UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Posgrado en Ciencia e Ingeniería de la Computación

Curso de aprendizaje profundo

DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation

Debesh Jha*†, Michael A. Riegler*, Dag Johansen†, Pål Halvorsen*‡, Håvard D. Johansen†

*SimulaMet, Norway †UiT The Arctic University of Norway, Norway ‡Oslo Metropolitan University, Norway

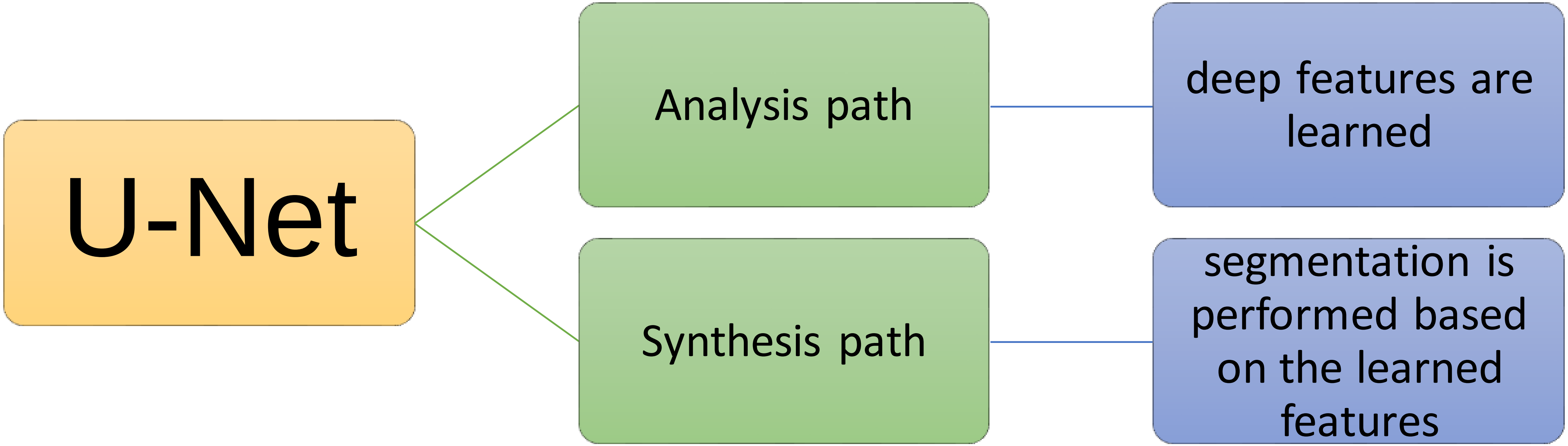
Email: debesh@simula.no

Equipo:
Dalia Y. Domínguez Jiménez
Miguel Ángel Veloz Lucas

Introduction
Segmentation task and U-Net network

In the state-of-the-art Convolutional Neural Networks (CNNs) have shown a good performance for automated medical image segmentation.

For example: U-Net



Related work in the field of medical image segmentation

U-Net Model

U-Net have gained significant popularity among semantic segmentation approach for 2D images.

In the medical imaging, there are many challenging images, which are usually missed out during colonoscopy examination and can develop into cancer if early detection is not performed. Therefore, there is a need for a more accurate medical image segmentation approach to deal with the challenging images

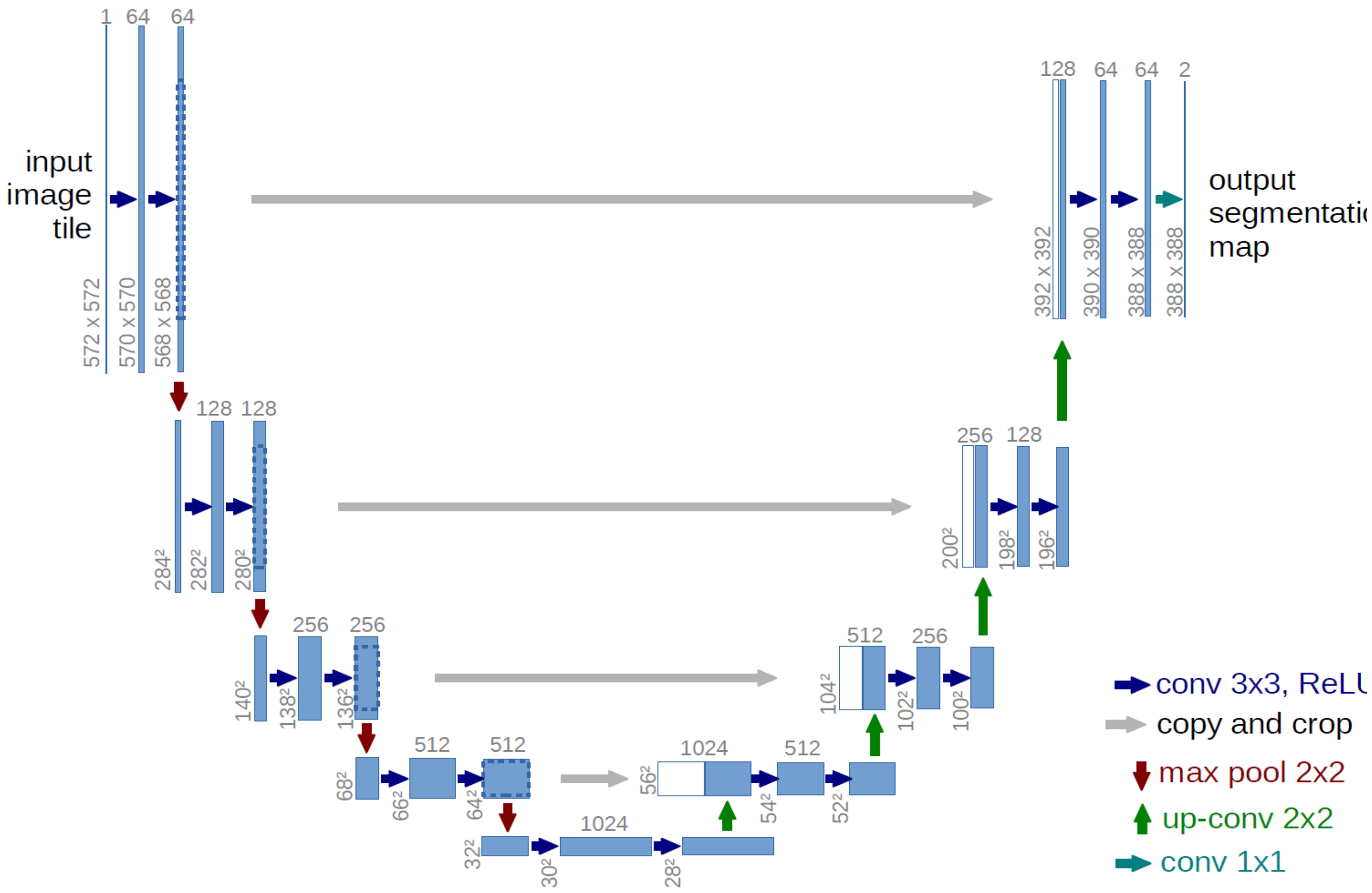
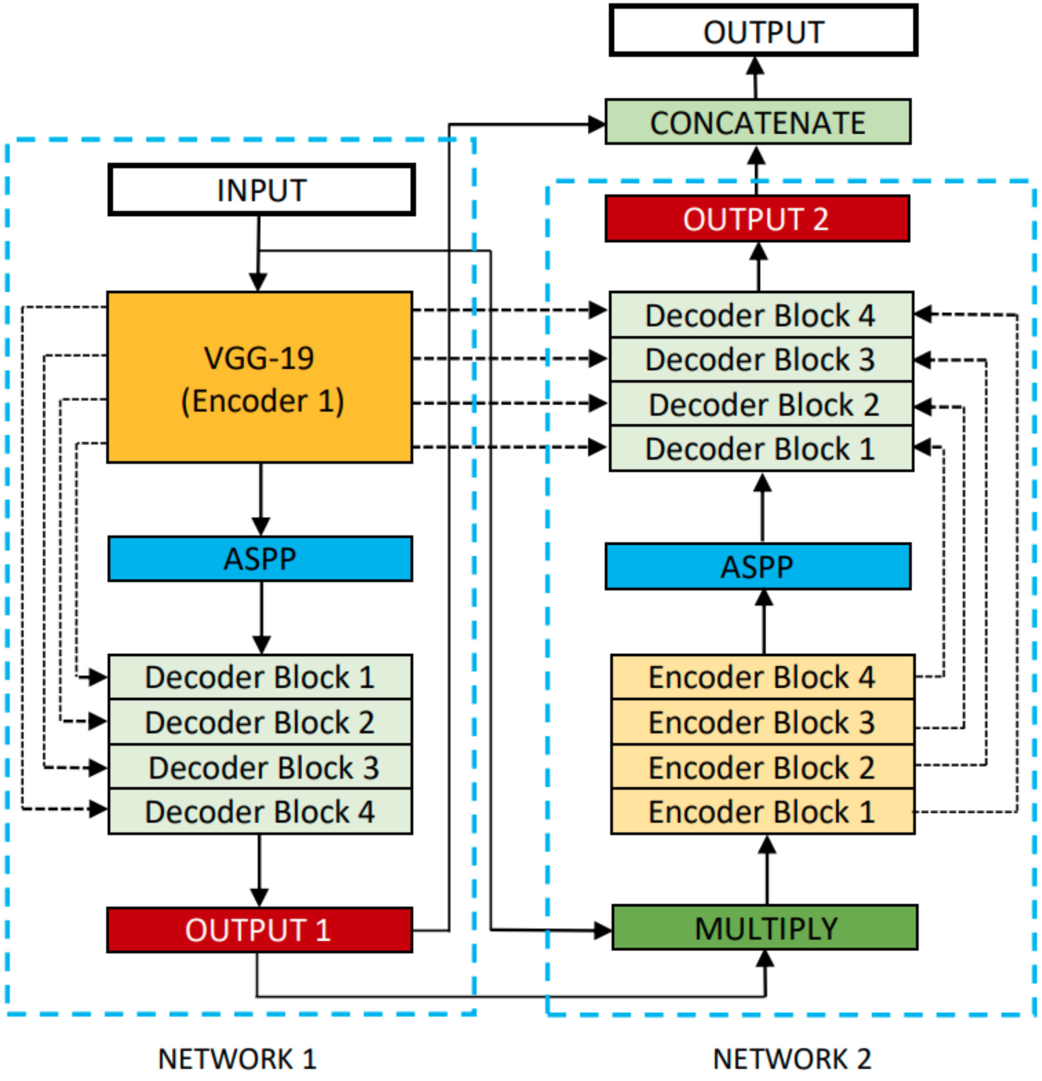


Image: https://paperswithcode.com/method/u-net

Double U-Net Architecture

The produced output feature map from NETWORK 1 can and concatenating with Output2 will produce a better segmentation mask than the previous one.



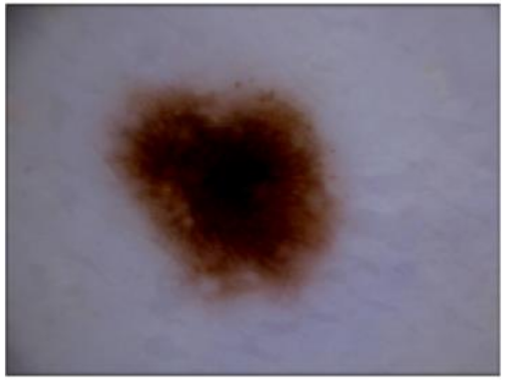
The experiments
Datasets

Dataset	Description	No. of Images	Input size	Aplication
a) 2015 MICCAI sub-challenge automatic polyp detection	It used the CVC-ClinicDB [1] for training and ETIS-Larib [2]	808	384 x 288	Colonoscopy
b) CVC-ClinicDB	For polyp segmentation	612	384 x 288	Colonoscopy
c) Lesion Boundary Segmentation dataset [3, 4]	It contains skin lesions and their corresponding annotations.	2594	Variable	Dermoscopy
d) 2018 Data Science Bowl challenge	For nuclei segmentation	670	256 x 256	Nuclei

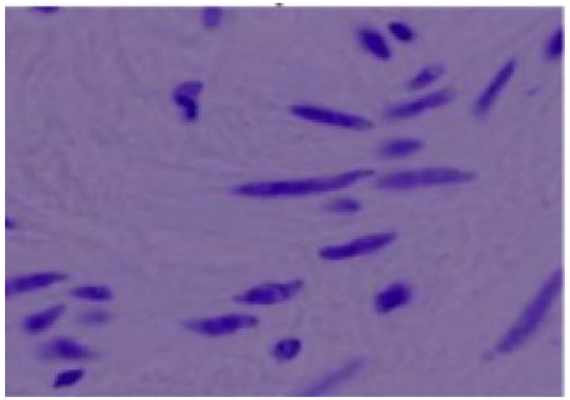
a)

b)

c)



d)



The experiments
Evaluation metrics

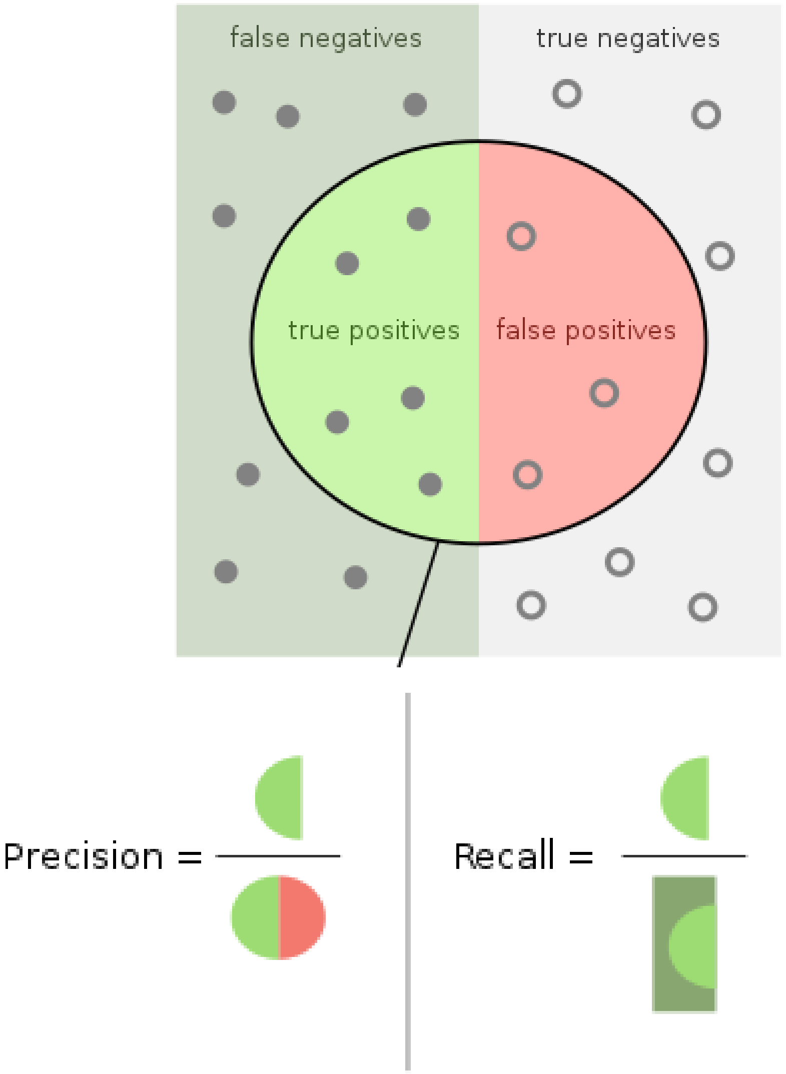
DoubleU-Net is evaluated on the basis of :

1. Precision:

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

2. Recall:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{P}} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

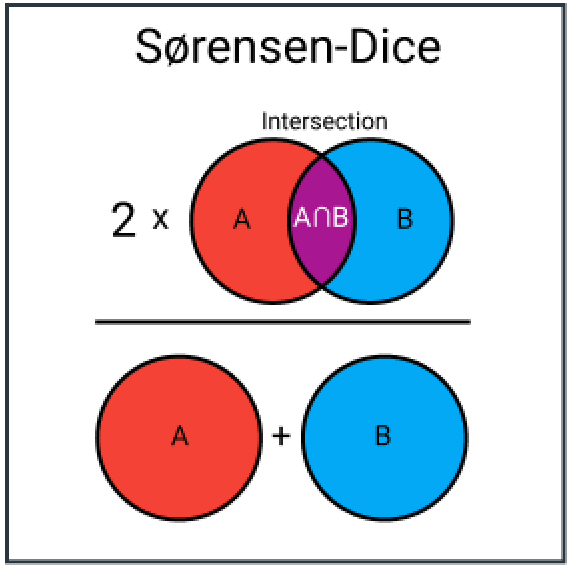


The experiments
Evaluation metrics

DoubleU-Net is evaluated on the basis of :

1. Sørensendice coefficient or Dice similarity coefficient (DSC) :

$$2\,|X \cap Y| \,/\, (|X| + |Y|)$$



2. Mean Intersection over Union (mIoU):

$$IoU = \frac{TP}{(TP + FP + FN)}$$

Ground Truth Mask

Predicted Mask

The experiments
Experiment setup and configuration

- All models were implemented using Keras framework [5] with Tensorflow 2.1.0 [6] as backend.
- A Volta 100 GPU and an Nvidia DGX-2 AI system were used.
- In all of the datasets:

- The original image size for the smaller dataset was used and resized the images to 384×512 for the Lesion Boundary segmentation challenge dataset to balance between training time and complexity.

Réplica de artículo

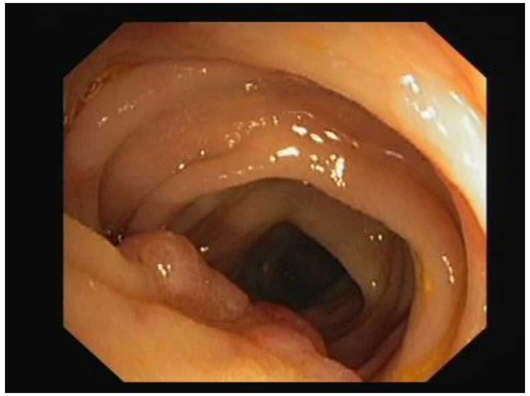
Base de datos ***CVC-ClinicDB***

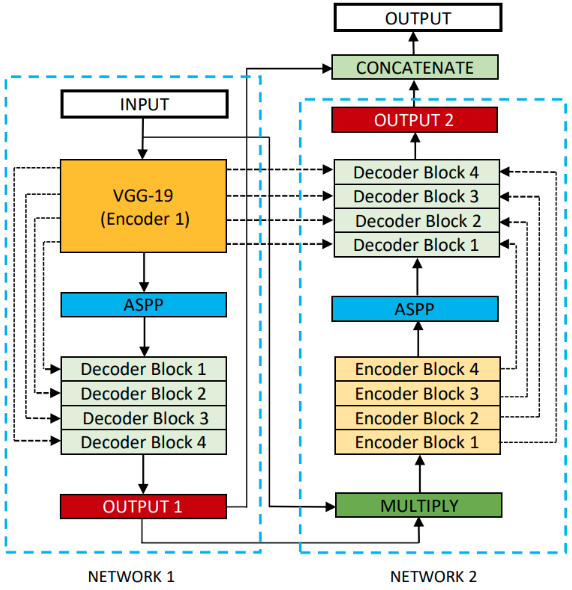
Configuración de los experimentos

- Se utilizó la entropía binaria cruzada.
- El optimizador Nadam fue empleado.
- El tamaño del batch es 16 y la tasa de aprendizaje de $1e{-5}$.
- Se entrenó el modelo con 100 épocas.

Réplica del modelo DoubleU-Net: Base de datos

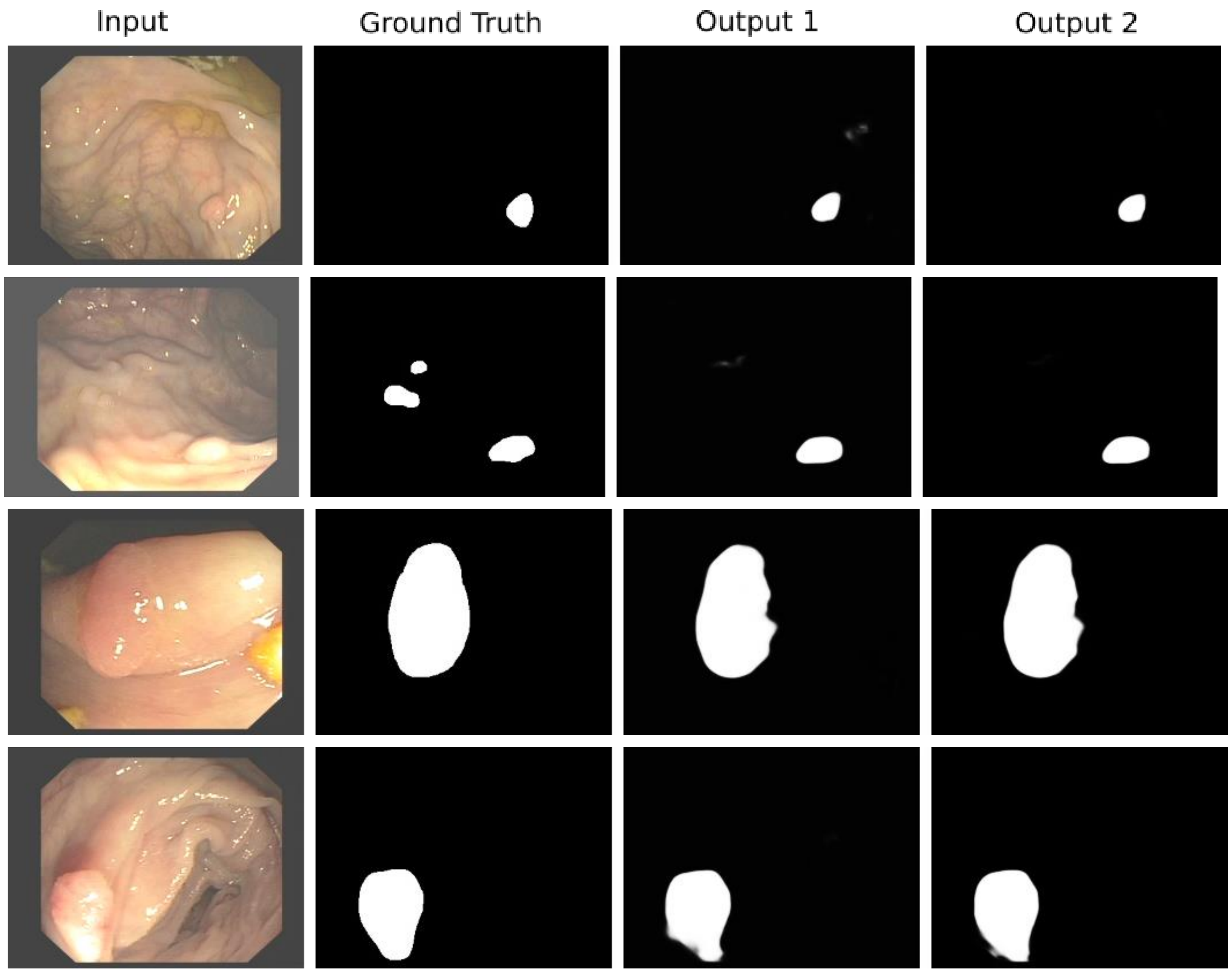
- Se utilizo la BD CVC-ClinicDB de colonoscopía.
- Resolución de imágenes resolución de 256x192 de un conjunto base 612:
  - 200 para entrenamiento ➔ 5 000 (25 transformaciones)
  - 61 para validación
  - 61 para prueba

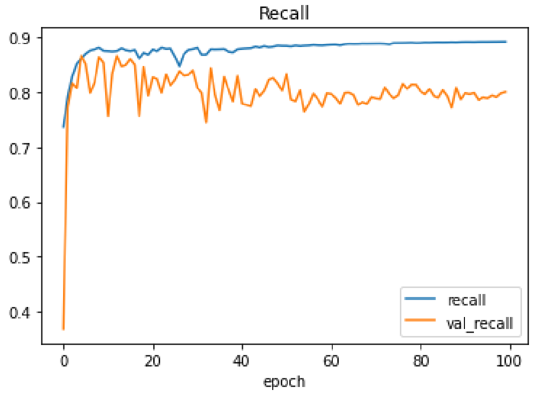


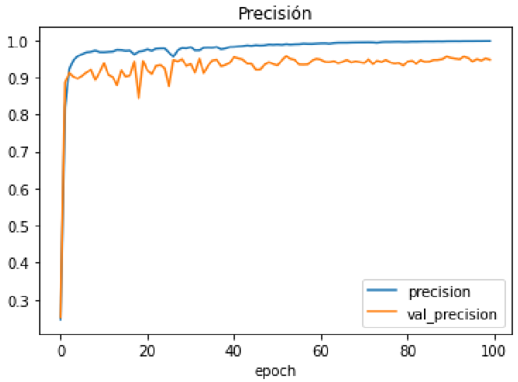


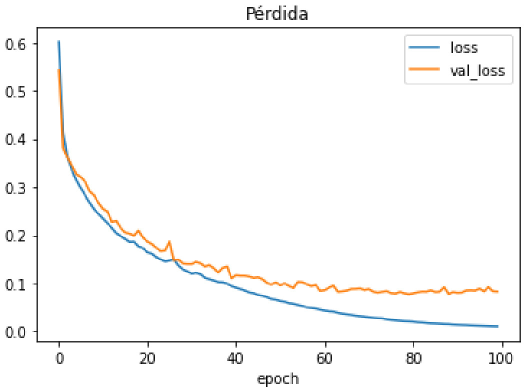

```python
from albumentations import (
    PadIfNeeded,
    HorizontalFlip,
    VerticalFlip,
    CenterCrop,
    Crop,
    Compose,
    Transpose,
    RandomRotate90,
    ElasticTransform,
    GridDistortion,
    OpticalDistortion,
    RandomSizedCrop,
    OneOf,
    CLAHE,
    RandomBrightnessContrast,
    RandomGamma,
    HueSaturationValue,
    RGBShift,
    RandomBrightness,
    RandomContrast,
    MotionBlur,
    MedianBlur,
    GaussianBlur,
    GaussNoise,
    ChannelShuffle,
    CoarseDropout
)
```

Desempeño DoubleUnet









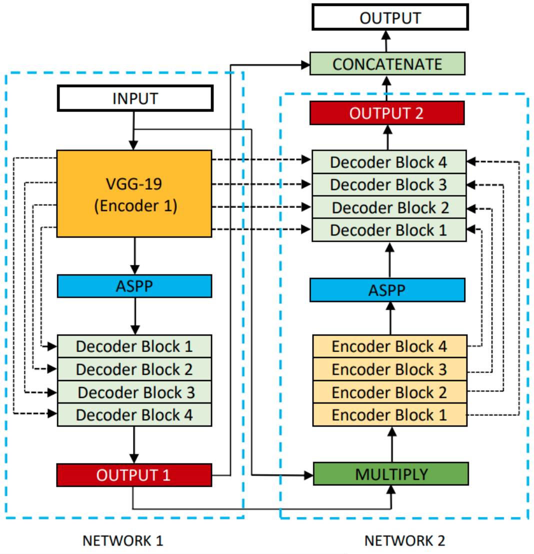
Modificación del modelo

Reducción de filtros en las dos redes que componen la arquitectura *DoubleUnet*

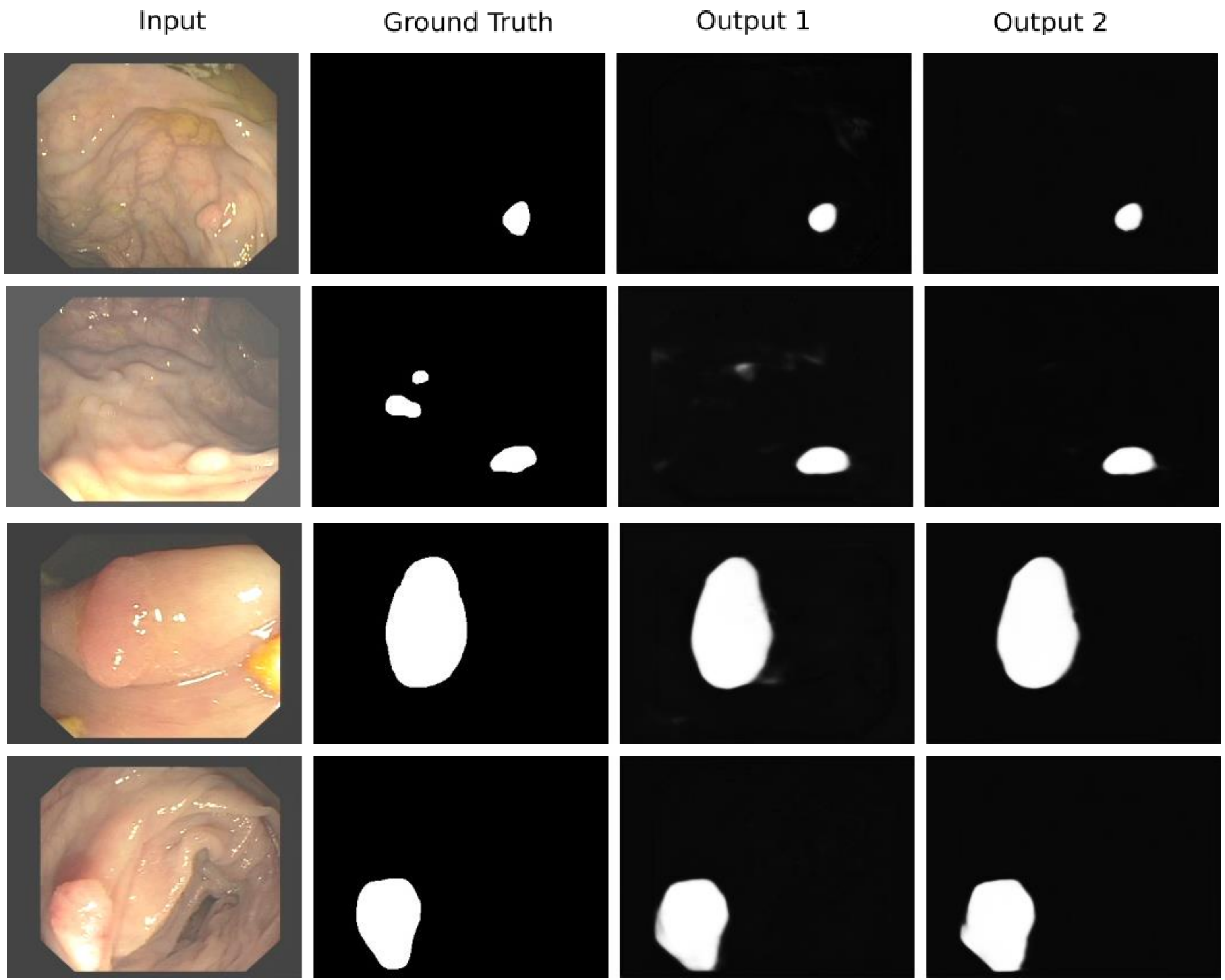
Descripción de la modificación

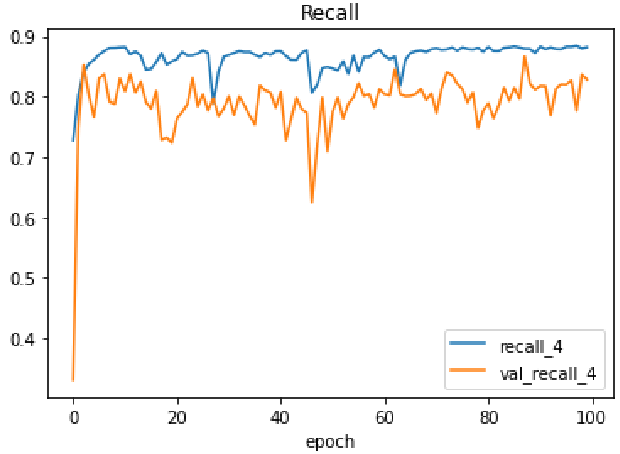
- Reducción de filtros aplicados por convolución en la étapa de codificación para el modelo UNET.
- Reducción de filtros en la etapa de decodificación para ambos modelos (VGG19, UNET).
- Resultados:
  - Mejora de velocidad de ejecución del modelo.
  - Desempeño parecido al original.

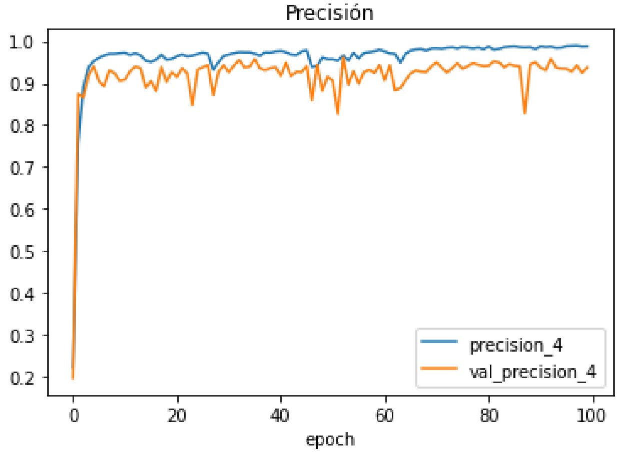
```python
def encoder2(inputs):
#     num_filters = [32, 64, 128, 256]
    num_filters = [16, 32, 64, 128]
    skip_connections = []
    x = inputs

    for i, f in enumerate(num_filters):
        x = conv_block(x, f)
        skip_connections.append(x)
        x = MaxPool2D((2, 2))(x)

    return x, skip_connections

def decoder2(inputs, skip_1, skip_2):
#     num_filters = [256, 128, 64, 32]
    num_filters = [128, 64, 32, 16]
    skip_2.reverse()
    x = inputs

    for i, f in enumerate(num_filters):
        x = UpSampling2D((2, 2), interpolation='bilinear')(x)
        x = Concatenate()([x, skip_1[i], skip_2[i]])
        x = conv_block(x, f)

    return x
```

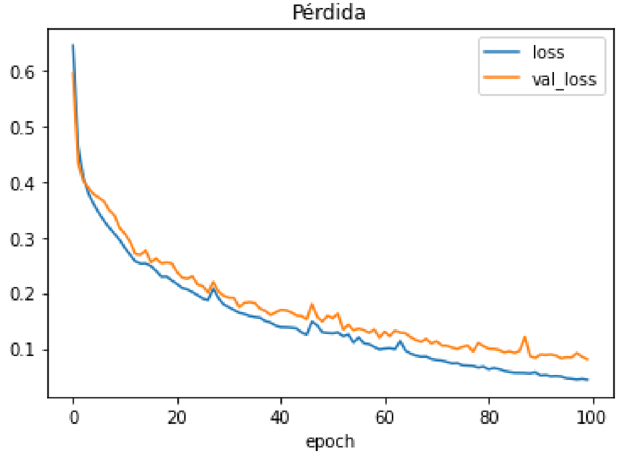


Desempeño DoubleUnet Modificada



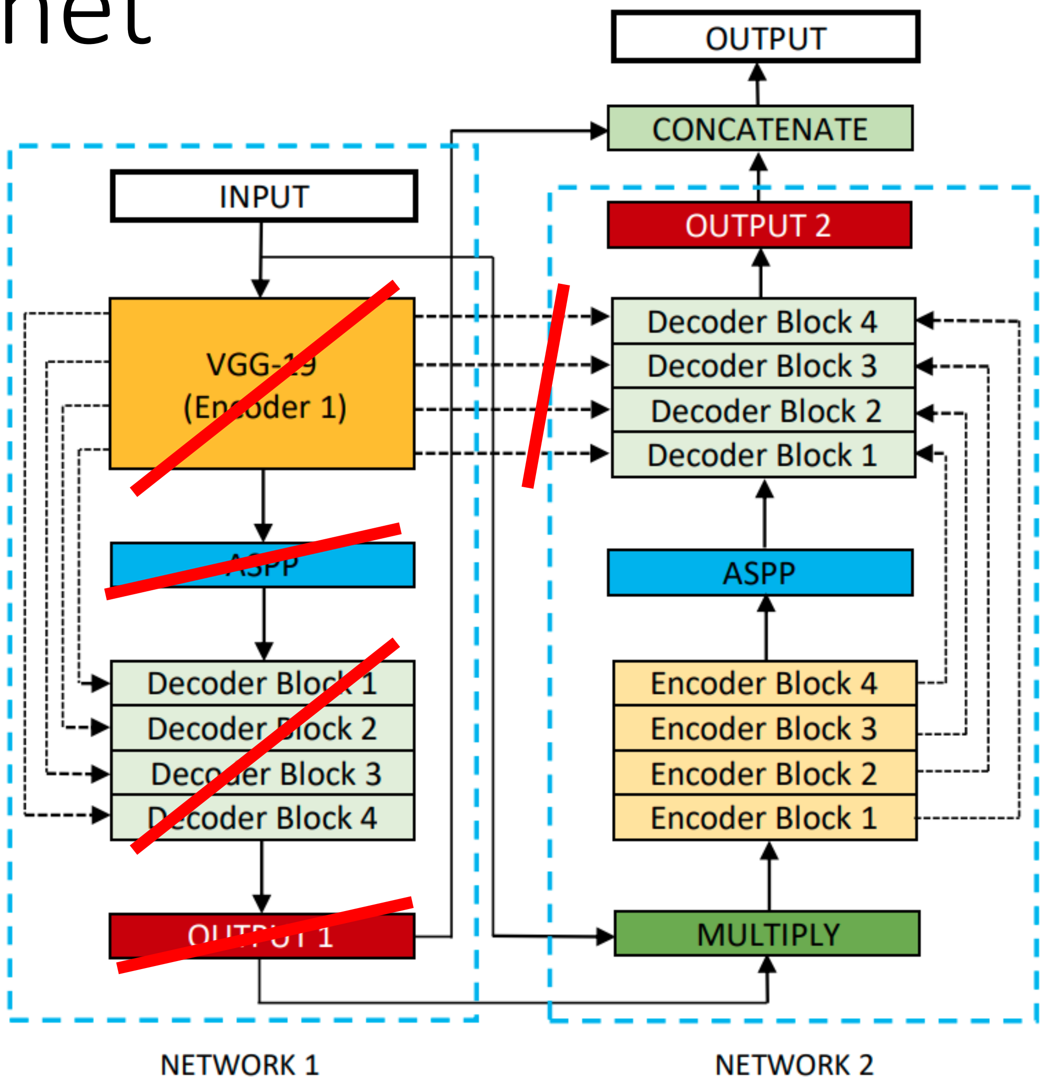




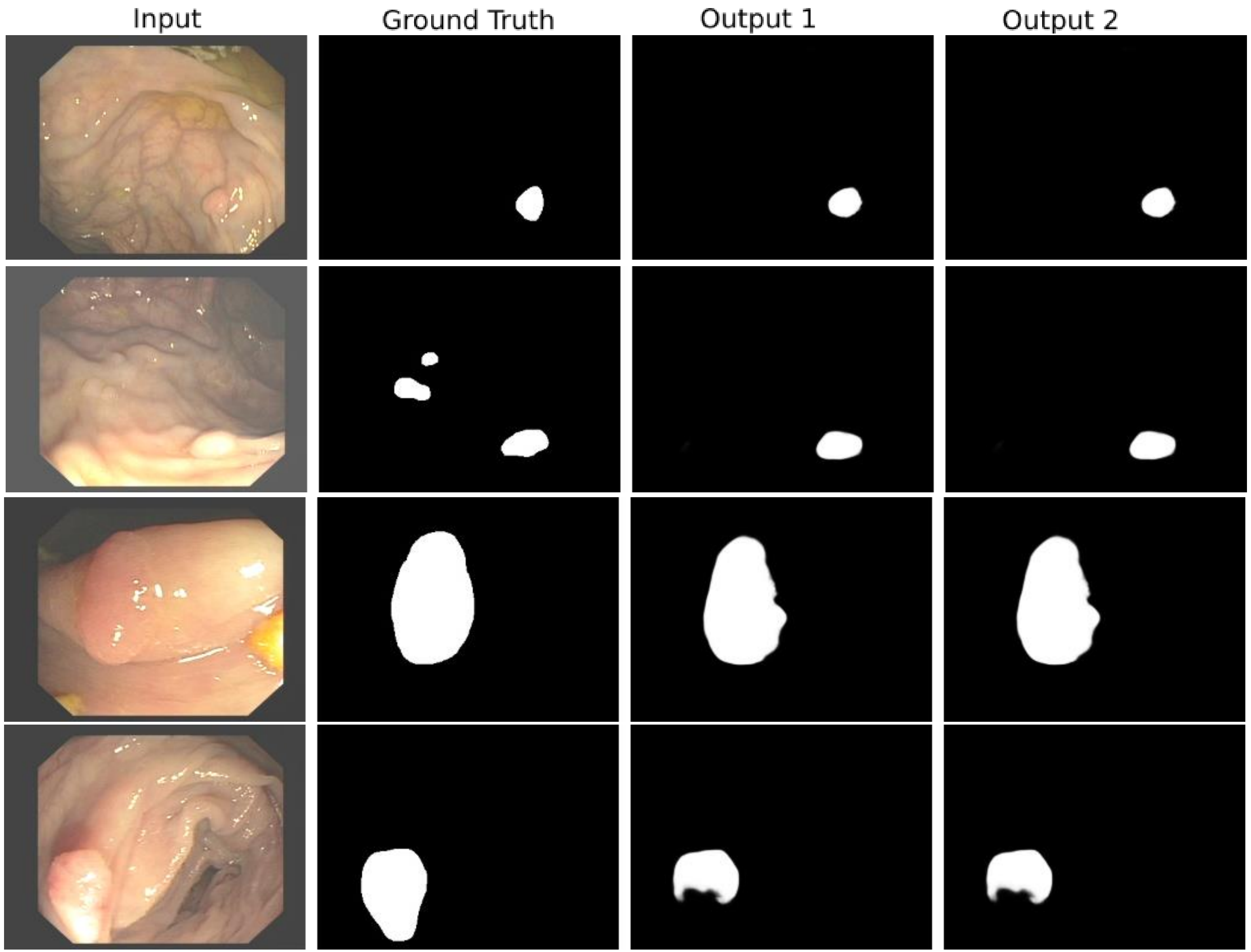


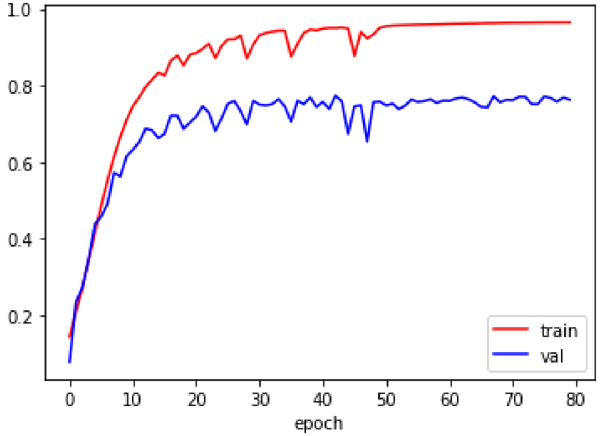
Modelo de referencia Unet

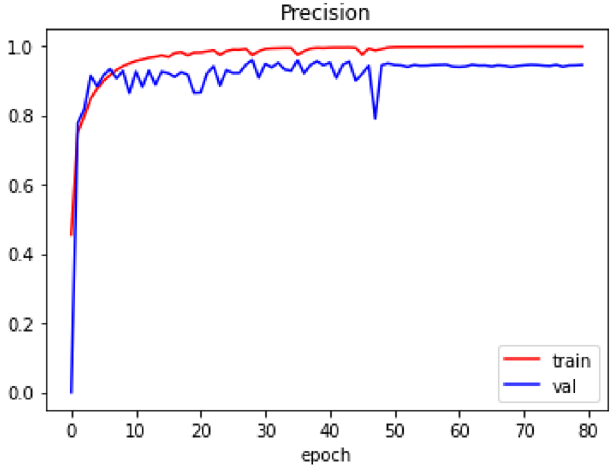
- Para tener una referencia de sobre el desempeño de la arquitectura DoubleUnet modificado, se entreno un modelo Unet basado en el número de filtros aplicados en las étapas de codificación de decodificación del modelo original.

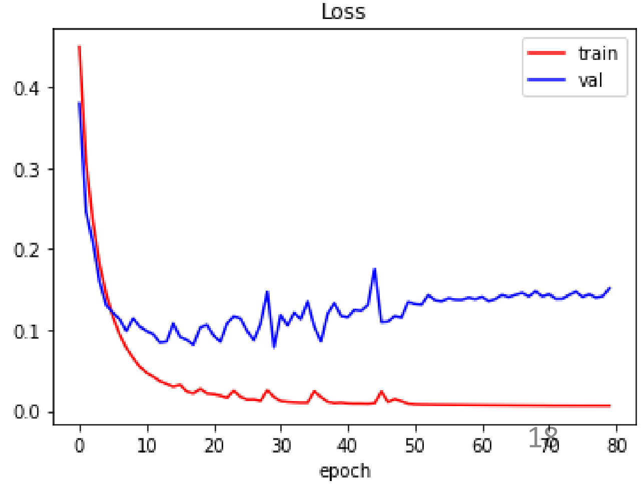


Desempeño modelo Unet de referencia







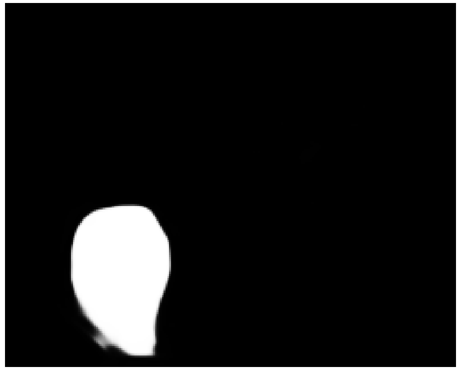


Discusión: Unet vs Double UNet

Input

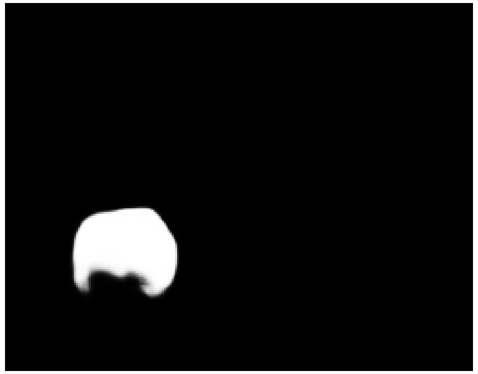
Ground Truth

Double Unet



Double Unet Mod

Unet



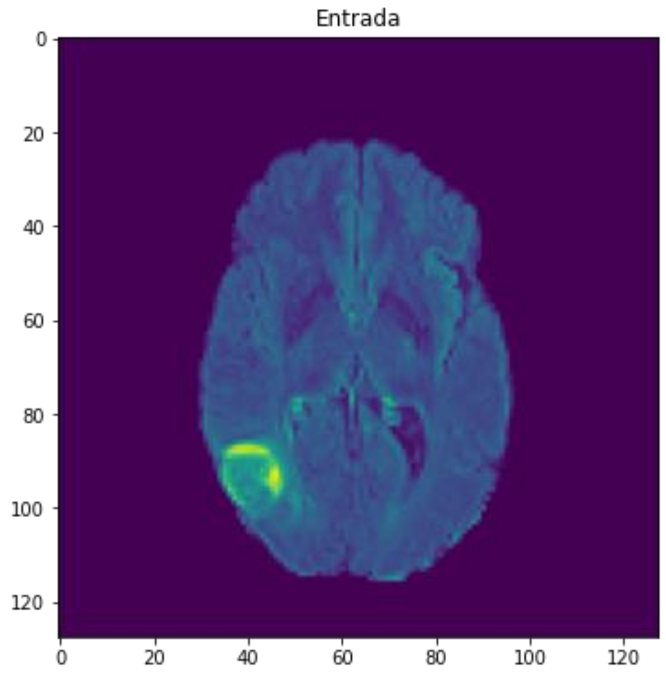
Modelo	loss	precision	val_loss	val_precision
DoubleUnet	**0,0113**	**0,998**	**0,0826**	**0,9475**
DoubleUnet Mod	0,0443	0,9879	0,0811	0,9376
Unet	0,0364	0,9683	0,1516	0,9461

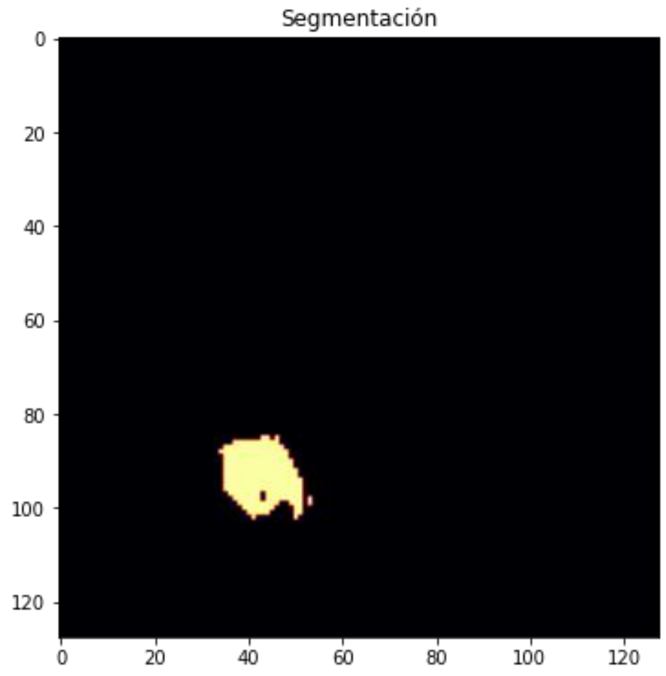
Segmentación BD BrainTS 2021

Implementación y desempeño de los modelos Unet y DoubleUnet en comparación con el modelo ViT Swin-UNETR (Swin-Unet Transformers) .

BD BrainTS 2021

- Se utilizo la BD BraTS 2021 Challenge con
- Resolución de imágenes resolución de 128x128 de un conjunto base 7000:
  - 5 600 para entrenamiento
  - 1 400 para validación
  - 1 para prueba

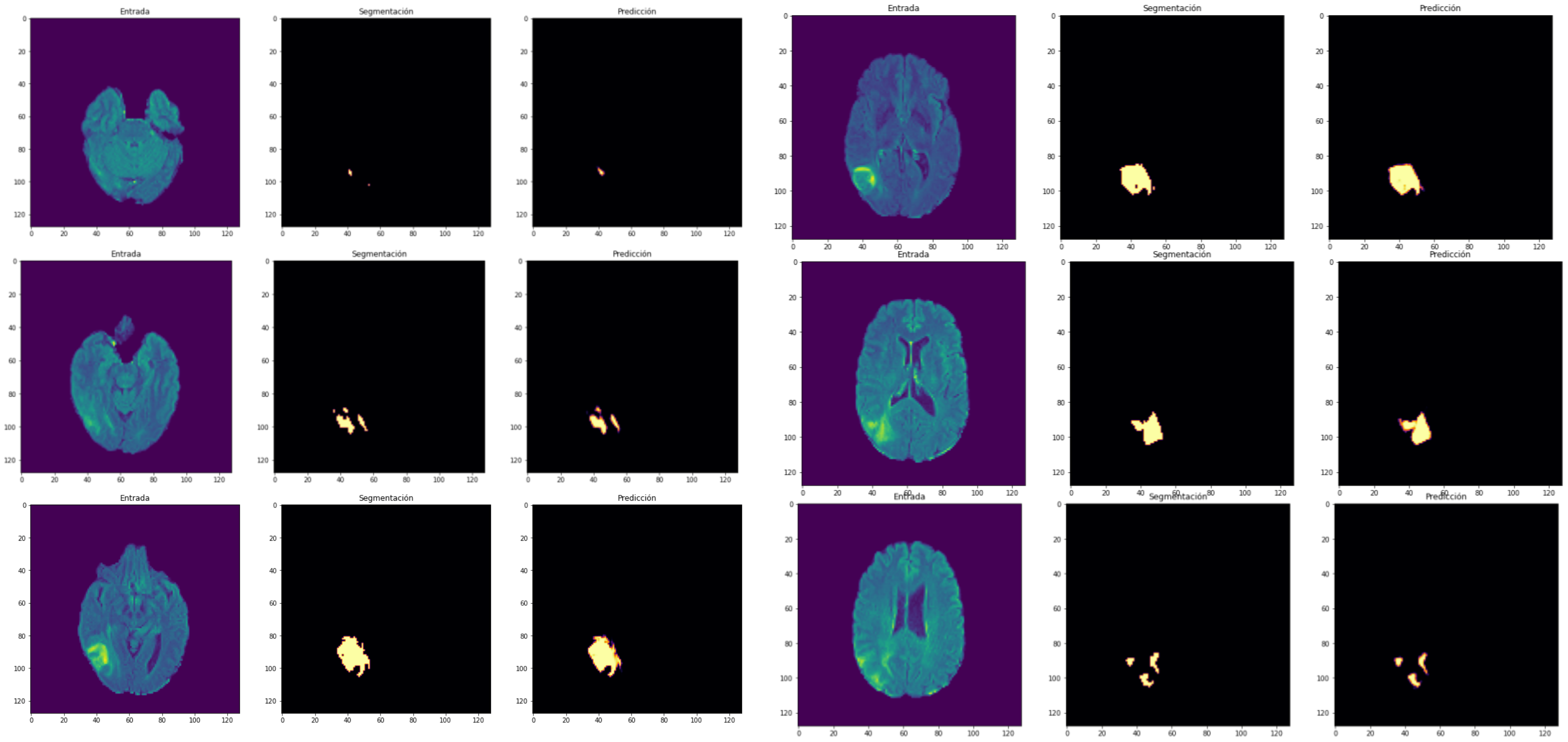


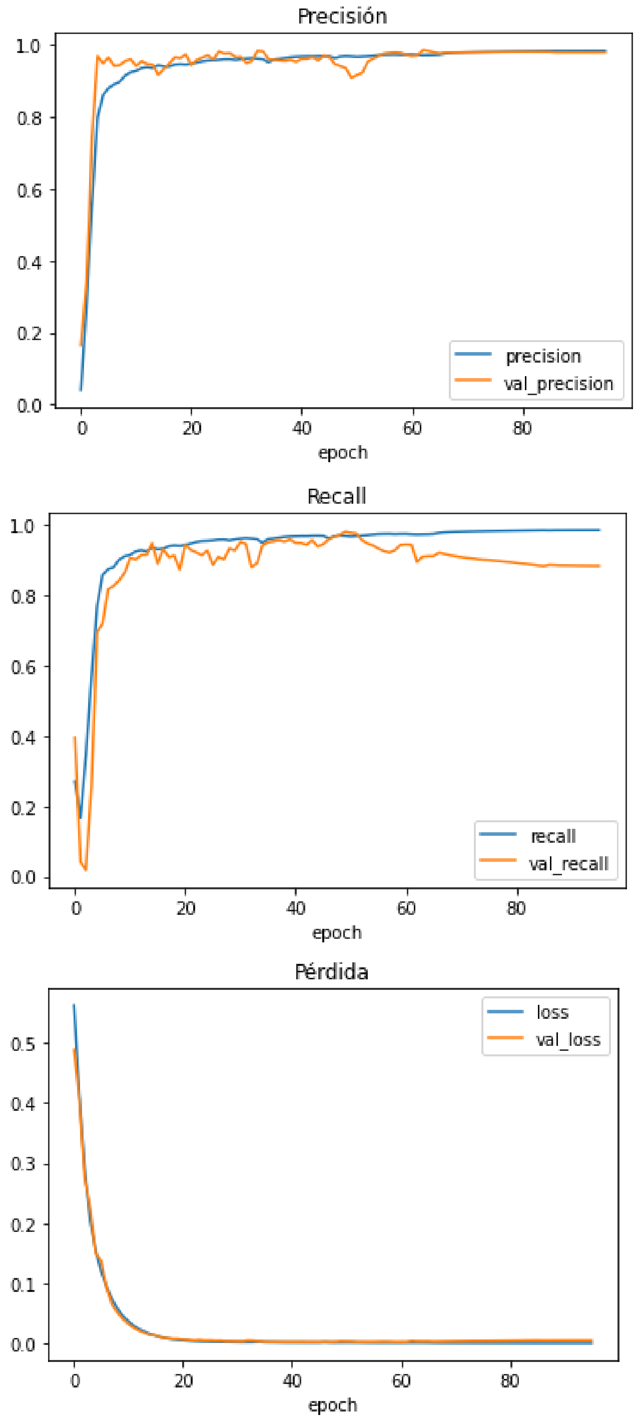


Configuración de los experimentos modelos DoubleUnet y Unet

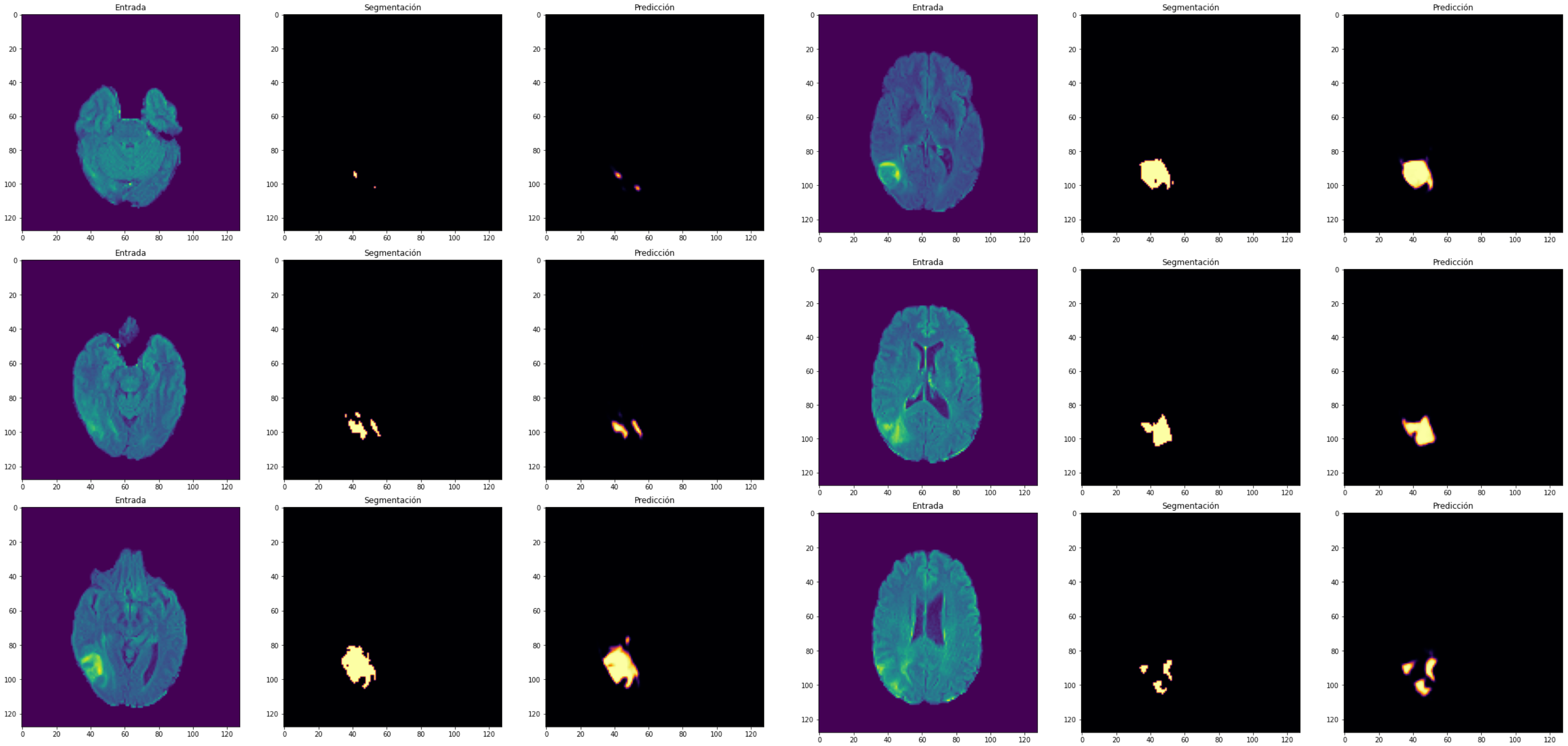
- Se utilizó la entropía binaria cruzada.
- El optimizador Adam fue empleado.
- El tamaño del batch es 16 y la tasa de aprendizaje de 1e−5 o 1e-4.
- Se entrenó el modelo con 100 épocas.

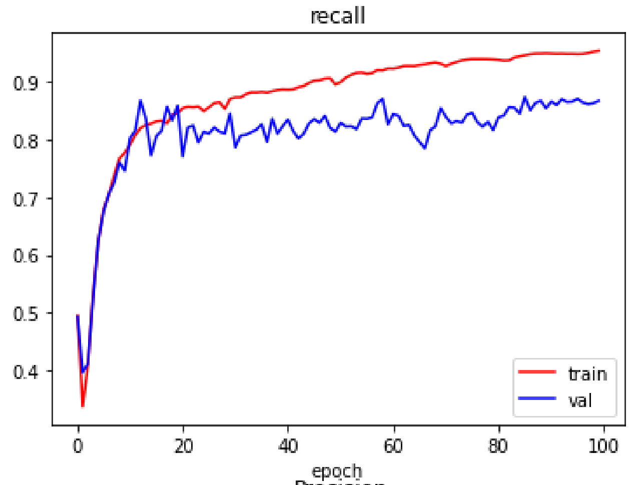
DoubleUNet-Brain

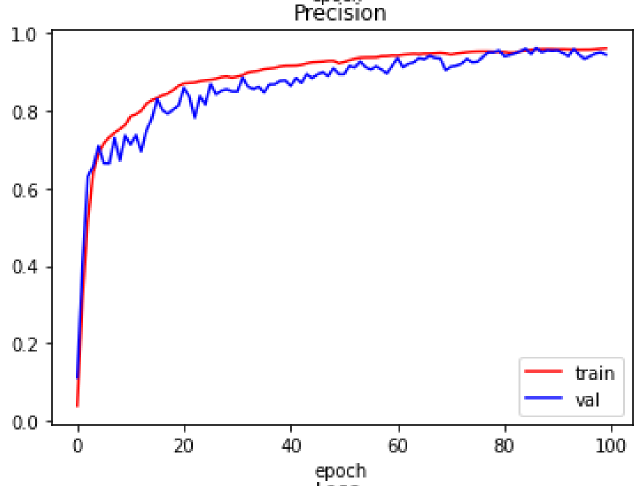


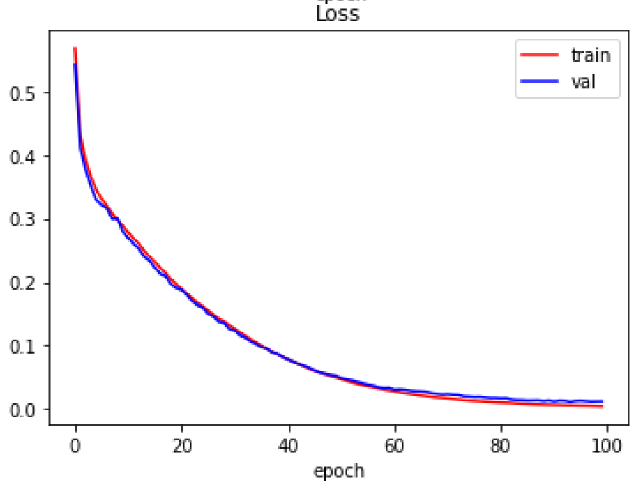


UNet-Brain


Entrada
Segmentación
Predicción

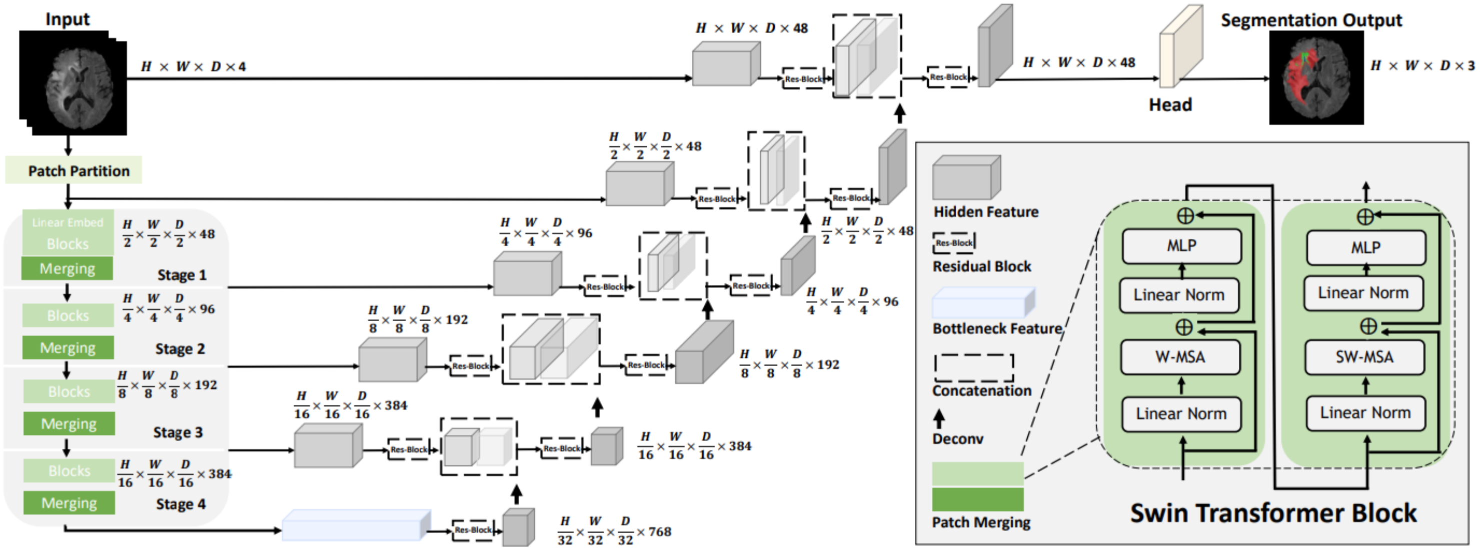

recall
train
val
epoch

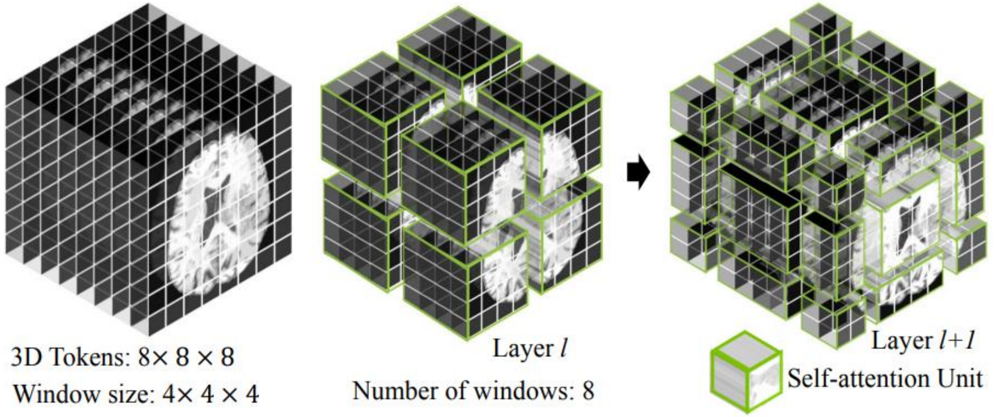

Precision
train
val
epoch


Loss
train
val
epoch

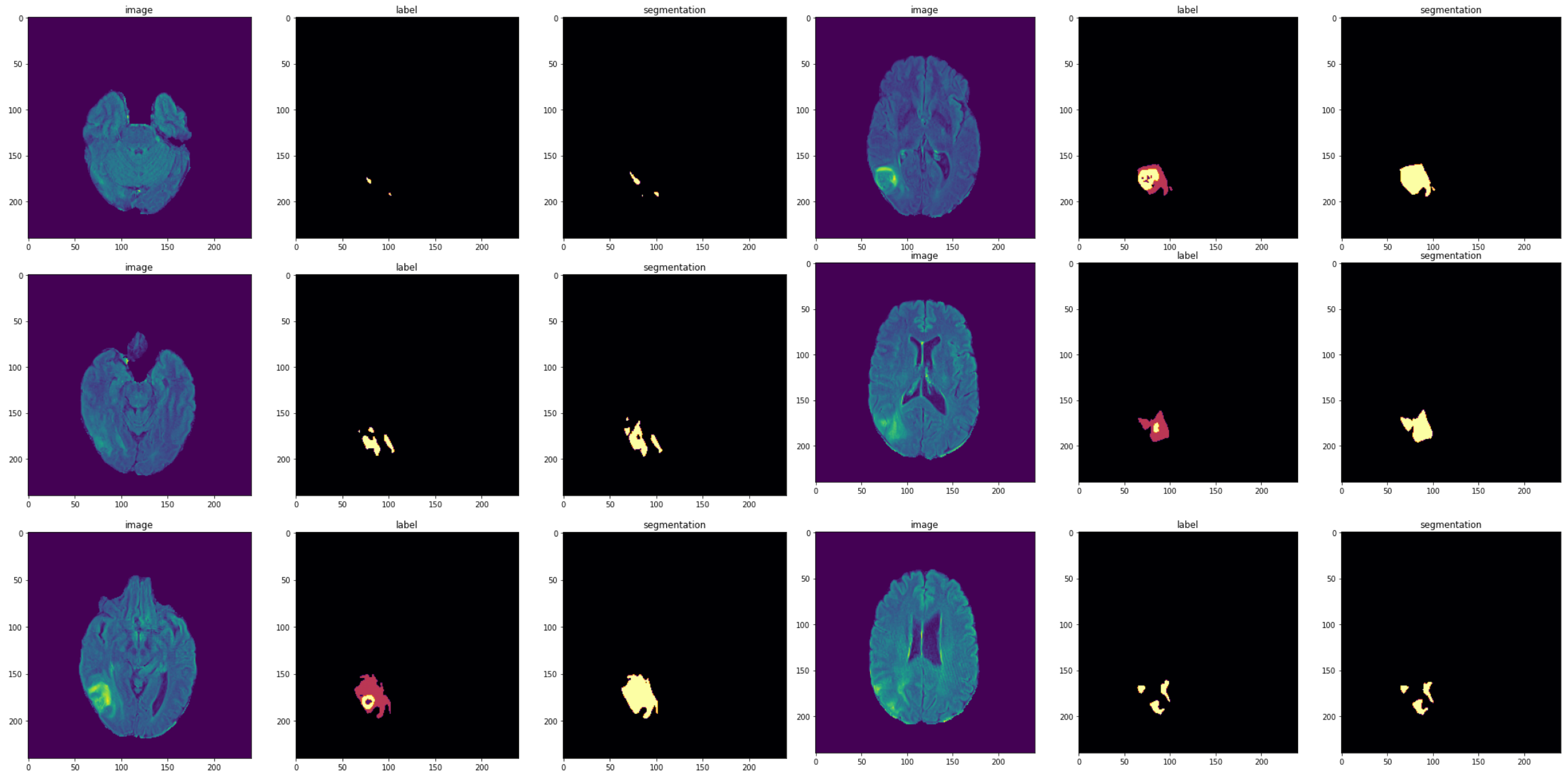
Configuración de los experimentos modelo Swin-UNETR

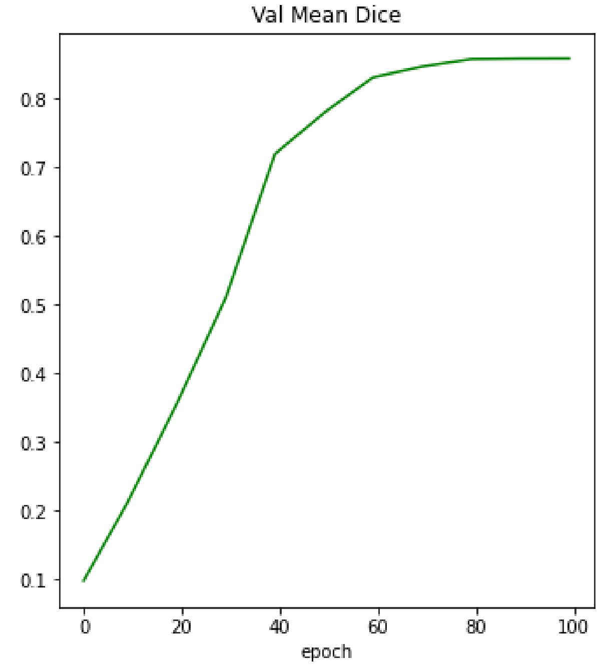
- Se utilizó la entropía binaria cruzada.
- El optimizador AdamW fue empleado.
- El tamaño del batch es 16 y la tasa de aprendizaje de 1e−5 o 1e-4.
- Transformaciones
  - RandFlipd
  - NormalizeIntensityd
  - RandScaleIntensityd
  - RandShiftIntensityd
  - ConvertToMultiChannelBasedOnBratsClassesd
  - RandSpatialCropd
  - CropForegroundd
- Se entrenó el modelo con 100 épocas.

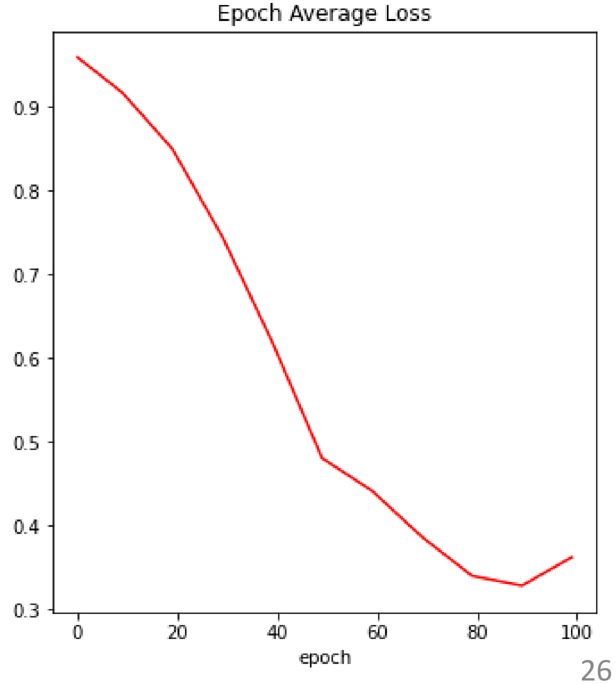




Swin-UNET Transformers

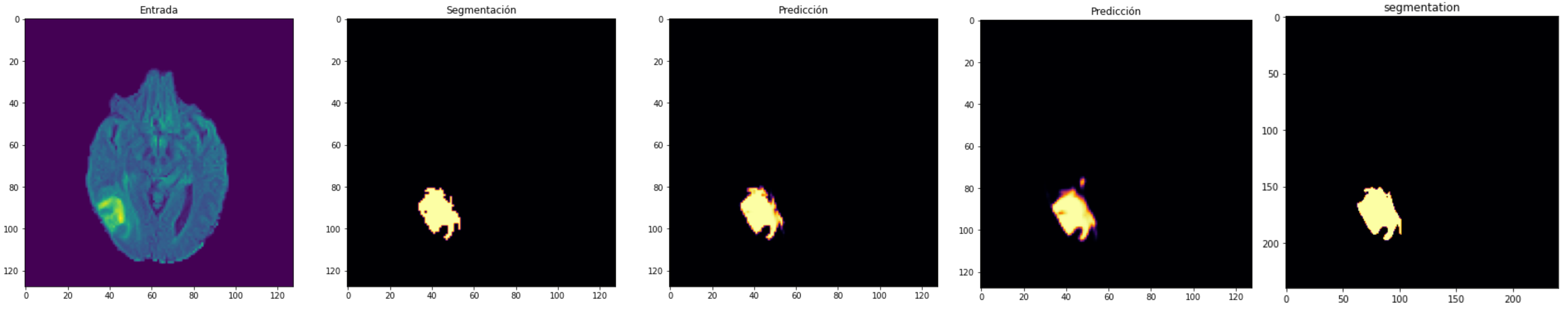






Discusión: Unet vs Double Unet VS *Swin UNETR*

Input | Ground Truth | Double Unet | Unet | Swin-UNETR



Modelo	loss	precision	val_loss	val_precision
DoubleUnet	**0.0093**	**0.9844**	**0.0052**	**0.9801**
Unet	0,0443	0,9610	0,0115	0,9444
Swin UNETR	0,0364	0,9263	--	0,8583

Referencia

- Jha, Debesh, et al. "Doubleu-net: A deep convolutional neural network for medical image segmentation." *2020 IEEE 33rd International symposium on computer-based medical systems (CBMS)*. IEEE, 2020.
- Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H. and Xu, D., 2022. Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. arXiv preprint arXiv:2201.01266.
- Tang, Y., Yang, D., Li, W., Roth, H.R., Landman, B., Xu, D., Nath, V. and Hatamizadeh, A., 2022. Self-supervised pre-training of swin transformers for 3d medical image analysis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 20730-20740).
- U.Baid, et al., The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification, arXiv:2107.02314, 2021.
- B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015) DOI: 10.1109/TMI.2014.2377694
- S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data, 4:170117 (2017) DOI: 10.1038/sdata.2017.117
- S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, et al., "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-GBM collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.KLXWJJ1Q

Referencias

[1] J. Bernal, F. J. S ́anchez, G. Fernández-Esparrach, D. Gil, C. Rodr
iguez,and F. Vilari ̃no, “Wdova maps for accurate polyp highlighting incolonoscopy: Validation vs. saliency maps fro m physicians,”Comput-erized Medical Imaging and Graphics, vol. 43, pp. 99–111, 2015
[2] .Silva, A. Histace, O. Romain, X. Dray, and B. Granado, “Towardembedded detection of polyps in wce images for early diagnosis of col-orectal cancer,”International Journal of Computer Assisted Radiologyand Surgery, vol. 9, no. 2, pp. 283–293, 2014.
[3] N. C. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti,S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittleret al.,
“Skinlesion analysis toward melanoma detection: A challenge at the 2017international symposium on biomedical i maging (isbi), hosted by theinternational skin imaging collaboration (isic),” inIEEE InternationalSymposium on Biomedical Imaging (ISBI), 2018, pp. 168–172.
[4] P. Tschandl, C. Rosendahl, and H. Kittler, “The ham10000 dataset,a large collection of multi-
source dermatoscopic images of commonpigmented skin lesions,”Scientific data, vol. 5, p. 180161, 2018.[26] F. Cholletet al., “Keras,” 2015
[5] F. Cholletet al., “Keras,” 2015.
[6] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M.
Devin,S. Ghemawat, G. Irving, M. Isardet al., “Tensorflow: A system forlarge-scale machine learning,” inProceeding of{USENIX}Symposiumon Operating Systems Design and Implementation ({OSDI}), 2016, pp.265–283
[7] Ronneberger, O., Fischer, P., and Brox, T., “U-Net: Convolutional Networks for Biomedical Image Segmentation”, 2015.

Gracias….