U-net semantic segmentation

(https://www.youtube.com/watch?v=oLvmLJkmXuc&list=PLhhyoLH6IjfwqKKZhVLp7diKFxTmj4Q6s&index=9)

-the desired output should include localization(class label is supposed to . be assigned to each label)

-input input image

-output segmentation map

-the model has an contraction path where the model is downsampled

-and then an expansion path where the image is upsampled

-between the paths there are some skip connections

-the input is grayscale (has a single channel)

-firstly it uses 2 3X3 convolutions

-after there is a downsampling with maxpooling of k size 2 and stride 2

-this step is repeated 4 times

-then the model is upsampling by using a transpose convolution and this is concatenated to the last convolution using skip connection

-the model continuos by using again a double convolution

-Then upsample and concat with a skip conecction to the same level doubleconvolution used before

-This step is repeated one more time with the same rule

-finally the model does a 2\* 2X2 conv again

-and the las one is a one by one conv so that this doesn t changed the input channel but changes the number of channels (in case of more classes needed)