

[Get started](#)[Open in app](#)

Sik-Ho Tsang

[Follow](#)

4.95K Followers

[About](#)

Review: UNet++ — A Nested U-Net Architecture (Biomedical Image Segmentation)

Outperforms U-Net and Wide U-Net



[Sik-Ho Tsang](#) Oct 1, 2019 · 5 min read

In this story, UNet++, by **Arizona State University**, is reviewed. UNet++ **uses the Dense block ideas from DenseNet to improve U-Net**. UNet++ differs from the original U-Net in three ways:

- **1) having convolution layers on skip pathways**, which bridges the semantic gap between encoder and decoder feature maps.
- **2) having dense skip connections on skip pathways**, which improves gradient flow.
- **3) having deep supervision, which enables model pruning** and improves or in the worst case achieves comparable performance to using only one loss layer.

This is a **2018 DLMIA** paper with more than **40 citations**. ([Sik-Ho Tsang @ Medium](#))

Outline

1. UNet++ Architecture

Get started

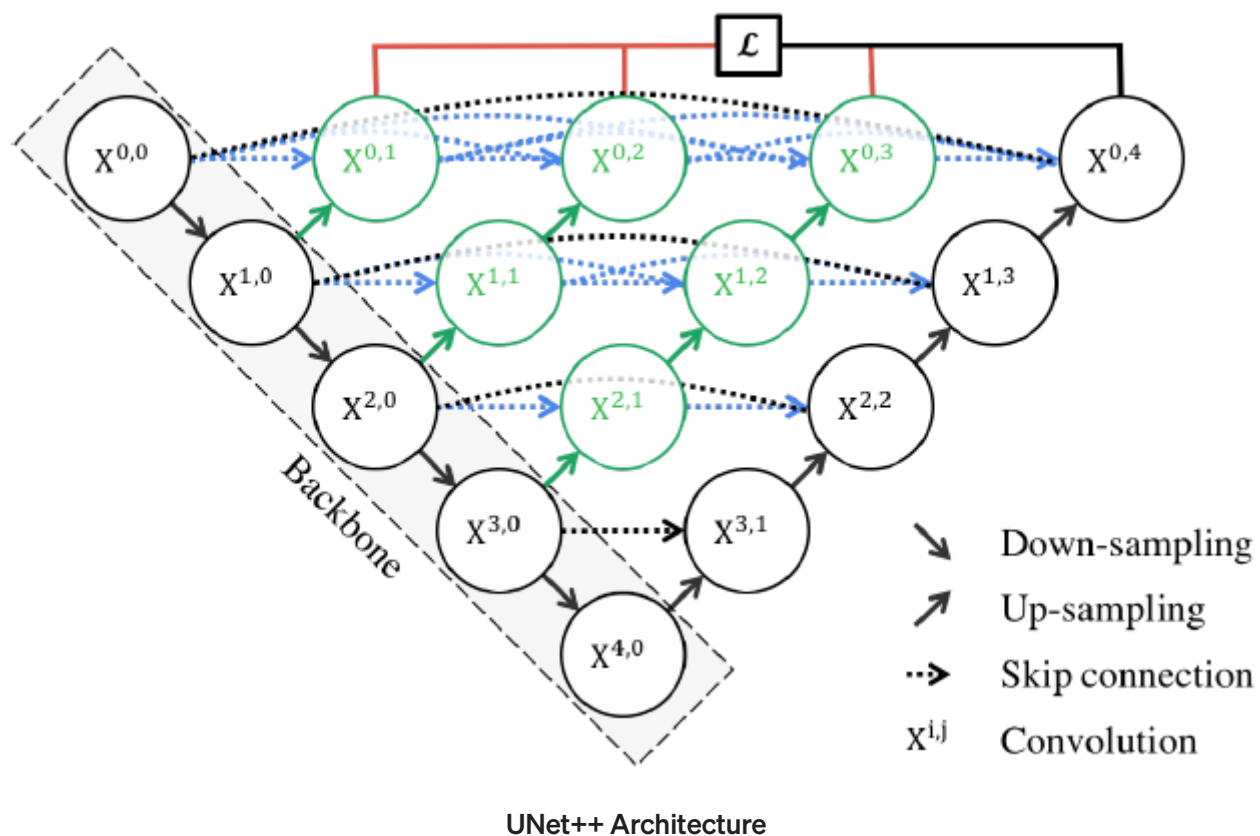
Open in app



3. Deep Supervision

4. Experimental Results

1. UNet++ Architecture

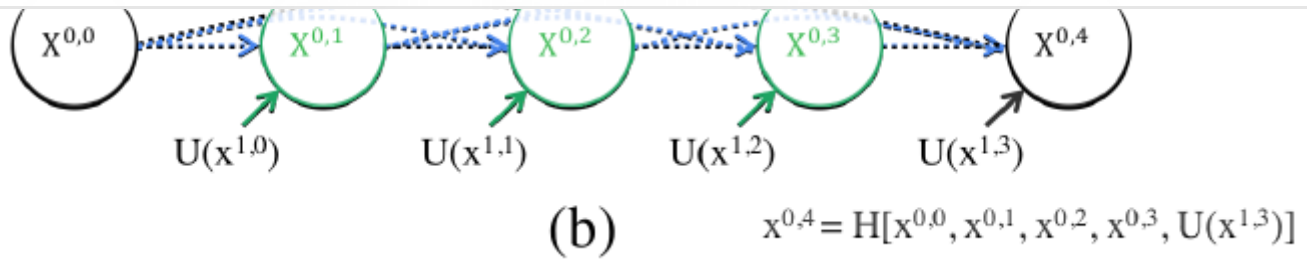


- UNet++ starts with an encoder sub-network or backbone followed by a decoder sub-network.
- There are **re-designed skip pathways (green and blue)** that connect the two sub-networks and the use of **deep supervision (red)**.

2. Re-designed Skip Pathways

Get started

Open in app



Re-designed Skip Pathways

- The above figure shows an example how the feature maps travel through the top skip pathway of UNet++.
- Another example, consider the skip pathway between nodes $X^{0,0}$ and $X^{1,3}$, as shown in the first figure. The skip pathway consists of a dense convolution block with three convolution layers.
- Each convolution layer is preceded by a concatenation layer that fuses the output from the previous convolution layer of the same dense block with the corresponding up-sampled output of the lower dense block.
- Formally, we can formulate as follows:

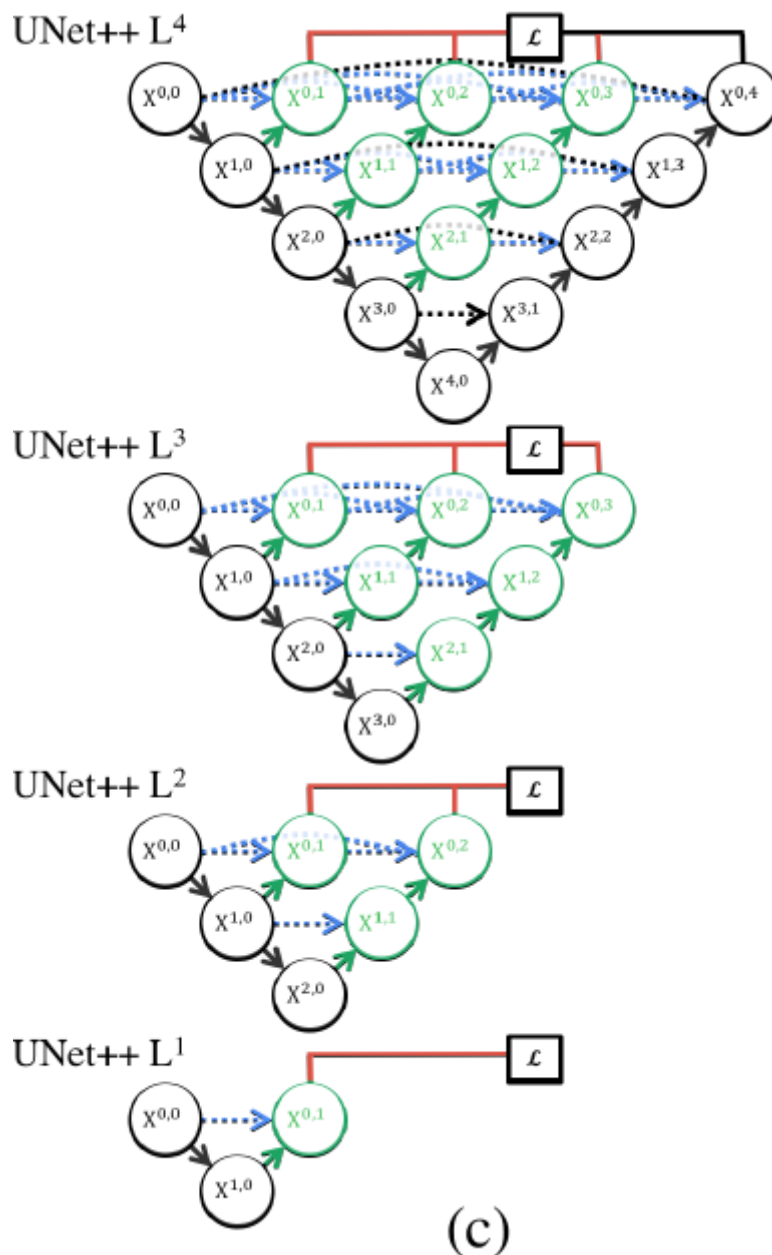
$$x^{i,j} = \begin{cases} \mathcal{H}(x^{i-1,j}), & j = 0 \\ \mathcal{H}\left(\left[\left[x^{i,k}\right]_{k=0}^{j-1}, \mathcal{U}(x^{i+1,j-1})\right]\right), & j > 0 \end{cases}$$

- where $H()$ is a convolution operation followed by an activation function, $U()$ denotes an up-sampling layer, and $[]$ denotes the concatenation layer.
- This is the idea from **DenseNet**.

The main idea behind is to bridge the semantic gap between the feature maps of the encoder and decoder prior to fusion.

Get started

Open in app



Deep Supervision

- With deep supervision:

accurate mode wherein the outputs from all segmentation branches are averaged.

Or fast mode wherein the nal segmentation map is selected from only one of the segmentation branches, the choice of which determines the extent of model pruning and speed gain.

[Get started](#)[Open in app](#)

- Also, a combination of binary cross-entropy and dice coefficient as the loss function:

$$\mathcal{L}(Y, \hat{Y}) = -\frac{1}{N} \sum_{b=1}^N \left(\frac{1}{2} \cdot Y_b \cdot \log \hat{Y}_b + \frac{2 \cdot Y_b \cdot \hat{Y}_b}{Y_b + \hat{Y}_b} \right)$$

- where N is the batch size.

4. Experimental Results

4.1. Datasets

Dataset	Images	Input Size	Modality	Provider
cell nuclei	670	96×96	microscopy	Data Science Bowl 2018
colon polyp	7,379	224×224	RGB video	ASU-Mayo [10,11]
liver	331	512×512	CT	MICCAI 2018 LiTS Challenge
lung nodule	1,012	64×64×64	CT	LIDC-IDRI [1]

Datasets

- Four medical imaging datasets are used for model evaluation, covering lesions/organs from different medical imaging modalities.

4.2. Baseline Models

encoder / decoder	$X^{0,0}/X^{0,4}$	$X^{1,0}/X^{1,3}$	$X^{2,0}/X^{2,2}$	$X^{3,0}/X^{3,1}$	$X^{4,0}/X^{4,0}$
U-Net	32	64	128	256	512
wide U-Net	35	70	140	280	560

Number of Convolutional Kernels

Wide U-Net is the modified U-Net with more layers such that it has similar number of parameters with the UNet++.

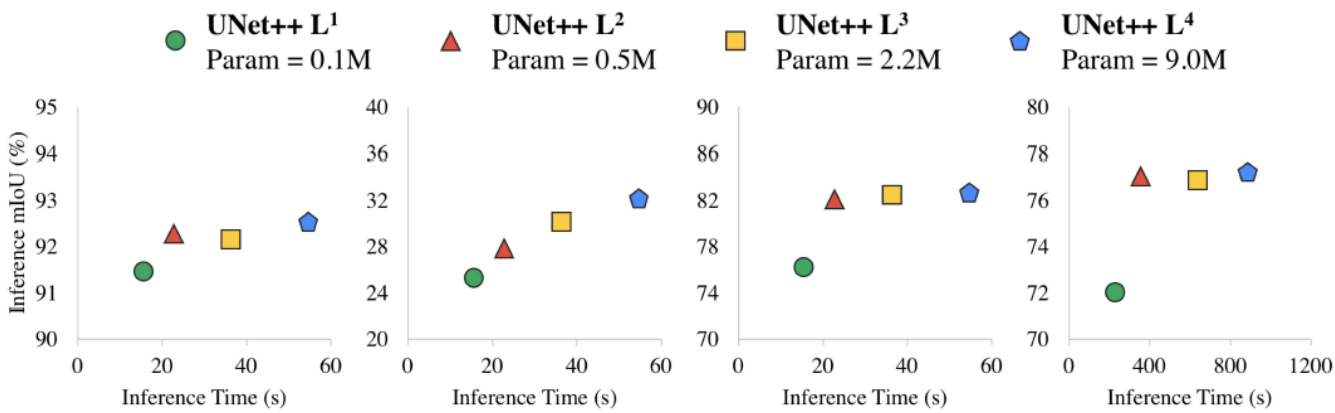
4.3. Results

Architecture	Params	Dataset			
		cell nuclei	colon polyp	liver	lung nodule
U-Net [9]	7.76M	90.77	30.08	76.62	71.47
Wide U-Net	9.13M	90.92	30.14	76.58	73.38
UNet++ w/o DS	9.04M	92.63	33.45	79.70	76.44
UNet++ w/ DS	9.04M	92.52	32.12	82.90	77.21

IoU (%), DS: Deep Supervision

- UNet++ without deep supervision achieves a significant performance gain over both U-Net and wide U-Net, yielding average improvement of 2.8 and 3.3 points in IoU.
- UNet++ with deep supervision exhibits average improvement of 0.6 points over UNet++ without deep supervision.

4.4. Model Pruning

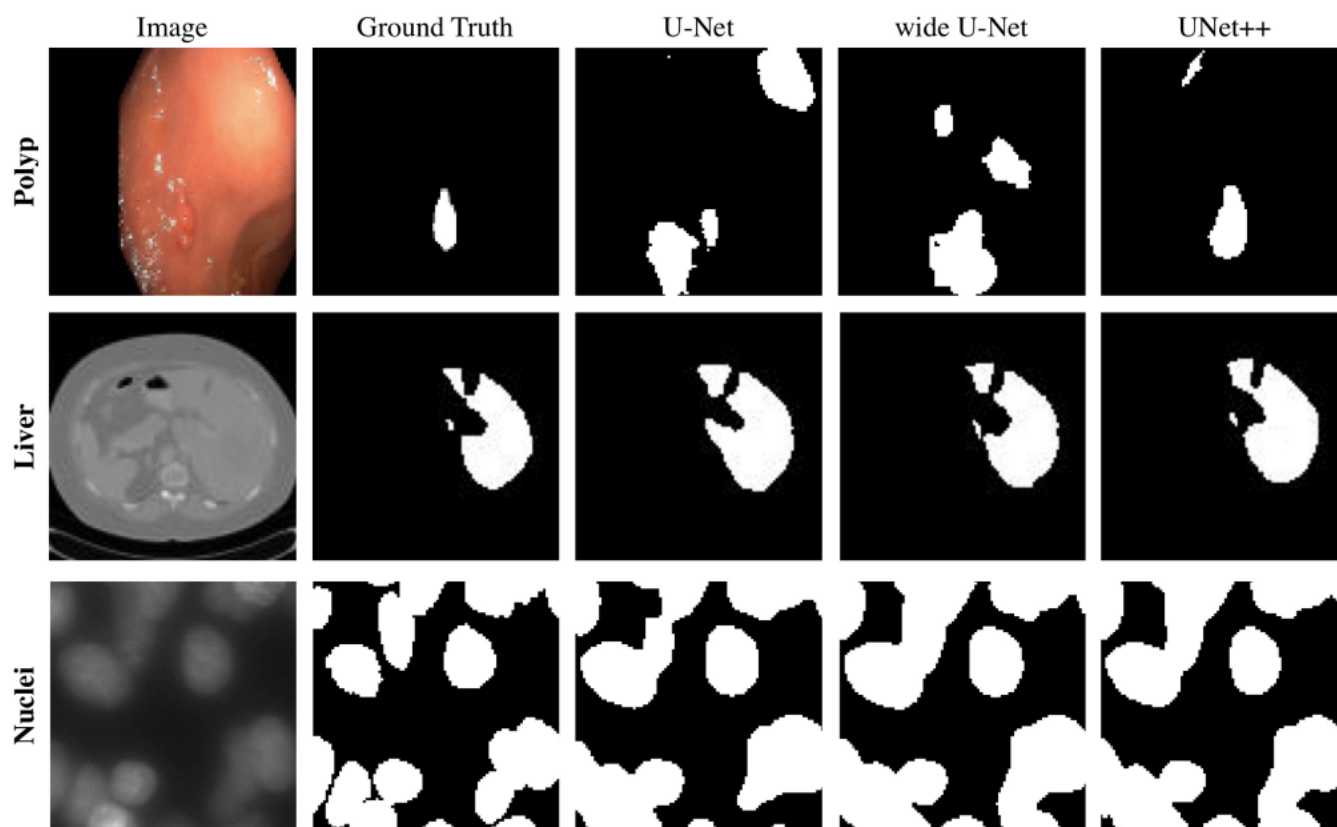


mIoU vs Inference Time for Model Pruning

- UNet++ L3 achieves on average 32.2% reduction in inference time while degrading IoU by only 0.6 points.

[Get started](#)[Open in app](#)

4.5. Qualitative Results



Qualitative Results

Around from 2017 to 2018 after [DenseNet](#), there are papers borrowed the [DenseNet](#) idea to improve the segmentation accuracy in Biomedical Image Segmentation including this paper and [DenseVoxNet](#).

Reference

[2018 DLMIA] [UNet++]

[UNet++: A Nested U-Net Architecture for Medical Image Segmentation](#)

[Get started](#)[Open in app](#)

[[SPPNet](#)] [[PReLU-Net](#)] [[STN](#)] [[DeepImage](#)] [[SqueezeNet](#)] [[GoogLeNet / Inception-v1](#)] [[BN-Inception / Inception-v2](#)] [[Inception-v3](#)] [[Inception-v4](#)] [[Xception](#)] [[MobileNetV1](#)] [[ResNet](#)] [[Pre-Activation ResNet](#)] [[RiR](#)] [[RoR](#)] [[Stochastic Depth](#)] [[WRN](#)] [[ResNet-38](#)] [[Shake-Shake](#)] [[FractalNet](#)] [[Trimps-Soushen](#)] [[PolyNet](#)] [[ResNeXt](#)] [[DenseNet](#)] [[PyramidNet](#)] [[DRN](#)] [[DPN](#)] [[Residual Attention Network](#)] [[DMRNet / DFN-MR](#)] [[IGCNet / IGCv1](#)] [[MSDNet](#)] [[ShuffleNet V1](#)] [[SENet](#)] [[NASNet](#)] [[MobileNetV2](#)]

Object Detection [[OverFeat](#)] [[R-CNN](#)] [[Fast R-CNN](#)] [[Faster R-CNN](#)] [[MR-CNN & S-CNN](#)] [[DeepID-Net](#)] [[CRAFT](#)] [[R-FCN](#)] [[ION](#)] [[MultiPathNet](#)] [[NoC](#)] [[Hikvision](#)] [[GBD-Net / GBD-v1 & GBD-v2](#)] [[G-RMI](#)] [[TDM](#)] [[SSD](#)] [[DSSD](#)] [[YOLOv1](#)] [[YOLOv2 / YOLO9000](#)] [[YOLOv3](#)] [[FPN](#)] [[RetinaNet](#)] [[DCN](#)]

Semantic Segmentation [[FCN](#)] [[DeconvNet](#)] [[DeepLabv1 & DeepLabv2](#)] [[CRF-RNN](#)] [[SegNet](#)] [[ParseNet](#)] [[DilatedNet](#)] [[DRN](#)] [[RefineNet](#)] [[GCN](#)] [[PSPNet](#)] [[DeepLabv3](#)] [[ResNet-38](#)] [[ResNet-DUC-HDC](#)] [[LC](#)] [[FC-DenseNet](#)] [[IDW-CNN](#)] [[DIS](#)] [[SDN](#)] [[DeepLabv3+](#)]

Biomedical Image Segmentation [[CUMedVision1](#)] [[CUMedVision2 / DCAN](#)] [[U-Net](#)] [[CFS-FCN](#)] [[U-Net+ResNet](#)] [[MultiChannel](#)] [[V-Net](#)] [[3D U-Net](#)] [[M²FCN](#)] [[SA](#)] [[QSA+QNT](#)] [[3D U-Net+ResNet](#)] [[Cascaded 3D U-Net](#)] [[VoxResNet](#)] [[DenseVoxNet](#)] [[Attention U-Net](#)] [[RU-Net & R2U-Net](#)] [[UNet++](#)]

Instance Segmentation [[SDS](#)] [[Hypercolumn](#)] [[DeepMask](#)] [[SharpMask](#)] [[MultiPathNet](#)] [[MNC](#)] [[InstanceFCN](#)] [[FCIS](#)]

Super Resolution [[SRCNN](#)] [[FSRCNN](#)] [[VDSR](#)] [[ESPCN](#)] [[RED-Net](#)] [[DRCN](#)] [[DRRN](#)] [[LapSRN & MS-LapSRN](#)] [[SRDenseNet](#)] [[SR+STN](#)]

Human Pose Estimation [[DeepPose](#)] [[Tompson NIPS'14](#)] [[Tompson CVPR'15](#)] [[CPM](#)]

Codec Post-Processing [[ARCNN](#)] [[Lin DCC'16](#)] [[IFCNN](#)] [[Li ICME'17](#)] [[VRCNN](#)] [[DCAD](#)] [[DS-CNN](#)]

Generative Adversarial Network [[GAN](#)]

[Get started](#)[Open in app](#)[Machine Learning](#)[Deep Learning](#)[Artificial Intelligence](#)[Data Science](#)[Biomedical Segmentation](#)[About](#) [Help](#) [Legal](#)

Get the Medium app

