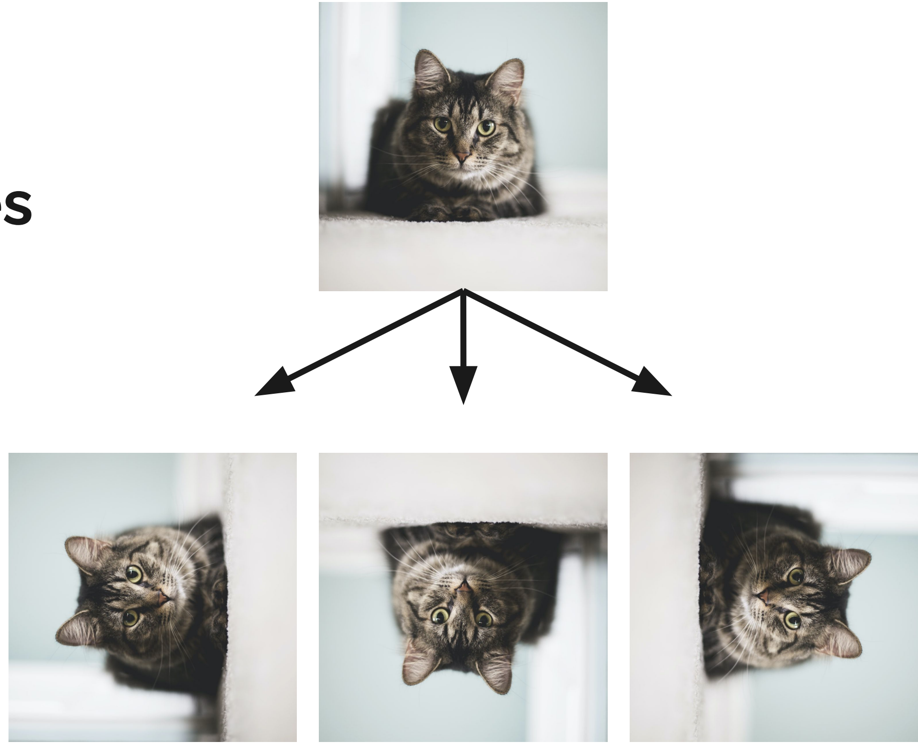
Learning Broken Symmetries with Resimulation and Encouraged Invariance

Edmund Witkowski, Daniel Whiteson

University of California, Irvine

Learning Broken Symmetries

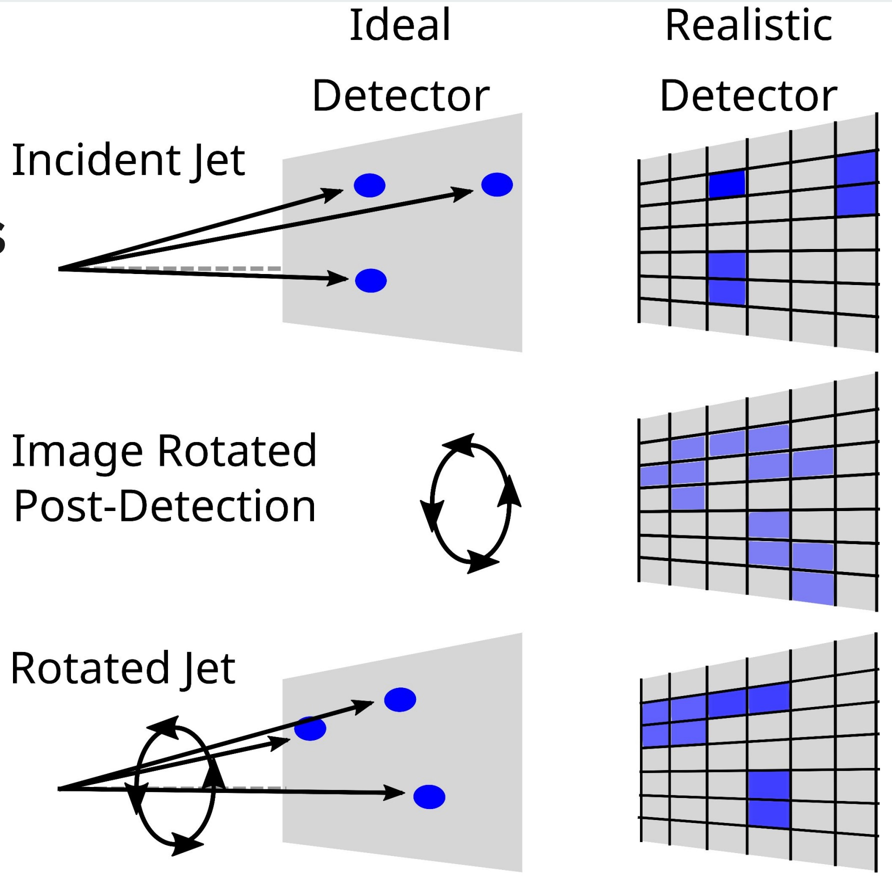
- Symmetries (e.g. rotational) in data can be leveraged to improve learning
- Augment dataset with transformations[1]
- Build symmetry into network[2] (e.g. network only performs invariant operations)



Synthesizing augmented data

Learning Broken Symmetries

- Symmetries can be broken in observed data
- Typical symmetry methods become less effective
- Detector effects can break rotational symmetry in jet images
- Transformations post-detection lead to non-physical artifacts



Learning Broken Symmetries

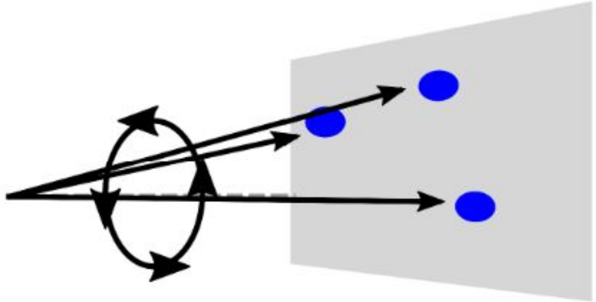
- Jet images tend to be sparse
- Artifacts disproportionately dominate information content in sparse images
- Can influence learned strategy relatively more than in non-sparse computer vision tasks

Not sparse - artifacts negligible[3]

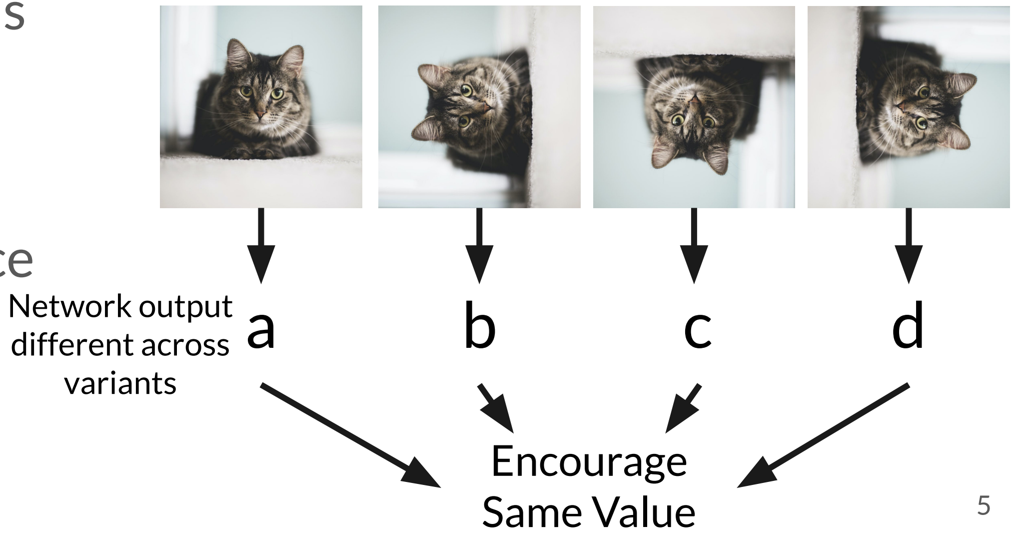
Sparse - artifacts prominent

Learning Broken Symmetries - Solutions

- Applying transformations pre-detection (in sim) avoids artifacts
- Can additionally explicitly encourage output invariance
- Transfer learning for use with real data (we don't explore this here)



Rotate jet instead of detected image



Learning Broken Symmetries - Encouraged Invariance

- Invariance encouraged in loss function
- Accumulate outputs for variants of same image
- Penalize for non-zero standard deviation in set

General form of loss

$$L = aL_{\text{cls}} + bL_{\text{inv}}$$

a, b scalar coefficients

L_{cls} Classification term

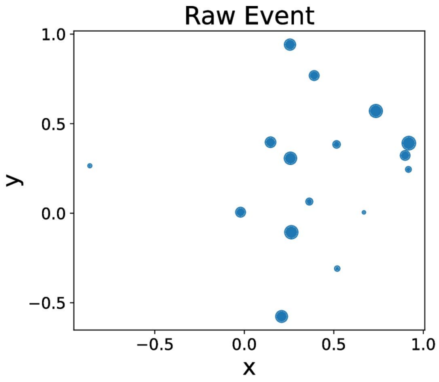
L_{inv} Invariance term

We use BCE for classification, MSE to penalize for non-zero std. dev.

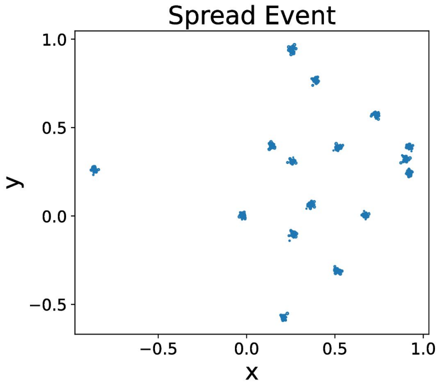
Learning Broken Symmetries - Toy Data

- Toy dataset generated to have similar characteristics to jet images
- Draw energy values from uniform dist.
- Background locations distributed uniformly on disk
- Signal locations are more clustered
- Deposits spread slightly so they may occupy multiple pixels when binned

Initial deposit locations and intensities

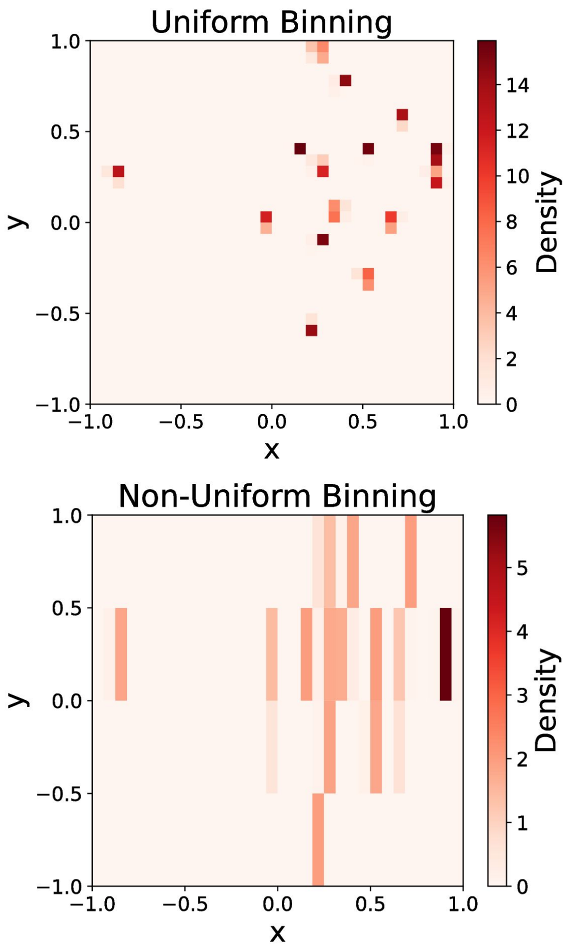


Deposits split into multiple, giving width for binning



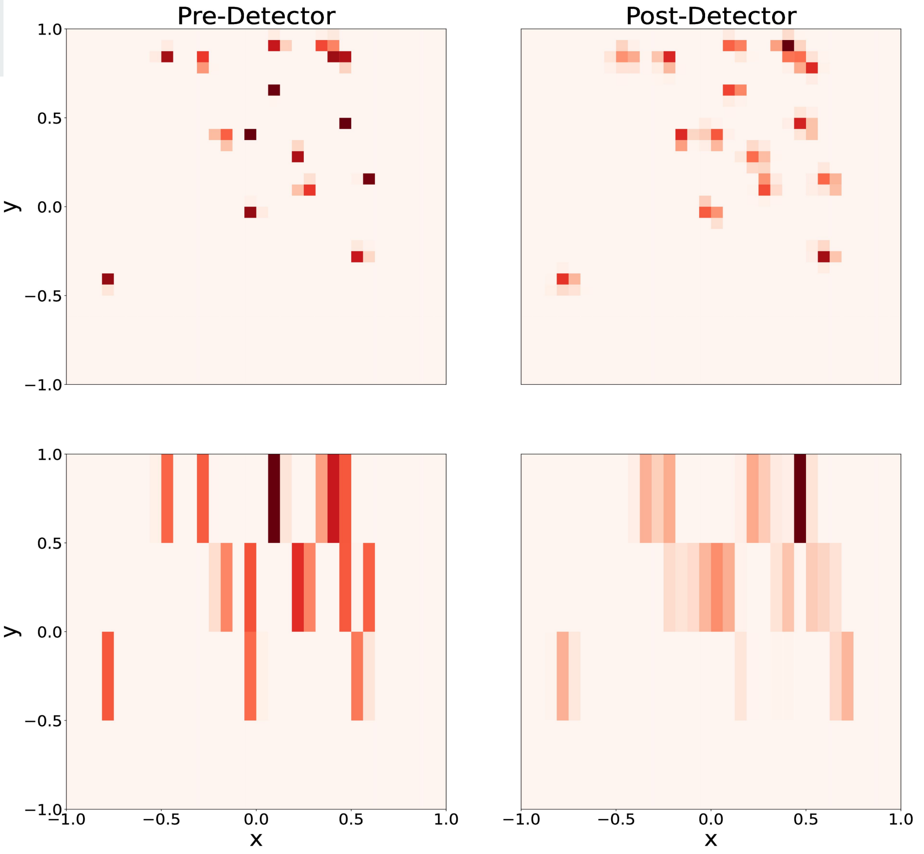
Learning Broken Symmetries - Toy Data

- Pixelate resulting data for basic detector response
- Use square and rectangular pixels
- Rectangular pixels exaggerate symmetry breaking



Learning Broken Symmetries - Toy Data

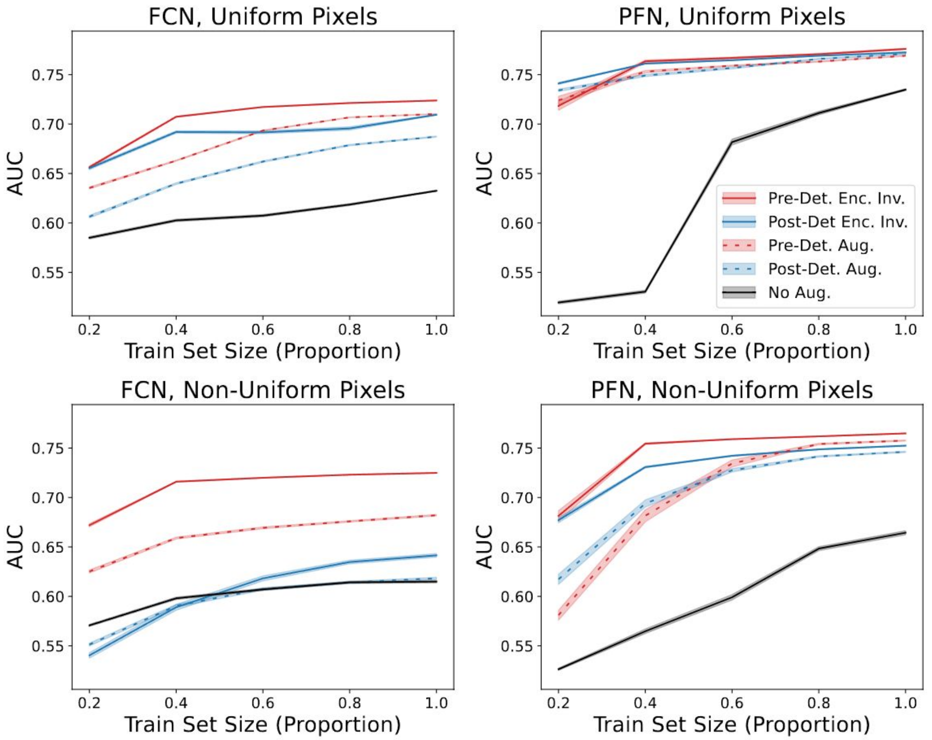
- 8 variants of each event are created, rotated at 45° increments
- 2 versions of these sets are created: transformed pre and post detector



45° rotated variants of event from previous slide. Artifacts in post-det. rotations are visually apparent - images are blurrier

Learning Broken Symmetries - Results

- Scan w/ each aug + pix method over training set sizes
- Fully Connected Networks (FCN) + Particle Flow (PFN)
- Pre-det. aug. similar to or better than post-det. aug.
- Encouraged invariance w/ pre-det. aug. does best



Learning Broken Symmetries - Results

- Pre-det. aug. has clear benefits in the case of a simpler network
- Encouraging invariance outperforms aug. alone
- Encouraging invariance pre-det. does best, especially when symmetry is more obscured (rectangular pix)
- Differences more apparent for smaller training data

		Uniform bins		Non-uniform bins	
		small	large	small	large
Arch.	Augm.	set	set	set	set
FCN	None	0.585(1)	0.632(1)	0.571(1)	0.615(1)
	post-det. aug.	0.606(2)	0.687(1)	0.551(1)	0.618(1)
	pre-det. aug.	0.635(1)	0.710(1)	0.625(2)	0.682(1)
	post-det. inv.	0.655(2)	0.709(1)	0.540(3)	0.642(2)
	pre-det. inv.	0.656(1)	0.724(1)	0.672(2)	0.725(1)
PFN	None	0.519(1)	0.735(1)	0.526(1)	0.664(2)
	post-det. aug.	0.734(1)	0.771(1)	0.617(5)	0.746(1)
	pre-det. aug.	0.724(4)	0.770(1)	0.581(5)	0.758(1)
	post-det. inv.	0.741(1)	0.772(1)	0.677(3)	0.752(1)
	pre-det. inv.	0.718(4)	0.776(1)	0.681(5)	0.765(1)

Sources

[1] C. Shorten and T. M. Khoshgoftaar. A survey on image data augmentation for deep learning. Journal of Big Data, 6(1):60, 2019.

[2] A. Butter, G. Kasieczka, T. Plehn, and M. Russell. Deep-learned Top Tagging with a Lorentz Layer. SciPost Phys., 5(3):028, 2018.

[3] A. Krizhevsky. Learning multiple layers of features from tiny images. pages 32–33, 2009.