Appendix

1. Neural Network Architecture

Our CNN baseline is AlexNet. Table 1∼3 show the architectural details of SqueezeNet, Ours_F34, and Ours_I789, respectively. More specifically, the input channel numbers, output channel numbers, kernel size, stride size, and padding size in each layer of a neural network are listed. For PPNNs, each $ReLU$ activation is approximated by a degree-2 polynomial, and each max pooling is converted to an average pooling.

SqueezeNet is shown in Table 1, where Conv2-1 and Conv2-2 in each Fire module are concatenated instead stacked. Ours_F34 shown in Table 2 replaces the last two fire modules, i.e., F3 and F4, by C3 and C4. Table 3 shows Ours_I789 replaces the last three Inception modules in InceptionNet by C3, C4, and C5 convolutions. We use BN. in Table 3 to represent Batch Normalization. And the BN.+$ReLU$ after a convolution in an Inception module is hidden in Table 3. The pair of Conv2-1 and Conv2-2 is stacked. Similarly, the pair of Conv3-1 and Conv3-2 also is stacked. But Conv1, Conv2, Conv3 and Conv4 in each Inception are concatenated.

Table 1. The network architecture of SqueezeNet on CIFAR-10.

Block	Descriptions	Input Channel	Output Channel	Kernel Size	Stride	Padding
C1	Convolution	3	64	3	1	0
P1	Avg. Pool	-	-	3	2	0
F1	Conv1.	64	32	1	1	0
	$ReLU1$	-	-	-	-	-
	Conv2-1	32	64	1	1	0
	$ReLU2$-1	-	-	-	-	-
	Conv2-2.	32	64	3	1	0
	$ReLU2$-2	-	-	-	-	-
F2	Conv1.	128	32	1	1	0
	$ReLU1$	-	-	-	-	-
	Conv2-1	32	64	1	1	0
	$ReLU2$-1	-	-	-	-	-
	Conv2-2.	32	64	3	1	0
	$ReLU2$-2	-	-	-	-	-
P2	Avg. Pool	-	-	3	2	0
F3	Conv1.	128	32	1	1	0
	$ReLU1$	-	-	-	-	-
	Conv2-1	32	128	1	1	0
	$ReLU2$-1	-	-	-	-	-
	Conv2-2.	32	128	3	1	0
	$ReLU2$-2	-	-	-	-	-
F4	Conv1.	256	32	1	1	0
	$ReLU1$	-	-	-	-	-
	Conv2-1	32	128	1	1	0
	$ReLU2$-1	-	-	-	-	-
	Conv2-2.	32	128	3	1	0
	$ReLU2$-2	-	-	-	-	-
C2	Convolution	256	10	1	1	0
P3	Avg. Pool	-	-	-	-	0

Table 2. The network architecture of Our_F34 on CIFAR-10.

Block	Descriptions	Input Channel	Output Channel	Kernel Size	Stride	Padding
C1	Convolution	3	64	3	1	0
P1	Avg. Pool	-	-	3	2	0
F1	Conv1.	64	32	1	1	0
	$ReLU1$	-	-	-	-	-
	Conv2-1	32	64	1	1	0
	$ReLU2$-1	-	-	-	-	-
	Conv2-2.	32	64	3	1	0
	$ReLU2$-2	-	-	-	-	-
F2	Conv1.	128	32	1	1	0
	$ReLU1$	-	-	-	-	-
	Conv2-1	32	64	1	1	0
	$ReLU2$-1	-	-	-	-	-
	Conv2-2.	32	64	3	1	0
	$ReLU2$-2	-	-	-	-	-
P2	Avg. Pool	-	-	3	2	0
C2	Convolution	128	256	3	1	0
	$ReLU$	-	-	-	-	-
C3	Convolution	256	256	3	1	0
	$ReLU$	-	-	-	-	-
C4	Convolution	256	10	1	1	0
P3	Avg. Pool	-	-	-	-	0

Table 3. The network architecture of Our_I789 on CIFAR-10.

Block	Descriptions	Input Channel	Output Channel	Kernel Size	Stride	Padding
C1	Convolution	3	192	3	1	1
B1	BN.+ $ReLU$	-	-	3	2	0
I1	Conv1.	192	64	1	1	0
	Conv2-1.	192	96	1	1	0
	Conv2-2.	96	128	3	1	0
	Conv3-1.	192	16	1	1	0
	Conv3-2.	16	32	5	1	0
	Conv4.	192	32	1	1	0
I2	Conv1.	256	128	1	1	0
	Conv2-1.	256	128	1	1	0
	Conv2-2.	128	192	3	1	0
	Conv3-1.	256	32	1	1	0
	Conv3-2.	32	96	5	1	0
	Conv4.	256	64	1	1	0
P1	Avg. Pool	-	-	3	2	1
I3	Conv1.	480	192	1	1	0
	Conv2-1.	480	96	1	1	0
	Conv2-2.	96	208	3	1	0
	Conv3-1.	480	16	1	1	0
	Conv3-2.	16	48	5	1	0
	Conv4.	480	64	1	1	0
I4	Conv1.	512	128	1	1	0
	Conv2-1.	512	128	1	1	0
	Conv2-2.	128	192	3	1	0
	Conv3-1.	512	32	1	1	0
	Conv3-2.	32	96	5	1	0
	Conv4.	512	64	1	1	0
I5	Conv1.	512	128	1	1	0
	Conv2-1.	512	128	1	1	0
	Conv2-2.	128	256	3	1	0
	Conv3-1.	512	24	1	1	0
	Conv3-2.	24	64	5	1	0
	Conv4.	512	64	1	1	0
I6	Conv1.	512	112	1	1	0
	Conv2-1.	512	144	1	1	0
	Conv2-2.	144	288	3	1	0
	Conv3-1.	512	32	1	1	0
	Conv3-2.	32	64	5	1	0
	Conv4.	512	64	1	1	0
C3	Convolution	528	832	3	1	1
C4	Convolution	832	832	3	1	1
C5	Convolution	832	1024	3	1	1
P2	Avg. Pool	-	-	3	2	1
D1	Dense	1024	10	1	-	-