Appendix A:

Table 5: Network architectures, in which we replace Conv with LinearConv in our experiments. $[.]^k$ represents k repetitions of blocks having shortcut connections, and C stands for the number of groups in group convolutions. Each Conv layer is represented as $h_l \times w_l, f_l$. Output is the spatial dimension of feature maps $H_l \times W_l$.

Output	Base	VGG11 [46]	AllConv [48]	ResNet-18 [19]	ResNeXt-29 [56]	MobileNetV2 [44]
32×32	-	-	$3 \times 3,\ 96$ $3 \times 3,\ 96$	$3 \times 3,\ 64$ $\left[\begin{array}{l} 3 \times 3,\ 128 \\ 3 \times 3,\ 128 \end{array}\right]^2$	$1 \times 1,\ 64$ $\left[\begin{array}{l} 1 \times 1,\ 128 \\ 3 \times 3,\ 128\ ;\ C = 2 \\ 1 \times 1,\ 256 \end{array}\right]^3$	$3 \times 3,\ 32$ $\left[\begin{array}{l} 1 \times 1,\ 32 \\ 3 \times 3,\ 32\ ;\ C = 32 \\ 1 \times 1,\ 16 \end{array}\right]^1$ $\left[\begin{array}{l} 1 \times 1,\ 96 \\ 3 \times 3,\ 96\ ;\ C = 96 \\ 1 \times 1,\ 24 \end{array}\right]^2$
16×16	$3 \times 3,\ 32$	$3 \times 3,\ 64$	$3 \times 3,\ 96$ $3 \times 3,\ 192$ $3 \times 3,\ 192$	$\left[\begin{array}{l} 3 \times 3,\ 128 \\ 3 \times 3,\ 128 \end{array}\right]^2$	$\left[\begin{array}{l} 1 \times 1,\ 256 \\ 3 \times 3,\ 256\ ;\ C = 2 \\ 1 \times 1,\ 512 \end{array}\right]^3$	$\left[\begin{array}{l} 1 \times 1,\ 144 \\ 3 \times 3,\ 144\ ;\ C = 144 \\ 1 \times 1,\ 32 \end{array}\right]^3$
8×8	$3 \times 3,\ 64$	$3 \times 3,\ 128$ $3 \times 3,\ 256$	$3 \times 3,\ 192$ $3 \times 3,\ 192$ $1 \times 1,\ 192$ $1 \times 1,\ 10$	$\left[\begin{array}{l} 3 \times 3,\ 256 \\ 3 \times 3,\ 256 \end{array}\right]^2$	$\left[\begin{array}{l} 1 \times 1,\ 512 \\ 3 \times 3,\ 512\ ;\ C = 2 \\ 1 \times 1,\ 1024 \end{array}\right]^3$	$\left[\begin{array}{l} 1 \times 1,\ 196 \\ 3 \times 3,\ 196\ ;\ C = 196 \\ 1 \times 1,\ 64 \end{array}\right]^4$ $\left[\begin{array}{l} 1 \times 1,\ 384 \\ 3 \times 3,\ 384\ ;\ C = 384 \\ 1 \times 1,\ 96 \end{array}\right]^3$
4×4	$3 \times 3,\ 128$	$3 \times 3,\ 256$ $3 \times 3,\ 512$	-	$\left[\begin{array}{l} 3 \times 3,\ 512 \\ 3 \times 3,\ 512 \end{array}\right]^2$	-	$\left[\begin{array}{l} 1 \times 1,\ 576 \\ 3 \times 3,\ 576\ ;\ C = 576 \\ 1 \times 1,\ 160 \end{array}\right]^3$ $\left[\begin{array}{l} 1 \times 1,\ 960 \\ 3 \times 3,\ 960\ ;\ C = 960 \\ 1 \times 1,\ 320 \end{array}\right]^1$ $1 \times 1,\ 1280$
2×2	$3 \times 3,\ 256$	$3 \times 3,\ 512$ $3 \times 3,\ 512$	-	-	-	-
1×1	10-d *fc*	$3 \times 3,\ 512$ 10-d *fc*	avg pool	avg pool 10-d *fc*	avg pool 10-d *fc*	avg pool 10-d *fc*