online inspection using the convolutional neural network (CNN) technique in the image processing approach. *Results in Engineering*, 19:101346, 2023.

[17] O Ongbali Samson, A Afolalu Sunday, and Y Salawu Enesi. Bottleneck problems arising in inter-industry production setting and vertical integration: a review. *Technology*, 10(5):606–612, 2019.

[18] Alireza Saberironaghi, Jing Ren, and Moustafa El-Gindy. Defect detection methods for industrial products using deep learning techniques: A review. *Algorithms*, 16(2), 2023.

[19] Somayeh Shahrabadi, Yusbel Castilla, Miguel Guevara, Luís G. Magalhães, Dibet Gonzalez, and Telmo Adão. Defect detection in the textile industry using image-based machine learning methods: a brief review. *Journal of Physics: Conference Series*, 2224(1):012010, apr 2022.

[20] Monica Puri Sikka, Alok Sarkar, and Samridhi Garg. Artificial intelligence (AI) in textile industry operational modernization. *Research Journal of Textile and Apparel*, 28(1):67–83, 2024.

[21] Karen Simonyan. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[22] Henrike Stephani, Thomas Weibel, Ronald Rösch, and Ali Moghiseh. Challenges and approaches when realizing online surface inspection systems with deep learning algorithms. *Discover Data*, 1(1):3, 2023.

[23] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[24] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.

[25] Xian Tao, Dapeng Zhang, Wenzhi Ma, Xilong Liu, and De Xu. Automatic metallic surface defect detection and recognition with convolutional neural networks. *Applied Sciences*, 8(9):1575, 2018.

[26] Mohammad Mustafa Taye. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers*, 12(5):91, 2023.

[27] Nuria Velasco-Pérez, Samuel Lozano-Juárez, Beatriz Gil-Arroyo, Juan Marcos Sanz, Nuño Basurto, Daniel Urda, and Álvaro Herrero. Defect detection in batavia woven fabrics by means of convolutional neural networks. In Pablo García Bringas, Hilde Pérez García, Francisco Javier Martínez de Pisón, Francisco Martínez Álvarez, Alicia Troncoso Lora, Álvaro Herrero, José Luis Calvo Rolle, Héctor Quintián, and Emilio Corchado, editors, *18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023)*, pages 205–215, Cham, 2023. Springer Nature Switzerland.

[28] Bing Wei, Bailing Xu, Kuangrong Hao, and Lei Gao. Textile defect detection using multilevel and attentional deep learning network (MLMA-Net). *Textile Research Journal*, 92(19-20):3462–3477, 2022.

[29] Hongwei Zhang, Wenbo Xiong, Shuai Lu, Mengdan Chen, and Le Yao. QA-USTNet: Yarn-dyed fabric defect detection via u-shaped swin transformer network based on quadtree attention. *Textile Research Journal*, 93(15-16):3492–3508, 2023.

[30] Hongwei Zhang, Weiwei Zhang, Yang Wang, Shuai Lu, Le Yao, and Xia Chen. Colour-patterned fabric-defect detection using unsupervised and memorial defect-free features. *Coloration Technology*, 138(6):602–620, 2022.

Received \jreceived

4.3. Artículo 3: "Deep Learning for Quality Control in Woven Fabrics"

DYNA Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

DEEP LEARNING FOR QUALITY CONTROL IN WOVEN FABRICS

Beatriz Gil-Arroyo[1], Marco Melgarejo[2], Abraham Casas[2], Alejandro López[2], Juan Marcos Sanz[3] and Daniel Urda[1]

1 Grupo de Inteligencia Computacional Aplicada (GICAP), Departamento de Digitalización, Escuela Politécnica Superior, Universidad de Burgos, Av.Cantabria s/n 09006, Burgos, Spain

2 Centro Tecnológico de Componentes-CTC, Parque Científico y Tecnológico de Cantabria (PCTCAN) c/ Isabel Torres, 1 - 39011 Santander (España)

3 Textil Santanderina. N-634, km 43 - 39500 Cabezón de la Sal (España).

DOI: https://doi.org/10.52152/D11517 | Recibido: 06/ago/2025 • Inicio Evaluación: 19/ago/2025 • Aceptado: 20/oct/2025

To cite this article: GIL-ARROYO, Beatriz; MELGAREJO, Marco; CASAS, Abraham; LÓPEZ, Alejandro; SANZ, Juan Marcos; URDA, Daniel. DEEP LEARNING FOR QUALITY CONTROL IN WOVEN FABRICS. DYNA. November December 2025. Vol. 100, n.6, pp. 545-551. DOI: https://doi.org/10.52152/D11517

ABSTRACT:

Automated defect detection in fabrics is a key challenge in quality control within the textile industry. This study proposes a deep learning-based methodology to identify defects in Batavia and Sarga fabrics. In the first stage, an autoencoder was used to filter anomalous images, enabling the creation of a dataset with sufficient defective cases, which are otherwise difficult to obtain in textile production.

Subsequently, convolutional neural networks (DenseNet121, EfficientNetB0/B3, Xception, and VGG) were trained using data augmentation techniques and stratified cross-validation. For Batavia fabrics, DenseNet121 achieved an AU-ROC of 0.88 and an AU-PR of 0.93, demonstrating high detection capability. For Sarga fabrics, three different references (42402, 45433, and 43105) were considered, showing more variable performance across models and datasets. Nonetheless, models such as ResNet101 and Xception achieved competitive results.

The results indicate that the combination of autoencoder and CNN facilitates the generation of balanced datasets and enables consistent defect detection, although performance depends on the type of fabric and the specific reference, suggesting that model selection should be adapted to the characteristics of each case.

Keywords: Defect detection, Textile, Industry 4.0, Deep Learning, Convolutional neural networks, Image analysis, Autoencoder

1. - INTRODUCTION

In recent years, soft computing techniques have addressed numerous challenges in industrial sectors [1]. Despite this, the textile industry remains largely traditional, with minimal adoption of new technologies. Increasing market pressures for higher product quality and faster throughput have exposed the limitations of traditional manual or rule-based inspection systems, which struggle with diverse textures and subtle defects. Recent work shows that deep learning (DL) methods—particularly optimized CNN/YOLO variants and transformer-based anomaly detectors—substantially improve detection accuracy and inference speed on industrial fabric datasets, although challenges such as false positives, limited annotated data, and cross-product generalization persist [2–6].

Currently, textile inspection relies heavily on specially trained workers performing visual checks at around 20 meters per minute, while production lines operate at about 42 meters per minute, creating a bottleneck. These limitations demand innovative, automated solutions [7,8]. This study proposes and validates the use of DL models for automatic quality control of fabrics produced by Textil

Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

Santanderina (TS) S.A., aiming to enhance efficiency, accuracy, and scalability, and to detect defects such as warp and weft flaws, impurities, bowing, skewing, or color gradients at any stage of production.

Artificial neural networks (ANNs), including convolutional neural networks (CNNs) and autoencoders (AEs), are capable of learning complex, non-linear patterns. CNNs are effective for real-time defect localization, while AEs and anomaly detection frameworks identify rare or subtle defects without extensive labeled datasets [2,4,5,6,9]. Integrating multi-scale feature extraction, attention mechanisms, and data-efficient training further improves robustness and generalization across fabric textures. AEs, with encoder–decoder architectures, learn an identity function and reconstruct input data, minimizing reconstruction loss [10]. Variants such as convolutional AEs (CAEs), variational AEs, and skip-connected AEs enhance performance for image processing and representation learning [11,12].

Traditional machine vision relies on handcrafted features, and classical machine learning methods like Support Vector Machines (SVM), Gaussian mixture models, and Bayesian networks require substantial labeled data and struggle with real-world variability [5]. In contrast, DL provides end-to-end models integrating feature extraction and classification. CNNs have demonstrated effectiveness in industrial inspection, including textiles, achieving accurate detection of pattern irregularities and localized defects in fabrics such as Batavia and Sarga [13–15]. Recent approaches also combine generative models, lightweight GANs, and data-efficient strategies to detect anomalies with competitive accuracy and inference speed, suitable for high-throughput production [4,6,9].

The rest of this paper is organized as follows: Section 2 describes the materials and methods, Section 3 presents the results, and Section 4 summarizes the conclusions and proposals for future work.

2. - MATERIAL & METHODS

In this study, AE and CNNs were applied to images collected from fabric production at TS. TS is a prominent European company specializing in the development and manufacturing of yarns, fabrics, finishes, and garments made from both pure and blended materials. TS is one of the few European textile companies with a vertical supply chain, encompassing the entire fabric manufacturing process: spinning, weaving, coating, finishing, garment manufacturing, and logistics. TS produces over 40 million meters of fabric and more than seven million garments annually.

To achieve the research objectives, TS has installed a Basler raL camera equipped with an Awaiba DR-12k-3.5 CMOS sensor, capable of delivering 8 kHz at a 12k resolution array. The camera operates within the VIS-NIR bandwidth and features Basler proprietary optics, offering a resolution of 10 px/mm. For uniform lighting, a LED array emitting at 850 nm with a field of 15mm x 1510mm and 360 infrared (IR) 850nm LEDs has been positioned 20 cm above the fabric at an incidence angle of 15º. IR illumination ensured that there was no interference with the D65 standard lighting systems used for visual inspection. The camera was synchronized with the fabric by using a 10-bit inductive encoder to achieve high-precision positioning (0.1 mm). The lighting conditions, setup geometry, optics, and CMOS sensor specifications were chosen to meet the additional requirements for defect inspection, such as fabric width, batch length, and inspection speed.

Using this setup, images of Batavia and Sarga fabrics were collected over several days in November 2022. A neutral greige color was selected, as it is commonly used in woven fabrics for the summer season and minimizes the variability introduced by different color combinations and patterns. A collection of 16-bit images consisting of approximately 130, 000 images (see examples in Figure 1) was collected for defect inference.

DYNA Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

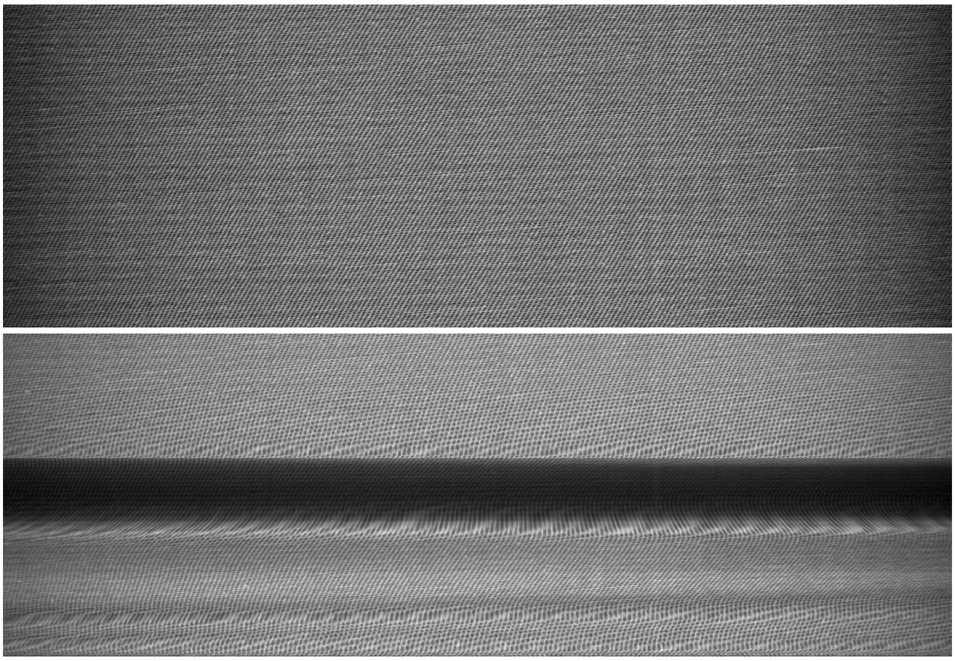


Fig. 1. Examples of controls (top) and cases (bottom) in the dataset

To address this task, two complementary deep learning approaches were explored: AE and CNNs. The following subsections describe the methodology and experimental setup of each approach.

2.1.- AE

AE was used for unsupervised anomaly detection and to filter images with a higher likelihood of defects, creating a balanced training set for CNNs. The AE architecture (Figure 2(a)) includes convolutional layers, ELU activation, batch normalization, and a final sigmoid layer producing grayscale values between 0 and 1. Kernels of 3×3 pixels (encoder) and 4×4 pixels (decoder) with padding 1 and stride 2 were used. The trained AE was validated with labelled images to ensure accurate reconstruction of non-defect samples.

Collecting a large number of textile images is necessary to capture a wide range of anomalies, but manual labelling is tedious and limits dataset size. Training the AE to automatically detect anomalies allows the generation of a larger, defect-rich dataset for CNN classification. The overall defect detection pipeline (Figure 2) predicts non-defect patches, which are subtracted from input images to highlight critical defects after thresholding.

Unsupervised learning focuses on unlabelled data without predefined targets, discerning patterns and relationships essential for clustering, anomaly detection, and exploratory analyses [16]. In contrast, reinforcement learning involves an agent learning to make decisions based on interactions with an environment, effective for complex decision-making tasks [17].

ISSN: 0012-7361 eISSN: 1989-1490 / DYNA. November December 2025. Vol. 100, n.6, pp. 545-551. DOI: https://doi.org/10.52152/D11517

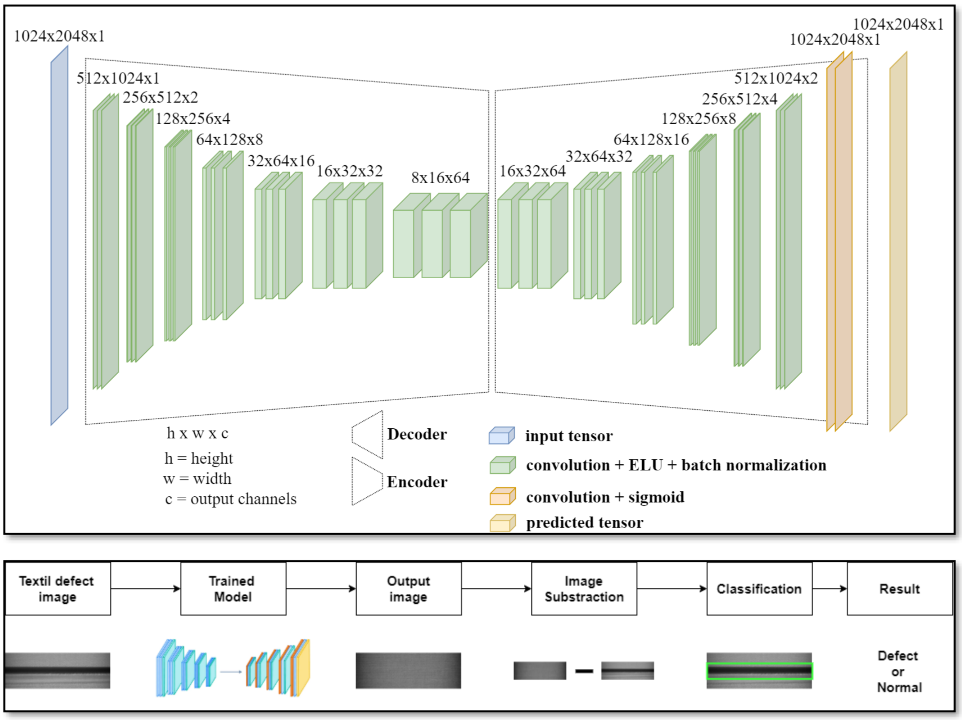


Fig. 2: (a)Trained AE architecture (top). (b) Defect inference protocol (bottom)

2.1.1.- Image preprocessing

Firstly, 16-bits original images were re-scaled to their corresponding 8-bits version, thus reducing the number of possible pixel values in an image from 65536 to 256. For defect inference task, the 8-bits images were cropped to a 1024x2048 size.

2.1.2.- Techniques used for defect inference

The 130,000 images available for defect inference were divided into distinct groups of images, each containing different types of textile fabrics. These groups were distinguished using reference numbers.

The first step was to filter the dataset to obtain defect-free images. The training phase started by running an AE with all the images belonging to a reference number and selecting only outputs that were reconstructed with an error lower than the mean error. This initial filtration ensured that we had potential defect-free images, which were used to train other AE models.

Once the model was trained, all images were introduced as input, with the expectation that a defect-free output image would always be generated. Another filtering step was then performed by selecting reconstructed images with a Structural Similarity Index (SSIM) loss greater than the mean plus three standard deviations. For this task, the class piqa.ssim.SSIM of PIQA[18] was used.

These images were considered anomalous. Therefore, a comparison between the input and output images was conducted, and the resulting grayscale image was binarized, assigning non-zero values to those pixels greater than twice the mean value of the grayscale maximum range.

ISSN: 0012-7361 eISSN: 1989-1490 / DYNA. November December 2025. Vol. 100, n.6, pp. 545-551. DOI: https://doi.org/10.52152/D11517

Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

After that, bounding boxes were generated in the binarized image to enclose and label the defect using cv.connectedComponentsWithStats [19], a function of OpenCV which implements a labelling algorithm of connected components in a binary image.Components with fewer than 150 pixels were ignored. The bounding box with the largest area was selected, and the results were annotated in . json file. Then, using the Label Studio tool [20], an expert reviews and re-labels the defects to validate the inference.

The training parameters of the AE structure were as follows:

- epochs: 20
- patience: 5
- learning rate: 1×10^{-3}
- weight decay: 1×10^{-8}

2.1.3.- Hardware and software used

The versions of the main software packages or libraries upon which the developed code for the defect inference task depends to run properly are:

- Python [21]: 3.11.5
- NumPy [22]: 1.26.4
- OpenCV: [23]: 4.9.0
- Torch [24]: 2.2.2
- Torchvision [25]: 0.17.2

The use of two GPUs was necessary to accelerate the computations. The main hardware specifications are:

- Graphics Card: 2x(Nvidia RTX 3090 24GB)
- Processor: AMD Ryzen 9 12-Core 5900X
- RAM: 2x(32 GB DDR4 2666Mhz)
- Hard Disk: WD Black SN850X 500GB SSD NVMe

2.2.- CNNs

CNNs were employed in this study due to their proven effectiveness in image processing for quality control [26], particularly for defect detection in fabrics. Industrial adoption of artificial vision systems remains limited by challenges such as production stoppages after defect detection, delays, wasted resources, and high reliance on human factors. While some image-processing systems exist in the textile industry, they generally lack artificial intelligence support, which is necessary for rapid training and adaptation to a high variability of products. In the reference textile company (TS), batch sizes range from 500 to 50,000 meters, with over 700 new references annually, emphasizing the need for AI-based solutions capable of generalizing quickly.

In this context, CNNs were applied for supervised binary classification, following a preprocessing stage and an unsupervised anomaly detection phase using autoencoders (AEs). This methodology allowed the evaluation and comparison of state-of-the-art CNN architectures, optimized through *transfer learning*, including:

- VGG16 & VGG19: Known for their simplicity and uniform structure using small 3×3 convolutions, they laid the foundation for deeper networks [27].
- InceptionV3: Introduced parallel multi-scale convolutions and factorization to reduce computational cost while maintaining accuracy [28].
- ResNet50V2 & ResNet101: Leveraged residual connections and pre-activation blocks to enable the training of very deep networks efficiently [29,30].
- Xception: Improved upon Inception by using depthwise separable convolutions for better computational efficiency [31].
- DenseNet121: Enhanced feature propagation through dense connections between layers [32].
- EfficientNetB0 & B3: Used compound scaling to balance depth, width, and resolution, achieving high accuracy with fewer parameters [33].

All models were trained with a learning rate of 0.002, batch size of 128, and two fully connected layers (8,192 neurons each) without dropout regularization. Training ran for up to 100 epochs, with early stopping applied after 10 epochs without validation improvement.

Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

2.2.1.- Dataset

CNNs were trained on image patches from Batavia and Sarga fabrics, taken from a publicly available curated dataset [34]. In Batavia, the twelve-per-image rule was not strictly followed due to manual review, while Sarga adhered to the rule. The final dataset included both defective and non-defective samples (exact numbers: 2,755 large images, 8,782 case patches, and 19,911 control patches for Batavia; 1,548 large images, 173 case patches, and 18,403 control patches for Sarga). All these patches were used for training, validation, and testing of the CNN models in defect classification.

2.2.2.- Data augmentation

To mitigate the class imbalance in the dataset, data augmentation was applied dynamically during training. Transformations included horizontal and vertical flips, as well as brightness and color adjustments, expanding the training set fivefold. This approach increased dataset diversity, reduced overfitting, and improved the models' generalization ability.

2.2.3.- Evaluation strategy

Model evaluation was performed using 5-fold stratified cross-validation, ensuring that each subset maintained the same class distribution. In each iteration, one fold was used as the test set and the remaining four folds for training and validation.

To address class imbalance, a class weighting mechanism was incorporated into the loss function (binary cross-entropy), increasing the importance of the minority class. This adjustment was implemented using the class_weight parameter in Keras' model.fit function.

2.2.4.- Performance metrics

Due to the dataset's imbalance, complementary metrics were considered depending on quality control priorities: detecting defective images, non-defective images, or both. Metrics included Accuracy, Precision, Recall, F1-Score, AU-PR, AU-ROC, and G-Mean, derived from True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

2.2.4.- Computational and implementation environment

Experiments were conducted on a high-performance GPU cluster funded by the InfraRed program of the Junta de Castilla y León, designed for large-scale AI tasks. The system featured 576 AMD Epyc CPU cores, 9.5 TB DDR4 RAM, 320 TB storage, and 30 NVIDIA Tesla A100 GPUs with 40 GB each, delivering 9,360 TFLOPS in FP16 and 4,680 TFLOPS in TF32. The environment included Python 3.9.18 [21] and libraries such as TensorFlow 2.19.0 [35], PyYAML 6.0.2 [36], Pandas 2.2.3 [37], NumPy 1.26.4 [22], Scikit-learn 1.6.1 [38], Matplotlib 3.9.2 [39] and Pillow 8.4.0 [40], enabling efficient GPU-accelerated training and evaluation.

3. - RESULTS

An autoencoder (AE) was first used to filter potential defects, followed by CNNs to classify the highlighted anomalies. Key findings are summarized below.

3.1.- AE

The results for the defect inference task, whose methodology was separately applied to different groups of images tagged with a reference number, are shown in Table 1. Both the inferred anomalies using the methodology based on AE and the true positive defects validated by the qualified personnel of TS are presented.

	UK Zhende Publishing Limited Unit 1804 South Bank Tower, 55 Upper Ground - London (UK) SE1 9EY Phone +86 13851794319 – www.revistadyna.com - email: dyna@revistadyna.com	Pag. 6 / 11

ISSN: 0012-7361 eISSN: 1989-1490 / DYNA. November December 2025. Vol. 100, n.6, pp. 545-551. DOI: https://doi.org/10.52152/D11517

DYNA Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

Reference	Total of images	Inferred anomalies	(%)	True positive defects	(%)
42410	12,786	950	7	187	20
44406	16,952	851	5	73	9
44423	4,146	551	13	22	4
42402	17,763	1091	6	265	24
40129	18,603	208	1	13	6
45433	29,969	449	1	60	13
43105	15,536	581	4	87	15
43112	14,018	784	6	34	4

Table 1. Quantity of image with defects inferred per groups of images of different textile fabric references. Only inferred defective images were validated by qualified personnel of TS

Reference numbers, such as 44406, 44423, 40129, and 43112, whose images show a relatively low percentage of true positive defects, belong to images of more complex textile fabrics with a hard-to-predict houndstooth pattern. Nevertheless, the rest of the

reference groups show more than a 10% true positive defects, easily increasing the database size and diminishing the effort of the qualified personnel, who only have to search for defects in the subgroup of images of the inferred anomalies.

3.2.- CNNs

It is important to note that the CNNs were trained using the entire published dataset [34]; however, the results presented here correspond exclusively to the predictions on the subset of images that had been previously detected as anomalous by the AE, as shown in Table 2. Only the segments where the AE identified potential defects were retained for further analysis. This approach allowed us to focus the evaluation on the most challenging cases, where potential defects were more likely to appear. Each large image in the aforementioned published dataset was generally divided into 12 patches for CNN analysis; however, in some datasets (e.g., Batavia), the number of patches per image was lower due to the removal of uncertain patches.

Reference	Fabric	Inferred anomalies by AE	Images results CNNs	Patches results CNNs
42410	Batavia	950	258	2,485
44406	Esterilla	851	0	0
44423	Sarga	551	0	0
42402	Sarga	1,091	133	1,596
40129	Sarga	208	0	0
45433	Sarga	449	25	300
43105	Sarga	581	72	864
43112	Sarga	784	0	0

Table 2. Summary of AE-inferred anomalies and CNN results at large-image and patch level from the published dataset.

The following tables 3-6 present the classification results for the Batavia and Sarga fabrics. For the Batavia fabric, the results are available for a single reference, identified as 42410. In contrast, for the Sarga fabric, results were obtained from three distinct references: 42402, 45433, and 43105.

The results presented in Table 3 correspond to reference 42410, which pertains to the Batavia fabric.

DYNA Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

CNN	Threshold	TP	FP	FN	TN	Acc.	Prec.	Rec.	F1	GMean	AU-ROC	AU-PR
DenseNet121	0.325	1427	311	202	545	**0.79**	0.82	0.88	**0.85**	**0.75**	**0.88**	**0.93**
VGG19	0.625	1428	419	201	437	0.75	0.77	0.88	0.82	0.67	0.83	0.91
ResNet50V2	0.375	1519	509	110	347	0.75	0.75	**0.93**	0.83	0.61	0.81	0.89
ResNet101	0.450	1227	297	402	559	0.72	0.81	0.75	0.78	0.70	0.78	0.88
EfficientNetB3	0.325	730	49	899	807	0.62	**0.94**	0.45	0.61	0.65	0.77	0.88
Xception	0.275	719	92	910	764	0.60	0.89	0.44	0.59	0.63	0.74	0.86
InceptionV3	0.500	1349	481	280	375	0.69	0.74	0.83	0.78	0.60	0.74	0.84
VGG16	0.375	1407	574	222	282	0.68	0.71	0.86	0.78	0.53	0.72	0.84
EfficientNetB0	0.450	886	209	743	647	0.62	0.81	0.54	0.65	0.64	0.70	0.83

Table 3. Evaluation results of CNN models for reference 42410, sorted by AU-ROC value

DenseNet121 model achieved the best overall performance, with an AU-ROC of 0.88 and an AU-PR of 0.93, indicating a strong ability to distinguish between defective and non-defective patches even in the presence of class imbalance. Its F1-Score of 0.85 and recall of 0.88, further confirm its robustness in detecting anomalies. VGG19 and ResNet50V2 also showed competitive results, with AU-ROC values of 0.83 and 0.81, respectively. Despite slight variations in precision and recall across the models, the top-performing networks maintained a balanced trade-off between false positives and false negatives. In contrast, architectures such as EfficientNetB3, Xception, and VGG16 yielded lower scores, particularly in recall, suggesting reduced sensitivity to subtle defect patterns within Batavia. This discrepancy highlights the importance of model selection when addressing fine-grained texture analysis in high-regularity fabrics.

Table 4 displays the evaluation outcomes for reference 42402, associated with theSarga fabric.

CNN	Threshold	TP	FP	FN	TN	Acc.	Prec.	Rec.	F1	GMean	AU-ROC	AU-PR
ResNet101	0.850	34	11	28	1523	**0.98**	**0.76**	0.55	**0.64**	0.74	**0.94**	**0.65**
InceptionV3	0.925	33	12	29	1522	0.97	0.73	0.53	0.62	0.73	0.93	0.62
ResNet50V2	0.950	35	67	27	1467	0.94	0.34	0.56	0.43	0.73	0.91	0.52
Xception	0.950	34	34	28	1500	0.96	0.50	0.55	0.52	0.73	0.89	0.45
VGG19	0.905	21	22	41	1512	0.96	0.49	0.34	0.40	0.58	0.85	0.30
DenseNet121	0.600	44	329	18	1205	0.78	0.12	**0.71**	0.20	**0.75**	0.75	0.19
VGG16	0.725	20	15	42	1519	0.96	0.57	0.32	0.41	0.57	0.74	0.27
EfficientNetB0	0.775	12	6	50	1528	0.96	0.67	0.19	0.30	0.44	0.72	0.27
EfficientNetB3	0.550	8	21	54	1513	0.95	0.28	0.13	0.18	0.36	0.66	0.16

Table 4. Evaluation results of CNN models for reference 42402, sorted by AU-ROC value

Among the evaluated models, ResNet101 demonstrated the best overall performance, achieving an AU-ROC of 0.94 and an AU-PR of 0.65. These results indicate a solid capacity to discriminate between defective and non-defective regions, despite a moderate class imbalance. Its F1-Score of 0.64 and recall of 0.55 reflect a good balance between sensitivity and precision. InceptionV3 followed closely with an AU-ROC of 0.93 and a slightly lower F1-Score of 0.62, confirming its effectiveness for this type of fabric. Models such as ResNet50V2 and Xception showed intermediate performance, with AU-ROC values of 0.91 and 0.89, respectively, although their precision and F1-Score were more modest. Conversely, architectures such as EfficientNetB3, DenseNet121, and VGG16 exhibited lower scores, particularly in recall and F1-Score, suggesting difficulties in correctly identifying subtle defects in the Sarga weave of this reference.

Table 5 shows the evaluation results of the different CNNs for reference 45433, corresponding to the Sarga fabric. The table was sorted by the AU-ROC value to facilitate the comparison of the model performance in terms of discriminative power.

UK Zhende Publishing Limited
Unit 1804 South Bank Tower, 55 Upper Ground - London (UK) SE1 9EY
Phone +86 13851794319 – www.revistadyna.com - email: dyna@revistadyna.com

ISSN: 0012-7361 eISSN: 1989-1490 / DYNA. November December 2025. Vol. 100, n.6, pp. 545-551. DOI: https://doi.org/10.52152/D11517

Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

CNN	Threshold	TP	FP	FN	TN	Acc.	Prec.	Rec.	F1	GMean	AU-ROC	AU-PR
Xception	*0.905*	*87*	*29*	*65*	*119*	***0.69***	*0.75*	*0.57*	*0.65*	***0.68***	***0.76***	*0.73*
ResNet101	*0.850*	*66*	*17*	*86*	*131*	*0.66*	*0.80*	*0.43*	*0.56*	*0.62*	*0.74*	***0.76***
InceptionV3	*0.925*	*62*	*15*	*90*	*133*	*0.65*	*0.81*	*0.41*	*0.54*	*0.61*	*0.74*	*0.75*
VGG19	*0.950*	*53*	*18*	*99*	*130*	*0.61*	*0.75*	*0.35*	*0.48*	*0.55*	*0.70*	*0.67*
VGG16	*0.725*	*56*	*17*	*96*	*131*	*0.62*	*0.77*	*0.37*	*0.50*	*0.57*	*0.67*	*0.66*
ResNet50V2	*0.950*	*68*	*37*	*84*	*111*	*0.60*	*0.65*	*0.45*	*0.53*	*0.58*	*0.67*	*0.68*
DenseNet121	*0.600*	*111*	*65*	*41*	*83*	*0.65*	*0.63*	***0.73***	***0.68***	*0.64*	*0.62*	*0.59*
EfficientNetB3	*0.550*	*24*	*5*	*128*	*143*	*0.56*	***0.83***	*0.16*	*0.27*	*0.39*	*0.60*	*0.64*
EfficientNetB0	*0.775*	*21*	*8*	*131*	*140*	*0.54*	*0.72*	*0.14*	*0.23*	*0.36*	*0.58*	*0.61*

Table 5. Evaluation results of CNN models for reference 45433, sorted by AU-ROC value

Among the models evaluated for Sarga fabric reference 45433, the Xception model achieved the best performance, reaching an AU-ROC of 0.76 and an AU-PR of 0.73, a F1-Score of 0.65 and a recall of 0.57. These results indicate a reasonable balance between the sensitivity and precision. Although other models such as ResNet101 and InceptionV3 reached an AU-ROC of 0.74, their F1-Score and recall values were lower (0.56 and 0.54, respectively), suggesting a reduced ability to correctly identify defective regions despite high precision. Models like ResNet50V2, VGG19, and VGG16 demonstrated intermediate performance, with AU-ROC values between 0.67 and 0.70, but limited defect detection capabilities as reflected in their F1-Score, ranging from 0.48 to 0.53. On the other hand, architectures such as EfficientNetB0 and EfficientNetB3 showed considerably lower results, especially in terms of recall (0.14 and 0.16, respectively) and F1-Score (0.23 and 0.27), highlighting substantial difficulties in detecting defects in this fabric.

Table 6 presents the evaluation results for the CNNs applied to reference 43105, which corresponds to the Sarga fabric.

CNN	Threshold	TP	FP	FN	TN	Acc.	Prec.	Rec.	F1	GMean	AU-ROC	AU-PR
Xception	*0.950*	*20*	*21*	*6*	*817*	*0.97*	*0.49*	***0.77***	***0.60***	***0.87***	***0.97***	*0.48*
InceptionV3	*0.925*	*13*	*5*	*13*	*833*	***0.98***	***0.72***	*0.50*	*0.59*	*0.70*	***0.97***	***0.62***
ResNet101	*0.850*	*13*	*9*	*13*	*829*	*0.97*	*0.59*	*0.50*	*0.54*	*0.70*	*0.93*	*0.39*
VGG19	*0.950*	*10*	*16*	*16*	*822*	*0.96*	*0.38*	*0.38*	*0.38*	*0.61*	*0.91*	*0.29*
ResNet50V2	*0.950*	*12*	*40*	*14*	*798*	*0.94*	*0.23*	*0.46*	*0.31*	*0.66*	*0.87*	*0.32*
VGG16	*0.725*	*11*	*12*	*15*	*826*	*0.97*	*0.48*	*0.42*	*0.45*	*0.65*	*0.77*	*0.32*
DenseNet121	*0.600*	*20*	*169*	*6*	*669*	*0.80*	*0.11*	***0.77***	*0.19*	*0.78*	*0.77*	*0.13*
EfficientNetB0	*0.775*	*4*	*5*	*22*	*833*	*0.97*	*0.44*	*0.15*	*0.23*	*0.39*	*0.64*	*0.17*
EfficientNetB3	*0.550*	*2*	*11*	*24*	*827*	*0.96*	*0.15*	*0.08*	*0.10*	*0.28*	*0.60*	*0.11*

Table 6. Evaluation results of CNN models for reference 43105, sorted by AU-ROC value

Xception model obtained the highest AU-ROC value (0.97), along with a recall of 0.77 and an F1-Score of 0.60, which reflects a strong ability to detect defective regions despite a relatively low precision (0.49). InceptionV3 also performed well in terms of AU-ROC (0.97), achieving a higher precision (0.72) but lower recall (0.50), leading to a comparable F1-Score of 0.59. ResNet101 ranked third in AU-ROC (0.93), although it showed a more balanced but slightly weaker performance across precision, recall, and F1-Score. Other architectures, such as VGG19, ResNet50V2, and VGG16, demonstrated moderate results, with AU-ROC values between 0.77 and 0.91, and limited effectiveness in terms of recall and F1-Score, indicating reduced sensitivity to defects. DenseNet121, despite identifying 20 true positives, generated a large number of false positives (169), leading to low precision (0.11) and an F1-Score of only 0.19. EfficientNetB0 and EfficientNetB3 showed the lowest performance metrics overall, with AU-ROC values below 0.65 and F1-Score under 0.25, highlighting their poor suitability for this particular fabric reference.

4. - CONCLUSIONS

This study has addressed the critical challenge of defect detection in textile fabrics using DL models to enhance the quality control process at TS, focusing on autoencoders (AEs) and CNNs to improve automated defect detection and dataset quality.

UK Zhende Publishing Limited
Unit 1804 South Bank Tower, 55 Upper Ground - London (UK) SE1 9EY
Phone +86 13851794319 – www.revistadyna.com - email: dyna@revistadyna.com
Pag. 9 / 11
ISSN: 0012-7361 eISSN: 1989-1490 / DYNA. November December 2025. Vol. 100, n.6, pp. 545-551. DOI: https://doi.org/10.52152/D11517

Ingeniería e Industria	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

The defect inference methodology proved useful for assisting qualified personnel in efficiently generating a database with a high number of defective images, improving model training and subsequent classification outcomes.

Pretrained CNN models effectively improved fabric defect detection, with DenseNet121 excelling in Batavia (AU-ROC 0.88), ResNet101 in Sarga 42402 (AU-ROC 0.94), and Xception in Sarga 45433/43105 (AU-ROC 0.76/0.97). Performance varies across Sarga references: DenseNet121 has high recall but many false positives, while InceptionV3 and ResNet101 balance AU-ROC with lower recall. VGG and EfficientNet are stable in accuracy but less robust to subtle or heterogeneous defects, showing no single model is optimal for all cases.

Combining multiple CNN architectures in a hybrid system could leverage complementary strengths, improving detection performance, reducing missed defects, and controlling false positives, though such strategies require experimental validation under real production conditions.

Key metrics like F1-score, G-mean, and AU-ROC were essential to evaluate model performance, guiding the selection of the best models for each fabric and demonstrating the advantages of pretrained neural networks over traditional methods like autoencoders. Selecting DL models should be customized for each fabric type to maximize precision and efficiency. CNNs can enhance precision and scalability, though integration into production lines requires real-world testing.

Future research should focus on real-time defect detection, exploring attention mechanisms, reinforcement learning, and semantic segmentation to improve robustness, interpretability, and precision. Additionally, analyzing false positives and negatives, and adapting models to different fabrics and lighting conditions, is essential for broader industrial applicability. Real-time detection combined with continuous feedback loops can further optimize quality control systems, but experimental validation in industrial environments is required to confirm these improvements.

OTHER SECTIONS:

REFERENCES

1. Diez-Olivan A, Del Ser J, Galar D, Sierra B. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. Information Fusion. 2019;50:92–111. https://doi.org/10.1016/j.inffus.2018.10.005
2. Zhou L, Ma B, Dong Y, Yin Z, Lu F. DCFE-YOLO: A novel fabric defect detection method. PLOS ONE. 2025;20(1):e0314525. https://doi.org/10.1371/journal.pone.0314525
3. Li X, Zhu Y. PEI-YOLOv5: A real-time and accurate convolutional neural network for fabric defect detection. Complex & Intelligent Systems. 2023;10(3):1–17. https://doi.org/10.1007/s40747-023-01317-8
4. Yang R, Guo N, Yu M, Tian B. Fabric defect detection via saliency model based on adjacent context coordination and transformer (ACCTNet). Measurement: Sensors.2024. https://doi.org/10.1177/15589250241258272
5. Carrilho R, Yaghoubi E, Lindo J, Hambarde K, Proença H. Toward automated fabric defect detection: A survey of recent computer vision approaches. Electronics. 2024;13(18):3728. https://doi.org/10.3390/electronics13183728
6. Yang X. Research on real-time detection of fabric defects based on an improved ELO rating algorithm. Scientific Reports. 2025;15:32065. https://doi.org/10.1038/s41598-025-17747-y
7. Ongbali OS, Afolalu AS, Salawu YE. Bottleneck problems arising in inter-industry production setting and vertical integration: a review. Technology. 2019;10(5):606–12.
8. Shahrabadi S, Castilla Y, Guevara M, Magalhães LG, Gonzalez D, Adão T. Defect detection in the textile industry using image-based machine learning methods: a brief review. J Phys: Conf Ser. 2022;1:012010. https://doi.org/10.1088/1742-6596/2224/1/012010
9. Gil-Arroyo B, Velasco-Pérez N, Sanz JM, Casas A, Arroyo A, Urda D. Defect detection in Batavia and Sarga woven fabrics by means of convolutional neural networks. J Appl Logics. 2025. Accepted for publication.
10. Chollet F. Building autoencoders in Keras. Keras Blog. 2016. Available from: https://blog.keras.io/building-autoencoders-in-keras.html [Accessed 2 Oct 2025].
11. Fast Forward Labs. Under the hood of the variational autoencoder in prose and code. Fast Forward Labs Blog. 2016. Available from: https://blog.fastforwardlabs.com/2016/08/12/introducing-variational-autoencoders-in-prose-and-code.html [Accessed 2 Oct 2025].
12. Kim J, Ko J, Choi H, Kim H. Printed circuit board defect detection using deep learning via a skip-connected convolutional autoencoder. Sensors. 2021;21(15):1–15. https://doi.org/10.3390/s21154968
13. Weimer D, Scholz-Reiter B, Shpitalni M. Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection. CIRP Ann. 2016;65(1):417–20. https://doi.org/10.1016/j.cirp.2016.04.072
14. Dai W, Mujeeb A, Erdt M, Sourin A. Soldering defect detection in automatic optical inspection. Adv Eng Inform. 2020;43:101004. https://doi.org/10.1016/j.aei.2019.101004
15. Velasco-Pérez N, Lozano-Juárez S, Gil-Arroyo B, Sanz JM, Basurto N, Urda D, et al. Defect detection in Batavia woven fabrics by means of convolutional neural networks. In: García Bringas P, Pérez García H, Martínez de Pisón FJ, Martínez Álvarez F, Troncoso Lora A, Herrero Á, et al., editors. 18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023); Cham: Springer Nature Switzerland; 2023. p. 205–15.
16. IBM. What is machine learning? 2018. Available from: https://www.ibm.com/topics/machine-learning [Accessed 2 Oct 2025].

	Deep Learning for Quality Control in Woven Fabrics	UNESCO Discipline 1203 Computer Sciences
RESEARCH ARTICLE	Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz and Daniel Urda	Sub-discipline 04 Artificial Intelligence

17. Sutton RS, Barto AG. Reinforcement Learning: An Introduction. 2nd ed. MIT Press; 2018.
18. Rozet F. piqa.ssim.SSIM class. PIQA: PyTorch Image Quality Assessment. 2023. Available from: https://piqa.readthedocs.io/en/stable/api/piqa.ssim.html [Accessed 2 Oct 2025].
19. OpenCV. connectedComponentsWithStats(). 2024. Available from: https://docs.opencv.org/ [Accessed 2 Oct 2025].
20. HumanSignal. label-studio. 2024. Available from: https://labelstud.io/ [Accessed 2 Oct 2025].
21. Python Software Foundation. Python: A dynamic, open source programming language. 2023. Available from: https://www.python.org/ [Accessed 2 Oct 2025].
22. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020;585(7825):357–62. https://doi.org/10.1038/s41586-020-2649-2
23. Bradski G. The OpenCV Library. Dr. Dobb's Journal of Software Tools. 2000. Available from: https://opencv.org/ [Accessed 2 Oct 2025].
24. Ansel J, Yang E, He H, Gimelshein N, Jain A, Vozne-sensky M, et al. PyTorch 2: Faster Machine Learning Through Dynamic Python Bytecode Transformation and Graph Compilation. In: 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2 (ASPLOS '24). ACM; April 2024.
25. TorchVision maintainers and contributors. TorchVision: PyTorch's Computer Vision library. November 2016. Available from: https://pytorch.org/vision/ [Accessed 2 Oct 2025].
26. Saberironaghi A, Ren J, El-Gindy M. Defect detection methods for industrial products using deep learning techniques: A review. Algorithms. 2023;16(2):1–30. https://doi.org/10.3390/a16020095
27. Simonyan K. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014. Available from: https://arxiv.org/abs/1409.1556 [Accessed 2 Oct 2025].
28. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 2818–26. https://doi.org/10.1109/CVPR.2016.308
29. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 770–8. https://doi.org/10.1109/CVPR.2016.90
30. He K, Zhang X, Ren S, Sun J. Identity mappings in deep residual networks. In: Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14. Springer; 2016. p. 630–45. https://doi.org/10.1007/978-3-319-46493-0_38
31. Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017. p. 1251–8. https://doi.org/10.1109/CVPR.2017.195
32. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017. p. 4700–8. https://doi.org/10.1109/CVPR.2017.243
33. Tan M, Le Q. EfficientNet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning; 2019. p. 6105–14. PMLR. https://proceedings.mlr.press/v97/tan19a.html
34. Gil-Arroyo B, Sanz JM, Arroyo Á, Urda D, Basurto N, Herrero Á. Dataset for defect detection in textile manufacturing. Data in Brief. 2025;59:111451. https://doi.org/10.1016/j.dib.2025.111451
35. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467. 2016. Available from: https://www.tensorflow.org/ [Accessed 2 Oct 2025].
36. Simonov A. PyYAML. 2006. Available from: https://pyyaml.org/ [Accessed 2 Oct 2025].
37. McKinney W. Data Structures for Statistical Computing in Python. In: van der Walt S, Millman J, editors. Proceedings of the 9th Python in Science Conference; 2010 Jun 28–Jul 3; Austin, TX, USA. p. 51–6. Available from: https://pandas.pydata.org/ [Accessed 2 Oct 2025].
38. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12:2825–30. Available from: https://scikit-learn.org/ [Accessed 2 Oct 2025].
39. Hunter JD. Matplotlib: A 2D Graphics Environment. Comput Sci Eng. 2007;9(3):90–5. Available from: https://matplotlib.org/ [Accessed 02 Oct 2025].
40. Clark A. Pillow (PIL Fork) Documentation. 2015. Available from: https://pillow.readthedocs.io/en/stable/ [Accessed 2 Oct 2025].

5. Discusión General y Conclusiones

Este capítulo presenta la discusión integrada de los resultados obtenidos a lo largo de la tesis, analizados en relación con el objetivo general y los objetivos específicos planteados. Asimismo, se sintetizan las principales aportaciones científicas y tecnológicas, se describen las limitaciones del estudio, se resaltan las principales conclusiones y se proponen líneas futuras de investigación derivadas de los hallazgos.

5.1. Discusión Integrada

La revisión del estado del arte muestra una clara evolución en la inspección automática de defectos textiles, desde métodos clásicos de visión por computador basados en filtros y transformadas (Liu et al., 2011 [7]; Modrângă et al., 2017 [8]), hasta sistemas de visión inteligente basados en DL, tanto supervisado (Biradar et al., 2021 [17]; Bing Wei et al., 2022 [18]; Li y Zhu, 2024 [19]; Yang et al., 2024 [20]; Zhou et al., 2025 [21]) como no supervisado (Koulali y Eskil, 2021 [22]; Zhang et al., 2022 [23], 2023 [24]; Si y Kim, 2024 [25]). Los enfoques supervisados han demostrado alta precisión cuando se dispone de suficientes datos etiquetados, mientras que los métodos no supervisados son esenciales en escenarios donde los defectos son escasos o difíciles de clasificar.

En este marco, los resultados de la tesis aportan varias ventajas comparativas frente al estado del arte, siguiendo la estructura de las tres publicaciones principales:

La primera publicación de la tesis destaca por la generación y liberación de un dataset original que supera varias de las limitaciones de los conjuntos de datos públicos más utilizados en la literatura, como PFD, TILDA o Lusitano. Mientras que estos datasets suelen presentar restricciones en cuanto al número de muestras, la diversidad de tejidos o el acceso a los datos, el conjunto desarrollado incorpora 2.755 imágenes originales de tejido Batavia y 1.548 de tejido Sarga, a partir de las cuales se generaron más de 47.000 patches cuidadosamente etiquetados. Esta mayor escala y diversidad permiten entrenar y evaluar modelos de DL en condiciones más realistas, favoreciendo análisis comparativos más robustos basados en métricas exigentes como AU-ROC y Area Under the Precision-Recall curve (AU-PR), y mejorando la capacidad de generalización de los modelos frente a escenarios industriales reales.

En la segunda publicación se presenta una implementación y comparativa de nueve arquitecturas CNN preentrenadas, evaluadas mediante métricas estandarizadas (accuracy, precision, recall, F1-score, AU-ROC y AU-PR). Esta aproximación demuestra que las CNN pueden detectar defectos en tejidos de manera precisa y consistente, superando las limitaciones de la inspección manual y de los métodos clásicos de visión por computador. La comparación detallada permitió identificar DenseNet121 como la arquitectura más eficaz para tejidos Batavia e InceptionV3 como la más adecuada para tejidos Sarga, ofreciendo criterios prácticos para la selección de modelos según las características del tejido, especialmente relevantes en escenarios con desbalance de clases, donde métricas globales como accuracy resultan insuficientes. Esto representa un valor añadido frente a estudios previos, que suelen evaluar arquitecturas sobre datasets más limitados o considerando menos modelos, mientras que aquí se comparan y analizan nueve CNN.

En la tercera publicación se combina un AE no supervisado para prefiltrado de imágenes potencialmente defectuosas con CNN supervisadas para la clasificación final. El AE permite reducir significativamente la carga de trabajo de los expertos y mejorar la calidad del dataset final, mientras que las CNN analizan las imágenes preseleccionadas, alcanzando un desempeño elevado incluso sobre las muestras más complejas. Este enfoque híbrido demuestra que la combinación de métodos no supervisados y supervisados puede aumentar la eficiencia y escalabilidad del sistema, abordando escenarios que los métodos estrictamente supervisados o no supervisados no cubren por completo.

En definitiva, la ausencia de métricas basadas en curvas (AU-ROC, AU-PR) en una parte relevante de los trabajos revisados en estado del arte, junto con el uso frecuente de datasets privados o de tamaño reducido, dificulta la comparación objetiva del rendimiento y la evaluación del comportamiento de los modelos en escenarios con desbalance de clases. Este hecho refuerza la relevancia de los resultados presentados en esta tesis, obtenidos sobre un dataset público de mayor escala y evaluados mediante métricas más exigentes.

La tesis ha cumplido de manera eficaz el objetivo general de desarrollar, analizar y evaluar técnicas de DL para la detección automatizada de defectos en tejidos. Tal como se muestra en la Tabla 3, existe una correspondencia directa entre los objetivos específicos de la tesis y las contribuciones de cada publicación.

Los resultados obtenidos a lo largo de las tres publicaciones consolidan la viabilidad de los enfoques basados en DL para la inspección automática de tejidos. La liberación de un dataset realista, la validación exhaustiva de arquitecturas CNN y la propuesta de una metodología híbrida AE + CNN constituyen aportaciones complementarias que refuerzan tanto el valor científico como el potencial tecnológico de la tesis. Aunque los sistemas desarrollados no han sido aún desplegados en tiempo real en una línea de producción, la estructura metodológica, la validación experimental y los resultados alcanzados sientan una base sólida para futuras implementaciones industriales y procesos de transferencia tecnológica.

5.2. Aportaciones Científicas y Tecnológicas

En esta sección se sintetizan las principales contribuciones de la tesis, tanto en el ámbito científico como en el tecnológico. Estas aportaciones ponen de manifiesto la novedad del enfoque propuesto y su aplicabilidad práctica, así como el conocimiento generado para la comunidad académica e industrial.

5.2.1. Aportaciones Científicas

Las principales contribuciones científicas derivadas de este trabajo son:

- **Dataset público y validado de tejidos Batavia y Sarga.** Recurso abierto con metadatos completos, estructura documentada y eliminación de muestras ambiguas. Constituye un benchmark de referencia para futuras investigaciones en detección de defectos textiles [48].

- **Análisis experimental detallado de múltiples arquitecturas CNN.** Se aporta evidencia comparada sobre DenseNet, Inception, Xception y otras variantes en tejidos con morfologías distintas, ofreciendo información valiosa para la selección de modelos.

- **Metodología híbrida AE + CNN.** Se propone y valida un enfoque que integra aprendizaje no supervisado para el prefiltrado y enriquecimiento de defectos. Además de los sólidos resultados de detección obtenidos con las CNN, la validación del AE como prefiltrado demuestra un beneficio adicional: permite reducir significativamente el coste de anotación manual al facilitar la identificación automática de imágenes potencialmente anómalas para su revisión por expertos.

- **Gestión del desbalance de clases y validación mediante métricas robustas.** Dado que los defectos textiles representan una proporción minoritaria en el dataset, se implementaron estrategias específicas para mitigar el impacto del desbalance durante el entrenamiento. Entre ellas se incluyen DA aplicado al vuelo y balanceo de clases mediante ajuste de la función de pérdida, ponderando más los errores en la clase minoritaria. La validación se realizó mediante k-fold cross-validation balanceada, asegurando que cada partición mantuviera la proporción de clases, lo que proporciona estimaciones más fiables del desempeño del modelo. Además, los modelos fueron evaluados mediante métricas robustas frente al desbalance, como AU-ROC , AU-PR y Geometric Mean (G-Mean), junto con medidas tradicionales (accuracy, precision, recall y F1-score), garantizando una valoración completa y consistente del desempeño de las CNN en escenarios con distribución de clases desequilibrada.

- **Impacto del tipo de tejido.** Se evidencia que las CNN presentan un desempeño diferenciado entre Batavia y Sarga, lo que sugiere que la estructura del tejido influye en el rendimiento.

5.2.2. Aportaciones Tecnológicas y Transferencia

Las aportaciones tecnológicas se caracterizan por su aplicabilidad directa en entornos industriales:

- **Sistema de inspección automático objetivo y trazable.** Las CNN permiten un proceso de detección más consistente que la inspección humana, con decisiones cuantificables y replicables.

- **Metodología completa de implementación.** Desde la adquisición de datos hasta la evaluación comparativa, proporcionando una guía aplicable a contextos industriales reales (ver figura 2).

- **Base técnica para su integración en línea.** La eficiencia de las arquitecturas seleccionadas y su estabilidad frente a variaciones permiten proyectar su uso en sistemas en tiempo real.

- **Transferencia al proyecto DECENT y a la empresa TS.** Mediante el desarrollo de un modelo de IA plenamente funcional, la tesis ha contribuido de manera directa a disponer de esta herramienta, que proporciona una base sólida para la mejora de procesos internos y el avance tecnológico en colaboración industrial, ofreciendo un recurso listo para su futura implantación en un entorno operativo real.

- **Impulso hacia la industria 4.0.** La automatización basada en DL facilita procesos más inteligentes, eficientes y escalables.

5.3. Limitaciones

Las limitaciones del estudio no deben interpretarse como restricciones estrictas, sino como aspectos que acotan el alcance del trabajo y que ofrecen margen de mejora para futuras investigaciones:

- **Clasificación binaria.** El enfoque actual distingue entre defecto y no defecto. Aunque esta aproximación resulta útil para la inspección automática, una clasificación más detallada —por tipología o severidad del defecto— podría aportar valor añadido en aplicaciones industriales más avanzadas.

- **Dependencia del patrón del tejido.** Se observó que el rendimiento varía entre los tejidos Batavia y Sarga, lo cual refleja la sensibilidad natural de los modelos a la estructura visual del material. Esto no limita su utilización, pero sí indica que la generalización a tejidos muy distintos requerirá investigar y realizar ajustes específicos.

- **Eliminación de muestras ambiguas.** La depuración del dataset permitió obtener un conjunto de entrenamiento más consistente. No obstante, es posible que algunos defectos extremadamente sutiles quedaran fuera, debido a la dificultad inherente a su identificación incluso por parte de expertos.

- **Dependencia del sistema de adquisición.** Los modelos se optimizaron para un sistema óptico específico; futuras pruebas en diferentes sistemas permitirán confirmar la robustez y la adaptabilidad del enfoque.

- **Tejidos en escala de grises.** El trabajo se ha centrado en tejidos grises, lo que facilita el control experimental. Sin embargo, los tejidos coloreados o estampados introducen variaciones adicionales que no han sido consideradas en esta fase. Explorar estos escenarios más complejos permitiría evaluar la aplicabilidad del sistema en un abanico más amplio de contextos industriales.

- **Validación operativa en línea.** Los modelos han mostrado un rendimiento elevado sobre los datos disponibles, si bien su comportamiento en un sistema completamente operativo en tiempo real aún no ha sido evaluado. Aspectos como la latencia o la estabilidad bajo condiciones industriales dinámicas requerirán una fase adicional de pruebas.

5.4. Conclusiones

En primer lugar, se completó y depuró un dataset amplio y representativo, compuesto por 2.755 imágenes de tejidos Batavia y 1.548 imágenes de tejidos Sarga, de las cuales se generaron 28.693 patches de Batavia y 18.576 patches de Sarga para entrenamiento y validación de los modelos de deep learning, optimizando la calidad de los datos y facilitando la evaluación robusta de los modelos.

En segundo lugar, la comparativa de nueve arquitecturas CNN preentrenadas mostró qué modelos alcanzan un rendimiento superior según el tipo de tejido. Para tejidos Batavia, DenseNet121 obtuvo los mejores resultados, con un AU-ROC de 0,87, un F1-score equilibrado y alta precisión, mientras que para tejidos Sarga, InceptionV3 alcanzó un AU-ROC de hasta 0,93.

En tercer lugar, mediante el enfoque híbrido de AE y CNN se alcanzó un AU-ROC de 0,88 para DenseNet121 en Batavia, 0,94 para ResNet101 en Sarga 42402 y hasta 0,97 para Xception en referencias Sarga 43105. Es importante destacar que estos resultados se obtuvieron exclusivamente sobre los patches previamente identificados como anómalos por el AE , es decir, sobre un subconjunto de muestras especialmente complejas y visualmente ambiguas, donde los modelos CNN tienden a presentar mayores dificultades de clasificación. A pesar de esta mayor complejidad, los modelos alcanzaron un desempeño elevado, lo que pone de manifiesto la robustez del enfoque propuesto. Esto evidencia que, mientras que para tejidos Batavia DenseNet121 sigue siendo el modelo óptimo incluso en escenarios de alta dificultad, para tejidos Sarga la selección del modelo debe adaptarse a las características de cada referencia, y no necesariamente coincide con el modelo identificado como óptimo en el análisis general del artículo 2.

5.5. Líneas de Trabajo Futuras

A partir de los resultados y limitaciones identificados, se proponen las siguientes líneas de investigación:

- **Generalización a nuevos tejidos y dominios.** Ampliación del dataset con nuevos patrones, materiales y condiciones (color, texturas complejas), incluyendo técnicas de *domain adaptation* que minimicen la necesidad de etiquetar grandes cantidades de datos adicionales. Entre las estrategias posibles se encuentran el fine-tuning con un conjunto reducido de imágenes etiquetadas, el aprendizaje de representaciones compartidas entre dominios y métodos adversariales que promuevan invariancia frente a los cambios de dominio, facilitando así la extensión de los sistemas automáticos de detección de defectos a otros tejidos industriales.

- ***Ensemble* de modelos especializados por tipo de tejido.** Entrenar múltiples CNN adaptadas a distintos tejidos y combinarlas mediante votación o ponderación (*ensemble*) para mejorar la generalización a nuevos tipos de tejido.

- **Clasificación multiclase.** Extensión del enfoque binario hacia modelos capaces de identificar tipologías concretas de defectos.

- **Localización precisa de defectos.** Desarrollo de modelos de segmentación que permitan identificar y delimitar defectos de forma precisa.
- **Validación en tiempo real.** Desarrollo de prototipos operativos en línea e integración de técnicas de *continual learning* que permitan adaptarse a variaciones en la producción.

Acrónimos y siglas

AE Autoencoder.

AU-PR Area Under the Precision-Recall curve.

AU-ROC Area Under the Receiver Operating Characteristic Curve.

BDI Beliefs Desires Intentions.

CNN Convolutional Neural Network.

CRediT Contributor Roles Taxonomy.

DA Data Augmentation.

DECENT Deep lEarning for automatiC tExtile iNspecTion.

DenseNet Densely Connected Convolutional Network.

DIH-LEAF Digital Innovation Hub on Livestock, Environment, Agriculture & Forest.

DL Deep Learning.

G-Mean Geometric Mean.

GICAP Grupo de Inteligencia Computacional APlicada.

IA Inteligencia Artificial.

IoT Internet of Things.

LSTM Long Short-Term Memory.

MLMA-Net Multilevel multi-attentional network.

PFD Patterned Fabric Dataset.

ResNet Residual Network.

RNN Redes Nueuronales Recurrentes.

SOCO Conference on Soft Computing Models in Industrial and Environmental Applications.

TS Textil Santanderina.

UBU Universidad de Burgos.

VGG Visual Geometry Group.

Xception Extreme Inception.

Bibliografía

[1] Biruk Gedif. Causes of Woven Fabric Defects and Cost Analysis: Case Study at Bahir Dar Textile Share Company (BDTSC). *Journal of Textile Science and Fashion Technology*, 7(3):1–7, 2020. URL https://doi.org/10.33552/JTSFT.2021.07.000664.

[2] Recep Ali Geze and Ayhan Akbaş. Detection and Classification of Fabric Defects Using Deep Learning Algorithms. *Politeknik Dergisi*, 27(1):371–378, 2023. URL https://doi.org/10.2339/politeknik.1386458.

[3] Textil Santanderina. Textil Santanderina. https://textilsantanderina.com/, 2025. Accedido: 10-12-2025.

[4] Walter A. Shewhart. *Economic Control of Quality of Manufactured Product*. D. Van Nostrand Company, New York, 1931.

[5] Douglas C Montgomery. *Introduction to Statistical Quality Control*. John wiley & sons, 2020.

[6] Thamiris de Souza Alves, Caterine Silva de Oliveira, Cesar Sanin, and Edward Szczerbicki. From knowledge based vision systems to cognitive vision systems: A review. *Procedia Computer Science*, 126:1855–1864, 2018. ISSN 1877-0509. URL https://doi.org/10.1016/j.procs.2018.08.077. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 22nd International Conference, KES-2018, Belgrade, Serbia.

[7] Yan Bei Liu, Zhi Tao Xiao, Fang Zhang, and Jun Wu. Fabric Defect Detection Method Based on Gabor Filters. *Advanced Materials Research*, 301:229–234, 2011. URL https://doi.org/10.4028/www.scientific.net/AMR.301-303.229.

[8] C Modrângă, R Brad, and R Brad. Fabric Defect Detection Using Fourier Transform and Gabor Filters. *J. Text. Eng. Fashion Technol.*, 3:00107, 2017.

[9] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by error propagation. In University of California, editor, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*, pages 318–362. MIT Press, 1986. URL https://doi.org/10.21236/ADA164453.

[10] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989. URL https://doi.org/10.1016/0893-6080(89)90020-8.

[11] Jeffrey L. Elman. Finding structure in time. *Cognitive Science*, 14(2):179–211, 1990. URL https://doi.org/10.1207/s15516709cog1402_1.

[12] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166, 1994. URL https://doi.org/10.1109/72.279181.

[13] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997. URL https://doi.org/10.1162/neco.1997.9.8.1735.

[14] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, 2006.

[15] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521 (7553):436–444, 2015. URL https://doi.org/10.1038/nature14539.

[16] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Cambridge, MA, 2016.

[17] Maheshwari S Biradar, BG Shiparamatti, and PM Patil. Fabric Defect Detection Using Deep Convolutional Neural Network. *Optical Memory and Neural Networks*, 30(3):250–256, 2021. URL https://doi.org/10.3103/S1060992X21030024.

[18] Bing Wei, Bailing Xu, Kuangrong Hao, and Lei Gao. Textile Defect Detection Using Multilevel and Attentional Deep Learning Network (MLMA-Net). *Textile Research Journal*, 92(19-20):3462–3477, 2022. URL https://doi.org/10.1177/00405175211073773.

[19] Xueshen Li and Yong Zhu. A Real-Time and Accurate Convolutional Neural Network for Fabric Defect Detection. *Complex & Intelligent Systems*, 10(3):3371–3387, 2024. URL https://doi.org/10.1007/s40747-023-01317-8.

[20] Ruimin Yang, Na Guo, Bo Tian, Junpu Wang, Shanliang Liu, and Miao Yu. Fabric Defect Detection via Saliency Model Based on Adjacent Context Coordination and Transformer. *Journal of Engineered Fibers and Fabrics*, 19:15589250241258272, 2024. URL https://doi.org/10.1177/15589250241258272.

[21] Lei Zhou, Bingya Ma, Yanyan Dong, Zhewen Yin, and Fan Lu. DCFE-YOLO: A Novel Fabric Defect Detection Method. *PLOS ONE*, 20:1–21, 01 2025. URL https://doi.org/10.1371/journal.pone.0314525.

[22] Imane Koulali and M. Taner Eskil. Unsupervised Textile Defect Detection Using Convolutional Neural Networks. *Applied Soft Computing*, 113:107913, 2021. URL https://doi.org/10.1016/j.asoc.2021.107913.

[23] Hongwei Zhang, Shuting Liu, Quanlu Tan, Shuai Lu, Le Yao, and Zhiqiang Ge. Colour-Patterned Fabric Defect Detection Based on an Unsupervised Multi-Scale U-Shaped Denoising Convolutional Autoencoder Model. *Coloration Technology*, 138(5):522–537, 2022. URL https://doi.org/10.1111/cote.12609.

[24] Hongwei Zhang, Wenbo Xiong, Shuai Lu, Mengdan Chen, and Le Yao. QA-USTNet: Yarn-Dyed Fabric Defect Detection via U-Shaped Swin Transformer Network Based on Quadtree Attention. *Textile Research Journal*, 93(15-16):3492–3508, 2023. URL https://doi.org/10.1177/00405175231158134.

[25] Jongwook Si and Sungyoung Kim. V-DAFT: Visual Technique for Texture Image Defect Recognition with Denoising Autoencoder and Fourier Transform. *Signal, Image and Video Processing*, 18(10):7405–7418, 2024. URL https://doi.org/10.1007/s11760-024-03403-x.

[26] Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://doi.org/10.48550/arXiv.1703.01780.

[27] Qihong Zhou, Jun Mei, Qian Zhang, Shaozong Wang, and Ge Chen. Semi-supervised fabric defect detection based on image reconstruction and density estimation. *Textile Research Journal*, 91(9-10):962–972, 2021.

[28] V. Hariharan and K. Sivaraman. Machine Learning Techniques for Quality Control in Textile Fabric Manufacturing. In *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, pages 902–907, 2024. URL https://doi.org/10.1109/ICSCSS60660.2024.10625436.

[29] Mariam Nasim, Rafia Mumtaz, Muneer Ahmad, and Arshad Ali. Fabric Defect Detection in Real World Manufacturing Using Deep Learning. *Information*, 15(8), 2024. URL https://doi.org/10.3390/info15080476.

[30] Rui Carrilho, Kailash A. Hambarde, and Hugo Proença. A Novel Dataset for Fabric Defect Detection: Bridging Gaps in Anomaly Detection. *Applied Sciences*, 14(12), 2024. URL https://doi.org/10.3390/app14125298.

[31] Chenkai Zhang, Shaozhe Feng, Xulongqi Wang, and Yueming Wang. Zju-leaper: A Benchmark Dataset for Fabric Defect Detection and a Comparative Study. *IEEE Transactions on Artificial Intelligence*, 1(3):219–232, 2021. URL https://doi.org/10.1109/TAI.2021.3057027.

[32] Tianchi. Smart Diagnosis of Cloth Flaw Dataset. https://tianchi.aliyun.com/dataset/79336, 2022. Visitado: 05-12-2025.

[33] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. MVTec AD–A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9592–9600, 2019.

[34] Irvin Andersen. TILDA Fabric Dataset v2, 2024. URL https://universe.roboflow.com/irvin-andersen/tilda-fabric/dataset/2. visitado: 05-12-2025.

[35] Priemsh Pathirana. Fabric Stain Dataset. https://www.kaggle.com/datasets/priemshpathirana/fabric-stain-dataset. visitado: 05-12-2025.

[36] Henry Y. T. Ngan. NKU Fabric Defect Dataset. URL https://ytngan.wordpress.com/codes/. visitado: 05-12-2025.

[37] Javier Silvestre-Blanes, Teresa Albero-Albero, Ignacio Miralles, Rubén Pérez-Llorens, and Jorge Moreno. A public fabric database for defect detection methods and results. *AUTEX Research Journal*, 19(4):363–371, 2019. URL https://doi.org/10.2478/aut-2019-0035.

[38] Industrial Automation Research Laboratory. Patterned Fabric Database, 2009. URL https://ytngan.wordpress.com/codes/. Visitado: 09-12-2025.

[39] C. E. Belman-Lopez, J. A. Jiménez-García, and S. Hernández-González. Análisis Exhaustivo de los Principios de Diseño en el Contexto de Industria 4.0. *Revista Iberoamericana de Automática e Informática Industrial*, 17(4):432–447, sep 2020. URL https://doi.org/10.4995/riai.2020.12579.

[40] GICAP — Grupo de Investigación Applied Computational Intelligence, Universidad de Burgos. Applied Computational Intelligence (GICAP) — Universidad de Burgos. https://investigacion.ubu.es/grupos/1830/detalle, 2025. Visitado: 10-12-2025.

[41] Universidad de Burgos – GICAP. Comienzo del Proyecto DECENT: Deep Learning Aplicado al Control de Calidad en el Sector Textil. https://www.ubu.es/noticias/comienzo-del-proyecto-decent-deep-learning-aplicado-al-control-de-calidad-en-el-sector-textil, 2025. Accedido: 10-12-2025.

[42] DIH-LEAF. DIH-LEAF – Digital Innovation Hub on Livestock, Environment, Agriculture & Forest. https://www.dih-leaf.eu/, 2025. Accedido: 10-12-2025.

[43] Marina Sokolova and Guy Lapalme. A Systematic Analysis of Performance Measures for Classification Tasks. *Information Processing & Management*, 45(4):427–437, 2009. URL https://doi.org/10.1016/j.ipm.2009.03.002.

[44] Tom Fawcett. An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27 (8):861–874, 2006. URL https://doi.org/10.1016/j.patrec.2005.10.010.

[45] Beatriz Gil-Arroyo, Juan Marcos Sanz, Ángel Arroyo, Daniel Urda, Nuño Basurto, and Álvaro Herrero. Dataset for Defect Detection in Textile Manufacturing. *Data in Brief*, 59:111451, 2025. URL https://doi.org/10.1016/j.dib.2025.111451.

[46] Beatriz Gil-Arroyo, Marco Melgarejo, Abraham Casas, Alejandro López, Juan Marcos Sanz, and Daniel Urda. Deep Learning for Quality Control in Woven Fabrics. *DYNA Ingeniería e Industria*, 100:545–551, 2025. URL https://doi.org/10.52152/D11517.

[47] Nuria Velasco-Pérez, Samuel Lozano-Juárez, Beatriz Gil-Arroyo, Juan Marcos Sanz, Nuño Basurto, Daniel Urda, and Álvaro Herrero. Defect Detection in Batavia Woven Fabrics by Means of Convolutional Neural Networks. In *18th International Conference on Soft Computing Models in Industrial and Environmental Applications*, pages 205–215, Cham, Switzerland, 2023. Springer Nature Switzerland. ISBN 978-3-031-42536-3. URL https://doi.org/10.1007/978-3-031-42536-3_20.

[48] Beatriz Gil-Arroyo, Nuria Velasco-Pérez, Nuño Basurto, Juan Marcos Sanz, Daniel Urda, and Álvaro Herrero. Original and Processed Dataset of Batavia and Sarga Woven Fabric Images. UBU Repository, v1, 2025. URL https://doi.org/10.36443/10259/9965.