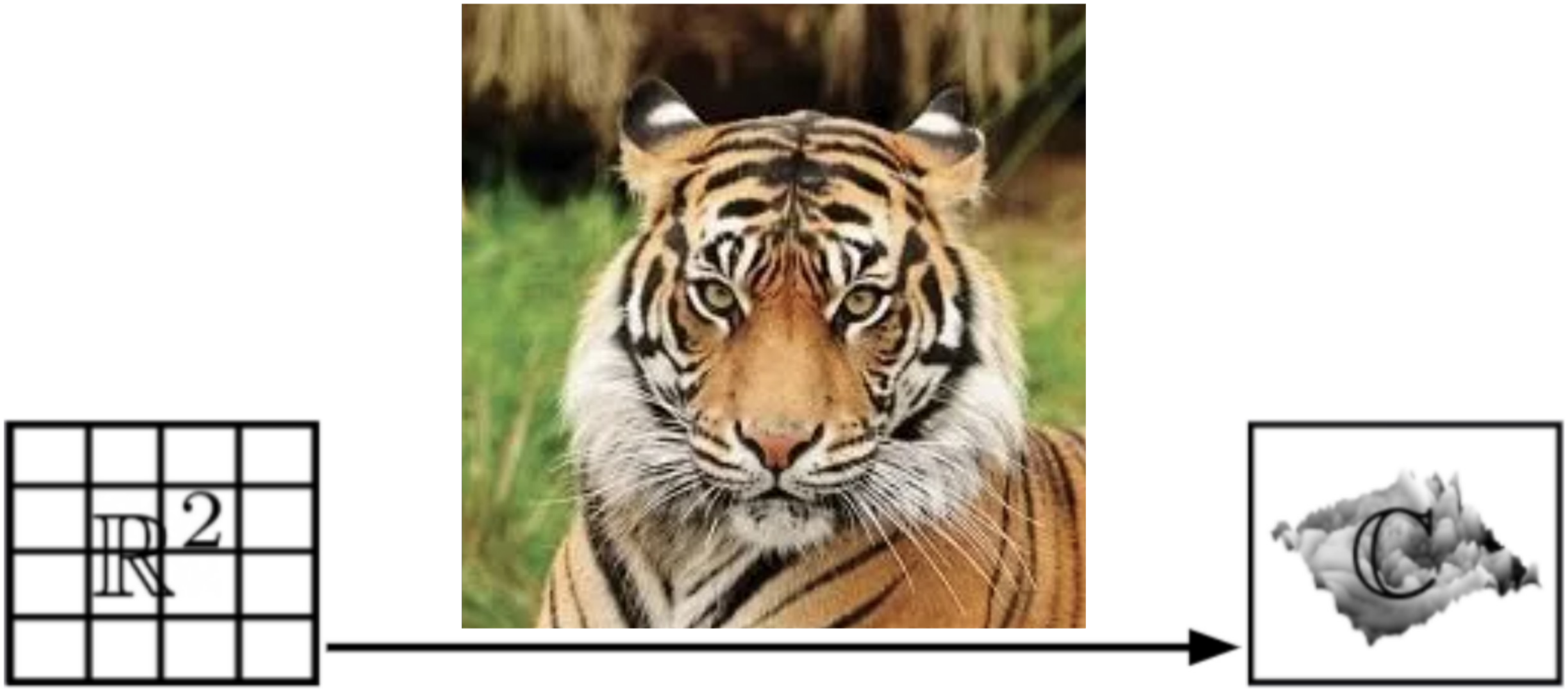
Co-Domain Symmetry for Complex-Valued Deep Learning

Utkarsh Singhal Yifei Xing Stella X. Yu

Berkeley UNIVERSITY OF CALIFORNIA ICSI

CVPR JUNE 19-24 2022 NEW ORLEANS LOUISIANA

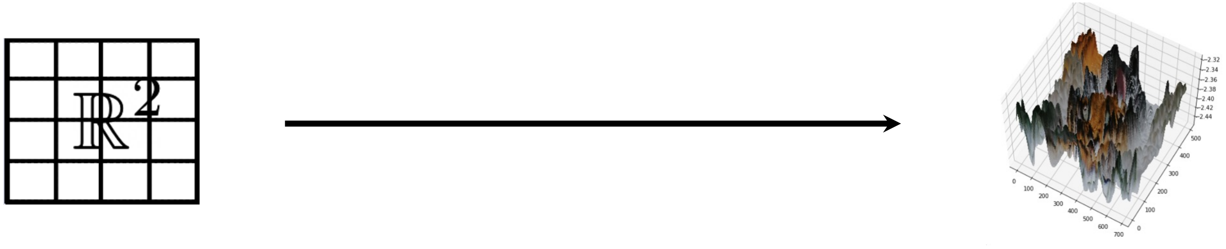
An Image is a Function from Domain to Co-Domain



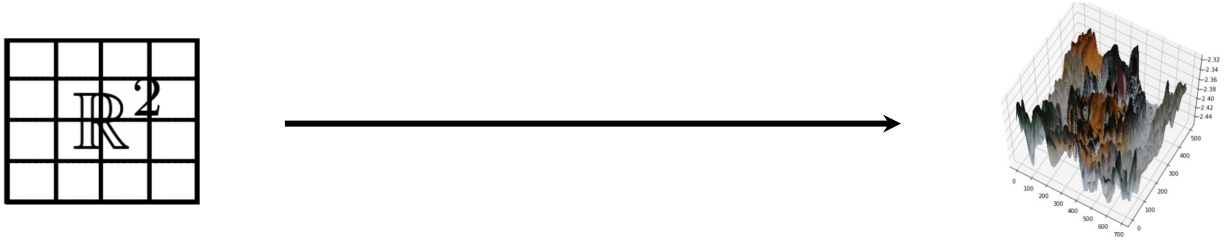
Domain: Pixel Locations

Co-Domain: Pixel Values

An Image is a Function from Domain to Co-Domain

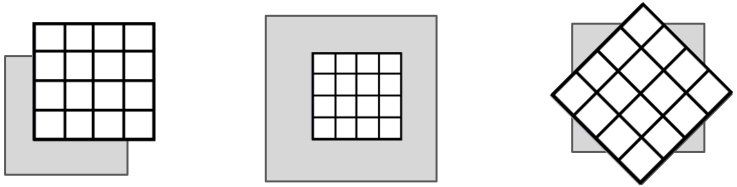


Domain Transformations Act on the Pixel *Coordinates*

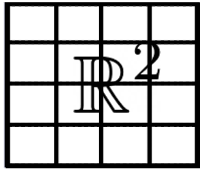


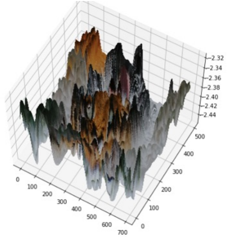
Domain Transformation

translation **scaling** **rotation**

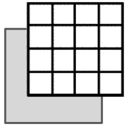
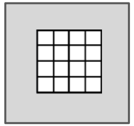
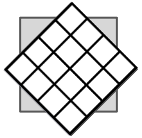


Domain Transformations Act on the Pixel *Coordinates*





Domain Transformation

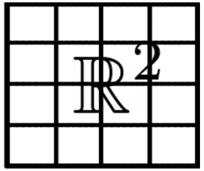
translation	scaling	rotation
		
CNN [1]	Scale-Invariant CNN [2]	E(2)-Steerable CNN [3]
		

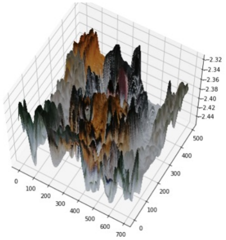
[1]: LeCun et al., Backpropagation Applied to Handwritten Zip Code Recognition

[2]: Xu et al., Scale-Invariance Convolutional Neural Network

[3]: Weiler et al., General E(2)-Equivariant Steerable CNNs

Co-Domain Transformations Act on the Pixel *Values*

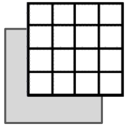




Domain Transformation

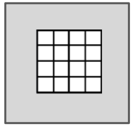
Co-domain transformation

translation



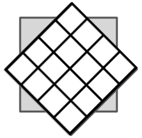
CNN [1]

scaling



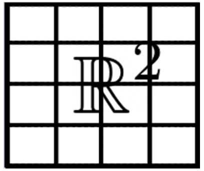
Scale-Invariant CNN [2]

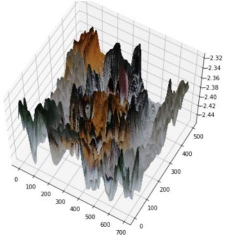
rotation



E(2)-Steerable CNN [3]

Co-Domain Transformations Act on the Pixel *Values*

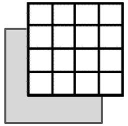




Domain Transformation

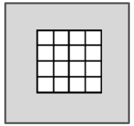
Co-domain Transformation

translation



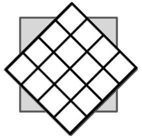
CNN [1]

scaling

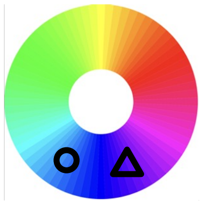


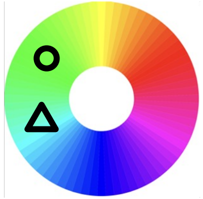
Scale-Invariant CNN [2]

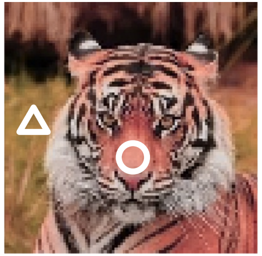
rotation



E(2)-Steerable CNN [3]



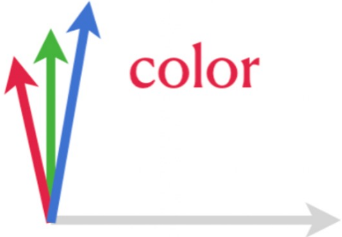




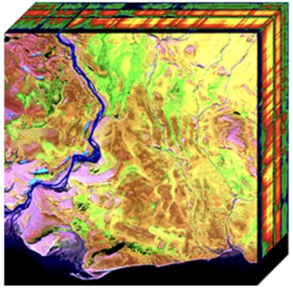
Co-Domain Encapsulates Diversity of Image Types

Thermal

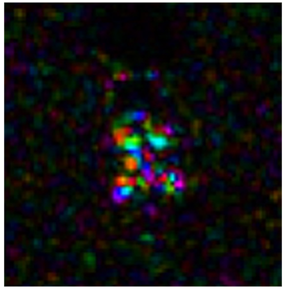
RGB



Hyperspectral

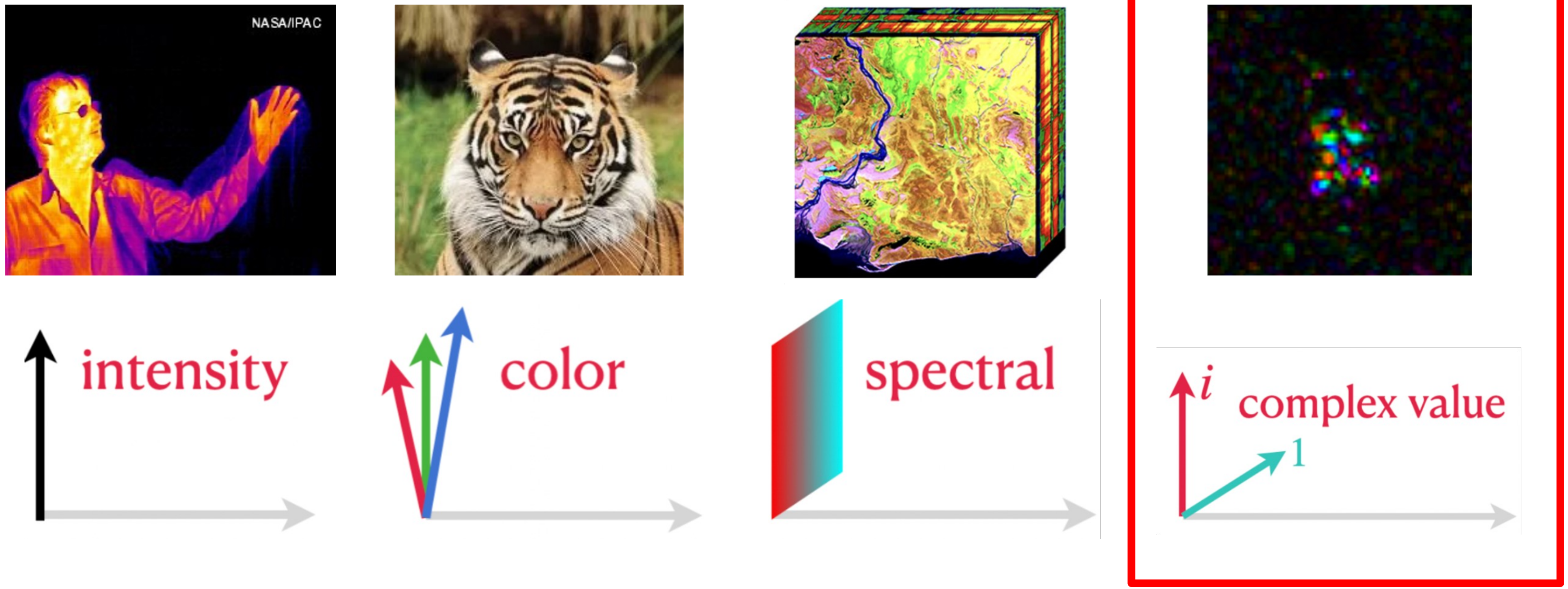


SAR

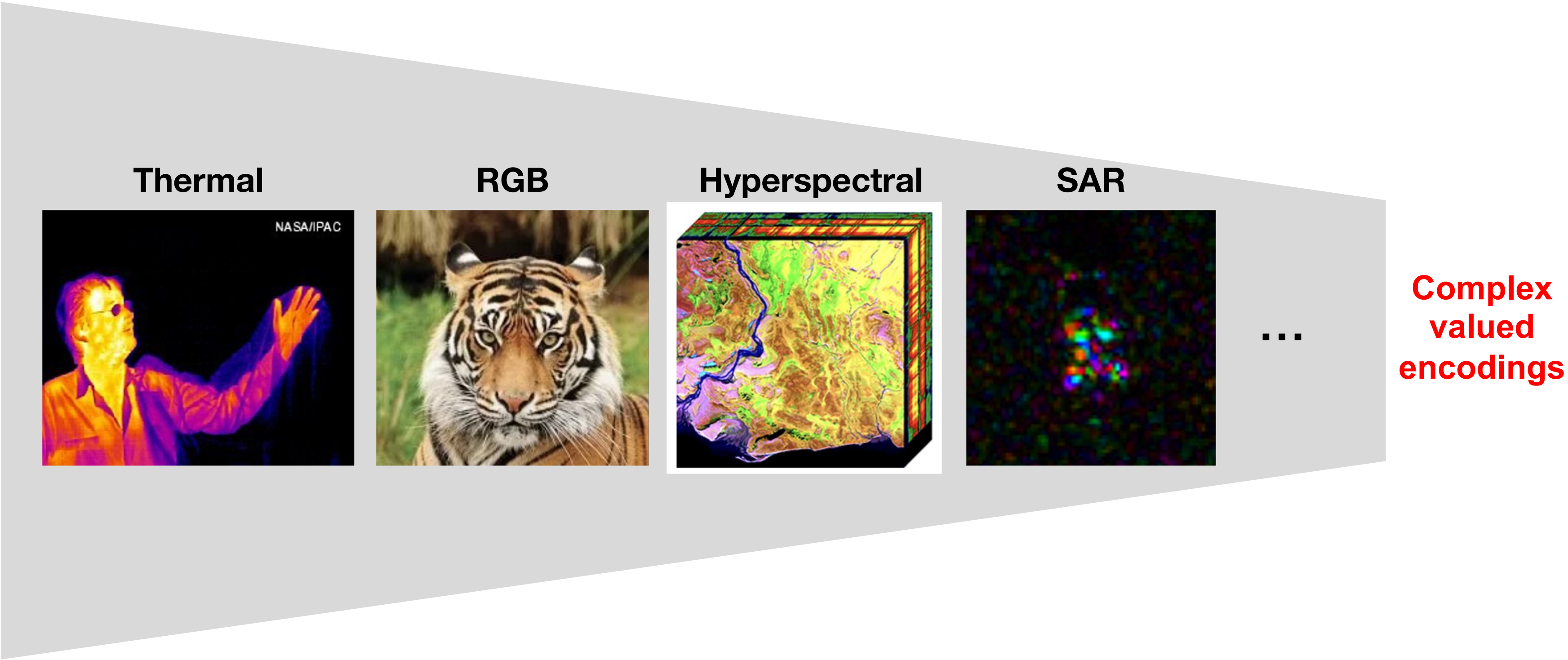


We Focus on Complex-Valued Data

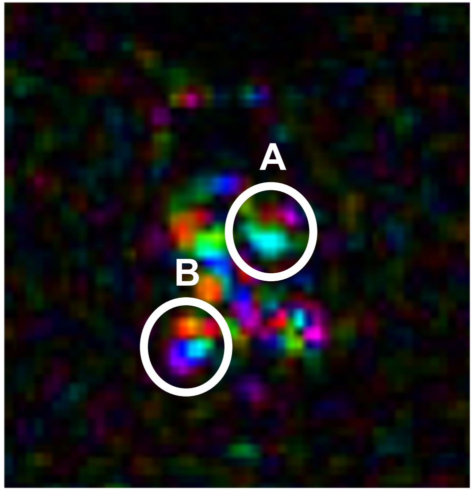
Thermal | RGB | Hyperspectral | SAR

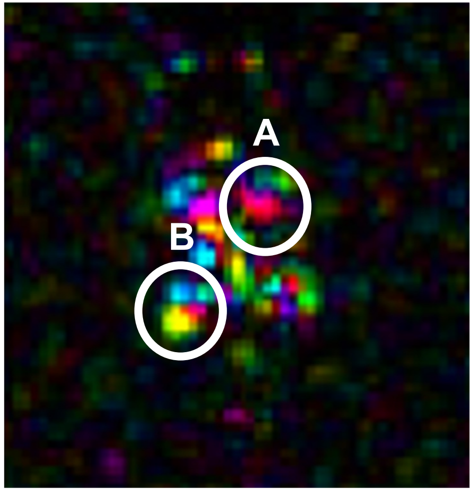


We Can Represent All These Data Types in Complex Values!

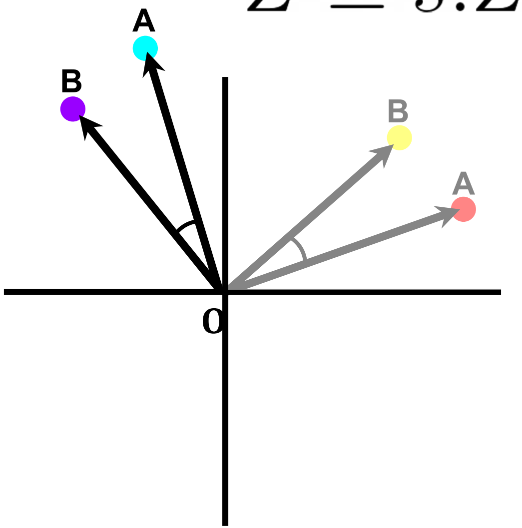


Property 1: Equivalence Under Complex-Valued Scaling



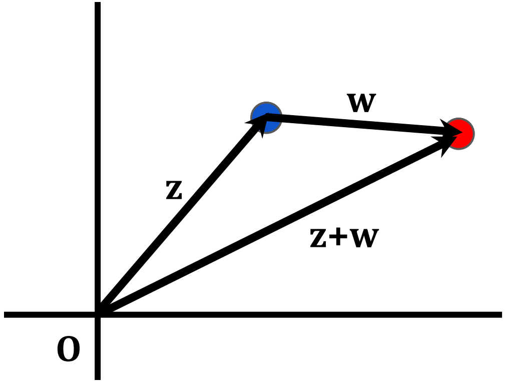


$$\vec{Z} \equiv s.\vec{Z}$$



Property 2: Rich Set of Algebraic Operations

addition



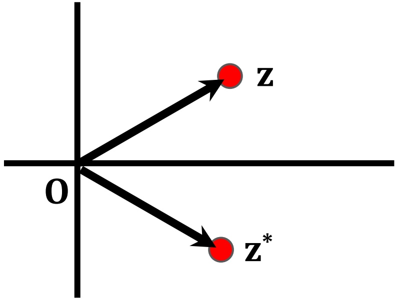
multiplication

$$(a+ib)(c+id) = (ac-bd) + i(ad+bc)$$

division

$$\frac{(a+bi)}{(c+di)} = \frac{ac+bd}{c^2+d^2} + \frac{bc-ac}{c^2+d^2}$$

conjugation

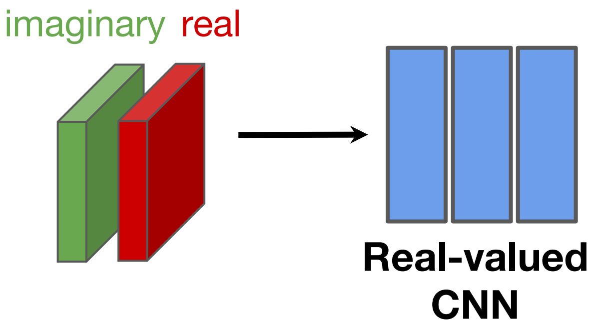


Methods For Handling Complex-Valued Data

Method	Complex-scaling?	Complex-valued algebra?

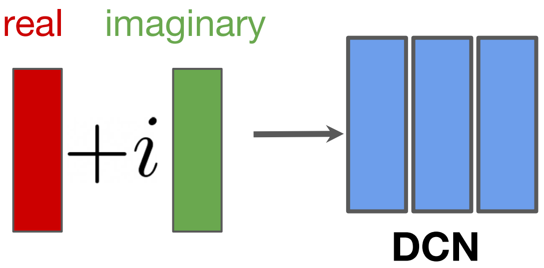
Methods For Handling Complex-Valued Data

Method	Complex-scaling?	Complex-valued algebra?
Real-valued CNN	✗	✗



Methods For Handling Complex-Valued Data

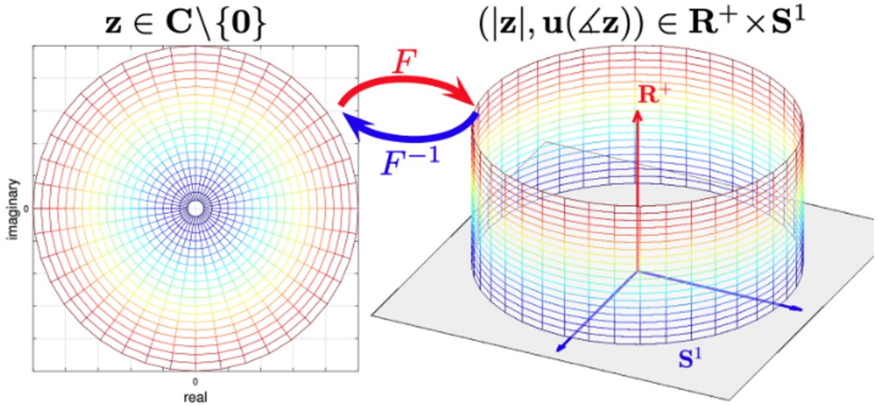
Method	Complex-scaling?	Complex-valued algebra?
Real-valued CNN	✗	✗
Deep Complex Nets	✗	✓



$$\mathbf{W} * \mathbf{h} = (\mathbf{A} * \mathbf{x} - \mathbf{B} * \mathbf{y}) + i(\mathbf{B} * \mathbf{x} + \mathbf{A} * \mathbf{y})$$

Methods For Handling Complex-Valued Data

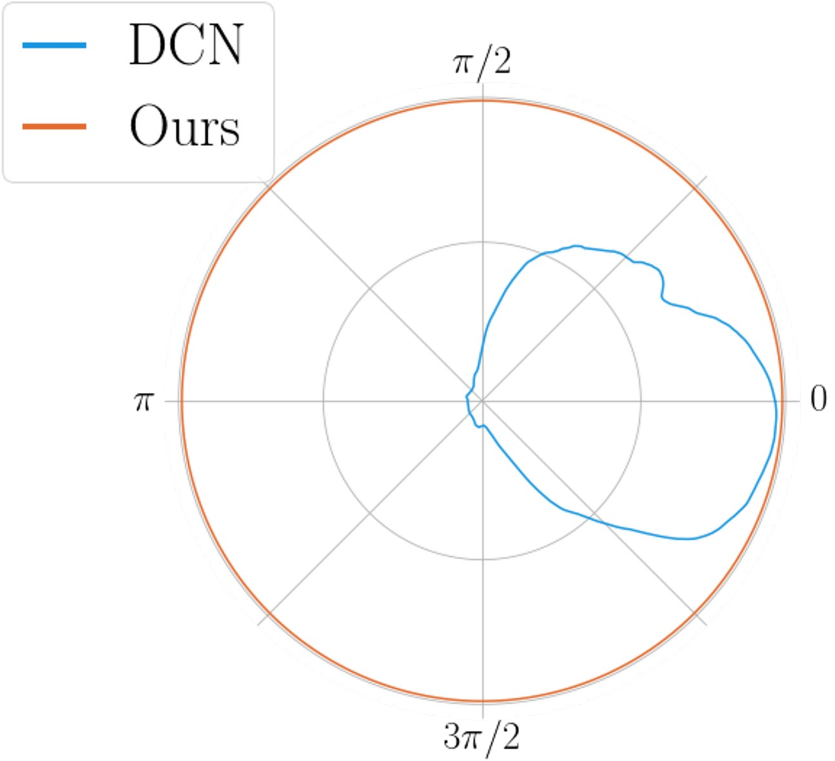
Method	Complex-scaling?	Complex-valued algebra?
Real-valued CNN	✗	✗
Deep Complex Nets	✗	✓
SurReal	✓	✗



Methods For Handling Complex-Valued Data

Method	Complex-scaling?	Complex-valued algebra?
Real-valued CNN	✗	✗
Deep Complex Nets	✗	✓
SurReal	✓	✗
Ours	✓	✓

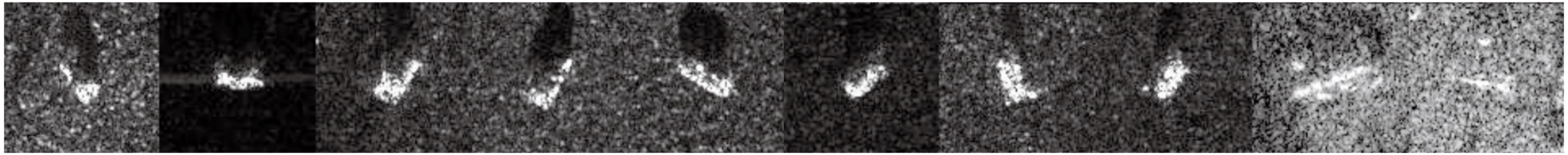
Benefits: Complex-Scaling Invariance



Our model makes predictions invariant to complex-valued scaling

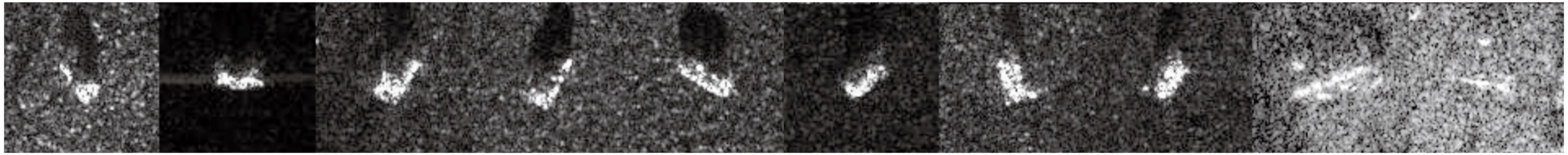
Benefits: Higher Accuracy with Leaner Models

MSTAR: Synthetic Aperture Radar Imaging



Benefits: Higher Accuracy with Leaner Models

MSTAR: Synthetic Aperture Radar Imaging

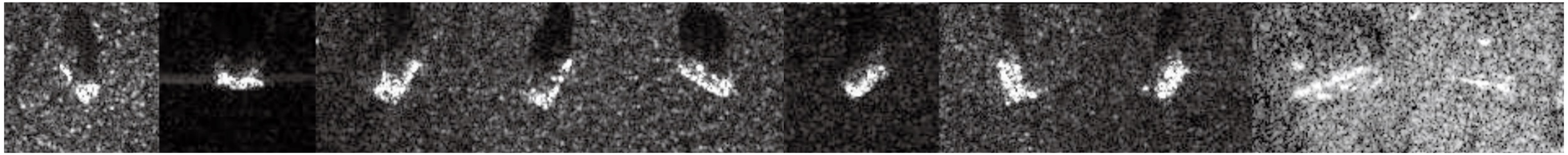


Model	# Params	Relative Params	Training dataset size (%)		
			5%	**10%**	**100%**
Complex CNN	863,587	1.00	49.8	47	89.1
Ours	29,536	**0.03**	**69.5**	**78.3**	**96.1**

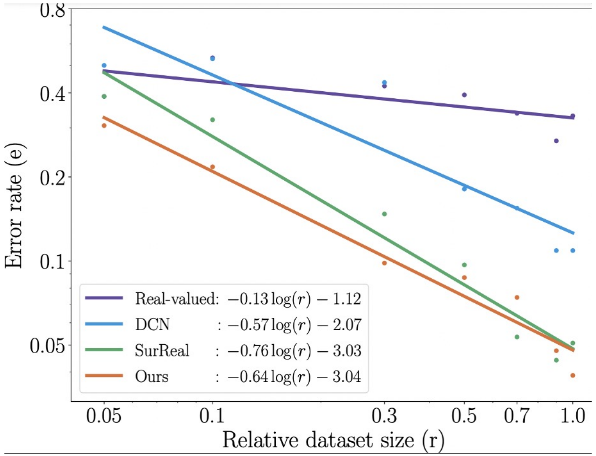
Higher accuracy with much leaner models

Benefits: Higher Accuracy with Leaner Models

MSTAR: Synthetic Aperture Radar Imaging



Model	# Params	Relative Params	Training dataset size (%)		
			5%	10%	100%
Complex CNN	863,587	1.00	49.8	47	89.1
Ours	29,536	**0.03**	**69.5**	**78.3**	**96.1**

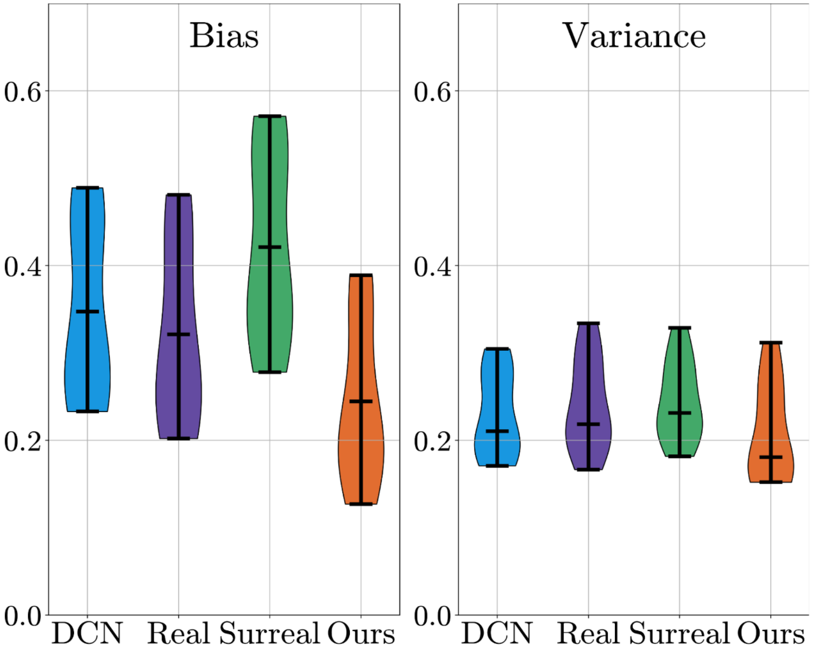


Higher accuracy with much leaner models

Consistently lower error rate

Benefits: Diverse Filters, Lower Bias/Variance

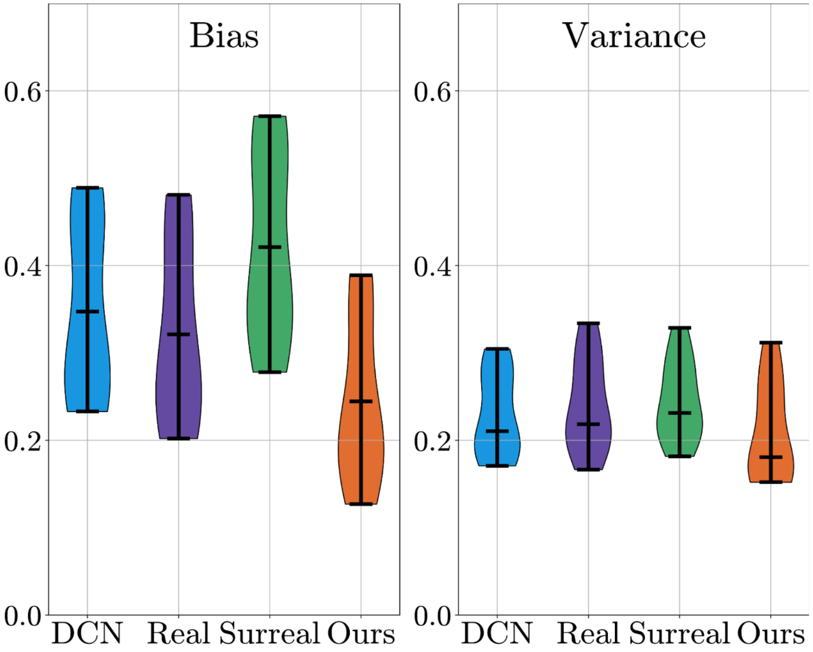
CIFAR 10



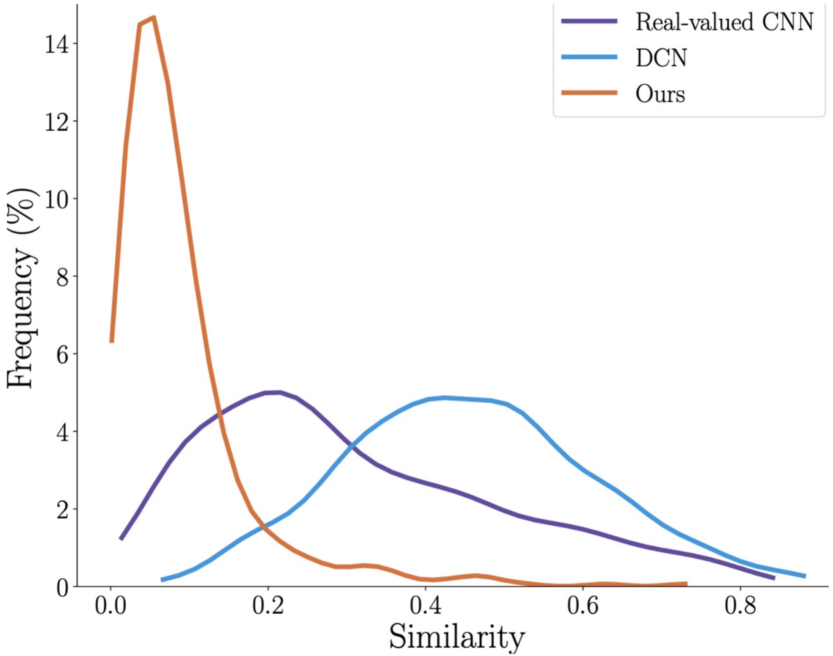
Lower Bias and Variance

Benefits: Diverse Filters, Lower Bias/Variance

CIFAR 10

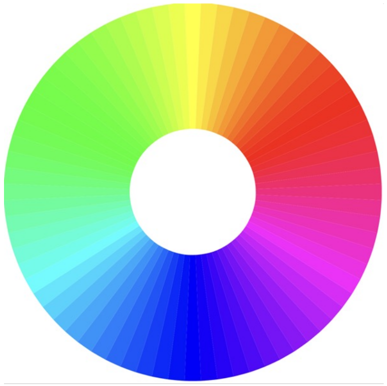


Lower Bias and Variance



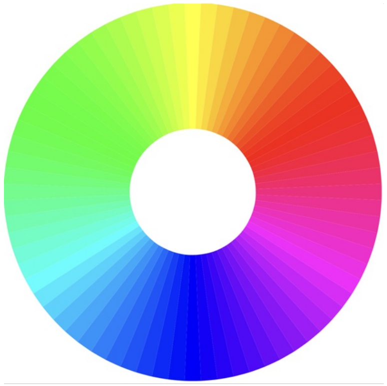
Less Redundant Filters

Benefits: Robustness Against Some Types of Color Distortion

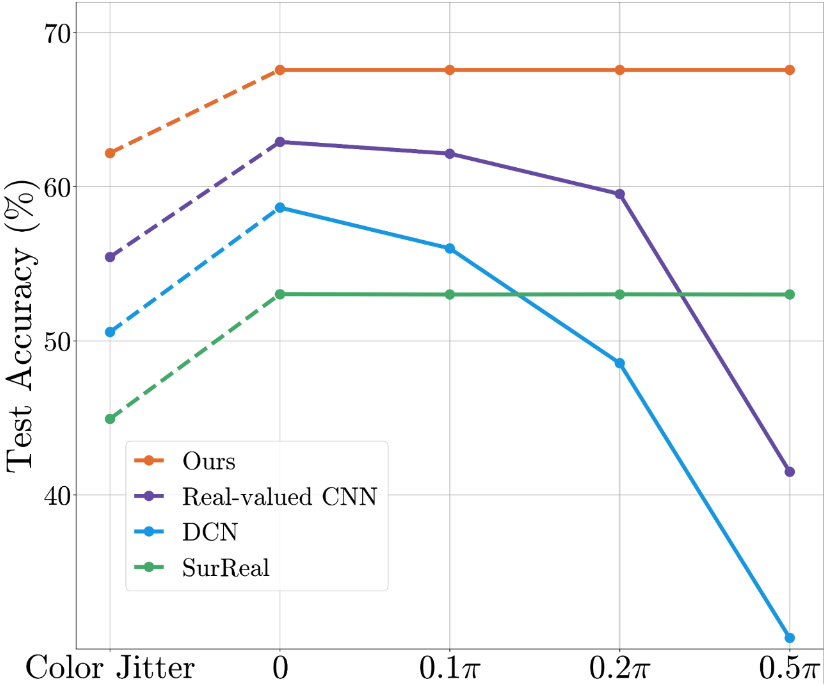


Encoding color with complex numbers

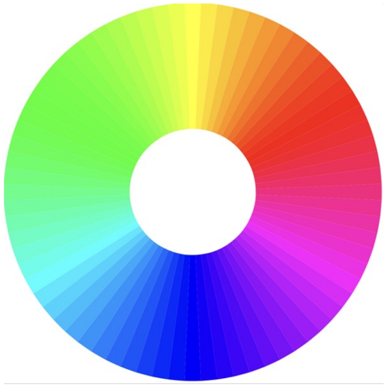
Benefits: Robustness Against Some Types of Color Distortion



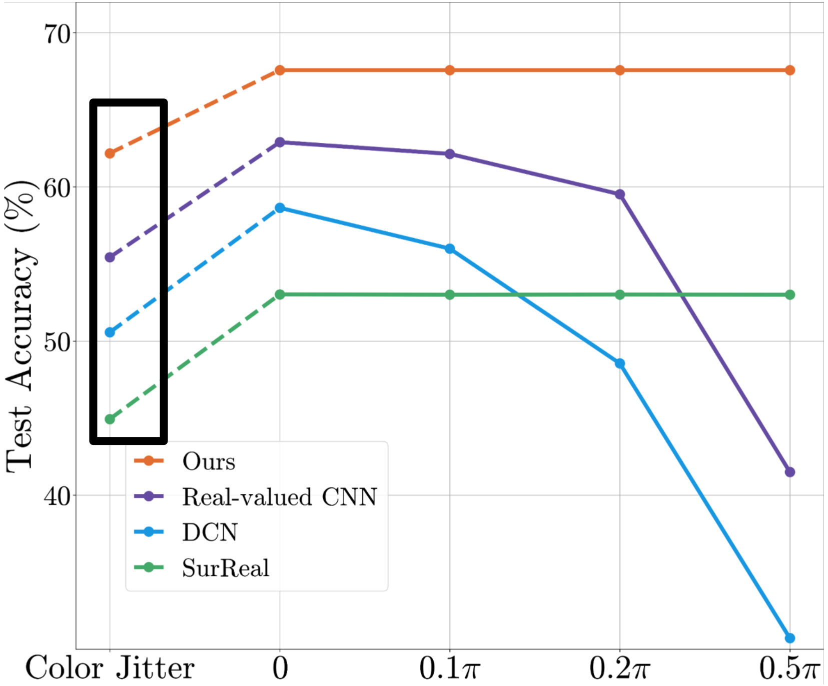
Encoding color with complex numbers



Benefits: Robustness Against Some Types of Color Distortion



Encoding color with complex numbers



Methods: Complex-Scale Equi-/In-variant Layers

Equivariant

Equivariant Convolution

Equivariant Batch-Norm

Equivariant Non-Linearity

Equivariant Pooling

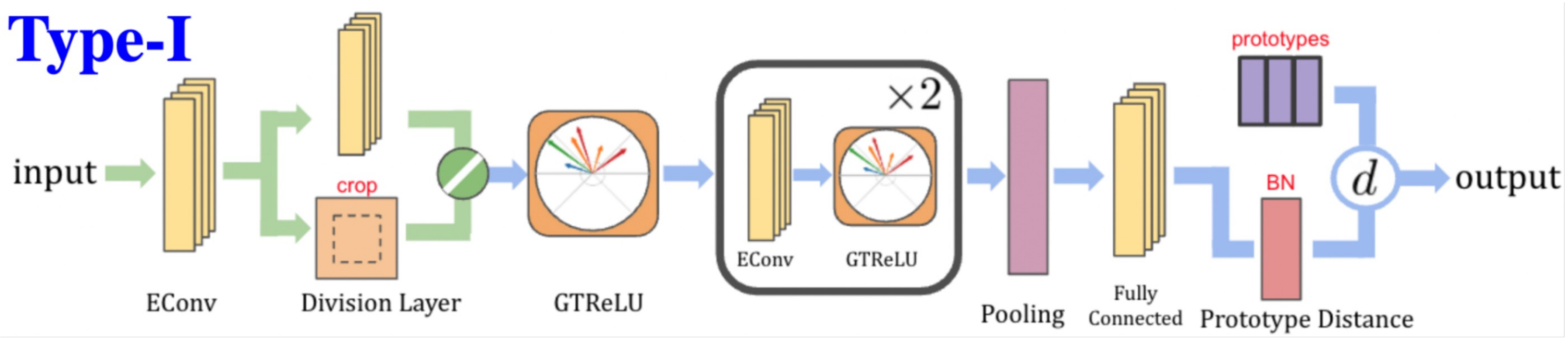
Invariant

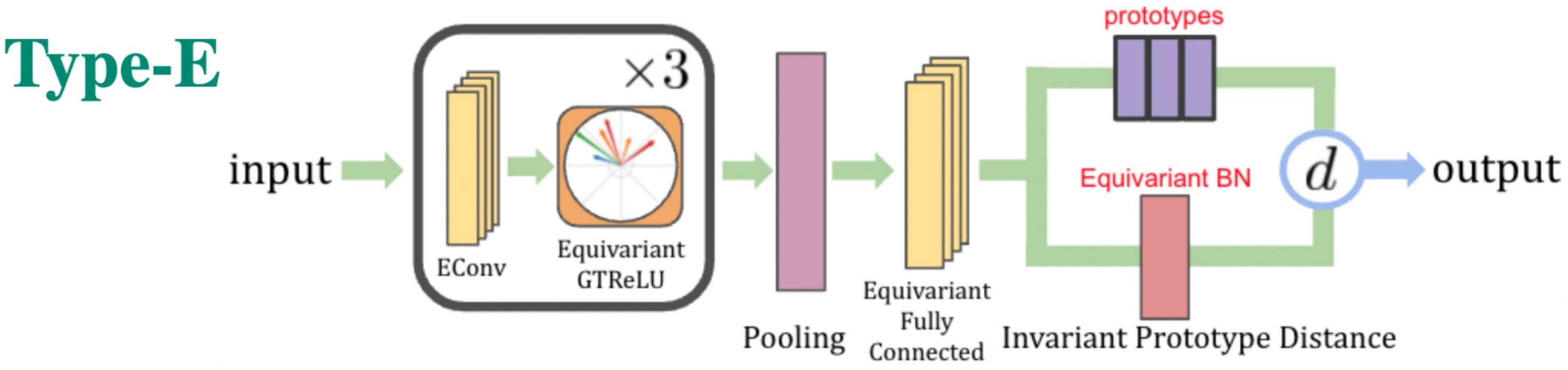
Conjugate Layer

Division Layer

Prototype-Distance Invariant Layer

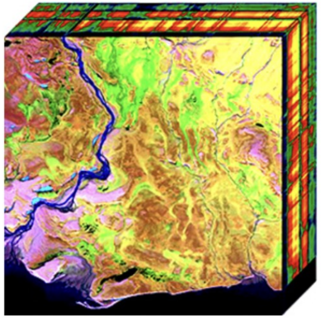
Methods: Two Architecture Styles





Methods: Our Proposed Complex-Valued Encodings

Hyperspectral



Sliding channel encoding

$$[x_1, x_2, x_3, x_4, \ldots]$$

$$\downarrow$$

$$[x_1 + ix_2, x_2 + ix_3, x_3 + ix_4, \ldots]$$

Color

LAB encoding

$$[L, a, b]$$

$$\downarrow$$

$$[L, a + ib]$$

Thank you!

Poster 68a, June 21st, 10AM-12:30PM

github.com/sutkarsh/cds

Co-domain Symmetry for Complex-Valued Deep Learning

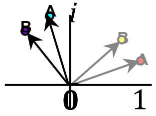
Utkarsh Singhal Yifei Xing Stella X. Yu

Contributions

1. New complex-valued learning method based on co-domain symmetry with respect to complex-valued scaling
2. New leaner classifiers with higher accuracy, better generalization, more robustness, lower model bias/variance
3. New complex-valued encodings of various types of images
4. Achieve color jitter robustness without any augmentation

Complex-Valued Data Properties

complex scaling



$\vec{Z} \equiv s.\vec{Z}$

complex algebra

$(a+bi)^* = a - bi$

$(a+bi) \times (c+di)$

$\vdots$

$(a+bi) + (c+di)$

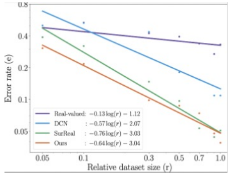
	Complex Algebra	ℂ-scale Invariance
Real	✗	✗
DCN	✓	✗
Surreal	✗	✓
Ours	✓	✓

Leaner, Better, More Robust Models

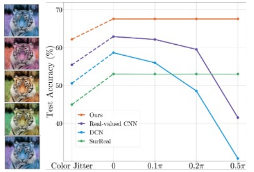
MSTAR	Params	5%	10%	50%	100%
Real	33k	47.4	46.6	60.6	66.9
DCN	863k	49.8	47.0	81.9	89.1
SurReal	63k	61.1	68.0	90.3	94.9
Ours	29k	69.5	78.3	91.3	96.1

xView	Params	Ratio	Acc (%)
Real	36k	1.00	59.9
DCN	69k	1.89	56.7
SurReal	36k	0.99	58.7
Ours	27k	0.75	67.7

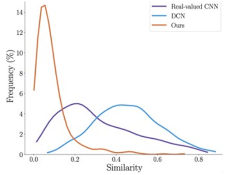
higher accuracy with fewer model parameters



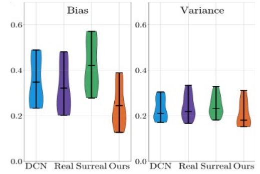
better generalization



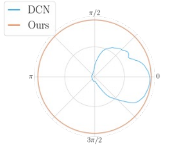
color robustness



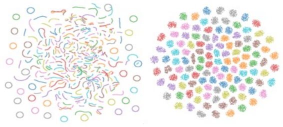
less redundant filters



lower bias/variance



invariance

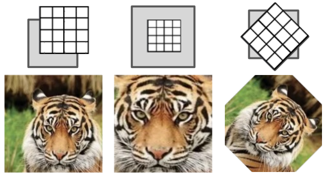


invariant representation

Co-Domain Image Transformations

Domain: Pixel locations

Co-domain: Pixel values



complex-valued encodings

Layer Functions for Co-Domain Symmetry

Equivariant

$$f(s \cdot \mathbf{z}) = s \cdot f(\mathbf{z})$$

$$\mathbf{f}_{BN} = BN(|\mathbf{f}|) \odot \frac{\mathbf{f}}{|\mathbf{f}|}$$

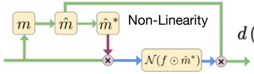
$$\text{Econv}(\mathbf{z};\mathbf{W}) = \mathbf{W} * \mathbf{z} = (\mathbf{X} + i\mathbf{Y}) * (\mathbf{a} + i\mathbf{b})$$

Invariant

$$f(s \cdot \mathbf{z}) = f(\mathbf{z})$$

$$\text{Div}(\mathbf{z_1},\mathbf{z_2}) = \frac{|\mathbf{z_1}|}{|\mathbf{z_2}| + \epsilon} \exp\{i(\measuredangle \mathbf{z_1} - \measuredangle \mathbf{z_2})\}$$

$$\text{Conj}(\mathbf{z_1},\mathbf{z_2}) = \mathbf{z_1}\mathbf{z_2^*}$$



$$d(z_1, z_2) = \sqrt{(\ln|z_1| - \ln|z_2|)^2 + \text{arc}(\measuredangle z_1, \measuredangle z_2)^2}$$

Model Architectures: Early or Late Invariance