Group Equivariant Convolutional Networks

Andrei Gasparovici

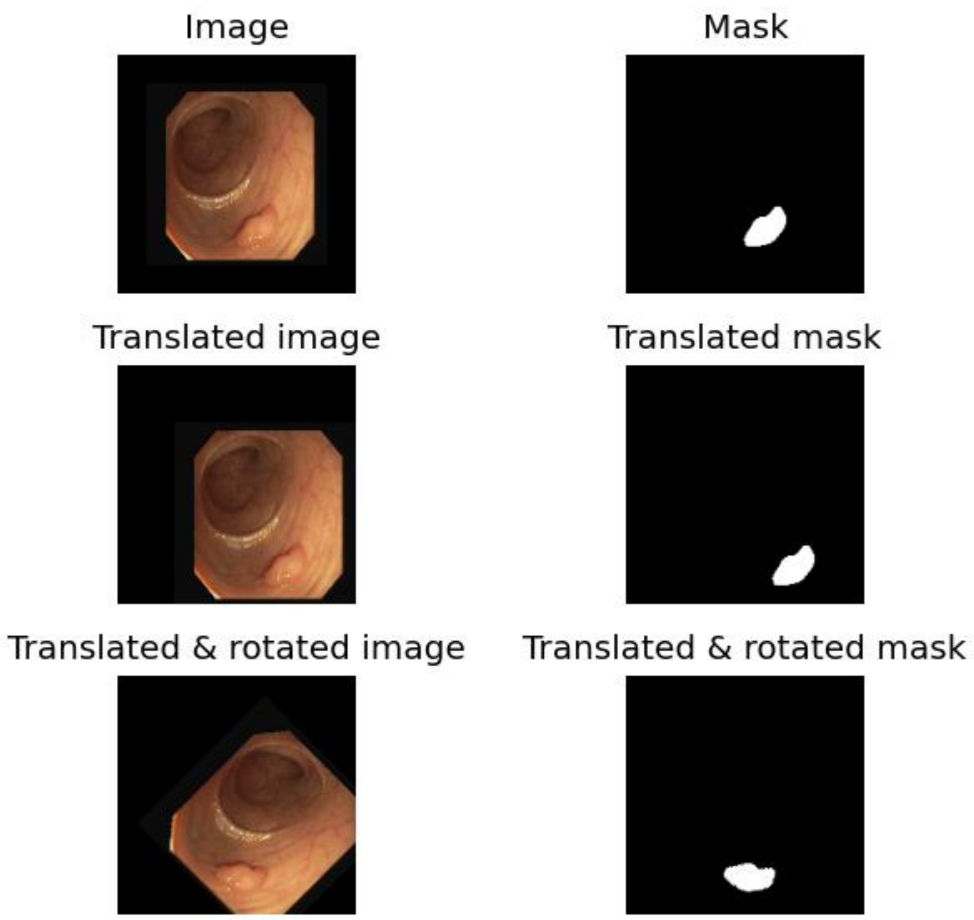
andrei.gasparovici.ext@siemens.com

Agenda

- Introduction to equivariance & group theory
- Methods for roto-translation equivariance:
  - $p4$-equivariant CNNs
  - Interpolation-based $SE(2)$-CNNs
  - Steerable $SE(2)$-CNNs
- Equivariance experiments on medical images
- Empirical results on medical images
- Implementation details
- Conclusions

Motivation

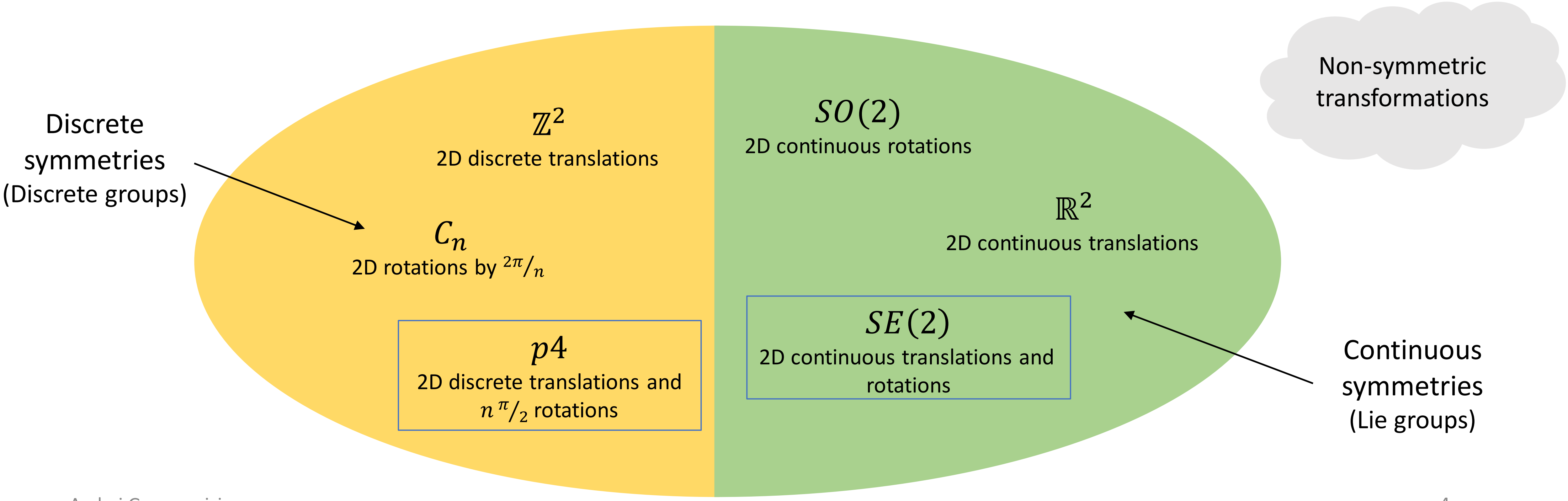
- The convolution operation is **equivariant** under **translation**, causing CNNs to share the same property
- Translation equivariance eliminates the need for **translation augmentation** and allows **parameter sharing**
- Generalizing CNNs for equivariance under larger groups of symmetries (such as rotations and reflections) **removes the need for data augmentation** and further enhances **robustness** and **weight sharing**



Example: Roto-translation equivariance

Symmetries and group theory

- Symmetries = reversible and composable transformations
- The theoretical framework for working with symmetries is **group theory**



Group action and equivariance

- A group G can **act** on images:

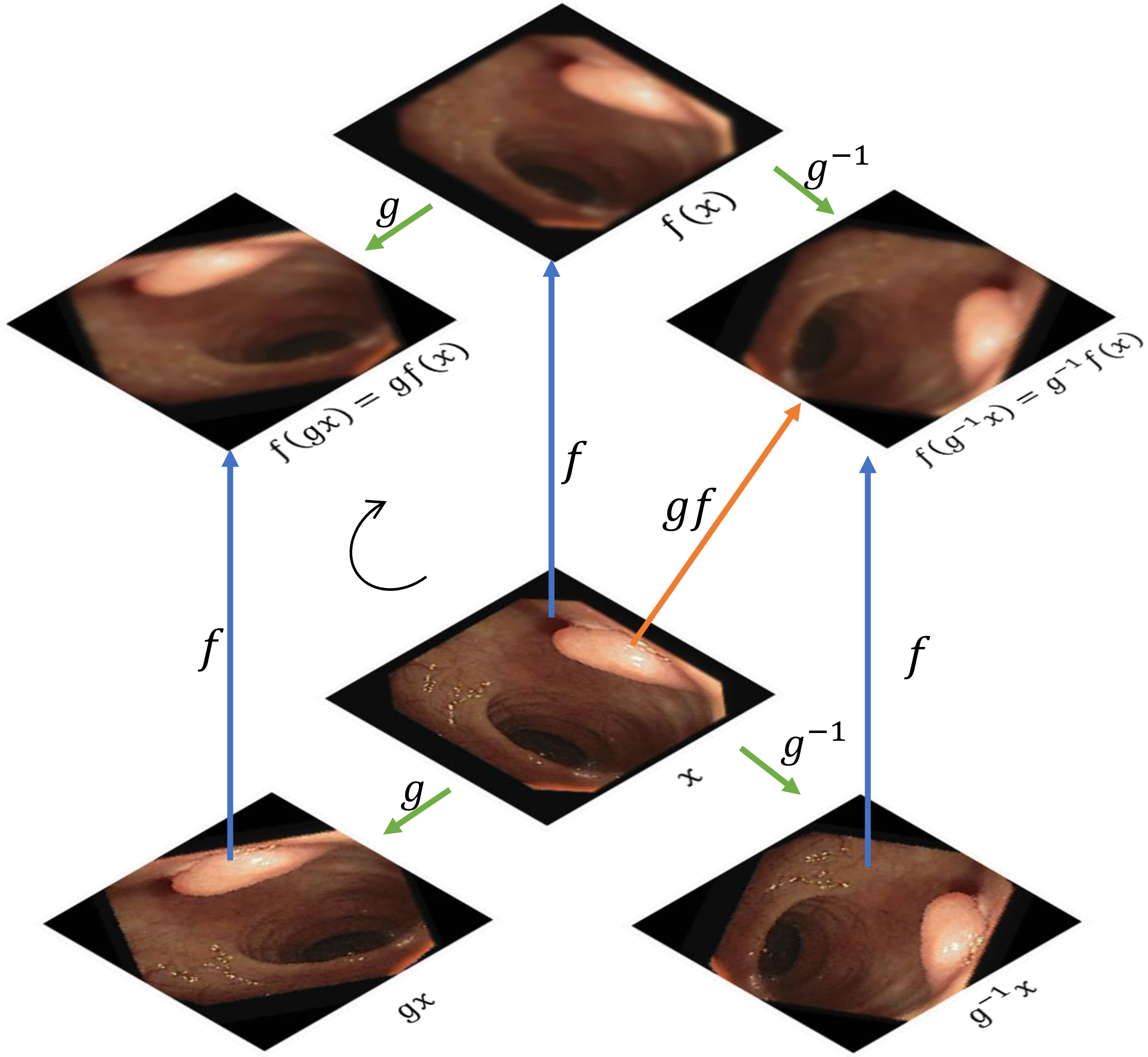
$$x \to gx$$

and on transformations:

$$f \to gf$$
$$(gf)(x) = f(g^{-1}x)$$

- A transformation f is ***G*-equivariant** if it commutes with the group action:

$$f(gx) = gf(x)$$



Group convolution

- The **convolutional structure** is not a preference, but a **requirement** for group equivariance
- To construct group equivariant networks, the regular (spatial) convolution can be generalized to a group convolution:

Translate kernel

Regular 2D convolution:
(translation equivariant)

$$(f * k)(x) = \int_{\mathbb{R}^2} f(t) \cdot k(x - t) dt$$

$$(f * k)(g) = \int_{\mathbb{R}^2} f(x) \cdot (gk)(x) dx = \int_{\mathbb{R}^2} f(x) \cdot k(g^{-1}x) dx$$

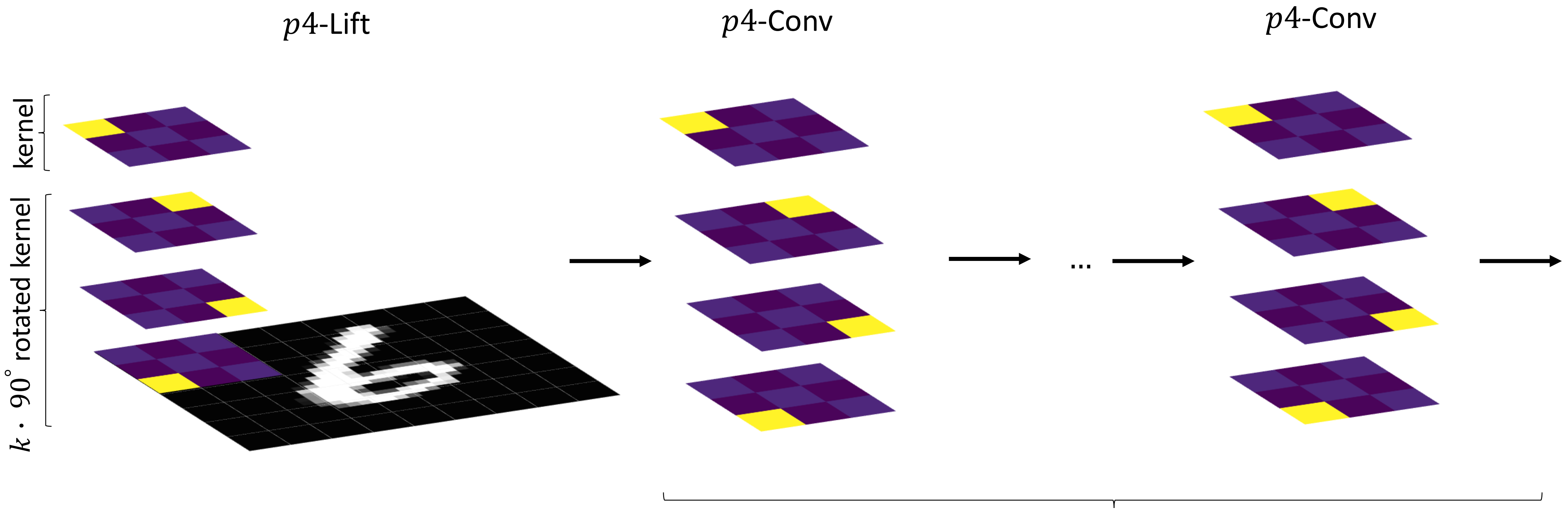
Group convolution:
(G equivariant)

Transform kernel by group action

$$(f * k)(g) = \int_G f(h) \cdot (gk)(h) dh = \int_G f(h) \cdot k(g^{-1}h) dh$$

Cohen et al. *A General Theory of Equivariant CNNs on Homogeneous Spaces* 2019

$n\,{}^{\pi}/_{2}$ rotations: G-CNNs



$$(k * f)(g) = \sum_{x \in \mathbb{Z}^2} k(g^{-1}x) f(x)$$

$$(k * f)(g) = \sum_{h \in G} k(g^{-1}h) f(h)$$

Cohen & Welling *Group Equivariant Convolutional Networks* 2016

Experiment: Polyp segmentation

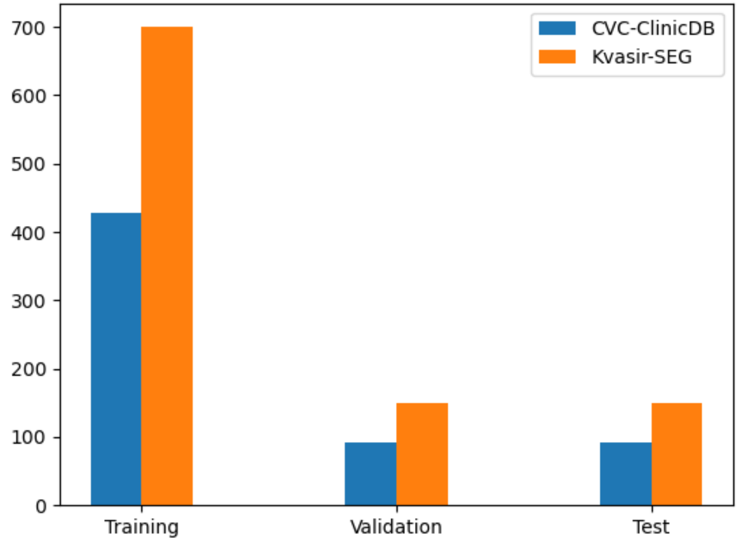
Datasets

CVC-ClinicDB

- Open-access dataset of polyp images and segmentation masks from colonoscopy sequences
- Size: 612 images
- Split: 70% training, 15% validation, 15% test

Kvasir-SEG

- Open-access dataset of gastrointestinal polyp images and corresponding segmentation masks
- Size: 1000 images
- Split: 70% training, 15% validation, 15% test

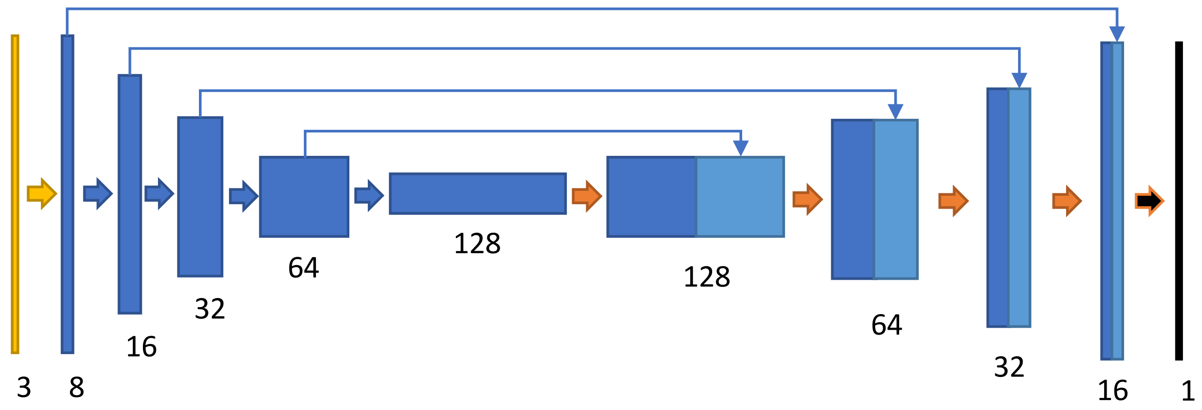


https://paperswithcode.com/dataset/cvc-clinicdb
https://paperswithcode.com/dataset/kvasir-seg

Experiment: Polyp segmentation

Architectures

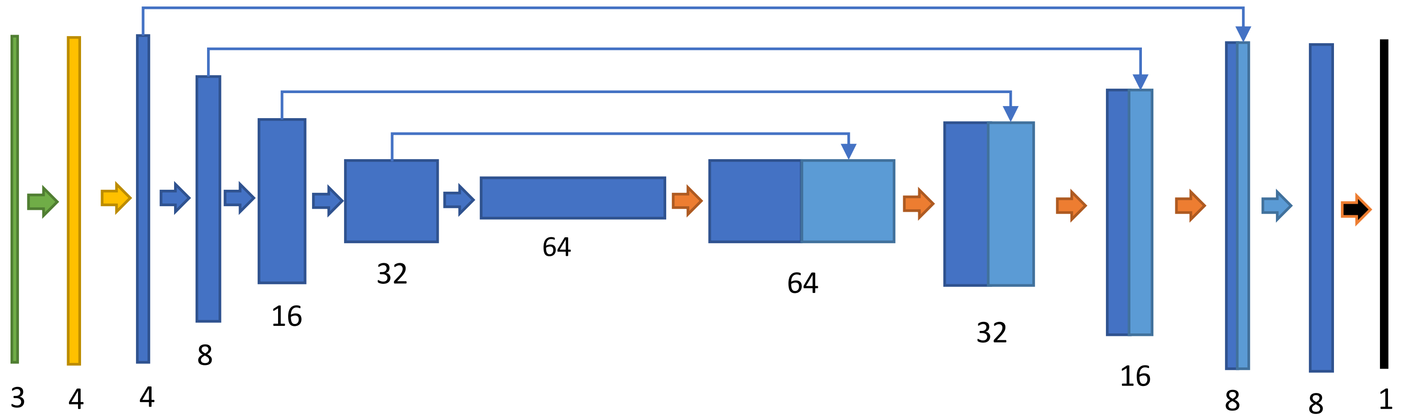
U-Net



- (Conv + BatchNorm + ReLU) × 2
- MaxPool + (Conv + BatchNorm + ReLU) × 2
- Copy
- Upsample + (Conv + BatchNorm + ReLU) × 2
- Conv

Number of parameters: 1,965,625

$p4$-equivariant U-Net



- $p4$-Lift (Conv + BatchNorm+ReLU)
- ($p4$-Conv + BatchNorm + ReLU) × 2
- Spatial MaxPool + ($p4$-Conv + BatchNorm + ReLU) × 2
- Copy
- Spatial Upsample + ($p4$-Conv + BatchNorm + ReLU) × 2
- Rotation MaxPool
- Conv

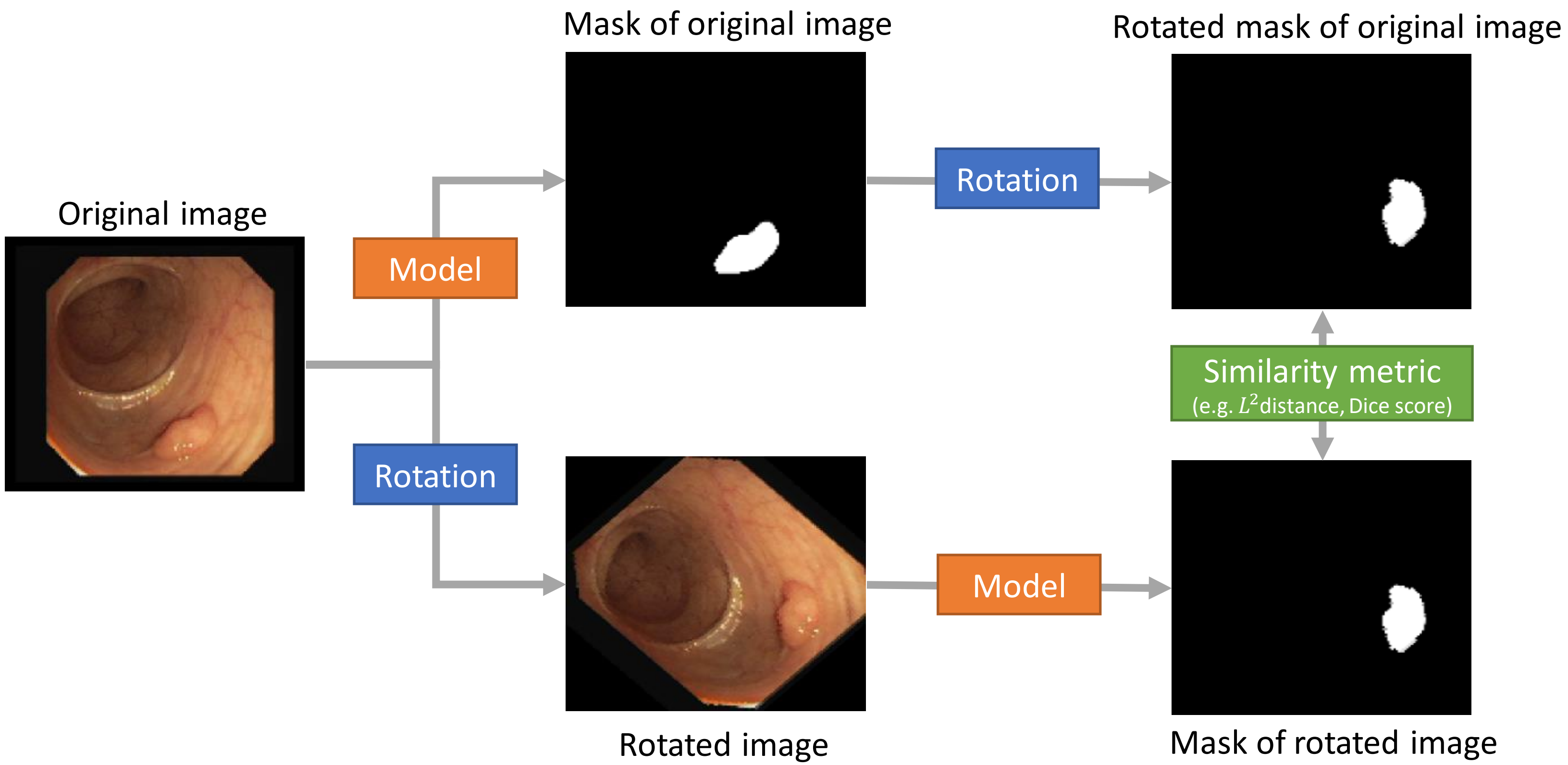
Number of parameters: 1,964,029

Ronneberger et al. *U-Net: Convolutional Networks for Biomedical Image Segmentation* 2015

Experiment: Polyp segmentation

Evaluating model equivariance

- To measure rotation equivariance, we compute a similarity metric between the **rotated mask** of the image and **the mask of the rotated image**:

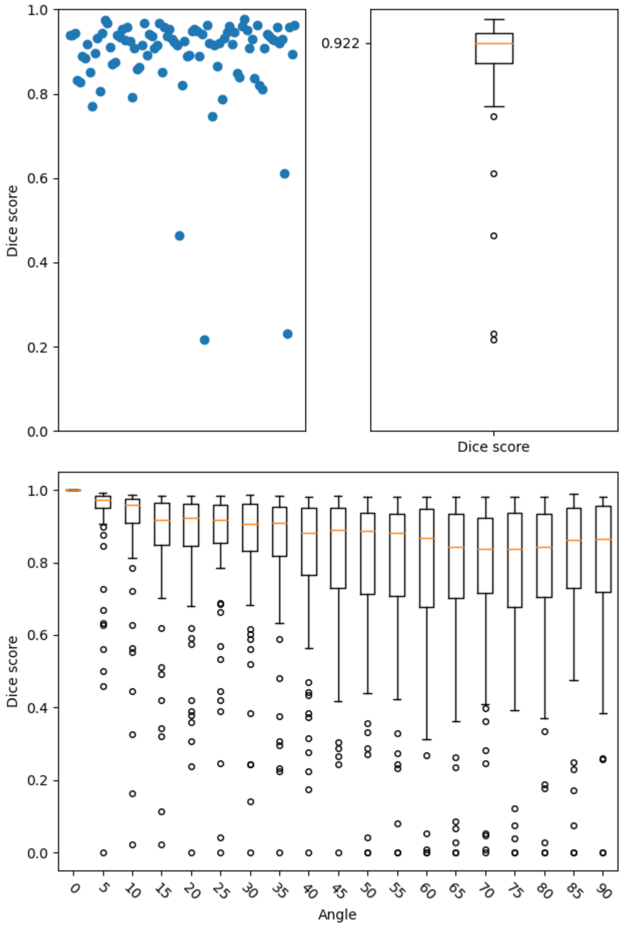


Experiment: Polyp segmentation

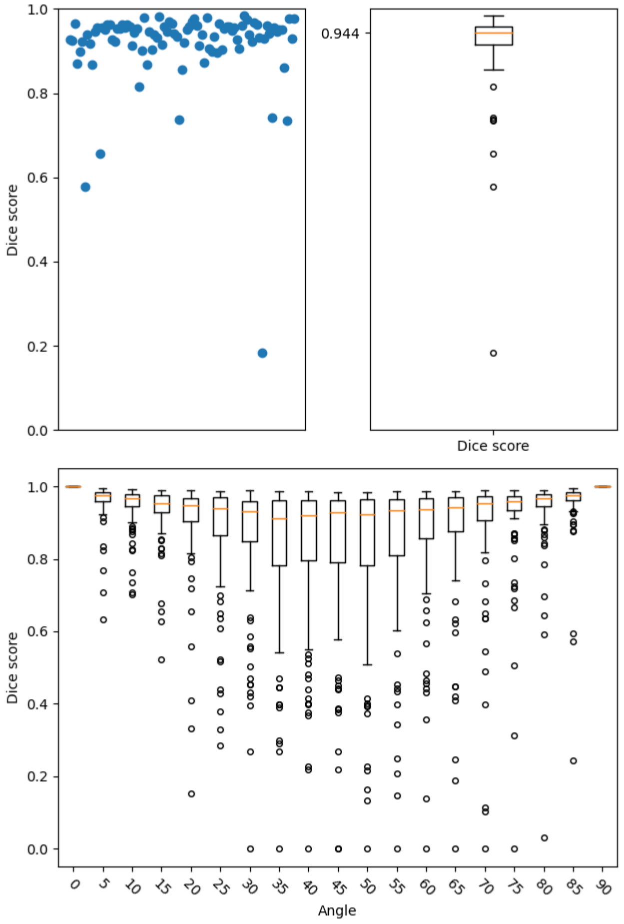
Results: CVC-ClinicDB

- The **equivariant** network obtains **better performance**
- **Perfect** equivariance at **0°** and **90°** (expected)
- **Good equivariance in between** – although kernels are rotated only by multiples of 90° (not expected)
- Final **rotation max pool** layer likely **improves equivariance** in the vicinity of 90°
- Plots show equivariance for 0°-90° although the **results are preserved for 90°-360°** (periodicity)

U-Net



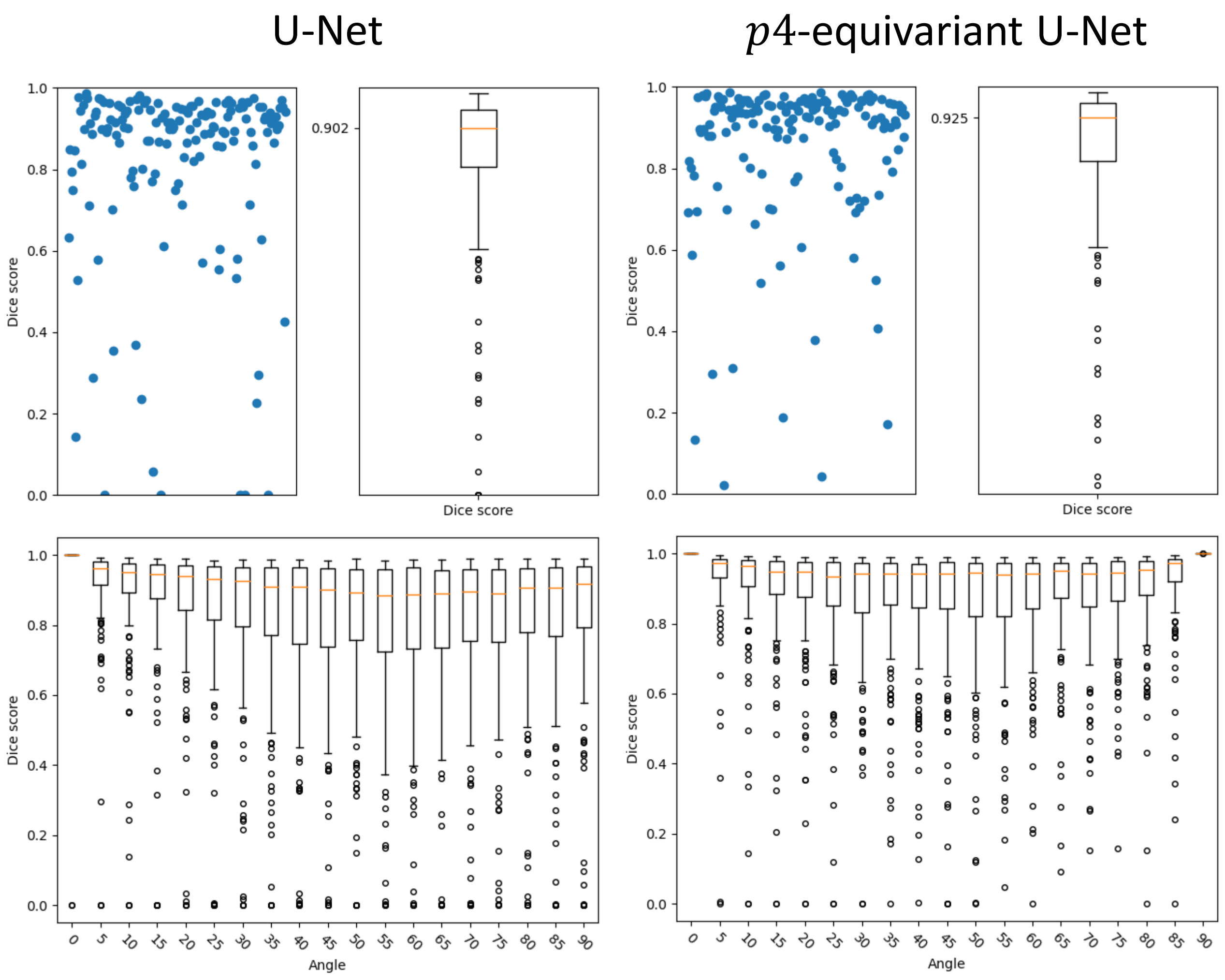
$p4$-equivariant U-Net



Experiment: Polyp segmentation

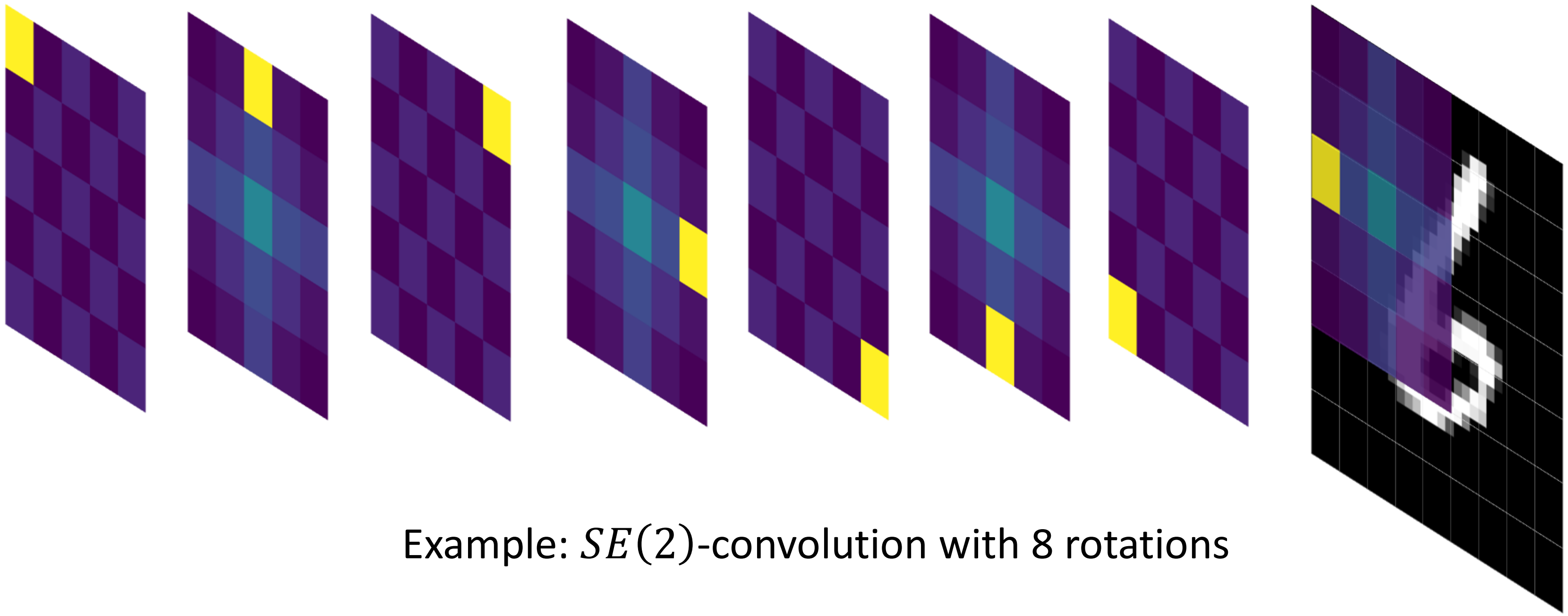
Results: **Kvasir-SEG**

- Kvasir-SEG is a more **complex** data set
- The **equivariant** network obtains **better performance**
- **Equivariance** results **preserve** across data sets



Arbitrary rotations: Angle discretization + G-CNN

- To rotate the (sampled) kernel by **arbitrary angles**, we construct a linear operator that performs rotation and **bilinear interpolation:**

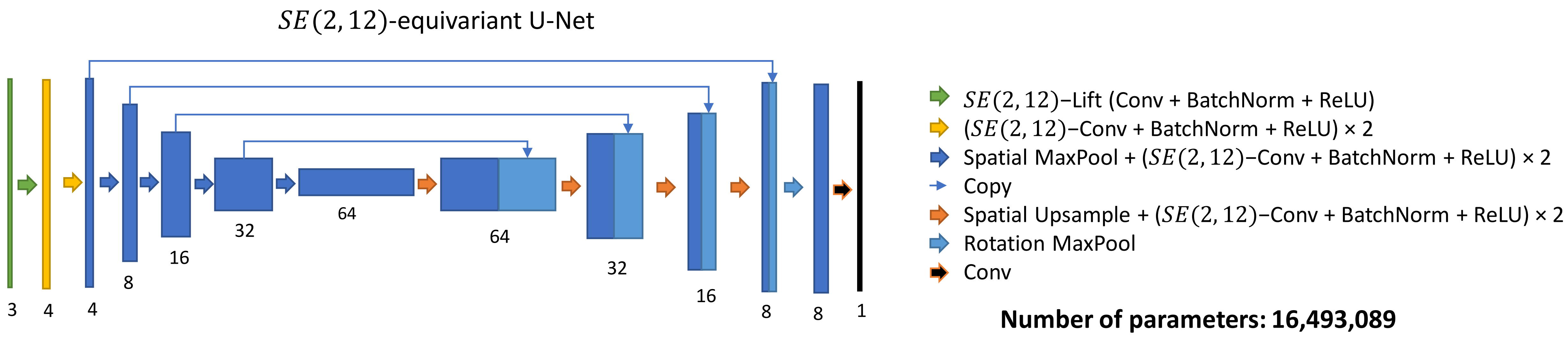


Example: $SE(2)$-convolution with 8 rotations

Bekkers et al. *Roto-Translation Covariant Convolutional Networks for Medical Image Analysis* 2018

Experiment: Polyp segmentation

Architecture

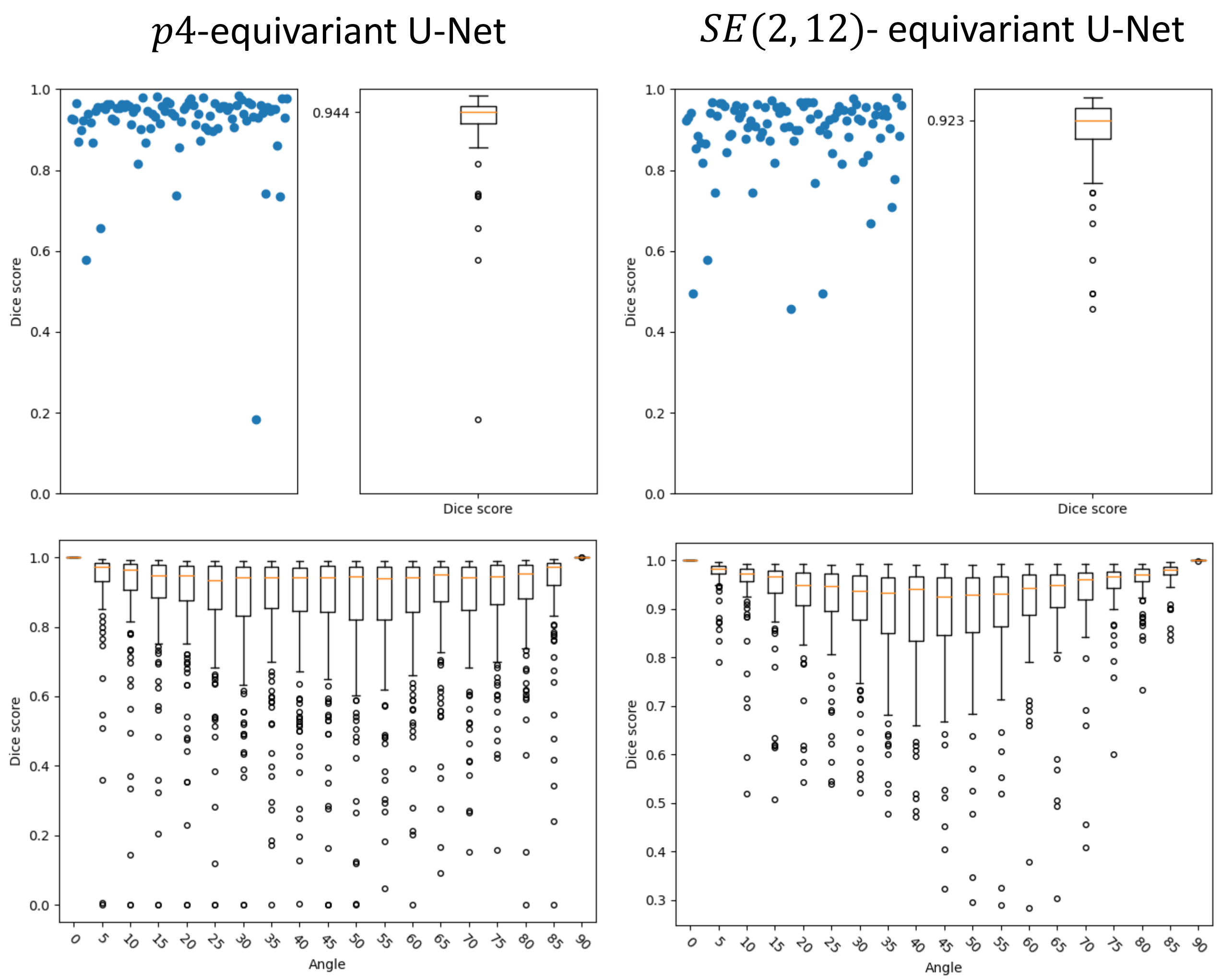


- Similar structure with the $p4$-equivariant U-Net
- Number of parameters **increases** with the number of **sampled angles**
- **Errors** due to interpolation are **expected**

Experiment: Polyp segmentation

Results: **CVC-ClinicDB**

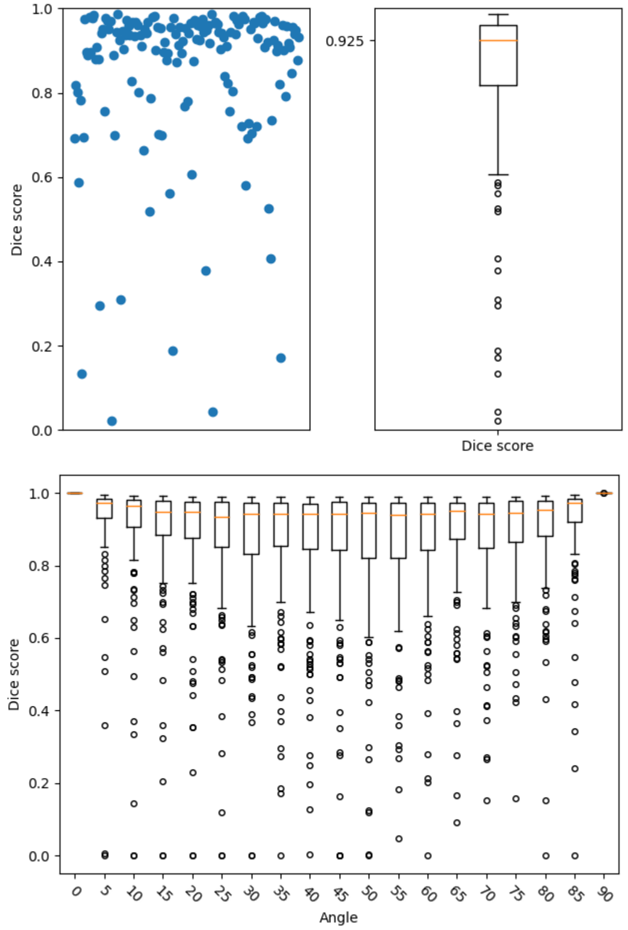
- **Perfect** equivariance at **90°** is preserved (expected)
- Equivariance at 30° and 60° not exact (expected due to interpolation)
- $p4$-equivariant network obtains better performance overall
- Interpolation introduces significant **errors**, although it **reduces outliers**
- Although employing only 90° rotations, $p4$-equivariant network achieves similar (even slightly better) equivariance



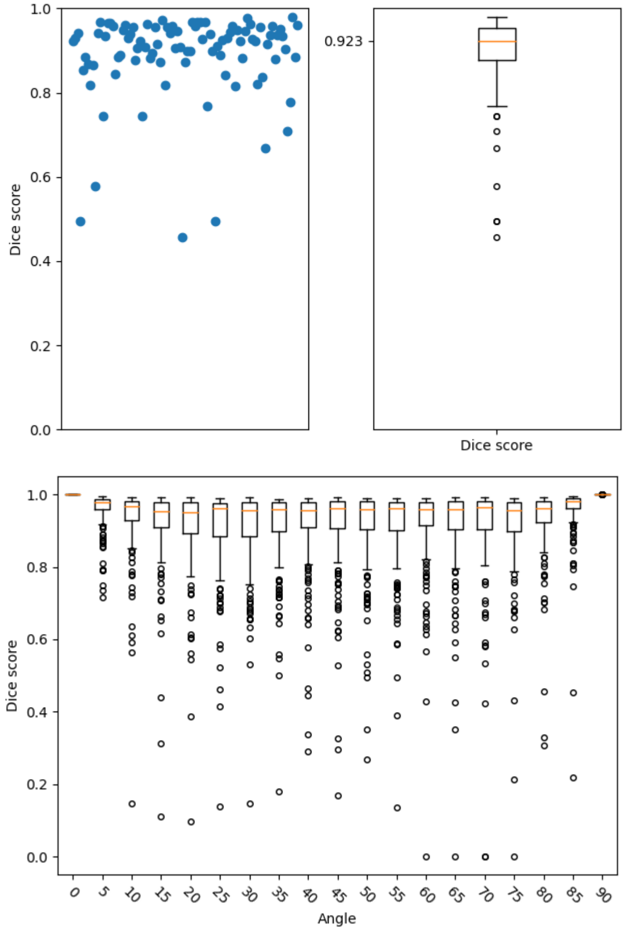
Experiment: Polyp segmentation

Results: **Kvasir-SEG**

$p4$-equivariant U-Net



$SE(2,12)$- equivariant U-Net



- $p4$-equivariant network attains better performance
- $SE(2,12)$- equivariant, while being surpassed in performance, has **fewer outliers**
- Robustness induced by $p4$-equivariance is not consistent over data sets

Continuous rotations: Steerable equivariant CNNs

Approximation by Fourier series and radial basis functions

- Considering a learnable filter $\widehat{\psi}$ to be complex-valued, we can approximate $\widehat{\psi}$ by its Fourier series as:

$$\widehat{\psi}(r,\theta) \approx \widehat{\psi}_N(r,\theta) = \sum_{n=-N}^{N} c_n(r) e^{in\theta}$$

- For every n, we can further approximate $c_n(r)$ using Gaussian radial basis functions:

$$c_n(r) \approx c_n^{J_n}(r) = \sum_{j=1}^{J_n} w_{nj}\, e^{-\frac{(r-j)^2}{2\sigma^2}}$$

- Finally, we obtain the following approximation of $\widehat{\psi}$:

$$\widehat{\psi}(r,\theta) \approx \sum_{n=-N}^{N} \sum_{j=1}^{J_n} w_{nj}\, e^{in\theta} e^{-\frac{(r-j)^2}{2\sigma^2}}$$

Arsenault & Sheng *Properties of the circular harmonic expansion for rotation-invariant pattern recognition* 1986

Continuous rotations: Steerable equivariant CNNs

Steerable filters

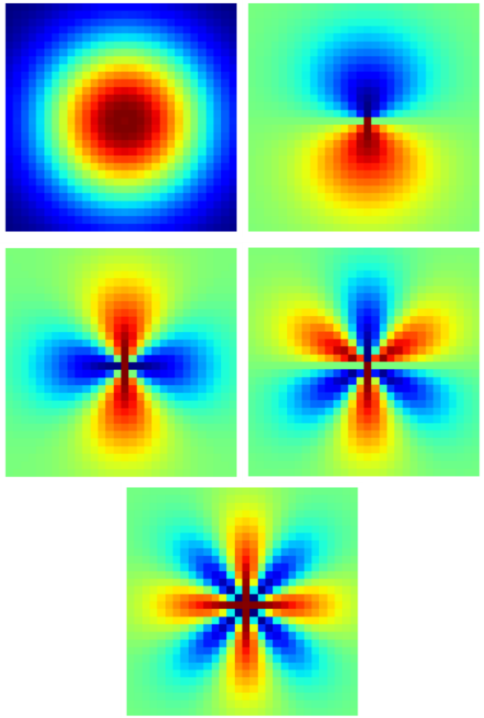
$$\hat{\psi}(r,\theta) \approx \sum_{n=-N}^{N} \sum_{j=1}^{J_n} w_{nj}\, e^{in\theta} e^{-\frac{(r-j)^2}{2\sigma^2}}$$

- The filters $\psi_{nj} = e^{in\theta} e^{-\frac{(r-j)^2}{2\sigma^2}}$ form a basis for the space of learnable filters:

$$\hat{\psi} = \sum_{n,j} \boxed{w_{nj}}\, \psi_{nj}$$

w_{nj} → Learnable parameters

- Such a filter $\hat{\psi}$ is called **steerable.**



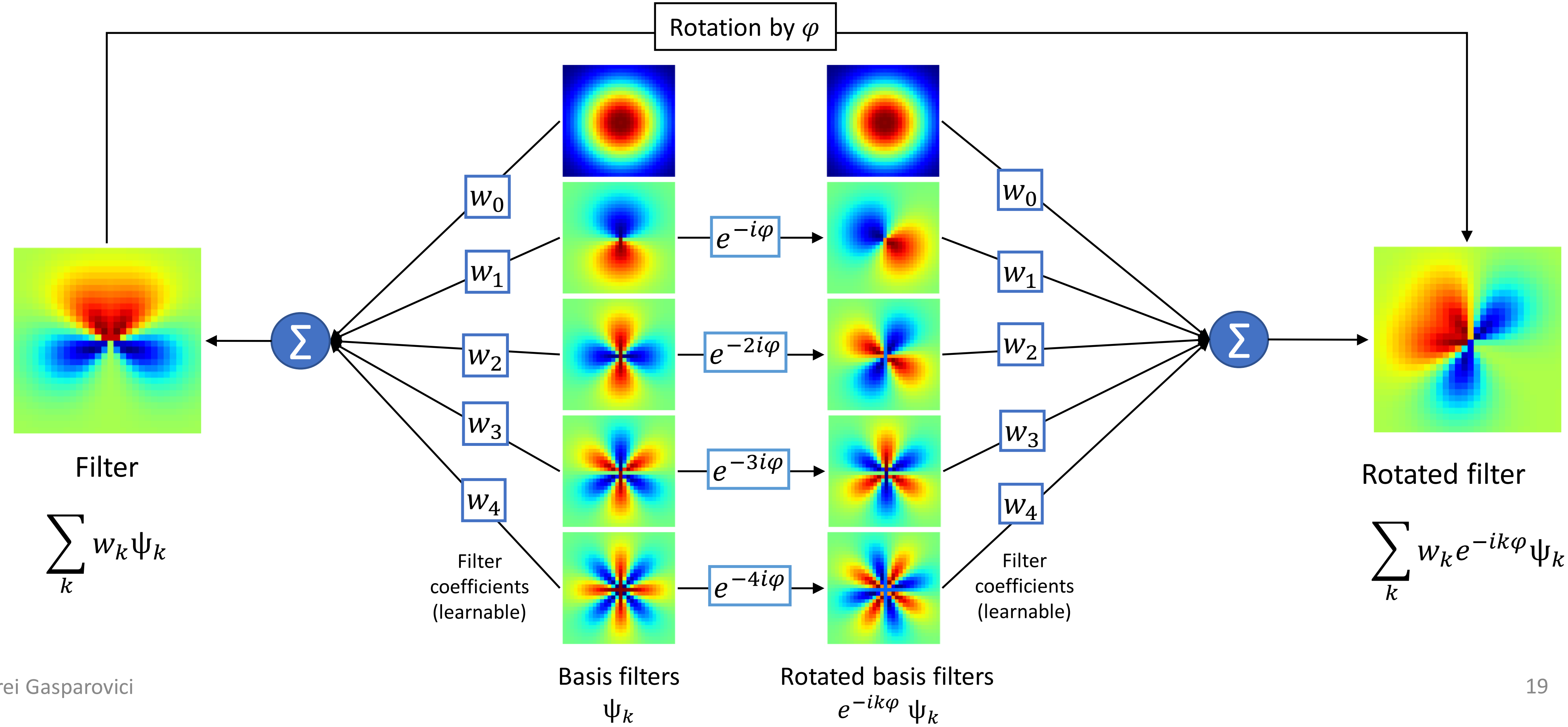
Example: Basis filters with $1 \leq k \leq 6$ and $j = 1$ sampled on a 30×30 grid (real part)

Weiler et al. *Learning Steerable Filters for Rotation Equivariant CNNs* 2017

Continuous rotations: Steerable equivariant CNNs

Rotation of steerable filters

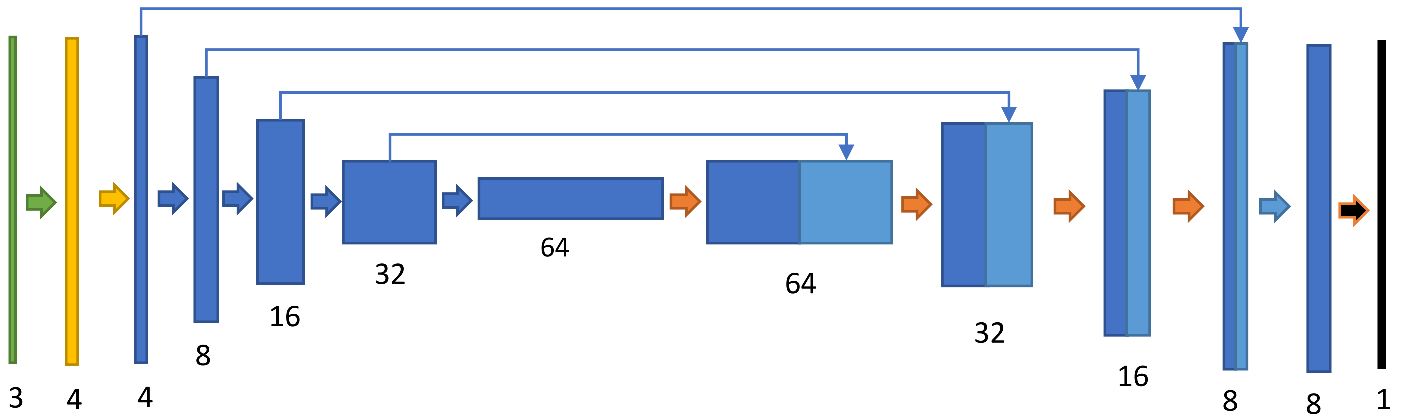
- **Rotations** of a steerable filter can be computed **analytically** by **rotating the basis filters**, thus removing the need for interpolation (rotation + sampling instead of sampling + rotation):

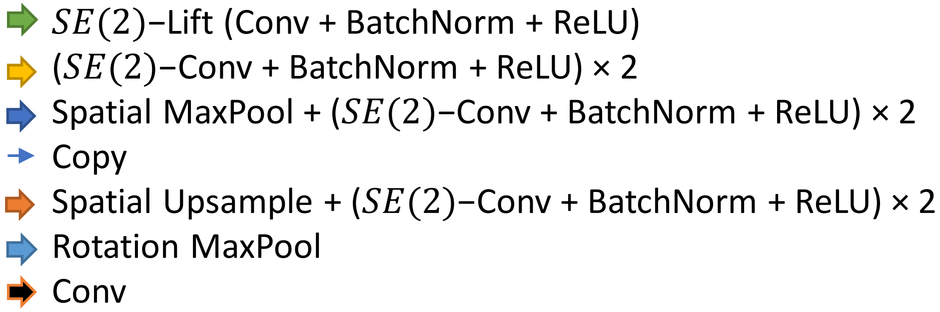


Experiment: Polyp segmentation

Architecture

Steerable $SE(2)$-equivariant U-Net





Number of parameters: 7,195,093

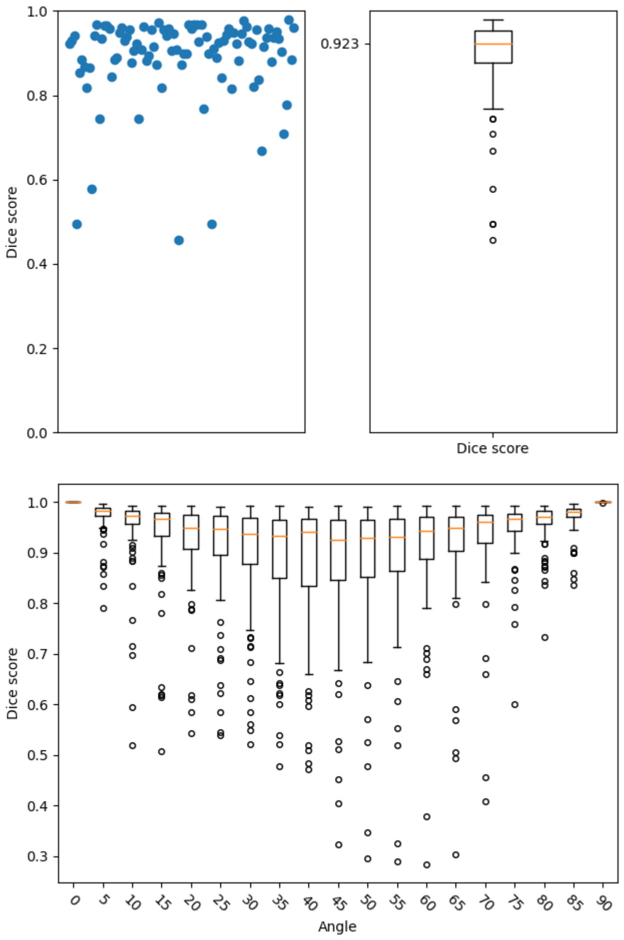
- Number of **parameters decreases significantly** and depends on the **number of basis filters chosen**

Experiment: Polyp segmentation

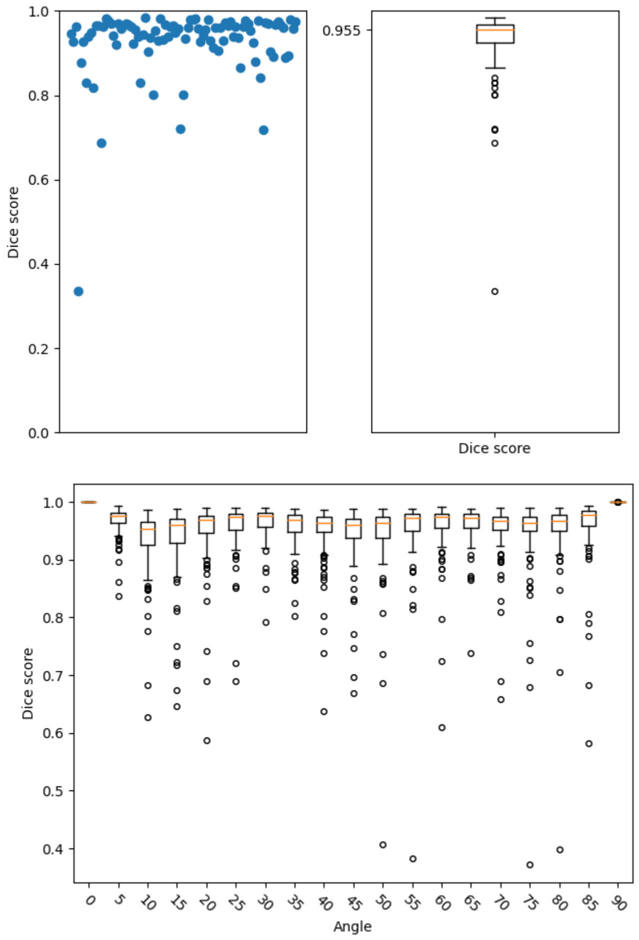
Results: **CVC-ClinicDB**

- The **steerable** equivariant network **performs better** and has **fewer outliers**
- **Equivariance** results are also **better** (with fewer parameters)

$SE(2,12)$- equivariant U-Net



Steerable $SE(2)$- equivariant U-Net

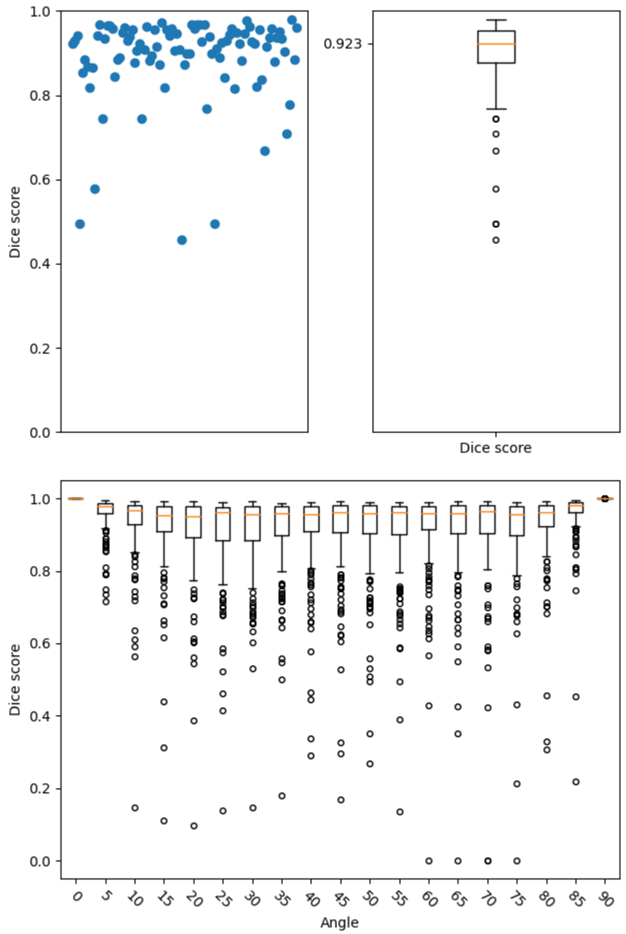


Experiment: Polyp segmentation

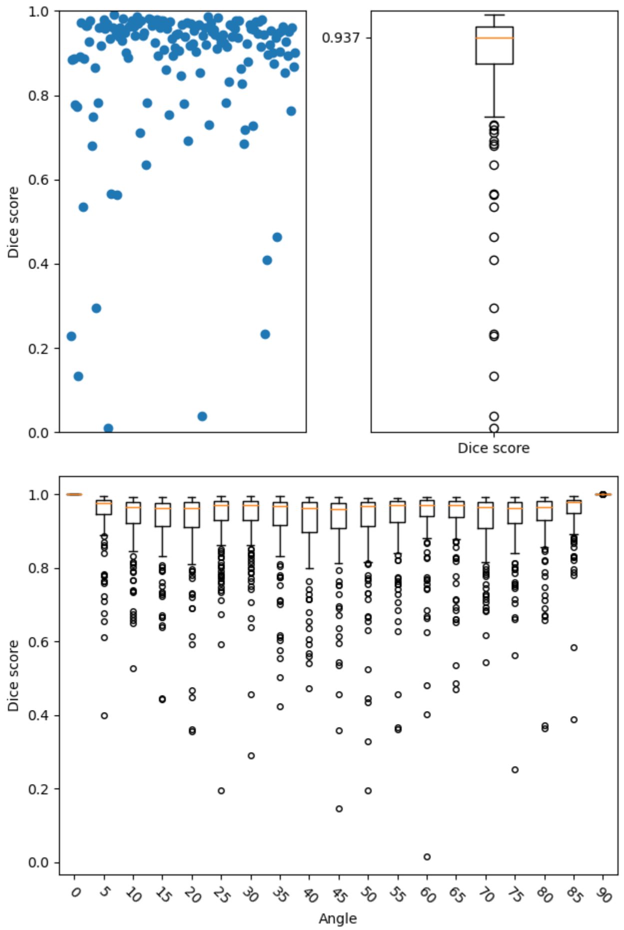
Results: **Kvasir-SEG**

- The **steerable** equivariant network **performs better**, although the number of outliers is higher
- **Equivariance** results are **preserved** (with fewer parameters)

$SE(2,12)$- equivariant U-Net

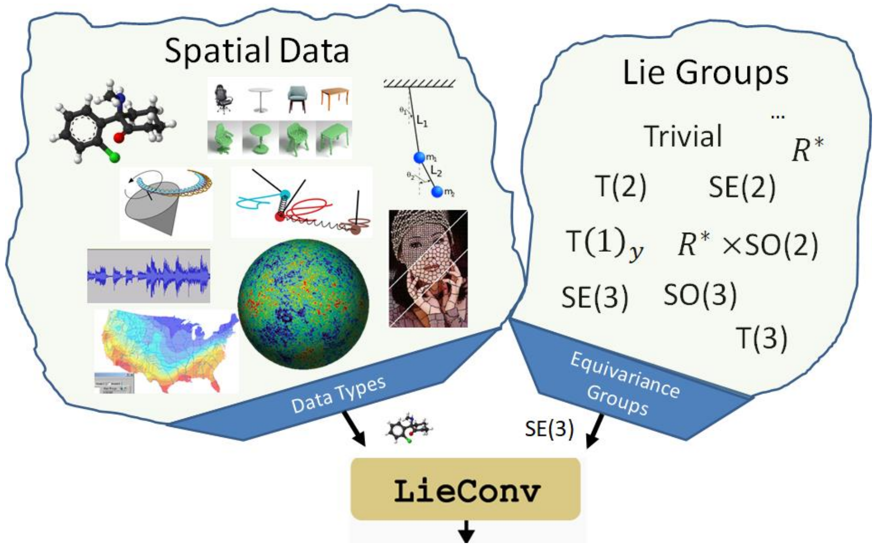


Steerable $SE(2)$- equivariant U-Net



General framework: $E(n)$-steerable CNNs & LieConv

- The steerable CNN method can be generalized to **any distance-preserving symmetry** of the n-dimensional Euclidean space
- For equivariance of other types of continuous data (molecules, point clouds, time series, videos, geostatistics etc.), one approach is **LieConv**, which discretizes the group convolution via Monte Carlo sampling



Finzi et al. *Generalizing Convolutional Neural Networks for Equivariance to Lie Groups on Arbitrary Continuous Data* 2017
Weiler & Cesa *General $E(2)$-Equivariant Steerable CNNs* 2019
Cesa at al. *A Program to Build $E(n)$-Equivariant Steerable CNNs* 2021

Implementation details

- Implementation of experiments is available at: code.siemens.com/andrei.gasparovici.ext/polyp-detection

- Building blocks:
  - $p4$-conv (easy to implement in PyTorch using `torch.rot90`)
  - Interpolation-based $SE(2)$-conv: Implemented from scratch in PyTorch (zoo/se2_conv.py) by constructing a rotation + interpolation operator as a sparse tensor
  - Steerable $SE(2)$-conv: Implemented using github.com/QUVA-Lab/e2cnn – easy to use following the examples

Conclusions

- The study of symmetry (group theory) is a very important tool in math and physics that was successfully brought to deep learning
- Applications of equivariant CNNs have not been fully explored for **medical imaging** (few papers available)
- Experimental results shows that equivariant CNNs have multiple benefits over classical CNNs, e.g., **removing** the need for **data augmentation**, improving **performance**, and enhancing **weight sharing** and **robustness**

- Future work:
  - Apply group-equivariant CNNs to other medical imaging problems
  - Explore the connection between group equivariance and other deep learning techniques (attention, generative models)
  - Investigate applications of group equivariance beyond vision (NLP, reinforcement learning)

References

- Group equivariant CNNs
  - Cohen, Taco S. / Welling, Max **Group Equivariant Convolutional Networks** 2016
  - Cohen, Taco S. / Welling, Max **Steerable CNNs** 2016
  - Weiler, Maurice / Hamprecht, Fred A. / Storath, Martin **Learning Steerable Filters for Rotation Equivariant CNNs** 2017
  - Cohen, Taco S. / Geiger, Mario / Weiler, Maurice **A General Theory of Equivariant CNNs on Homogeneous Spaces** 2019
  - Weiler, Maurice / Cesa, Gabriele **General E(2)-Equivariant Steerable CNNs** 2019
- Group equivariant CNNs in medical imaging
  - Bekkers, Erik J. / Lafarge, Maxime W. / Veta, Mitko / Eppenhof, Koen A. J. / Pluim, Josien P. W. / Duits, Remco **Roto-Translation Covariant Convolutional Networks for Medical Image Analysis** 2018
  - Winkels, Marysia / Cohen, Taco S. **3D G-CNNs for Pulmonary Nodule Detection** 2018
  - Worrall, Daniel Ernest **Equivariance For Deep Learning And Retinal Imaging** 2019
- Group theory
  - Alperin, J. L. / Bell, R. B. **Groups and Representations** 2012
  - Fulton, William / Harris, Joe **Representation Theory: A First Course** 2013

"Wherever groups disclosed themselves, or could be introduced, simplicity crystallized out of comparative chaos."

– *Eric Temple Bell*