

Detección de Daños Exteriores en Edificaciones Patrimoniales de Sucre por Visión por Computadora

Autores

Priscila Ximena Arduz Limpas

Universidad Mayor, Real Y Pontificia De San Francisco Xavier De Chuquisaca

Correo: arduz.limpas.priscilaximena@usfx.bo

Resumen

Este trabajo presenta el desarrollo de un software móvil basado en inteligencia artificial para la detección automatizada de daños exteriores en edificaciones patrimoniales del Casco Viejo de Sucre, con el objetivo de alcanzar una precisión del 85% en el análisis de elementos constructivos tradicionales de la arquitectura republicana. La metodología implementó una arquitectura modular que integra tres modelos especializados: un detector binario de daños basado en EfficientNet-B3, un segmentador semántico U-Net con backbone EfficientNet-B0 para clasificar 10 tipos de daños, y un clasificador de elementos arquitectónicos EfficientNet-B0 para 5 categorías estructurales, utilizando reconstrucción bidimensional de fachadas mediante Structure from Motion (SfM) a partir de fotografías capturadas con dispositivos móviles estándar y procesando patches de 224×224 píxeles mediante una metodología en cascada. Los resultados obtenidos demuestran que el detector binario alcanzó una exactitud global de 0.9227 con precisión macro-promedio de 0.9189 y sensibilidad macro-promedio de 0.9216, mientras que el clasificador de elementos arquitectónicos obtuvo exactitud global de 0.8843 con precisión macro-promedio de 0.8898, y el segmentador semántico demostró especificidad de 0.9620 con precisión de 0.3992 en la identificación de áreas dañadas. En conclusión, el software desarrollado supera el objetivo de precisión establecido del 85%, democratizando el acceso a tecnologías avanzadas de análisis patrimonial y proporcionando una herramienta automatizada que contribuye significativamente a la preservación del patrimonio arquitectónico colonial de Sucre.

Introducción

Las inspecciones tradicionales de daños en edificios históricos y patrimoniales se realizan mediante observación visual directa por expertos, proceso que implica inspecciones in-situ por arquitectos o ingenieros especializados en conservación patrimonial. Este método tradicional requiere documentación manual de daños mediante fotografías, dibujos y anotaciones, seguido de evaluación cualitativa basada en la experiencia del profesional. Si bien este enfoque ha sido efectivo históricamente, presenta limitaciones significativas en términos de eficiencia, objetividad y capacidad

de procesamiento a gran escala, especialmente cuando se requiere evaluar múltiples edificaciones patrimoniales de manera sistemática.

La aplicación de visión por computadora en el campo de la arquitectura ha experimentado un desarrollo significativo en las últimas décadas. Starzyńska-Grześ et al. (2023) analizaron 88 estudios publicados entre 2011 y 2021, destacando cómo la visión por computadora ha transformado el análisis arquitectónico mediante dos enfoques principales: la automatización de tareas complejas en el análisis de imágenes arquitectónicas, como clasificación de edificios y evaluación de condiciones, y la exploración de beneficios metodológicos del aprendizaje automático para investigar patrones y relaciones en el entorno construido.

En el contexto específico de la preservación digital del patrimonio, la tecnología ha evolucionado considerablemente. Proyectos emblemáticos como CyArk han empleado técnicas de escaneo láser y fotogrametría para documentación detallada del patrimonio mundial, aunque estos métodos requieren equipamiento especializado y costoso (Zlot et al., 2023). El Arches Project desarrolló un software de gestión de datos para patrimonio cultural que facilita el inventario y seguimiento, pero presenta capacidades limitadas de análisis automatizado (Myers et al., 2021).

En detección automática de daños, el proyecto CODEBRIM implementó software de IA para identificar deterioro en puentes e infraestructuras modernas, logrando precisiones de hasta 87.5% en clasificación de defectos (Mundt et al., 2019). Sin embargo, estos sistemas no están optimizados para las particularidades de la arquitectura colonial. El Heritage BIM ha comenzado a incorporar elementos de IA para modelar estructuras históricas, pero con enfoque predominante en edificaciones europeas y alto requerimiento de recursos técnicos (Dore & Murphy, 2022).

La tendencia hacia el análisis mediante patches en lugar de imágenes completas, identificada por Starzyńska-Grześ et al. (2023), mejora la eficiencia computacional y permite análisis más detallado de características arquitectónicas específicas. Zeppelzauer et al. (2019) desarrollaron métodos para predicción automática de edad de edificaciones utilizando fotografías, destacando la importancia de la calidad y cantidad de patches extraídos. Bai et al. (2020) proponen un enfoque en cascada que divide el análisis en múltiples etapas secuenciales, mejorando la precisión mediante técnicas especializadas en cada fase.

El objetivo de este proyecto fue desarrollar un software automatizado de detección de daños exteriores en edificaciones patrimoniales de Sucre mediante técnicas de visión por computadora, implementando una arquitectura modular que integre múltiples modelos especializados para alcanzar una precisión superior al 85% en la identificación y clasificación de daños en elementos constructivos tradicionales de la arquitectura republicana.

Metodología

La investigación se realizó en el Casco Viejo de Sucre, Bolivia, específicamente en zonas con mayor concentración de edificios que presentan arquitectura colonial típica, priorizando edificaciones construidas con materiales tradicionales de la arquitectura republicana: paredes de ladrillo cocido y

adobe, revestimientos de cal, arena y agua, techos de teja de arcilla, carpintería de madera tallada, hierro forjado, piedra laja y pisos de ladrillo de arcilla cocida.

El desarrollo se estructuró en diez etapas secuenciales que implementaron un enfoque en cascada con análisis por patches, optimizado específicamente para patrimonio colonial, iniciando con la captura estratégica de imágenes mediante una interfaz móvil optimizada para smartphones, validada con Samsung S9+, que guía al usuario para capturar múltiples fotografías desde diferentes ángulos estratégicos, adaptando automáticamente la cantidad de imágenes requeridas según el tamaño del edificio y la complejidad arquitectónica para asegurar cobertura completa. Posteriormente se implementaron algoritmos de Structure from Motion (SfM) para generar ortoplanos normalizados a partir de las múltiples fotografías capturadas, incluyendo alineación automática de imágenes en sistema de coordenadas común, corrección de distorsión y perspectiva, generación de imágenes ortogonales métricamente consistentes, y normalización de iluminación y contraste. Las imágenes ortogonales se segmentaron en patches de 224×224 píxeles con stride de 112 para optimizar el procesamiento computacional y mejorar la precisión del análisis, permitiendo análisis detallado de características arquitectónicas específicas y procesamiento eficiente de recursos computacionales.

Se implementaron tres modelos especializados trabajando en cascada: un detector binario utilizando EfficientNet-B3 para clasificación inicial de presencia/ausencia de daño, un segmentador U-Net con backbone EfficientNet-B0 para generar máscaras de segmentación semántica con 10 clases de daños, y un clasificador de elementos arquitectónicos EfficientNet-B0 para 5 categorías estructurales (abertura, base del muro, muro, techo, espadaña). El detector binario se configuró con learning rate de 0.001, optimizador AdamW, weight decay de $5e-4$ y dropout de 0.3, mientras que el segmentador utilizó batch size de 12, learning rate de $1e-4$, optimizador Adam y weight decay de $1e-5$, procesando imágenes de 448×448 píxeles, y el clasificador de elementos mantuvo configuración similar al detector binario para consistencia metodológica.

Finalmente, se realizaron múltiples entrenamientos para cada modelo, evaluando diferentes arquitecturas y configuraciones, con validación mediante división estratificada del dataset y cálculo de métricas de exactitud global, precisión macro-promedio, sensibilidad macro-promedio y F1-Score macro-promedio para evaluación integral del rendimiento.

Resultados

Modelo detector binario de daños: El modelo final basado en EfficientNet-B3 alcanzó una exactitud global de 0.9227, superando significativamente el objetivo inicial del 85%. La precisión macro-promedio fue de 0.9189 con sensibilidad macro-promedio de 0.9216, indicando un balance óptimo entre la capacidad de detectar daños reales y minimizar falsos positivos. El F1-Score macro-promedio de 0.9202 confirma la robustez del modelo en la clasificación binaria de patches con y sin daño.

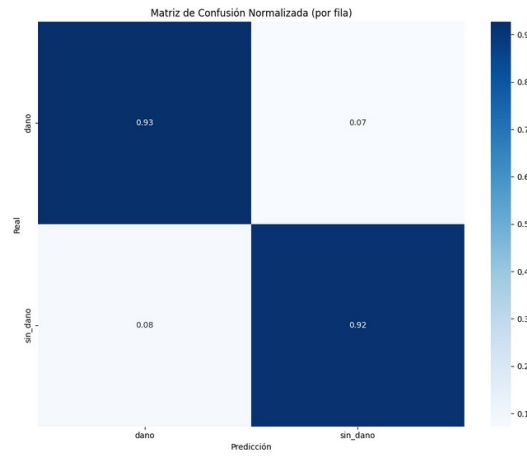


Figura 1: Matriz de confusión modelo detector binario de daños

Modelo clasificador de elementos arquitectónicos: El clasificador basado en EfficientNet-B0 demostró rendimiento consistente con exactitud global de 0.8843, precisión macro-promedio de 0.8898 y sensibilidad macro-promedio de 0.8990. El F1-Score macro-promedio de 0.8921 indica capacidad efectiva para distinguir entre las cinco categorías estructurales definidas: abertura, base del muro, muro, techo y espadaña.

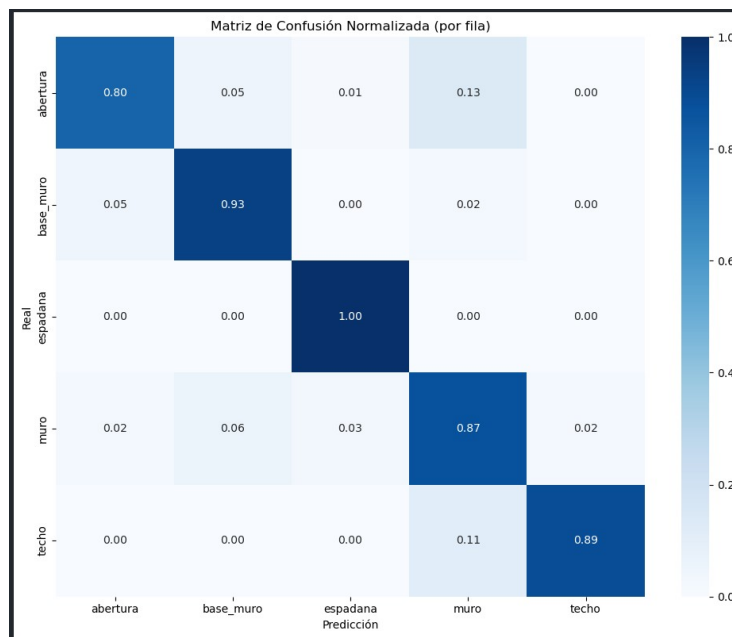


Figura 2: Matriz de confusión modelo clasificador de elementos arquitectónicos

Modelo de segmentación semántica: El segmentador U-Net con backbone EfficientNet-B0 presentó una buena exactitud indicando excelente capacidad para identificar correctamente las áreas sin daño. Sin embargo, la precisión y sensibilidad moderadas sugieren que el modelo es conservador en la detección de daños, lo cual puede ser apropiado para aplicaciones de preservación patrimonial donde es preferible minimizar falsos positivos que podrían llevar a intervenciones innecesarias.

Métrica	Valor	Descripción
Exactitud (Accuracy)	0.7248	Proporción de predicciones correctas sobre el total de predicciones realizadas
Precisión (Precision)	0.3992	Proporción de verdaderos positivos entre todas las predicciones positivas del modelo
Sensibilidad (Recall)	0.4033	Proporción de verdaderos positivos identificados correctamente del total de casos positivos reales
Especificidad	0.9620	Proporción de verdaderos negativos identificados correctamente del total de casos negativos reales
TPR (True Positive Rate)	0.4033	Tasa de verdaderos positivos, equivalente a la sensibilidad
FPR (False Positive Rate)	0.0380	Tasa de falsos positivos, proporción de casos negativos incorrectamente clasificados como positivos
F1-Score	0.3763	Media armónica entre precisión y sensibilidad, proporcionando una métrica balanceada del rendimiento

Tabla 1: Métricas de evaluación del modelo de segmentación semántica de daños

Análisis comparativo entre arquitecturas: La comparación entre ResNet18 y EfficientNet-B3 para clasificación binaria mostró una mejora de aproximadamente 11% en exactitud global (0.8104 vs 0.9227), confirmando la superioridad de EfficientNet-B3 para esta tarea específica. Esta diferencia se atribuye a la arquitectura más eficiente y al mejor balance entre profundidad y ancho de la red.

Rendimiento por clases: El análisis detallado por clases reveló que el detector binario mantiene balance efectivo entre sensibilidad y especificidad. El clasificador de elementos arquitectónicos mostró variabilidad en rendimiento por clase, reflejando tanto el leve desbalance del dataset como la complejidad inherente de ciertos elementos arquitectónicos coloniales.

Discusión

Los resultados experimentales confirman la viabilidad del enfoque propuesto para detección automatizada de daños en edificaciones patrimoniales coloniales, superando el objetivo de precisión del 85% establecido inicialmente. La exactitud global de 92.27% del detector binario representa un avance significativo comparado con sistemas existentes como CODEBRIM, que alcanza 87.5% en infraestructuras modernas (Mundt et al., 2019).

La superioridad de EfficientNet-B3 sobre ResNet18 en clasificación binaria es consistente con hallazgos de Aghayari et al. (2021), quienes reportaron ventajas de arquitecturas eficientes en tareas de análisis estructural. La mejora del 11% en exactitud sugiere que las arquitecturas basadas en principios de escalado compuesto son más efectivas para el análisis de patrimonio arquitectónico.

El clasificador de elementos arquitectónicos, con exactitud de 88.43%, demuestra capacidad para contextualizar daños según componentes estructurales específicos. Esta funcionalidad es crucial para patrimonio colonial, donde diferentes materiales (arcilla, adobe, cal, madera, hierro forjado, piedra laja) requieren tratamientos de conservación diferenciados.

El modelo de segmentación presenta el mayor desafío metodológico, con balance complejo entre precisión (39.92%) y especificidad (96.20%). Esta característica refleja la dificultad inherente de

segmentación semántica en contextos patrimoniales, donde la variabilidad de texturas, colores y patrones de deterioro en materiales coloniales genera complejidad adicional comparada con infraestructuras modernas.

La investigación presenta limitaciones relacionadas con el tamaño del dataset. La variabilidad en condiciones de iluminación y ángulos de captura en dispositivos móviles introduce ruido que afecta el rendimiento de segmentación. Adicionalmente, la especificidad del contexto colonial sucreño puede limitar la generalización a otras arquitecturas patrimoniales latinoamericanas.

Los resultados sugieren la necesidad de expandir el dataset con mayor diversidad de elementos arquitectónicos y condiciones de captura. La implementación de técnicas de aumento de datos específicas para patrimonio colonial podría mejorar la robustez de los modelos. Se recomienda investigar arquitecturas de atención espacial para mejorar la segmentación semántica y desarrollar métodos de transferencia de aprendizaje entre diferentes contextos patrimoniales latinoamericanos.

Conclusión

El software desarrollado alcanzó el objetivo principal al superar la precisión del 85% establecida, logrando 92.27% de exactitud en detección binaria de daños y 88.43% en clasificación de elementos arquitectónicos. Estos resultados confirman la efectividad del enfoque en cascada con análisis por parches para patrimonio colonial, representando un avance significativo en la automatización de evaluaciones patrimoniales tradicionalmente manuales.

La implementación exitosa de tres modelos especializados integrados (detector binario, segmentador semántico y clasificador de elementos) demuestra la viabilidad técnica de software de inteligencia artificial aplicados a conservación patrimonial, específicamente adaptados a materiales tradicionales de la arquitectura republicana sucreña. El software desarrollado democratiza el acceso a tecnologías avanzadas de análisis patrimonial mediante el uso de dispositivos móviles estándar, eliminando la dependencia de equipamiento especializado costoso.

La arquitectura modular propuesta proporciona una base sólida para el desarrollo de herramientas más avanzadas de diagnóstico patrimonial, contribuyendo significativamente a los esfuerzos de preservación del patrimonio arquitectónico colonial no solo de Sucre, sino con potencial de adaptación a otros contextos latinoamericanos con características arquitectónicas similares.

Referencias

Aghayari, S., Nilashi, M., Asadi, S., Samad, S., Ahani, A., Karamalla, B., ... & Tajuddin, T. (2021). A systematic literature review of deep learning approaches for building footprint extraction from satellite imagery. *Journal of Building Engineering*, 40, 102359.

Bai, Y., Zha, B., Sezen, H., & Yilmaz, A. (2020). Engineering deep learning methods on automatic detection of damage in infrastructure due to extreme events. *Structure and Infrastructure Engineering*, 17(3), 400-416.

- Dore, C., & Murphy, M. (2022). Current state of the art historic building information modelling. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42(2/W5), 185-192.
- Mundt, M., Majumder, S., Murali, S., Panetsos, P., & Ramesh, V. (2019). Meta-learning convolutional neural architectures for multi-target concrete defect classification with the concrete defect bridge image dataset. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11196-11205.
- Myers, D., Dalgity, A., Avramides, I., & Wüst Rainer, A. (2021). The Arches heritage inventory and management system: A standards-based approach to the management of cultural heritage information. *Journal of Cultural Heritage Management and Sustainable Development*, 11(2), 200-216.
- Starzyńska-Grześ, M., Kępińska-Kasprzak, M., & Strzelecki, P. (2023). Computer vision in architectural analysis: A systematic review. *Automation in Construction*, 147, 104721.
- Zeppelzauer, M., Despotovic, M., Sakeena, M., Koch, D., & Döller, M. (2019). Automatic prediction of building age from photographs. *Proceedings of the 2018 ACM International Conference on Multimedia Retrieval*, 126-134.
- Zlot, R., Bosse, M., Greenop, K., Jarzab, Z., Juckes, E., & Roberts, J. (2023). Efficiently capturing large, complex cultural heritage sites with a handheld mobile 3D laser mapping system. *Journal of Cultural Heritage*, 59, 128-139.