Paper Review: UNet++: A Nested U-Net Architecture for Medical Image Segmentation
By: Lam Nguyen

1. SUMMARY

This paper proposes an improved version of the U-Net and wide U-Net architectures. The U-Net architectures can be used for image processing and segmentation. These U-Net architectures consist basically of an encoder which takes a whole image and reduces it down to its bare essence and a decoder section which takes that bare essence of an image and attempts to convert it back to the original image. Although the regular U-Net and wide U-Net perform adequately well…. When it comes to medical imaging, adequate is not good enough because if there is a misidentification or failure to identify something in an image intended for diagnosing a patient… That patient could receive the wrong treatment or no treatment and have their well-being harmed. Due to these higher stakes… A more accurate image-segmentation architecture is proposed. The UNet++ architecture is a deeply supervised encoder-decoder network. The difference this new architecture has versus the older UNet architectures are that UNet++ has nested and dense skip connections. Skip connections skip over certain layers in a neural network to prevent vanishing gradients, to provide past information to further layers in the network and other benefits.

What distinguishes the UNet++ architecture from U-Net is the redesigned skip pathways that connect the subnetworks in the model as well as deep supervision. In normal U-Net architecture, the feature maps from the encoder go directly to the decoder. However, the new architecture has the feature maps go through a dense convolutional layer before reaching the decoder. The extra convolution makes it easier for the optimizer part of the code to do its job. These convolution maps sort of "even things out"

Deep Supervision is the next enhancement in UNET++. It allows the model to operate in a slower and more accurate mode or a faster and less accurate mode. Also, it enables model pruning to increase efficiency and performance.

To test the performance of UNET++ against U-NET and wide U-NET. Medical imaging datasets covering lesions/organs were used. Wide U-NET consistently outperformed regular U-NET due to having more parameters. U-NET++ without supervision yielded an average improvement of 2.8-3.3 points in IoU over the previous base models. And UNET++ with supervision yielded an extra 0.6 points improvement on top of that.

The experimenters were able to prune the model, reducing inference time by 32.2% while degrading the IoU score by 0.6 points. This was the maximum amount of pruning without degrading the accuracy excessively.

2. STRENGTHS

- The author(s) did a good job in clearly stating why and how this model was an improvement over the other U-NET architectures which was the use of modified skip connections and deeply supervised learning

- The relevance of UNET++ in creating a more accurate method of image segmentation was given. It was made clear that medical imaging really needs an accurate algorithm or lese lives would be at stake.
- The new proposed method uses less parameters than the Wide U-Net while offering a greater performance when segmenting cell nuclei, liver, polyps and liver nodules.
- The new segmentation method can be run on a relatively small amount of computing resources as of 2025, which is 12GB of VRAM and a Nvidia Titan GPU.
- Pruning the model from 9.0 million parameters down to 0.5 million parameters only drops the accuracy down a couple of percentage points for all the tested image types except for Liver Polyps images.

3. WEAKNESSES

- If there is are any downsides in terms of computing costs or something else, this could be further discussed.
- It might be beyond the scope of the paper but it would be good to have a link to the code and GitHub repository in order to use this and test things out on our own computers
- The effectiveness of UNET++ against other image segmentation models should be discussed.

4. TECHNICAL EXTENSIONS
- It would be worth testing U-NET against other image segmentation models.
- An explanation could be given on why all the U-NET models did poorly on Colon Polyps. This would reveal new realms of research on why the algorithms work as they do and how to improve them.
- The algorithm already runs on reasonable hardware. Maybe further research could be done on model pruning and supervision in order to further allow these models to run on less and less powerful hardware while still offering great performance.

5. OVERALL REVIEW

The paper offers a new and incremental improvement in image segmentation using U-NET. It uses less parameters, while offering more performance compared to other UNET architectures.

It would be good to include the code for this architecture in order to reproduce the results. The experiment design was simple and easy to follow although it could be discussed why a certain image type did so poorly on segmentation and if other algorithms would do better for that image type.

Enhanced Skip Connections and/or Deep Supervision which was used to enhance UNET, can be applied to other algorithms to test their effectiveness.