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.

# Referencias

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*.

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*.

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.

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

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*.