

Layer	Type	Filters	Size/Stride	Output Size
0	Convolutional	32	$3 \times 3 / 1$	$416 \times 416 \times 32$
1	Maxpool		$2 \times 2 / 2$	$208 \times 208 \times 32$
2	Convolutional	64	$3 \times 3 / 1$	$208 \times 208 \times 64$
3	Maxpool		$2 \times 2 / 2$	$104 \times 104 \times 64$
4	Convolutional	128	$3 \times 3 / 1$	$104 \times 104 \times 128$
5	Convolutional	64	$1 \times 1 / 1$	$104 \times 104 \times 64$
6	Convolutional	128	$3 \times 3 / 1$	$104 \times 104 \times 128$
7	Maxpool		$2 \times 2 / 2$	$52 \times 52 \times 64$
8	Convolutional	256	$3 \times 3 / 1$	$52 \times 52 \times 256$
9	Convolutional	128	$1 \times 1 / 1$	$52 \times 52 \times 128$
10	Convolutional	256	$3 \times 3 / 1$	$52 \times 52 \times 256$
11	Maxpool		$2 \times 2 / 2$	$26 \times 26 \times 256$
12	Convolutional	512	$3 \times 3 / 1$	$26 \times 26 \times 512$
13	Convolutional	256	$1 \times 1 / 1$	$26 \times 26 \times 256$
14	Convolutional	512	$3 \times 3 / 1$	$26 \times 26 \times 512$
15	Convolutional	256	$1 \times 1 / 1$	$26 \times 26 \times 256$
16	Convolutional	512	$3 \times 3 / 1$	$26 \times 26 \times 512$
17	Convolutional	1024	$3 \times 3 / 1$	$26 \times 26 \times 1024$
18	Convolutional	1024	$3 \times 3 / 1$	$26 \times 26 \times 1024$
19	Passthrough		$10 \rightarrow 20$	$26 \times 26 \times 1024$
20	Convolutional	1024	$3 \times 3 / 1$	$26 \times 26 \times 1024$
21	Convolutional	$N_f$	$1 \times 1 / 1$	$26 \times 26 \times N_f$