Multi-Spectral Image Classification with Ultra-Lean Complex-Valued Models

Utkarsh Singhal   Stella X. Yu   Zackery Steck   Scott Kangas   Aaron A. Reite

AI + HADR 2022

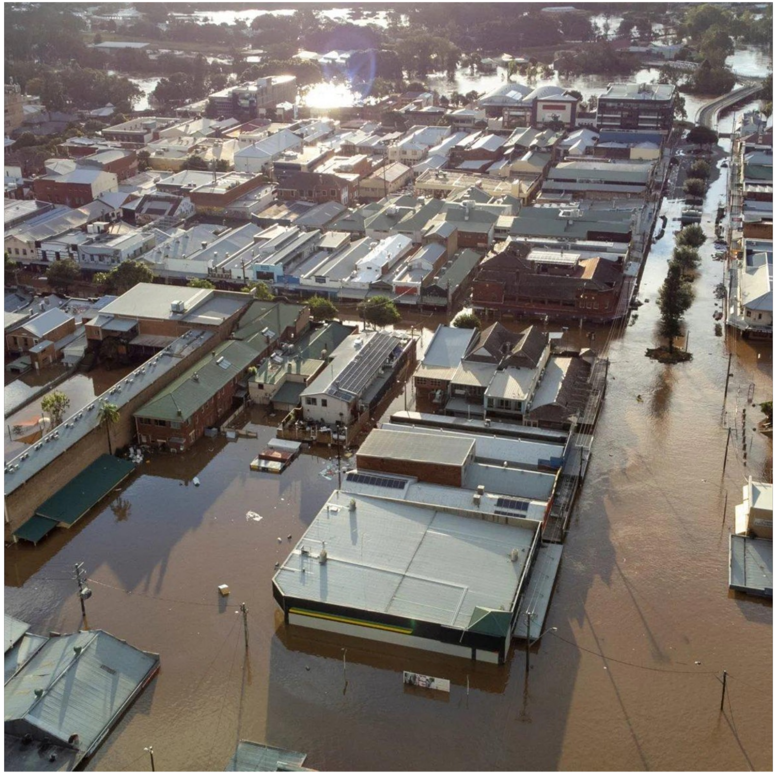
NEURAL INFORMATION PROCESSING SYSTEMS

New South Wales Floods, November 2022

NSW floods become most expensive natural disaster on record with $5.5b in claims

By Melinda Hayter, Holly Tregenza, and Indiana Hansen

Posted Mon 21 Nov 2022 at 2:07pm, updated Tue 22 Nov 2022 at 1:15am



Satellite Image (RGB)

Source: planet.com

Satellite Image (RGB + Infrared, false color)

water *absorbs* IR, so appears *dark blue*
vegetation *reflects* IR, so appears *red*

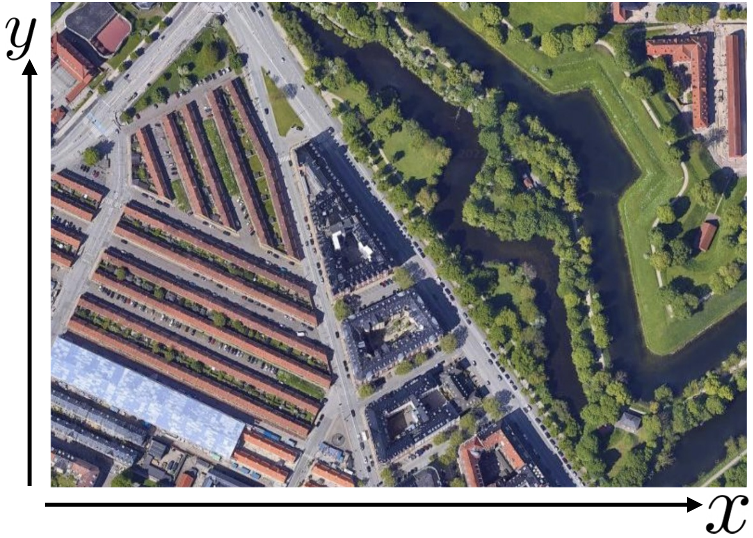
Source: planet.com

Satellite Image (RGB + Infrared, false color)

Extra EM bands (e.g. infrared) can reveal changes invisible in RGB

Source: planet.com

Multi-Band Imaging

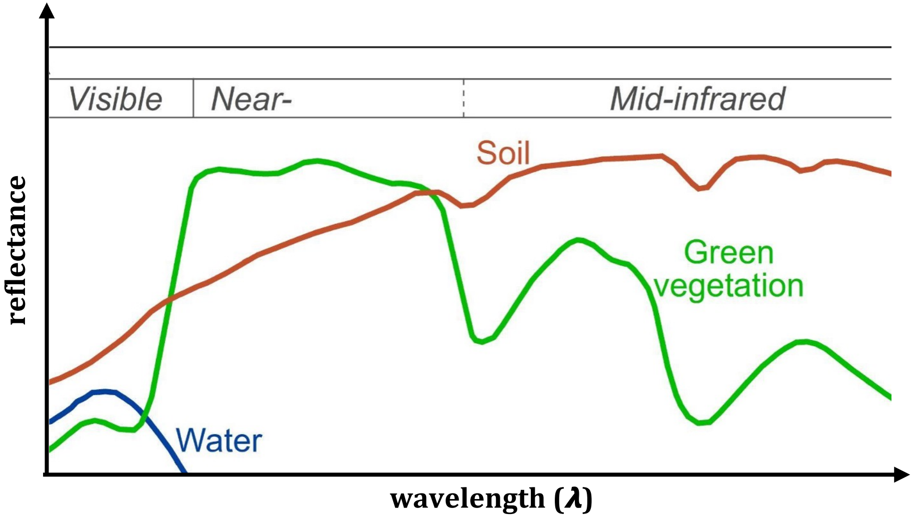


Source: https://seos-project.eu/classification/classification-c01-p05.html

Multi-Band Imaging

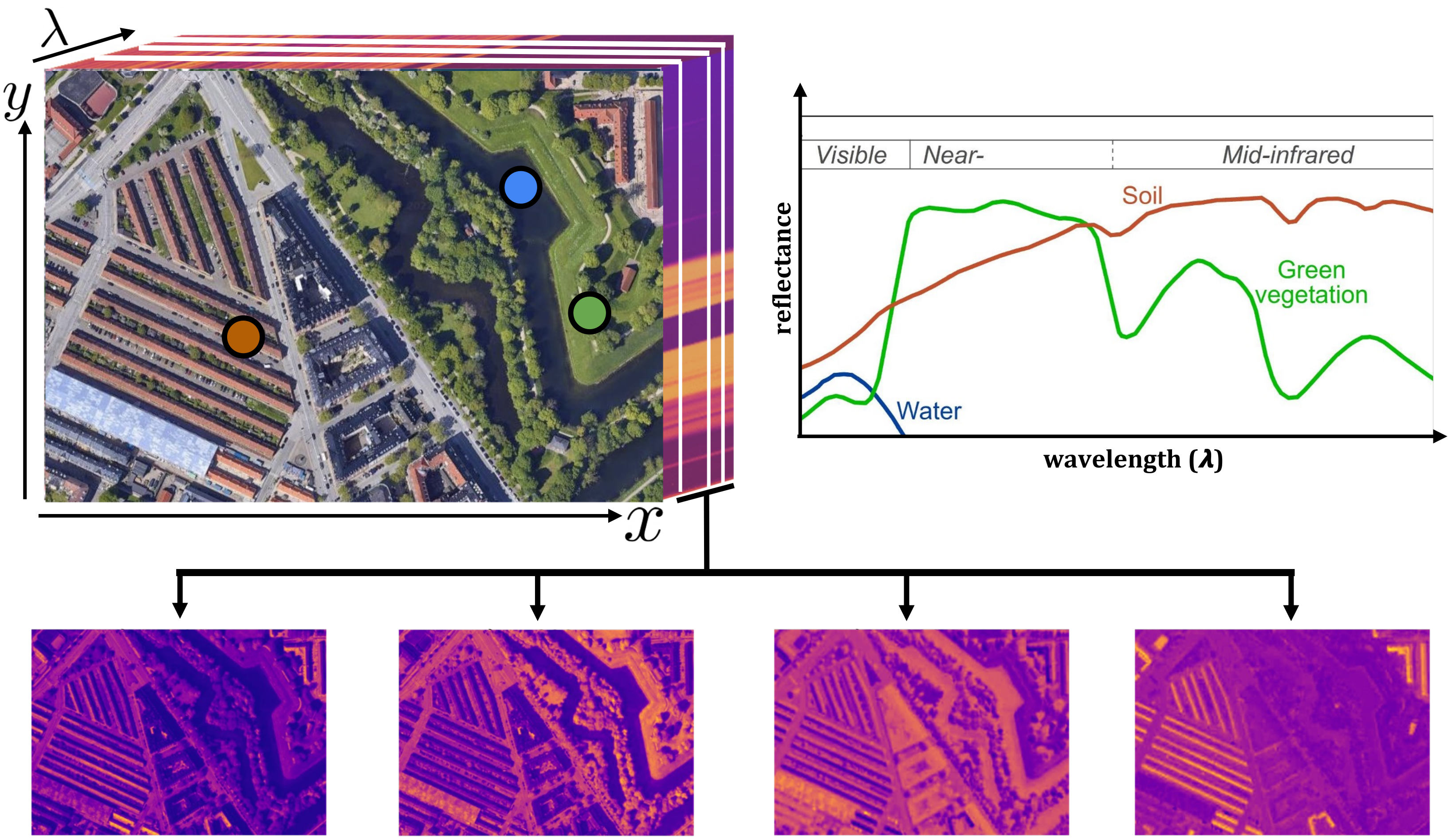
Source: https://seos-project.eu/classification/classification-c01-p05.html

Multi-Band Imaging



Source: https://seos-project.eu/classification/classification-c01-p05.html

Multi-Band Imaging



Multi-Band Imaging for HADR

Disaster Assessment

Multi-Band Imaging for HADR

Disaster Assessment

Environmental Impact Monitoring

Multi-Band Imaging for HADR

Disaster Assessment

Environmental Impact Monitoring

Agricultural Health Measurement

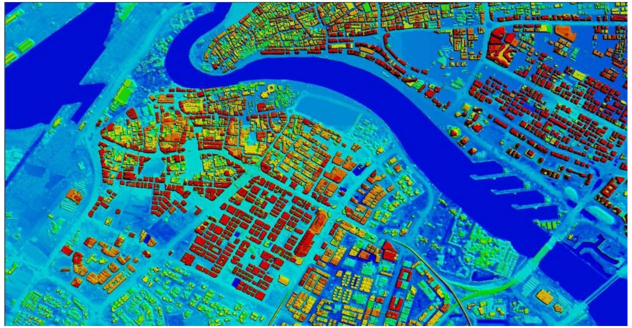
Multi-Band Imaging for HADR

Disaster Assessment

Environmental Impact Monitoring

Agricultural Health Measurement

Urban Planning



Common Strategy for Dealing with New Datasets

- **Large dataset → supervised learning from scratch**

Common Strategy for Dealing with New Datasets

- **Large dataset → supervised learning from scratch**

Extremely successful for datasets like ImageNet

Common Strategy for Dealing with New Datasets

- **Large dataset → supervised learning from scratch**

  Extremely successful for datasets like ImageNet

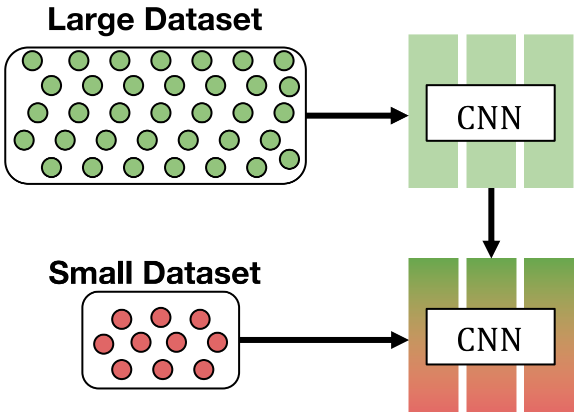
- **Small dataset → transfer learning**

Common Strategy for Dealing with New Datasets

- **Large dataset → supervised learning from scratch**

  Extremely successful for datasets like ImageNet

- **Small dataset → transfer learning**
  - ❑ Neural Net pre-trained on, e.g, ImageNet
  - ❑ Fine-tune on the smaller dataset
  - ❑ Extensive data augmentations



How to Handle a Multi-band Dataset?

- **Supervised training from scratch?**

How to Handle a Multi-band Dataset?

- ~~**Supervised training from scratch?**~~
  Relatively limited labels

How to Handle a Multi-band Dataset?

- ~~**Supervised training from scratch?**~~
  Relatively limited labels

- **Transfer learning from a large RGB dataset?**

How to Handle a Multi-band Dataset?

- ~~**Supervised training from scratch?**~~
  Relatively limited labels

- ~~**Transfer learning from a large RGB dataset?**~~
  Not 3 channel, encourages reduction to RGB

How to Handle a Multi-band Dataset?

- ~~**Supervised training from scratch?**~~
  Relatively limited labels
- ~~**Transfer learning from a large RGB dataset?**~~
  Not 3 channel, encourages reduction to RGB
- **Convert back to RGB?**

How to Handle a Multi-band Dataset?

- **~~Supervised training from scratch?~~**
  Relatively limited labels
- **~~Transfer learning from a large RGB dataset?~~**
  Not 3 channel, encourages reduction to RGB
- **~~Convert back to RGB?~~**
  Loses the original benefits of multi-band data

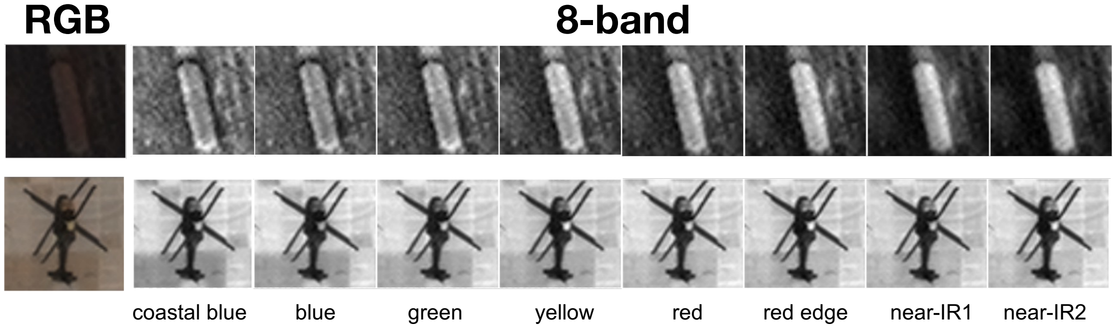
How to Handle a Multi-band Dataset?

- ~~Supervised training from scratch?~~
  Relatively limited labels

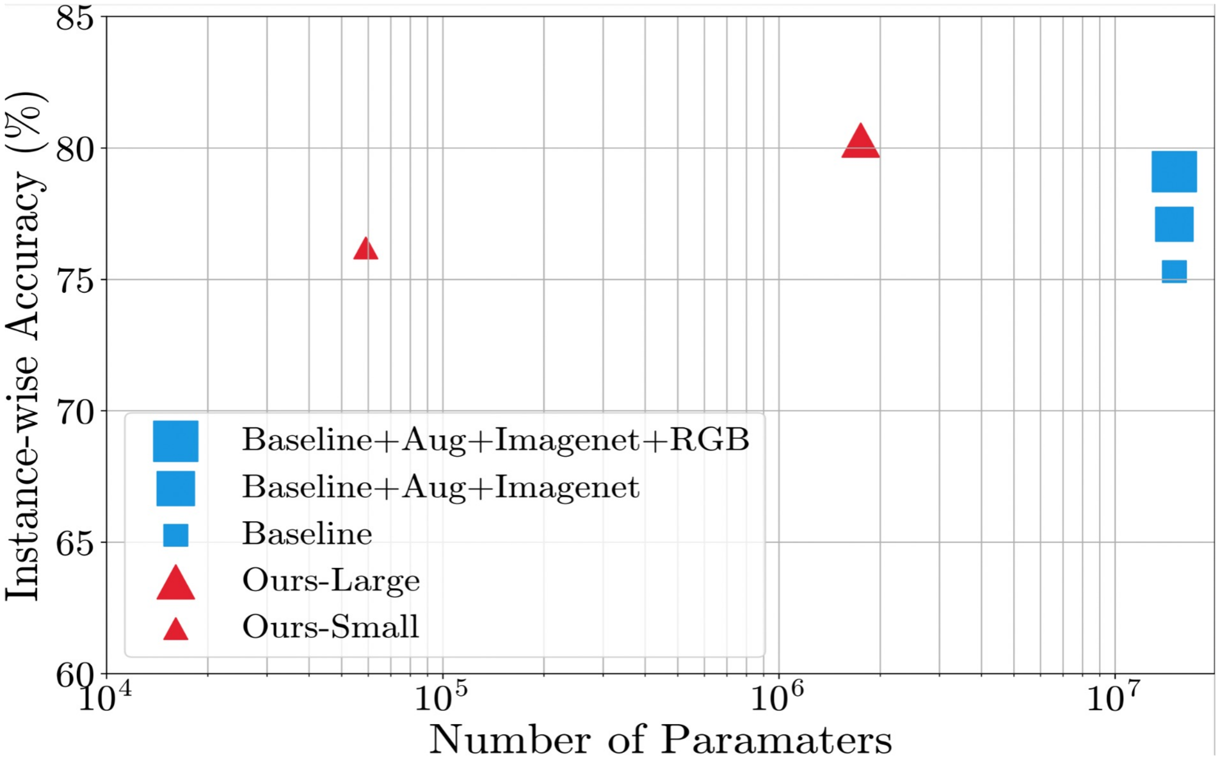
This Work: Complex-valued Deep Learning as an alternative

- ~~Convert back to RGB?~~
  Loses the original benefits of multi-band data

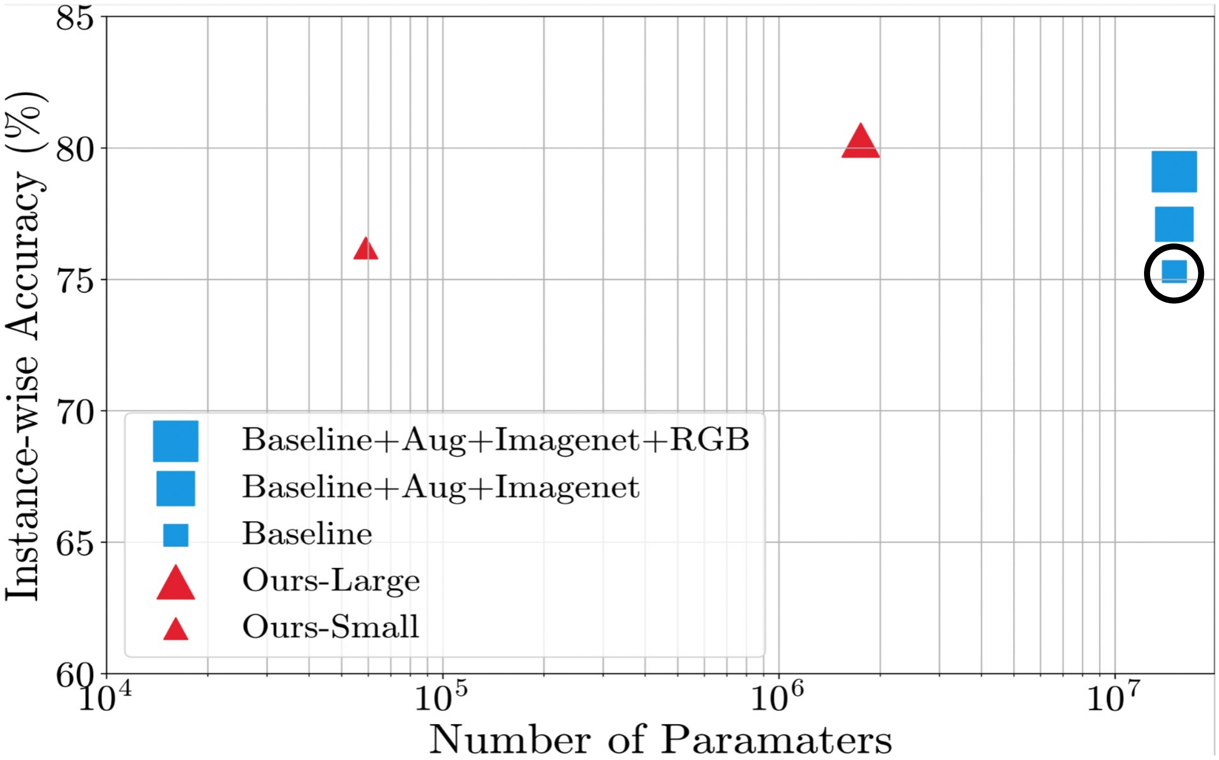
xView Multi-Band Image Dataset



Results: Simpler and Better Ultra-lean Models

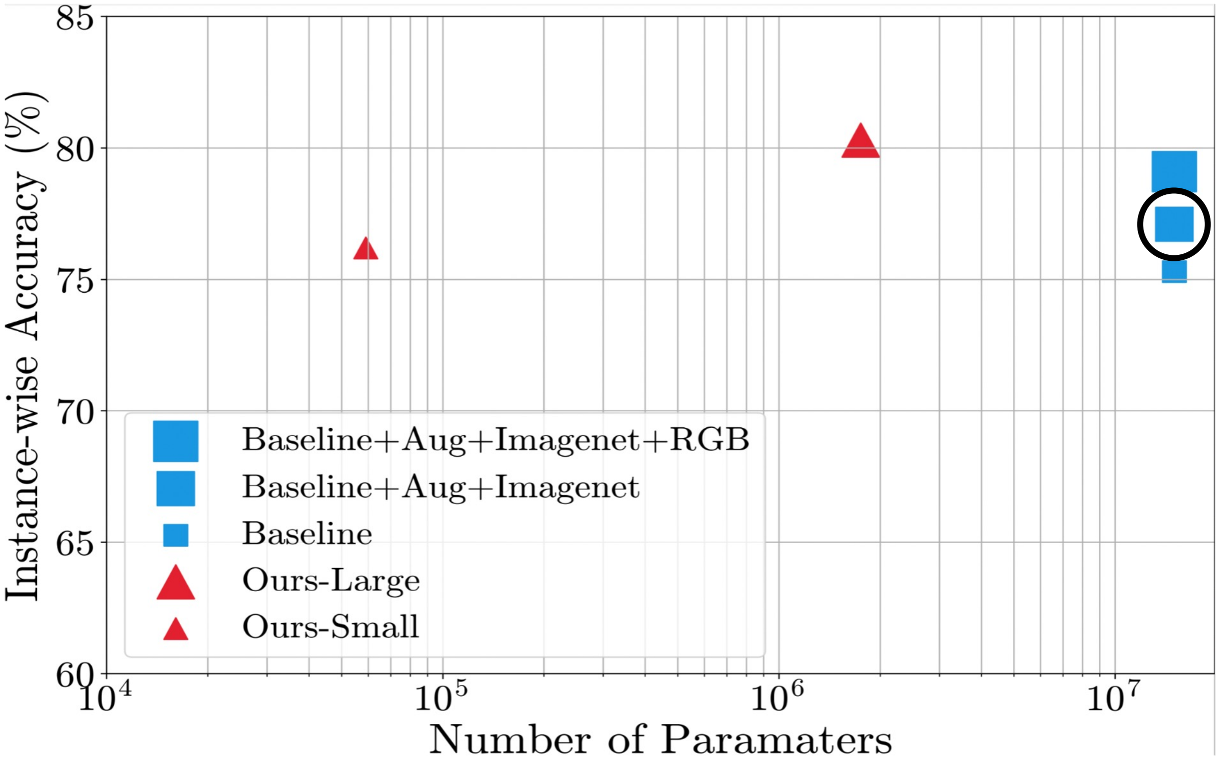


Results: Simpler and Better Ultra-lean Models



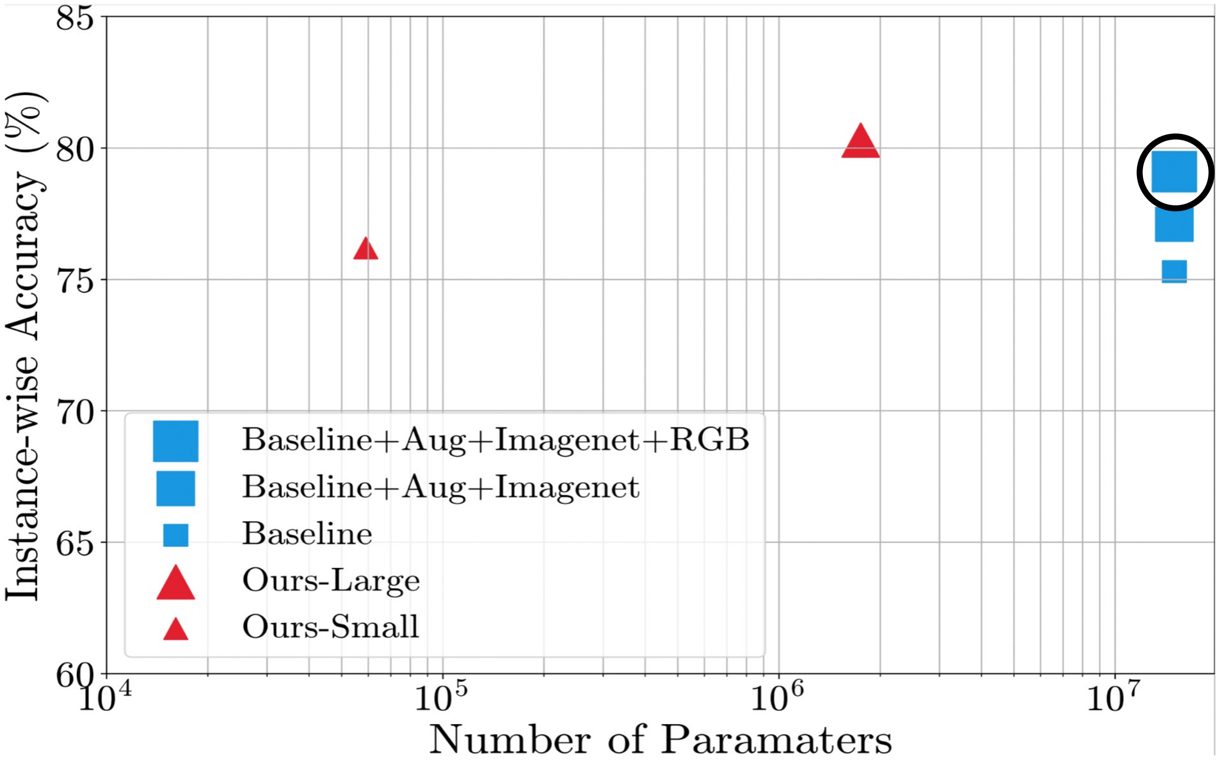
Baseline: ResNet18 trained from scratch

Results: Simpler and Better Ultra-lean Models



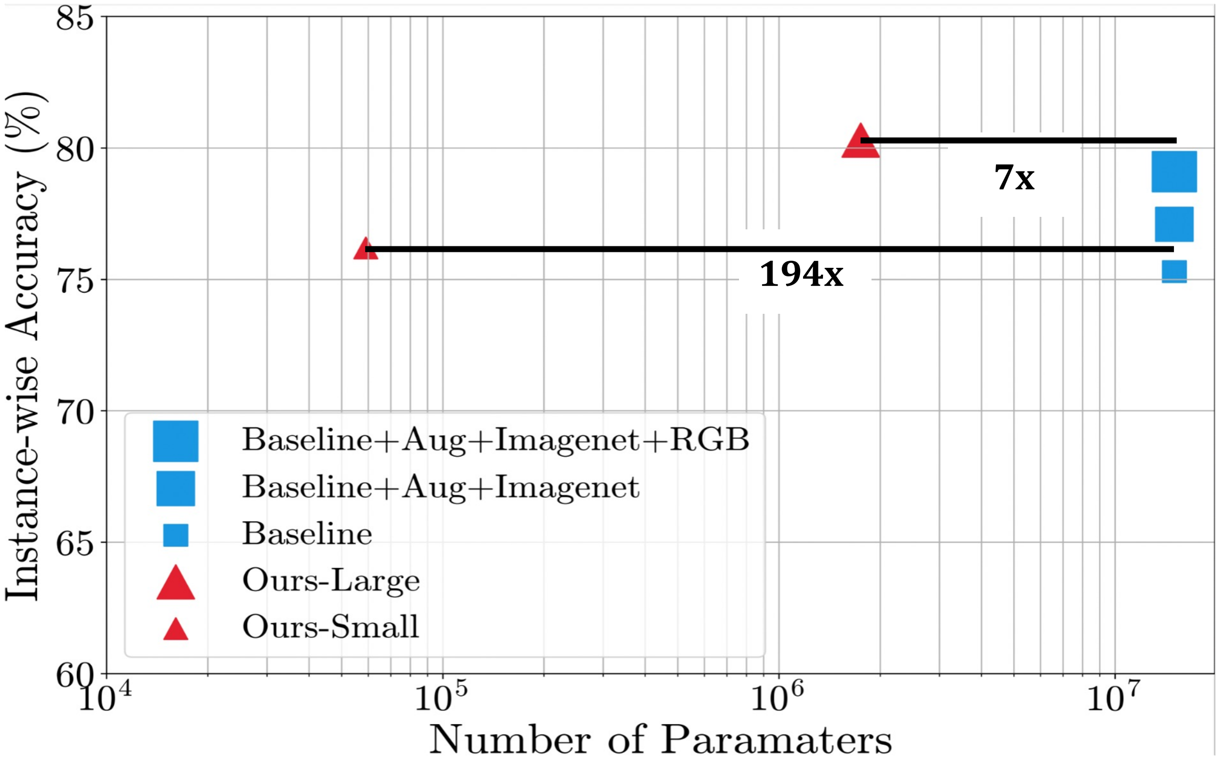
Baseline: ResNet18 with ImageNet pre-training and data augmentation

Results: Simpler and Better Ultra-lean Models



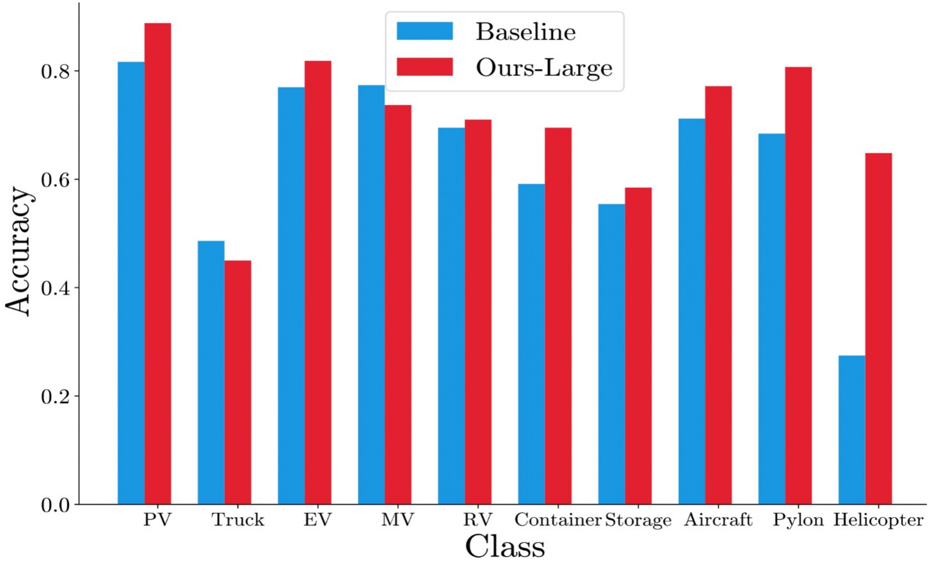
Baseline: Reduce down to RGB + ImageNet pre-training + data augmentation

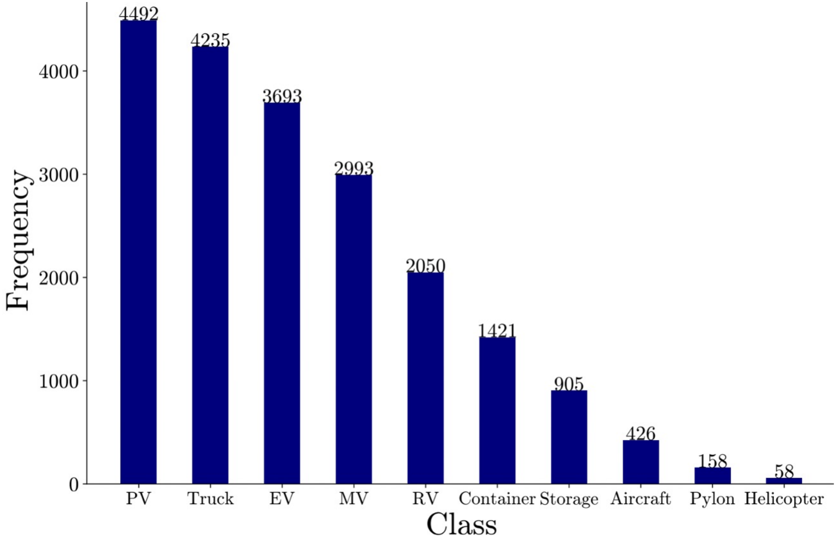
Results: Simpler and Better Ultra-lean Models



Higher accuracy, 194x smaller, no augmentation/pre-training, no RGB conversion

Imbalanced Classification Results



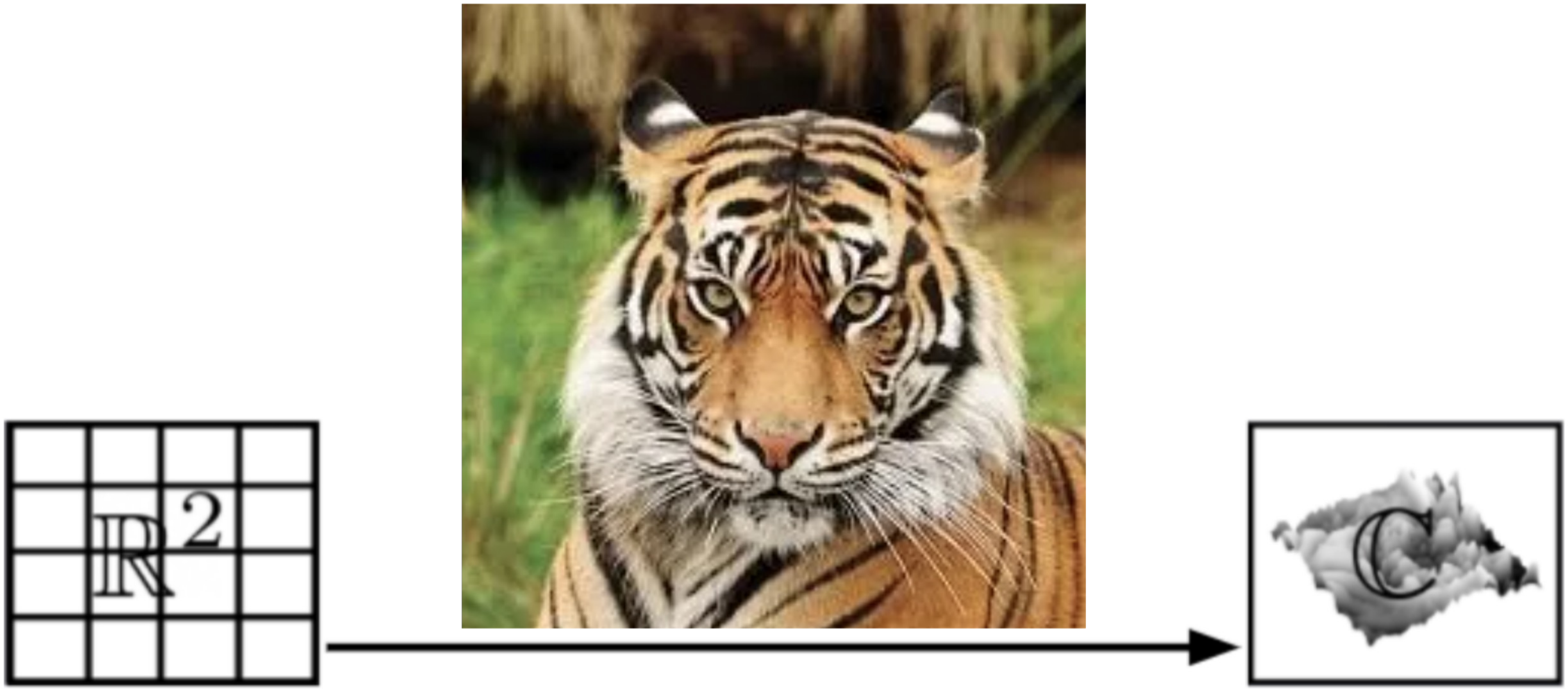


Higher Accuracy for 8 out of 10 classes

Methods: Co-domain Symmetric Models (CDS)[1]

[1]: *Co-domain Symmetry for Complex-Valued Deep Learning*, U. Singhal, Y. Xing, S.X. Yu, CVPR 2022

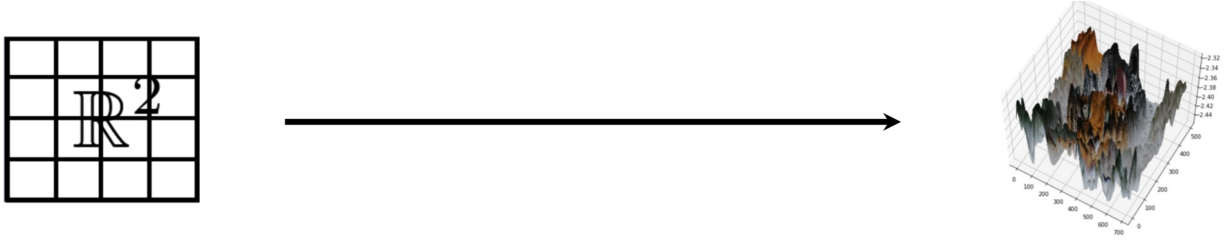
An Image is a Function from Domain to Co-Domain



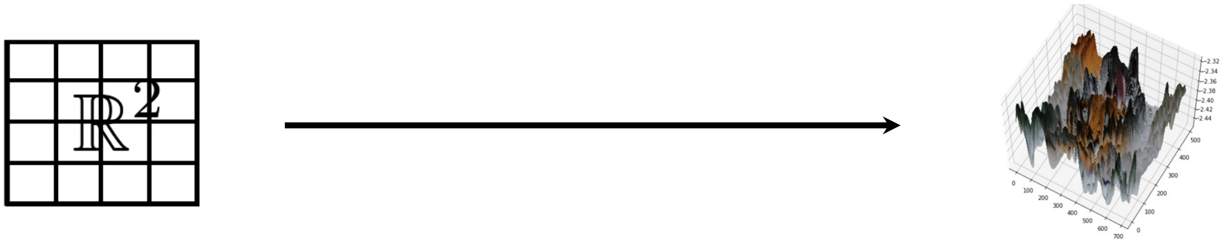
Domain: Pixel Locations

Co-Domain: Pixel Values

An Image is a Function from Domain to Co-Domain

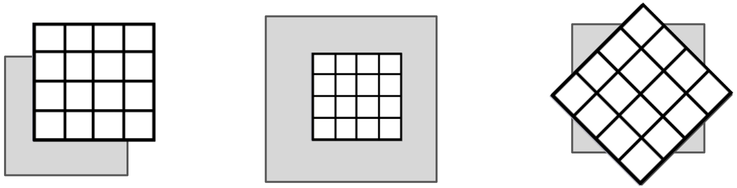


Domain Transformations Act on the Pixel *Coordinates*

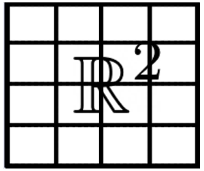


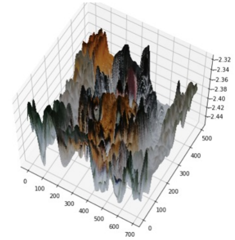
Domain Transformation

translation **scaling** **rotation**



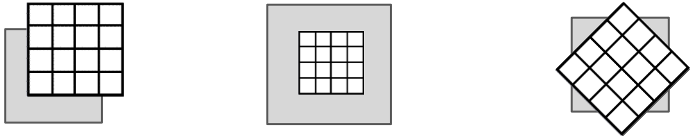
Domain Transformations Act on the Pixel *Coordinates*





Domain Transformation

translation | **scaling** | **rotation**



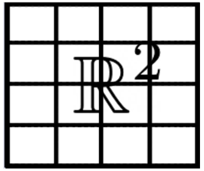
CNN [1]

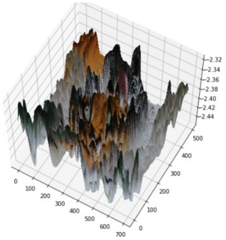
Scale-Invariant CNN [2]

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*

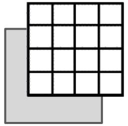




Domain Transformation

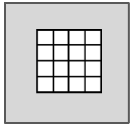
Co-domain Transformation

translation



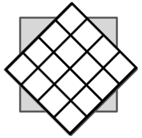
CNN [1]

scaling



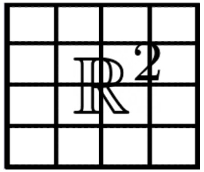
Scale-Invariant CNN [2]

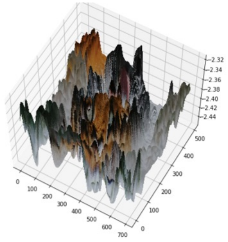
rotation



E(2)-Steerable CNN [3]

Co-Domain Transformations Act on the Pixel *Values*

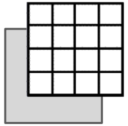




Domain Transformation

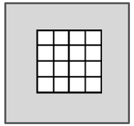
Co-domain Transformation

translation



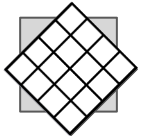
CNN [1]

scaling

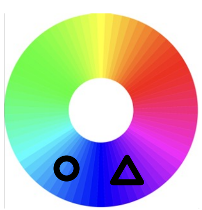


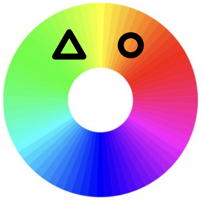
Scale-Invariant CNN [2]

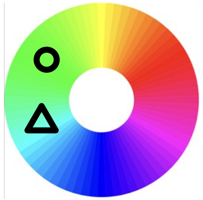
rotation

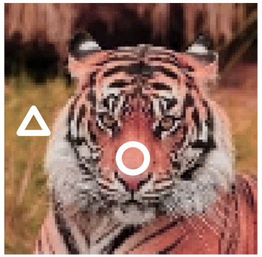


E(2)-Steerable CNN [3]





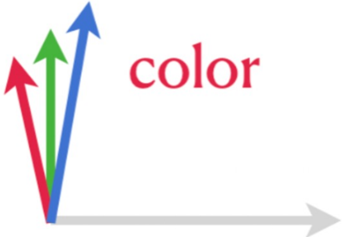




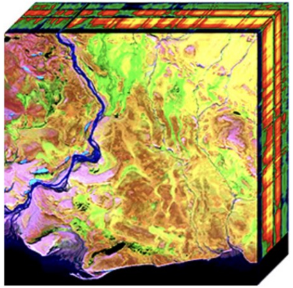
Co-Domain Encapsulates Diversity of Image Types

Thermal

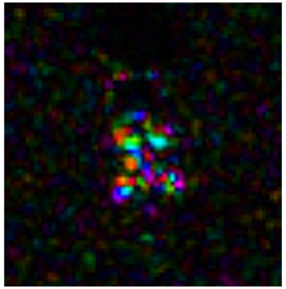
RGB



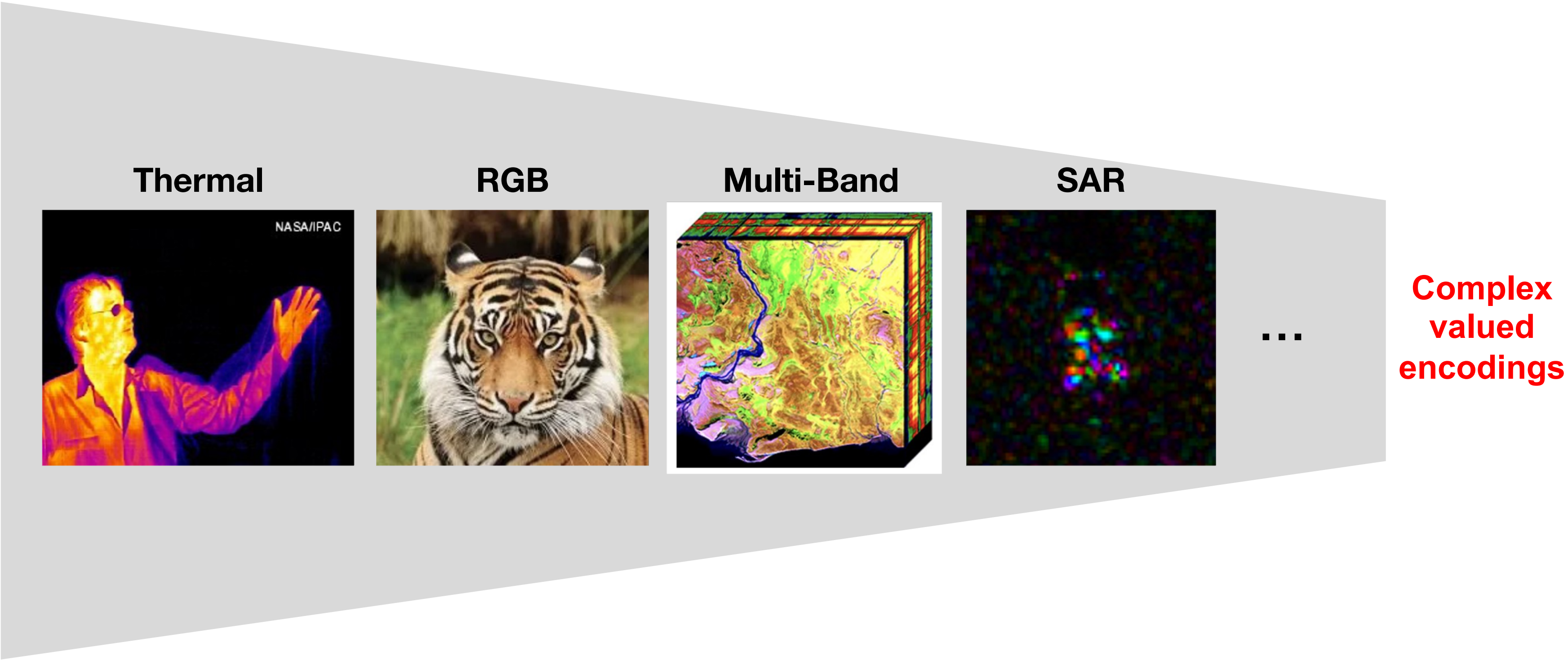
Multi-Band



SAR

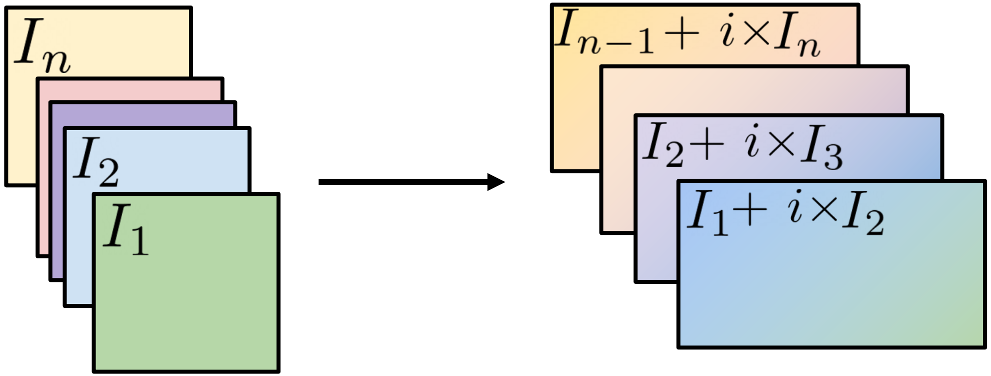


We Can Represent All These Data Types in Complex Values!



Complex-Valued Encoding for MSI Data

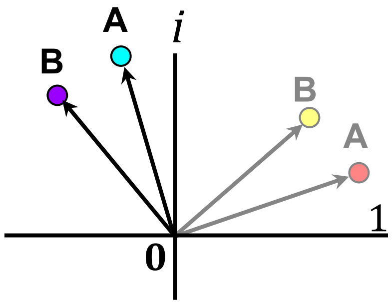
$$I = [I_1, I_2, \ldots, I_m] \rightarrow [I_1 + iI_2,\ \ I_2 + iI_3,\ \ \ldots,\ \ I_{m-1} + iI_m]$$



- Adjacent channels are paired into the real/imaginary parts of a complex number.
- Ratio of adjacent channels is represented by the phase.
- Imparts an ordering to the input channels

Robustness to Co-Domain Transformations

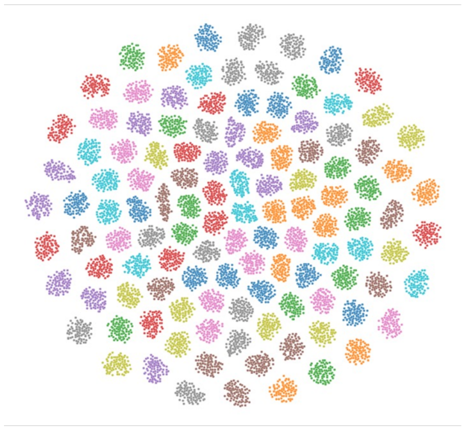
complex scaling



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

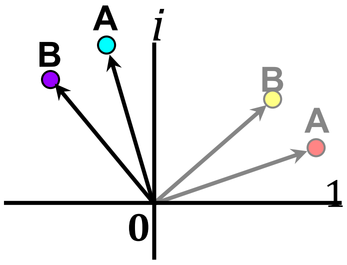
non-invariant

invariant



Robustness to Co-Domain Transformations

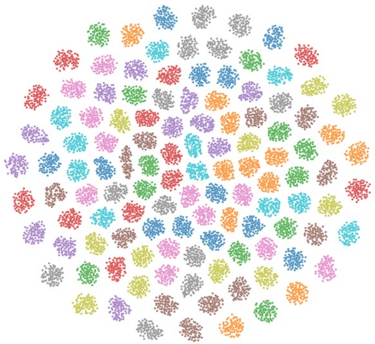
complex scaling



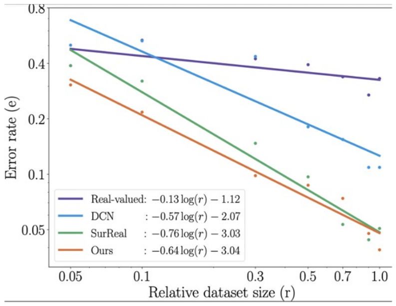
$$\vec{Z} \equiv s.\vec{Z}$$

non-invariant

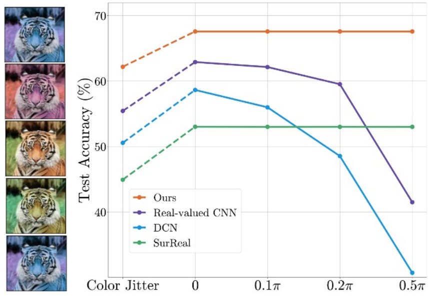
invariant



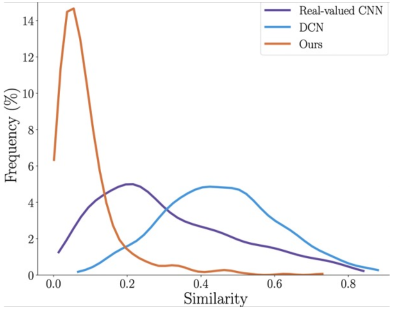
Previously on CIFAR 10:



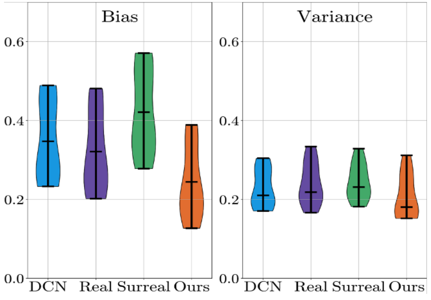
better generalization



color robustness



less redundant filters



lower bias/variance

Complex-Scale Equi-/In-variant Layers

Equivariant

Equivariant Convolution

Equivariant Batch-Norm

Equivariant Non-Linearity

Equivariant Pooling

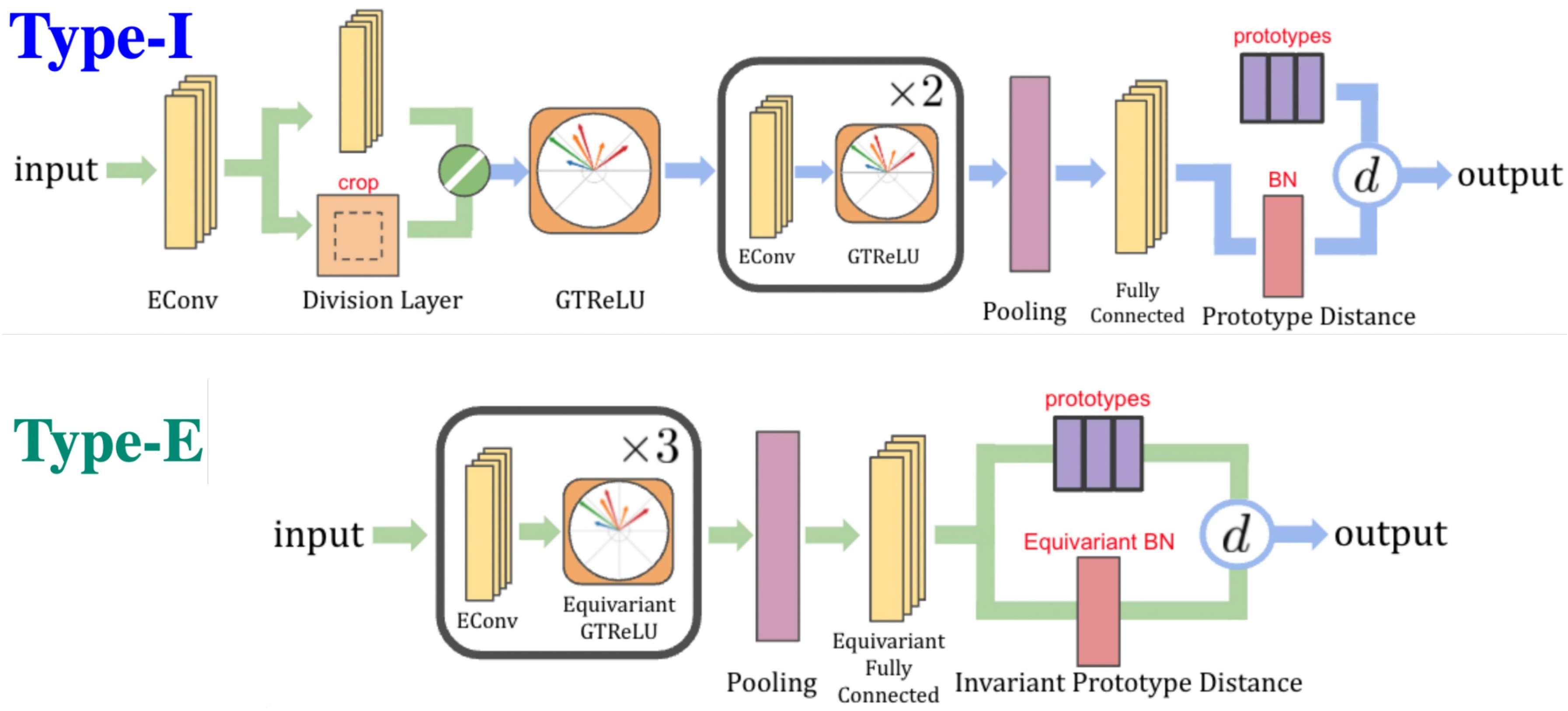
Invariant

Conjugate Layer

Division Layer

Prototype-Distance Invariant Layer

Two Architecture Styles



Summary

- Multi-Band imaging is invaluable for HADR applications.
- Traditional transfer learning approaches are not readily applicable.
- We propose using co-domain symmetric models trained from scratch.
- We propose a complex-valued encoding and use complex-scale invariant models.
- The resulting models have higher accuracy, significantly fewer parameters, no augmentation, no pre-training, and no RGB conversion

Thank you!