Urine sediment classification using deep learning models

Lara Musuruana[1] and Mauro Pacchiotti [1]

[1] Centro de I+D de Ingeniería en Sistemas de Información, UTN-FRSF
{lmusuruana, mpacchiotti}@frsf.utn.edu.ar

Abstract. Urine sediment analysis is a fundamental practice in the clinical diagnosis of renal and urological pathologies. However, its manual examination is operator-dependent, time-consuming, and subject to subjective variability, while commercial automated systems are often cost-prohibitive for medium-scale laboratories. In this context, this paper addresses the application of deep learning techniques for the automatic identification and classification of particles in microscopic urine sediment images. Under the adaptation of the CRISP-DM methodology, various multiclass datasets are integrated through a format standardization process compatible with object detection architecture such as YOLO (You Only Look Once). Data treatment prioritizes the integrity of the original samples, avoiding techniques that could alter the natural class distribution. Class imbalance is addressed through the configuration of loss functions and data augmentation strategies applied during training. Performance evaluation is conducted using metrics such as F1-Score, Recall, and confusion matrices, allowing a detailed analysis of model behavior across different classes. The results obtained demonstrate satisfactory performance, suggesting that the use of pretrained models combined with fine-tuning strategies constitutes a promising approach for the automation of urine sediment analysis in clinical settings.

Keywords: urine sediment, deep learning, classification.

Clasificación de sedimentos urinarios mediante modelos de aprendizaje profundo

Lara Musuruana[1] y Mauro Pacchiotti [1]

[1] Centro de I+D de Ingeniería en Sistemas de Información, UTN-FRSF
{lmusuruana, mpacchiotti}@frsf.utn.edu.ar

Resumen. El análisis de sedimentos urinarios es una práctica fundamental en el diagnóstico clínico de patologías renales y urológicas. Sin embargo, su ejecución manual es dependiente del operador, consume tiempo y está sujeta a variabilidad subjetiva, mientras que los sistemas automatizados comerciales suelen resultar costosos para laboratorios de mediana escala. En este contexto, este trabajo aborda la aplicación de técnicas de aprendizaje profundo para la identificación y clasificación automática de partículas en imágenes microscópicas de sedimento urinario. Bajo una adaptación de la metodología CRISP-DM, se integran diversos conjuntos de datos multiclase mediante un proceso de estandarización de formatos compatibles con arquitecturas de detección de objetos como YOLO (You Only Look Once). El tratamiento de los datos prioriza la integridad de las muestras originales, evitando técnicas que pudieran alterar la distribución natural de las clases. El desbalance de clases se aborda mediante la configuración de las funciones de pérdida y las estrategias de aumento de datos aplicadas durante el entrenamiento. La evaluación del desempeño se realiza utilizando métricas como F1-Score, Sensibilidad y matrices de confusión, permitiendo un análisis detallado del comportamiento del modelo en distintas clases. Los resultados obtenidos evidencian un desempeño satisfactorio, lo que sugiere que el uso de modelos preentrenados junto con estrategias de ajuste fino constituye un enfoque prometedor para la automatización del análisis de sedimentos urinarios en contextos clínicos.

Palabras clave: sedimento urinario, aprendizaje profundo, clasificación.

1 Introducción

Las enfermedades renales y del tracto urinario afectan a millones de personas en todo el mundo y causan aproximadamente 830.000 muertes al año, según Yildirim et al. (2023). Una parte fundamental para el diagnóstico, monitoreo y posterior tratamiento de estas patologías es el análisis de sedimentos urinarios.

El término sedimento urinario se refiere a las partículas sólidas que se depositan en el fondo de una muestra de orina después de un proceso de centrifugación. El mismo se examina mediante microscopía, con el objetivo de identificar y cuantificar elementos como glóbulos rojos, glóbulos blancos, células epiteliales, entre otros. La presencia, cantidad y características de estos elementos pueden ofrecer información crítica sobre

el estado funcional del sistema urinario y la presencia de posibles alteraciones patológicas.

Este análisis puede realizarse de forma manual o automatizada. El método manual, ampliamente utilizado en laboratorios clínicos tradicionales, requiere una considerable inversión en términos de tiempo y esfuerzo por parte del profesional, quien debe identificar y clasificar visualmente los distintos elementos presentes en la muestra. Este trabajo, aunque efectivo, es altamente demandante, lo que puede retrasar la entrega de resultados precisos y afectar la eficiencia del diagnóstico. Existen equipos automatizados que optimizan este proceso y mejoran notablemente las limitaciones del análisis manual, pero su alto costo representa una barrera considerable. En muchos laboratorios de tamaño medio o pequeño, donde no se procesan grandes volúmenes de muestras, la inversión en estos equipos no se justifica económicamente. En este marco, la integración de la inteligencia artificial y, particularmente, del aprendizaje profundo en los flujos de trabajo clínicos representa una oportunidad para optimizar procesos que aún dependen en gran medida de la intervención humana (Blezek et al., 2021).

El presente trabajo busca contribuir al avance del conocimiento en el campo del análisis automatizado de imágenes microscópicas, evaluando el desempeño de modelos de aprendizaje profundo para automatizar y optimizar el procesamiento de muestras de sedimento urinario. Mediante una metodología robusta se confecciona un conjunto de datos y se ajusta una arquitectura YOLO con el fin de identificar y clasificar seis elementos de alta relevancia clínica: eritrocitos, leucocitos, células epiteliales, cristales, cilindros y levaduras.

El resto de este trabajo se organiza como sigue: la Sección 2 presenta fundamentos y trabajos relacionados, la Sección 3 describe la metodología propuesta, la Sección 4 desarrolla las pruebas y presenta los resultados obtenidos, y la Sección 5 expone las conclusiones y propone los trabajos futuros.

2 Trabajos Relacionados

Los primeros avances en el análisis de imágenes médicas se apoyan en modelos de aprendizaje profundo, más específicamente redes neuronales convolucionales (CNN). En este contexto, diversos estudios han demostrado que las CNN permiten extraer características de las imágenes que luego sirven a distintos modelos de clasificación (Litjens et al., 2017).

En particular, Liang et al. (2018) desarrollaron un sistema de extremo a extremo basado en CNNs para el reconocimiento de partículas en sedimento urinario. Su enfoque permitió identificar múltiples clases dentro de un mismo marco de aprendizaje, alcanzando una precisión media superior al 84%.

Posteriormente, Ji et al. (2019) proponen un método combinado que utiliza la arquitectura AlexNet —una red neuronal convolucional preentrenada en la clasificación de imágenes a gran escala— junto con un algoritmo de atención a las características del área (AFA). Utilizando un conjunto de datos de más de 300.000

imágenes, junto con aumento de datos y transferencia de aprendizaje, alcanzan una precisión del 97% y un tiempo de inferencia cercano a 6 ms por imagen.

Yildirim et al. (2023) presentan un enfoque híbrido que combina descriptores de textura LBP, (Local Binary Pattern) con una arquitectura ResNet50, una red neuronal convolucional preentrenada para la clasificación de imágenes. Finalmente, al evaluar diversos clasificadores, el mejor rendimiento se obtiene con SVM (Support Vector Machine), logrando una precisión del 96% sobre un conjunto de datos de 8.509 imágenes distribuidas en 8 clases. En la misma línea, Erten et al. (2023) proponen Swin-LBP, una arquitectura que integra Swin Transformer con descriptores de textura LBP. El sistema opera en un pipeline de cinco fases que abarca desde el preprocesamiento hasta la selección de características mediante NCA (Neighborhood Component Analysis), finalizando con una clasificación basada en SVM y votación mayoritaria. Esta propuesta utiliza un conjunto de datos de 12.330 imágenes distribuidas en 7 clases.

Posteriormente, Akhtar et al. (2024) introducen un enfoque híbrido que integra estrategias centradas en datos y en el modelo, mediante un proceso de limpieza iterativa del conjunto de datos y aumento de datos. La propuesta emplea una arquitectura en cascada basada en VGG19 —una red neuronal convolucional profunda— modificada con submódulos especializados. Sobre un conjunto de datos clínico de 12 clases, el sistema reporta una precisión micro y macro del 100% con un tiempo de inferencia de 61 ms por imagen.

Con el auge de los detectores de una sola etapa, la familia YOLO ha sido ampliamente adoptada en este campo. En particular, diversos estudios han aplicado variantes de YOLO para la detección de partículas urinarias en imágenes microscópicas. Por ejemplo, Alvarado Jaya (2024) evalúa un sistema que combina modelos YOLOv8 y MMDetection — recurso de código abierto basado en PyTorch para la detección de objetos y la segmentación de imágenes — con validación humana. Utilizando un conjunto de datos público de 5.376 imágenes con 7 clases, obtiene un mAP de 90,3% y 89,9% respectivamente, concluyendo que “el sistema de reconocimiento de imágenes desarrollado es extremadamente eficaz…” (Alvarado Jaya, 2024, p. 85) y destacando mejoras en eficiencia y calidad diagnóstica mediante la integración humano-IA.

A partir del análisis de los antecedentes revisados, se evidencian avances significativos en la clasificación de sedimentos urinarios. No obstante, persisten desafíos relevantes, entre los que se destacan el desbalance de clases y la variabilidad en la calidad de las imágenes, factores que impactan directamente en la robustez de los modelos propuestos. En este contexto, la presente investigación plantea un enfoque que aprovecha los beneficios de arquitecturas consolidadas, con el objetivo de alcanzar desempeños competitivos mediante el uso de modelos de pequeña escala y bajo requerimiento computacional. De este modo, se busca lograr un adecuado equilibrio entre eficiencia y precisión, favoreciendo su eventual implementación en entornos con recursos limitados.

3 Metodología

3.1 Marco de Trabajo

La estrategia metodológica implementada en este trabajo se fundamenta en el estándar CRISP-DM (Cross-Industry Standard Process for Data Mining), seleccionado por su robustez en la estructuración de proyectos de análisis de datos. Este marco proporciona un flujo operativo que garantiza un ciclo de ejecución ordenado e iterativo.

No obstante, con el propósito de alinear el flujo de trabajo a los objetivos específicos del estudio y a los requisitos del ámbito clínico, se ha diseñado una adaptación estructural de seis fases. Como se observa en la Figura 1, esta configuración incorpora una etapa inicial de exploración del dominio y una fase de cierre dedicada a la documentación y al desarrollo del prototipo de software. Las flechas bidireccionales entre las fases 2 (Preparación de los Datos), 3 (Modelado) y 4 (Evaluación) representan la naturaleza iterativa del proceso: el bucle entre las fases 2 y 3 permite retornar a la preparación de datos cuando durante el modelado se identifican inconsistencias o problemas en las anotaciones, mientras que el bucle entre las fases 3 y 4 habilita el reajuste de parámetros o la exploración de configuraciones alternativas cuando los resultados de la evaluación no alcanzan los criterios de desempeño establecidos. Las flechas externas representan iteraciones de mayor alcance, posibilitando el retroceso hasta las fases iniciales cuando los problemas detectados son de naturaleza más profunda, lo cual resulta necesario dadas las variaciones en calidad, contraste y condiciones de adquisición propias de las imágenes clínicas.

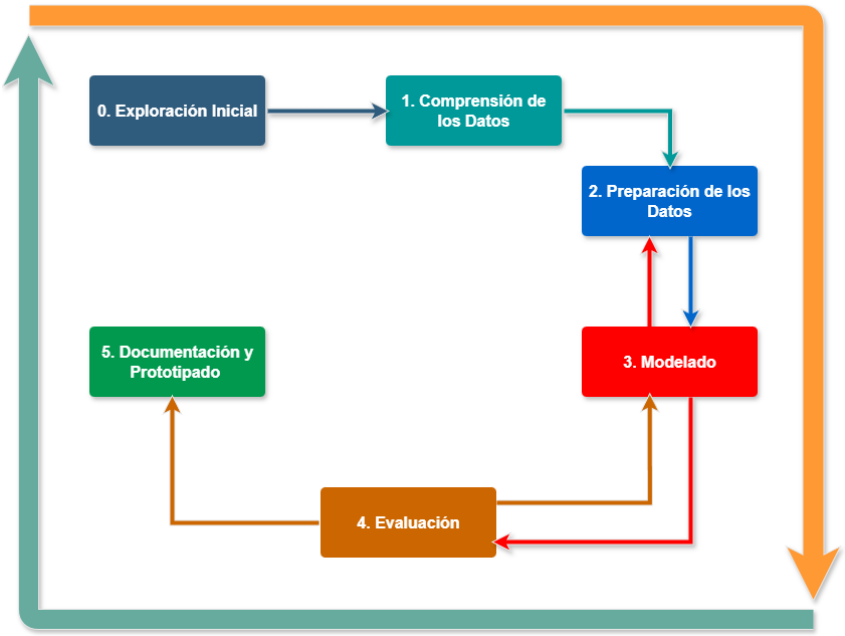

Fig. 1. Diagrama de la metodología propuesta, basada en el estándar CRISP-DM.

3.2 Conjunto de Datos

El análisis de sedimentos urinarios mediante microscopía conlleva desafíos tales como la baja densidad de partículas en determinadas muestras, la presencia de ruido de fondo y la variabilidad en las condiciones de iluminación durante la captura. Ante estas dificultades, se optó por conformar un conjunto de datos integrando fuentes de diversos orígenes. Por un lado, se utilizó el conjunto UMID (Urine Microscopic Image Dataset), propuesto por Goswami et al. (2021), el cual se compone de 367 imágenes con una resolución de 1280x720 píxeles. Este conjunto de datos identifica tres categorías de relevancia clínica: eritrocitos, leucocitos y células epiteliales. La mayor parte de las instancias se encuentran definidas mediante cuadros delimitadores, los cuales especifican las coordenadas mínimas y máximas en los ejes x e y para cada partícula individual. No obstante, aquellos elementos que presentan solapamientos fueron etiquetados con anotaciones de puntos. Estas marcas consisten en una coordenada única (x,y) situada en el centroide de la partícula, una técnica implementada por los autores para registrar la presencia y ubicación de cada elemento en áreas de alta densidad donde la superposición impide delimitar bordes precisos.

Complementariamente, se integró el repositorio de acceso público USE-Dataset (Liang et al., 2018). Dicho conjunto presenta 5.376 imágenes anotadas con una resolución de 800x600 píxeles, abarcando siete categorías diagnósticas fundamentales para el entorno clínico: cilindros, cristales, células epiteliales, núcleos epiteliales, eritrocitos, leucocitos y levaduras.

A partir de la disponibilidad de datos en las fuentes seleccionadas y la consulta a profesionales de laboratorio sobre su práctica diaria, se definió una categorización final de seis clases de interés. Esta selección se fundamenta en la frecuencia y relevancia diagnóstica de los elementos observados en el entorno clínico, estableciendo como categorías prioritarias a los eritrocitos, leucocitos, células epiteliales, cilindros, cristales y levaduras. En la Figura 2 se exhiben ejemplares de cada clase, mientras que en la Tabla 1 se detalla la distribución de las instancias según su origen, así como la composición final del conjunto de datos resultante.

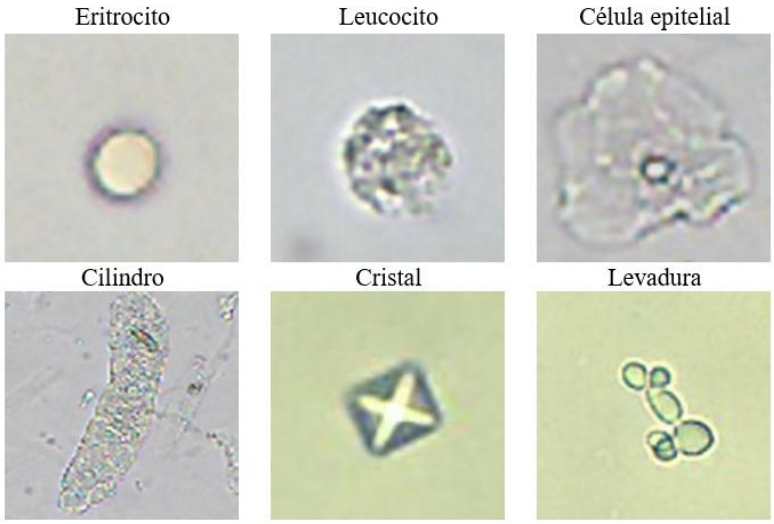

Fig. 2. Muestras representativas de las seis categorías de interés obtenidas del conjunto de datos integrado.

Tabla 1. Distribución de instancias por conjunto de datos de origen y total consolidado.

Clase	UMID-Dataset	USE-Dataset	Total
Eritrocitos	1.556	21.815	23.371
Leucocitos	986	6.169	7.155
Células epiteliales	437	6.862	7.299
Cilindros		3.662	3.662
Cristales		1.644	1.644
Levaduras		2.083	2.083

3.3 Preprocesamiento y Estandarización

La etapa de preprocesamiento consistió en la ejecución de tareas orientadas a la homogeneización de las diversas fuentes de datos. El objetivo fue garantizar que las imágenes, independientemente de su origen clínico, presenten un formato, resolución y estructura de etiquetado compatibles con la arquitectura de aprendizaje profundo seleccionada.

En lo que respecta al conjunto UMID, se identificó una limitación en sus anotaciones de puntos para la implementación del modelo YOLO. Dado que este algoritmo requiere cajas delimitadoras para su funcionamiento, las anotaciones que contienen únicamente el centroide de la partícula resultan insuficientes para el entrenamiento. Ante este inconveniente, se implementó un procedimiento de

generación de cajas delimitadoras artificiales, consistente en la creación de cajas cuadradas centradas en las coordenadas de los puntos originales, cuyas dimensiones fueron establecidas a partir del análisis estadístico de la morfología característica de cada clase (Figura 3).

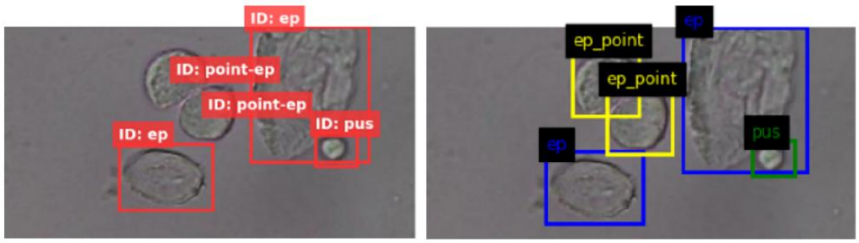

Fig. 3. La imagen de la izquierda muestra dos anotaciones de punto (point) pertenecientes a células epiteliales (ep), que en la imagen de la derecha fueron reemplazadas por cajas.

La estrategia adoptada contempló criterios diferenciados según la clase de partícula, con el fin de reducir el sesgo morfológico que introduciría el uso de un tamaño uniforme. Para los eritrocitos, se utilizó la mediana de las cajas disponibles en la misma imagen. Esta decisión se fundamenta en la morfología altamente homogénea de los eritrocitos: durante su maduración terminal, los reticulocitos pierden volumen y organelas residuales adoptando la forma de disco bicóncavo (Ovchynnikova et al., 2018), lo que resulta en un diámetro típico de entre 7.5 y 8.7 µm (Diez-Silva et al., 2010). En el caso de los leucocitos se aplicó un factor de escala de 0.9 sobre dicha mediana, justificado por su mayor dispersión morfológica respecto a los eritrocitos. Finalmente, para las células epiteliales se aplicó un factor de 0.5 sobre la mediana y el percentil 25 global en ausencia de anotaciones con caja, dado que en el conjunto de datos UMID estas células tienden a aparecer agrupadas y fueron anotadas con cajas únicamente cuando se presentaban de forma aislada, siendo estas visualmente más grandes que las agrupadas. En todos los casos, se recurrió a la mediana global de la clase cuando la imagen no contaba con instancias anotadas con caja.

Por otro lado, el procedimiento para el USE-Dataset consistió en la conversión del conjunto originalmente etiquetado en formato PASCAL VOC al estándar requerido por YOLO. En el formato PASCAL VOC, la información de cada imagen se almacena en archivos XML que describen las cajas delimitadoras mediante coordenadas absolutas en píxeles, definiendo los puntos específicos de las esquinas. Para adaptar estos datos a la arquitectura YOLO, se realizó una transformación hacia archivos de texto plano donde cada objeto se identifica por un índice numérico seguido de coordenadas relativas y normalizadas.

Debido a la disparidad en las dimensiones originales de las imágenes, se procedió a su normalización a una resolución de 800 x 800 píxeles. En la Figura 4 se presenta un ejemplo de imagen en su formato original. Este ajuste permitió la integración uniforme de las fuentes de datos, logrando un equilibrio entre la preservación de los rasgos morfológicos de las partículas pequeñas y la eficiencia computacional del modelo.

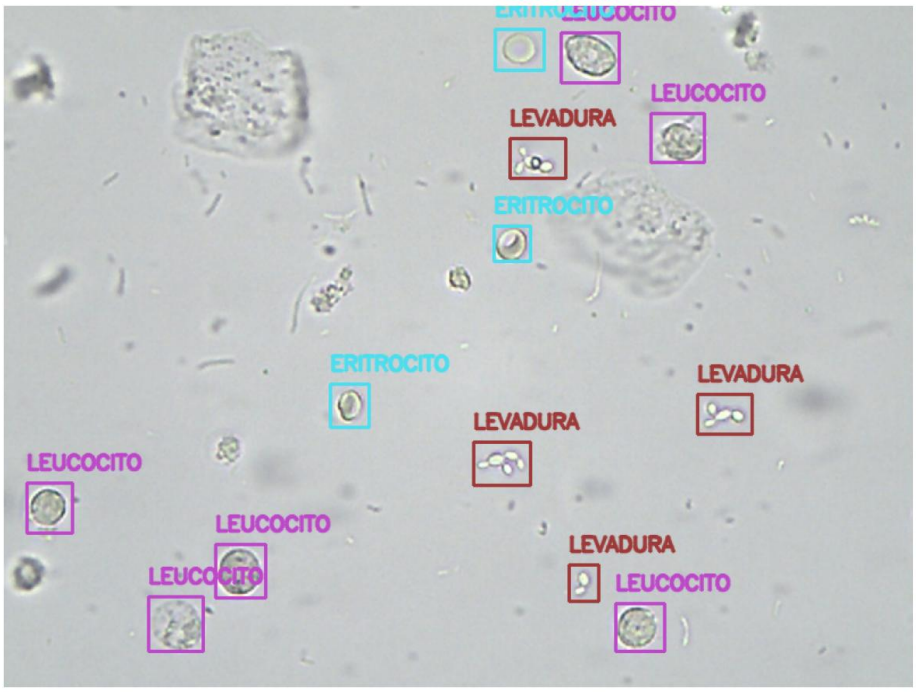

Fig. 4. Representación de elementos sedimentarios etiquetados en su formato original.

3.4 Selección y Ajuste de Modelos

Para este trabajo se seleccionó la arquitectura YOLOv8 en su variante Small (YOLOv8s), que integra en una única red neuronal convolucional la extracción de características, la localización y la clasificación de partículas, analizando la imagen completa en una sola pasada. Se adoptó la variante Small por ofrecer mayor profundidad de red que la variante Nano, sin incurrir en los requerimientos computacionales de las variantes Medium o Large, lo que la hace adecuada para entornos con recursos limitados y para capturar morfologías complejas sin demandar hardware excesivo.

Para el manejo del desbalance de clases se implementaron estrategias de aumento de datos y se configuraron explícitamente los hiperparámetros de entrenamiento. Se utilizó el optimizador SGD con momento de 0.937 y decaimiento de pesos de 0.0005, tasa de aprendizaje inicial $lr_0 = 0.01$ y final $lrf = 0.01$, con una fase de calentamiento de 3 épocas. La función de pérdida combinó tres componentes: pérdida de caja (7.5), clasificación (0.5) y distribución focal (1.5). Las técnicas de aumento incluyeron volteo horizontal aleatorio, variaciones HSV, escalado, traslación y composición en mosaico. El entrenamiento se realizó en la plataforma Kaggle[1] con GPU NVIDIA Tesla T4, con un tamaño de imagen de 800 píxeles y lote de 16 muestras. Se configuró un

[1] https://www.kaggle.com/

entrenamiento inicial de 50 épocas; al analizar las curvas de aprendizaje, que indicaban margen de mejora sobre el conjunto de validación, se realizó un segundo entrenamiento con un límite de 70 épocas. En ambos casos se implementó parada temprana con paciencia de 10 épocas. La Tabla 2 presenta los resultados de ambos entrenamientos, que llevaron a seleccionar el modelo de 70 épocas por su mejor desempeño en Sensibilidad y mAP50.

Tabla 2. Resultados globales en el conjunto de validación para ambos entrenamientos.

Épocas	Batch Size	Img.Size	Sensibilidad	mAP50
50	16	800x800	.837	.900
70	16	800x800	**.859**	**.907**

4 Evaluación y Resultados Preliminares

La evaluación del modelo se fundamenta en métricas orientadas a garantizar la fiabilidad diagnóstica. Se emplea la Sensibilidad (Recall) para asegurar que no se omitan partículas de relevancia clínica, el F1-Score para analizar el desempeño en clases con menor representación, y el mAP50-95 como medida más exigente de la precisión en la localización espacial de las detecciones. Complementariamente, la Matriz de Confusión permite identificar patrones de error entre clases morfológicamente similares.

Los indicadores se obtuvieron a partir del conjunto de prueba correspondiente al 10% del conjunto de datos final, derivado de una división aleatoria a nivel de imagen. Cabe destacar que, dado el desbalance de clases y que cada imagen puede contener múltiples instancias en cantidades variables, la distribución por clase no es uniforme entre splits, lo cual debe considerarse al interpretar los resultados.

Tal como ilustra la curva Sensibilidad-Confianza (Recall-Confidence) en la Figura 5, se evidencia que el modelo mantiene una alta sensibilidad en un amplio rango de umbrales, destacándose especialmente en la zona del “codo” de la curva, donde logra un buen equilibrio entre Recall y nivel de confianza. En este punto, el modelo conserva una elevada capacidad de detección sin requerir umbrales excesivamente bajos.

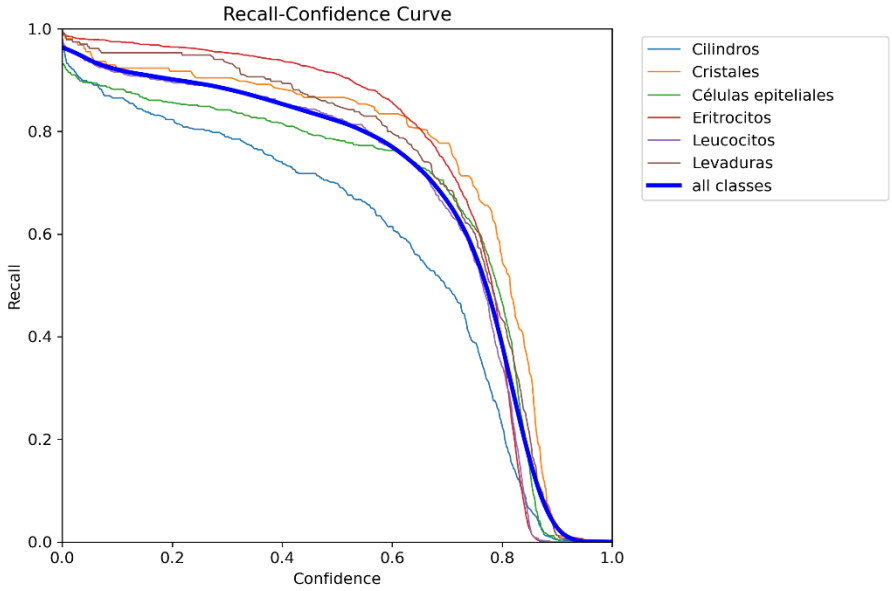

Fig. 5. Evaluación de exhaustividad mediante la curva Recall-Confidence.

Este rendimiento se mantiene estable para la mayoría de las categorías, las cuales conservan un Recall superior a 0.8 incluso ante umbrales de confianza cercanos a 0.6. Particularmente, los eritrocitos y cristales se consolidan como las clases más robustas, comportamiento que puede atribuirse a la consistencia de sus características morfológicas que facilita una extracción de patrones más precisa. No obstante, el análisis individualizado permite identificar que los cilindros presentan una mayor complejidad diagnóstica para el modelo. Mientras que a un nivel de confianza de 0.4 la mayoría de las partículas mantienen una sensibilidad cercana al 90%, la detección de cilindros se ubica aproximadamente en el 75%. Esta disparidad puede responder a la naturaleza de estos elementos, los cuales suelen presentar mayores dimensiones, una apariencia traslúcida y bordes menos definidos en comparación con otros elementos.

Complementariamente, la evaluación del equilibrio entre la precisión y la sensibilidad se analiza mediante la curva F1Score-Confianza (F1Score-Confidence) de la Figura 6. Los resultados demuestran una estabilidad notable, alcanzando un valor máximo de F1 de 0.86 para un umbral de confianza de 0.384. Esta amplia meseta observada en la métrica global sugiere que el modelo mantiene un desempeño consistente. En concordancia con los hallazgos previos, la clase de eritrocitos destaca con el índice de F1 más elevado, lo que ratifica la eficacia del modelo en la discriminación de este elemento frente al ruido del fondo. Por otro lado, la categoría de cilindros presenta el desempeño más desafiante para el modelo, con una curva que se sitúa persistentemente por debajo del promedio global. Bajo este umbral óptimo, la

Tabla 3 presenta el desglose de Precisión, Sensibilidad, F1-Score, mAP50 y mAP50-95 para cada clase individual sobre el conjunto de prueba.

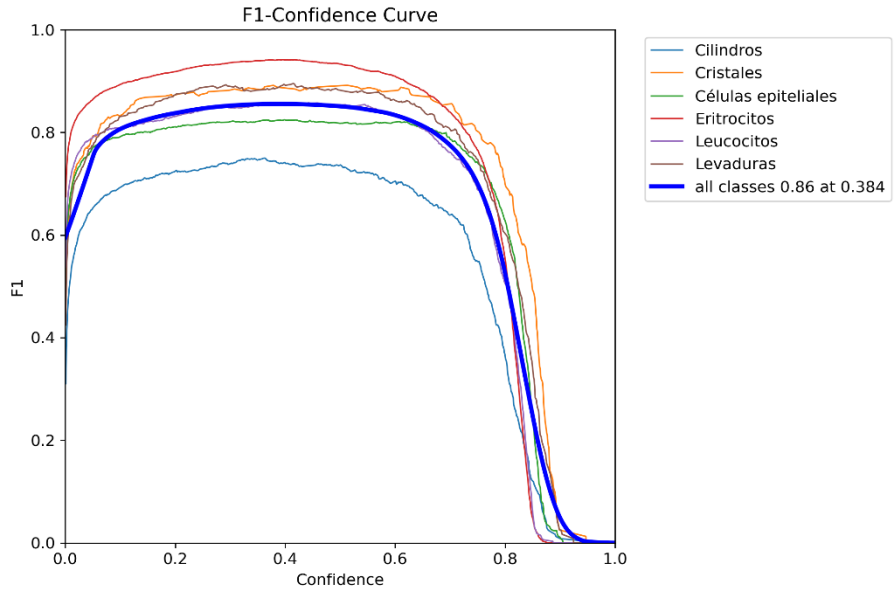

Fig. 6. Análisis del equilibrio entre precisión y sensibilidad a través de la curva F1-Confidence.

Tabla 3. Métricas de desempeño por clase en el conjunto de prueba.

Clase	Precisión	Sensibilidad	F1-Score	mAP50	mAP50-95
Eritrocitos	.942	.940	.941	.972	.574
Leucocitos	.850	.863	.856	.905	.500
Células epiteliales	.829	.818	.823	.872	.518
Cilindros	.736	.745	.742	.804	.414
Cristales	.891	.885	.889	.939	.563
Levaduras	.876	.898	.886	.948	.600
Global	.854	.859	.856	.903	.579

Para profundizar en el análisis de resultados, la matriz de confusión normalizada de la Figura 7 permite detallar los errores y aciertos obtenidos. Esta fue generada bajo un umbral de confianza de 0.25, valor por defecto que permite una inspección exhaustiva del comportamiento del modelo al ser más permisivo en la aceptación de detecciones, reduciendo así la omisión de partículas y facilitando la identificación de patrones de confusión entre clases. Los resultados corroboran la alta especificidad en la

identificación de eritrocitos y levaduras, con una precisión diagnóstica del 97% y 95% respectivamente. Un hallazgo relevante se observa en la relación entre el modelo y el ruido de fondo: la columna de falsos positivos revela que un porcentaje significativo de las detecciones de eritrocitos (27%) y cilindros (24%) corresponden en realidad a elementos del fondo que el modelo interpreta como partículas, mientras que la fila inferior muestra que el 18% de los cilindros reales y el 14% de las células epiteliales son omitidos, siendo clasificados como fondo. Este patrón indica que el principal desafío del modelo reside en la discriminación de partículas de bajo contraste respecto al ruido de fondo, más que en la confusión entre tipos celulares, lo que orienta las líneas de mejora futuras hacia estrategias de preprocesamiento y aumento de datos que refuercen dicha distinción.

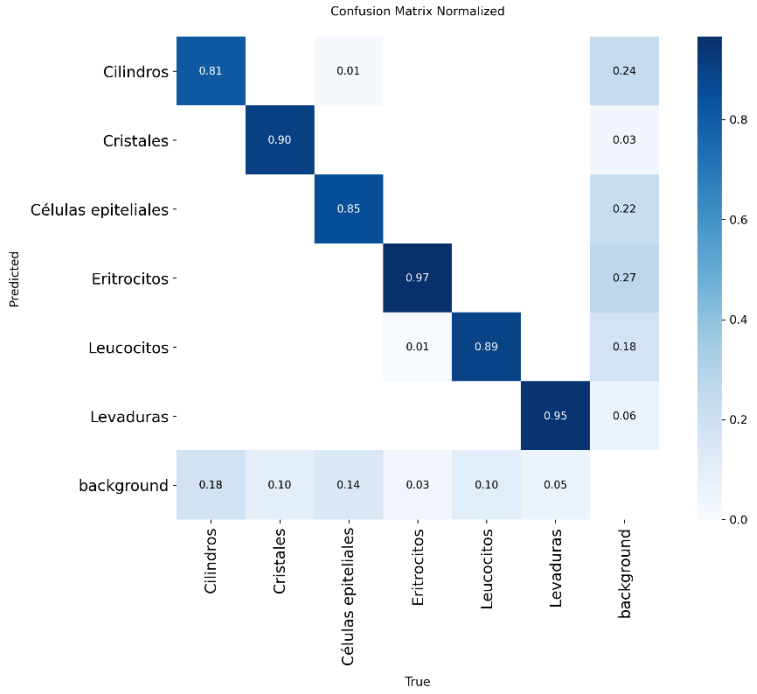

Fig. 7. Matriz de confusión normalizada por clase real (True).

En síntesis, los resultados obtenidos validan la capacidad del modelo para la detección efectiva de elementos presentes en sedimentos urinarios, alcanzando una sensibilidad global de alto impacto clínico. La evidencia sugiere que, si bien la arquitectura ha logrado robustez en la clasificación de partículas con rasgos definidos, persisten desafíos en la discriminación de elementos de bajo contraste frente al ruido del fondo. No obstante, el alto nivel de acierto en la identificación de categorías críticas demuestra que el modelo planteado constituye una base sólida y confiable para aportar a la automatización del análisis de sedimentos urinarios.

5 Conclusiones y Trabajos Futuros

Los resultados obtenidos en esta investigación confirman que las soluciones basadas en aprendizaje profundo constituyen una alternativa prometedora para la automatización del análisis de sedimentos urinarios. Se han logrado desempeños satisfactorios mediante un modelo de pequeña escala, lo que habilita la continuidad del desarrollo mediante prototipos y la ampliación de las pruebas con datos reales.

No obstante, el estudio identifica ciertas limitaciones que deben ser consideradas para su implementación en entornos reales. Entre ellas, se destaca la sensibilidad del modelo ante la variabilidad en la calidad de las imágenes procesadas, donde factores como el ruido de fondo y el bajo contraste de ciertos elementos pueden dificultar una discriminación precisa. Estas restricciones técnicas marcan la pauta para las próximas etapas del estudio, las cuales se centran en fortalecer la robustez frente a condiciones de captura no ideales.

Como parte de las líneas de trabajo futuro, se está avanzando en la incorporación de un conjunto de datos validado por personal de laboratorio, lo que permitirá entrenar el modelo con imágenes de la misma calidad técnica que las del entorno real de operación. Del mismo modo, se proyecta el diseño de algoritmos específicos para mejorar la identificación de cilindros, junto con la ampliación de la capacidad del modelo para reconocer otros elementos de interés clínico como bacterias, parásitos, espermatozoides y variedades más complejas de cristales y células. Finalmente, se explorará la integración directa del software con equipos de microscopía existentes, consolidando una herramienta que responda a las necesidades de la práctica clínica diaria.

References

Alvarado Jaya, H. G. (2024). Eficiencia del sistema de reconocimiento de imágenes en el análisis de sedimento urinario frente al método de conteo y categorización manual en laboratorios clínicos [Tesis de maestría, Escuela de Posgrado Newman]. Tacna, Perú.

Akhtar, S., et al. (2024). An optimized data and model centric approach for multi-class automated urine sediment classification. IEEE Access, 12, 59500–59520. https://doi.org/10.1109/ACCESS.2024.3385864

Blezek, D. J., Olson-Williams, L., Missert, A., & Korfiatis, P. (2021). AI Integration in Clinical Workflow. Journal of digital imaging, 34(6), 1435–1446. https://doi.org/10.1007/s10278-021-00525-3

Diez-Silva, M., Dao, M., Han, J., Lim, C. T., & Suresh, S. (2010). Shape and biomechanical characteristics of human red blood cells in health and disease. MRS Bulletin, 35(5), 382–388. https://doi.org/10.1557/mrs2010.571

Erten, M., Barua, P. D., Tuncer, I., et al. (2023). Swin-LBP: A competitive feature engineering model for urine sediment classification. Neural Computing and Applications, 35, 21621–21632. https://doi.org/10.1007/s00521-023-08919-w

Goswami, D., Aggrawal, H. O., Gupta, R., y Agarwal, V. (2021). Urine microscopic image dataset. arXiv. https://doi.org/10.48550/arXiv.2111.10374

Ji, Q., Li, X., Qu, Z., & Dai, C. (2019). Research on urine sediment images recognition based on deep learning. IEEE Access, 7, 166711–166720. https://doi.org/10.1109/ACCESS.2019.2953775

Liang, Y., Tang, Z., Yan, M., & Liu, J. (2018). Object detection based on deep learning for urine sediment examination. Biocybernetics and Biomedical Engineering, *38*(3), 661–670.

Liang, Y., Zhang, Y., Lian, C., & Kang, R. (2018). An end-to-end system for automatic urinary particle recognition with convolutional neural network. Journal of Medical Systems, 42(9). https://doi.org/10.1007/s10916-018-1011-8

Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Ovchynnikova, E., Aglialoro, F., von Lindern, M., & van den Akker, E. (2018). The shape shifting story of reticulocyte maturation. Frontiers in Physiology, 9, 829. https://doi.org/10.3389/fphys.2018.00829

Yildirim, M., Bingol, H., Cengil, E., Aslan, S., & Baykara, M. (2023). Automatic classification of particles in the urine sediment test with the developed artificial intelligence-based hybrid model. Diagnostics, 13(7), 1299