

Name	Type	Output Size
input	RGB image	$3 \times 512 \times 512$
encoder1	2× conv→ReLU	$64 \times 512 \times 512$
	max-pooling	$64 \times 256 \times 256$
encoder1	<i>repeat</i> 3×	$512 \times 32 \times 32$
	Dropout with $p_i = 0.5$	$512 \times 32 \times 32$
center	2× conv→ReLU	$1024 \times 28 \times 28$
	Dropout with $p_i = 0.5$	$512 \times 28 \times 28$
	up-conv→ReLU	$512 \times 56 \times 56$
decoder1	concat. encoder maps	$1024 \times 56 \times 56$
	2× conv→ReLU	$512 \times 52 \times 52$
	up-conv→ReLU	$256 \times 104 \times 104$
decoder1	<i>repeat</i> 3× (no last up-conv)	$64 \times 388 \times 388$
final	conv	$C \times 388 \times 388$
	bilinear upsample	$C \times 512 \times 512$