
Efficient Equivariant Network

Supplementary Materials

Anonymous Author(s)

Affiliation

Address

email

1 A Details of Data Efficient Analysis

2 As shown in table 5 and Figure 3, the test error of all models increase as training data size decrease.
3 Let’s take a closer look at the gap between different models. First, the gap between R18 and $p4$ -R18
4 is 2.17% at 5k and increases to 6.02% at 1k, while the gap between $p4$ -R18 and $p4m$ -R18 is 1.7% at
5 5k and increases to 3.69% at 1k. This indicates symmetry prior of model plays a more important
6 role when data is less. Then, we compare our models with their corresponding counterparts. The gap
7 between $p4$ -R18($p4m$ -R18) and $p4$ - E^4 R18($p4m$ - E^4 R18) is 1.11%(0.87%) at 5k, and it increase to
8 5.3%(5.22%) at 1k. Furthermore, in Figure 3, the red line($p4$ - E^4 R18) and blue dash line($p4m$ -R18)
9 intersect, which means that $p4$ - E^4 R18 is even more superior than $p4m$ -R18 when data is less although
10 the latter one has more capacity when we increase training data. The above results demonstrate
11 greater data efficiency of E^4 model compared with G-CNNs, or it can better utilize the symmetry
12 prior of data.

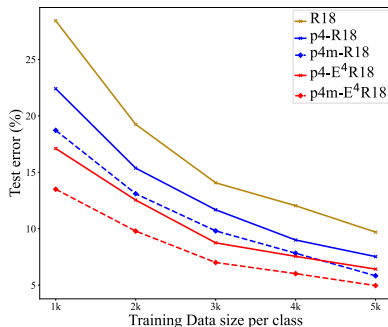


Figure 3: Trend of test error(%) on various training data sizes.

Table 5: Test error(%) of models trained on different training data sizes

Module/Dataset size	1k	2k	3k	4k	5k
R18	28.43	19.25	14.07	12.04	9.70
$p4$ -R18	22.41	15.37	11.68	9.01	7.53
$p4$ - E^4 R18(Ours)	17.11	12.54	8.75	7.55	6.42
$p4m$ -R18	18.72	13.10	9.81	7.68	5.83
$p4m$ - E^4 R18(Ours)	13.50	9.79	7.01	6.02	4.96

13 B MNIST-rot Model Architecture

Please refer to Table 6.

Table 6: Architecture of E^4 -Net on Mnist-rot classification, p means dropout rate.

Layer	Kernel size	Output channels
Group convolution BatchNorm+ReLU	3×3	16
E^4 -layer BatchNorm+ReLU Spatial-wise max pooling	5×5 2×2	16
E^4 -layer BatchNorm+ReLU+Dropout($p=0.1$)	5×5	16
E^4 -layer BatchNorm+ReLU+Dropout($p=0.1$) Spatial-wise max pooling	5×5 2×2	16
E^4 -layer BatchNorm+ReLU+Dropout($p=0.1$)	5×5	16
E^4 -layer BatchNorm+ReLU+Dropout($p=0.1$) lobal max group pooling layer	5×5	16
Fully connected+Softmax		10

14

15 Here, Group-wise Global max pooling is the operation act on each channel:

$$f_i = \max_{g \in \mathcal{G}} f_i(g), \quad (13)$$

16 and i denotes the channel index.

17 C Details of CIFAR10 and CIFAR100 Experiments

18 The ResNet-18 is composed of an initial convolution layer, followed by 4 stage Res-Blocks and one
 19 final classification layer. Each stage contain 2 Res-Blocks and each block contain 2 convolution
 20 layers. In each stage, the channel dimension change and spatial-wise pooling only occur at and after
 21 the first layer of the first block. From initial convolution layer to 4th stage Res-Blocks, the change of
 22 the output channel dimension is $64 \rightarrow 64 \rightarrow 128 \rightarrow 256 \rightarrow 512$, and spatial-wise pooling is carried out at
 23 2nd, 3rd, and 4th stage Res-Blocks. In our implementation, we replace all the conventional layers
 24 with $p4$ ($p4m$) convolutions layers without changing kernel size, and the output channel dimension
 25 is $32 \rightarrow 32 \rightarrow 64 \rightarrow 128 \rightarrow 256$ ($22 \rightarrow 22 \rightarrow 44 \rightarrow 88 \rightarrow 176$). Besides, we increase the width of each layer
 26 by $\sqrt{4}$ ($\sqrt{8}$) to keep the number of parameters almost invariant. This is the construction of $p4$ -R18
 27 ($p4m$ -R18). Then, for building our model $p4$ - E^4 R18 ($p4m$ - E^4 R18), the second group convolution
 28 layer in each Res-Block of $p4$ -R18 ($p4m$ -R18) is replaced by our E^4 -layer with $k = 3$, $r = 2$, and
 29 $s = C_l/2$, where C_l is the channel dimension at current layer.

30 References