Name	Type	Output Size
input	RGB image	$3 \times 512 \times 512$
encoder1	$2\times$ conv$\to$BN$\to$ReLU	$64 \times 512 \times 512$
	max-pooling	$64 \times 256 \times 256$
encoder1	*repeat*	$128 \times 128 \times 128$
encoder2	$3\times$ conv$\to$BN$\to$ReLU	$256 \times 128 \times 128$
	max-pooling	$256 \times 64 \times 64$
encoder2	*repeat* $2\times$	$512 \times 16 \times 16$
decoder2	$3\times$ conv$\to$BN$\to$ReLU	$512 \times 16 \times 16$
	max-unpool	$512 \times 32 \times 32$
decoder2	*repeat* $2\times$	$128 \times 128 \times 128$
decoder1	$2\times$ conv$\to$BN$\to$ReLU	$64 \times 128 \times 128$
	max-unpool	$64 \times 256 \times 256$
decoder1	*repeat* (without last BN)	$64 \times 512 \times 512$
final	softmax	$C \times 512 \times 512$