

Diagram illustrating the U-Net architecture for image segmentation. The network consists of an encoder (left), a bottleneck (middle), and a decoder (right). The encoder processes the input image (128x128) through four stages of convolution and pooling, resulting in feature maps of sizes 64x64, 32x32, 16x16, and 8x8. The bottleneck is a single layer of 8x8 feature maps. The decoder then upsamples the feature maps through four stages, combining them with skip connections from the encoder to produce the final output (128x128). The output is a segmented image with different colors representing different classes.

Diagram illustrating a U-Net architecture for image segmentation. The input is a 128x128 image. The encoder path consists of four stages: 1) 64x64 convolution (3 filters, 12, 48, 48), 2) 32x32 convolution (2 filters, 96, 96), 3) 16x16 convolution (2 filters, 192, 192), and 4) 32x32 convolution (3 filters, 96, 96, 48). The decoder path consists of four stages: 1) 64x64 convolution (3 filters, 48, 48, 12), 2) 32x32 convolution (2 filters, 96, 96), 3) 16x16 convolution (2 filters, 192, 192), and 4) 64x64 convolution (3 filters, 12, 24, 24). The output is a 128x128 image. Skip connections are shown as lines connecting the output of each encoder stage to the corresponding decoder stage.

Diagram illustrating a U-Net architecture for image segmentation. The input is a 128x128 image. The encoder path consists of four stages: 1) 128x128 input, 2) 64x64 feature maps, 3) 32x32 feature maps, and 4) 16x16 feature maps. The decoder path consists of four stages: 1) 16x16 feature maps, 2) 32x32 feature maps, 3) 64x64 feature maps, and 4) 128x128 feature maps. Skip connections are shown as lines connecting the output of each encoder stage to the corresponding decoder stage. The final output is a 128x128 image.

Diagram illustrating a 3D U-Net architecture for video denoising. The input is a 128x128x128 volume. The encoder consists of four 3D convolutional layers with kernel sizes 1, 3, 12, and 24, resulting in feature maps of sizes 48, 48, 96, and 96. The bottleneck is a 32x32x32 volume. The decoder consists of four 3D convolutional layers with kernel sizes 96, 96, 96, and 247, resulting in feature maps of sizes 96, 96, 96, and 247. The output is a 128x128x128 volume.

The diagram illustrates the U-Net architecture, which consists of an encoder, a bottleneck, a decoder, and skip connections. The input is a 128x128 image. The encoder consists of a series of convolutional layers with decreasing spatial dimensions and increasing feature maps: 128x128 (1 channel), 64x64 (3 channels), 32x32 (12 channels), 16x16 (24 channels), 8x8 (48 channels), 4x4 (96 channels), 2x2 (96 channels), 1x1 (96 channels), 1x1 (96 channels), 1x1 (48 channels), and 1x1 (48 channels). The bottleneck is a single 1x1 convolutional layer with 48 channels. The decoder consists of a series of convolutional layers with increasing spatial dimensions and decreasing feature maps: 1x1 (48 channels), 2x2 (96 channels), 4x4 (48 channels), 8x8 (24 channels), 16x16 (12 channels), 32x32 (6 channels), 64x64 (3 channels), and 128x128 (1 channel). Skip connections are shown as arrows from the encoder to the decoder, connecting corresponding spatial dimensions. The output is a 128x128 image.

