Exploring U-SegNet and its variants for medical image segmentation

Antonine Batifol

Based on: [1]U-SEGNET: Fully convolutional neural network based automated brain, Pulkit Kumar Pravin Nagar Chetan Arora Anubha Gupta

Introduction and Motivation

Deep Learning has become a useful tool for automating medical image segmentation. Indeed manual segmentation from experts is time-intensive and prone to error due to the challenging architecture some organs. We focus here on the segmentation of blood vessels in human organs which may help diagnose and monitor various vascular diseases. State-of-the-art DL segmentation architectures are often computationally heavy and require a large number of parameters to be learned. U-Net and SegNet architectures require lesser training data and are widely used for image segmentation, however they both present challenges:

- **U-Net** excels at capturing multiscale features using skip connections but requires a high number of learning parameters (due to learnable upsampling), and so high computational memory.
- **SegNet** is more memory efficient (passing pooling indices to upsampling layers reduce the number of parameters to learn) but struggles to capture fine boundary details and multiscale information.

To address these limitations, the paper [3] proposed a new hybrid architecture combining U-Net and SegNet. We implemented **U-SegNet** architectures (with 3 levels) and variants to evaluate its perfomance for blood vessel segmentation as well as baseline models like U-Net (with 4 levels), SegNet (with 5 levels) and Res-Unet for comparison. We try to evaluate the trade-off between model size and performance using metrics such as Dice Score and accuracy, as well as visual results, across two datasets.

Methods

Architecture

The paper proposed a new architecture taking the SegNet architecture and incorporating in the last deconvolutional layer of the decoder a skip connection inspired by Unet. Introducing one skip connection at the end helps incorporate fine information without increasing too much the number of parameters. The last convolutional block thus combines coarser and finer information for the segmentation task. The diagram below illustrates the U-SegNet design with its downsampling path, the pass of pooling indices, and one skip connection from U-Net.

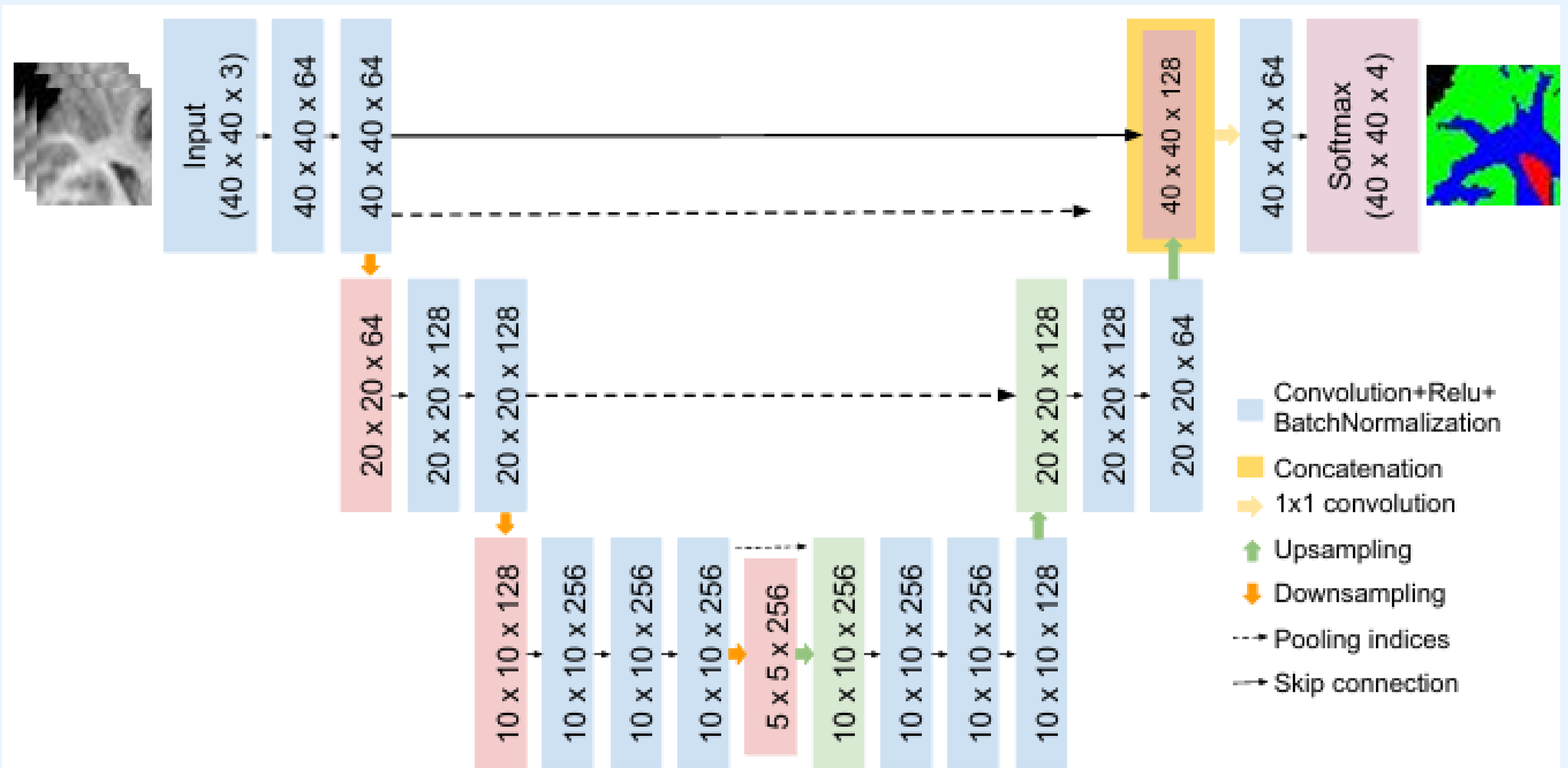


Figure 1. The original U-SegNet architecture proposed in [3] combining skip connections from U-Net, pooling indices pass from SegNet

We explore this U-SegNet architecture further by testing 3 architectures:

- 1 skip connection applied at the highest level of the decoder
- 2 skip connections applied at the highest and middle level of the encoder
- 3 skip connections, one for each decoder block

The goal was to observe a trade-off between the number of parameters and the performance of the model (Dice score, convergence speed, accuracy). We also develop another architecture: the residual U-SegNet, a hydrib architecture between the Residual U-Net and the U-SegNet combining residual blocks in the encoder and two skip connections in the decoder.

Datasets

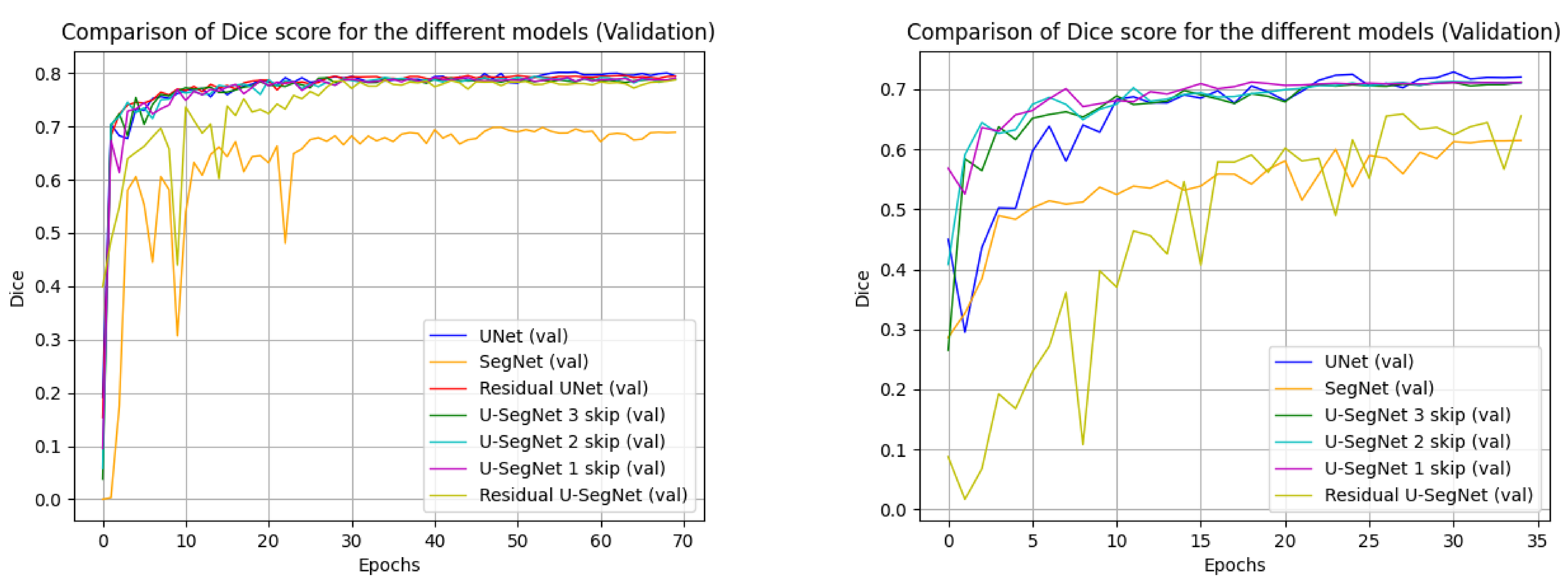
- "Retina Blood Vessel Segmentation" [2] containing 100 images (512x512x3) with manual mask annotations (challenges: fine blood vessel details require precise boundary delineation and small dataset)
- "Kidney Blood Vessel: 3D hierarchical phase-contrast tomography (HiP-CT) scans of human kidneys [1]: we use 285 grayscale images (256x256)

Implementation

- **Preprocessing**: We apply data augmentation on the training set (using rotation, flipping, cropping, bluring,..) to improve generalization, we split the datasets in 80% training and 20% validation
- **Loss Function**: DiceBCE Loss (a combination between dice loss and Binary Cross entropy which may be better for imbalanced datasets)

Results and discussion

To validate the efficacy of the new models we first observe the Dice score and accuracy across all models and over epochs. Due to GPU-time constraints we only trained the models for 35 epochs on the Kidney Dataset. However, we can observe a faster convergence for the U-SegNet models with 2 and 3 skip connections compared to the U-Net, SegNet and Residual U-Segnet models. Notably, all models, except the SegNet and Residual U-SegNet, achieved comparable Dice scores (mean value over batches) after 10 epochs on both datasets.



(a) Validation Dice score for the Retina Dataset (b) Validation Dice score for the Kidney Dataset

Figure 2. Comparison of model Dice Score on the validation set

Quantitative Results : Models comparison

We evaluated the performance of the models using the Dice score and accuracy metrics and balanced it with the number of parameters of the model.

Table 1. (a) Retina Blood Vessel Dataset

Model	Dice	Accuracy	Parameters
U-Net (4 levels)	**0.7856**	0.9620	31,043,521
Residual U-Net	0.7788	0.9631	13,043,009
SegNet (5 levels)	0.6876	0.9580	29,443,585
U-SegNet (3 skips)	0.7763	0.9631	4,247,809
U-SegNet (2 skips)	0.7776	**0.9640**	3,657,985
U-SegNet (1 skip)	0.7752	0.9633	3,510,529
Residual U-SegNet	0.7724	0.9576	8,126,665

Table 2. (b) Kidney Blood Vessel Dataset

Model	Dice	Accuracy	Parameters
U-Net (4 levels)	0.7392	0.9971	31,042,369
Residual U-SegNet	0.6684	0.9957	8,124,357
SegNet (5 levels)	0.6287	0.9959	29,442,433
U-SegNet (3 skips)	0.7231	0.9969	4,246,657
U-SegNet (2 skips)	0.7231	0.9969	3,656,833
U-SegNet (1 skip)	0.7230	0.9969	3,509,377

Table 3. Comparison of Models on Retina and Kidney Datasets: Trade-off Between Performance and Number of Parameters.

We implemented a U-Net with 4 levels (4 encoder and 4 decoder blocks), and a bottleneck, as well as a SegNet with 5 levels. While U-Net achieved the highest overall Dice score (which may be due to its greater depth), the U-SegNet (3 levels) with two skip connections demonstrated the best accuracy and a higher Dice score compared to other U-SegNet variants. However, the Dice scores across all U-SegNet variants are close and approximately 1% lower than U-Net. On the other hand, SegNet shows a significantly lower Dice score. When balancing these results with the number of parameters, the U-SegNet with 2 skip connections demonstrates a highly competitive performance with 8 times fewer parameters than the U-Net model and just 1% lower Dice score. Thus, this architecture offers a better balance between model complexity and segmentation performance, making it more suitable for usage with limited memory and/or computational resources.

Visual results

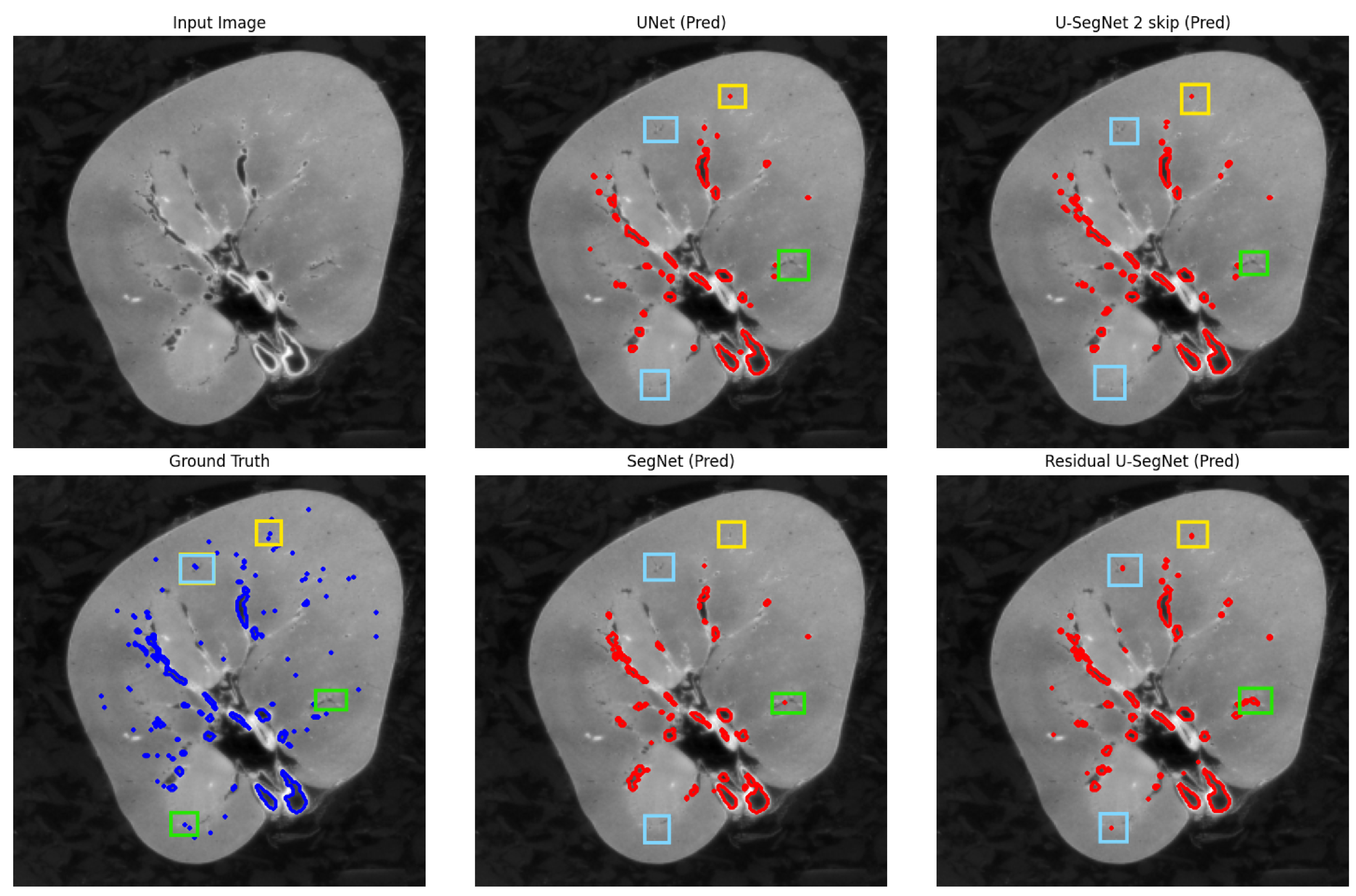


Figure 3. Visualization of the predictions on the Kidney Vessel Dataset. U-SegNet variants capture more details than SegNet (yellow rectangles). We also observed that Residual U-Segnet captures even more details than U-Net (and other models) (see blue rectangles) however it also captures false blood vessel (green rectangles)

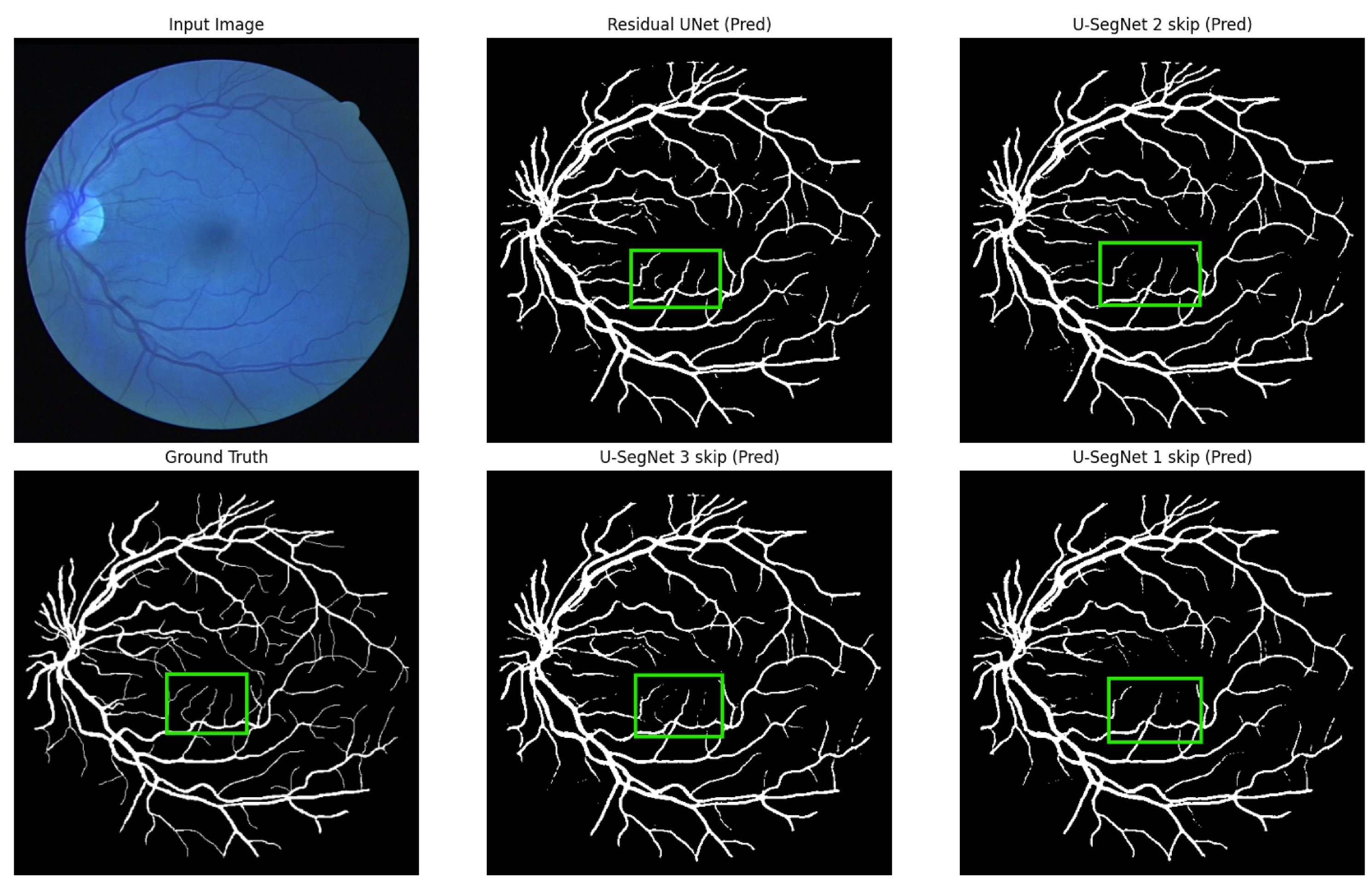


Figure 4. Visualization of the predictions on the Retina Vessel Dataset. Among U-SegNet variants U-SegNet with 2 skip connections seems to capture finer details compared to U-Segnet with 1 and 3 connections. However it does not seem to perform better than U-Net

By visualizing the results on the Retina and Kidney Dataset, we observed that U-SegNet models capture finer details than SegNet. The Residual U-SegNet may capture even finer details than the U-SegNet models, while having lower Dice score, but seems more prompt to hallucinations. Overall U-Segnet succeeds in capturing finer details similar to the U-Net.

Conclusion and further work

The U-SegNet architecture offers a favorable trade-off between performance and the number of parameters, effectively capturing finer details with the added skip connections. However using 3 skip connections may not improve the model further. Unlike in the paper, we did not utilize any pretrained models; however, to enhance our results, we would incorporate pretrained weights and fine-tune our models layer by layer, starting from the last. To further reduce the number of parameters, we could consider using a smaller model, using many overlapping patches from the original images with a smaller (40x40patches instead of 256x256 images), as suggested in the article, since local information may be more important than global information.

[1] https://www.kaggle.com/competitions/blood-vessel segmentation/data. Segment vasculature in 3d scans of human kidney.

[2] https://www.kaggle.com/datasets/abdallahwagih/retina-blood vessel. Retina blood vessel segmentation.

[3] Pulkit Kumar, Pravin Nagar, Chetan Arora, and Anubha Gupta. U-segnet: Fully convolutional neural network based automated brain tissue segmentation tool. 2018.