Equivariant Neural Networks for Dark Matter Morphology with Strong Gravitational Lensing

Apoorva Vikram Singh

Mentors

K Pranath Reddy

Jeremy Quijano

Michael Toomey

Anna Parul

ML4SCI

Goal

- Current ***convolutional neural networks are only capable of translational equivariance***. However, in a number of application (including ours), a ***larger groups of symmetries, including rotations and reflections*** are present in the data as well that needs to be exploited. This gives rise to the notion of ***Equivariant Convolutional Networks***.

- ***E2-CNNs*** (proposed by Weiler et. al) propose a solution to this by guaranteeing a specified transformation behavior of their feature spaces under transformations of their input.

- ***E2-CNNs*** are ***equivariant under all isometries E(2)*** of the image plane i.e. under *translations, rotations and reflections.*

- We design a network that uses equivariant convolutional layers instead of a vanilla convolutional layer to capture the symmetries present in the data.

Datasets

- We use two different datasets to train and test our network:
  - Model F (corresponding to model A in the paper)
  - Model J (corresponding to model B in the paper)

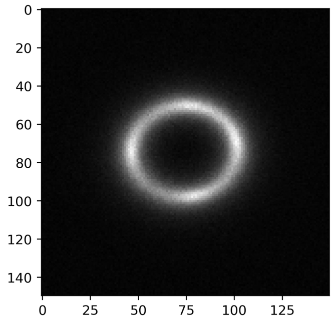
- The paper mentioned above: *Alexander, Stephon, et al. "Deep Learning the Morphology of Dark Matter Substructure." The Astrophysical Journal 893.1 (2020): 15.*

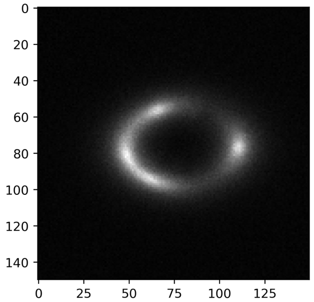
- Both model F and J contain 75000 images for training and 7500 images for testing. The images are evenly divided with each of the three class having 25000 training images and 2500 testing images.

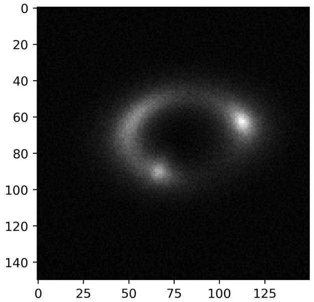
Datasets

No Substructure | **Spherical Substructure** | **Vortex Substructure**

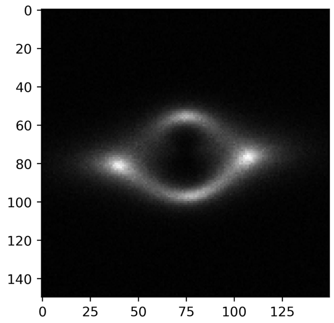
MODEL F

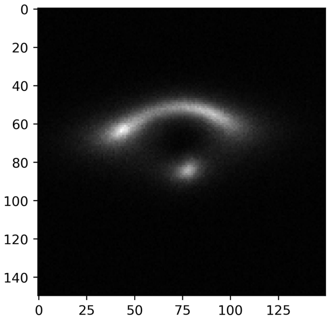


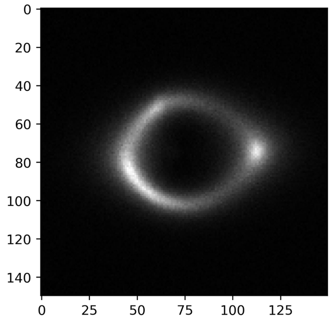




MODEL J







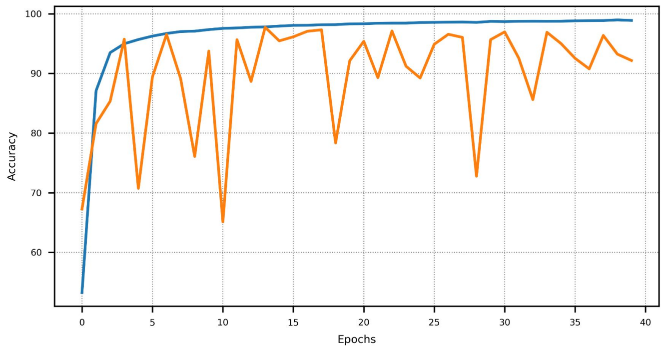
Preprocessing

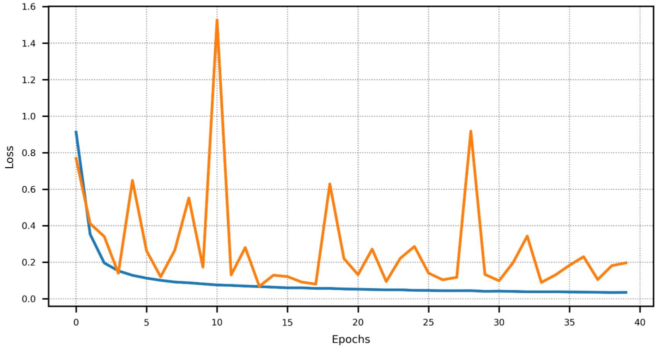
- ***Cropped*** to size 128 $\times$ 128 from original size 150 $\times$ 150.

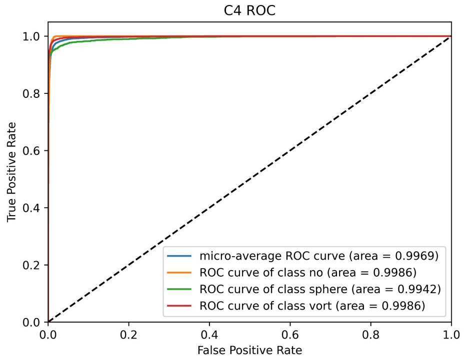
- ***Padding*** to increase the size to 129 $\times$ 129. This allows us to use odd-size filters with stride 2 when downsampling a feature map in the model.

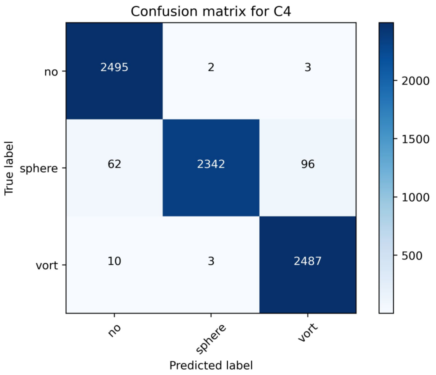
- We ***upsample*** an image by a factor of 3, ***rotate*** it and finally ***downsample*** it again. This allows us to reduce interpolation artifacts (e.g. when testing the model on rotated images).

Results: Model F (C4 symmetry)

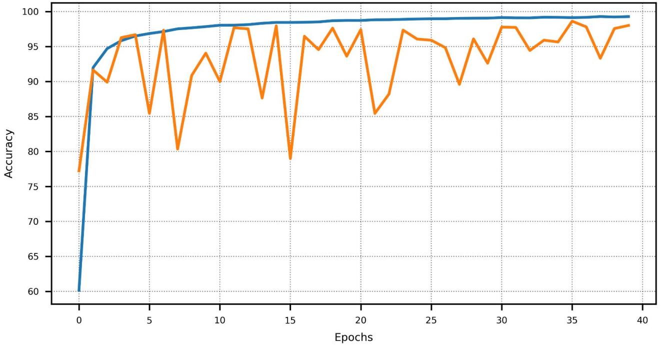


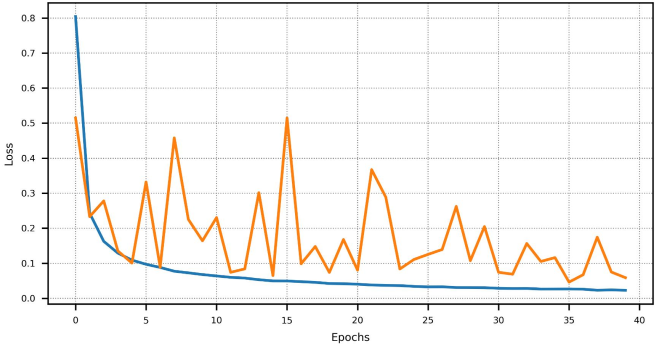


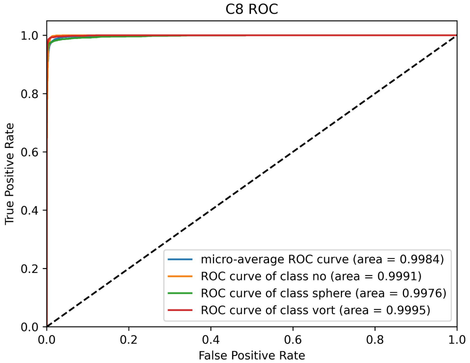


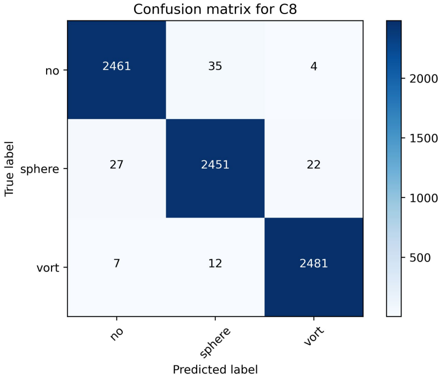


Results: Model F (C8 symmetry)

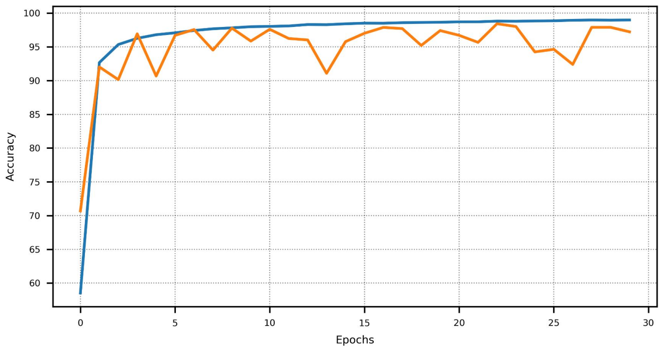


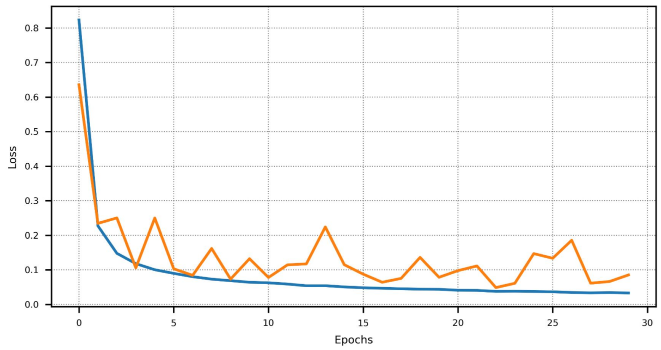


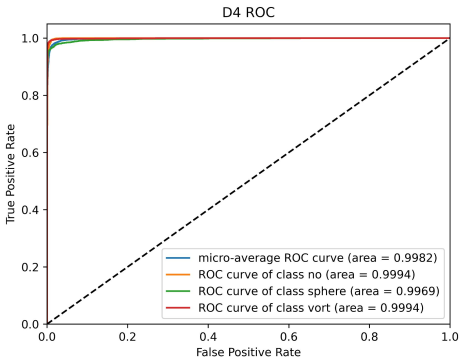


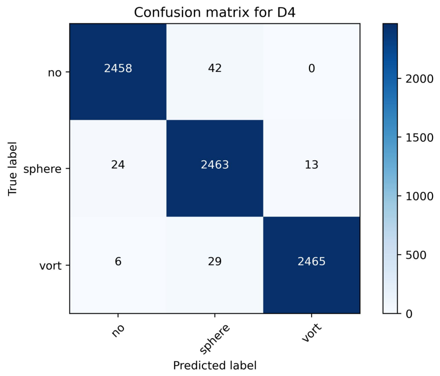


Results: Model F (D4 symmetry)

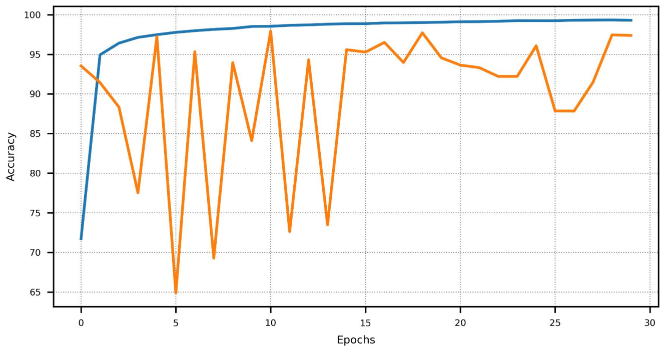


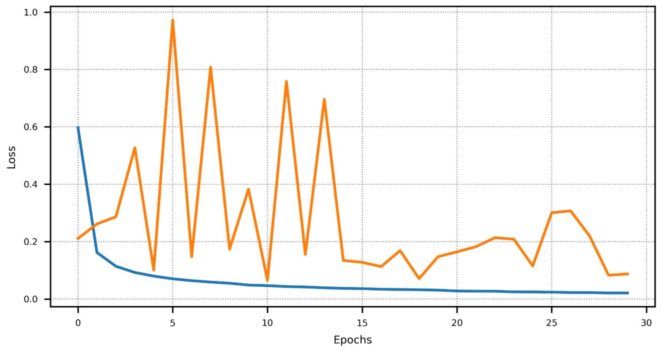


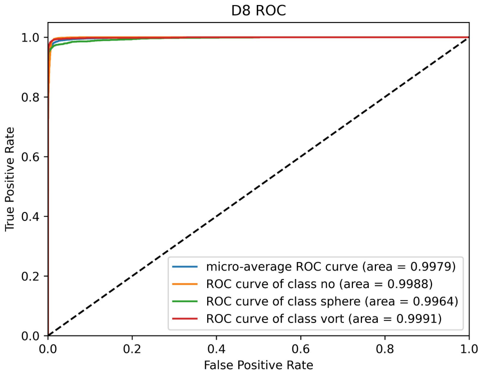


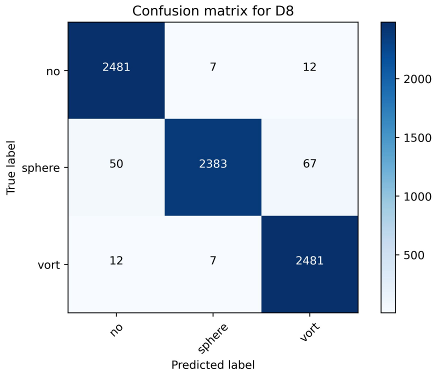


Results: Model F (D8 symmetry)

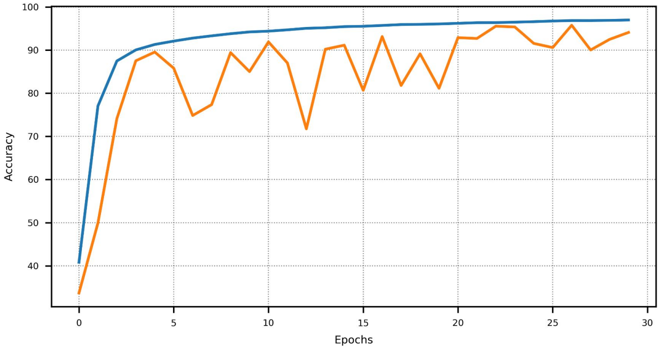


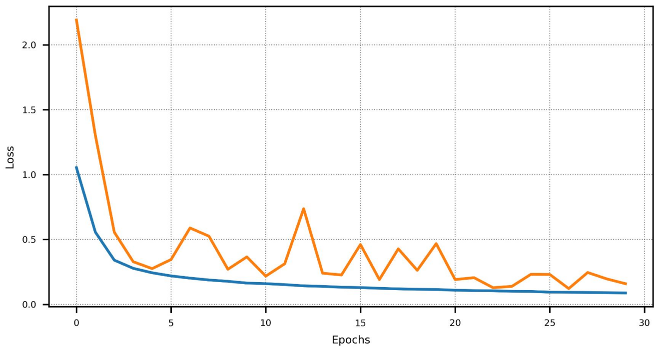


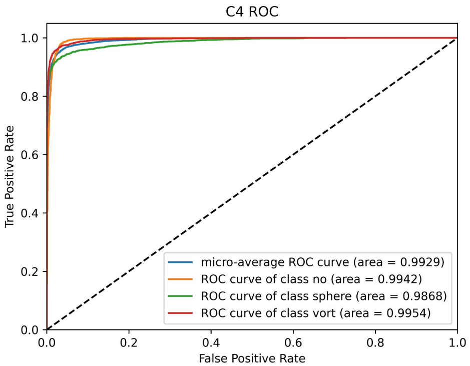


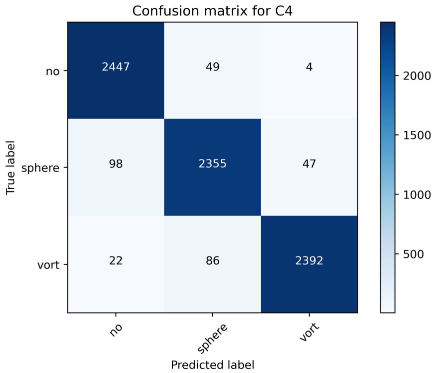


Results: Model J (C4 symmetry)

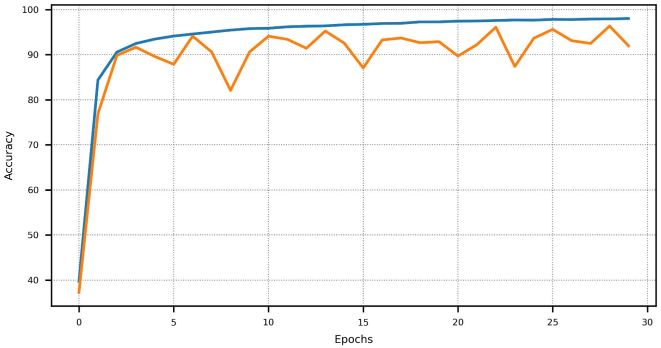


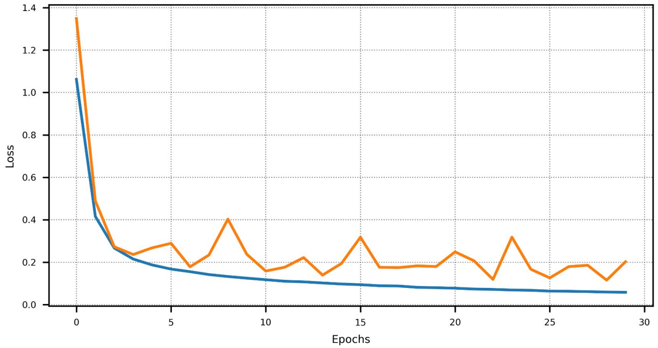


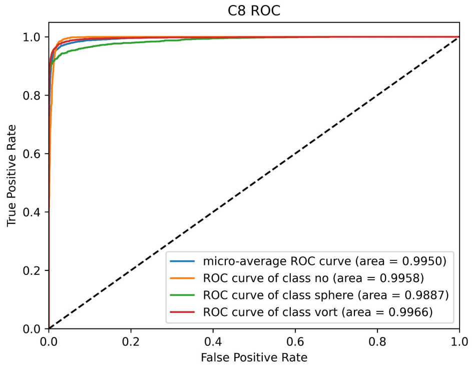


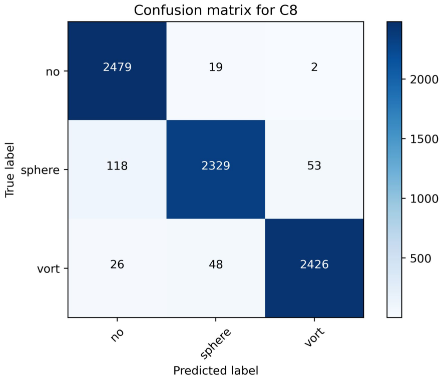


Results: Model J (C8 symmetry)

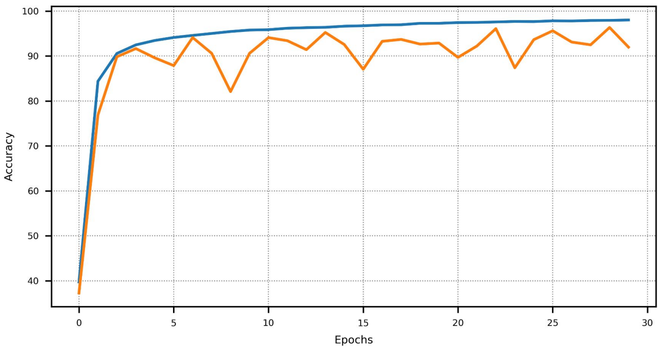


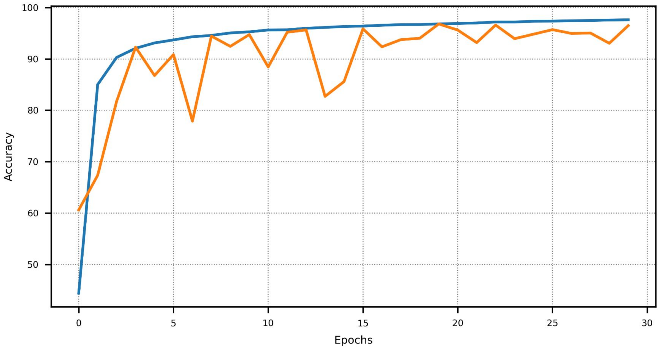


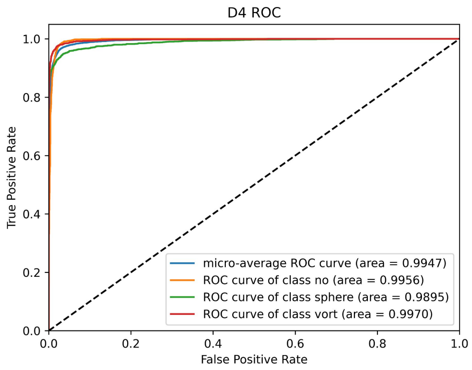


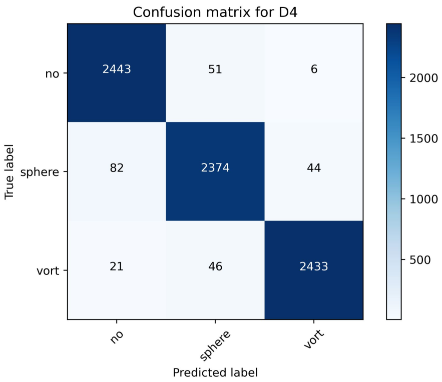


Results: Model J (D4 symmetry)

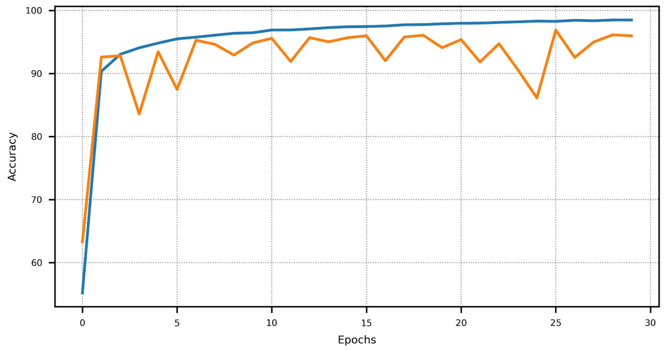


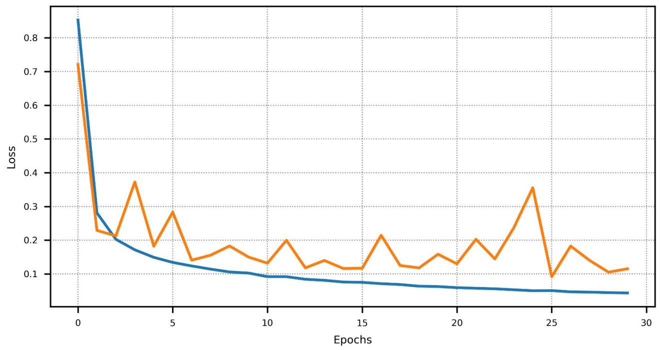


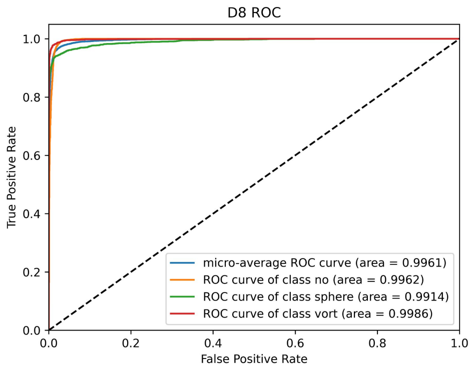


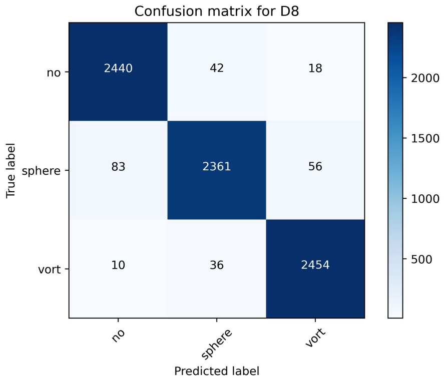


Results: Model J (D8 symmetry)

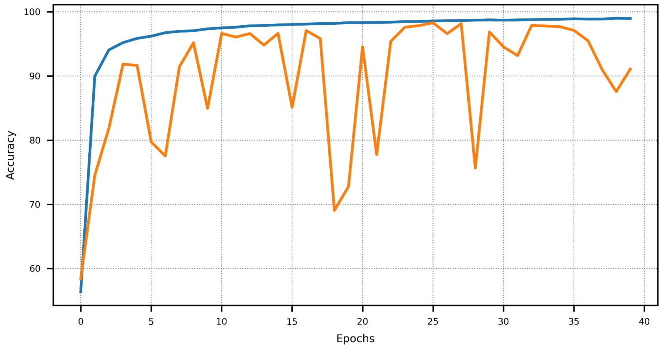


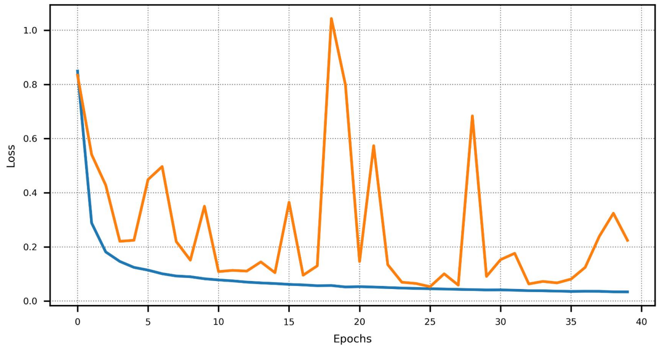


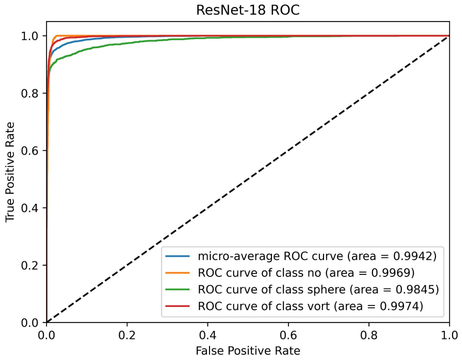


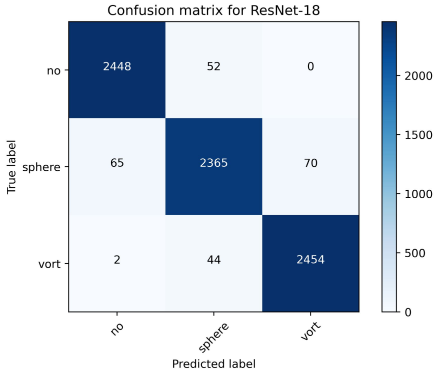


Results: Model F (CNN Model: ResNet-18)

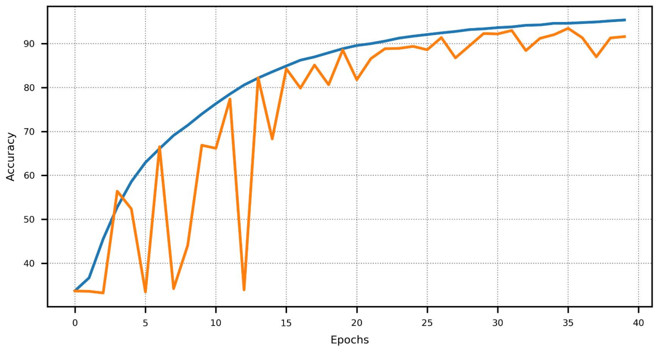


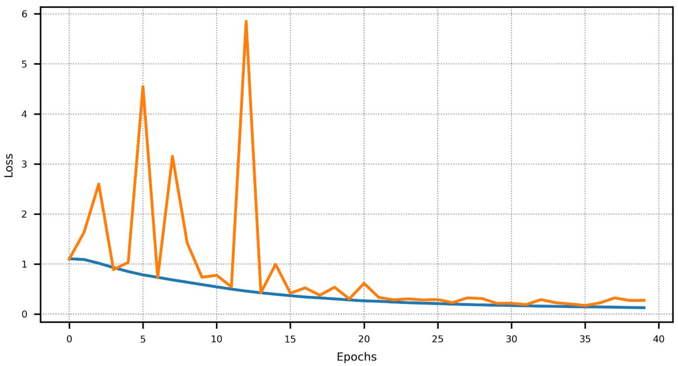


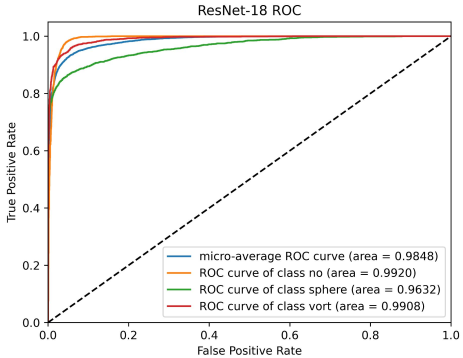


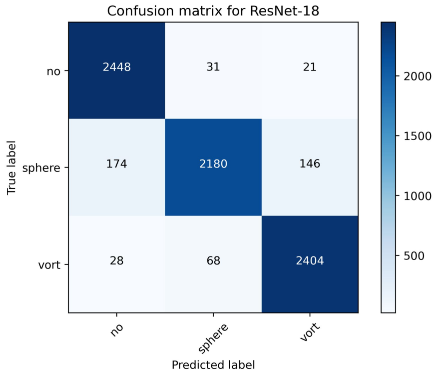


Results: Model J (CNN Model: ResNet-18)



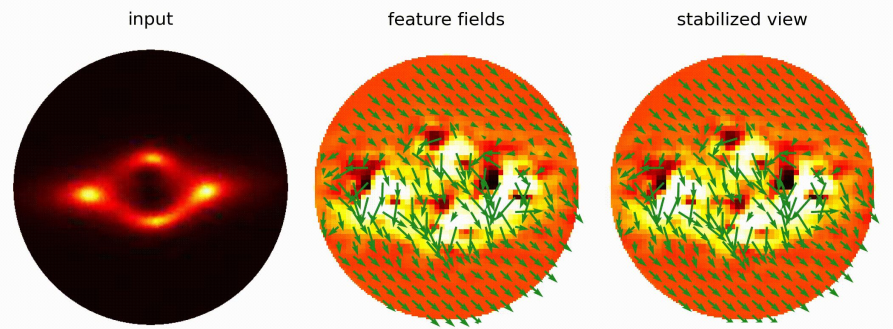




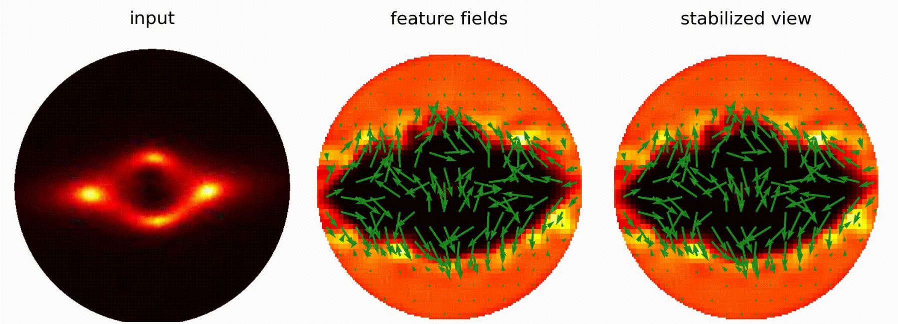


Feature Field Visualization (No sub)

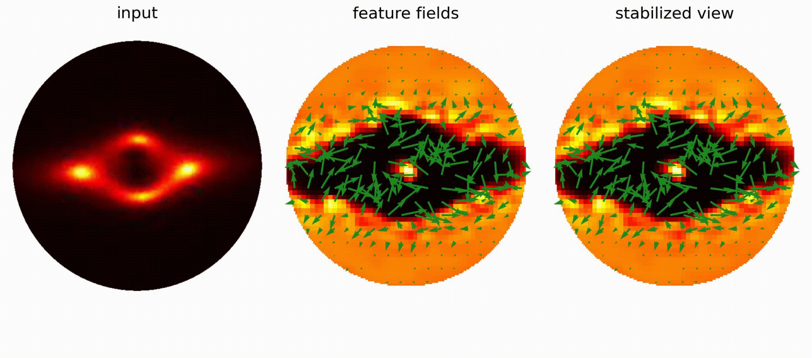
CNN



C4

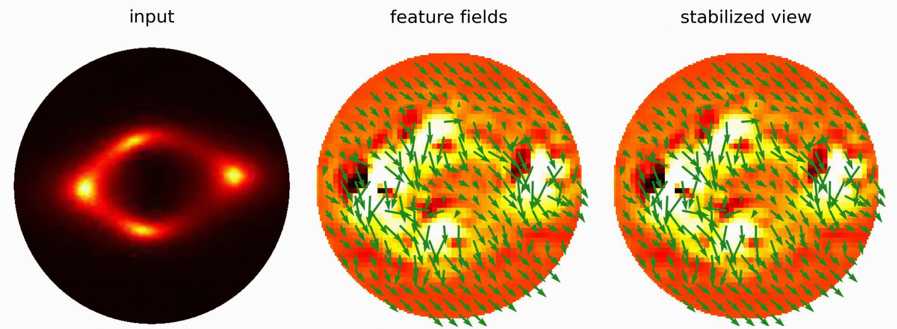


C8

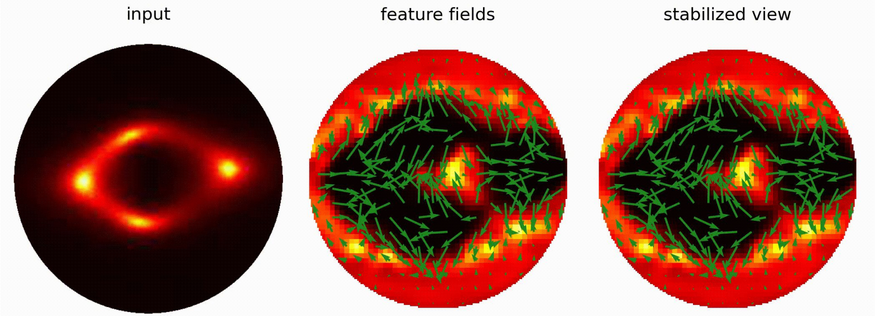


Feature Field Visualization (Spherical)

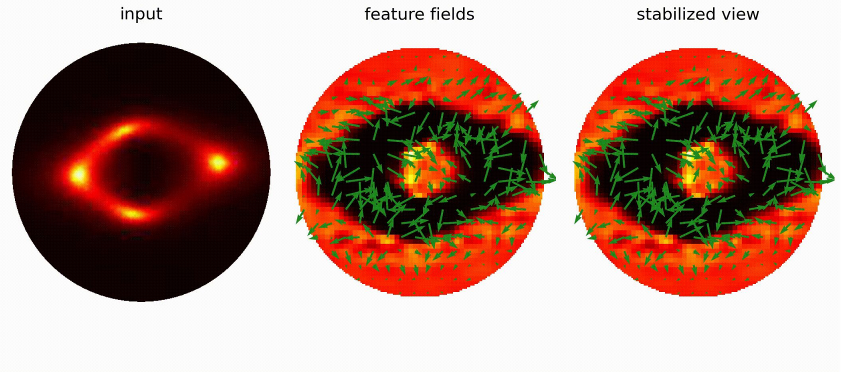
CNN



C4

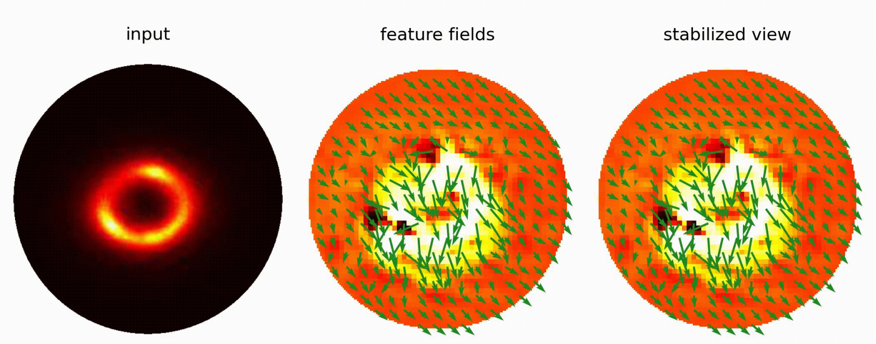


C8

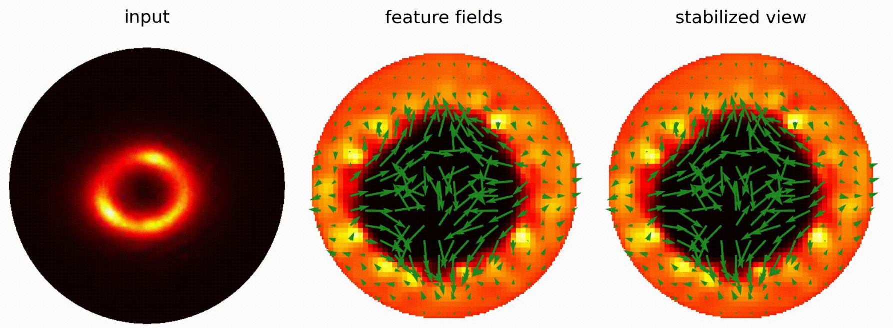


Feature Field Visualization (Vortex)

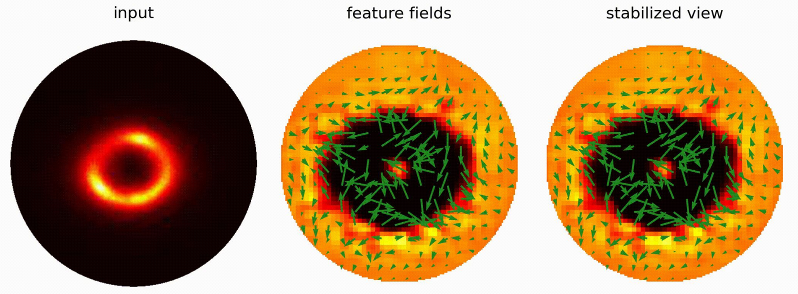
CNN



C4



C8



Thank You
for your attention.