

ARTÍCULO ORIGINAL

Detección de la úlcera del pie diabético utilizando Inteligencia Artificial

Detection of diabetic foot ulcers using artificial intelligence

Edgar Moya Cáceres

edgarcmoyma@gmail.com • <https://orcid.org/0009-0004-6197-430>

Yusely Ruíz González

yuselyr@uclv.edu.cu • <https://orcid.org/0000-0002-6011-1947>

María Matilde García Lorenzo

mmgarcia@uclv.edu.cu • <https://orcid.org/0000-0002-1663-5794>

UNIVERSIDAD CENTRAL “MARTA ABREU” DE LAS VILLAS, CUBA

Recibido: 2025-11-05 • Aceptado: 2025-12-18

RESUMEN

Las úlceras del pie diabético (UPD) son una de las complicaciones más frecuentes y devastadoras de la diabetes, representando un desafío considerable para los sistemas de salud y afectando significativamente la calidad de vida de los pacientes. El objetivo de este artículo es abordar la aplicación de técnicas de inteligencia artificial para la detección de UPD. Se implementaron modelos de aprendizaje profundo (DL), específicamente YOLOv8 en sus variantes m y l, y Faster R-CNN estándar y con convoluciones deformables. Estos modelos se optimizaron mediante técnicas de aumento de datos y ajuste de hiperparámetros. Para el entrenamiento, se empleó el dataset proporcionado por el Diabetic Foot Ulcers Grand Challenge 2020 (DFUC 2020). Adicionalmente, se utilizó una base de casos de imágenes nacionales para evaluar la eficacia del modelo en el contexto local. En la etapa de entrenamiento y validación se obtuvo un mAP superior a 0.70 para todos los modelos entrenados, siendo el YOLOv8l el de mejor resultado (mAP=0.799). La prueba externa con los datos nacionales mostró poder de generalización al obtenerse valores de precisión de 0,856, sensibilidad de 0,789, puntuación F1 de 0,821 y un mAP50 de 0,839. También se desarrolló una aplicación web, utilizando Streamlit, que proporciona una interfaz interactiva para el análisis de imágenes de UPD. Los modelos de DL para detección de objetos implementados demostraron ser herramientas eficaces para la identificación precisa de UPD. La aplicación desarrollada integra el mejor modelo de forma efectiva, ofreciendo una interfaz intuitiva para los profesionales de la salud.

Palabras clave: Úlceras del pie diabético; Detección de objetos; Aprendizaje profundo; DFUC2020.

ABSTRACT

Diabetic foot ulcers (DFUs) are one of the most common and devastating complications of diabetes, representing a considerable challenge for healthcare systems and significantly affecting patients' quality of life. This paper is aimed to address the application of artificial intelligence techniques for DFU detection. Deep learning (DL) models were implemented, specifically YOLOv8, in its m and l variants, and standard and deformable convolutions Faster R-CNN. These models were optimized using data augmentation and hyperparameter tuning techniques. The dataset provided by the Diabetic Foot Ulcers Grand Challenge 2020 (DFUC 2020) was used for training. Additionally, a national image database was used to evaluate the model's effectiveness in the local context. In the training and validation stages, mAP greater than 0.70 was obtained for all trained models, being the YOLOv8l model the best performed (mAP=0.799). The external test with national data demonstrated generalization power, granting accuracy values of 0.856, sensitivity of 0.789, F1 score of 0.821, and an mAP50 of 0.839. A web application was also developed using Streamlit, providing an interactive interface for analyzing DFU images. The object detection DL models implemented proved to be effective tools for accurately identifying DFUs. The developed application effectively integrates the best performed model, offering an intuitive interface for healthcare professionals.

Keywords: Diabetic foot ulcers; Object detection; Deep learning; DFUC2020.

INTRODUCCIÓN

En los últimos años, la diabetes (Diabetes Mellitus) se ha convertido en una de las enfermedades crónicas más comunes, con una amplia gama de complicaciones y un notable impacto en la salud. El pie diabético (PD) es una de las complicaciones más frecuentes y devastadoras de la diabetes. Según la Organización Mundial de la Salud (OMS), el PD se caracteriza por la presencia de ulceración, infección y/o gangrena en el pie, asociada a neuropatía diabética y diferentes grados de enfermedad vascular periférica. Esta condición es el resultado de la interacción compleja de múltiples factores inducidos por una hiperglucemia mantenida (Juarez-Salinas, Solis-Sánchez, Castañeda-Miranda, Ortiz-Rodríguez & Rosales, 2014).

Una úlcera del pie diabético (UPD) es una llaga o herida abierta que se desarrolla en los pies de las personas con diabetes, y representa un desafío significativo para el manejo clínico debido a su tendencia a la infección y a la cicatrización deficiente. El tratamiento de las UPD es desafiante y costoso, afectando a los pacientes y sus familias tanto psicológica como económicamente, y planteando desafíos significativos para los sistemas de salud en todo el mundo (Sun, Heng, Liu, Geng, & Liang, 2023). En el pasado, el conocimiento médico práctico se obtenía a través de libros de texto, revistas, opiniones de expertos y la relación maestro-aprendiz. Hoy en día, la combinación adecuada de "datos" y "experiencia" puede mejorar la atención si se integran correctamente, lo cual es esencial para aplicar la inteligencia artificial (IA) en la medicina (Tekkeşin, 2019).

Anualmente, se celebra una competencia internacional de clasificación de imágenes médicas organizada por la Medical Image Computing and Computer Assisted Intervention (MICCAI). El objetivo del Diabetic Foot Ulcers Grand Challenge 2020 (DFUC 2020), dentro del MICCAI 2020, fue mejorar la precisión en la detección de UPD en entornos clínicos y promover el uso de técnicas avanzadas de aprendizaje automático basadas en datos, proporcionando a la comunidad investigadora un conjunto de datos significativo etiquetado con precisión para la identificación de UPD. A pesar de los avances logrados, estos retos siguen siendo una problemática abierta, ya que la detección precisa y eficiente de las UPD continúa representando un desafío importante en la práctica clínica. Las ediciones recientes de competencias similares, como las destacadas por MIUA 2024 (Manchester Metropolitan University. 2024, July 2), siguen impulsando la innovación y la colaboración en este campo crucial. En 2021, tras la culminación del DFUC 2020, se publicó un artículo [5] que actualiza sobre los modelos de detección de objetos de aprendizaje profundo (del inglés, "Deep Learning") de última generación aplicados a este problema y con los cuales se logró una precisión promedio media (mAP) de 0,6940 y una puntuación F1 de 0,7434.

En Cuba la medición de las áreas de las lesiones de UPD ha sido un proceso manual o semiautomático. Se han realizado varios estudios y desarrollos tecnológicos para mejorar la precisión y eficiencia en la evaluación de las UPD. Un ejemplo notable es el uso del dispositivo FrameHeber 03® (Cabal Mirabal, González Dalmau, Oramas Díaz, Herrera Martínez, Berlanga Acosta & Fernández Montequín, 2019), desarrollado por el Centro de Ingeniería Genética y Biotecnología (CIGB), que permite la toma estandarizada de imágenes fotográficas de las úlceras y facilita evaluaciones cuantitativas del tratamiento aplicado a las UPD. No obstante, aún no se ha reportado la existencia de software cubano que realice de manera totalmente automatizada la medición de las áreas de las lesiones (García Gómez, 2019). Automatizar este proceso mediante un sistema basado en IA sería de gran ayuda para Cuba, ya que aliviaría la carga sobre los profesionales de la salud y les permitiría enfocarse en el tratamiento y la prevención de otras complicaciones relacionadas con la diabetes. En este contexto, en la presente investigación se implementaron modelos de aprendizaje profundo, específicamente YOLOv8 en sus variantes m y l, y Faster R-CNN en su variante estándar y con convoluciones deformables. Estos modelos se optimizaron mediante técnicas de aumento de datos y ajuste de hiperparámetros para maximizar su rendimiento. Para facilitar la interacción de los profesionales de la salud con el modelo de mejor desempeño, YOLOv8l, comenzó el desarrollo de una aplicación web utilizando Streamlit que proporciona una interfaz interactiva e intuitiva y posibilita su implementación en el entorno clínico.

METODOLOGÍA

Conjunto de datos

Se utilizaron las 2000 imágenes etiquetadas del conjunto DFUC2020, liberadas durante la fase 1 de la competencia pues son de las que se brinda la localización de las lesiones. Estas se distribuyeron en 1800 imágenes para el proceso de entrenamiento y 200 para la validación, las cuales presentan un total de 2235 y 261 úlceras respectivamente.

Para la prueba externa del modelo, se utilizó un conjunto de 36 imágenes de origen nacional, las cuales contienen un total de 38 úlceras. Las imágenes fueron anotadas por profesionales especializados en el manejo del pie diabético, lo que asegura la precisión en la localización exacta de las úlceras. Para la anotación de las imágenes se empleó el VGG Image Annotator (Dutta, Zisserman, & Gupta, 2024). Todas las imágenes fueron etiquetadas con coordenadas xmin, ymin, xmax y ymax para indicar la ubicación precisa de las úlceras del pie.

Técnicas de aumento de datos

El aumento de datos, o data augmentation, es una técnica fundamental en el aprendizaje profundo para mejorar la capacidad de generalización de los modelos. Consiste en aplicar diversas transformaciones a los datos de entrenamiento, generando nuevas versiones de los mismos sin modificar sus etiquetas (Liu, L., et al., 2020). En el desarrollo de esta investigación se implementan varias técnicas de aumento de datos utilizando la librería Albumentations (Buslaev, Iglovikov, Khvedchenya, Parinov, Druzhinin, & Kalinin, 2020), que permite aplicar diversas transformaciones tanto a las imágenes como a sus cajas delimitadoras. En la Figura 1 se presentan ejemplos de imágenes originales y generadas para ilustrar el impacto de estas técnicas en el conjunto de entrenamiento.



Figura 1. Ejemplos de imágenes originales y generadas con el aumento de datos.

Implementación de los modelos

Se implementaron 2 modelos basados en la arquitectura Faster R-CNN, que obtuvo los mejores resultados en la competencia DFU 2020, junto con la última versión de YOLO.

Faster R-CNN

La implementación del modelo de Faster R-CNN se llevó a cabo utilizando la biblioteca MMDetection (Chen, et al., 2019). Se implementaron dos variantes, ambas basadas en la arquitectura ResNeXt-101, preentrenada con el checkpoint "resnext101_64x4d" disponible en OpenMMLab. (2024, October 15). La primera variante utiliza ResNeXt-101 en su configuración estándar, mientras que la segunda variante integra Deformable Convolutional Networks v2 (DCNv2). En esta segunda variante, la arquitectura fundamental de la red se mantiene alineada con la de Faster R-CNN, sin embargo, se sustituye la capa de convolución estándar por una capa de convolución deformable en el segundo, tercer y cuarto bloque del extractor de características de ResNeXt. Con el objetivo de optimizar el rendimiento de los modelos, se realizó una búsqueda de hiperparámetros, considerando enfoques y resultados de investigaciones previas en la detección de UPD [5].

YOLOv8

Para la implementación del modelo YOLOv8, fue utilizada la librería (Ultralytics Jocher, Chaurasia, & Qiu, 2023). Las variantes YOLOv8m y YOLOv8l fueron implementadas debido a su capacidad para mantener un buen balance entre precisión y rendimiento. Estas versiones medianas se consideraron adecuadas, ya que la salida del modelo será empleada en una etapa posterior de clasificación, lo que requiere una detección precisa sin comprometer la

eficiencia. Con el fin de mejorar el rendimiento del modelo, se llevó a cabo una búsqueda de hiperparámetros inspirada en trabajos previos como [5] que aplicaron versiones anteriores de este modelo en la detección de UPD.

Posprocesamiento

Se incorporó una etapa de posprocesamiento que consta de dos fases. Primeramente, se filtran las predicciones con un umbral de confianza de 0,3 lo que significa que se consideraron detecciones con una probabilidad de al menos 30% de ser correctas, reduciendo así las falsas detecciones al eliminar aquellas con baja probabilidad de ser correctas. Esta probabilidad corresponde al nivel de confianza asignado por el modelo a cada predicción, el cual se calcula a partir de la salida de la red neuronal. Posteriormente, se aplica Supresión de No-Máximos (NMS, por sus siglas en inglés) con un umbral de intersección sobre unión (IoU) de 0,5 permitiendo eliminar las detecciones redundantes y asegurar que cada objeto detectado se represente de manera única en la imagen.

Métricas utilizadas

La evaluación del rendimiento de los modelos de detección de UPD se basa en métricas fundamentales como son: el número de verdaderos positivos (TP), el número de falsos positivos (FP), la precisión, la sensibilidad, la puntuación F1 y la precisión media promedio (mAP). Estas métricas permiten evaluar la capacidad predictiva de cada modelo de detección entrenado con el conjunto de datos, proporcionando una visión integral de su desempeño (Yap, 2021).

La puntuación F1 se considera particularmente relevante en este contexto debido a su enfoque en los falsos negativos y falsos positivos. Dado que los falsos negativos pueden conllevar riesgos significativos de complicaciones adicionales para los pacientes, y los falsos positivos pueden resultar en costos adicionales y cargas de tiempo para las clínicas, la puntuación F1 ofrece una medida más precisa del rendimiento del modelo que el simple porcentaje de predicciones correctas (Sarmun, et al. 2024). Por su parte, el mAP, y en particular el mAP50, permite evaluar el grado de superposición entre las predicciones y la verdad de terreno, considerando como verdaderas detecciones aquellas con una IoU igual o superior a 0,5 (Cassidy, B., et al. 2021). Estas métricas en conjunto permiten una valoración integral de la capacidad predictiva del modelo.

RESULTADOS Y DISCUSIÓN

Los resultados que se presentan en la Tabla 1 corresponden al conjunto de validación y muestran el mejor desempeño alcanzado por cada modelo, obtenidos al aplicar técnicas de aumento de datos.

Tabla 1. Resultados obtenidos en el conjunto de validación por los modelos implementados, solo se registraron las métricas mAP50 y mAP50-95 para los modelos R-CNN debido a que MMDetection solo proporciona estas métricas.

Modelos	Sensibilidad	Precisión	F1	mAP50	mAP50-95
Faster R-CNN	-	-	-	0,725	0,392
Faster R-CNN Deformable	-	-	-	0,715	0,392
YOLOv8m	0,739	0,774	0,756	0,782	0,447

YOLOv8l	0,716	0,805	0,758	0,799	0,474
----------------	-------	-------	-------	-------	-------

Los resultados obtenidos con Faster R-CNN fueron similares en ambos casos, mostrando que el modelo con convoluciones deformables no ofreció una mejora significativa respecto al modelo estándar. El modelo estándar alcanzó un mAP50 de 0,725, mientras que la variante con convoluciones deformables obtuvo un mAP50 de 0,715.

En la comparación entre las variantes de YOLOv8, ambas ofrecieron buenos resultados. YOLOv8l obtuvo un mejor rendimiento general, destacándose en precisión (0,805), puntuación F1 (0,758), mAP50 (0,799) y mAP50-95 (0,474). Sin embargo, YOLOv8m presentó una sensibilidad ligeramente superior (0,739 frente a 0,716).

El modelo YOLOv8l, optimizado mediante técnicas de aumento de datos, fue seleccionado para evaluar las imágenes de la base de casos nacional debido a que demostró un mayor desempeño en el conjunto de validación. Al evaluar el modelo se obtuvieron métricas consistentes: precisión de 0,856, sensibilidad de 0,789, puntuación F1 de 0,821 y un mAP50 de 0,839. Estas métricas reflejan la efectividad del modelo en la detección correcta de úlceras y su capacidad para identificar la mayor cantidad posible de úlceras presentes en las imágenes.

Implementación de aplicación web para la detección de UPD

La implementación de una aplicación web intuitiva para la detección de UPD no solo facilita el acceso a los resultados del modelo, sino que también tiene un impacto directo en la capacidad de los profesionales de la salud para tomar decisiones más informadas y precisas. Con esta herramienta los resultados generados por el modelo son más accesibles y utilizables en la práctica diaria ofreciendo un soporte adicional para confirmar o reevaluar diagnósticos clínicos, mejorando así la calidad del cuidado del paciente.

Para el desarrollo de la aplicación (Figura 2) fue utilizado en Streamlit (2024) un framework de código abierto para Python. Los usuarios pueden acceder a la aplicación desde sus dispositivos móviles o computadoras, siempre que cuenten con un navegador web y utilicen una VPN desde Cuba, a través de la dirección <https://upd-ai.streamlit.app/>.

La interfaz de la aplicación “Cuidado inteligente del pie diabético” es intuitiva y fácil de usar. Incluye una barra lateral para ajustar el nivel de certeza de las predicciones, un botón para cargar imágenes (JPG, JPEG, PNG) y otro para iniciar la inferencia. Los resultados se presentan en imágenes anotadas con las áreas detectadas. Tras el procesamiento, se habilita un botón para exportar las predicciones en un archivo comprimido con las imágenes anotadas y un CSV con las coordenadas. Si se cargan varias imágenes, se puede seleccionar cuál visualizar.

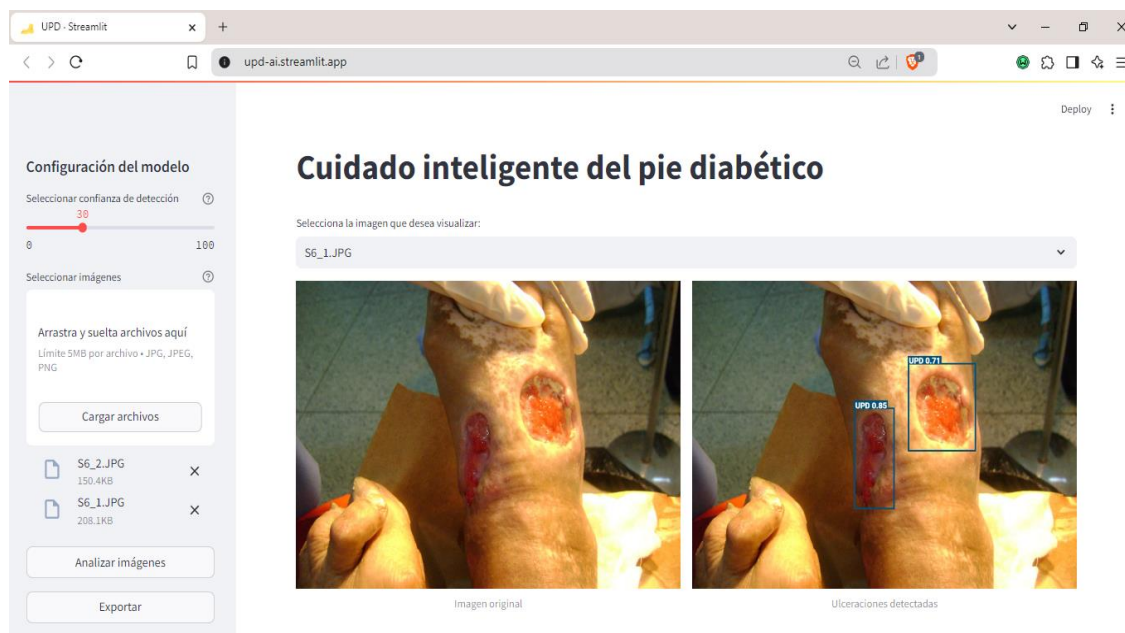


Figura 2. Interfaz de usuario de la aplicación.

CONCLUSIONES

Los modelos de detección de objetos basados en aprendizaje profundo implementados, como YOLOv8 y Faster R-CNN, así como variantes de ellos, demostraron ser herramientas eficaces para la identificación precisa de UPD, y el uso de técnicas de aumento de datos mejoró significativamente sus resultados.

YOLOv8l obtuvo métricas de rendimiento superiores, alcanzando un mAP50 de 0,799, lo que lo posiciona como el más eficaz entre las arquitecturas evaluadas. La evaluación de YOLOv8l en la base de casos nacionales mostró un mAP50 de 0,839, destacando su potencial en la práctica clínica para la detección de UPD.

La aplicación web desarrollada con Streamlit integra YOLOv8l de forma efectiva, ofreciendo una interfaz intuitiva para los profesionales de la salud y facilitando futuras mejoras e incorporación de clasificación de imágenes.

Evaluar el rendimiento en distintos tipos de piel es clave para medir la generalización del modelo. Dado que el entrenamiento usó mayormente pieles blancas, existe riesgo de sesgo en la detección de úlceras en pieles no blancas, afectando su precisión y equidad en entornos clínicos diversos.

AGRADECIMIENTOS

La investigación que da origen a los resultados presentados en la presente publicación recibió fondos de la Oficina De Gestión de Fondos y Proyectos Internacionales bajo el código PN223LH004-039 y fondos del Programa Territorial del CITMA Salud Humana bajo el código PT241VC001-009.

REFERENCIAS

- Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A. A. (2020). Albuementations: Fast and flexible image augmentations. *Information (Switzerland)*, 11(2). <https://doi.org/10.3390/info11020125>
- Cabal Mirabal, C. A., González Dalmau, E. R., Oramas Diaz, L., Herrera Martinez, L. S., Berlanga Acosta, J. A., & Fernández Montequín, J. I. (2019). Organización Mundial de la Propiedad Intelectual. Marco estereotáxico para extremidades. Patente de Cuba PCT/CU2016/050002, 2017. Acceso, 31(07).
- Cassidy, B., et al. (2021). The DFUC 2020 dataset: Analysis towards diabetic foot ulcer detection. *touchREVIEWS in Endocrinology*, 17(1), 5. <https://doi.org/10.17925/EE.2021.17.1.5>
- Chen, K., et al. (2019). MMDetection: Open MMLab Detection Toolbox and Benchmark.
- Dutta, A., Zisserman, A., & Gupta, A. (2024, September 19). VGG Image Annotator (VIA). GitHub. <https://github.com/ox-vgg/via>
- García Gómez, G., et al. (2019). Mejoramiento de contraste y segmentación en imágenes de úlceras del pie diabético. *Revista Cubana de Angiología y Cirugía Vascular*, 20.
- Jocher, G., Chaurasia, A., & Qiu, J. (2023, August 24). YOLO by Ultralytics. <https://github.com/ultralytics/ultralytics>
- Juarez-Salinas, J. J., Solis-Sánchez, L. O., Castañeda-Miranda, R., Ortiz-Rodriguez, J. M., & Rosales, H. G. (2014). Algoritmo de procesamiento digital de imágenes para la detección y evaluación de heridas de pie diabético. *Difu100ci@*, *Revista de difusión científica, ingeniería y tecnologías*, 8(1), 11–16.
- Liu, L., et al. (2020). Deep learning for generic object detection: A survey. *International Journal of Computer Vision*, 128(2), 261–318. <https://doi.org/10.1007/s11263-019-01247-4>
- Manchester Metropolitan University. (2024, July 2). 28th UK Conference on Medical Image Understanding and Analysis - MIUA. MIUA. <https://miua2024.github.io/#challenges>
- OpenMMLab. (2024, October 15). OpenMMLab: Open Source Computer Vision Toolbox. <https://openmmlab.com/>
- Sarmun, R., et al. (2024). Diabetic foot ulcer detection: Combining deep learning models for improved localization. *Cognitive Computation*, 16(3), 1413–1431. <https://doi.org/10.1007/s12559-024-10267-3>
- Streamlit. (2024, October 22). Streamlit: The fastest way to build and share data apps. <https://streamlit.io/>
- Sun, H., Heng, H., Liu, X., Geng, H., & Liang, J. (2023). Evaluation of the healing potential of short-term ozone therapy for the treatment of diabetic foot ulcers. *Frontiers in Endocrinology*, 14. <https://doi.org/10.3389/fendo.2023.1304034>
- Tekkeşin, A. İ. (2019). Artificial intelligence in healthcare: Past, present and future. *Anatolian Journal of Cardiology*, 22(Suppl 2), 8–9.
- Yap, M. H., et al. (2021). Deep learning in diabetic foot ulcers detection: A comprehensive evaluation. *Computers in Biology and Medicine*, 135, 104596. <https://doi.org/10.1016/j.compbiomed.2021.104596>

Copyright © 2025, Autores: Moya Cáceres, Edgar, Ruíz González, Yusely, García Lorenzo, María Matilde



Esta obra está bajo una licencia de Creative Commons Atribución-No Comercial 4.0 Internacional