

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 1 \times 32$	$96 \times 96 \times 1$
Conv dw / s1	$3 \times 3 \times 32$ dw	$48 \times 48 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$48 \times 48 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$48 \times 48 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$24 \times 24 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 24 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$24 \times 24 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$24 \times 24 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$12 \times 12 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$12 \times 12 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$12 \times 12 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$12 \times 12 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$6 \times 6 \times 256$
5×Conv dw / s1	$3 \times 3 \times 512$ dw	$6 \times 6 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 512$	$6 \times 6 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$6 \times 6 \times 512$
Avg Pool / s1	Pool 6×6	$6 \times 6 \times 512$
FC / s1	$512 \times l$	$1 \times 1 \times 512$
FC / s1	$l \times c$	$1 \times 1 \times l$
Softmax / s1	Classifier	$1 \times 1 \times c$