

Segmentación y clasificación de cabezas de espermatozoides para diferenciar su normalidad utilizando deep learning

Pilar Nilo V.^{1,3}, Víctor Castañeda^{2,3}, Carlos F. Navarro^{1,3}

pilar.nilo@ug.uchile.cl

¹Departamento de Ingeniería Eléctrica, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile.

³Technology for Medical Image-Based Clinical Support Systems Lab(Tech4Medics-Lab), Facultad de Medicina, Universidad de Chile.

²Departamento de Tecnología Médica, Facultad de Medicina de la Universidad de Chile.

Introducción

La infertilidad afecta al 15% de las parejas a nivel mundial (OMS) [1], con un tercio de los casos vinculados a factores masculinos, principalmente alteraciones en la morfología espermática. El análisis morfológico tradicional es manual, subjetivo y lento, mientras que los sistemas ASAC son costosos y de acceso limitado. La 6ª edición del manual de la OMS enfatiza evaluaciones estandarizadas. Las tecnologías de deep learning ofrecen soluciones objetivas, precisas y accesibles para automatizar este proceso.

Objetivo

Desarrollar y evaluar algoritmos de deep learning para segmentar y clasificar cabezas de espermatozoides en imágenes de microscopía óptica de campo claro teñidas con Hematoxilina-Eosina, considerando baja resolución espacial, con el fin de diferenciar su normalidad.

Materiales y métodos

- Base de datos:** SCIAN (Universidad de Chile) con 2665 imágenes de 40 x 40 píxeles. [2]
- Modelos de Segmentación:** U-Net

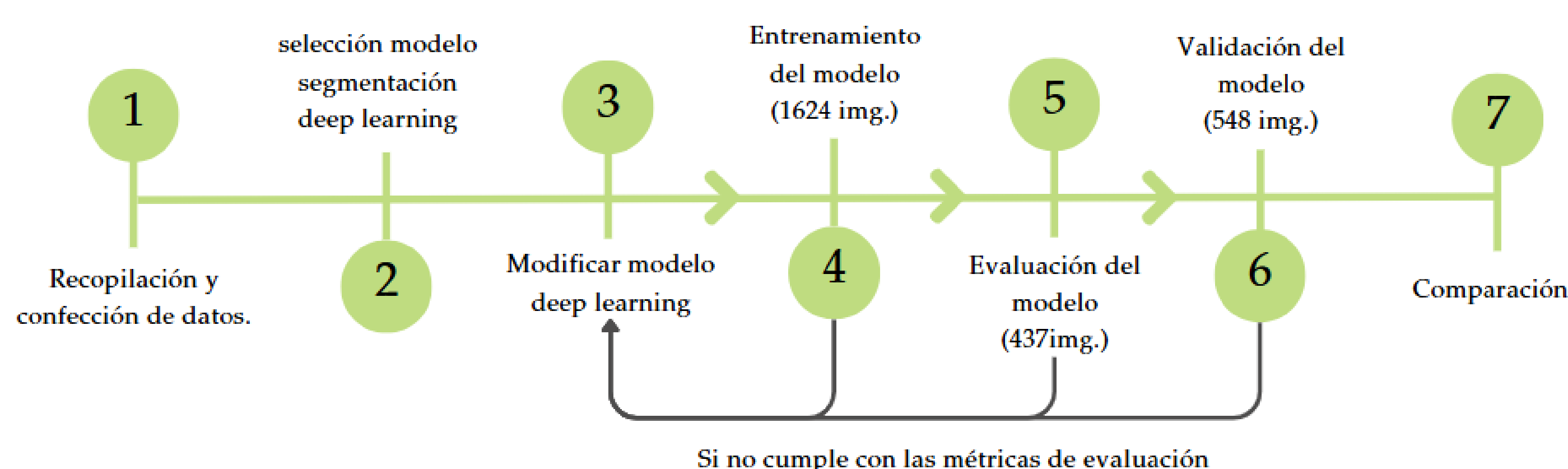


Figura 1: Flujo de segmentación.

- Modelos de clasificación:** ShuffleNetV2, CNN, ResNet-18, Red combinada de espacios latentes con características morfológicas.
- Funciones de perdida clasificación:** CrossEntropyLoss, BCEWithLogitsLoss y Focal Loss.

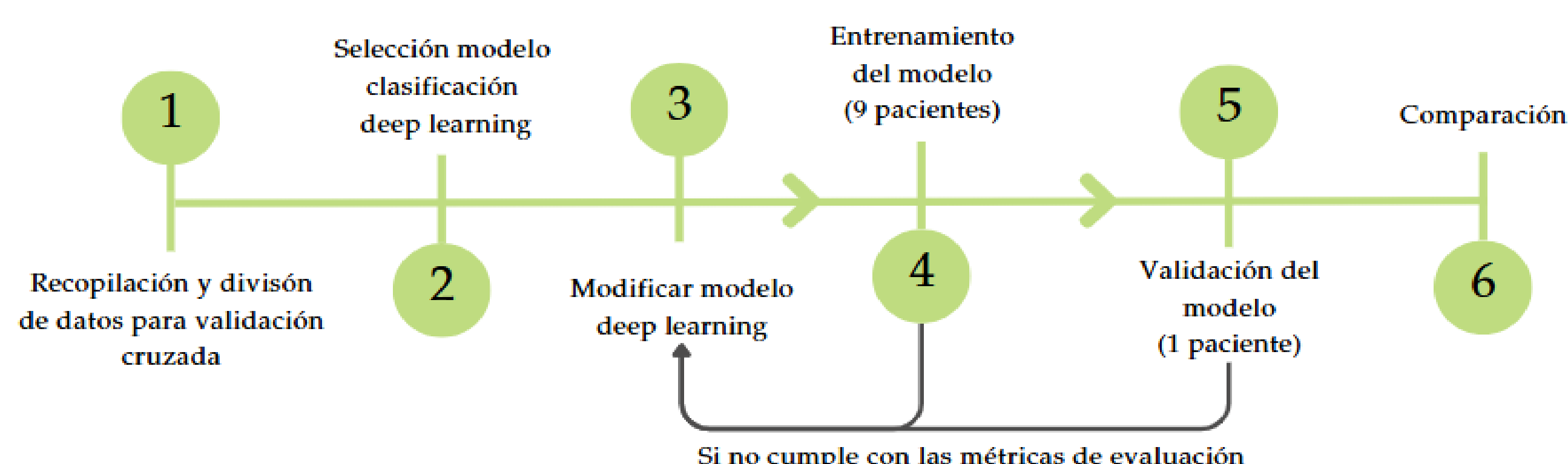


Figura 2: Flujo de clasificación.

Resultados

El modelo de segmentación U-Net logro un coeficiente DICE de 0,9530, un Intersection over Union de 0,91 y distancia Hausdorff de 1,9 píxeles, acercándose a las métricas reportadas por el estado del arte (DICE 0,96) sin preentrenamiento [3]. Este rendimiento destaca la robustez de U-Net para aprender características relevantes en la base de datos SCIAN, aunque errores locales en contornos, debido a la baja resolución (40 x 40 píxeles) y máscaras imperfectas (ej. superposición de espermatozoides no segmentados), afecta ligeramente la precisión.

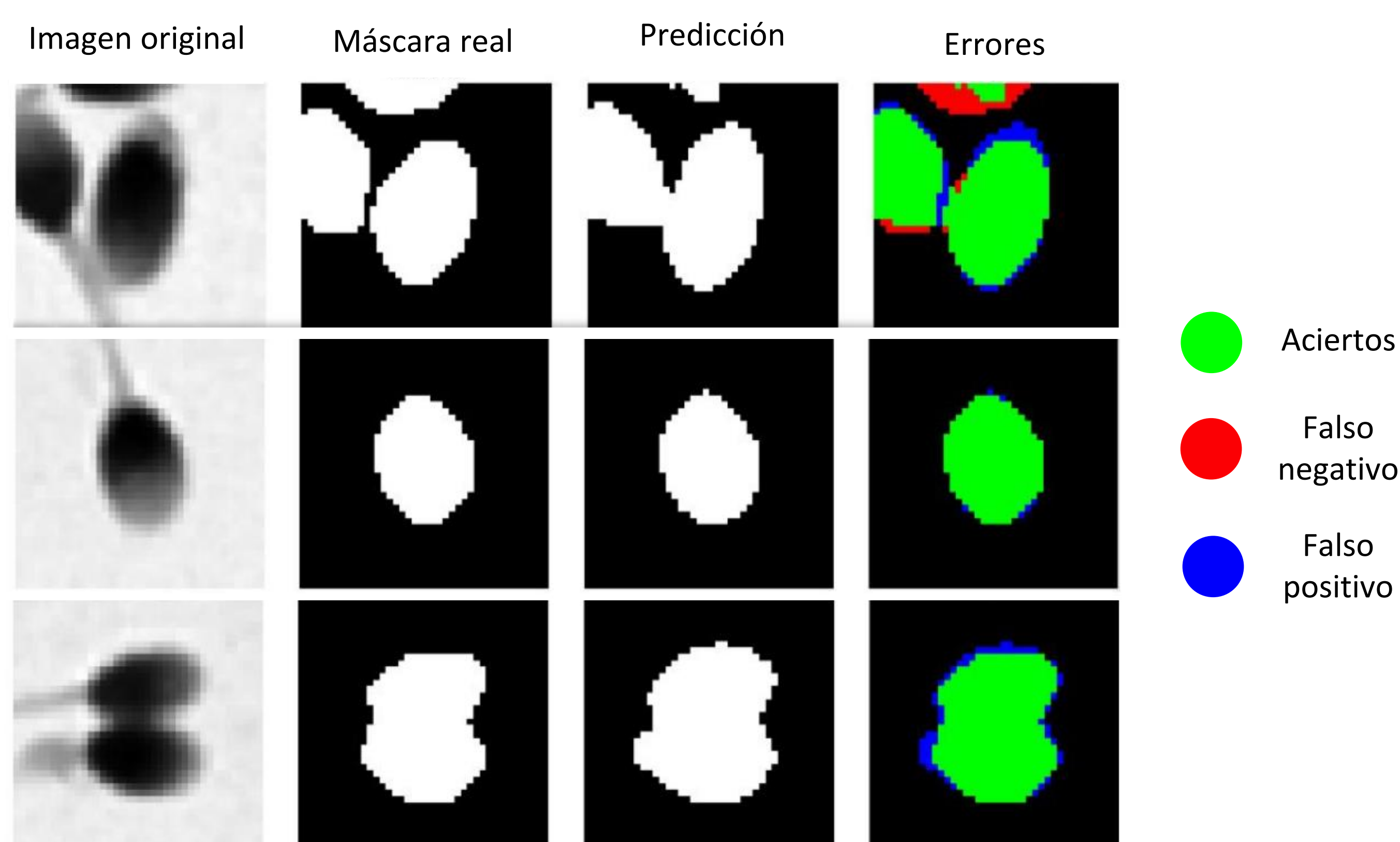


Figura 3: Ejemplo de resultados de segmentación.

Resultados

Tabla 1: Métricas de evaluación en conjunto de validación y evaluación del modelo U-Net.

Conjunto	Coeficiente DICE	IoU	Distancia Hausdorff
Validación	0,9471 ± 0,0584	0,9038 ± 0,0783	2,3705 ± 3,1628
Evaluación	0,9530 ± 0,0491	0,9132 ± 0,0660	1,9104 ± 2,6166

Tabla 2: Métricas de evaluación del modelo U-Net diseñado y el estado del arte.

Tipo de red	Coeficiente DICE
U-Net con Transfer Learning	0,96
U-Net Diseñada	0,953
U-Net	0,92
Mask R-CNN con Transfer Learning	0,92
Mask R-CNN	0,87

En clasificación, ShuffleNetV2 alcanzó un accuracy de 0,77, superando a la CNN personalizada, ResNet-18 y la red combinada de espacios latentes con características morfológicas. Su diseño eficiente optimiza la generalización en datasets pequeños como SCIAN. La función de pérdida Focal Loss ($\alpha = 0,75$; $\gamma = 4,0$) fue la más efectiva para el desbalance de clases, superando a CrossEntropyLoss y BCEWithLogitsLoss. Sin embargo, la aplicabilidad de ShuffleNetV2 a bases como HuSHeM o SMIDs requiere preprocesamiento adicional debido a diferencias en tinciones (Método Diff-Quick y Hematoxilina-Eosina modificada, respectivamente). Comparado con el estado del arte (ej: SHMC-Net [4] y Meta-classifier [5]), shuffleNetV2 tiene menor precisión por la falta de preentrenamiento y aumento de datos, pero su eficiencia lo posiciona como una herramienta prometedora de apoyo en laboratorios, aunque no alcanza estándares clínicos ($>90\%$ accuracy).

Tabla 3: Métricas de evaluación del mejor modelo de clasificación diseñado y el estado del arte.

Algoritmo	Accuracy	Precision	Recall	F1-Score
ShuffleNetV2	0,77	0,815	0,77	0,78
SHMC-Net	0,869	0,867	0,834	0,85
Meta-classifier	0,862	0,7695	0,9177	0,8124

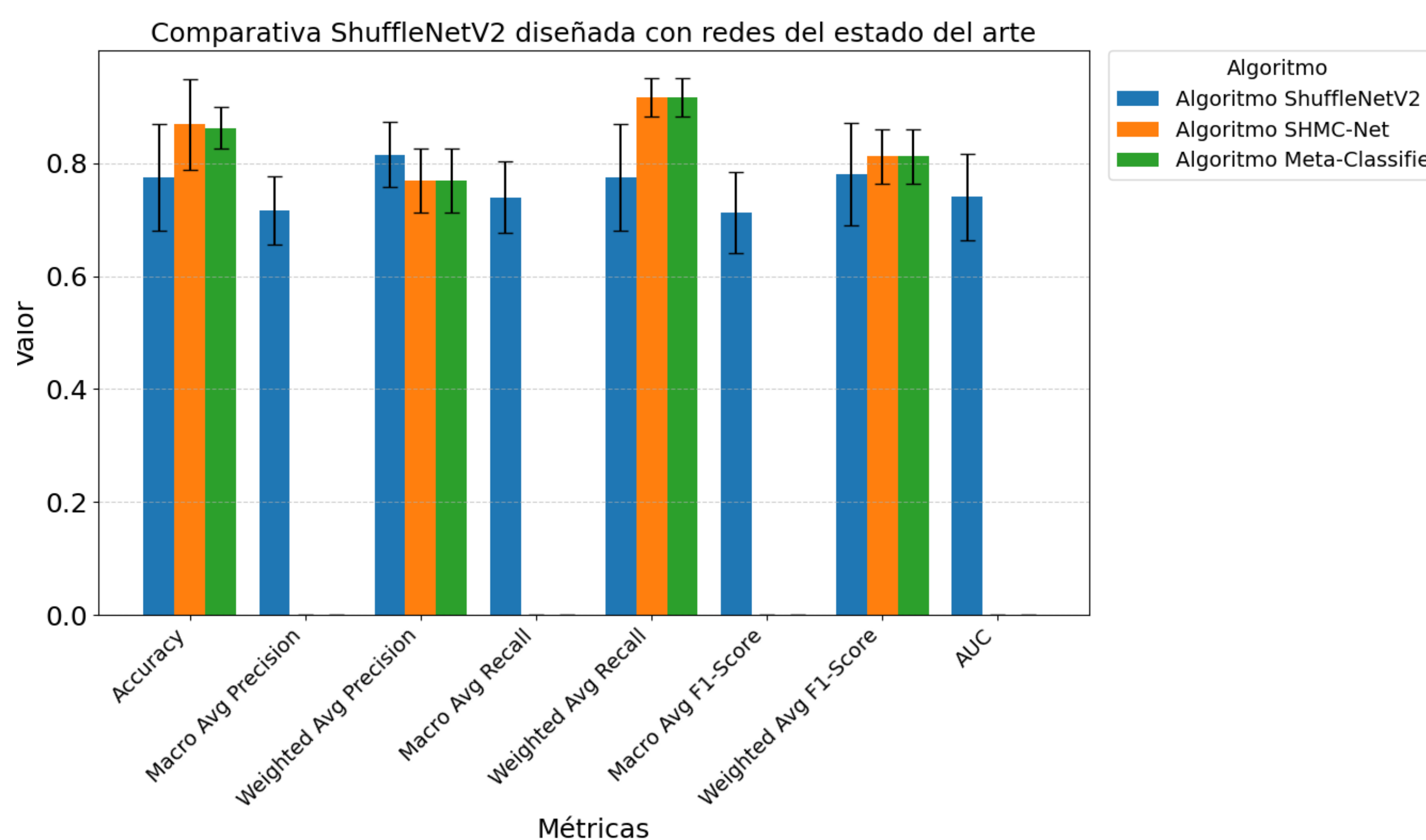


Figura 4: Comparación de métricas de evaluación entre el mejor modelo de clasificación diseñado y el estado del arte.

Conclusión

Los algoritmos de deep learning implementados muestran un alto rendimiento en la segmentación de espermatozoides, con U-Net apto para uso clínico. Sin embargo, la clasificación requiere de mejoras para alcanzar estándares clínicos. Este trabajo valida el potencial de deep learning para automatizar el diagnostico de infertilidad masculina, ofreciendo una alternativa objetiva y accesible al análisis tradicional.

Referencias

- [1] Cui, W., "Mother of nothing: the agony of infertility", Bull World Health Org, no. 12, pp. 881–882, 2010, doi:10.2471/BLT.10.011210.
- [2] Chang, V., Garcia, A., Hirschfeld, N., y Härtel, S., "Gold-standard for computer-assisted morphological sperm analysis", Computers in Biology and Medicine, vol. 83, pp. 143–150 2017, doi:10.1016/j.compbiomed.2017.03.004.
- [3] Marin, R. y Chang, V., "Impact of transfer learning for human sperm segmentation using deep learning", Computers in Biology and Medicine, 2021, doi:10.1016/j.compbiomed.2021.104687.
- [4] Sapkota, N., Zhang, Y., Li, S., Liang, P., y otros., "Shmc-net: A mask-guided feature fusion network for sperm head morphology classification", en 2024 IEEE International Symposium on Biomedical Imaging (ISBI), 2024, doi:10.1109/ISBI56570.2024.10635339.
- [5] Spencer, L., Fernando, J., Akbaridoust, F., Ackermann, K., y Nosrati, R., "Ensembled deep learning for the classification of human sperm head morphology", Advanced Intelligent Systems, 2022, doi:10.1002/aisy.202200111.