Appendices

A. Proof to Theorem 1

Proof. From Definition 1, $\mathcal{A}$ is an isometry if and only if

$$\langle \mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{x}' \rangle = \langle \boldsymbol{x}, \boldsymbol{x}' \rangle, \ \ \forall \{\boldsymbol{x}, \boldsymbol{x}'\} \subseteq \mathbb{R}^{C\times H\times W}. \tag{19}$$

By the property of the adjoint operator, (19) is equivalent to

$$\langle \mathcal{A}^*\mathcal{A}\,\boldsymbol{x}, \, \boldsymbol{x}' \rangle = \langle \boldsymbol{x}, \boldsymbol{x}' \rangle, \ \ \forall \{\boldsymbol{x}, \boldsymbol{x}'\} \subseteq \mathbb{R}^{C\times H\times W}, \tag{20}$$

which holds if and only if

$$\mathcal{A}^*\mathcal{A}\,\boldsymbol{x} = \boldsymbol{x}, \ \ \forall \boldsymbol{x} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_C) \in \mathbb{R}^{C\times H\times W}. \tag{21}$$

By using (7) and (8), we rewrite $\mathcal{A}^*\mathcal{A}\,\boldsymbol{x}$ as

$$\begin{aligned} &\sum_{m=1}^{M}\sum_{c=1}^{C}\Big(\boldsymbol{\alpha}_{m1} * \big(\boldsymbol{\alpha}_{mc} \star \boldsymbol{\xi}_c\big), \ldots, \boldsymbol{\alpha}_{mC} * \big(\boldsymbol{\alpha}_{mc} \star \boldsymbol{\xi}_c\big)\Big) \\ = &\sum_{m=1}^{M}\sum_{c=1}^{C}\Big(\big(\boldsymbol{\alpha}_{mc}\star\boldsymbol{\alpha}_{m1}\big)*\boldsymbol{\xi}_c, \ldots, \big(\boldsymbol{\alpha}_{mc}\star\boldsymbol{\alpha}_{mC}\big)*\boldsymbol{\xi}_c\Big). \end{aligned} \tag{22}$$

In (22), we have used the fact that for arbitrary 2D signals $\boldsymbol{\alpha}, \boldsymbol{\alpha}'$ and $\boldsymbol{\xi}$,

$$\begin{aligned} &\boldsymbol{\alpha} * (\boldsymbol{\alpha}' \star \boldsymbol{\xi}) = (\boldsymbol{\alpha}' \star \boldsymbol{\xi}) * \boldsymbol{\alpha} \\ &\quad = \boldsymbol{\alpha}' \star (\boldsymbol{\alpha} * \boldsymbol{\xi}) = \boldsymbol{\alpha}' \star (\boldsymbol{\alpha} * \boldsymbol{\xi}) = (\boldsymbol{\alpha}' \star \boldsymbol{\alpha}) * \boldsymbol{\xi}, \end{aligned} \tag{23}$$

which follows from the commutative property of convolution (i.e., $\boldsymbol{\alpha} * \boldsymbol{\xi} = \boldsymbol{\xi} * \boldsymbol{\alpha}$) and the associative property of convolution and correlation (i.e., $\boldsymbol{\alpha}\star(\boldsymbol{\alpha}'*\boldsymbol{\xi}) = (\boldsymbol{\alpha}\star\boldsymbol{\alpha}')*\boldsymbol{\xi}$). By equating the last line of (22) with $\boldsymbol{x} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_C)$, we get that $\mathcal{A}$ is an isometry if and only if

$$\sum_{m=1}^{M}\sum_{c=1}^{C}\big(\boldsymbol{\alpha}_{mc}\star\boldsymbol{\alpha}_{mc'}\big)*\boldsymbol{\xi}_c = \boldsymbol{\xi}_{c'} \ \ \forall \boldsymbol{c}' \in \{1, \ldots, C\} \tag{24}$$

holds for all $\boldsymbol{x} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_C)$, which is equivalent to (9). Analogously, we can show that $\mathcal{A}^*$ is an isometry if and only if (10) holds. □

B. Additional Experiments

Isometric Components on ResNet. The isometric components are not heuristic engineering components but are specifically designed to promote neural network's isometric property. To validate this hypothesis, we show that naively adding those components to a standard ResNet34 will only have marginal or even negative impact on its performance (see Table 7). Standard ResNet, which has the BatchNorm layers, is not so sensitive to Delta initialization, as shown in Table 7 (b). Applying SReLU will however even decrease the performance since the nonlinearity effect is compromised by BatchNorm due to its effect of forcing output to have zero mean. In addition, imposing orthogonal regularization improves the performance by only a small margin.

	Method	SReLU	Delta Init.	Ortho. Reg.	Top-1 Accuracy (%)
(a)	ResNet				73.29
(b)		✓			72.92
(c)			✓		71.80
(d)				✓	73.60
(e)		✓	✓	✓	71.53

Table 7. ImageNet Top-1 accuracy for standard ResNet34 with additional isometric components. The results show that these components are specifically designed for imposing isometry of the network, instead of heuristic engineering components.