

Modelo basado en YOLOv8 para la detección automática de daños en tejados residenciales.

A. S. Silva^{1*} , A. R. Azevedo² , F. H. A. M. Neto² , P. H. Ferreira² 

*Autor de Contacto: so_alisson@hotmail.com

DOI: <https://doi.org/10.21041/ra.v15i1.783>

Recibido: 17/09/2024 | Correcciones recibidas: 19/11/2024 | Aceptado: 10/12/2024 | Publicado: 01/01/2025

RESUMEN

Este estudio desarrolló un modelo automatizado de reconocimiento de imágenes para la inspección de techos residenciales utilizando la arquitectura YOLOv8, con el objetivo de identificar tres tipos de daños. La metodología incluyó imágenes de 167 edificios, capturadas por drones y anotadas en CVAT, que se utilizaron para entrenar y probar el modelo. YOLOv8 se aplicó para la detección y clasificación de anomalías, logrando un 79% de precisión. Las principales limitaciones fueron el pequeño tamaño del conjunto de datos y la limitada variedad de ángulos de captura. La originalidad del trabajo radica en el uso innovador de YOLOv8 para la inspección de techos. Investigaciones futuras se centrarán en el desarrollo de las arquitecturas YOLOv9 y YOLOv10, además de ampliar el conjunto de datos y las clases de daños.

Palabras clave: gestión de mantenimiento; aprendizaje automático; You Only Look Once (YOLO); inspección de techos; evaluación de techos.

Citar como: Silva, A. S., Azevedo, A. R., Neto, F. H. A. M., Ferreira, P. H. (2025), “Modelo basado en YOLOv8 para la detección automática de daños en tejados residenciales.”, Revista ALCONPAT, 15 (1), pp. 50 – 63, DOI: <https://doi.org/10.21041/ra.v15i1.783>

¹ Departamento de Ingeniería Estructural y de la Construcción, Universidad Federal de Bahía, Salvador, Brasil.

² Departamento de Matemáticas y Estadística, Universidad Federal de Bahía, Salvador, Brasil.

Contribución de cada autor

A.S.S: conceptualización, curación de datos, análisis formal, adquisición de financiamiento, investigación, metodología, administración del proyecto, supervisión, validación, visualización, redacción del borrador original, revisión y edición. A.R.A: curación de datos, análisis formal, investigación, metodología, validación, visualización. F.H.A.M.N: curación de datos, análisis formal, investigación, metodología, validación, programación, visualización, redacción del borrador original, revisión y edición. P.H.F: curación de datos, análisis formal, investigación, metodología, programación, validación, visualización, redacción del borrador original, revisión y edición.

Licencia Creative Commons

Los derechos de autor (2025) son propiedad de los autores. Este trabajo es un artículo de acceso abierto publicado bajo los términos y condiciones de una licencia internacional Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

Discusiones y correcciones posteriores a la publicación

Cualquier discusión, incluyendo la réplica de los autores, se publicará en el tercer número del año 2025 siempre y cuando la información se reciba antes del cierre del segundo número del año 2025.

YOLOv8-based model for automatic detection of residential roof damage.

ABSTRACT

This study developed an automated image recognition model for inspecting residential roofs using the YOLOv8 architecture to identify three types of damage. The methodology involved images from 167 buildings captured by drones and annotated in CVAT, which were used to train and test the model. YOLOv8 was applied for anomaly detection and classification, achieving 79% precision. The limitations were the small dataset and the limited variety of capture angles. The originality of the work lies in the innovative use of YOLOv8 for roof inspection. Future research will focus on developing the YOLOv9 and YOLOv10 architectures and expanding the dataset and damage classes.

Keywords: maintenance management; machine learning; You Only Look Once (YOLO); roof inspection; rooftop assessment.

Modelo baseado em YOLOv8 para detecção automática de danos em telhados residenciais.

RESUMO

Este estudo desenvolveu um modelo automatizado de reconhecimento de imagens para inspeção de telhados residenciais, utilizando a arquitetura YOLOv8, com o objetivo de identificar três tipos de danos. A metodologia contou com imagens de 167 edifícios, capturadas por drones e anotadas no CVAT, utilizadas para treinar e testar o modelo. A YOLOv8 foi aplicada na detecção e classificação de anomalias, alcançando 79% de precisão. As principais limitações foram o tamanho reduzido do conjunto de dados e a pouca variedade de ângulos de captura. A originalidade do trabalho está no uso inovador da YOLOv8 para a inspeção de telhados. Pesquisas futuras incluirão o desenvolvimento das arquiteturas YOLOv9 e YOLOv10, bem como a ampliação do conjunto de dados e das classes de problemas.

Keywords: gestão da manutenção; aprendizado de máquina; You Only Look Once (YOLO); inspeção de telhado; avaliação de telhado.

Información Legal

Revista ALCONPAT es una publicación cuatrimestral de la Asociación Latinoamericana de Control de Calidad, Patología y Recuperación de la Construcción, Internacional, A. C., Km. 6, antigua carretera a Progreso, Mérida, Yucatán, C.P. 97310, Tel.5219997385893, alconpat.int@gmail.com, Página Web: www.alconpat.org

Reserva de derechos al uso exclusivo No.04-2013-011717330300-203, eISSN 2007-6835, ambos otorgados por el Instituto Nacional de Derecho de Autor. Editor responsable: Dr. Pedro Castro Borges. Responsable de la última actualización de este número, Unidad de Informática ALCONPAT, Ing. Elizabeth Sabido Maldonado.

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor.

La reproducción total o parcial de los contenidos e imágenes de la publicación se realiza en apego al código COPE y a la licencia CC BY 4.0 de la Revista ALCONPAT.

1. INTRODUCCIÓN

La envolvente del edificio juega un papel fundamental en el rendimiento general de una estructura, actuando como una capa protectora que protege al edificio de la degradación ambiental externa (Silva et al., 2015). Los techos y revestimientos protegen principalmente a los edificios, creando ambientes interiores protegidos de agentes externos como temperatura, humedad, lluvia, viento y ruido (Garcez et al., 2012; Conceição et al., 2017). El mantenimiento de este sistema es un componente crucial en los programas de mantenimiento de edificios, ya que el sistema de techo es la parte más vulnerable de los edificios debido a su exposición directa al medio ambiente (Rosenberger, 2012). Cuando se exponen a condiciones climáticas severas, estos sistemas a menudo experimentan un deterioro acelerado, lo que genera un mayor consumo de energía, compromete la integridad estructural del techo y otros componentes del edificio y reduce la vida útil del edificio (Alashari et al., 2022). Por tanto, las evaluaciones periódicas son esenciales, ya que pueden conducir a reducciones sustanciales de costes (Liu et al., 2020; Alzarrad et al., 2022).

La inspección manual es el método principal para evaluar periódicamente los requisitos estructurales y funcionales para garantizar el cumplimiento de los estándares básicos de desempeño (Dang et al., 2021). Los métodos de inspección tradicionales dependen del trabajo humano para identificar problemas y defectos. Sin embargo, estas tareas suelen ser tediosas y agotadoras, lo que da lugar a errores debidos a factores humanos como falta de atención, ilusiones ópticas, evaluaciones subjetivas y fatiga de los inspectores (Radopoulou y Brilakis, 2017; Hoang et al., 2018). Además, los inspectores humanos requieren formación y experiencia, lo que requiere mucho tiempo. La rotación de empleados es otro inconveniente importante del trabajo manual (Olimov et al., 2022). Más allá de estos desafíos, dos factores adicionales complican las evaluaciones periódicas: el riesgo de caerse del techo y los costos adicionales derivados de la falta de precisión y el error humano (Brown et al., 2021).

Con base en estos hallazgos, podemos clasificar las limitaciones de las inspecciones de techos en dos áreas principales: adquisición de datos y análisis de datos. Para abordar la primera limitación relacionada con la adquisición de datos, destacan los drones, también conocidos como vehículos aéreos no tripulados (UAV). Los drones son muy flexibles, con movimientos versátiles, lo que permite una exploración rápida de entornos construidos y requieren costos laborales mínimos, ya que solo necesitan cambios de batería y rutas planificadas previamente que eliminan la necesidad de operaciones manuales durante el vuelo (Chu et al., 2024). Muchos estudios han utilizado esta tecnología para inspecciones de tejados (Gajjar & Burgett, 2020; Silveira et al., 2020; Staffa et al., 2023; Santos et al., 2023).

En cuanto al análisis de datos, los algoritmos automatizados de Inteligencia Artificial (IA) han mostrado resultados prometedores para superar las limitaciones de las evaluaciones subjetivas realizadas por inspectores individuales (Alizadeh & Ma, 2021; Avola et al., 2022). Trabajos notables en esta área incluyen estudios de Alzarrad et al. (2022), Hou et al. (2023), y Wang et al. (2023). La combinación de estas dos estrategias puede mejorar significativamente los resultados de la inspección, permitiendo inspecciones periódicas y abordando las limitaciones mencionadas anteriormente, como lo demuestra la investigación de Mostafa et al. (2023), Santos et al. (2023) y Staffa et al. (2023).

Sin embargo, a pesar de la relevancia de estos estudios, persisten algunas limitaciones. Mostafa et al. (2023) desarrollaron un modelo de red neuronal convolucional (CNN) de dos etapas para la detección y clasificación de defectos en tejados utilizando imágenes 2D de más de 5500 imágenes de 21 tejados. El modelo logró una precisión del 95% en la detección y del 97% en la clasificación. Sin embargo, la ausencia de defectos como grietas y ampollas en las imágenes de entrenamiento limita su eficacia, lo que pone de relieve la necesidad de realizar más investigaciones y mejoras. En el estudio de Santos et al. (2023), los autores optimizaron el mantenimiento de techos planos

utilizando 330 imágenes capturadas por drones y probaron seis algoritmos de aprendizaje profundo para la detección de objetos, incluida CNN basada en regiones (Faster R-CNN), RetinaNet y Selección de muestras de entrenamiento adaptativo (ATSS). R-CNN más rápido logró los mejores resultados con un 93% de recuperación y puntuación F1, mientras que ATSS alcanzó un 95% de precisión. El análisis de errores destacó la complejidad de los tejados planos, sugiriendo la necesidad de mejorar los algoritmos para manejar sus variaciones. Aunque el estudio detectó eficazmente unidades condensadoras, es necesario ampliar el enfoque a otros problemas, como baldosas adicionales y oxidación en canalones, que también requieren inspección.

Para abordar estas lagunas de investigación, este trabajo propone un modelo de reconocimiento de imágenes para techos residenciales. Se utilizó un conjunto de datos completo de imágenes de 167 edificios residenciales recopiladas por vehículos aéreos no tripulados para detectar y clasificar tres tipos de daños en el tejado: tejas adicionales en el tejado, oxidación en canalones y tapas de trampillas abiertas. Las anotaciones se realizaron utilizando CVAT y el estudio empleó la octava generación de la arquitectura You Only Look Once (YOLO), YOLOv8.

2. APLICACIONES DE YOLO EN CONSTRUCCIÓN

En visión por computadora, la desafiante tarea de la detección de objetos ha experimentado una evolución significativa, impulsada por la revolucionaria introducción de la arquitectura YOLO en 2016 (Redmon et al., 2016). El enfoque innovador de YOLO se diferenciaba de las arquitecturas tradicionales de detección de objetos en dos etapas al proponer una arquitectura unificada capaz de predecir simultáneamente cuadros delimitadores y probabilidades de clase, satisfaciendo las demandas de procesamiento en tiempo real (Koubaa et al., 2021). YOLOv8 es una de las últimas versiones de la serie YOLO, que utiliza una arquitectura de red que incorpora una red piramidal de funciones (FPN) y conexiones de capas para integrar información de funciones en múltiples escalas de manera eficiente. Esta versión combina mecanismos de atención y estrategias de optimización para mejorar la precisión y el rendimiento en la detección de objetos. La estructura central incluye una red troncal para la extracción de características de imágenes, típicamente usando CNN profunda como Darknet o ResNet, y un cabezal de detección compuesto por capas convolucionales y completamente conectadas responsables de predecir el cuadro delimitador y la probabilidad de clase de los objetos (Cao et al., 2024).

En este contexto, importantes estudios en construcción han explorado las capacidades avanzadas de YOLOv8. Sohaib et al. (2024) desarrollaron un modelo para detectar y segmentar grietas en concreto, mientras que Han et al. (2024) automatizaron la detección y clasificación de pequeñas herramientas en obras de construcción. Zhao y cols. (2024) crearon un sistema para el seguimiento automatizado de defectos en hormigón impreso en 3D. Jiang et al. (2024) propusieron un método para optimizar la sincronización de la vibración del hormigón con robots, utilizando YOLOv8 para identificar las mejores posiciones de vibración y eliminar la interferencia visual de las barras de refuerzo. Ribeiro et al. (2024) propusieron un método basado en YOLOv8 para detectar fisuras en edificios. Las imágenes utilizadas para la capacitación se recopilaron de Internet y el modelo logró una precisión del 80 % en la validación con imágenes no incluidas en la capacitación. Sin embargo, el modelo se centró únicamente en fisuras, lo que limita su aplicabilidad a otros casos de uso.

En un contexto más específico, para evaluar cubiertas y cubiertas de edificios, Hou et al. (2023) desarrollaron un algoritmo de detección mejorado basado en YOLOv3 para identificar elementos decorativos conocidos como "bestias de cresta" en los techos de edificios antiguos chinos. Esta tecnología es crucial para la reconstrucción 3D de edificios históricos, datación histórica y servicios culturales y turísticos. Mejoraron la red YOLOv3 incorporando características locales y convoluciones separables profundas para expandir el campo receptivo y capturar características más diversas. Además, optimizaron la estructura residual de la red de extracción de características

y rediseñaron el cabezal de predicción del modelo introduciendo módulos de extrusión y excitación y utilizando convoluciones de varios tamaños. Los experimentos demostraron que el algoritmo mejorado logró una precisión promedio (mAP) media del 86,48 %, un 3,05 % más que el YOLOv3 original, al tiempo que redujo los parámetros del modelo en un 70 % y mejoró el rendimiento de detección. Aunque esta investigación se centró en detectar elementos decorativos, el modelo también podría aplicarse para identificar problemas comunes en edificios antiguos, como tejas rotas o faltantes, así como musgo y oxidación en elementos de tejados, que son esenciales para el correcto mantenimiento de estos históricos. estructuras.

Además de este estudio, Alzarrad et al. (2022) desarrollaron un método que utiliza el aprendizaje profundo para la evaluación automatizada de techos residenciales, centrándose en identificar áreas dañadas, como tejas faltantes. Emplearon una metodología de aprendizaje por transferencia utilizando YOLOv5 para procesar imágenes e identificar estos daños. El estudio encontró que a pesar de los datos limitados, el modelo logró aproximadamente un 81% de precisión y un 86% de exactitud. Sin embargo, si bien el estudio mostró buenos resultados en la detección de problemas con las cubiertas de techos residenciales, se centró sólo en un tipo de problema. En la práctica, es necesario identificar muchos otros problemas con los techos residenciales durante las inspecciones para su resolución adecuada, como la oxidación en canalones y tapajuntas. Descuidar estos problemas puede provocar un deterioro estructural, aumentando significativamente los costos de reparación.

Qiu et al. (2024) propusieron un método de dos etapas basado en YOLOv8. En la primera etapa, se utiliza el modelo YOLOv8 para identificar edificios de importancia histórica en el área, logrando una tasa de mAP del 97,2%. El mismo modelo se emplea en la segunda etapa para segmentar las imágenes seleccionadas, detectando posibles defectos del techo, como colapsos, tejas faltantes, adiciones arquitectónicas inapropiadas e invasión de vegetación. Aunque esta investigación representa un avance significativo y utiliza una de las últimas versiones de YOLO, todavía se centra en los techos de edificios culturalmente históricos, como los templos budistas, que generalmente tienen sistemas de techos específicos diferentes a los utilizados en los edificios residenciales modernos.

Teniendo esto en cuenta, está claro que la detección y clasificación inmediata de defectos y problemas en los tejados de edificios residenciales son cruciales. Si bien los estudios mencionados han avanzado significativamente en la automatización de la detección de ciertos problemas en los techos, aún se necesita más investigación para abordar otros tipos de problemas. Teniendo en cuenta los avances en las arquitecturas YOLO, existe una clara oportunidad de aplicar estas redes en este contexto. En este estudio, evaluamos el uso de YOLOv8 para detectar y clasificar tres problemas en los techos de edificios residenciales. Como no encontramos estudios previos en la literatura que aborden este tema, destacamos la aplicación de una de las últimas versiones de YOLO para automatizar el proceso de inspección. Además, proporcionamos los códigos utilizados en esta investigación para permitir que otros investigadores mejoren el modelo y agreguen más clases de problemas, ya que también es necesario identificar y abordar diferentes tipos de defectos con igual importancia.

3. MÉTODO DE INVESTIGACIÓN

Este estudio adoptó una estrategia de investigación de simulación experimental, desarrollada en cuatro etapas (Figura 1): (1) refinar la base de datos de imágenes de techos residenciales con daños superficiales; (2) crear el modelo de reconocimiento automatizado de daños utilizando la arquitectura YOLOv8; (3) entrenar el modelo utilizando el 70% de la base de datos; y (4) probar y validar el modelo entrenado con nuevas imágenes (30%). Los indicadores de precisión, recuperación y mAP evaluaron el desempeño del modelo.

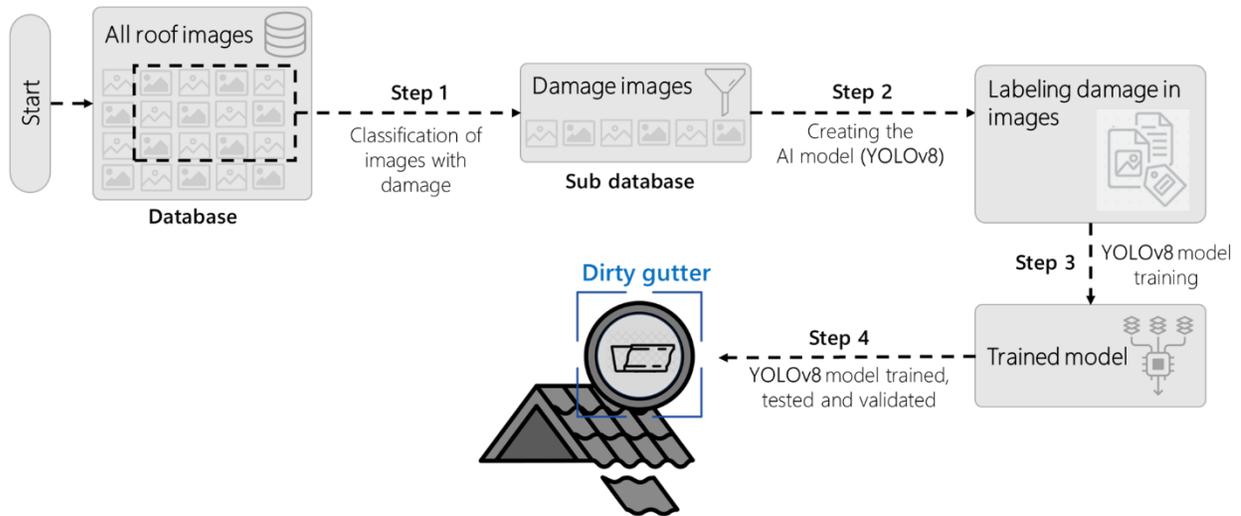


Figura 1. Diseño de investigación.

En este trabajo, utilizamos YOLOv8. Según Lalinia y Sahafi (2024), una de las principales ventajas de incorporar YOLOv8 en proyectos de visión por computadora es su mayor precisión en comparación con los modelos YOLO anteriores. YOLOv8 admite múltiples tareas, como detección de objetos, segmentación y clasificación de imágenes, lo que mejora su versatilidad para diversas aplicaciones. La arquitectura YOLOv8 se muestra en la Figura 2; Se pueden encontrar más detalles sobre la arquitectura en Terven et al. (2023).

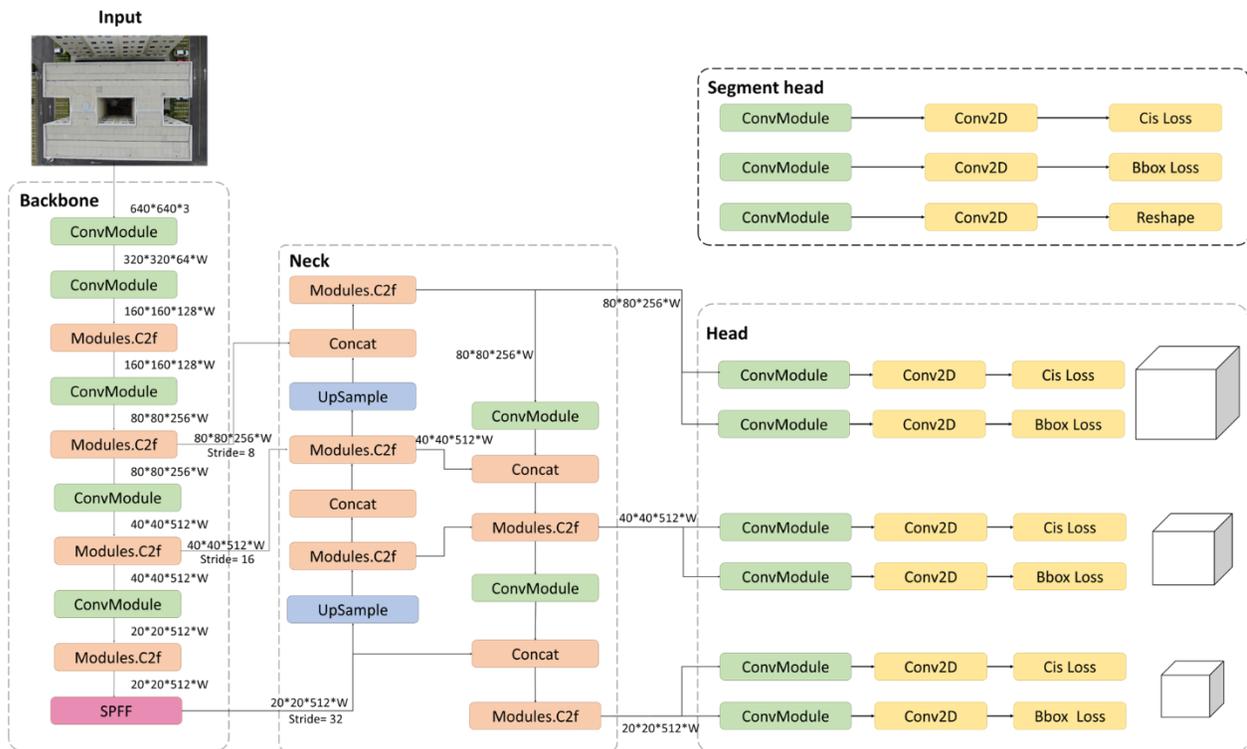


Figure 2. YOLOv8 architecture used.

La Figura 2 muestra que la arquitectura está organizada en tres componentes principales: columna vertebral, cuello y cabeza, cada uno de los cuales desempeña un papel específico en el procesamiento de imágenes y la generación de predicciones.

La imagen de entrada es el punto de partida del proceso, donde la red comienza a extraer características relevantes para identificar posibles defectos. Backbone extrae características básicas de la imagen a través de varias capas convolucionales que detectan patrones iniciales, como bordes y texturas. A medida que la imagen avanza a través de estas capas, módulos más complejos, conocidos como *Modules.C2f*, refinan las características extraídas, lo que permite un análisis de la imagen más detallado. Backbone también incluye el módulo *Spatial Pyramid Pooling Fast (SPPF)*, que agrega información a diferentes escalas y mejora la solidez de la detección.

El Cuello es un intermediario entre la Columna Vertebral y la Cabeza, refinando y combinando las características extraídas mediante Concatenación (Concat), Upsampling y capas convolucionales adicionales. El objetivo de Neck es preparar estas características para una detección multiescala más precisa, asegurando que en la predicción final se tengan en cuenta tanto los detalles finos como las estructuras más grandes de la imagen.

Finalmente, el Head es el responsable de generar las predicciones finales. Los rasgos refinados por el Cuello se utilizan para identificar y clasificar objetos o defectos en la imagen. El cabezal consta de dos tipos principales: el cabezal de segmento, que maneja tareas de segmentación identificando áreas exactas donde hay defectos, y el cabezal de detección, que se centra en la detección y clasificación de objetos generando cuadros delimitadores y clases de objetos. Además, las pérdidas (Cis Loss y Bbox Loss) asociadas con estas predicciones se calculan para optimizar el rendimiento del modelo.

Esta arquitectura permite a YOLOv8 procesar imágenes de manera eficiente, desde la extracción de características iniciales hasta la generación de predicciones finales, lo que garantiza una detección precisa de defectos en techos residenciales.

3.1 Base de datos

GETEC-UFBA ("Grupo de Pesquisa e Extensão em Gestão e Tecnologia das Construções da Escola Politécnica da Universidade Federal da Bahia", en portugués) proporcionó el conjunto de datos utilizado en esta investigación. Estas imágenes fueron recopiladas de edificios residenciales con edades comprendidas entre uno y cinco años. Se inspeccionaron más de 167 tejados residenciales mediante vuelos manuales de drones, con una altitud aproximada de 5 metros sobre el edificio analizado. La adquisición de datos siguió los protocolos de inspección propuestos por Silveira et al. (2021). Después de analizar las imágenes en este conjunto de datos, se seleccionaron 394 imágenes, distribuidas entre 425 anotaciones para la clase "Presencia de teja extra en la dirección transversal del techo", 473 anotaciones para la clase "Integridad del canal (oxidación)" y 29 anotaciones para la clase "Abrir trampilla". El conjunto de datos se dividió en aproximadamente un 70 % para capacitación, un 20 % para validación y un 10 % para pruebas. Después del aumento de datos en el conjunto de entrenamiento, se obtuvieron 498 imágenes del conjunto de entrenamiento, y en la Figura 3 se muestran ejemplos de estas clases.



Figura 3. Ejemplos de daños en el techo abordados en esta investigación: a) Presencia de teja extra en la dirección transversal del techo; b) Integridad del canal (oxidación); c) Abrir trampilla.

3.2 Creando el modelo YOLOv8

Después de definir las anomalías, las imágenes de entrenamiento se etiquetaron utilizando Roboflow. Un ejemplo de este etiquetado se muestra en la Figura 4a, donde se etiquetan algunas tejas adicionales en el techo, oxidación y un tragaluz abierto. La Figura 4b ilustra el flujo de trabajo completo desde la entrada de datos hasta el entrenamiento del modelo y el posterior proceso de inferencia, donde se realizan las anotaciones; luego, el preprocesamiento, el aumento de datos y, finalmente, los datos se exportan en formato YOLOv8 para su análisis en Google Colab.

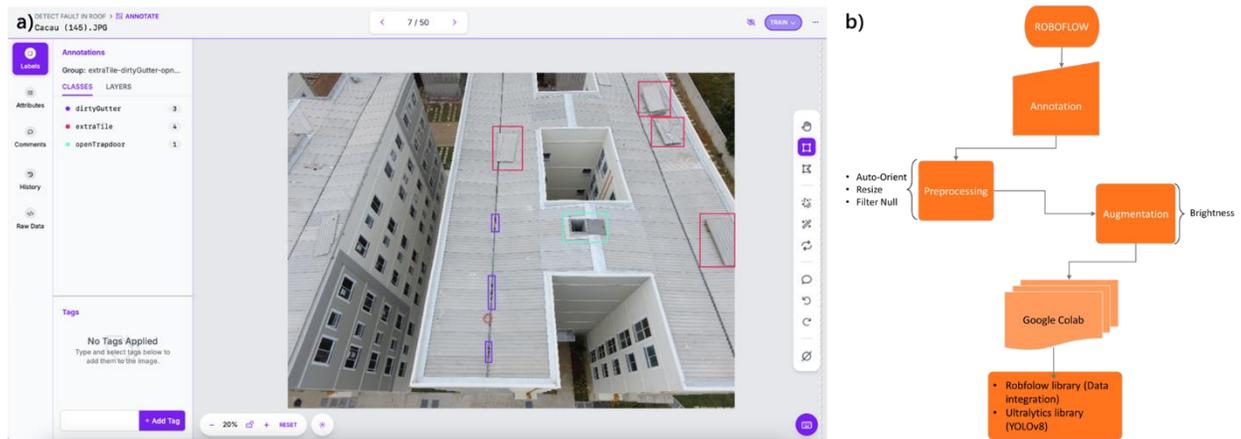


Figura 4. a) Notas de daños; segundo) flujo de trabajo

3.3 Entrenamiento, prueba y validación del modelo

Precision, Recall y mAP son los métodos de evaluación utilizados para evaluar los modelos creados en las plataformas Custom Vision y Roboflow. La precisión representa el número de casos correctos de todos los casos disponibles. La alta precisión indica que las anomalías se están identificando con precisión. En consecuencia, recordar es el número de casos correctos de todos los casos positivos. Un modelo con alta recuperación identifica efectivamente la mayor cantidad de anomalías correctas. De manera similar, mAP es la media de la precisión promedio, que es la precisión obtenida en varios niveles de recuperación en todas las clases (anomalías). Las siguientes fórmulas representan los indicadores clave de rendimiento (KPI):

$$(1) \quad Precision = \frac{TP}{TP + FP}$$

$$(2) \quad Recall = \frac{TP}{TP + FN}$$

$$(3) \quad mAP = \frac{1}{n} \sum_{k=1}^n AP_k$$

Donde TP (verdaderos positivos) son aciertos en instancias positivas, FN (falsos negativos) son errores en instancias positivas, TN (verdaderos negativos) son aciertos en instancias negativas y FP (falsos positivos) son errores en instancias negativas y AP_k es el promedio. Precisión para cada clase k. Estas métricas proporcionan información sobre la clasificación y la calidad del modelo.

4. RESULTADOS Y DISCUSIÓN

Durante el entrenamiento del modelo, la arquitectura YOLOv8 realiza activaciones de funciones en diferentes etapas de la red neuronal. Estas activaciones demuestran cómo la red procesa información visual para detectar patrones y características importantes en las imágenes de entrada (Figura 5). Estas visualizaciones ilustran cómo YOLOv8 realiza la detección de objetos, desde la extracción de características básicas hasta la combinación de información en diferentes escalas y resoluciones, lo que permite una identificación precisa de los defectos del techo.

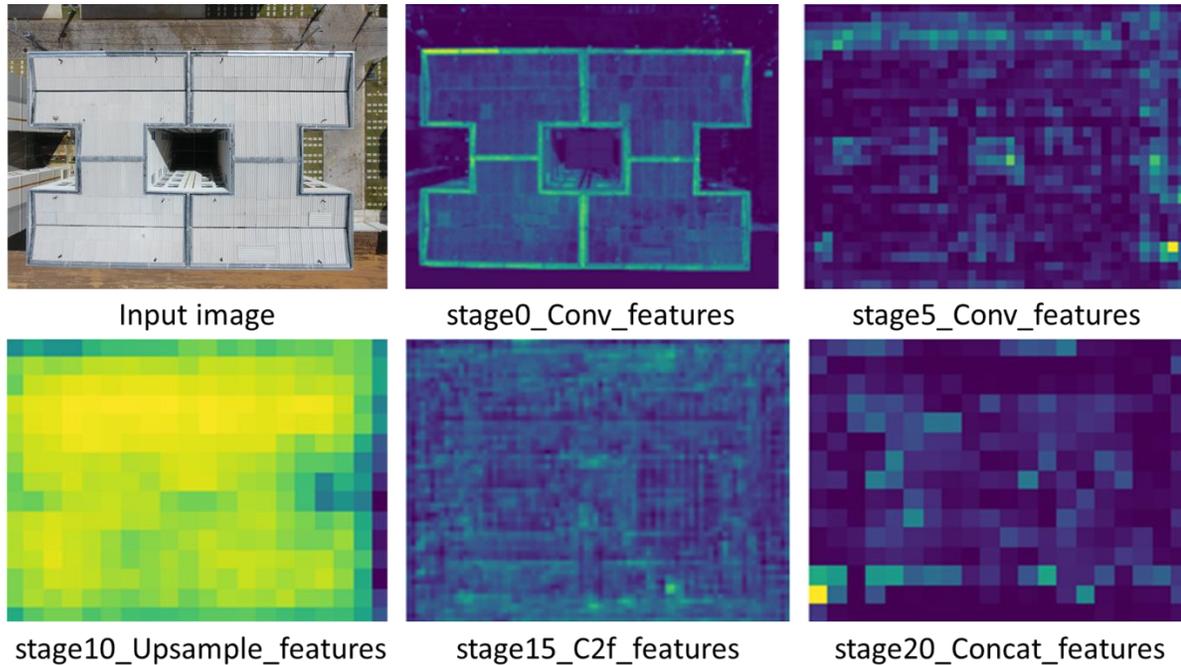


Figura 5. Estructura de las capas YOLOv8 entrenadas.

Las "Conv_features" son las salidas de las capas convolucionales, responsables de identificar patrones locales como bordes y texturas. A medida que avanzamos por las diferentes etapas, desde "stage0_Conv_features" hasta "stage20_Concat_features", las activaciones reflejan representaciones cada vez más complejas y abstractas de las características de la imagen. Las "Upsample_features" están relacionadas con el proceso de upsampling, donde se aumenta la resolución espacial de las activaciones, permitiendo a la red detectar objetos a diferentes escalas. Este proceso es crucial para detectar detalles más pequeños en los tejados.

Las "C2f_features" se refieren a combinaciones de convoluciones seguidas de funciones de activación diseñadas para capturar interacciones más complejas entre características extraídas en capas anteriores, mejorando la capacidad de la red para identificar patrones. Finalmente, las "Concat_features" resultan de operaciones de concatenación, que combinan activaciones de diferentes capas o etapas de la red. Esta técnica permite a la red integrar información de otros niveles de abstracción, fusionando características extraídas en varias etapas para producir la predicción final.

El modelo YOLOv8 se entrenó utilizando Google Colab. La capacitación del modelo y el código de desarrollo de la aplicación Streamlit están disponibles en: (OMITIDO PARA REVISIÓN). El modelo entrenado tiene 168 capas y aproximadamente 3 millones de parámetros. Las imágenes de entrada para el modelo eran de 640×640 píxeles. Las métricas de rendimiento de YOLOv8 en el conjunto de prueba se presentan en la Tabla 1, donde mAP50 representa el mAP calculado con un umbral de Intersección sobre Unión (IoU) de 0,50.

Tabla 1. Métricas de evaluación del modelo YOLOv8 para el conjunto de prueba.

Class	Images	Precision (%)	Recall (%)	mAP50 (%)
All	73	0.79	0.59	0.66
oxidación	73	0.69	0.44	0.51
teja extra	73	0.78	0.66	0.68
abrir trampilla	73	0.90	0.67	0.78

Para tejas adicionales en el techo, el modelo logró una precisión del 78 %, una recuperación del 66 % y mAP50 del 68 %. El modelo identifica efectivamente mosaicos adicionales, aunque es ligeramente menos preciso que detectar tragaluces abiertos. Finalmente, para la integridad de los canales, el modelo tuvo el peor desempeño, con una precisión del 69 %, una recuperación del 44 % y mAP50 del 51 %, lo que indica una mayor dificultad para detectar y clasificar con precisión este problema. La Figura 6 muestra los gráficos de pérdida de entrenamiento y validación, junto con las métricas de Precisión y Recuperación; estos indican que el modelo está aprendiendo de manera sólida y constante (con pérdidas decrecientes y métricas crecientes).

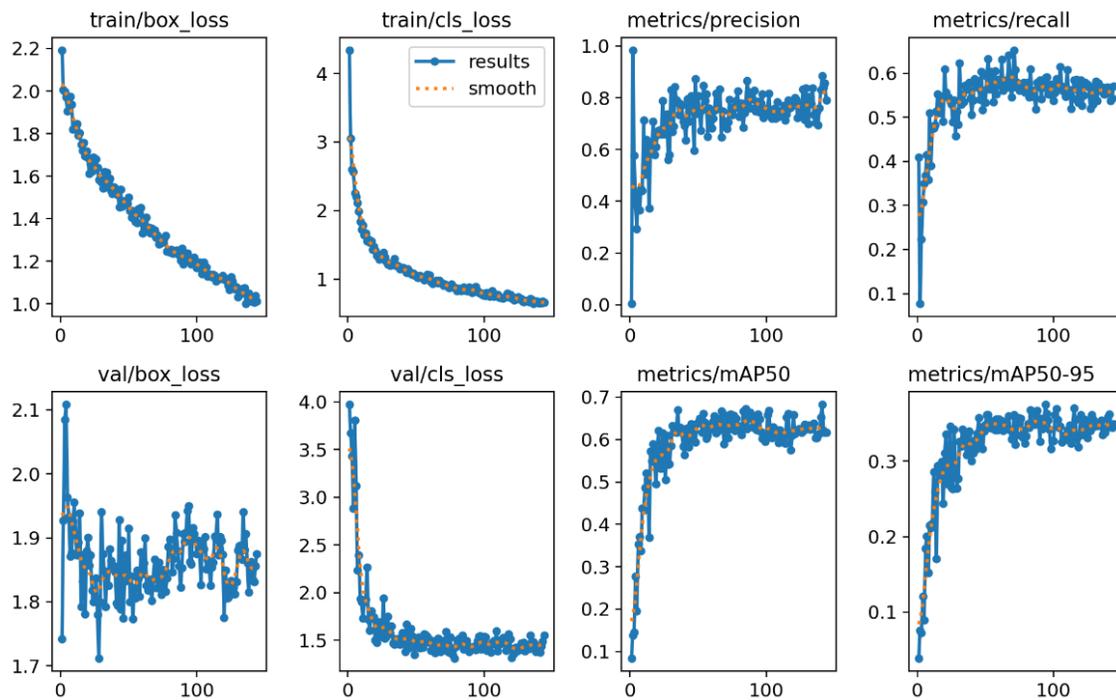


Figura 6. Pérdidas de entrenamiento y validación, métricas de precisión y recuperación de YOLOv8.

Las Figuras 7 y 8 muestran los resultados de clasificación del conjunto de prueba. A partir de estas imágenes, queda claro que el modelo puede identificar tejas adicionales dispuestas en diferentes orientaciones (la Figura 7a muestra tejas adicionales; la Figura 7b muestra una teja adicional y oxidación en el canalón) y tragaluces abiertos en varias posiciones (las Figuras 8a y 8b muestran una claraboya abierta y algunos azulejos extra). Estas clasificaciones son vitales para el mantenimiento preventivo, asegurando la integridad estructural y la adecuada funcionalidad de las cubiertas. A pesar de los éxitos del modelo, algunos elementos no fueron identificados en las Figuras 7 y 8.

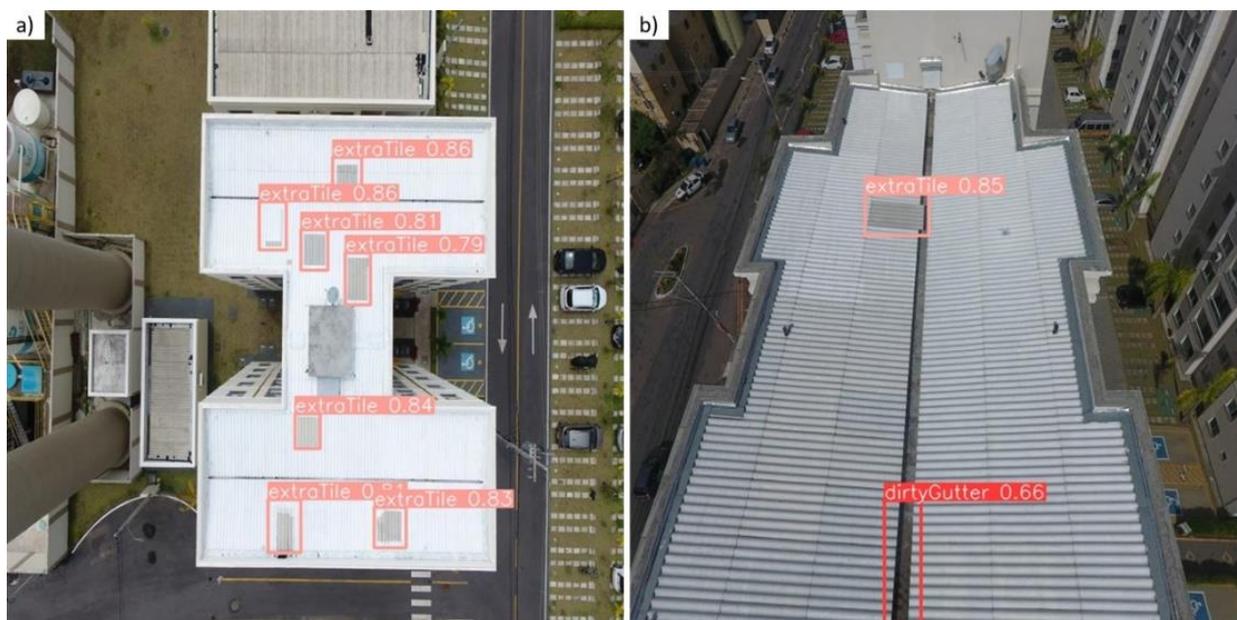


Figura 7. Clasificación del conjunto de pruebas para teja extra y oxidación.

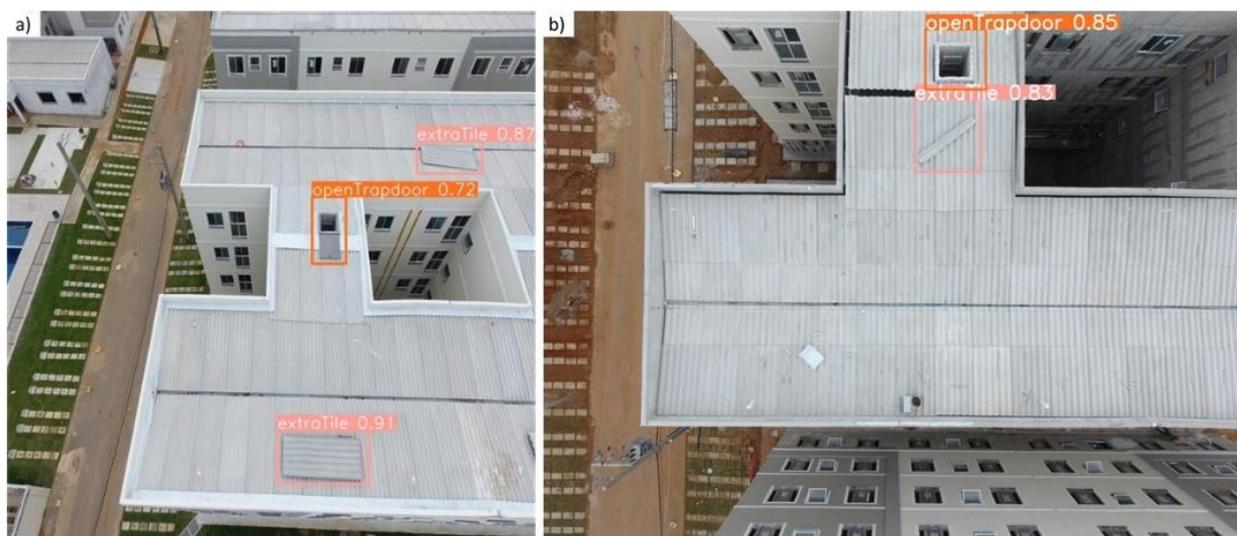


Figura 8. Clasificación del conjunto de prueba para trampilla abierta y teja extra.

La revisión de la literatura sobre la detección y clasificación de defectos en cubiertas inclinadas de edificios residenciales destacó la necesidad de estudios centrados en el uso de algoritmos YOLOv8, dados los resultados prometedores de versiones anteriores para propósitos similares (Alzarrad et al., 2022; Hou et al., 2023; Wang y otros, 2023; Este estudio abordó esta brecha desarrollando un modelo de reconocimiento automatizado para tres defectos en techos residenciales utilizando imágenes recopiladas por drones de 167 techos. El modelo permitió identificar tejas adicionales, oxidación en canalones y tragaluces abiertos, que son cruciales para mantener la integridad del techo. Los resultados lograron una precisión del 90% en la detección de estos defectos.

Los resultados de este estudio fueron positivos en comparación con estudios similares (Alzarrad et al., 2022; Hou et al., 2023; Qiu et al., 2024). Además, si bien esta investigación se centra en defectos en tejados residenciales, estudios anteriores se concentraron en detectar elementos decorativos en tejados de edificios históricos. Sin embargo, todos los estudios subrayan la importancia de la identificación temprana de los problemas y la necesidad de realizar inspecciones de rutina para garantizar la integridad física de los edificios. Además, otros estudios, como los de

Santos et al. (2023), utilizan YOLO para detectar condensadores de aire acondicionado, señalando que estos elementos son más distintos en los techos, lo que los hace más fáciles de visualizar que detectar canaletas oxidadas. Los resultados de este estudio mostraron que el modelo reconoció este daño con hasta un 78% de precisión durante las pruebas.

Además, comparar este estudio con otra literatura que no utilizó arquitecturas YOLO revela algunos puntos en común. En primer lugar, los resultados de este estudio se alinean con los hallazgos de Staffa et al. (2023) respecto a la necesidad de realizar inspecciones en techos residenciales, considerando la mantenibilidad de estas estructuras. Otro punto importante es que aunque Staffa et al. (2023) abordaron un mayor número de clases a detectar en tejados, utilizaron algoritmos de una Interfaz de Programación de Aplicaciones (API) que no permiten ajustes de hiperparámetros, complicando las mejoras del modelo.

5. CONCLUSIÓN

Este estudio propuso un modelo de reconocimiento de imágenes para techos residenciales utilizando la octava generación de la arquitectura YOLO, YOLOv8. Se utilizó un rico conjunto de datos que consta de imágenes de 167 edificios recopilados por sistemas aéreos no tripulados (UAS) para evaluar de manera integral las condiciones reales de los techos de los edificios residenciales. Estas imágenes fueron anotadas en CVAT para detectar y clasificar tres tipos de daños en los tejados: tejas extra en dirección transversal, oxidación en canalones y claraboyas abiertas. El modelo se construyó utilizando 394 imágenes, de las cuales el 70% se utilizó para entrenamiento, el 20% para validación y el 10% para pruebas. A pesar del número limitado de imágenes utilizadas para el entrenamiento, los resultados fueron prometedores.

La arquitectura YOLOv8 demuestra ventajas significativas sobre sus versiones anteriores. YOLOv8 permite una detección y clasificación rápidas, ofreciendo una fácil implementación y eficiencia en el análisis de imágenes. Además, YOLOv8 se puede escalar para clasificación en tiempo real. Una limitación de este estudio fue el pequeño conjunto de datos de sólo 167 edificios, lo que puede haber afectado el rendimiento final del modelo. Además, la captura de imágenes mediante drones puede limitar la variedad de ángulos de captura. El trabajo futuro podría considerar un enfoque de recopilación de datos más diverso.

Los estudios futuros se centrarán en el desarrollo de arquitecturas YOLOv9 y YOLOv10 para la inspección de tejados. Además, se desarrollarán modelos de reconocimiento de objetos específicos de cubiertas para mejorar la gestión del mantenimiento. Estos modelos integrarán los datos generados en el sistema de gestión de mantenimiento y evaluarán posibles mejoras. También se explorarán oportunidades para incorporar otras tecnologías al proceso, optimizar la automatización e incluir una variedad más amplia de tipos y materiales de techos. Para fortalecer el modelo, los autores agregarán más imágenes y ampliarán el conjunto de clases de problemas en techos residenciales, con el objetivo de aumentar la solidez y precisión del análisis.

6. AGRADECIMIENTOS

Los autores desean agradecer a las agencias financiadoras CAPES y CNPq.

7. REFERENCIAS

Alashari, M., El-Rayes, K., Attalla, M., Al-Ghzawi, M. (2022). *Multivariate time series and regression models for forecasting annual maintenance costs of EPDM roofing systems*. Journal of Building Engineering, 54, 104618.

- Alizadeh, M., Ma, J. (2021). *A comparative study of series hybrid approaches to model and predict the vehicle operating states*. Computers & Industrial Engineering, 162, 107770. <https://doi.org/10.1016/j.cie.2021.107770>
- Alzarrad, A., Awolusi, I., Hatamleh, M. T., Terreno, S. (2022). *Automatic assessment of roofs conditions using artificial intelligence (AI) and unmanned aerial vehicles (UAVs)*. Frontiers in Built Environment, 8, 1026225.
- Avola, D., Cascio, M., Cinque, L., Fagioli, A., Foresti, G. L., Marini, M. R., Rossi, F. (2022). *Real-time deep learning method for automated detection and localization of structural defects in manufactured products*. Computers & Industrial Engineering, 172, 108512. <https://doi.org/10.1016/j.cie.2022.108512>
- Brown, S., Harris, W., Brooks, R. D., Dong, X. S. (2021). *Fatal injury trends in the construction industry*.
- Cao, Y., Pang, D., Zhao, Q., Yan, Y., Jiang, Y., Tian, C., Li, J. (2024). *Improved yolov8-gd deep learning model for defect detection in electroluminescence images of solar photovoltaic modules*. Engineering Applications of Artificial Intelligence, 131, 107866.
- Chu, J. C., Shui, C. S., & Lin, K. H. (2024). *Optimization of trucks and drones in tandem delivery network with drone trajectory planning*. Computers & Industrial Engineering, 189, 110000. <https://doi.org/10.1016/j.cie.2024.110000>
- Conceição, J., Poça, B., De Brito, J., Flores-Colen, I., Castelo, A. (2017). *Inspection, diagnosis, and rehabilitation system for flat roofs*. Journal of Performance of Constructed Facilities, 31(6), 04017100.
- Dang, L. M., Kyeong, S., Li, Y., Wang, H., Nguyen, T. N., Moon, H. (2021). *Deep learning-based sewer defect classification for highly imbalanced dataset*. Computers & Industrial Engineering, 161, 107630. <https://doi.org/10.1016/j.cie.2021.107630>
- Fan, C. L. (2024). *Using convolutional neural networks to identify illegal roofs from unmanned aerial vehicle images*. Architectural Engineering and Design Management, 20(2), 390-410.
- Gajjar, D., Burgett, J. (2020). *Evaluating the Use of Unmanned Aerial Systems (UAS) to Perform Low-Slope Roof Inspections*. EPiC Series in Built Environment, 1, 214-222.
- Garcez, N., Lopes, N., de Brito, J., Silvestre, J. (2012). *System of inspection, diagnosis and repair of external claddings of pitched roofs*. Construction and Building Materials, 35, 1034-1044.
- Han, S., Park, W., Jeong, K., Hong, T., Koo, C. (2024). *Utilizing synthetic images to enhance the automated recognition of small-sized construction tools*. Automation in Construction, 163, 105415. <https://doi.org/10.1016/j.autcon.2024.105415>
- Hou, M., Hao, W., Dong, Y., Ji, Y. (2023). *A detection method for the ridge beam based on improved YOLOv3 algorithm*. Heritage Science, 11(1), 167. <https://doi.org/10.1186/s40494-023-00995-4>
- Jiang, D., Kong, L., Wang, H., Pan, D., Li, T., Tan, J. (2024). *Precise control mode for concrete vibration time based on attention-enhanced machine vision*. Automation in Construction, 158, 105232. <https://doi.org/10.1016/j.autcon.2023.105232>
- Koubaa, A., Ammar, A., Kanhouch, A., AlHabashi, Y. (2021). *Cloud versus edge deployment strategies of real-time face recognition inference*. IEEE Transactions on Network Science and Engineering, 9(1), 143-160. <https://doi.org/10.1109/TNSE.2021.3055835>
- Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Wei, X. (2022). *YOLOv6: A single-stage object detection framework for industrial applications*. arXiv preprint arXiv:2209.02976. <https://doi.org/10.48550/arXiv.2209.02976>
- Liu, Y. F., Nie, X., Fan, J. S., Liu, X. G. (2020). *Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction*. Computer-Aided Civil and Infrastructure Engineering, 35(5), 511-529.

- Ma, X., Li, Y., Yang, Z., Li, S. (2024). *Lightweight network for millimeter-level concrete crack detection with dense feature connection and dual attention*. Journal of Building Engineering, 109821. <https://doi.org/10.1016/j.jobe.2024.109821>
- Mostafa, K., Hegazy, T., Hunsperger, R. D., Elias, S. (2023). *Using image analysis to quantify defects and prioritize repairs in built-up roofs*. Facilities, 41(7/8), 498-509.
- Ni, Y. H., Wang, H., Mao, J. X., Xi, Z., Chen, Z. Y. (2024). *Quantitative detection of typical bridge surface damages based on global attention mechanism and YOLOv7 network*. Structural Health Monitoring, 14759217241246953. <https://doi.org/10.1177/14759217241246953>
- Olimov, B. A., Veluvolu, K. C., Paul, A., Kim, J. (2022). *UzADL: Anomaly detection and localization using graph Laplacian matrix-based unsupervised learning method*. Computers & Industrial Engineering, 171, 108313. <https://doi.org/10.1016/j.cie.2022.108313>
- Qiu, H., Zhang, J., Zhuo, L., Xiao, Q., Chen, Z., Tian, H. (2024). *Research on intelligent monitoring technology for roof damage of traditional Chinese residential buildings based on improved YOLOv8: taking ancient villages in southern Fujian as an example*. Heritage Science, 12(1), 231.
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). *You only look once: Unified, real-time object detection*. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
- Ribeiro, W. S., Zanetti, J., Totola, L. B., Junqueira, S. Â. C., Lauff, P. H. P. (2024). *Crack detection in buildings using the YOLO v8 network*. Revista ALCONPAT, 14(3), 288-298. <https://doi.org/10.21041/ra.v14i3.76>.
- Rosenberger, T. (2012), *Roofing Systems in Costa Rica Need Careful Consideration*, The Costa Rica Star, <https://news.co.cr/roofing-systems-in-costa-rica-need-careful-consideration/7258/>. Acessado em outubro de 2023.
- Santos, L. M. A., Zanoni, V. A. G., Bedin, E., Pistori, H. (2023). *Deep learning applied to equipment detection on flat roofs in images captured by UAV*. Case Studies in Construction Materials, 18, e01917.
- Silva, C. M., Flores-Colen, I., Coelho, A. (2015). *Green roofs in Mediterranean areas—Survey and maintenance planning*. Building and Environment, 94, 131-143.
- Silveira, B., Melo, R., Costa, D. B. (2020). *Using UAS for roofs structure inspections at post-occupational residential buildings*. International Conference on Computing in Civil and Building Engineering (pp. 1055-1068). Cham: Springer International Publishing.
- Sohaib, M., Jamil, S., Kim, J. M. (2024). *An ensemble approach for robust automated crack detection and segmentation in concrete structures*. Sensors, 24(1), 257. <https://doi.org/10.3390/s24010257>
- Solawetz, J. (2023). *What is yolov8? the ultimate guide*. Roboflow Blog, 25.
- Staffa, L. B. J., Costa, D. B., Nogueira, J. L. T., and Silva, A. S. (2023). *Web platform for building roof maintenance inspection using UAS and artificial intelligence*. International Journal of Building Pathology and Adaptation. <https://doi.org/10.1108/IJBPA-12-2022-0186>.
- Wang, C. Y., Bochkovskiy, A., Liao, H. Y. M. (2023). *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 7464-7475).
- Wang, W., Chen, J., Huang, Z., Yuan, H., Li, P., Jiang, X., Lin, Q. (2023). *Improved YOLOv7-based algorithm for detecting foreign objects on the roof of a subway vehicle*. Sensors, 23(23), 9440. <https://doi.org/10.3390/s23239440>
- Zhao, H., Wang, X., Sun, J., Wang, Y., Chen, Z., Wang, J., Xu, X. (2024). *Artificial intelligence powered real-time quality monitoring for additive manufacturing in construction*. Construction and Building Materials, 429, 135894. <https://doi-org/10.1016/j.conbuildmat.2024.135894>