Layer	Activation size
Input	$64(128) \times 16 \times 16$
$1 \times 1 \times 512$ conv, pad 0, stride 1	$512 \times 16 \times 16$
Residual block, 512 filters	$512 \times 16 \times 16$
Residual block, 512 filters	$512 \times 16 \times 16$
Depth to Space, stride 2	$128 \times 32 \times 32$
$3 \times 3 \times 256$ conv, pad 1, stride 1	$256 \times 32 \times 32$
Residual block, 256 filters	$256 \times 32 \times 32$
Depth to Space, stride 4	$128 \times 32 \times 32$
$3 \times 3 \times 32$ conv, pad 1, stride 1	$32 \times 128 \times 128$
$3 \times 3 \times 3$ conv, pad 1, stride 1	$3 \times 128 \times 128$