1. (Ronneberger, Fischer, & Brox, 2015) Propusieron una arquitectura supervisada, a la cual denominaron U-Net, basada en técnicas de Deep Learning, específicamente en redes neuronales convolucionales, para realizar segmentación semántica de imágenes de índole médico. Se utilizaron tres **datasets**; el primero contiene 30 imágenes del *EM segmentation challenge,* el segundo contiene 35 imágenes y es parte del *ISBI cell tracking challenge* *2014 and 2015,* el tercer dataset (*DIC-HeLa*) contiene 20 imágenes. La **metodología** consiste en una rama de codificación, la cual obtiene características de la imagen original y va reduciendo su tamaño, y una rama de expansión, en la cual se utiliza una convolución inversa para expandir la imagen y combinar la información espacial con las características obtenidas. Para los dos últimos datasets se utilizó aumento de datos para hacer más robusto el modelo. Los **resultados** en el primer dataset, acorde a las métricas *warping error* y *rand error,* fueron de 0.0003529 y 0.0382 respectivamente, para el segundo y tercer dataset se evaluó la métrica IOU, obteniendo valores del 92% y 77.5% respectivamente. En el trabajo original se específica contra que algoritmos se probó la U-Net, la cual obtuvo los mejores resultados en todos los datasets.
2. (Milletari, Navab, & Ahmadi, 2016)Se propone un modelo supervisado el cual es una generalización de la U-net para procesar imagen médicas volumétricas, aunque la arquitectura base es la misma se cambian los kernels de las capas convolucionales para extraer información en tres dimensiones. El **dataset** contiene 80 imágenes de resonancia magnética de próstata, el cual forma parte de *PROMISE2012 challenge dataset.* Como parte de la **metodología** se utilizó aumento de datos para hacer más robusto el modelo, utilizando 50 imágenes para el entrenamiento, al final de la capa *soft-max* se consideran los valores mayores a 0.5 como parte de la segmentación. Se utilizó un *momentum* de 0.9, un *learning rate* de 0.0001. Los **resultados** se evaluaron en términos de las métricas *Dice Coefficent, Hausdorff Distance* y el valor obtenido en el *challenge data*, obteniendo valores de 0.869, 5.71 y 82.39, respectivamente. Estos valores son similares a los obtenidos por otros algoritmos, obteniendo un menor tiempo de cómputo.
3. (Xia & Kulis, 2017)Se propone una arquitectura en forma de W, utilizando dos U-Nets concatenadas, para formar un modelo no supervisado el cual no requiere etiquetas previas para lograr la segmentación. El **dataset** utilizado para el entrenamiento fue el *PASCALVOC2012*, el cual contiene 11,530 imágenes (de las cuales solo 6,929 están segmentadas). La **metodología** consiste en obtener una primera segmentación con la U-Net que codifica la imagen y posteriormente refinar dicho resultado con la segunda U-Net. Se utilizaron los datasets BSDS300 y BSDS500, con 300 y 500 imágenes respectivamente, para evaluar la segmentación obtenida. Los **resultados** obtenidos se evaluaron a través de las métricas *Variation of Information, Probabilistc Rand Index* y *Segmentation Covering*. Se reportaron los valores utilizando *Optimal Dataset Scale (ODS)* y *Optimal Image Scale (OIS)* para cada métrica. Se obtuvieron parejas de valores para cada métrica, correspondientes al ODS y OIS, de (0.60,0.65), (0.82, 0.86), (1.63, 1.45) para el dataset BSDS300 y (0.59, 0.64), (0.82, 0.85), (167, 147) para el dataset BSDS500. Los cuales se compararon contra 7 algoritmos para cada dataset.
4. (He, Zhang, Ren, & Sun, 2016)Proponen una red profunda añadiendo conexiones directas entre la entrada y salida en las capas convolucionales, a la cual se nombró ResNet. Se probó la arquitectura con **dataset** ImageNet2012 *classification dataset*, el cual consiste en 1,000 clases en más de 1.28 millones de imágenes. Como parte de la **metodología** se hicieron varias pruebas con la arquitectura, variando el número de capas, además de utilizar un modelo sin conexiones residuales. Posteriormente se utilizó el dataset CIFAR-10, el cual contiene 60,000 imágenes de 10 clases para comparar el modelo propuesto. Como **resultado** se obtuvo un error de clasificación del 6.43%, para la ResNet con 110 capas. Se comparó el modelo contra diferentes versiones de la propia ResNet, así como otros 6 trabajos previos.
5. (Zeng, Xie, Zhang, & Lu, 2019)Se propone una arquitectura basada en la U-Net a la que se añaden bloques residuales y mecanismos de canales y escalas múltiples, por lo que se nombró como RIC-Unet por las siglas *Residual-Inception-Channel attention U-Net.* Se utilizó el **dataset** TCGA, el cual contiene 30 imágenes con tamaño 1,000 x 1,000 de siete distintos órganos. Como parte de la **metodología** se usan bloques RI en la rama de codificación y en la rama decodificadora se utiliza la metodología propuesta en la U-Net original añadiendo bloques que preservan la información de características de baja resolución. Adicionalmente se hizo un preprocesamiento de las imáges y un postprocesamiento para mejorar la segmentación obtenida. Para reportar los **resultados** se utilizaron las métricas *F1 Score, Dice Coefficient* y *Jaccard Index*. Obteniendo valores de 0.8278, 0.8008 y 0.5635, respectivamente, los cuales fueron los mejores valores comparados con 3 algoritmos previos y 3 versiones de la U-Net.
6. (Alom, Yakopcic, Taha, & Asari, 2018)Se propone una modificación de la U-Net original, añadiendo bloques recurrentes a las capas convolucionales y concatenando la entrada a cada bloque con su respectiva salida de manera residual, es por ello que se nombró R2U-Net a la arquitectura. El **dataset** utilizado fue el *2018 Data Science Bowl Grand Challenges,* el cual contiene 735 imágenes. Como parte de la **metodología**, se usan 650 imágenes para entrenar al modelo, de las cuales el 20% se utiliza para validación, y las 65 restantes se utilizan como prueba. Se utilizaron las métricas *Dice Coefficent* y *Mean Squared Error*. En la sección de **resultados** se reporta un valor del 92.15% para *Dice Coefficent*
7. (Ibtehaz & Rahman, 2019)Se propone modificar los bloques convolucionales de la U-Net original introduciendo un bloque denominado *MultiRes Block*, el cual concatena tres capas convolucionales de 3x3 con la entrada original de cada bloque para obtener información en varias escalas de la misma entrada. Adicionalmente cada capa convolucional tiene un número distinto de filtros para igualar el número de parámetros de la U-Net original, así como capas convolucionales y una conexión residual extra al momento de concatenar las características del codificador con las imágenes de mayor tamaño en el decodificador. Se utilizaron cinco **datasets**: 97 imágenes microscópicas de fluorescencia, 30 imágenes 2D EM images (del ISBI-2012), 2594 imágenes dermoscópicas (del ISIC-2018, ISIC-2017 y HAM10000), 612 imágenes endoscópicas (del CVC.ClinicDB) y 210 imágenes de resonancia magnética (del BraTS17 Competititon). Como parte de la **metodología** se compararon los resultados con la U-Net original, y su versión para imágenes 3D la cual se comparó con una versión 3D de la MultiResUnet. Se utilizó la métrica *Jaccard Index* y validación cruzada de 5 pliegues para llevar acabo los entrenamientos. En término de porcentajes, los **resultados** obtenidos para cada dataset, siguiendo su aparición en la parte superior, fueron de 91.65, 87.94, 80.29, 82.05 y 78.19, respectivamente. En todos los datasets se obtuvo un mejor resultado contra la U-Net
8. (Gu, y otros, 2019)Se propone modificar la U-Net utilizando bloques a los que denominan *Dense Atrous Convolution (DAC)* así como el *Residual Multi-kernel Pooling (RMP)*, a la cual denominaron CE-Net*.* La idea detrás del bloque DAC es extraer información a diferentes escalas cambiando el tamaño del kernel de las capas convolucionales, además se integran caminos paralelos los cuales van variando el número de convoluciones así como la tasa de espacio. En el bloque RMP también se tienen capas de pooling paralelas con distintos tamaños de kernel, las cuales se concatenan a través de una convolución. Cabe señalar que se utilizó como base la ResNet34 para la rama de codificación. Se utilizaron tres **datasets**: ORIGA con 650 imágenes, Messidor con 1,200 imágenes y RIM-ONE-R1 con 169 imágenes, DRIVE con 40 imágenes, LUNA Competition con 534 imágenes, EM Challenge (del ISBI2012) con 30 imágenes*, retinal OCT segmentation* con 20 imágenes en 3D. Como parte de la **metodología** se utilizó el aumento de datos para hacer más robusto al modelo. Se utilizaron distintas métricas, dependiendo de que dataset se evalúa, las cuales incluyen *overlapping error, sensitivity, accuracy, foreground-restricted information theoretic scoring, rand Split score, rand merge score,* entre otras. Los **resultados** son extensos ya que se compararon distintos algoritmos dependiendo del dataset a segmentar, siendo constante la comparación contra la U-Net orignal en todos; sin embargo, los mejores índices reportados son obtenidos con la arquitectura propuesta.
9. (Wollmann, y otros, 2019)Se propone utilizar la forma de la U-Net agregando el refinamiento iterativo de las características aprendidas a través de unidades recurrentes cerradas (GRU por sus siglas en inglés); además se propone agregar unidades densas de resolución completa (FRDU por sus siglas en inglés), combinadas con capas de pooling y unpooling para combinar las características de las GRU y obtener la segmentación. Es por ello que se denomina a esta arquitectura GRUU-Net. Se usaron varios **datasets**: 50 imágenes de células de glioblastomas y el Cell Tracking Challenge el cual contiene un gran conjunto de imágenes de diferentes problemas. Se pueden consultar los detalles en el artículo. Como parte de la **metodología** se utilizó aumento de datos para hacer más robusto al modelo, las métricas que se consideran para evaluar las segmentaciones son *Dice Coefficent, Jaccard Index* y *Hausdorff Distance.* En la sección de **resultados** se compara la GRUU-Net contra cuatro algoritmos, así como tres versiones de la U-Net y la ASPP-Net. Puesto que se reportan las métricas para cada tipo de imagen en el segundo dataset, se omite escribir todos los índices; sin embargo, se reporta que el modelo propuesto obtiene los mejores resultados en la mayoría de los subconjuntos.
10. (Zeng, Chen, Zhang, Bai, & Han, 2018)Se propone una modificación de la U-Net en la cual se conectan a pares las capas convolucionales usando una capa de *batch normalization*, asi como una capa ReLU, terminando con una concatenación de dos capas convolucionales. Por este motivo se le denomnó Dense-U-Net a esta arquitectura. Los **datasets** utilizados son: ISBI 2012 con 30 imágenes, Glioblastoma-astrocytroma U373 con 34 imágenes y HeLa con 18 imágenes. Como parte de la **metodología** se utilizaron las métricas *Accuracy, foreground-restricted rand scoring* y *foreground-restricted information theoretic scoring.* Obteniendo como **resultados** mejores valores para los tres datasets, los cuales se compararon contra tres algoritmos (incluida la U-Net). Siguiendo los datasets mencionados, se obtuvieron los valores (96.62%, 0.7412, 0.9206), (99.54%, 0.9950, 0.9874) y (98.28, 0.6017, 0.5725) para accuracy, rand scoring e information theoretic scoring, respectivamente, para cada dataset. Cabe señalar que fueron los mejores valores de los cuatro algoritmos evaluados.
11. (Jia, Liu, & Tai, 2019) Se propone modificar la función de activación (Soft-max) de la U-Net agregando un término que regula espacialmente la segmentación, es por ello que se denomina RU-Net a la arquitectura. Los **datasets** utilizados son: *WBC datasets*, el cual consta de un dataset con 100 imágenes y otro dataset con 300 imágenes. Además se usan los datasets: CamVid el cual consta de videos que se analizan en tandas de imágenes y SUN-RGBD dataset, del cual se seleccionaron 5,285 imágenes al azar. Estos dos últimos no contienen imágenes médicas por lo que omitiremos estos resultados. La **metodología** consta del formalismo matemático para añadir el término regulador en la capa de activación y en la capa ReLU; sin embargo, para los experimentos solo se utiliza la modificación en la función de activación ya que el tiempo de cómputo crece demasiado. Este nuevo término se auto regula a lo largo del entrenamiento durante la etapa de *backpropagation*, lo cual también se demuestra, pero solo se toma una iteración para ajustar este nuevo parámetro para reducir el tiempo de cómputo. Las métricas a evaluar son *Accuracy, mean Intersection Over Union* y *Regularization Effect.* Los **resultados** se comparan solo contra la U-Net original, pero se hizo un entrenamiento añadiendo ruido en las imágenes originales. Se demuestra que en todos los casos el modelo propuesto obtiene los mejores índices.
12. (Acevedo, Alférez, Merino, Puigvi, & Rodellar, 2019) Se propone utilizar redes neuronales convolucionales (CNN por sus siglas en inglés) y una máquina de soporte vectorial (SVM) para clasificar los elementos presentes en muestras de sangre. Se usó un **dataset** con 17,092 imágenes divididas en ocho clases de células presentes en las muestras de sangre. La **metodología** consiste en utilizar dos CNNs, la VGG-16 e InceptionV3, para extraer las características y posteriormente se añade una SVM para la clasificación. Posteriormente se ajustan los parámetros de las CNNs para llevar a cabo la clasificación añadiendo una capa Soft-Max de activación. Además se usó aumento de datos para balancear el número de imágenes en cada clase. La métrica de evaluación fue *Accuracy*, además se muestra la matriz de confusión. Los **resultados** para el modelo entrenado con el dataset propio obtiene un accuracy de 96.2% y 95% para vgg-16 e InceptionV3, respectivamente. Se lograron valores altos del índice promedio sobre todas las clases, además se toman en cuenta más clases de las que normalmente se utilizan en este tipo de problemas.
13. (Fan, y otros, 2019)Se propone modificar una red neuronal convolucional profunda, basada en la Mask-RCNN, para localizar las regiones de interés (ROI) en la imagen y solo segmentar dichas regiones en lugar de hacer la segmentación de la imagen completa, se denominó a esta arquitectura LeukocyteMask. Los **datasets** utilizados son: dataset WBC, el cual contiene dos conjuntos con 100 y 300 imágenes respectivamente, BCISC, el cual contiene 268 imágenes, y LISC, el cual contiene 257 imágenes. Como parte de la **metodología** para la extracción de características se usa una red piramidal de características (FPN por sus siglas en inglés) la cual a su vez está inspirada en la ResNet, posteriormente se usa una modificación de la red de región propuesta (RPN) para localizar las ROIs como un problema binario, finalmente se usan dos ramas paralelas para realizar la predicción final. Las métricas usadas son *Presicion, Dice Coefficent, mean Intersection Over Union, false positive rate, false negative rate* y *misclassification error.* Además se utilizó aumento de datos para hacer más robusto el modelo y validación cruzada de cinco pliegues. En la sección de **resultados** se compraron contra tres arquitecturas, incluida la U-Net, así como contra tres versiones de la arquitectura propuesta. LeukocyteMask, utilizando deformaciones elásticas en el aumento de datos, obtuvo en su mayoría los mejores índices en los cuatro datasets.
14. (Zhou, Wang, Li, & Zhang, 2019)Se propone un método denominado umbral de histograma adaptativo (AHT por sus siglas en inglés) el cual primero segmenta el núcleo de la región de interés (ROI), posteriormente localiza el citoplasma para obtener la segmentación final. Se usa un **dataset** con 60 imágenes con tinción estándar. En la **metodología** se dan los detalles acerca de la combinación del espacio de color HSI y canales RGB, así como el criterio de actualización de los pesos de cada componente, para lograr segmentar los núcleos. Para segmentar el citoplasma se usa *median filter* en escala de grises y se dan detalles de la actualización de umbrales. Las métricas usadas son *Misclassification Error (ME), false positive rate (FPR), false negative rate (FNR)* y *kappa index (KI).* Se compara contra otros dos algoritmos en la sección de **resultados,** obteniendo 0.048, 0.051, 0.038 y 0.944 para ME, FPR, FNR y KI respectivamente. Estos valores fueron los mejores evaluando la segmentación total, también se midieron las métricas para evaluar la segmentación del núcleo y el método propuesto obtuvo los mejores resultados (excepto en FPR) comparando con los otros dos métodos.
15. (Sengür, Akbulut, Budak, & Cömer, 2019)Se propone para un algoritmo obtener la clasificación de células blancas en muestras de sangre. Los **datasets** utilizados se obtuvieron del citio web Kaggle, el primero contiene 410 imágenes y el segundo contiene 12,500 imágenes, en las últimas se usó aumento de datos. La **metodología** consiste en una combinación de preprocesamiento, obtención de características con una técnica de Deep Learning y un algoritmo basado en la forma, análisis de componentes principales (PCA) y una red neuronal recurrente. Se utilizó la métrica *Accuracy* para evaluar las clasificaciones. En la sección de **resultados** se obtiene un 85.7% de accuracy, también se reporta la matriz de confusión donde se explica el rendimiento de clasificación de cada clase.
16. (Al-Dulaimi, Al-Saabawi, & Resen, 2019)Se propone un método no supervisado al utilizar un algoritmo basado en la distribución de color y texturas, así como un crecimiento adaptativo de región, para contar y segmentar células blancas en muestras de sangre.Se propone mezclar los siguientes **datasets**: Cellavision dataset, del cual se usaron 90 imágenes, ALL-IDB, del cual se usaron 50 imágenes, y *the Wadsworth Centre database*, del cual se usaron 60 imágenes. Como parte de la **metodología** se usa el algoritmo JSEG para cuantizar los colores presentes en cada imagen, posteriormente se mapea la imagen a un espacio de clases (donde cada clase es un color distinto de la etapa anterior), finalmente se utiliza el algoritmo *Fuzzy C Mean* para realizar la segmentación, acorde a los pixels cercanos a la semilla, y el conteo de células. Las métricas usadas para evaluar al modelo son *Jaccard Distance error, Rand Index, Boundary Displacement Error* y *F-Index.* En la sección de **resultados** se reportan valores de 0.002, 0.93, 10.11 y 0.93 para las métricas mencionadas en el respectivo orden. Se compararon las métricas contra 13 algoritmos, obteniendo los mejores resultados. También se añaden los resultados en el conteo de células por cada clase.
17. (Xheng, Wang, Guang, & Liu, 2018)Proponen un método no supervisado para la segmentación y clasificación de células blancas en muestras de sangre utilizando. Se utilizan dos **datasets**, el primero con 100 imágenes y el segundo con 300 imágenes. En la **metodología** se especifican los algoritmos propuestos. Primero se retira el fondo utilizando agrupamiento por *K-means*, posteriormente se recupera el área de interés mediante un análisis de concavidad (lo cual retira las células rojas adyacentes a las células blancas). Posteriormente se extraen las características y se clasifica cada imagen utilizando una máquina de soporte vectorial (SVM). Las métricas usadas para evaluar los resultados son *over segmentation rate, under segmentation rate, overall error rate, relative distance error.* En la sección de **resultados** se muestra la mejora en el proceso de segmentación en cada etapa del algoritmo. Los valores reportados de las métricas en su mayoría son los mejores, comparados contra cuatro algoritmos (incluida la U-Net). A su vez se reporta el tiempo promedio por segmentación en la etapa de evaluación, el cual es el menor de todos los algoritmos (90 ms para el dataset 1 y 630 ms para el dataset 2)

# Referencias

Acevedo, A., Alférez, S., Merino, A., Puigvi, L., & Rodellar, J. (2019). Recognition of peripheral blood cell images usinf convolutional neural networks. *Computer Methods and Programs in Biomedicine*.

Al-Dulaimi, K., Al-Saabawi, A., & Resen, R. (2019). Using Adapted JSEG Algorithm with Fuzzy C Mean for Segmentation and Counting of White Blood Cell and Nuleus Images. *Asia-Pacific Conference on Computer Science and Data Engineering*.

Alom, M., Yakopcic, C., Taha, T., & Asari, V. (2018). Nuclei Segmentation with Recurrent Residual Convolutional Neural Networks based U-Net (R2U-Net). *IEEE National Aerospace and Electronics Conference*.

Fan, H., Zhang, F., Xi, L., Li, Z., Liu, G., & Xu, Y. (2019). LukocyteMask: An auomated localization and segmentation method for leukocyte in blood smear images using deep neural networks. *Jorunal of Biophotonics*.

Gu, Z., Cheng, J., Fu, H., Zhou, K., Hao, H., Zhao, Y., . . . Liu, J. (2019). CE-Net: Context Encoder Network for 2D Medical Image Segmentation. *IEEE transactions on medical imaging*, 2281,2292.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition*.

Ibtehaz, N., & Rahman, M. (2019). MultiResUnet: Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation. *ArXiv*.

Jia, F., Liu, J., & Tai, X.-c. (2019). A Regularized Convolutional Neural Network for Semantic Image Segmentation. *arXiv*.

Milletari, F., Navab, N., & Ahmadi, S.-A. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *Fourth International Conference on 3D vision*, 565,571.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention*, 234, 241.

Sengür, A., Akbulut, Y., Budak, Ü., & Cömer, Z. (2019). White Blood Cell Classification Based on Shape and Deep Features. *International Artificial Intelligence and Data Processing Symposium*.

Wollmann, T., Gunkel, M., Chung, I., Erfle, H., Rippe, K., & Rohr, K. (2019). GRUU-Net: Integrated convolutional and gated recurrent neural network for cell segmentation. *Medical Image Analysis*, 68,79.

Xheng, X., Wang, Y., Guang, G., & Liu, J. (2018). Fast and robust segmentation of white blood cell images by self-supervised learning. *Micron*.

Xia, X., & Kulis, B. (2017). W-Net: A Deep Model for Fully Unsupervised Image Segmentation. *ArXiv*.

Zeng, Y., Chen, X., Zhang, Y., Bai, L., & Han, J. (2018). Dense-U-Net: densely connected convolutional network for semantic segmentation with a small number of samples. *Tenth International Conference on Graphic and Image Processing*.

Zeng, Z., Xie, W., Zhang, Y., & Lu, Y. (2019). RIC-Unet: An Improved Neural Network Based on Unet for Nuclei Segmentation in Histology Images. *IEEE Access*.

Zhou, X., Wang, C., Li, Z., & Zhang, F. (2019). Adaptative Histogram Thersholding-based for Leukocyte Segmentation. En *Advances in Intelligent Information Hiding and Multimedia Signal Processing* (págs. 451,459). Jilin, China: Springer.