

The diagram illustrates the proposed 3D-UNet architecture for medical image segmentation. It starts with a 3D volume of input slices (labeled R, G, B) for slices $n-1$, n , and $n+1$. These are combined into a 2D color image, which serves as the 3D-like input for slice n . The input is processed by a series of convolutional layers (blue arrows) and max pooling layers (red arrows). The main path consists of:

- Initial convolution: 3 channels, 1x1 kernel, repeated 16 times.
- Convolution: 64 channels, 3x3 kernel, repeated 16 times.
- Max pooling: 2x2 kernel.
- Convolution: 128 channels, 3x3 kernel, repeated 16 times.
- Max pooling: 2x2 kernel.
- Convolution: 256 channels, 3x3 kernel, repeated 16 times.
- Max pooling: 2x2 kernel.
- Convolution: 512 channels, 3x3 kernel, repeated 16 times.
- Max pooling: 2x2 kernel.
- Convolution: 512 channels, 3x3 kernel, repeated 16 times.

 The output of the final convolutional layer is a 16x16x16 volume. This is followed by a series of up-convolutional layers (green arrows) with skip connections (light blue arrows) from the corresponding convolutional layers:

- Up-convolution: 16x16x16 to 8x8x8.
- Up-convolution: 8x8x8 to 4x4x4.
- Up-convolution: 4x4x4 to 2x2x2.
- Up-convolution: 2x2x2 to 1x1x1.

 The final output is the segmentation map for slice n , shown as a white shape on a black background. A legend on the right defines the symbols:

- Blue arrow: conv 3x3, ReLU
- Red arrow: max pool 2x2
- Light blue arrow: concat + conv 1x1
- Purple arrow: soft max
- Green arrow: conv 3x3
- Green arrow with 'a': up-conv axa + crop 1xr

